Enhanced User Context-Aware Reputation Measurement of Multimedia Service

SHANGGUANG WANG and AO ZHOU, Beijing University of Posts and Telecommunications
WEI LEI, Chinese Academy of Sciences
ZHIWEN YU, Northwestern Polytechnical University
CHING-HSIEN HSU, Chung Hua University
FANGCHUN YANG, Beijing University of Posts and Telecommunications

Reputation plays an important role for users in choosing or paying for multimedia applications or services. Some efficient multimedia reputation-measurement approaches have been proposed to achieve accurate reputation measurement based on feedback ratings that users give to a multimedia service after invoking. However, the implementation of these approaches suffers from the problems of wide abuse and low utilization of user context. In this article, we study the relationship between user context and feedback ratings according to which one user often gives different feedback ratings to the same multimedia service in different user contexts. We further propose an enhanced user context-aware reputation-measurement approach for multimedia services that is accurate in two senses: (1) Each multimedia service has three reputation values with three different user context levels when its feedback ratings are sufficient and (2) the reputation of a multimedia service with different user context levels is found using user context sensitivity and user similarity when its feedback ratings are limited or not available. Experimental results based on a real-world dataset show that our approach outperforms other approaches in terms of accuracy.

CCS Concepts: • **Information systems** → **Reputation systems**

Additional Key Words and Phrases: Multimedia service, user context, reputation, feedback rating, user similarity

ACM Reference Format:
Shangguang Wang, Ao Zhou, Wei Lei, Zhiwen Yu, Ching-Hsien Hsu, and Fangchun Yang. 2016. Enhanced user context-aware reputation measurement of multimedia service. ACM Trans. Multimedia Comput. Commun. Appl. 12, 4s, Article 59 (October 2016), 18 pages.
DOI: http://dx.doi.org/10.1145/2978569

1. INTRODUCTION

With the increasing popularity of mobile networks recently, the widespread use of smartphones and other mobile devices contributes to unprecedented subscriptions of online multimedia content or services (e.g., movies, videos on demand, video sharing), sharing of mobile multimedia on social networking sites such as Facebook,

---

This work was supported by the NSFC (No. 61472047), the Major Project supported by National Social Science Foundation of China (Grant No.15ZDA063), and the Open Research Fund Program of Beijing Key Laboratory on Integration and Analysis of Large-scale Stream Data.
Authors' addresses: S. Wang, A. Zhou, and F. Yang, State Key Laboratory of Networking and Switching Technology, Beijing University of Posts and Telecommunications, Beijing, 100876, China; emails: {sgwang, aozhou, fcyang}@bupt.edu.cn; W. Lei, Institute of Information Engineering, Chinese Academy of Sciences, Beijing, 10093, China; email: 13426013647@126.com; Z. Yu, School of Computer Science, Northwestern Polytechnical University, Xi'an, 710072, China; email: zhiwenyu@nwpu.edu.cn; C.-H. Hsu (corresponding author), School of Mathematics and Big Data, Foshan University; Department of Computer Science and Information Engineering, Chung Hua University, 30012, Taiwan; email: chh@chu.edu.tw.

DOI: http://dx.doi.org/10.1145/2978569

and streaming on websites such as YouTube. Additionally, the web and mobile multimedia converge, as the mobile networks become an integral part of the Internet [Kovachev et al. 2014; Dong et al. 2014, 2015].

In the face of the huge number of multimedia services, reputation of a service plays an important role and users often select and purchase high-reputation multimedia content or service. Hence, for multimedia-service providers, it is important to ensure that all of their multimedia services have a high reputation value, because it can potentially increase their overall profit.

Reputation is the collective perception of a multimedia service by its users. The reputation of an invoked multimedia service is the collective feedback rating of the users that have interacted with or used the multimedia service in the past [Wang et al. 2015a]. Accurate reputation measurement of multimedia services on the Internet is important in identifying good multimedia-service providers. Hence, the ability to obtain an accurate reputation score of each multimedia service is also important [Lee and Oh 2013; Wang et al. 2015b].

Most reputation measurement schemes of multimedia services rely on the aggregation of user feedback ratings over a specific period of time (a sample interval). However, as it is not realistic to assume that user feedback ratings are fairly accurate [Wang et al. 2015b], several studies have recognized the importance of improved and accurate reputation measurements of multimedia services. The proposed solutions [Atrey et al. 2008; Lua et al. 2011; Lages et al. 2007; Liu and Shi 2010; Malik and Bouguettaya 2009a; Wang et al. 2008, 2011; Xu et al. 2007; Wang and Lin 2008] employ different techniques to measure reputation based on user feedback ratings. Although previous work has explored the efficiency and robustness of various measurement approaches, most of them suffer from the weaknesses described, as follows.

Wide abuse of user context. Almost all existing approaches are based on all historical feedback ratings, although some of these approaches consider user context. Wide abuse of user context can often result in the reputation value of a multimedia service to be the same for all consumers, which is obviously inaccurate. It is well known that consumers often have different experiences regarding the quality of a multimedia service [Li et al. 2014]. The differences may be caused by several factors, such as network environment and user terminal's ability. Even an individual user has different experiences for the same multimedia service in different user contexts. For example, consider Sam, who, after using a smartphone on a 3G network to watch a video clip, gives a medium feedback rating. However, after using the same smartphone on a 4G network to watch the same video again, he gives a high rating. Obviously, any difference in the user context may affect the user's feedback rating. Then, for one multimedia service, feedback ratings in a certain user context are not suitable for calculating the reputation value for these users in another user context. Hence, it is more practical and accurate that, if users have different user contexts of using one multimedia service, the reputation value of this service is not a single value, but differs according to the context.

Low utilization of user context. It is well known that reputation systems rely on past information to establish trust among unknown participants. Therefore, we refer to the aggregated perception that the user community has toward a given multimedia service as its reputation. However, user perceptions might not always be available, for example, when a multimedia service is initially created for business profit, no consumer has interacted with it, and no feedback rating exists for the past performance of the service. Consequently, consumers cannot assess its reputation, and questions about its trustworthiness are left unanswered, which could cause users to overlook the multimedia service for future transactions. Therefore, it is crucial for reputation systems to assign reputation for newly created multimedia services even when no feedback rating on their performance exists, so that they can compete with existing multimedia services for market share [Malik and Bouguettaya 2009b]. However, little attention has been given to reputation measurement with few or no feedback ratings, and the approaches that consider the problem [Maximilien and Singh 2001, 2002; Zacharia et al. 2000] often adopt solutions (assigning neutral or default reputation values to newly created multimedia services) that are not fair to all multimedia services. Hence, they cannot obtain the inherent feedback rating of the multimedia service by eliminating user context effect according to other similar multimedia services, thus fail in assigning a fair reputation when feedback ratings are limited or not available.

To address these two problems related to user context, we study the relationship between user context and feedback ratings according to which user often gives different feedback ratings to the same multimedia service in different user contexts. We further propose an enhanced user context–aware reputation measurement approach that achieves accurate reputation measurement of multimedia services regardless of the number of feedback ratings (sufficient, low, or zero). The contributions of this article are as follows:

1) Aiming at the problem that different user contexts affect user feedback ratings differently for multimedia services, we propose the two concepts of reputation vector and user-context sensitivity, such that we can obtain different reputation values with different user-context levels. To the best of the our knowledge, this is the first effort in considering the reputation of a multimedia service as a vector composed of reputation value and user context level.
2) We propose an enhanced user context–aware reputation measurement approach. This approach first models, formalizes, and normalizes user context. Then, it clusters user context in order to use a quantitative value to represent user-context level and user-context sensitivity. Third, (a) when the number of feedback ratings is sufficient, it uses feedback ratings in a given user context to measure multimedia service reputation; (b) when the number of feedback ratings is low or zero, by adopting user-context sensitivity to weaken the impact of user context on user-feedback ratings and obtain user-inherent feedback ratings without user-context effect, this approach employs user-similarity computing to measure the reputation of a multimedia service or newly published multimedia service.
3) To evaluate our approach, we implement all approaches based on a real feedback rating dataset and compare our approach with others. Experimental results show that our proposed approach can obtain higher accuracy than other approaches.

The remainder of this article is organized as follows: Section 2 reviews related work in the area. Our proposed reputation measurement approach is detailed, including user-context computing, user-similarity computing, and reputation measurement, in Section 3. Experiments to compare our proposal against prevalent approaches are described in Section 4. We offer our conclusions as well as an outlook for future work in the area in Section 5.

2. RELATED WORK

A number of schemes have been proposed for reputation measurement of multimedia or web services. However, in this article, we review only selected notable works, which consist of reputation measurement when feedback ratings are sufficient and reputation measurement when feedback ratings are limited or not available.

From the perspective of reputation measurement when feedback ratings are sufficient, Atrey et al. [2008] present a method that dynamically computes the reputation of a multimedia service on the basis of its association with other multimedia services in a composition task to overcome the dependency on users' feedback ratings. Unfortunately, Atrey et al. [2008] fail to consider user context. Similarly, Conner et al. [2009] propose a trust framework of managing services on the basis of reputation. Their main idea is trust management service (TMS). TMS not only supports a trust relationship between several entities, but also allows each entity to use its own scoring function to perform reputation measurement. The main advantage of TMS is that it can support different scoring functions so that each entity can use its own effective and specific functions. It is more accurate than directly measuring reputation on the basis of feedback ratings. However, different user contexts will lead to different feedback ratings after using a multimedia service; thus, the approach of Conner et al. [2009] and other similar studies [Caverlee et al. 2008; Lua et al. 2011; Kamvar et al. 2003; Lages et al. 2007; Xiong and Liu 2004; Liu and Shi 2010; Malik and Bouguettaya 2009a; Wang et al. 2008, 2011; Xu et al. 2007; Wang and Lin 2008; Zhou and Hwang 2007] had a lack of accuracy and objectivity for reputation measurement. Our previous work [Wang et al. 2015b] proposed a reputation measurement approach for web service recommendations. This approach first detects malicious feedback ratings by adopting the cumulative sum control chart, then reduces the effect of subjective user feedback preferences by employing the Pearson correlation coefficient (PCC). Although this approach is effective when malicious feedback ratings exist and can consider user context, it results in an abuse of user context, and the reputation of a multimedia service is identical for all users regardless of user context. Ghaffarinejad and Akbari [2013] proposed a reputation mechanism based on a number of special reputation centers (SRCs) in service-oriented environments. Each SRC is responsible for gathering feedback on a specific service offered by different service providers. They appeal to different service users with a common interest to form a community and collect their feedback as well as feedback from other sources. However, Ghaffarinejad and Akbari [2013] and other similar works [Alnemr et al. 2009; Lee et al. 2012; Wen et al. 2012; Yan et al. 2015] rarely take advantage of user context to measure reputation, and most fail in finding the inherent reputation of multimedia services.

From the perspective of reputation measurement when feedback ratings are limited or not available, traditional works [Maximilien and Singh 2001, 2002; Zacharia et al. 2000] often assign neutral or default reputation values to newly published services. This assignment favors either existing services or new services. If the initial reputation is set to high, existing services are left at a disadvantage, as the newcomers would get preference over existing services who may have worked hard to attain their reputation. If low initial values are assigned, as a new service, it may not be able to win consumers' favor with its low reputation. Unlike traditional works [Maximilien and Singh 2001, 2002; Zacharia et al. 2000], Bagheri et al. [2009] propose a reputation estimation model for multicontext environments. The model is suitable for online communities that constitute multiple contexts, and focuses on the propagation of already observed contextual reputation to unobserved contexts. However, this approach cannot easure the reputation of newly published multimedia services. Malik and Bouguettaya [2009b] provide two techniques for bootstrapping the reputation of newly deployed services. The first technique proposes the use of a dishonest transactions ratio to guide the consumer in initializing service reputations. The second technique proposes obtaining help from community providers in assigning an initial reputation for new services. However, this approach cannot guarantee fair or accurate reputation initialization because variation in user context affects the deserved reputation.

In contrast to the existing approaches, which cannot achieve fair and accurate reputation measurement for multimedia services because of wide abuse and low utilization of user context, the proposed approach, which is different from our previous work [Wang et al. 2015a; Li et al. 2014], focuses on achieving accurate reputation measurement in two senses: (1) the reputation is a vector that contains different reputation values with three different user-context levels when feedback ratings are sufficient and (2) the reputation with different user-context levels is found using user-context sensitivity when feedback ratings are limited or not available.

3. REPUTATION MEASUREMENT APPROACH

3.1. System Architecture

Multimedia service providers will likely not realize their affirmed performance because of several factors, including network and reliability. This will raise the question of how users can believe that a multimedia service can provide high-quality service as promised. To solve this problem, the concept of reputation is proposed, which is calculated using feedback ratings from users. For example, when a user invokes a multimedia service, that user will give a feedback rating to represent the degree of satisfaction with the service. Then, a feedback platform, such as Amazon, will collect many users' feedback ratings. In this way, for one service, the platform will gather numerous historical feedback ratings. By using an algorithm, we can calculate one value to represent the reputation of this multimedia service on the basis of all the historical feedback ratings. This is very similar to when a user watches a film on Youku (a popular video website in China); the user may give the service a rating to indicate the level of user satisfaction.

In this study, for the jth invoked multimedia service ms_j, a user provides a feedback rating that indicates the level of satisfaction with the multimedia service after each interaction over a specific period of time (a sample interval). A feedback rating is simply an integer that ranges from 1 to R, R representing extreme satisfaction and 1 representing extreme dissatisfaction. (Many platform websites, e.g., Youku, allow ratings on a scale of 1 to 10. In most cases, 10 represents the highest score, and 1 represents the lowest.) Then, users maintain n feedback ratings that represent their perception of ms_j performance. We take $R(ms_j)$ to represent the reputation score of ms_j over a global time period. Then, $R(ms_j)$ can be calculated as

$$R(ms_j) = \frac{1}{n}\sum_{i=1}^{n} r_i, \tag{1}$$

where r_i represents the i-th feedback rating and n represents the number of feedback ratings. The higher the score, the more the user likes this service.

Until now, platform websites have measured a value for each service, namely, for all users, such that the service reputation is identical to all users. However, in a practical service environment, a service will perform differently in the eyes of different users. Therefore, giving each service only a single reputation value seems to be inaccurate. Moreover, when the number of feedback ratings is low or zero, it is very difficult to give a fair or accurate reputation value to these services.

In this section, we propose the system architecture of our approach. As shown in Figure 1, when a user invokes a multimedia service, the user reports a feedback rating for the service regarding its performance. The reputation system collects the feedback rating and its user-context data with a collector, calculates the reputation (scores) by using user-context computing, user-similarity computing, and reputation measurement, updates these scores, and provides the scores when other users want to use the multimedia service or newly created services. Note that, in this system architecture, if one multimedia service is used by many users, the collector collects sufficient feedback ratings with different user contexts. Moreover, there are many similar services that are used by other similar users in this system. This system is not effective against malicious feedback ratings [Wang et al. 2015b].

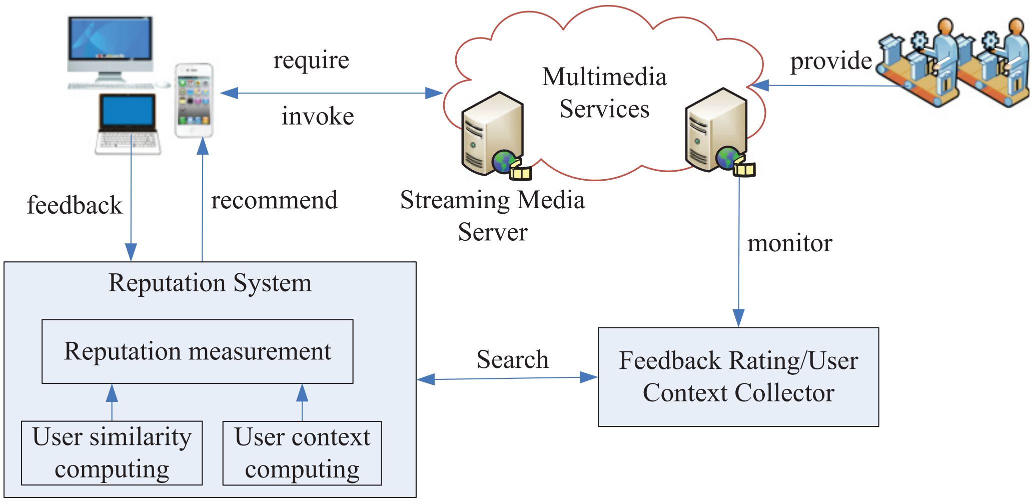

Fig. 1. System architecture of our approach. The main modules of this system are user-context computing, user-similarity computing, and reputation measurement.

Table I. Notations

Symbol	Meaning
C	User context level, and $C = \{L,M,H\}$ contains low user-context level (L), middle user-context level (M), and high user-context level (H)
ms_j	jth used multimedia service
r_i	ith feedback rating
$r_{i,j,C}$	Feedback rating provided by the ith user on using ms_j with the user context level C
$\bar{r}_{i,j}$	Inner feedback rating of the ith user u_i that used the jth service ms_j
Δ_i^{MH}	Difference between $r_{i,j,M}$ and $r_{i,j,H}$
Δ_i^{LM}	Difference between $r_{i,j,L}$ and $r_{i,j,M}$
S_i^{MH}	Set of multimedia services that user u_i has invoked with both middle- and high-context levels
S_i^{LM}	Set of multimedia services that user u_i has invoked with both low- and middle-context levels
$Sim(a, u, C)$	Similarity of two users (a and u) with the user-context level C
K	Number of similar users
$FS^K(a, u, C)$	Feedback similarity of two users (a and u) with the user-context level C
$\bar{r}_{i,j}$	Inherent feedback rating of ms_j from the ith user
$\bar{r}_{n,C}$	Adjusted nth feedback rating with the user-context level C

In order to make it easier to understand our proposed approach, we first introduce user-context computing, including user-context level and user-context sensitivity (Section 3.2). Then, we adopt the PCC to calculate user similarity according to a set of commonly rated multimedia services by other users (Section 3.3). Finally, regardless of the number of feedback ratings, a fair and accurate reputation measurement can be obtained (Section 3.4). Note that the notations in Table I will be used throughout the article.

3.2. User-Context Computing

Definition 1 (*User Context*). User context is the ability of a user terminal and network when a user is requesting or using a multimedia service. In this article, we consider

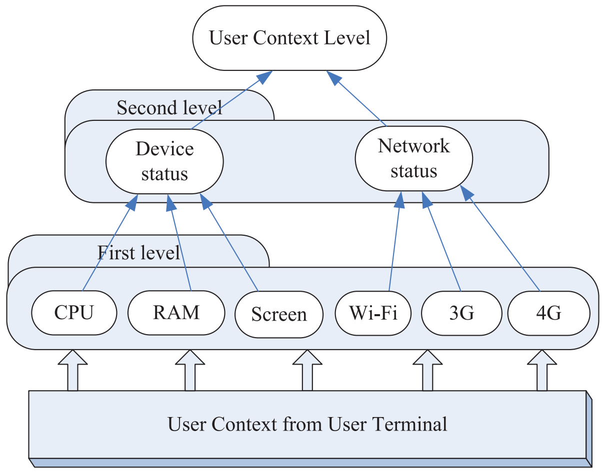

Fig. 2. Hierarchical fuzzy system for calculating user-context level.

the CPU, RAM, screen, and network of the user terminal (3G, 4G, Wi-Fi, and so on) as user-context data.

Because user context is a qualitative concept, in order to measure reputation in different user-context environments, we have to quantify user context. However, the challenge in this task is how to perform the quantization of user context to obtain a comprehensive value to represent a certain user context. In this article, we take user context level to denote the quantization of user context.

3.2.1. User-Context Level.

Definition 2 (*User-Context Level*). User-context level denotes the quantization value of user context. As shown in Figure 2, with the technology of Chuang et al. [2008], the user-context level C can be obtained using a hierarchical fuzzy system by the following four steps:

Step 1 (*Membership*). We set the user-context data as input, with the user-context level as output. In this hierarchical fuzzy system, we adopt a triangular membership function that is specified by three parameters {a, b, c}:

$$f(x; a, b, c,) = \begin{cases} 0, & x \le a \\ \dfrac{x-a}{b-a}, & a \le x \le b \\ \dfrac{c-x}{c-b}, & b \le x \le c \\ 0 & c \le x \end{cases}. \tag{2}$$

Before inputting the user-context data, we adopt the min-max normalization method to map all these data into the same interval [0, 1].

Step 2 (*Fuzzification*). By using the defined membership functions, we translate the input values into a set of linguistic values and assign a membership degree to each linguistic value.

Step 3 (*Inference*). A fuzzy rule can be defined as a conditional statement in the form: IF x is A, THEN y is B, where x and y are linguistic variables including; and A and B are linguistic values determined by fuzzy sets on the universe of discourses X

Table II. Example of the Fuzzy Rules Defined

ID	Rule
1	If (Device is L) and (Network is L), then (ucl is L).
2	If (Device is L) and (Network is M), then (ucl is L).
3	If (Device is L) and (Network is H), then (ucl is M).
4	If (Device is M) and (Network is L), then (ucl is L).
5	If (Device is M) and (Network is M), then (ucl is M).
6	If (Device is M) and (Network is H), then (ucl is H).
7	If (Device is H) and (Network is L), then (ucl is M).
8	If (Device is H) and (Network is M), then (ucl is H).
9	If (Device is H) and (Network is H), then (ucl is H).

and Y, respectively. The inference engine makes decisions based on fuzzy rules with three fuzzy sets: "low (L)," "medium (M)," and "high (H)." Each rule is an IF–THEN [Van Broekhoven and De Baets 2009] clause in nature, which determines the linguistic value of all user-context data. Table II lists the nine most important rules.

Step 4. (*Defuzzification*). Defuzzification transforms the linguistic values of C into crisp values. We adopt the most common defuzzification method, called center of gravity [Van Broekhoven and De Baets 2009].

Through these four steps, the user-context level can be obtained to denote the user context of multimedia services. Then, we adopt the k-means clustering algorithm [Modha and Spangler 2003] to classify all feedback ratings of each multimedia service into p levels (in this article, $p = 3$) according to the user-context level C. Let $C = \{L, M, H\}$ denote a classified low user-context level, a middle user-context level, and a high user-context level, respectively. Thus, we divide all feedback ratings of each multimedia service into k sets according to the range of feedback ratings; then, the feedback ratings of each set have the same user context level.

After obtaining the user-context level, unlike the traditional feedback rating, which is the only value to represent the feedback rating of one user, we take $r'_{i,j,C} = (r_{i,j,C}, C)$ to represent the feedback rating provided by the ith user on using the jth multimedia service with the user-context level C. Then, for the same multimedia service, if the user-context level of one user changes, then the measured reputation should also change.

3.2.2. User-Context Sensitivity. It is well known that, for a given user, different user contexts will result in different feedback ratings. However, a given user context may have different effects on the feedback ratings of different users. For example, for a given service, user u may give a higher feedback rating if u used a faster CPU, and may give a lower rating in the case of a slower CPU. This means that the CPU of the user terminal significantly affects user u. In contrast, another user may not give a much higher feedback rating with a faster CPU than with a slower CPU. This means that the CPU does not significantly affect this user. In this article, we call this case a *user-context sensitivity* problem.

To overcome this problem, we transform the feedback rating of each multimedia service into a vector $r_{i,j,C} = (\bar{r}_{i,j}, CS_{i,j})$. $\bar{r}_{i,j}$ represents the inner feedback rating of the ith user u_i who used the jth service ms_j, which is the user's natural opinion without any context effect. $CS_{i,j}$ relies on user context. For example, for a given multimedia service, a user with high user-context sensitivity will give a high feedback rating with a high user-context level and a low feedback rating with a low user-context level. In contrast, a user who is not very sensitive to user context will likely give similar feedback ratings with different user-context levels. Hence, it is crucial to calculate the user-context sensitivity of each user for accurate reputation measurement.

Definition 3 (*User-Context Sensitivity*). If the ith user u_i invokes the jth multimedia service ms_j three times with different user contexts, then the feedback rating $\bar{r}_{i,j}$ to ms_j that u_i gives is $r_{i,j,L}$ with a low user-context level, $r_{i,j,M}$ with a middle user-context level, and $r_{i,j,H}$ with a high user-context level. We denote the difference between $r_{i,j,L}$ and $r_{i,j,M}$ as Δ_i^{LM}, which represents the difference in feedback rating by user u_i between low and middle user-context levels. Similarly, the difference between $r_{i,j,M}$ and $r_{i,j,H}$ is denoted as Δ_i^{MH}. We call $(\Delta_i^{LM}, \Delta_i^{MH})$ the user-context sensitivity of u_i, which can be calculated by the following:

$$\Delta_i^{LM} = \frac{\sum_{ms_j \in S_i^{LM}} (r_{i,j,M} - r_{i,j,L})}{\left|S_i^{LM}\right|}, \tag{3}$$

$$\Delta_i^{MH} = \frac{\sum_{ms_j \in S_i^{MH}} (r_{i,j,H} - r_{i,j,M})}{\left|S_i^{MH}\right|}, \tag{4}$$

where S_i^{LM} and S_i^{MH} represent the sets of multimedia services that user u_i has invoked with both low and middle user-context levels and with both middle and high context levels, respectively; $r_{i,j,L}$, $r_{i,j,M}$, and $r_{i,j,H}$ are the feedback ratings that user u_i gives to the multimedia service ms_j with a low user-context level, middle user-context level, and high user-context level, respectively; and $|S_i^{LM}|$ and $|S_i^{MH}|$ represent the total number of two multimedia service sets, respectively.

Once the user-context sensitivity of each user context is obtained, it can be used to measure reputation when the feedback rating number is low. It can accurately measure the reputation of a multimedia service by calculating the feedback rating with any user-context level according to the historical feedback ratings with other user-context levels. For example, if one multimedia service has only one feedback rating with a low user-context level, then we can obtain two different feedback ratings with middle and high user-context levels. Then, we can provide an accurate reputation for a user (who wants to invoke the multimedia service) with a high user-context level.

3.3. User-Similarity Computing

User-similarity computing is proposed to find similar users who have used a set of commonly related multimedia services by using the PCC to measure reputation when the number of feedback rings is low or zero.

We assume that there are m users and n multimedia services; the relationship between users and multimedia services is denoted with an $m \times n$ matrix. Each entry $r_{a,i,C}$ in the matrix denotes the feedback rating of the multimedia service i rated by the user a with the user context level C.

The PCC uses the following equation to compute the similarity between user a and user u on the basis of their commonly rated multimedia services:

$$Sim(a, u, C) = \frac{\sum_{i \in I_a, C \cap I_{u,C}} (r_{a,i,C} - \overline{r_{a,C}})(r_{u,i,C} - \overline{r_u, C})}{\sqrt{\sum_{i \in I_{a,C} \cap I_{u,C}} (r_{a,i,C} - \overline{r_{a,C}})^2} \sqrt{\sum_{i \in I_a \cap I_u} (r_{u,i,C} - \overline{r_{u,C}})^2}}, \tag{5}$$

where $I_{a,C} \cap I_{u,C}$ is a set of commonly rated multimedia services by both users a and u, $r_{a,i,C}$ is the feedback rating of multimedia service i rated by the user a, and $\bar{r}_{a,C}$ represents the average feedback rating of all multimedia services rated by user a with user-context level C. A larger $Sim(a, u, C)$ indicates a higher similarity of the two users.

After calculating and ranking the PCC similarity values between the current user and other users, a set of similar $S(a, C)$ with user-context level C can be identified as follows:

$$S(a,C) = \left\{u | Sim(a,u,C) \geq Sim_K, Sim(a,u,C) > 0, a \neq u\right\}, \tag{6}$$

where Sim_K is the Kth largest PCC value with the current user u (K denotes the number of similar users, that is, the K users who have larger PCC values than others will be selected as similar users by setting a parameter K), and $Sim(a,u,C) > 0$ is a condition to prevent dissimilar users (e.g., with negative PCC values) from influencing the reputation measure accuracy.

After obtaining the set of similar users, according to a set of community multimedia services $\mathbb{S}^K$ that are used by all K users, we can calculate the feedback similarity [Wang et al. 2015b] between user a and user u by using

$$FS^K(a,u,C) = \begin{cases} 1 - \sqrt{\frac{\sum_{u \in S(a,C)} (r^K_{a,i,C} - r^K_{u,i,C})^2}{|\mathbb{S}^K|}}, & if \ \ |\mathbb{S}^K| \neq 0 \\ 0 & , \ if \ \ |\mathbb{S}^K| = 0 \end{cases}, \tag{7}$$

where $FS^K(a,u,C) \in [0,1]$ represents the feedback similarity of the two users with the user-context level C and $|\mathbb{S}^K|$ is the number of multimedia services in $\mathbb{S}^K$. A larger $FS^K(a,u,C)$ indicates a higher similarity.

3.4. Reputation Measurement Algorithm

If the total number of feedback ratings of a multimedia service is larger than the threshold ∂, the feedback ratings are sufficient; otherwise, they are limited. Then, we propose a reputation measurement algorithm, as written in Algorithm 1.

In Algorithm 1, if the feedback ratings of the jth multimedia service are sufficient, we can calculate its reputation $R(ms_j, C)$ with the user-context level C by the following:

$$R(ms_j, C) = \frac{1}{n}\sum_{i=1}^{n} r_{i,C}, \tag{8}$$

where $r_{i,C}$ represents the ith feedback rating with the user-context level C of the multimedia service and n represents the total number of feedback ratings with the user-context level C.

If the feedback ratings of the jth multimedia service are limited, we can calculate its reputation $R(ms_j, C)$ by the following three steps:

Step 1. By using user-context sensitivity, we obtain the inherent feedback rating by each user for the multimedia service ms_j by the following:

$$\bar{r}_{i,j} = \begin{cases} r_{i,j,L} + \Delta_i^{LM} + \Delta_i^{MH}, & Low\ user\ context \\ r_{i,j,M} + \Delta_i^{MH} & , Middle\ user\ context \\ r_{i,j,H} & , High\ user\ context \end{cases}, \tag{9}$$

where $\bar{r}_{i,j}$ represents the inherent feedback rating of ms_j; $r_{i,j,H}$, $r_{i,j,M}$, and $r_{i,j,L}$ represent the feedback ratings of ms_j given by the ith user with a high user-context level, middle user-context level, and low user-context level, respectively; Δ_i^{MH} represents the user-context sensitivity of the ith user between middle and high user-context levels; and Δ_i^{LM} represents the user-context sensitivity of the ith user between low and middle high user-context levels.

Step 2. By using Equation (1), we obtain the inherent reputation of ms_j as follows:

$$R(ms_j) = \frac{1}{n}\sum_{i=1}^{n} \bar{r}_{i,j}. \tag{10}$$

ALGORITHM 1: Reputation Measurement Algorithm

Input: $r_{0,C}, r_{i,j,C}, \partial$,
Output: $R(ms_j, C)$
1: **If** the number of feedback ratings $\geq \partial$ \\ *Reputation measurement based on sufficient feedback ratings.*
2: $R(ms_j, C) = \frac{1}{n}\sum_{i=1}^{n} r_{i,C}$;
3: **End if**
4: **If** the number of feedback ratings $< \partial$ \\ *Reputation measurement based on limited feedback ratings.*
5: Calculate $\bar{r}_{i,j}$ by Equation (9);
6: Obtain the inherent reputation $R(ms_j)$ by Equation (10);
7: Obtain the final reputation by Equation (11);
8: **End if**
9: **If** the number of feedback ratings is zero \\ *Reputation measurement based on limited feedback ratings.*
10: Initialize the feedback rating $r_{n,C}$ of ms_j;
11: **Repeat**
12: Calculate $\bar{r}_{u,i,C}$ by Equation (12);
13: **If** the number of newly arrived ratings $> K/2$
14: Discard the created ratings from $K/2$ lowest similar users
15: **End if**
16: Go to 4;
17: **Until** the number of feedback ratings $= \partial - 1$
18: **If** the number of feedback ratings $\geq \partial$;
19: Go to 2;
20: **End if**
21: **End if**

Step 3. Having obtained the inherent reputation of service ms_j without any user-context effect, to provide an accurate reputation measurement for different user-context levels, we need to add the user-context level to the reputation. Then, the final $R(ms_j, C)$ with the user-context level C can be obtained by the following:

$$R(ms_j, C) = \begin{cases} R(ms_j) - \Delta_i^{LM} - \Delta_i^{MH}, & Low\ user\ context \\ R(ms_j) - \Delta_i^{LM} & , Middle\ user\ context \\ R(ms_j) & , High\ user\ context \end{cases}. \tag{11}$$

If feedback ratings are not available for the jth multimedia service, we can calculate its reputation $R(ms_j, C)$ by the following three steps:

Step 1. According to different user-context levels, we assign three average feedback ratings $r_{0,L}, r_{0,M}, r_{0,H}$ to ms_j.

Step 2. When one user a provides the first feedback rating $\hat{r}_{a,i,C}$ for ms_j, we can obtain $K - 1$ ($K < \partial$) similar users by using Equation (6). Then, we can use $FS^K(a, u, C)$ to create the feedback ratings of these similar users by the following:

$$\bar{r}_{u,i,C} = \sum_{u \in S(a)} \frac{FS^k(a, u, C)}{\sum_{u \in S(a)} FS^k(a, u, C)} \times \hat{r}_{a,i,C}, \tag{12}$$

where $\bar{r}_{u,i,C}$ is the created feedback rating for user u.

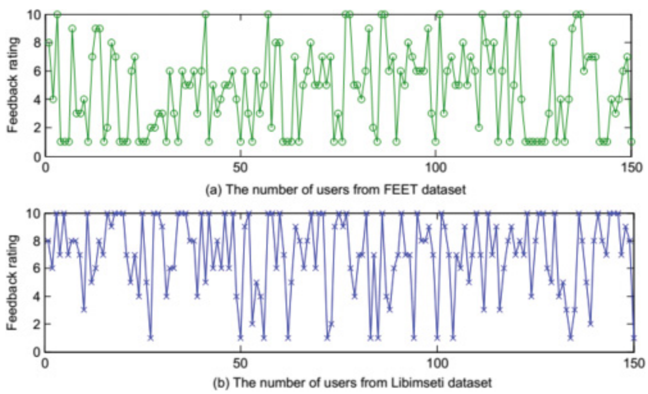

Fig. 3. Parts of feedback ratings from two real datasets.

Step 3. By using Equation (12), we can transform the reputation measurement with no feedback ratings into reputation with limited or sufficient feedback ratings. When the total number of feedback ratings of the multimedia service is not zero and smaller than the threshold ∂, we use Equations (9), (10), and (11) to calculate the reputation. Then, if new feedback ratings arrive, we will replace these created feedback ratings with low feedback similarity with the newly arrived feedback ratings. If the number of newly arrived feedback ratings is larger than $K/2$, all created feedback ratings will be discarded. If the total number of feedback ratings of the multimedia service is larger than the threshold ∂, we can use Equation (8) to calculate the reputation.

Finally, regardless of the number of feedback ratings, our reputation measurement approach can provide an accurate reputation of a multimedia service.

4. PERFORMANCE EVALUATION

Experiments were conducted to evaluate the performance of our proposed approach by using a real-world feedback rating dataset of multimedia services. Moreover, a real feedback rating dataset of an online dating service was also used in the experiments. We also chose to use a simulation to generate feedback ratings because it enabled us to study limited or no feedback ratings for the reputation measurement of multimedia services. Finally, we compared our approach with other approaches [Conner et al. 2009; Ghaffarinejad and Akbari 2013; Malik and Bouguettaya 2009b; Zacharia et al. 2000] in term of accuracy.

4.1. Experiment Setup

For the experiments, we adopted a real feedback rating dataset called FEET, which consists of a user-feedback rating data and user-context data from a real multimedia service feedback system[1]. Overall, the dataset contains 150 users, who provided 200 feedback rating records with 9 attributes, such as username, terminal, screen, CPU, RAM, network, feedback rating, feedback time, and IP. Parts of the feedback ratings are shown in Figure 3(a), and Table III gives the attribute description of the FEET dataset. To the best of our knowledge, the FEET dataset is the largest public dataset that includes feedback rating and user context. Another actual feedback rating dataset (called Libimseti) [Brozovsky and Petricek 2007] consisting of data from a real online dating service was also adopted. Overall, this dataset contains 194,439 users, who provided 11,767,448 feedback ratings. Only users who provided at least 20 feedback ratings were

[1]http://sguangwang.com/projects.htm.

Table III. FEET Attributes and Units

ID	Parameter	Description	Units
1	User name	Name of the user	None
2	Terminal	Device used to run multimedia services by user	None
3	Screen	Screen size of the user terminal	None
4	CPU	Central processing unit of the user terminal	GHz
5	RAM	Random access memory of the user terminal	MB
6	Network	Access network of the user terminal	None
7	Feedback rating	A score given by a user of the multimedia services	None
8	Feedback time	Time at which a user rated the multimedia service	None
9	IP	IP address of the user terminal	None

included. In both datasets, feedback ratings are on a 1 to 10 integer scale, for which "10" is best and "1" is least. Parts of the feedback ratings are shown in Figure 3(b).

It is worth noting that, due to the current limited availability of feedback rating data, many existing reputation systems [Conner et al. 2009; Xiong and Liu 2004; Maarouf et al. 2009; Malik and Bouguettaya 2009b; Li et al. 2014] often used simulation data for performance evaluation. Similarly, this study also employed simulation to generate feedback ratings to evaluate our proposed approach. In order to measure reputation for different feedback ratings with low, middle, and high user contexts, we simulated 200 multimedia services and 500 users. These users provided their feedback ratings on a scale of 1 to 10 with all three user contexts.

Unless otherwise noted, the parameters were set to ($\partial = 20, K = 5$) according to the experimental results in Section 4.3. Note that the two parameters denote only the threshold value. In the experimental comparisons, all test cases and runtime environments were the same. There were three user contexts, that is, low user context (L), middle user context (M), and high user context (H). All results were collected as average values after each method was applied 10 times.

We conducted our experimental results on a PC with an Intel Core 2.0GHz processor and 8.0GB of RAM. The machine was running Windows 8.1, JDK 7.0 and Eclipse 4.3, and MATLAB 7.6. We compared our approach with the reputation measurement approach in Conner et al. [2009], Ghaffarinejad and Akbari [2013], Malik and Bouguettaya [2009b], and Zacharia et al. [2000] in terms of the accuracy of reputation measurement for multimedia services. For illustration purposes, BOOT represents the approach in Malik and Bouguettaya [2009b], which was used to bootstrap the reputation of newly deployed services; RAA represents the approach in Zacharia et al. [2000], which assigned neutral or default reputation values to newly created multimedia services; TMS represents the approach in Conner et al. [2009], which was a trust framework of managing services on the basis of reputation; and DRM represents the approach in Ghaffarinejad and Akbari [2013], which proposed a reputation mechanism based on a number of special reputation centers in service-oriented environments. More details about these approaches can be found in Section 2.

4.2. Experiment on Accuracy

In this experiment, we compared the accuracy of our approach in reputation measurement with BOOT, TMS, and DRM under the conditions of sufficient, limited, and no feedback ratings.

Definition 4 (*Accuracy*). We adopted the widely used mean absolute error (*MAE*) as the accuracy metric for our experiments. MAE is frequently used to measure the difference between values measured by a model or algorithm and actual values. MAE

Table IV. Accuracy Comparisons with Sufficient Feedback Ratings

Approach	Training Data=20%			Training Data=40%			Training Data=60%			Training Data=80%		
	L	M	H	L	M	H	L	M	H	L	M	H
DRM	2.4285	3.4350	3.9249	2.3245	3.1250	3.5387	2.0247	3.1248	3.0002	1.4224	1.247	1.3214
TMS	3.5294	3.4517	3.6234	3.1952	3.2356	3.4925	2.2414	2.6954	2.5491	1.0052	1.052	1.0219
Our Approach	**0.9253**	**0.8513**	**0.7425**	**0.6294**	**0.5864**	**0.6248**	**0.4328**	**0.5200**	**0.3524**	**0.0152**	**0.0954**	**0.0148**

is defined as

$$MAE = \frac{\sum \left| R_{ij} - R_{ij} \right|}{N}, \tag{13}$$

where R_{ij} denotes the actual reputation value (based on all feedback ratings) of multimedia service j for user i, R_{ij} is the measured reputation value (based on a part of feedback ratings or not based on ratings), and N is the number of measured values.

4.2.1. Accuracy with Sufficient Feedback Ratings. Because BOOT and RAA are only used to measure reputation when the number of feedback ratings is zero, in this experiment, we compared the accuracy of reputation measurement with only two other representative approaches, that is, DRM and TMS.

We selected 20%, 40%, 60%, and 80% data with three different user contexts (i.e., L, M, H) as training data; the remainder was used as test data. Training data were used to compute the reputation of multimedia services. The test data were used to verify the accuracy of reputation measurement.

The accuracies of our approach and the comparisons with the other approaches are presented in Table IV. With reference to Table IV, we can see that our approach is the most accurate. As the training data increase from 20% to 80%, the MAE values become smaller.

From Table IV, for the 80% training data, we can also see that the MAE value of our approach is smaller than 1, whereas all the MAE values of DRM and TMS are larger than 1. With increasing training data, although the measured reputations of DRM and TMS are more accurate, they are still less accurate than those of our approach (all their MAE values are larger than 1). The measured reputation scores by DRM and TMS are inaccurate, which masks the actual reputation of the multimedia service and makes the reevaluated multimedia service fail to compete with existing multimedia services for market share. In contrast, our approach works well with different user contexts.

In summary, from Table IV, with different training data sizes, when feedback ratings are sufficient, the accuracy of our approach is much higher than those of the other approaches in all three user-context environments.

4.2.2. Accuracy with Limited Feedback Ratings. In this experiment, we compared the accuracy of reputation measurement with DRM and TMS. The number of feedback ratings of each multimedia service with a certain user context is smaller than ∂.

We selected 2%, 4%, 6%, and 8% data with three different user contexts as training data; the remainder was used as test data. Other settings were similar to the previous comparison described in Section 4.2.1.

The accuracies of our approach and the comparisons with other approaches are presented in Table V. We can see that our approach is again the most accurate with different numbers of feedback ratings and different user contexts.

From Table V, regardless of the training data size, the measured reputations of DRM and TMS are less accurate than those of our approach. Most MAE values of our approach range from 1.0 to 2.0 with three user contexts, whereas almost all MAE values of the other approaches are higher than 2.0. This clearly shows that the measured reputation values by DRM and TMS are less accurate.

Table V. Accuracy Comparisons with Rare Feedback Ratings

	Training Data=2%			Training Data=4%			Training Data=6%			Training Data=8%		
Approach	L	M	H	L	M	H	L	M	H	L	M	H
DRM	4.2591	4.9520	5.6254	5.2846	5.2647	6.2546	4.2654	4.2189	3.9995	3.6548	2.5486	3.4856
TMS	5.0245	5.4623	5.9128	4.9574	6.2154	4.0215	4.8751	4.2868	4.5897	5.2642	3.2457	2.4590
Our Approach	**1.5846**	**1.5483**	**1.9524**	**1.8462**	**1.6240**	**1.2247**	**1.0521**	**1.0093**	**1.1250**	**1.2156**	**1.0029**	**1.3241**

Table VI. Accuracy Comparison with no Feedback Ratings

	NEW=1			NEW=3			NEW=5			NEW=7		
Approach	L	M	H	L	M	H	L	M	H	L	M	H
RAA	6.5845	7.2546	5.6842	6.2142	7.2541	4.5387	6.2514	6.245	4.2156	3.2540	3.0194	2.1548
BOOT	4.5423	5.5710	5.9587	4.0025	5.2590	4.9854	3.2165	2.9517	2.9520	1.3523	1.0548	0.8682
Our Approach	**2.4261**	**2.4654**	**1.9958**	**1.4215**	**1.3286**	**1.4125**	**0.9954**	**0.9085**	**0.9240**	**0.5124**	**0.2659**	**0.3415**

Hence, according to Table V, with different training data sizes, when feedback ratings are limited, the accuracy of our approach still is much higher than those of the other approaches.

4.2.3. Accuracy with No Feedback Ratings. In this experiment, because DRM and TMS cannot support reputation measurement when the number of feedback ratings is zero, we compared the accuracy of reputation measurement with BOOT and RAA.

The total number of multimedia services was 20, and their numbers of feedback ratings were zero. After assigning initial feedback ratings, the number of newly arrived feedback ratings (called NEW) was from 1 to 7, with a step value of 2.

The accuracies of our approach and the comparisons with other approaches are presented in Table VI. From Table VI, we can see that our approach is again the most accurate. As NEW increases from 1 to 7, the MAE values of all approaches become smaller.

From Table VI, when NEW $= 7$, we find that the MAE value of our approach is smaller than 1, whereas all MAE values of the other two approaches are larger than 1. This means that, when the number of feedback ratings is zero and if the number of newly arrived feedback ratings is not smaller than 7, our approach can measure the reputation of each multimedia service accurately.

The results show that even when no feedback ratings are available, our approach can obtain more accurate reputation scores of multimedia services.

4.3. Study of Parameters

We studied the effect of the parameters ∂ and K of our proposed approach on accuracy. Other settings were the same as in the experiments described in Section 4.2.

4.3.1. Effect of the Parameter ∂. Figure 4 shows the effect of the parameter ∂ for our approach, in which we varied the value of ∂ from 2 to 20 with a step value of 2.

We set $K = 4$ in the experiment. The figure was obtained by taking the average of 10 runs. The figure shows that: (1) the MAE values are significantly reduced when the value of α is increased from 2 to 20. This observation indicates that accuracy will increase as the number of feedback ratings increases; (2) accuracy is significantly influenced by the value of the parameter ∂; (3) the best performance of the approach is for values of ∂ larger than approximately 7.

4.3.2. Effect of the Parameter K. Figure 5 shows the effect of the parameter K, in which we varied the value of K from 1 to 10, with a step value of 1. We set $\partial = 5$ in the experiment. As before, the figure is obtained by taking the average of 10 runs. The figure shows that: (1) the MAE values are reduced when the value of K is increased from 1 to 10. This observation indicates that accuracy will increase as the number of

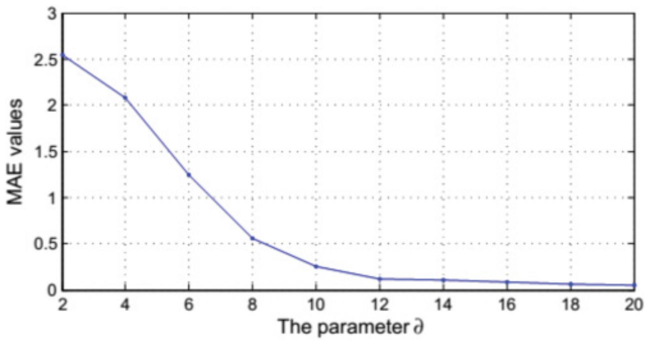

Fig. 4. Effect of the parameter ∂.

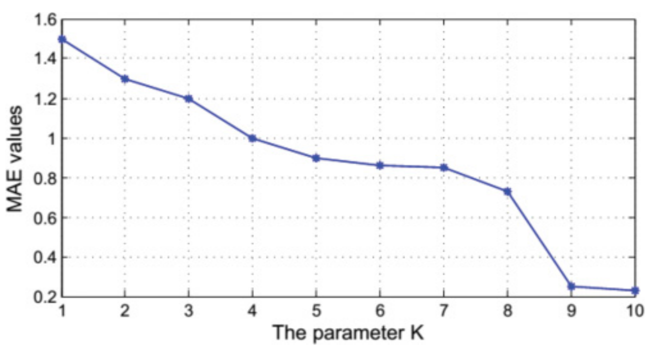

Fig. 5. Effect of the parameter K.

similar users increases; (2) the higher the value of K, the better the performance of the approach, that is, more accurate reputation score; (3) accuracy is not significantly influenced when the value of the parameter K is larger than 8.

4.4. Limitations of Our Proposed Approach

- Our approach may fail when user-context data are limited or not available. With increasing user-context data, the performance of our proposed approach improves.
- If there are limited similar users, our approach is not suitable for reputation measurement when the number of feedback ratings is low or zero. Because the number of similar users is low, our approach cannot create suitable feedback ratings.
- There is a trade-off between the accuracy of reputation measurement and computational load, which is heavy when the number of feedback ratings is zero because of the complex computation.

5. CONCLUSIONS

In this article, we studied the relationship between user context and feedback ratings, and proposed an enhanced user context–aware reputation measurement approach. This approach first models, formalizes, and normalizes user context. Then, it clusters user context in order to use a quantitative value to represent user-context level and user-context sensitivity. Finally, a reputation measurement algorithm is proposed, which achieves accurate measurement in three cases: (1) sufficient existing feedback ratings; (2) limited feedback ratings, and (3) no feedback ratings.

To evaluate our approach, we implemented all approaches based on real feedback rating datasets, and compared our approach with four other approaches. The experimental results showed that our approach can obtain higher accuracy than the other approaches.

In our approach, we assume that each feedback rating is applicable to all three user contexts. Thus, when user-context data are limited, our approach may not be effective. Therefore, our future work will focus on how to tackle limited user-context data.

REFERENCES

Rehab Alnemr, Justus Bross, and Christoph Meinel. 2009. Constructing a context-aware service-oriented reputation model using attention allocation points. In *Proceedings of IEEE International Conference on Services Computing (SCC'09)*. 451–457. DOI: http: //dx.doi.org/10.1109/SCC.2009.55

Pradeep K. Atrey, M. Anwar Hossain, and Abdulmotaleb El Saddik. 2008. A method for computing the reputation of multimedia services through selection and composition. In *Proceedings of Multimedia Computing and Networking Conference (MMCN'08)*. 1–8. DOI: http://dx.doi.org/10.1117/12.775143

Ebrahim Bagheri, Reza Zafarani, and M. Barouni-Ebrahimi. 2009. Can reputation migrate? On the propagation of reputation in multi-context communities. *Knowledge-Based Systems* 22, 6, 410–420. DOI: http://dx.doi.org/10.1016/j.knosys.2009.05.007

Lukas Brozovsky and Vaclav Petricek. 2007. Recommender system for online dating service. In *Proceedings Conference Znalosti*. 1–12. DOI: http://dx.doi.org/10.1.1.188.8232

James Caverlee, Ling Liu, and Steve Webb. 2008. Socialtrust: Tamper-resilient trust establishment in online communities. In *Proceedings of the 8th ACM/IEEE-CS Joint Conference on Digital Libraries (JCDL'08)*. 104–114. DOI: http://dx.doi.org/10.1145/1378889.1378908

Siu-Nam Chuang and Alvin T. S. Chan. 2008. Dynamic QoS adaptation for mobile middleware. *IEEE Transactions on Software Engineering* 34, 6, 738–752. DOI: http://dx.doi.org/10.1109/TSE.2008.44

William Conner, Arun Iyengar, Thomas Mikalsen, Isabelle Rouvellou, and Klara Nahrstedt. 2009. A trust management framework for service-oriented environments. In *Proceedings of the 18th International Conference on World Wide Web (WWW'09)*. 891–900. DOI: http://dx.doi.org/10.1145/1526709.1526829

Mianxiong Dong, Takashi Kimata, Komei Sugiura, and Koji Zettsu. 2014. Quality-of-experience (QoE) in emerging mobile social networks. *IEICE Transactions on Information and Systems* 97, 10, 2606–2612. DOI: http://dx.doi.org/10.1587/transinf.2013THP0011

Mianxiong Dong, Xiao Liu, Zhuzhong Qian, Anfeng Liu, and Tao Wang. 2015. QoE-ensured price competition model for emerging mobile networks. *IEEE Wireless Communications* 22, 7, 50–57. DOI: http://dx.doi.org/10.1109/MWC.2015.7224727

Ali Ghaffarinejad and Mohammad K. Akbari. 2013. An incentive compatible and distributed reputation mechanism based on context similarity for service oriented systems. *Future Generation Computer Systems* 29, 3, 863–875. DOI: http://dx.doi.org/10.1016/j.future.2012.03.021

Sepandar D. Kamvar, Mario T. Schlosser, and Hector Garcia-Molina. 2003. The eigentrust algorithm for reputation management in P2P networks. In *Proceedings of the 12th International Conference on World Wide Web (WWW'03)*. 640–651. DOI: http://dx.doi.org/10.1145/775152.775242

Dejan Kovachev, Yiwei Cao, and Ralf Klamma. 2014. Building mobile multimedia services: A hybrid cloud computing approach. *Multimedia Tools and Applications* 70, 2, 977–1005. DOI: http://dx.doi.org/10.1007/s11042-012-1100-6

Alexandre G. Lages, Luci Pirmez, Paulo F. Pires, and Flávia C. Delicato. 2007. Satya: A reputation-based approach for service discovery and selection in service oriented architectures. In *Proceedings of the 9th Annual ACM International Workshop on Web Information and Data Management (Widm'07)*. 153–160. DOI: http://dx.doi.org/10.1145/1316902.1316927

Jinhwan Lee, Jing Zhang, Zhengqiu Huang, and Kwei-Jay Lin. 2012. Context-based reputation management for service composition and reconfiguration. In *Proceedings of IEEE 14th International Conference on Commerce and Enterprise Computing (CEC'12)*. 57–61. DOI: http://dx.doi.org/10.1109/CEC.2012.18

JooYoung Lee and Jae C. Oh. 2013. A model for recursive propagations of reputations in social networks. In *Proceedings of IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM'13)*. 666–670. DOI: http://dx.doi.org/10.1145/2492517.2492663

Wei Li, Qibo Sun, and Shangguang Wang. 2014. Context-based web service reputation measurement. In *Proceedings of IEEE 17th International Conference on Computational Science and Engineering (CSE'14)*. 1489–1496. DOI: http://dx.doi.org/10.1109/CSE.2014.277

Ling Liu and Weisong Shi. 2010. Trust and reputation management. *IEEE Internet Computing* 14, 5, 10–13. DOI: http://dx.doi.org/10.1109/MIC.2010.124

Eng Keong Lua, Ruichuan Chen, and Zhuhua Cai. 2011. Social trust and reputation in online social networks. In *Proceedings of IEEE 17th International Conference on Parallel and Distributed Systems (ICPADS'11)*. 811–816. DOI: http://dx.doi.org/10.1109/ICPADS.2011.123

Ismat Maarouf, Uthman Baroudi, and Abdurahim R. Naseer. 2009. Efficient monitoring approach for reputation system-based trust-aware routing in wireless sensor networks. *IET Communications* 3, 5, 846–858. DOI: http://dx.doi.org/10.1049/iet-com.2008.0324

Zaki Malik and Athman Bouguettaya. 2009a. RATEWeb: Reputation assessment for trust establishment among web services. *VLDB Journal* 18, 4, 885–911. DOI: http://dx.doi.org/10.1007/s00778-009-0138-1

Zaki Malik and Athman Bouguettaya. 2009b. Reputation bootstrapping for trust establishment among web services. *IEEE Internet Computing* 13, 1, 40–47. DOI: http://dx.doi.org/10.1109/MIC.2009.17

Michael E. Maximilien and Munindar P. Singh. 2001. Reputation and endorsement for web services. *SIGecom Exchanges* 3, 1, 24–31. DOI: http://dx.doi.org/10.1145/844331.844335

Michael E. Maximilien and Munindar P. Singh. 2002. Conceptual model of web service reputation. *SIGMOD Record* 31, 4, 36–41. DOI: http://dx.doi.org/10.1145/637411.637417

Dharmendra S. Modha and Scott W. Spangler. 2003. Feature weighting in k-means clustering. *Machine Learning* 52, 3, 217–237. DOI: http://dx.doi.org/10.1023/A:1024016609528

Ester Van Broekhoven and Bernard De Baets. 2009. Only smooth rule bases can generate monotone Mamdani-Assilian models under center-of-gravity defuzzification. *IEEE Transactions on Fuzzy Systems* 17, 5, 1157–1174. DOI: http://dx.doi.org/10.1109/TFUZZ.2009.2023328

Ping Wang, Kuo-Ming Chan, Chi-Chun Lo, and Yinsheng Li. 2008. QoS-aware service selection under consumer's reputation. In *Proceedings of IEEE International Conference on E-Business Engineering (ICEBE'08)*. 627–632. DOI: http://dx.doi.org/10.1109/ICEBE.2008.22

Shangguang Wang, Lin Huang, Ching-Hsien Hsu, and Fangchun Yang. 2015a. Collaboration reputation for trustworthy web service selection in social networks. *Journal of Computer and System Sciences* DOI: http://dx.doi.org/10.1016/j.jcss.2015.06.009

Yan Wang and Kwei-Jay Lin. 2008. Reputation-oriented trustworthy computing in e-commerce environments. *IEEE Internet Computing* 12, 4, 55–59. DOI: http://dx.doi.org/10.1109/MIC.2008.84

Shangguang Wang, Qibo Sun, Hua Zou, and Fangchun Yang. 2011. Reputation measure approach of web service for service selection. *IET Software* 5, 5, 466–473. DOI: http://dx.doi.org/10.1049/iet-sen.2010.0077

Shuhui Wang, Zhengguang Zheng, Zhisheng Wu, Michael R. Lyu, and Fei Yang. 2015b. Reputation measurement and malicious feedback rating prevention in web service recommendation systems. *IEEE Transactions on Services Computing* 8, 5, 755–767. DOI: http://dx.doi.org/10.1109/TSC.2014.2320262

Shiting Wen, Qing Li, Lihua Yue, An Liu, Chaogang Tang, and Farong Zhong. 2012. CRP: Context-based reputation propagation in services composition. *Service Oriented Computing and Applications* 6, 3, 231–248. http://dx.doi.org/10.1007/s11761-012-0105-3

Li Xiong and Ling Liu. 2004. PeerTrust: Supporting reputation-based trust for peer-to-peer electronic communities. *IEEE Transactions on Knowledge and Data Engineering* 16, 7, 843–857. DOI: http://dx.doi.org/10.1109/TKDE.2004.1318566

Ziqiang Xu, Patrick Martin, Wendy Powley, and Farhana Zulkernine. 2007. Reputation-enhanced QoS-based web services discovery. In *Proceedings of IEEE International Conference on Web Services (ICWS'07)*. 249–256. DOI: http://dx.doi.org/10.1109/ICWS.2007.152

Su-Rong Yan, Xiao-Lin Zheng, Yan Wang, William W. Song, and Wen-Yu Zhang. 2015. A graph-based comprehensive reputation model: Exploiting the social context of opinions to enhance trust in social commerce. *Information Science* 318, 10, 51–72. DOI: http://dx.doi.org/10.1016/j.ins.2014.09.036

Giorgos Zacharia, Alexandros Moukas, and Pattie Maes. 2000. Collaborative reputation mechanisms for electronic marketplaces. *Decision Support Systems* 29, 4, 371–388. DOI: http://dx.doi.org/10.1016/S0167-9236(00)00084-1

Runfang Zhou and Kai Hwang. 2007. PowerTrust: A robust and scalable reputation system for trusted peer-to-peer computing. *IEEE Transactions on Parallel Distributed Systems* 18, 4, 460–473. DOI: http://dx.doi.org/10.1109/TPDS.2007.1021

Received October 2015; revised January 2016; accepted February 2016