Detecting Malicious Feedback Rating using Similarity Computation for Measuring Web Service Reputation

Subodh Kumar¹, Amardeep Kumar²

¹M.tech Scholar, BM College of Technology, Indore, INDIA
²Assistant Professor, BM College of Technology, Indore, INDIA

ABSTRACT

In the field of service computing, reputation of a Web service is usually calculated using feedback ratings provided by service users. However, the existing malicious ratings and the changing preferences of different service users often lead to a bias towards positive or negative ratings. A reputation computation approach is proposed for accurately measuring the reputation of Web services. We first detect malicious feedback ratings using similarity computation and then adjust the malicious rating with the average rating for computing the actual service reputation. Extensive experiments have been conducted based on a real-world Web service dataset to indicate that the proposed approach significantly improves the quality of the recommendation results compared with existing methods. It can also enhance the reliability of service selection.

Keywords--- Web service; Feedback rating; Reputation; Similarity computation

I. INTRODUCTION

Web services are becoming more and more popular to enable organizations to use the Web as a market for selling their own Web services and consuming existing Web services from others. The explosive growth of Web services increases the difficulty for users to choose among a large number of Web services. Therefore, selecting and ranking Web services effectively becomes a key challenge for the Web service community. Traditional Web service selection approaches are usually based on the promised qualities offered by the service providers. However, service providers may make quality promises on the published services but fail partially or fully to deliver on these promises at runtime [1]. It is not an easy task since some service providers may not perform what they guarantee. Reputation of the providers needs to be considered when making selection. Hence, accurate reputation measure of Web services is crucial for business applications. The discovered services can be evaluated and ranked according to the feedbacks from service users.

Several research works have recognized the importance of reputation measure of Web services. Although previous work on existing solutions has explored the efficacy and robustness of measure approaches, most of them suffer from certain weaknesses. Firstly, it is difficult to ensure the purity of feedback ratings because of existing malicious service users. Malicious service users might provide some malicious feedbacks to affect the measure results for commercial benefits. Hence, malicious service users can report malicious feedback ratings for a Web service. There is a large variety of service users on the Internet. Service users can express their preferences over Quality of Service (QoS) attributes of services [1]. The user ratings are often subject to service users’ preferences. As some service users provide low ratings on various Web services, whereas some others may provide higher ratings [7]. Hence, different service users often give different ratings to the same used service. To address the aforementioned weaknesses, an approach is proposed to detect the malicious feedback rating and determine the accurate service reputation for reliable service selection. An algorithm is also presented to show the overall working of proposed mechanism.

This paper is organized as follows: Section II describes literature survey carried out. Section III describes proposed approach for feedback rating detection and reputation computation, which contains malicious rating detection and adjustment. Section IV presents an algorithm for the proposed approach. Section V presents the implementation and experimental evaluation of proposed work. Finally, section VI concludes the paper.

II. LITERATURE SURVEY

Reputation is one of the primary issues among the various business communities. The service users select the provider or services based on their reputation. Many researchers have worked on trust and reputation using various mechanisms, incentive based, association based, previous invocation based etc. In this section,
various reputation computation approaches are identified.

A popularity based service search method is adopted, which consists of tracking all kind of services. It stores important information into database related to discovered, invoked or published services which is helpful in finding frequency of used web services, their availability, usage history etc [3]. The trust among two or more users in the web of trust increases the quality of recommendation in two ways. Firstly, the trust metrics reduce the computability of similarity assessment of users or items. Secondly, the reputation of users may be computed using trust propagation. A novel approach is proposed that dynamically recommends web services that fit users’ interest [4]. This approach is a hybrid approach which combines collaborative filtering and content-based recommendation. In particular, this approach considers simultaneously both rating data and content data of Web services using a three-way aspect model.

An approach for trust-based recommender system is proposed to solve the problem of traditional recommender system which is often ineffective and is not able to compute a user similarity weight for many of the users [5]. It accepts trust metrics and rating matrix as input and generates neighbors using trust metrics and user similarity respectively. In this approach, trust-based issues are discussed to solve the problem of traditional recommender system such as, data sparsity, cold-start users, malicious attacks on recommender systems and centralized architectures. A semantic enhanced personaliser (SEP) framework of hybrid recommender system is proposed for web personalization [6]. SEP comprised of three techniques of recommendation, such as original, semantic and category-based recommendation. The original recommendation consists of three components such as user-based collaborative filtering, item-based collaborative filtering and item-based contextual filtering. This recommendation is based on explicit feedback and contextual information provided by the web users while semantic and category-based recommendation is based on implicit feedback using data mining techniques such as, association-rule-mining (ARM), similarity measures and clustering. The SEP is capable to solve the problem of scalability, sparsity, quality of recommendation, synonymy, etc.

A novel Web service recommendation approach is proposed which incorporates user’s potential QoS preferences and diversity feature of user interests on Web services [8]. User’s interests and QoS preferences on Web services are first mined by exploring the Web service usage history. Then we compute scores of Web service candidates by measuring their relevance with historical and potential user interests, and their QoS utility. A diversified Web service ranking algorithm is also proposed to find the top-k diversified Web service ranked list based on their functional relevance including historical user interest relevance and potential user interest relevance, non-functional relevance such as QoS utility, and diversity feature. In their future work, they will study Web service clustering methods to improve the similarity computation and conduct real user survey to evaluate the usefulness of their method further. In addition, proposed diversified ranking measure mainly focuses on the immediate neighbour- hood information in the Web service graph. More tests will be performed by our diversified ranking measure with k-hop nearest neighbors in the future work.

A QoS prediction framework, called WSPred, is proposed to pro-vide time-aware personalized QoS value prediction for different service users [9]. QoS performance of services is highly related to the service status and network environments which are variable against time. A neighborhood based collaborative filtering approach is presented to predict unknown values for QoS-based service selection [10]. A hybrid service recommendation approach is presented by combining collaborative filtering with content-based features of Web services [13]. This approach exploited both rating data and content data of services using a three-way aspect model. In their work, user interests are represented by a set of latent variables, which is developed offline. However, QoS preferences of users are not considered in these works.

Although previous work on existing solutions has explored the efficacy and robustness of measure approaches, most of them suffer from the certain weaknesses. Firstly, it is difficult to ensure the purity of feedback ratings because of existing malicious service users. Malicious service users might provide some malicious feedbacks to affect the measure results for commercial benefits. Secondly, previous approaches fail to ensure the accuracy of feedback ratings. There is a large variety of service users on the Internet. Service users can express their preferences over Quality of Service (QoS) attributes of services [10]. The user ratings are often subject to service users’ preferences. As some service users are conservative (tend to provide low ratings on various Web services), whereas some others may be aggressive or neutral [7]. Hence, different service users often give different ratings to the same used service.

In order to address aforesaid issues, in this paper an approach is proposed for detecting the malicious users’ feedback rating and computing the service reputation. An algorithm is also presented for the proposed approach. The proposed approach detects malicious feedback ratings by using cumulative sum method and reduces the effect of user feedback preferences by Pearson correlation coefficient. The limitation with this approach is that only negative feedback ratings are dropped because they cause malicious ratings.

III. USER RATING BASED REPUTATION COMPUTATION APPROACH

Reputation of a service can also be evaluated according to feedback ratings provided by service users.
A collaborative filtering based approach is proposed in which user feedback plays a vital role in the assessment of reputation prediction accuracy of web services. Therefore, feedback ratings provided by service users cannot be completely avoided. But the major issue with user feedback rating is the trustworthiness and accuracy of their feedback. This issue can be solved by providing an automatic mechanism for detecting malicious user feedback which assesses the service reputation by discarding the malicious users.

In this mechanism, feedback data is separated into two parts. The first part contains all the malicious feedback ratings and the second part contains the actual feedback ratings. It can be achieved by setting up the maximum and minimum threshold value for malicious and normal feedback ratings. By comparing both the parts, if the majority of users provided negative feedback rating then the service reputation is computed based on negative feedback ratings and considered to be of poor reputation. If majority of users provides actual feedback ratings then the service reputation is computed based on actual feedback ratings. Here, the simple averaging method is used to compute the service reputation based on negative or actual user feedback ratings. If the service reputation is computed on the basis of actual feedback ratings then this computed value is compared against the ideal reputation value in order to exclude or include in composition. In this phase, only the feedback ratings from trustworthy users are considered for reputation assessment.

Here, manual user feedback is collected from group of consumers for their usage experience with the service. A log file is maintained for user feedback data, with columns user_id and user_feedback rating.

Feedback rating is any numeric value between 1 to 5, in which 1 means poor service, 2 means average, 3 means good, 4 means very good and 5 means excellent. Here, feedback data from 15 consumers are collected and analyzed to determine whether it contains any malicious user rating or not. It is necessary to consider only that users rating, which are accurately provided in order to improve the accuracy of reputation of a service.

The detection of malicious user rating for online dating service, the feedback data is analyzed to know the frequency of positive or negative malicious rating provider. For example, if a service user provides the feedback of a service as excellent and majority of users provide the ratings as poor, then the malicious positive rating is discarded for further computation of reputation of a service to predict and improve reputation accuracy. This phase works in two steps, first step identifies the list of non malicious users by performing similarity computation of previous users feedback and second step compares the percentage of non malicious users with the specified threshold value.

Similar feedback ratings are identified with similarity computation of previous user's feedback using Pearson correlation coefficient (PCC). Similar feedback ratings are stored in similarity coefficient variable. We assume that there are m users and n web services, and the relationship between users and web services is denoted with an m*n matrix. Then each entry rx,i in the matrix denotes the feedback rating of the web service i rated by the user x where rx,i is a normal feedback rating. It uses following equation to compute the similarity between user x and user y based on their commonly rated web services:

\[
Sim(x,y) = \frac{\sum_{i \in I_x \cap I_y} (r_{x,i} - \bar{r}_x)(r_{y,i} - \bar{r}_y)}{\sqrt{\sum_{i \in I_x} (r_{x,i} - \bar{r}_x)^2} \cdot \sqrt{\sum_{i \in I_y} (r_{y,i} - \bar{r}_y)^2}}
\]

Where, Sim(x,y)[{-1,1}] represents the similarity of two users.
I_x \cap I_y is a set of commonly rated web services by both users a and u.
\(r_{x,i}\) and \(r_{y,i}\) are the two feedback ratings of web service i rated by user a and u.
\(r_c\) represents the average feedback rating of all of the web services that are rated by user x with values in the set of all \(\{y | \text{sim}(x,y) = \text{sim}_k\}\) where \(x, y \neq x\).

After calculating and ranking the PCC similarity values between the current user and the other users, a set of similar users S(x) can be identified, as follows:

\[
FS_k(x,y) = \begin{cases} 
1 - \frac{\sum_{i \in I_x \cap I_y} (r_{x,i} - \bar{r}_x)(r_{y,i} - \bar{r}_y)}{|SS_k|}, & \text{if } |SS_k| \neq 0 \\
0, & \text{if } |SS_k| = 0 
\end{cases}
\]

Where, \(\text{sim}_k\) is the kth largest PCC value with the current user y.
\(k\) represents the number of the similar users (i.e., they have larger PCC values than others), and
Sim(x,y)\(>0\) is to exclude the dissimilar users (dissimilar users, e.g., those with negative PCC value, will influence the reputation measurement accuracy).

After obtaining the set of similar users, according to a set of community web services SS_k = \{s_1, s_2, …, s_{1k} \} which contains 1 services used by the \(K\) users, we can calculate the feedback similarity between user a and user u as:

Copyright © 2016. Vandana Publications. All Rights Reserved.
IV. PROPOSED ALGORITHM

This algorithm detects the malicious user rating and evaluates the service reputation based on positive feedback rating. Here, \( R \) is the user feedback rating, \( R_l \) is the lowest rating level. \( R_{avg} \) is the average of all rating given by user for particular service; \( R_{th} \) is the maximum no. of user given that rating called threshold rating. List1 and list2 are the list for storing feedback ratings for actual and malicious users.

**Begin**

**Step 1:** Input:
- \( R_L \) = Input given by user for Qualify for rating of services i.e lower rating
- \( R \) = User rating for particular service.

Calculate:
- \( R_{avg} \) = Average of all rating given by user for particular service i.e \( R_{avg}(R) \).
- \( R_{th} \) = Maximum no. of user given that rating called threshold rating.
- \( N_{OpAboveRth} \) = Total no. of people which has given the rating equal and above to \( R_{th} \).
- \( N_{OpBelowRth} \) = Total No. Of People Which Has Given Below The \( R_{th} \).

**Step 2:**
- If \( R \geq R_{th} \) then
  - \( N_{OpAboveRth}++ \)
- Else
  - \( N_{OpBelowRth}++ \)

**Step 3 :**
- If \( (N_{OpAboveRth} \geq 80\%) \) And \( (R_{th} \geq R_L) \) Then
  - \( List1[] = \) No. Of Positive User .
- Else
  - \( List2[] = \) No. Of Malicious User.

Else
- If \( R_{avg} > R_L \) Then
  - Final Rating = \( R_{avg} \) (It is Avg Rating)
- Else
  - Final Rating = \( R_L \) (Below Your Expectation)

**Step 4 :** Separate the List1[] for Positive User
Separate the List2[] for Malicious User

**Step 5:**
- [Reputation Computation ]
- Calculate Average Of Positive User List1 i.e Avg[List1].

**End.**

V. IMPLEMENTATION AND EXPERIMENTAL EVALUATION

The proposed algorithm is implemented using Microsoft dynamic NAV 2009 and SQL Server 2008 R2 tools. Fig. 1 shows the login interface through which the service provider can register their services. Fig. 2 shows the service rating interface through which the service consumer can rate the particular service if it fulfils their desired requirements. With the help of this interface the service consumer can browse the service to provide its rating as excellent, very good, good, average and poor. This rating can further be helpful to determine the final web service reputation. Fig. 3 shows the interface for service reputation assessment through which one can compute the service reputation without including the malicious rating.

Online dating dataset consists of data from real online dating service of 194439 users for providing 11767448 feedback ratings. The dataset consists of three fields, such as user id, profile id and ratings. User id is user who provided rating, profile id is user who has been rated and ratings are on a 1-10 scale. Every feedback rating is limited to natural number from 1 to 10 where feedback rating 1 and 10 represents week and strong feedback ratings.
A set of experiments have been conducted on proposed approach to improve the accuracy of service reputation. To measure the performance of proposed approach, it is compared with the approaches [14-16]. Proposed approach uses real world datasets to assess the service reputation based on user feedback rating. This dataset consists of user feedback ratings from online dating services [16].

This experiment is performed to compare proposed approach with existing approaches on the basis of deviation of actual reputation from the ideal reputation to determine the accuracy of user feedback. In this experiment, existing approaches that uses user feedback rating for reputation assessment of web services are compared with the proposed one in order to determine the improved reputation assessment and prediction accuracy [10][13][1][4]. As the existing approaches use user feedback ratings for the service reputation assessment without excluding the malicious feedback ratings, proposed approach detects malicious feedback ratings and computes the service reputation by discarding it. The performance of proposed approach in measuring the reputation accuracy is better than the existing approaches as it does not consider malicious ratings for reputation assessment.

In Figure 4, the ideal reputation is represented by ideality line through which the performance of a service can be effectively evaluated. Figure 4 shows the results of service reputation assessment in which the malicious ratings are discarded to provide improved reputation accuracy. Figure 4 shows the increased reputation accuracy in proposed approach from existing approaches. This means that smaller the deviation is, the approach is more close to the ideality line. As shown in Figure 4, the deviation in proposed approach is smaller than the existing approaches from ideality line.
VI. CONCLUSION

Reputation is one of the major concerns among the online as well as offline business community. User feedback rating plays vital role in assessing the web service reputation. We have addressed the issue of purity of user feedback rating which may be maliciously provided by some users in order to provide bad reputation of a web service. In this paper, we have proposed an approach for assessing web service reputation to address the limitations of existing reputation computation approaches which uses malicious feedback ratings. An algorithm is also presented for the proposed approach. Implementation and experimental evaluation is also performed to show the efficiency of proposed approach. The approach is able to detect the malicious feedback rating and assess the accurate web service reputation for reliable web service discovery.

REFERENCES