DETECTION AND PREVENTION OF MALICIOUS FEEDBACK RATING IN WEB SERVICE RECOMMENDATION SYSTEM

[1]Ms.D.Ramya, P.Manjula, K.Preethi and S.S.Monica
[1]Assistant Professor, Department of Computer Science and Engineering
Sri Krishna College of Engineering and Technology,
ramya@skcet.ac.in,saranmanjula1995@gmail.com,preethisues@gmail.com, monicasenvaraj@gmail.com

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 different preferences of different service users often lead to a bias towards positive or negative ratings. In this paper, we propose a novel reputation computation approach for accurately measuring the reputation of Web services. The proposed approach uses three phases to measure and compute service reputation such as, feedback data collection, malicious rating detection and reputation computation for improving the reputation computation accuracy. A framework is also proposed to detect the malicious feedback rating and determine the accurate service reputation for reliable service selection.

Index Terms – Web service,malicious feedback, Cumulative Sum Control Chart, Pearson Correlation Coefficient.

1.INTRODUCTION

Web service technologies create an environment where users and applications can search and compose services in an automatic and seamless manner. In the service oriented environment where everybody is allowed to offer services, it is natural that there will be numerous offers of services providing equivalent or similar functionality. Moreover, web services that span diverse organizations and computing platforms can be composed to create new, value-added service-oriented applications efficiently. However, some web services may act maliciously. Hence, a key requirement is to provide an effective mechanism in recommending trustworthy services for users.

Web Service recommendation systems can be employed to recommend the optimal web service for satisfying user's requirements. Service recommendation is helpful for users when two or more web services have the same functionality but different quality-of-service (QoS) performance. QoS is defined as a set of non-functional properties, including reputation, response time, reliability, etc. Web service recommendation can provide the user with necessary information to help decide which web service should be selected. Most QoS-aware web service recommendation schemes are based on the qualities promised by service providers. However, service providers may fail partially or fully in delivering the promised quality at runtime. It is not an easy task since some service providers may not fulfill their promised service quality. The reputation of web service needs to be considered when making a service selection.

Web service reputation is regarded as a metric of its future behavior. It is a collective measurement of the opinions of a community of users regarding their actual experience with the web service. It is computed as an aggregation of users' feedback ratings over a specific period of time (a sample interval) and reflects the reliability, trustworthiness, and credibility of the web service and its provider. With the web service reputation taken into consideration, the probability of recommending the optimal service and the success ratio of the composite services can be increased. However, as it is not realistic to assure that the user feedback ratings are fairly accurate and non-malicious, several studies have recognized the importance of reputation measurements of web services. The proposed solutions employ different techniques to measure web service reputations 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. First, it is difficult to ensure the purity of user feedback ratings because of the existence of malicious users. Malicious users could provide malicious feedback ratings to impair the measurement results for commercial benefit. In open service-oriented environments, there are no widely employed user verification mechanisms. Participating users are usually represented by a pseudonym. In such environment, a special threat comes from Sybil attacks. This attack allows a single malicious user to be represented by an arbitrary number of forged users. Hence, malicious users can initiate a flood of malicious feedback ratings to subvert the reputation system of web services. Second, previous approaches fail to ensure the accuracy of feedback ratings. There is a large variety of users on the Internet. Users have different feedback rating styles. Different users often give different feedback ratings to the same service. For a reputation mechanism to be fair and objective, it is essential to measure reputation on the basis of fair and objective feedback ratings. Finally, most previous research focused on various feedback rating aggregation schemes of reputation measurement, and little work investigated preventing malicious feedback ratings. If the web service recommendation system cannot prevent malicious feedback ratings, any effective reputation

measurement approach will become invalid since these malicious feedback ratings suppress benign feedback ratings.

Hence, an effective malicious feedback rating prevention scheme is very essential for the reputation measurement of web services. In our previous work, we briefly analyze the importance of a reputation measurement in service computing, which lacks of deep research on reputation measurement and malicious feedback rating prevention. To address these weaknesses, this paper extends our previous work by proposing a reputation measurement approach to reduce the deviation of the reputation measurement of web services and to improve the success ratio of the service recommendation.

Moreover, to prevent malicious users from suppressing benign feedback ratings, this paper presents a malicious feedback rating prevention scheme. This paper makes the contributions: 1) we adopt the cumulative sum control chart (called CUSUM) to identify malicious feedback ratings to lessen the influence of malicious feedback ratings on the trusted reputation measurement; 2) we devise feedback similarity computation to shield the different preferences in feedback ratings of users using the Pearson Correlation Coefficient (PCC); 3) we propose a malicious feedback rating prevention scheme to prevents malicious users from suppressing benign feedback ratings using a standard Bloom filter; 4) we validate our proposed malicious feedback rating prevention scheme through theoretical analysis, and also evaluate our proposed measurement approach experimentally on a real feedback rating data set involving 1.5 million real-world web service invocation records.

2.RELATED WORKS

Web content mining is the mining, extraction and integration of useful data, information and knowledge from Web page content. The heterogeneity and the lack of structure that permits much of the ever-expanding information sources on the World Wide Web, such as hypertext documents, makes automated discovery, organization, and search and indexing tools of the Internet and the World Wide Web such as Lycos, Alta Vista, WebCrawler, ALIWEB, MetaCrawler, and others provide some comfort to users, but they do not generally provide structural information nor categorize, filter, or interpret documents. In recent years these factors have prompted researchers to develop more intelligent tools for information retrieval, such as intelligent web agents, as well as to extend database and data mining techniques to provide a higher level of organization for semi-structured data available on the web. The agent-based approach to web mining involves the development of sophisticated AI systems that can act autonomously or semi-autonomously on behalf of a particular user, to discover and organize web-based information.

Web content mining is differentiated from two different points of view: Information Retrieval View and Database View summarized the research works done for unstructured data and semi-structured data from information retrieval view. It shows that most of the researches use bag of words, which is based on the statistics about single words in isolation, to represent unstructured text and take single word found in the training corpus as features. For the semi-structured data, all the works utilize the HTML structures inside the documents and some utilized the hyperlink structure between the documents for document representation. As for the database view, in order to have the better information management and querying on the web, the mining always tries to infer the structure of the web site to transform a web site to become a database.

There are several ways to represent documents; vector space model is typically used. The documents constitute the whole vector space. This representation does not realize the importance of words in a document. To resolve this, tf-idf (Term Frequency Times Inverse Document Frequency) is introduced.

By multi-scanning the document, we can implement feature selection. Under the condition that the category result is rarely affected, the extraction of feature subset is needed. The general algorithm is to construct an evaluating function to evaluate the features. As feature set, Information Gain, Cross Entropy, Mutual Information, and Odds Ratio are usually used. The classifier and pattern analysis methods of text data mining are very similar to traditional data mining techniques. The usual evaluative merits are Classification Accuracy, Precision, Recall and Information Score. Web mining is an important component of content pipeline for web portals. It is used in data confirmation and validity verification, data integrity and building taxonomies, content management, content generation and opinion mining.

3.EXISTING SYSTEM

As it is not realistic to assure that the user feedback ratings are fairly accurate and non-malicious, several studies have recognized the importance of reputation measurements of QOS Web services.

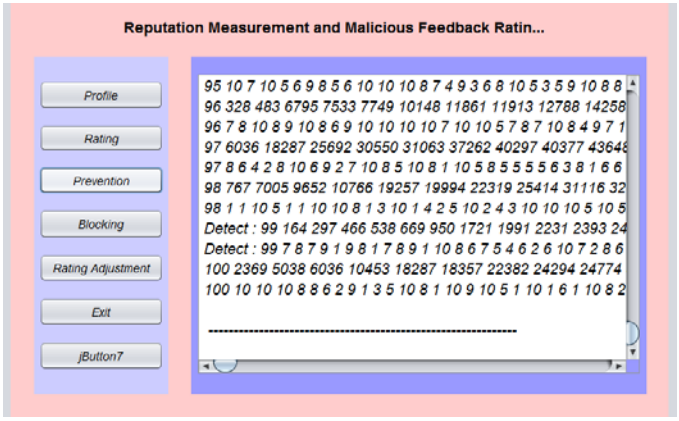
To provide accurate reputation measurement for Web service recommendation, some notable reputation measurement schemes have been proposed. To proposed a reputation-based trust management framework that supports the synthesis of trust-related feedback ratings from multiple services that are hosted within an infrastructure. The core of the framework is a trust management service (TMS). The feedback computation model, derived from the expectancy disconfirmation theory from market science, was used to generate a feedback from service utility and cost, and then a reputation derivation model had also been proposed to

aggregate feedbacks into a reputation value that better reflects the behavior of the service at selection time.

4.PROPOSED SYSTEM

We proposed a reputation measurement approach to reduce the deviation of the reputation measurement of Web services and to improve the success ratio of the service recommendation. The reputation represents a collective perception of the users in the community about a web service, that is, the reputation of a given service is a collective feedback rating of the users that have interacted with or used the service in the past. To prevent malicious feedback ratings from reaching the QoS repository of service brokers, we propose a malicious feedback rating prevention scheme. Its aim is to cooperate with the proposed reputation measurement approach to enhance the performance of the recommendation system.

It makes the contributions: 1) we adopt the Cumulative Sum Control Chart to identify malicious feedback ratings to lessen the influence of malicious feedback ratings on the trusted reputation measurement; 2) we devise feedback similarity computation to shield the different preferences in feedback ratings of users using the Pearson Correlation Coefficient; 3) we propose a malicious feedback rating prevention scheme to prevents malicious users from suppressing benign feedback ratings using a standard Bloom filter; 4) we validate our proposed malicious feedback rating prevention scheme through theoretical analysis, and also evaluate our proposed measurement approach experimentally.



5.PERFORMANCE EVALUATION

- **QOS Recommendation**

The user sends a service request to the recommendation system. With a service level agreement (SLA) between a user and a service provider, the user selects a web service that satisfies his QoS requirements and then invokes the service. After the service is consumed, the user reports a feedback rating for the service regarding the performance of the web service.

- **Reputation Measure**

The reputation measurement approach that is based on a feedback rating evaluation for the web service recommendation Reputation measure approaches, the trust an entity has in another entity is usually linked to a pseudonym that influences the accuracy of the reputation measurement. The reputation represents a collective perception of the users in the community about a web service, that is, the reputation of a given service is a collective feedback rating of the users that have interacted with or used the service in the past.

- **Data Collector**

Data collector, calculates the reputation (scores), updates these scores in a QoS repository, and provides the scores when recommending services to the users. The Data collector is using the cumulative sum control chart.

- **Malicious feedback Ratings**

It contains two phases, i.e., a malicious feedback rating detection and a feedback rating adjustment. The first phase involves detecting malicious feedback ratings collected by a data collector using the cumulative sum control chart. The second phase involves computing the feedback similarity of different users using the Pearson Correlation Coefficient to adjust the feedback ratings. Finally, the repository stores the reputation measured scores and provides the scores when requested by the recommendation system.

- **Malicious blocking**

This identifies the IP addresses with the offending feedback ratings and filters them out. In order to achieve this, we employ a standard Bloom filter to prevent the anomalous feedback ratings. The Bloom Filter of our approach stores each IP address with hash function IP addresses with the offending feedback ratings and block them using a standard Bloom filter.

6.CONCLUSION

The proposed reputation measurement approach utilizes malicious feedback rating detection and feedback similarity computation to measure the reputation of web services. The efficiency of our proposed approach is evaluated and validated by the theoretical analysis and extensive experiments. The experimental results show that our proposed approach can accomplish a trustworthy reputation measurement of web services and greatly improve the service recommendation process. The proposed prevention scheme can identify the IP addresses with the offending feedback ratings and block them using a standard Bloom filter. The theoretical analysis indicates the efficiency of the proposed prevention scheme in blocking malicious feedback ratings within the web service recommendation system. Our on-going research includes investigating the parameters of sampling interval according to the number of feedback ratings, the number of sampling, duration and storage space, and constructing a common malicious feedback rating prevention scheme for web service recommendation systems.

7.REFERENCE

[1] X. Chen, Z. Zheng, X. Liu, Z. Huang, and H. Sun, "Personalized QoS-aware web service recommendation and visualization," IEEE Trans. Serv. Comput., vol. 6, no. 1, pp. 35–47, First Quarter 2013.
[2] Z. Zheng and M. R. Lyu, "Personalized reliability prediction of web services," ACM Trans. Softw. Eng. Methodol., vol. 22, no. 2, pp. 1–28, 2013.
[3] E. M. Maximilien and M. P. Singh, "Conceptual model of web service reputation," ACM SIGMOD Rec., vol. 31, no. 4, pp. 36–41, 2002.
[4] Z. Malik and A. Bouguettaya, "Evaluating rater credibility for reputation assessment of web services," in Proc. 8th Int. Conf. Web Inform. Syst. Eng., 2007, pp. 38–49.
[5] Z. Xu, P. Martin, W. Powley, and F. Zulkernine, "Reputationenhanced QoS-based web services discovery," in Proc. IEEE Int. Conf. Web Serv., 2007, pp. 249–256.
[6] D. Ardagna and B. Pernici, "Adaptive service composition in flexible processes," IEEE Trans. Softw. Eng., vol. 33, no. 6, pp. 369–384, Jun. 2007.
[7] W. Conner, A. Iyengar, T. Mikalsen, I. Rouvellou, and K. Nahrstedt, "A trust management framework for service-oriented environments," in Proc. 18th Int. Conf. World Wide Web, 2009, pp. 891–900.
[8] S. Nepal, Z. Malik, and A. Bouguettaya, "Reputation propagation in composite services," in Proc. IEEE Int. Conf. Web Serv., 2009, pp. 295–302.