Reputation measurement of cloud services based on unstable feedback ratings

Shangguang Wang*, Lei Sun, Qibo Sun, Jie Wei and Fangchun Yang

State Key Laboratory of Networking and Switching Technology,
Beijing University of Posts and Telecommunications,
Haidian, Beijing, 100876, China
Email: sguang.wang@gmail.com
Email: lsun@bupt.edu.cn
Email: qbsun@bupt.edu.cn
Email: jwei@bupt.edu.cn
Email: fcyang@bupt.edu.cn
*Corresponding author

Abstract: With the rapid development of cloud computing, more and more service providers could provide cloud services (applications) to users. Faced with mass cloud services, trust and reputation mechanisms offer a promising way to solve the trust evaluation of cloud services. Hence, trust and reputation play an important role in evaluation of cloud services. In this paper, we propose a lightweight reputation measurement approach for cloud services based on (user) feedback ratings. The proposed approach first adopts cloud model to obtain the trust vector of each cloud service by exploiting feedback ratings. The trust vector consists of expected value, entropy value and hyper-entropy value. Then we use the fuzzy set theory to calculate the reputation scores of cloud services. Simulation results show that the proposed approach is significantly effective for unstable feedback ratings.

Keywords: cloud service; feedback rating; trust; reputation.

Reference to this paper should be made as follows: Wang, S., Sun, L., Sun, Q., Wei, J. and Yang, F. (2015) 'Reputation measurement of cloud services based on unstable feedback ratings', *Int. J. Web and Grid Services*, Vol. 11, No. 4, pp.362–376.

Biographical notes: Shangguang Wang is an Associate Professor at Beijing University of Posts and Telecommunications. He received his PhD degree in Computer Science and Technology at Beijing University of Posts and Telecommunications in 2011. His research interests include service computing and cloud computing. He served as Guest Editor for *Journal of Computational Science*, *EURASIP Journal on Wireless Communications and Networking*, etc. He also served as PC Chair for IOV 2014 and SC2 2014.

Lei Sun is a Postgraduate student at Beijing University of Posts and Telecommunications, China. His research interests include service selection and cloud service.

Qibo Sun is currently an Associate Professor at Beijing University of Posts and Telecommunication, China. He received his PhD in Communication and Electronic System from the Beijing University of Posts and Telecommunication in 2002. He is a Member of the China Computer Federation. His research interests include services computing, Internet of Things and network security.

Jie Wei is a PhD candidate at Beijing University of Posts and Telecommunications, China. His research interests include cloud service and cloud computing.

Fangchun Yang is currently a Professor at Beijing University of Posts and Telecommunications, China. He received his PhD in Communication and Electronic System from the Beijing University of Posts and Telecommunications in 1990. He has published six books and more than 80 papers. His research interests include network intelligence, services computing, communications software, softswitching technology and network security. He is a Fellow of the IET.

This paper is a revised and expanded version of a paper entitled 'Reputation measurement of cloud services based on unstable feedback ratings' presented at the '19th IEEE International Conference on Parallel and Distributed Systems (ICPADS)', 15–18 December 2013, Seoul, South Korea.

1 Introduction

With the increase in the number of cloud service providers, more and more cloud services/applications with different quality have gradually flooded the web (Zibin et al., 2010). For example, according to Amazon, the company's S3 cloud storage service jumped 192% year-over-year at the end of 2011, holding 762 billion objects. That is up from only 2.9 billion objects held when it launched in 2006! Thus, faced with a huge number of cloud services, end users frequently consume or invoke high-reputation (scores) cloud services. Hence, how to accurately measure the reputation of each cloud service is an imperative research problem in cloud computing.

Trust and reputation also play an important role in the trust evaluation of cloud services. It is natural that a cloud consumer would like to choose a service which is of high quality or a service with a high reputation. That is just why trust and reputation mechanisms are used in making good selections for cloud users.

Trust and reputation systems have been recognised as playing an important role in decision-making in the internet world, which are also used in cloud computing systems. The trust and reputation measurement of cloud services has been discussed in some recent literature (Abawajy, 2011; Braithwaite and Woodman, 2011; Brock and Goscinski, 2010; Goscinski and Brock, 2010; Habib et al., 2011; Rehman et al., 2011; Wang et al., 2011b; Wei-Tek et al., 2011; Wu et al., 2009, Zibin et al., 2010), which enables best cloud services to be found according to the QoS performance. For instance, Habib et al. (2011) proposed an architecture of a multifaceted trust management system for cloud marketplaces providing means to efficiently differentiate between good-quality and poor-quality (beyond the performance issues) cloud service providers. The system is designed to provide a customised trust score of a provider based on the attributes selected by end users. This system has a special feature allowing users to select from various sources and roots of trust information as a basis for the computation of the trust scores. The trust score will not be represented only in plain numbers, but the presentation will be supported by an intuitive graphical interface. Abawajy (2011) presented a distributed framework that enables interested parties to determine the trustworthiness of hybrid cloud computing entities. Two algorithms were developed to determine a user's trust value based on a sequence of her interactions and past behaviours. They can detect and filter out dishonest feedbacks by using a personalised similarity measure to compute the credibility of feedback through personalised experience and a variable tolerance threshold.

However, previous reputation measurement schemes (Abawajy, 2011; Braithwaite and Woodman, 2011; Brock and Goscinski, 2010; Goscinski and Brock, 2010; Habib et al., 2011; Rehman et al., 2011; Wang et al., 2011b; Wei-Tek et al., 2011; Wu et al., 2009; Zibin et al., 2010) often assume that all the feedback ratings are stable seriously. Unfortunately, the feedback ratings cannot precisely represent the actual performance of each cloud service used. Why? For example, the feedback ratings from users are often subject to users' preferences. Some users may be conservative, whereas some may be aggressive or neutral (Limam and Boutaba, 2010a). Consequently, different users give different feedback ratings to the same cloud service used. Moreover, some existing dishonest or malicious users might provide some malicious feedback ratings to affect the reputation measurement results for commercial benefits. As a result, the reputation scores of cloud services often deviate from their deserved values, which results in users failing to find the best or the most suitable cloud service. Hence, it is worth noting that a cloud service with consistent feedback ratings is typically more desirable than a cloud service with a large variance on its feedback ratings. Therefore, the consistency of feedback ratings should be considered as an important criterion for the reputation measurement of cloud services.

To provide accurate reputation scores of cloud services for end users, we present a reputation measurement approach for cloud service. Our work is quite different from existing approaches (Abawajy, 2011; Braithwaite and Woodman, 2011; Brock and Goscinski, 2010; Goscinski and Brock, 2010; Habib et al., 2011; Rehman et al., 2011; Wang et al., 2011b; Wei-Tek et al., 2011; Wu et al., 2009; Zibin et al., 2010), since we employ the reputation to evaluate the performance of cloud services. Our approach only collects feedback ratings that do not require cloud service invocation, which will save a lot of time and resources. The core of the scheme is how to accurately measure reputation scores when collected feedback ratings are unstable. In this paper, we adopt expected value, entropy value and hyper-entropy value to represent the trust vector of each cloud service by using a cloud model algorithm. The trust vector is able to show the instability of feedback ratings from users. And then we employ the fuzzy set theory to measure the reputation of cloud services, by inputting expected value, entropy value and hyper-entropy value into fuzzy logic of the fuzzy set theory. To the best of our knowledge, this is the first work that fuzzy logic is used to measure the reputation of cloud services. We evaluate our approach experimentally on the real-world feedback rating data set as well as on randomly generated feedback rating. Simulation results show our approach outperforms other approaches.

The remainder of this paper is organised as follows. Related work is discussed in Section 2. Section 3 describes our proposed reputation measurement approach of cloud services. Section 4 shows the simulation results. Finally, Section 5 concludes the paper.

2 Related work

Some trust and reputation measurement approaches of cloud services (Braithwaite and Woodman, 2011; Brock and Goscinski, 2010; Goscinski and Brock, 2010; Rehman et al., 2011, Wang et al., 2011b; Wei-Tek et al., 2011; Wu et al., 2009; Zibin et al., 2010) have been proposed, which enable best cloud services to be found according to the QoS performance. Our work is quite different from these approaches, since we employ trust and reputation to evaluate the performance of cloud services. Our approach only collects users' feedback that does not require cloud service invocation, which will save a lot of time and resources.

The trust and reputation scheme helps build trust based on users' experiences and feedback (Abawajy, 2011; Gutowska and Buckley, 2008; Habib et al., 2011; Liu and Shi, 2010). Research in the area of trust and reputation systems has put a lot of efforts in developing various trust models and associated trust update algorithms that support users or their agents with different behavioural profiles (Abawajy, 2011). Trust and reputation systems have been recognised as playing an important role in decision-making in the internet world, which are also used in cloud computing systems (Abawajy, 2011; Habib et al., 2011).

Habib et al. (2011) proposed an architecture of a multifaceted trust management system for cloud marketplaces providing means to efficiently differentiate between good-quality and poor-quality (beyond the performance issues) cloud service providers. The system is designed to provide a customised trust score of a provider based on the attributes selected by the cloud customers. In addition, the system aims at providing trust scores of the cloud providers based on trustworthy behaviour of the underlying systems and the service providers' answers to the TVA Consensus Assessment Initiative Questionnaire (https://cloudsecurityalliance.org/may12.html). This system has a special feature allowing cloud users to select from various sources and roots of trust information as a basis for computation of trust scores. The trust score will not be represented in plain numbers only, but the presentation will be supported by an intuitive graphical interface.

Noor et al (2013a) proposed a credibility model that not only identified credible trust feedbacks from fake ones, but also preserved the privacy of cloud service consumers. The model can also tackle the feedback collusion issue by identifying credible trust feedbacks from fake ones, and identify fake trust feedbacks from malicious cloud service consumers who use multiple identities to manipulate trust results. Noor et al (2013b) proposed a generic framework that considered a holistic view of the issues related to trust management for interactions in cloud environments. They in particular differentiate the trust management perspectives and classify trust management techniques into four categories, and compared 30 representative trust management research prototypes in cloud computing and the relevant research areas using the proposed analytical framework.

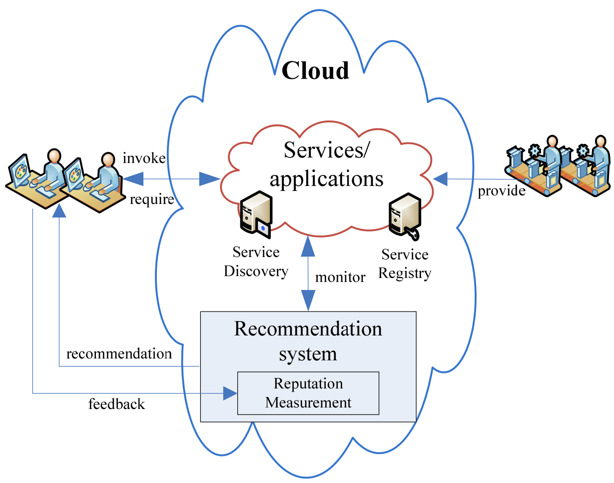
Abawajy (2011) presented a distributed framework that enable interested parties to determine the trustworthiness of hybrid cloud computing entities. The framework contains cloud resource manager, inter-cloud broker and trust manager. The cloud resource manager is responsible for resource provisioning and allocation at the individual cloud level, while the virtual infrastructure engine is the resource manager for the local cluster and can start, pause, resume and stop virtual machines on the physical resources. The inter-cloud broker is responsible for mediating the resource exchange between peering clouds, which also provides external cloud selection capabilities by selecting suitable clouds that are able to provide the required resources to users' requests. The trust manager is responsible for getting users' feedback, verifying users' feedback and updating the same value in the feedback repository. Moreover, two algorithms were developed to determine a user's trust value based on a sequence of his/her interactions and past behaviours. They can detect and filter out dishonest feedback by using a personalised similarity measure to compute the credibility of feedback through personalised experience and a variable tolerance threshold.

There is limited work in the literature that employs trust and reputation methods for reputation evaluation of cloud services, since trust is a concept with many uncertainties, among which, randomness and fuzziness are the two most important uncertainties. Most well-known reputation models use probability or fuzzy set theory to hold randomness or fuzziness, respectively. In order to overcome the problem and fully consider the randomness and fuzziness of the reputation of cloud service, in this paper, we propose an effective and novel reputation measurement approach for cloud services based on the trust and reputation computation.

3 Proposed reputation measurement approach

As shown in Figure 1, in a service-oriented cloud system, a cloud user provides his/her requirements. The requirements may vary from a description of the service and QoS parameters such as the max price and the reputation score. The cloud service providers would need to register their cloud services in Service Registry. These service descriptions will contain the semantic service profile and QoS parameters such as price and response time. The service provider would also be required to describe a cloud service. The Service Discovery matches the user's request with registered service descriptions and provides a list of available cloud services that match the requirements. This list will be given to the recommendation system. The recommendation system, based on its learning through user feedback, orders the list and presents it to the cloud user. The cloud user can then select a cloud service from this list. After the execution is over, the user may provide a feedback rating to this service using a given metric that indicates the user's satisfaction level. The feedback information from cloud users can be collected, published, managed, stored as reputations in a repository and, finally, used as inputs in the recommendation system.

Figure 1 Framework of feedback rating monitoring for cloud services (see online version for colours)



Because reputation influences the recommendation of an interaction partner, some dishonest cloud service providers misuse the system. These cloud service providers might have a direct interest in improving the chances of a certain candidate's selection or in diminishing the chances of others. Moreover, the feedback can be that of an individual because it is based on service users' personal expectations and opinions. Different cloud users that invoke the same cloud service may provide feedback with highly varied ratings. Therefore, the main challenge is addressing services that attempt to provide unstable ratings. Hence, a recommendation system needs appropriate mechanisms for filtering and weighting cloud services with a reputation score.

3.1 Motivation

Feedback rating is the perception of each user about services he/she has invoked. It could be a single value representing an overall perception or a vector representing a value for the overall performance of each cloud service. In this study, for each cloud service s_i invoked, a cloud user provides a feedback rating indicating the level of satisfaction with a service after each interaction with the service. A rating is simply an integer ranging from 1 to 5, where 5 means extreme satisfaction and 1 means extreme dissatisfaction. Then users maintain n feedback ratings representing their perception of s_i's performance (Wang et al., 2011a).

We take $q(s_i)$ to represent the reputation score of s_i in a global time. Then, $q(s_i)$ can be calculated by using

$$q(s_i) = \frac{1}{n}\sum_{i=1}^{n} r_i \tag{1}$$

where r_i represents the ith feedback rating. However, as shown in Table 1, because of the instability of feedback ratings, many studies are limited in their ability to support the reputation measurement of cloud services. They do not cater for the reputation of each cloud service, which makes the reputation of cloud service deviate from its deserved value in the cloud service system.

Table 1 Traditional reputation measurement

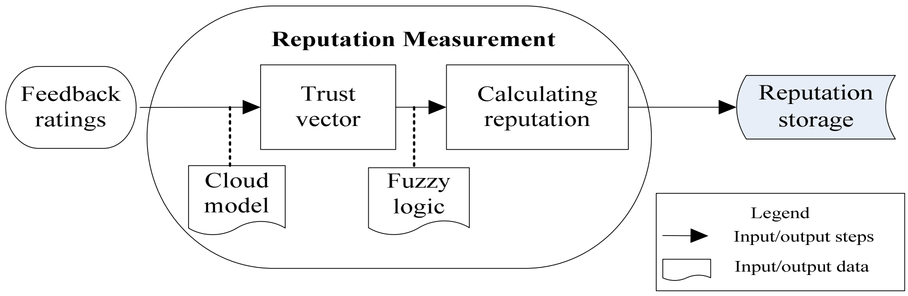
Cloud service: a		*Cloud service: b*	
ID	*Feedback rating*	*ID*	*Feedback rating*
1	9	1	6
2	1	2	5
3	2	3	5
4	10	4	6
5	2	5	5
6	8	6	1
qa	5.3	qb	4.7

As shown in Table 1, we give an example to show the instability of feedback ratings. In Table 1, some feedback ratings from end users are collected. The reputation scores of the two cloud services (*qa* and *qb*), obtained by averaging all feedback ratings according to (1), are given in the last row of Table 1. The reputation score of *a* is larger than that of b, i.e. *qa* > *qb*.

In the traditional approach, cloud service *a* is usually selected because its reputation is high, i.e. 5.3 > 4.7. However, by analysing each feedback rating, we find two disadvantages of the traditional approaches. Firstly, although the reputation of cloud service *b* is slightly lower than that of *a*, the three feedback ratings of cloud service *a* are fewer than that of *b*, i.e. 1 < 5, 2 < 5 and 2 < 5. This means that the feedback rating of cloud service *a* is less than that of cloud service *b* in most users' experience. Secondly, the feedback rating of cloud service *a* is more volatile than that of cloud service *b*. This means that cloud service *b* is with consistently good feedback ratings, and the feedback rating of cloud service *a* may contain malicious or dishonest feedback ratings. Hence, in order to solve the problem, we propose a novel reputation measurement approach for analysing unstable feedback ratings from end users.

As shown in Figure 2, the proposed reputation measurement approach of cloud services contains two phases. The first phase is trust vector, in which we adopt a cloud model to analyse the instability level of feedback rating. The second phase is calculating reputation, in which we adopt fuzzy logic to calculate the reputation score of each cloud service. Eventually, the reputation scores are stored, which is an important criterion for cloud service systems.

Figure 2 Procedures of the reputation measurement approach



3.2 Trust vector

Cloud model (Li et al., 1998) is based on probability. It is a model that contains the transferring procedure of uncertainty between quality concept and quantity data representation by using natural language.

Suppose that U is set as the universe of discourse, and C is a qualitative concept associated with U. The membership degree of quantitative numerical representation x in U to the concept C, $\mu(x) \in [0,1]$, is a random number with a stable tendency, i.e. $\mu: U \rightarrow [0,1]$, $\forall x \in U$, $x \rightarrow \mu(x)$. The distribution of x in the universe of discourse U is called cloud $C(X)$ and x is called a cloud drop. If the distribution of $C(X)$ is normal, it is named normal cloud model. It is a random number set that obeys normal distributive rule and has stable tendency, and it is determined by its three numerical characteristics: expected value (Ex), entropy (En) and hyper-entropy (He).

Ex is the position corresponding to the centre of the cloud gravity whose elements are fully compatible with the linguistic concept; En is a measure of the concept coverage, i.e. a measure of the fuzziness, which indicates how many elements could be accepted to the qualitative linguistic concept; and He is a measure of the dispersion on these cloud drops, which can also be considered as the entropy of *En*. Then, the vector $\{Ex, En, He\}$ is called the eigenvector of cloud model. For $\{x_1, x_2, \ldots, x_n\}$, we can get $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} x_i$ and $S^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(x_i - \bar{X}\right)^2$; $Ex = \bar{X}$; $En = \frac{\sqrt{\pi/2}}{N}\sum_{i=1}^{N}\left|x_i - Ex\right|$; and $He = \sqrt{S^2 - En^2}$.

In this paper, we use the vector $\{Ex, En, He\}$ to denote the trust vector of feedback ratings. Firstly, according to the feedback rating of each cloud service, x_i, we can compute the sample mean $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} x_i$ and the sample variance $S^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(x_i - \bar{X}\right)^2$. Secondly, the expected value of x_i's feedback ratings can be calculated by $Ex = \bar{X}$. Thirdly, the entropy can be calculated by $En = \frac{\sqrt{\pi/2}}{N}\sum_{i=1}^{N}\left|x_i - Ex\right|$. Fourthly, the hyper-entropy can be obtained by $He = \sqrt{S^2 - En^2}$. Finally, the trust vector of feedback ratings of the cloud service can be denoted by an eigenvector (*Ex, En, He*), where *En* and *He* determine the instability level of feedback ratings.

3.3 Calculating reputation

Having obtained the trust vector of feedback ratings of one cloud service, we are able to calculate the reputation score of the cloud service according to fuzzy logic of fuzzy set theory.

Figure 3 Procedures of fuzzy logic for trust computing

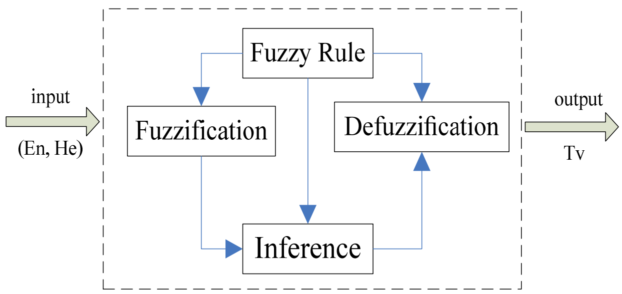
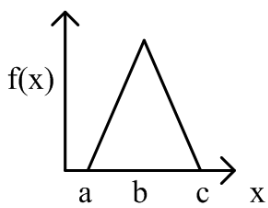


Figure 4 The triangular membership function



As shown in Figure 3, in the second phase, we apply fuzzy logic of fuzzy set theory to obtain the trust value of each cloud service. In this phase, we take *En* and *He* of the trust vector as inputs, with the trust value (*TV*) of the cloud service as output. Then *TV* and *Ex* are used to calculate the reputation of each cloud service. The fuzzy logic mainly contains the following five steps.

1 *Fuzzification*. In this study, as shown in Figure 4, we adopt a triangular membership function. A triangular membership function is specified by three parameters $\{a, b, c\}$:

$$f(x;a,b,c,) = \left\{ \begin{array}{cc} 0, & x \le a \\ \dfrac{x-a}{b-a}, & a \le x \le b \\ & \\ \dfrac{c-x}{c-b}, & b \le x \le c \\ 0 & c \le x \end{array} \right\} \quad (2)$$

By using the defined membership functions, we translate the input values into a set of linguistic values and assign a membership degree for each linguistic value using triangular membership functions.

2 *Fuzzy rule*. A fuzzy rule can be defined as a conditional statement in the following form: IF *x* is *A*, THEN *y* is *B*, where *x* and *y* are linguistic variables; and *A* and *B* are linguistic values determined by fuzzy sets on the universe of discourses *X* and *Y*, respectively. In this study, the fuzzy logic system is represented with three fuzzy sets: 'low (*L*)', 'low medium (*LM*)', 'medium (*M*)', 'high medium (*HM*)' and 'high (*H*)'. These fuzzy sets determine the shape and location of the membership functions. As shown in Figure 5, we give the fuzzy rules of *En*.

3 *Inference*. The inference engine makes decisions based on fuzzy rules. Each rule is an IF-THEN clause in nature, which determines the linguistic value of *TV* according to the linguistic values of the inputs. For example, if (*En* is low) and (*He* is medium), then (*TV* is high). In this study, the inference engine makes decisions based on 25 fuzzy inference rules as shown in Table 2.

4 *Defuzzification* (transforming the linguistic value of *TV* into crisp trust values). We adopt the most common defuzzification method called centre of gravity (Van Broekhoven and De Baets, 2009) to obtain the trust of each cloud service with a value in the range [0, 1].

5 Finally, the reputation score, *RS*, can be obtained by the following:

$$RS = Ex \cdot (1 + TV) \quad (3)$$

In this way, can the reputation of each cloud service be obtained with fuzzy logic. The reputation can help end users to select and avail high-quality cloud services.

Figure 5 The fuzzy rules of *En*

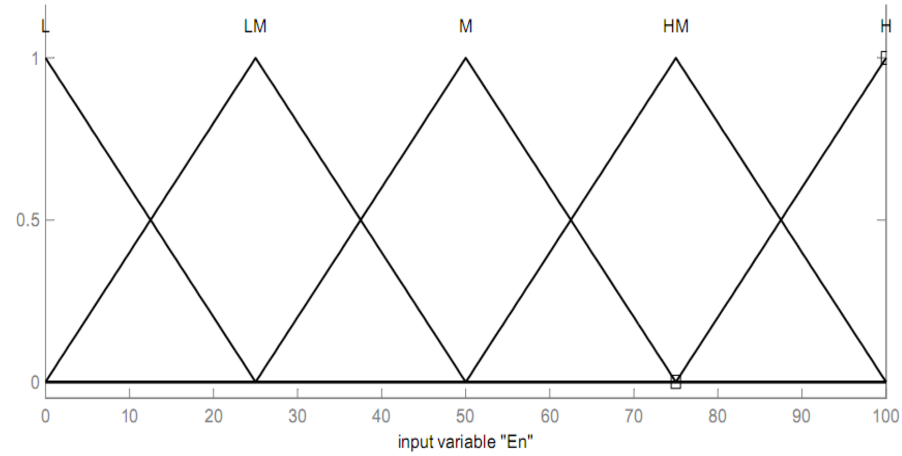


Table 2 The fuzzy rulers defined

ID	*Rulers*
1	If (*En* is *L*) and (*He* is *L*), then (*TV* is *H*)
2	If (*En* is *L*) and (*He* is *LM*), then (*TV* is H)
3	If (*En* is *L*) and (*He* is *M*), then (*TV* is H)
4	If (*En* is *L*) and (*He* is *MH*), then (*TV* is H)
5	If (*En* is *L*) and (*He* is H), then (*TV* is *HM*)
6	If (*En* is *LM*) and (*He* is *L*), then (*TV* is *H*)
7	If (*En* is *LM*) and (*He* is *LM*), then (*TV* is *H*)
8	If (*En* is LM) and (*He* is *M*), then (*TV* is *H*)
9	If (*En* is LM) and (*He* is *MH*), then (*TV* is *HM*)
10	If (*En* is *LM*) and (*He* is *H*), then (*TV* is *HM*)
11	If (*En* is *M*) and (*He* is *L*), then (*TV* is *HM*)
12	If (*En* is *M*) and (*He* is *LM*), then (*TV* is *M*)
13	If (*En* is *M*) and (*He* is *M*), then (*TV* is *M*)
14	If (*En* is *M*) and (*He* is *MH*), then (*TV* is *LM*)
15	If (*En* is *M*) and (*He* is *H*), then (*TV* is *L*)
16	If (*En* is *MH*) and (*He* is *L*), then (*TV* is *HM*)
17	If (*En* is *MH*) and (*He* is *LM*), then (*TV* is *LM*)
18	If (*En* is *M*H) and (*He* is *M*), then (*TV* is *L*)
19	If (*En* is *MH*) and (*He* is *MH*), then (*TV* is *L*)
20	If (*En* is *MH*) and (*He* is *H*), then (*TV* is *L*)
21	If (*En* is *H*) and (*He* is *L*), then (*TV* is *L*)
22	If (*En* is *H*) and (*He* is *LM*), then (*TV* is *L*)
23	If (*En* is *H*) and (*He* is *M*), then (*TV* is *L*)
24	If (*En* is *H*) and (*He* is *MH*), then (*TV* is *L*)
25	If (*En* is *H*) and (*He* is *H*), then (*TV* is *L*)

4 Simulations

In this section, we conduct experiments to evaluate our proposed approach by using a real-world feedback rating data set in the experiment. Then, we investigate the performance of our proposed approach by comparing it with other approaches (Conner et al., 2009; Limam and Boutaba, 2010b) and traditional average approach in terms of accuracy. We also evaluate the time cost and satisfactory rate of our approach.

4.1 Simulation set-up

For the experiments on the deviation, we use a real-world feedback rating data set. The data set consists of data from 20 Google Cloud Service (Google Play). Overall, the data set contains 723,471 feedback ratings. Ratings are on a 1–5 scale, where ‘5’ is the best (integer feedback ratings only).

It is worth noting that many existing reputation systems (Li and Ling, 2004; Maarouf et al., 2009; Malik and Bouguettaya, 2009) use simulation data for performance evaluation, because the limited availability of current feedback rating data. In this paper, the simulated feedback rating data can reflect the real situations by setting the magnitude (e.g. 1, 2, …, 5) of feedback ratings and the density (e.g. 10%, 20%, …, 50%) of unstable feedback ratings (Conner et al., 2009; Li and Ling, 2004; Limam and Boutaba, 2010b; Malik and Bouguettaya, 2007; Xu et al., 2007). Hence, in our experiments, we also employ simulation to generate unstable feedback ratings to evaluate the proposed approach.

Some unstable feedback ratings are generated synthetically, which allows us to control and study the characteristics of the feedback ratings. Hence, to investigate the performance of the reputation measurement for unstable feedback ratings, we simulated 10,000 feedback ratings of cloud services. Every feedback rating is also limited to an integer feedback rating from 1 to 5.

We study our experimental results from a PC with an Intel Core2 2.0 GHz processor and 2.0 GB of RAM. The machine runs Windows XP SP3, Matlab 7.6 and Java 1.4.8. We compare our approach with the reputation measurement approaches of Conner et al. (2009) and Limam and Boutaba (2010b), with respect to the accuracy. The approach takes the client, the service, the normalised transaction feedback rating and the set of optional attributes to create a service invocation history record that is used to measure the reputation (Conner et al., 2009). Based on a combination of a perception function and a disconfirmation function, the approach designed a feedback rating computation model, and then adopted the simple exponential smoothing approach to compute reputation scores (Limam and Boutaba, 2010b). For illustration purposes, TMS represents the approach by Conner et al. (2009), ARM represents the approach by Limam and Boutaba (2010b) and TAA represents the traditional average approach.

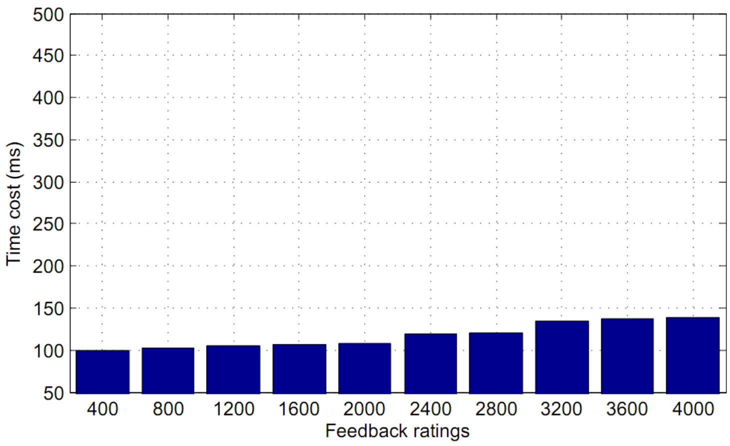
In the comparison, all of the test cases and the runtime environment are the same. Each experimental result is collected as an average after each approach is run 50 times.

4.2 Experiment on time cost

In order to evaluate the time cost of our proposed approach, we use 20 cloud services with 723,471 feedback ratings to verify its time cost.

As shown in Figure 6, we find the time cost of our approach is very low and is much less than 150 milliseconds (ms) on average. Obviously, the complexity of our approach is a linear time complexity, i.e. the time cost of our approach is a linear relationship between the number of feedback ratings and the time cost of our approach. Hence, our proposed approach has a good scalability in time cost.

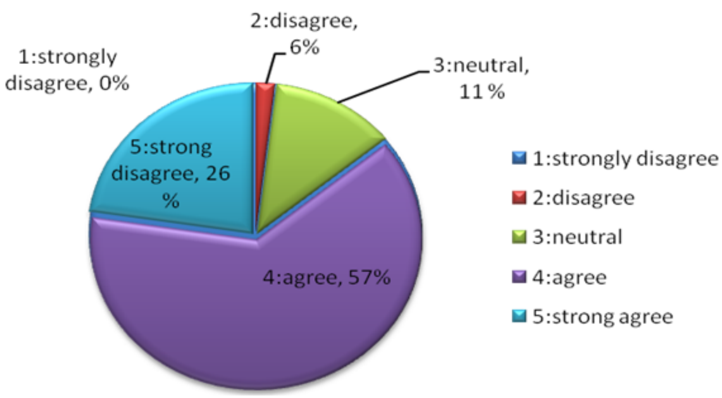
Figure 6 Time cost with different numbers of feedback ratings (see online version for colours)



4.3 Experiment on satisfaction rate

Figure 7 shows the satisfaction rate of this approach on a five-point Likert scale according to whether users felt that the reputation measurement approach helped them in choosing a cloud service and improve the accuracy of selection. Approximately 26% of users strongly agreed that our approach was helpful in improving the accuracy in choosing a cloud service and 57% of the users agreed that it was helpful. Most people think that the proposed approach can help improve the accuracy in choosing the best cloud services. Hence, it is an effective reputation measurement approach for cloud services.

Figure 7 Satisfaction rate with our approach



4.4 Experiment on accuracy

In this experiment, we evaluate the accuracy of our proposed approach by using 20 cloud services and their 723,471 feedback ratings.

The experiment contains two phases. In Phase 1, all approaches are used to calculate the reputation score of the 20 cloud services. The calculated score ranged from 1 to 5. In Phase 2, we select 100 feedback ratings of one cloud service randomly and compare the differences between each feedback rating and the calculated score. Then in this paper, the difference can be used to evaluate the accuracy of reputation calculation of each approach.

Definition 1 (Accuracy): *The accuracy is how often the ratio of the difference between the feedback ratings (r) and the reputation of the cloud service is less than or equal to a threshold (th) in n selected feedback ratings, i.e.*

$$\text{Accuracy} = \frac{\sum_{i=1}^{n} \begin{cases} 1, \left| r_i - \text{reputation} \right| \le th \\ 0, \text{otherwise} \end{cases}}{n} \times 100\%$$, *where th is set to 0.3.*

According to Definition 1, Table 3 shows the experimental results. The results are reported on average with 50 running.

Table 3 Competition results on accuracy

Instability level	*Accuracy*			
	TAA	*TMS*	*ARM*	*Our approach*
10%	54%	74%	68%	95%
20%	47%	62%x	53%	95%
30%	36%	48%	46%	96%
40%	23%	36%	41%	93%
50%	19%	29%	36%	94%

Table 3 shows that with different instability levels, the accuracy of our approach is the highest among all approaches. The accuracy of our approach is higher than 90%, which means that our approach is significantly better than other approaches. Why? The main reason is that our approach can find the unstable feedback ratings from users' feedbacks for each cloud services. Of course, other approaches failed to sense the issue.

5 Conclusion

Cloud computing provides cost-efficient opportunities for enterprises by offering a variety of dynamic, scalable and shared services (Flahive et al., 2013a; Flahive et al., 2013b). Usually, cloud service providers supply assurances by specifying technical and functional descriptions in Service Level Agreements (SLAs) for the cloud services they offer (Habib et al., 2011). End users, for instance, might have an SLA with a cloud service provider concerning how much bandwidth, CPU and memory the consumer can use at any given time throughout the day (Pallis, 2010). However, the descriptions in SLAs are not consistent among cloud service providers even though they offer cloud services with similar functionality. End users are not sure whether they can identify a good cloud service based only on its SLA. Hence, reputation measurement approach of cloud services based on feedback ratings is necessary to address the demand of the high-reputation cloud services when mass unstable feedback ratings exist.

In this paper, we present a reputation measurement approach of cloud services based on cloud model and fuzzy logic for unstable feedback ratings of cloud services. The main contributions of this paper can be summarised as follows: (1) we find the instability of feedback ratings for cloud services; (2) we adopt cloud model to compute the instability of feedback ratings and then employ fuzzy logic to calculate the reputation score of cloud service; (3) we compare our approach with other three approaches experimentally on a real-world feedback rating data set as well as on randomly generated feedback rating. The future work will focus on the dynamic computation of reputation to evaluate cloud services.

Acknowledgements

The work presented in this study is supported by NSFC (61202435), NSFC (61272521), and Beijing Natural Science Foundation (4132048).

References

Abawajy, J. (2011) 'Establishing trust in hybrid cloud computing environments', *Proceedings of the 10th IEEE International Conference on Trust, Security and Privacy in Computing and Communications*, 16–18 November, Changsha, China, pp.118–125.

Braithwaite, F. and Woodman, M. (2011) 'Success dimensions in selecting cloud software services', *Proceedings of the 37th EUROMICRO Conference on Software Engineering and Advanced Applications*, 30 August–2 September, Oulu, Finland, pp.146–154.

Brock, M. and Goscinski, A. (2010) 'Toward ease of discovery, selection and use of clusters within a cloud', *Proceedings of the 3rd IEEE International Conference on Cloud Computing*, 5–10 July, Miami, FL, pp.289–296.

Conner, W., Iyengar, A., Mikalsen, T., Rouvellou, I. and Nahrstedt, K. (2009) 'A trust management framework for service-oriented environments', *Proceedings of the 18th International Conference on World Wide Web*, 20–24 April, Madrid, Spain, pp.891–900.

Flahive, A., Taniar, D. and Rahayu, W. (2013a) 'Ontology as a Service (OaaS): a case for sub-ontology merging on the cloud', *The Journal of Supercomputin*g, Vol. 65, No. 1, pp.185–216.

Flahive, A., Taniar, D. and Rahayu, W. (2013b) 'Ontology as a Service (OaaS): extracting and replacing sub-ontologies on the cloud', *Cluster Computing*, Vol. 16, No. 4, pp.947–960.

Goscinski, A. and Brock, M. (2010) 'Toward dynamic and attribute based publication, discovery and selection for cloud computing', *Future Generation Computer Systems*, Vol. 26, No. 7, pp.947–970.

Gutowska, A. and Buckley, K. (2008) 'Computing reputation metric in multi-agent e-commerce reputation system', *Proceedings of the 28th International Conference on Distributed Computing Systems Workshops*, 17–20 June, Beijing, China, pp.255–260.

Habib, S.M., Ries, S. and Muhlhauser, M. (2011) 'Towards a trust management system for cloud computing', *Proceedings of the IEEE 10th International Conference on Trust, Security and Privacy in Computing and Communications*, 16–18 November, Changsha, China, pp.933–939.

Li, D., Cheung, D., Shi, X. and Ng, V. (1998) 'Uncertainty reasoning based on cloud models in controllers', *Computers and Mathematics with Applications*, Vol. 35, No. 3, pp.99–123.

Li, X. and Ling, L. (2004) 'PeerTrust: supporting reputation-based trust for peer-to-peer electronic communities', *IEEE Transactions on Knowledge and Data Engineering*, Vol. 16, No. 7, pp.843–857.

Limam, N. and Boutaba, R. (2010a) 'Assessing software service quality and trustworthiness at selection time', *IEEE Transactions on Software Engineering*, Vol. 36, No. 4, pp.559–574.

Limam, N. and Boutaba, R. (2010b) 'Assessing software service quality and trustworthiness at selection time', *IEEE Transactions on Software Engineering*, Vol. 36, No. 4, pp.559–574.

Liu, L. and Shi, W. (2010) 'Trust and reputation management', *IEEE Internet Computing*, Vol. 14, No. 5, pp.10–13.

Maarouf, I., Baroudi, U. and Naseer, A.R. (2009) 'Efficient monitoring approach for reputation system-based trust-aware routing in wireless sensor networks', *IET Communications*, Vol. 3, No. 5, pp.846–858.

Malik, Z. and Bouguettaya, A. (2007) 'Evaluating rater credibility for reputation assessment of web services', *Proceedings of the 8th International Conference on Web Information Systems Engineering*, 3–7 December, Nancy, France, pp.38–49.

Malik, Z. and Bouguettaya, A. (2009) 'Reputation bootstrapping for trust establishment among web services', *IEEE Internet Computing*, Vol. 13, No. 1, pp.40–47.

Noor, T.H., Sheng, Q.Z., Alfazi, A., Law, J. and Ngu, A.H.H. (2013a) 'Identifying fake feedback for effective trust management in cloud environments', *Proceedings of ICSOC Workshops*, 12–15 November, Shanghai, China, pp.47–58.

Noor, T.H., Sheng, Q.Z., Zeadally, S. and Yu, J. (2013b) 'Trust management of services in cloud environments: obstacles and solutions', *ACM Computing Surveys*, Vol. 46, No. 1, pp.1–35.

Pallis, G. (2010) 'Cloud computing: the new frontier of internet computing', *IEEE Internet Computing*, Vol. 14, No. 5, pp.70–73.

Rehman, Z., Hussain, F.K. and Hussain, O.K. (2011) 'Towards multi-criteria cloud service selection', *Proceedings of the 5th International Conference on Innovative Mobile and Internet Services in Ubiquitous Computing*, 30 June–2 July, pp.44–48.

Van Broekhoven, E. and De Baets, B. (2009) 'Only smooth rule bases can generate monotone Mamdani–Assilian models under center-of-gravity defuzzification', *IEEE Transactions on Fuzzy Systems*, Vol. 17, No. 5, pp.1157–1174.

Wang, S.G., Zheng, Z.B, Sun, Q.B., Zou, H. and Yang, F.C. (2011a) 'Evaluating feedback ratings for measuring reputation of web services', *Proceedings of the IEEE International Conference on Services Computing*, 4–9 July, Washington, DC, pp.192–199.

Wang, S.G., Zheng, Z.B., Sun, Q.B., Zou, H. and Yang, F.C. (2011b) 'Reliable web service selection via QoS uncertainty computing', *International Journal of Web and Grid Service*s, Vol. 7, No. 4, pp.410–426.

Wei-Tek, T., Peide, Z., Balasooriya, J., Yinong, C., Xiaoying, B. and Elston, J. (2011) 'An approach for service composition and testing for cloud computing', *Proceedings of the 10th International Symposium on Autonomous Decentralized Systems*, 23–27 March, Rio de Janeiro, Brazil, pp.631–636.

Wu, J., Liang, Q. and Bertino, E. (2009) 'Improving scalability of software cloud for composite web services', *Proceedings of IEEE International Conference on Cloud Computing*, 21–25 September, Bangalore, India, pp.143–146.

Xu, Z., Martin, P., Powley, W. and Zulkernine, F. (2007) 'Reputation-enhanced QoS-based web services discovery', *Proceedings of the IEEE International Conference on Web Services*, 9–13 July, Salt Lake City, UT, pp.249–256.

Zibin, Z., Yilei, Z. and Lyu, M. R. (2010) 'CloudRank: a QoS-driven component ranking framework for cloud computing', *Proceedings of the 29th IEEE Symposium on Reliable Distributed Systems*, 31 October–3 November, New Delhi, India, pp.184–193.