UIP: Estimating True Rating Scores of Services through Online User Communities

Xiaodi Huang
Charles Sturt University, Australia
xhuang@csu.edu.au

Weidong Huang
University of Tasmania, Australia
tony.huang@utas.edu.au

Wei Lai
Swinburne University of Technology
wlia@swin.edu.au

Abstract— **As many online systems rely on user ratings for making decisions such as recommendations, the quality of such rating scores are increasingly important. On the other hand, users interact with each other via online communities. How such interactions affect the trueness of their ratings? Can we obtain the true rating scores that exclude the influences among users? This paper presents a conceptual framework that characterizes the influences on quality of services among users, and an algorithm that estimates the true rating scores by minimizing the influence among users. In other words, the influence on users' ratings due to their interactions is minimized so as to obtain the more accurate rating scores. The proposed approach has been validated by experimenting on real data sets. The results of the experiments have demonstrated that our approach is capable of estimating true ratings.**

Keywords—services; online social network; ratings; propagation

I. Introduction

Relying on ratings by users is a very common approach to assessing quality of online services [14,15]. With the proliferation of online services, the estimation of such accurate ratings has become a critical issue for service providers[16]. In reality, users can easily share their opinions or experience of using a service and purchasing a product with other community members who have no direct experience with them.

As we know, mutual reinforcement takes place among hyperlinks [6, 8]. The Web is evolving into Web 2.0, which is saliently characterized by facilitating collaboration and sharing information between users in an unprecedented way. Web 2.0 provides both the hyperlinks of Web pages and a platform in which users are able to interact with each other. Not only is useful the hyperlinks, but user interactions are also informative. The members of an online user community can interact with each other in various ways such as reading, commenting, posting, and discussing ratings. As a result, information and influence are spread among its members. Approaches that incorporate social influence into recommender systems or online marketing in E-commerce are capable of improving the performance. They, however, focus on the influence on recommending products. It seems to be increasingly important to improve the accuracy of estimating the ratings scores of services by measuring online social influence. In a word, link analysis ranking makes use of the hyperlink relations among Web pages, while Web 2.0 makes it possible to exploit the interactive information among a set of users in order to indirectly quantify the quality of services. In this paper, our focus is to explore the extent to which user ratings are affected by their peers from the same online community.

The basic idea of our approaches is to employ indirect computations. Relying on a user social network, the indirect approach makes use of the influence mechanism so as to indirectly measure the quality of services. In order to obtain the true rating score of a service, we take into consideration the influence between users with respect to their evaluations on services.

To the best of our knowledge, this is the first work on integrating social influence on user ratings into measurement in social networks for estimating the true quality of services. In particular, the contributions of this research are twofold.

—We describe a new conceptual model that characterizes the influence among users on rating the quality of a service.

—We present an approach to compute user influences on ratings and quality scores of services, together with the experiment results.

The remainder of this paper is organized as follows. Related work is given in the following section. A conceptual model for services is presented in Section 3. Section 4 presents the preliminaries and the algorithms for computing the true ratings of services, followed by reporting experiments in Section 5. Section 6 concludes this paper..

II. Related work

In this section, we review the relevant literatures on influences in online communities and estimations of QoS parameters, as well as compare our algorithm with other relevant ones.

The recent proliferation of online networks has aroused much research interests in computer science. Roughly, there are two categories of relevant research on online networks. One area is in characterizing an online community itself and its members, such as discovery of communities [2, 4, 9], community evolution [7], and pattern mining in network data. The interests of users within a social network are inferred using social neighbors [12]. Modeling the dynamics and topic dependency of social influence, Tang et al. [11] measures the topic-level social influence on large-scale networks.

Building upon social influence and correlation in online communities, the improvement on the quality of services or recommendations is the second area. Social network researchers have extensively studied social influence and correlation among people's behavior. Researchers in computer sciences are recently interested in exploring these issues under the context of online communities [3, 11]. As an example, [1] studies the correlation between social similarity and influence. The correlation in online social networks is reported in [36]. Lots of research attempt to model the social influence so as to describe how such influences are spread in a network.

E-commerce such as recommender systems takes advantage of the social influence between consumers to increase sales by recommending products. Such systems build predictions based on the behavior and opinions of others who share similar behaviors and opinions. The similarity among user ratings is leveraged to improve product items recommendation. The trust-based approach to recommendation has emerged [13], with the advent of online social networks. Relying on a trust network, recommendations on purchasing produces to a user are made on basis of the ratings by those users who are directly or indirectly trusted by this particular user.

Existing approaches count on user ratings without considering the influence in the course of giving ratings. Our approach reported in this paper attempts to quantify user influence as a result of user interactions over rating services. We focus on examining influence impact on rating services, rather than recommending products.

III. CONCEPTUAL MODEL

In this section, we present a conceptual model on how to influence rating the quality of a service among users.

The way of scoring services takes advantage of the fact that the Web offers a rich context of implicit information. Such information is expressed through interactions between users in Web 2.0. In order to model online interactive groups, we introduce the concept of a Web social graph (WSG), which is treated as a graph. A node in the graph represents a user while an edge in a WSG indicates the influential relations between users with respect to a service. Such typical relations include collaborations, sharing, and hyperlinks. In fact, a WSG is an online community in which a number of end users are able to share their experience through the use of a particular service.

The propagation within the a WSG is called influence propagation, meaning that a set of users in a WSG influence each other on scoring a service. A user interacts with others from the same WSG. Such interactions can happen in different forms such as reading, posting, and commenting on reviews. Users may share their personal experiences on the use of a service by writing reviews, and chatting about it among their trusted peers. All these interactions impose indirect influence on users' rating decisions. Furthermore, users may directly exchange their opinions on ratings. In some cases, a number of ratings are already available before a user is required to rate a service. The simple observation of these existing ratings may affect the user's rating. In brief, the rating of a user may be affected by others, during social interactions, in a way that users' opinions and evaluations are propagated via a WSG. Existing ratings exert an influence on ratings by subsequent users.

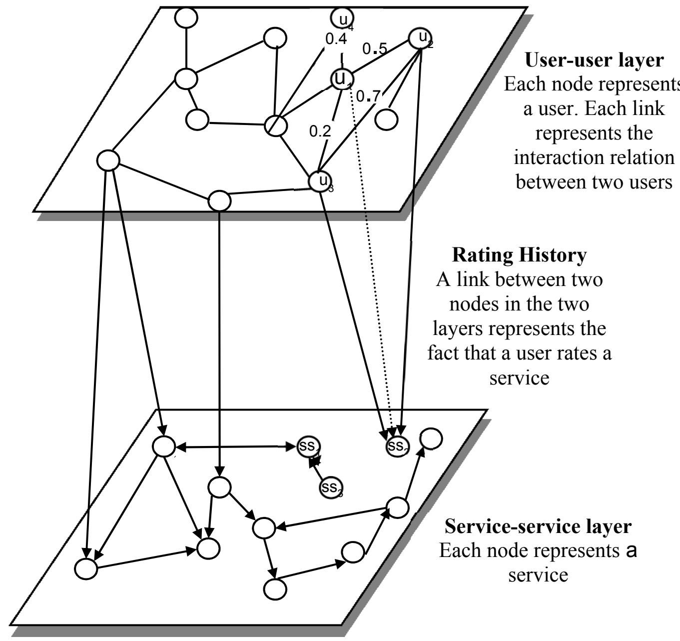

Fig.1. Two-layer graph model of users and services. The user-user layer is a weighted graph where nodes are users, edges are influence relations between users, and the weights are the degree of influence. The service-servicer layer is a graph where nodes are service providers, and edges indicate the relations between services. User ratings represent users' evaluation of services.

The framework for the propagation approach is illustrated in Fig. 1. There are two layers, namely a user-user layer and a service-service layer. Recall that a node in the two layers represents a human user and a service, respectively. Each link between two human users indicates the interactions between them with respect to rating services. Note that an edge in a user-user WSG indicates two-way interactions between two users. A user in a user-user WSG may be associated with the two layers. Some service providers are rated by a number of human users in a user-user WSG, such as the s2 in Fig. 1. In other words, an interlayer link between two layers, shown as the dotted line, is formed to associate a particular service with end-users who have already used it. For illustration, the quality of s1 in Fig. 1 is enhanced by the fact that both s2 and s3 invoke its service. The three users denoted as u1, u2, and u3 are part of the user-user WSG where they influence each other in terms of assessing s2. We assume that users u2 and u3 have already provided the rating scores to s2 before user u1 rates it. In order to rate s2, user u1 may observe the existing scores, and discuss this with users u2 and u3. As a consequence, all these interactions will influence user u1's rating on s2. In the following section, we will present a detailed model by which to calculate how much the influence is.

In summary, we conduct the influence propagation within a user-user WSG. By doing so, the scores of services are adjusted or computed.

This framework not only serves as the idea of the overall approach, but also helps to identify and tie together the theoretical aspect of this research.

IV. Computation approach

A. Preliminaries

In this section, we provide several definitions that are used for describing our approach and algorithms in the following section.

We rephrase the Web Social Graph in a formal way, which was introduced in Section 4.

Definition 1 (**Web Social Graph** (**WSG**)): a multiple-attribute graph $G = (V, E)$ modeling a Web community. It is either a user network where a number of users interact with each other in evaluating a set of services, or a service network where a number of service providers are connected to form a service supply chain. In the graph, V is a set of users, and E is a set of edges. An edge is built between nodes i and j, if user i is associated with user j regarding rating a service

We formalize the propagation problems in the following. The propagation computation model consists of three sets of different entities: a set of $WS = \{ws_1, ws_2, \cdots, ws_n\}$ with the number of n services, a set of $R = \{r_1, r_2, \cdots, r_m\}$ with m ratings over some or all these services, and a set of $U = \{1, 2, \cdots, l\}$ with the number of l users, some of whom have rated the services.

We show relationships between these sets of entities. There is a function of $M: R \rightarrow WS$ that maps each rating r to a unique service, i.e., $ws_r = M(r)$; an authorship function of $A: R \rightarrow U$ that maps each rating r to a unique user, i.e.,$u = A(r)$. Except for these functions, there is a relation of $UWSG \subset U \times U$ that defines the social network relations between users. We model a user social network relation as a weighted graph of $UWSG$ with an influence matrix $\boldsymbol{U} = \{u_{ij}\}$, where u_{ij} is the degree of influence between users i to j in terms of rating services. We formalize the problems in the following.

Definition 2 (**Problem of measuring rating influence on services**): Given a user-user WSG with a set of users U, a set of ratings R (observed scores) on a service, and the influence matrix $\boldsymbol{U}$ (constructed from user history rating data), we need to derive true scores of user ratings by reducing the amount of influence among ratings in R on this particular service.

In this following, we present our approach that computes the true score of a service in accordance with the concept model described in the previous section.

B. Algorithm

In a user-user WSG community, a number of users influence each other with respect to their assessment on a service. It is likely that a user has not used a service, thereby having no way of rating the service. If other peers in a community have already consumed this service, then their evaluations make it possible that the user who has not used the service indirectly evaluates it. In essence, the relations among a group of users enable us to compute the rating of a service in an indirect way. The existing evaluations are propagated within the community.

The common approach to scoring the quality of a service is to average user ratings as the score of this service. This straightforward approach does not take into account the interactions between users during their evaluations.

As mentioned before, the basic idea of influence propagation is that the evaluation of a service by a user is influenced by existing evaluations of other users within the same user-user WSG. The key issue is how to quantify the extent of the influence on affecting a user's rating on a service. The strength of such an influence between users depends on the score discrepancy, as well as on how easily the user is influenced by her peers. It should be proportional to their score difference when the difference is relatively small. As the score discrepancy is getting large, the influence is diminishing or even becoming zero. In other words, after reaching a peak when the score discrepancy is a critical one, the influence could be decreased gradually. As such, our model incorporates all these nonlinear coupling with users, which acts to bring rating scores closer together.

In our model, the rating of an individual user on a particular attribute of a service is shaped by two factors (1) a linear difference, which tends to make the rating towards the true value of the attribute that captures the inherent characteristics of the service; and (2) a nonlinear coupling with other individual users, which increase or decrease the affect as a result of the score difference. The strength of the influence between two users is nonlinearly in proportion to the difference between their initial rating scores. An observed score consists of two components: the true score, and influenced score. A true score is a rating given by a user who was not affected by other users. It is based on the user's own experience with the use of a service. Using user past interactive historical data in a community, we estimate the interesting true score of a user's rating on a service from observed scores by excluding the influenced score.

Given a user-user WSG with the number of users with respect to scoring a service, the observed scores of users j and i are w_j and w_i, respectively, in which w_i contains the component of the influenced score. This is because a number of w_js is available before a user gives score w_i. Suppose that the true rating given by user i is x_i. According the two above factors attributing to a user rating, the average degree of the rating score of user i affected by her neighbor users can be modeled as:

$$\frac{1}{|N_i|}\sum_{j \in N_i} \eta_{ij} \left(w_j - x_i\right) exp\left(-\frac{\left(w_j - x_i\right)^2}{2{\delta_i}^2}\right) \quad (1)$$

The error between the true score x_i and observed scores is a cost function of x_i

$$f(x_i) = \frac{1}{2}\left(x_i + \frac{1}{|N_i|}\sum_{j \in N_i} \eta_{ij} \left(w_j - x_i\right) exp\left(-\frac{\left(w_j - x_i\right)^2}{2{\delta_i}^2}\right) - w_i\right)^2 \quad (2)$$

In Eq.(2), we have

x_i: A true score of the attribute vector of a service by user i,
w_j: A component score of the attribute vector of a service by user j,
N_i: A neighborhood set of user i in a user-user WSG,
η_{ij}: The strength of the influence of user j upon user i on the ratings, and

$2\delta_i^{\,2}$ – the bandwidth for the Gaussian kernel density estimation.

In order to estimate the true score, we minimize the error of Eq.(3) with respect to x_i:

$$\text{Minimize} \quad f(x_i) \qquad \text{s.t.} \quad 0 \le x_i \le 1 \tag{3}$$

The coupling strength η_{ij} can be considered to be the product of two factors: the interaction rate at which j sends persuasive messages to i, and the regard of i for j that accounts for how susceptible i is to influence from j due to factors such as j's perceived credibility on the issue of evaluation of a service.

We can control the bandwidth to embed prior knowledge about a user. For example, we might treat a particular user who is difficult to accept other's opinions by changing the corresponding $2\delta_i^{\,2}$.

With the gradient of Eq.(2) given in Eq.(4), we can solve Eq.(4) using the gradient descent technique.

$$\frac{\partial f(x_i)}{\partial x_i} = \left(x_i + \eta_{ij}(w_j - x_i)\, exp\left(-\frac{(w_j - x_i)^2}{2\delta_i^{\,2}}\right) - w_i\right) \times \left(1 - \eta_{ij} exp\left(-\frac{(w_j-x_i)^2}{2\delta_i^{\,2}}\right) + \frac{\eta_{ij}(w_j-x_i)^2}{\delta_i^{\,2}} exp\left(-\frac{(w_j-x_i)^2}{2\delta_i^{\,2}}\right)\right)$$
$$=\left(x_i - w_i + \eta_{ij}(w_j - x_i)\, exp\left(-\frac{(w_j-x_i)^2}{2\delta_i^{\,2}}\right)\right) \times \left(1+ \quad \eta_{ij}\frac{(w_j-x_i)^2-\delta_i^{\,2}}{\delta_i^{\,2}} exp\left(-\frac{(w_j-x_i)^2}{2\delta_i^{\,2}}\right)\right) \tag{4}$$

The solution to Eq.(4) is the estimation of the true score given by user i that eliminates the influences from other users.

Using an influence function $I(i)$ that quantifies the influence strengths on user i by all other users, we have:

$$w_i = \text{aggr}(w_i, \text{aggr}(I(i))) \tag{5}$$

where the first aggr function combines the original rating by user i (w_i) into the amount of the influence she receives $(\text{aggr}(I(i)))$, and the second aggr aggregates the individual influence by all other users. The benefits of using the influence function are of two fold. First, we are able to choose other aggregation functions by which to accumulate the influences from others. Second, two aggregation functions can be different. A simple weighted average function is employed in Eq.(1).

The last question is how to acquire the values of η_{ij} in Eq.(1). Based on the similarities of the ratings, we generate a user-user WSG with η_{ij} as the weights of the edges from user past rating data (no edges for two users who have no common rating items). We cannot directly measure user interactions, and exchange opinion on evaluating a service. However, we can make an assumption that the more similar two users' scores on rating a service, the more likely the users influence each other. At the same time, the number of services both of two users have rates is also an indicator of their interaction. As such, we combine the Jaccard coefficient into the cosine similarity metric:

$$\eta_{ij} = \frac{1}{|S_i \cup S_j|} \sum_{i,j \in S_i \cap S_j} \frac{\boldsymbol{w_i} \cdot \boldsymbol{w_j}}{|\boldsymbol{w_i}|_2 |\boldsymbol{w_j}|_2} \tag{6}$$

where S_i and S_j are a set of services rated by users i and j, respectively. The vectors of $\boldsymbol{w_i}$ and $\boldsymbol{w_j}$ are the respective score vector of services rated by both users i and j.

We present our algorithm called UIP that calculates the user influence propagation on scoring services.

UIP Algorithm

Input: a current user-service rating matrix ***R***, a past user-service rating matrix ***M***, and the parameter $2\delta^2$.

Output: the estimated true scores of the services by users in ***R***

Generate a user-user matrix $\boldsymbol{U} = [u_{ij}]$ from the matrix ***M*** where $u_{ij} = \eta_{ij}$ according to Eq. (6)

for each service s in ***R***

for each user i in ***R***

find her neighbor set N_i

obtain the observed scores of user i and user $j \epsilon N_i$ from ***R***, and w_{ij} from ***U***

solve Eq. (3), and find the true score of user i on rating service s

end

end

In the current user-service rating matrix ***R***, and the past user-service rating matrix ***M***, their rows indicate different users, while their columns refer to different items, the element (i, j), for example, is a normalized rating score on item j given by user i. The element (i, j) of the user-user matrix ***U*,** which is a representation of the user-user WSG, is η_{ij}, namely $w_{ij} = \eta_{ij}$, reflecting the connection degree of the two users.

The time complexity of the UIP algorithm is $O(mlk)$ where m is the number of services, l the number of users, and k the number of 1-nearest neighbors of users.

We have presented the approaches to calculating the scores of different kinds of attributes.

We have presented the approach that computes the scores of the two attribute types of services. the approach relies on a Web social network. Normally, a service with a high score computed by the aggregation approach should receive a good rating from users using the propagation approach. In other words, the two aspects of a service should be highly related to each other, although each evaluation focuses on different perspectives.

V. Experiments

In this section, we describe several experiments for validating our algorithms that estimate the true scores of services. The experiments consist of two parts. The first part applied different aggregation functions to the same dataset. The second one evaluates the UIP algorithm.

We implemented the UIP algorithm. The experiments were conducted on a machine with 2 quad core 3 GHz CPUs, 32 GB RAM, and the Red Hat Linux system.

C. Evaluation metric

Before describing the experiments, we need to choose an evaluation metric that is able to capture the diversity of a number of user rating scores on a service. Using the normalized Shannon entropy that is commonly used in the literature, we quantify the rating diversity of an individual service *i* as

$$D(i) = \frac{-\sum p log(p)}{log(k)} \quad (7)$$

where k is the number of ratings on service i, and p is the probability of a possible rating scale. High diversity scores imply that users' opinions on evaluating a service are diverse. If all user ratings on SERVICE i are the same, then $D(i)$=0. In contrast, we have $D(i)$=1 if all ratings are different.

D. Dataset

We use the widely cited dataset of epinions instead (http://www.trustlet.org/wiki/Extended_Epinions_dataset). Although no datasets on user ratings of services are currently available for general use, we believe that the rating scores on items in epinions are similar to those on services. Using the first part of UIP algorithm, we are able to construct a user-user WSG with 49289 users. Two users are connected to each other in the graph only if they rate at least one common item. The η_{ij}weight of their edge connecting them is equal to the similar score calculated by Eq.(6).

The statistics of the generated graph includes the max degree (1-nn, the number of incident edges) of 322, average degree of 1.5158, and the variance of the degrees of 15.8736.

E. Experiment procedure and results

The experimental procedure is as follows. We test the UIP algorithm against every item of 139738 items in total in the dataset. Each test on each item is done using the four respective aggregation functions of average, max, min, and product that aggregate the influence on each user's rating from others. For each item, the algorithm thereby produces the four lists of true scores corresponding to a number of users who have rated this particular item. We then calculate the entropies of the four true score lists and the observed score list, as well as the correlation coefficients between them. Due to the limited space, part of resulting data is listed in Table 1.

Part of resulting data of the dataset using the UIP algorithm by using the ave, max, min, and prod. functions to aggregate the influence from others on a particular user's rating. The diversities of the rating score before and after the adjustments are listed with respect to different aggregation functions.

We note that 18844 (57.14%) among 32978 observed scores are bigger than their corresponding true scores. The mean of the entropies of the observed scores for all items is $\frac{1}{32978}\sum_i D(i) = 0.6607$, with the variance of 0.0385, while that of their true scores is 0.6233 with the variance of 0.0571. The variance among the observed scores is obviously reduced. This is consistent with the intuition that user ratings reach the consensus equilibrium after their interactions. We further validate this in the following.

For comparison, the entropies of observed scores on all items are ranked in ascending order, and the entropies of their corresponding true scores are accordingly repositioned in the lists. The results on 139738 items including part of data in Table 1 are illustrated in Fig.2.

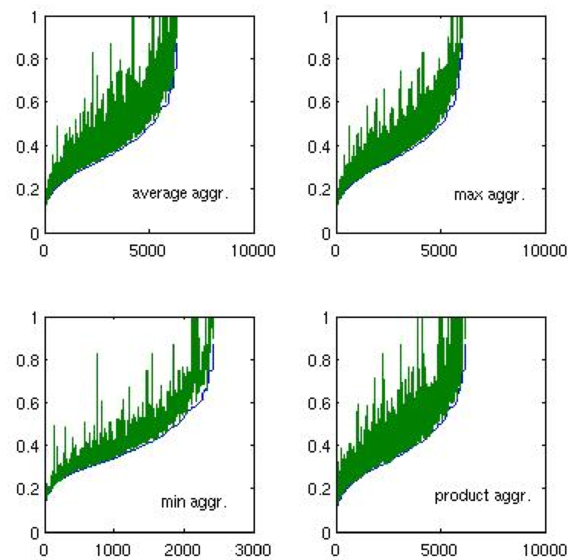

Fig.2. Entropies of resulting scores on 139738 items using aggregation functions of average, max, min, and product.

Having observed Fig.2, we can reach two conclusions regardless of which aggregation function is employed.

1) All vertical lines that show the entropies of true scores are positioned above the curves of the entropies of observed scores. In other words, the normalized entropy of true scores of each item is bigger than that of its corresponding observed scores. This means that true scores have more diversity than their observed scores. The reason for this is that users' ratings converge to a consensus by reducing discord after user interactions; and

TABLE 1

PART OF RESULTING DATA OF THE DATASET USING THE UIP ALGORITHM BY USING THE AVE, MAX, MIN, AND PROD. FUNCTIONS TO AGGREGATE THE INFLUENCE FROM OTHERS ON A PARTICULAR USER'S RATING. THE DIVERSITIES OF THE RATING SCORE BEFORE AND AFTER THE ADJUSTMENTS ARE LISTED WITH RESPECT TO DIFFERENT AGGREGATION FUNCTIONS.

Item Index (*i*)	**# 1-NN**	**Running Time (sec.)**	**Rating diversity *D*(*i*)**					**Correlation (O and T)**			
			Observed Scores(O)	***True scores (T)***							
				Ave.	Max.	Min	Prod.	Ave.	Max.	Min.	Prod.
1	13	0.1251	0.4935	0.6182	0.5508	0.5610	0.5372	0.9945	0.9939	0.9983	0.8451
47	17	0.1376	0.4588	0.5733	0.5213	0.5107	0.5733	0.9962	0.9982	0.9981	0.9962
61	60	0.4832	0.2316	0.3436	0.2770	0.2996	0.3619	0.9899	0.9951	0.9951	0.9898
78	9	0.0734	0.6230	0.7632	0.6931	0.6931	0.7632	0.9937	0.9960	0.9957	0.9937
187	41	0.3220	0.3670	0.4081	0.3868	0.3883	0.4081	0.9979	0.9990	0.9989	0.9979
233	98	0.8038	0.2868	0.3016	0.2940	0.2944	0.3052	0.9961	0.9959	0.9976	0.9962
252	9	0.0753	0.6495	0.8333	0.7196	0.6931	0.7196	0.9185	0.9293	0.9696	0.9293
309	162	1.6221	0.1492	0.1806	0.1758	0.1543	0.1947	0.9947	0.9962	0.9987	0.8827
358	140	1.5260	0.1947	0.2377	0.2311	0.2082	0.2140	0.9929	0.9845	0.9977	0.9604
361	57	0.4514	0.1910	0.2160	0.2160	0.2063	0.2160	0.9953	0.9800	0.9967	0.9963
107721	5	0.0447	0.6555	1.0000	0.8277	0.8277	1.0000	0.9758	0.9860	0.9910	0.9758
108268	6	0.0558	0.7421	0.8710	0.8710	0.8710	0.8710	0.9753	0.9929	0.9848	0.9751
126687	6	0.0504	0.7421	0.8710	0.8710	0.8710	0.8710	0.5773	0.5773	0.9723	0.5773
128127	6	0.0559	0.5645	0.6934	0.6934	0.4842	0.6934	0.9057	0.8589	0.8440	0.4042
136215	4	0.0350	0.4056	0.7500	0.7500	0.7500	0.7500	0.7229	0.7229	0.9623	0.7229

2) On average, the adjusted (observed) ratings do not deviate from than their original (true) ratings too far. The respective true scores by using four aggregation functions are strongly correlated with the observed scores (the coefficients listed in the last four columns in Table 1). This implies that despite of being affected by their peers' opinions, to some extent users still insist on their own evaluations. It is the presence of this self-bias force that prevents the users from coming to exact agreement.

In brief, user ratings on an item may begin with much disagreement. After reading other comments, discussing, or reading others' ratings, users may adjust their true ratings. Due to this, the true ratings evolve toward clusters of homogeneous ratings as the observed ones.

F. Effects of parameters

In the following, we examine to what extent the size of the kernel width $2\delta^2$ in Eq.(2) affects the true scores found. For this purpose, we systematically vary the kernel size such that the algorithm can produce a true score from its lower bound to the upper bound. For each size of the kernel width, the UIP algorithm outputs a true score of a user rating on an item using the average aggregation of all influences from other users. The experimental data is plotted in Fig.3.

From Fig.3, we can make an interesting observation. All three curves show nearly flat lines in their middle parts. Examining Eq.(2), we know that the model is approximately linear, if all score differences $w_j - w_i$ from a user's influence set are all small compared with the latitudes of the kernel width $2\delta_i^2$. It is this reason why Fig.3 shows the "flat" plateau. Furthermore, we deduce that the difference reaches a peak when the score difference $w_j - w_i$ is equal to $2\delta_i^2$.

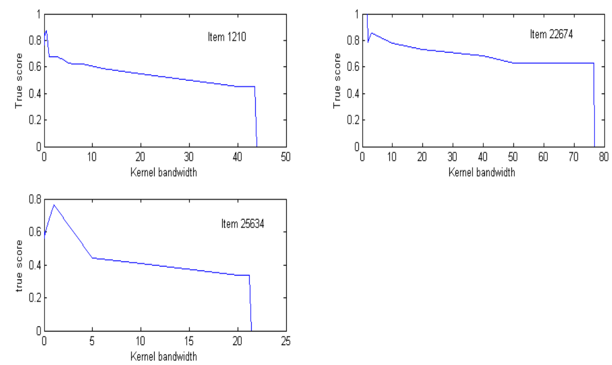

Fig.3. Kernel bandwidth and the entropies of true scores on three items.

G. Running time

We run the UIP algorithm against the varied sizes of users with different items. The running time is averaged over three runs of the algorithms. Parts of running time data with respect to the number of neighbors are listed in the third column in Table 1, as illustrated in Fig.4. Fig.4 shows some curves. This may result from the diverse sizes of the nearest neighbors of each item, and users have varying degrees of influence in a social network.

As we know, social influence results in the deviation from users' original ratings. The experiments have demonstrated that the UIP algorithm is able to find the true scores of user ratings from the observed scores. After obtaining these true scores, we will use them to further obtain the final scores of services.

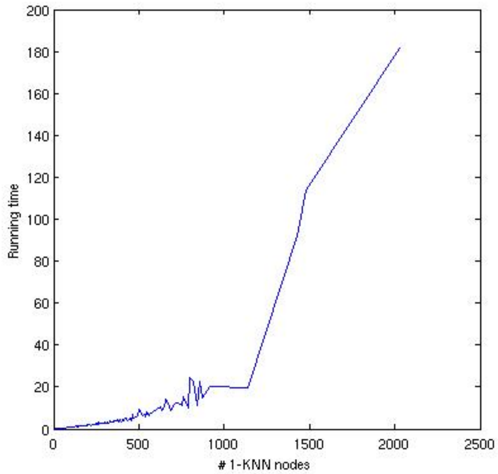

Fig.4. Running time and the number of nodes.

In this section, we have presented our approach to computing scores of services accurately. This accuracy has been achieved by the following way: our approach uses the adjusted true scores of human user ratings, which, by removing influenced scores, can reflect users' original evaluations more accurately than observed scores; The estimated scores therefore accurately capture the inherent qualities of the services.

V. CONCLUSION

The proliferation of services unavoidably leads to competition among the large numbers of services that offer similar functions. User rating scores are commonly used as a key metric in distinguishing between competing services. Therefore, it is important to accurately estimate rating scores by minimizing the user influence.

In this paper, we have presented a conceptual model that describes user influences on rating of services. Based on this model, the algorithm of UIP has been presented. Considering the influence impact on ratings among users within an online community, the UIP algorithm computes users' true rating scores of services. The proposed approach takes advantage of the evaluations from the communities of human users, in order to estimate true rating scores. As such, it is capable of estimating scores of services accurately. This has also been validated by the experiments.

REFERENCES

[1] A. Anagnostopoulos, R. Kumar, and M. Mahdian. "Influence and correlation in social networks," Proceeding of the 14th ACM SIGKDD international conference on Knowledge discovery and data mining , pp. 7–15, 2008

[2] A. Clauset, M. E. J. Newman, and C. Moore, "Finding community structure in very large networks," Phys. Rev. E, vol. 70, no. 6, p. 066111, Dec 2004.

[3] D. Crandall, D. Cosley, D. Huttenlocher, J. Kleinberg, and S. Suri., "Feedback effects between similarity and social influence in online communities," Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining , pp. 160-168,2008.

[4] I. F. Gergely Palla, Imre Dernyi and T. Vicsek, "Uncovering the overlapping community structure of complex networks in nature and society," Nature, vol. 435, pp. 814–818, 2005.

[5] R. Guha , R Kumar, P Raghavan, A. Tomkins,"Propagation of Trust and Distrust," Proc. ACM World Wide Web conference, ACM Press, pp. 403–412,2004.

[6] J. Kleinberg ," Authoritative sources in a hyperlinked environment," Journal of ACM, vol.46, no.5 ,pp. 604-632,1999.

[7] P.J. Mucha, T. Richardson, K. Macon, M.A. Porter, J.P. Onnela, "Community structure in time-dependent, multiscale, and multiplex networks," Science, vol.328, pp. 876–878, 2010.

[8] L. Page, S. Brin, R. Motwani, T. Winograd , "The pagerank citation ranking: Bringing order to the web," SIDL-WP-1999-0120, Stanford University.

[9] F. Radicchi, C. Castellano, F. Cecconi, V. Loreto, and D. Parisi, "Defining and identifying communities in networks," Proceedings of the National Academy of Sciences of the United States of America, vol. 101, no. 9, pp. 2658–2663, 2004.

[10] S. Rosario, A. Benveniste, S. Haar, and C. Jard, "Probabilistic QoS and soft contracts for transaction based Web services orchestrations," IEEE Transactions on Service Computing, vol.1, no.4, pp.1-14, 2008.

[11] J. Tang, J. Sun, C. Wang, and Z. Yang," Social influence analysis in large-scale networks," In KDD, pp.807-816, 2009.

[12] Z. Wen, C.Y. Lin," On the Quality of Inferring Interests From Social Neighbors," Proceeding of the 16th ACM SIGKDD international conference on Knowledge discovery and data mining, pp.373–382, 2010.

[13] M. Jamali and M. Ester. TrustWalker,"A random walk model for combining trust-based and item-based recommendation," KDD, pp.397–406. ACM, 2009.

[14] A. Doan, R. Ramakrishnan, and A.Y. Halevy. Crowdsourcing Systems on the World-Wide Web. Commun. ACM. 54(4), pp. 86-96, 2011.

[15] Q. Feng, L. Liu, and Y. Dai. Vulnerabilities and Countermeasures in Context-Aware Social Rating Services. ACM Trans. Internet Technol. 11(3), pp. 11:1-11:27, 2012.

[16] M.. Allahbakhsh, A. Ignjatovic, An Iterative Method for Calculating Robust Rating Scores, IEEE Transactions on Parallel and Distributed Systems, 26(2),2015.