Web Service Reputation-Semantics and Assessment based on Customer Feedback Forecasting Model

Jebrin AL-SHARAWNEH¹, Mary-Anne WILLIAMS¹, David GOLDBAUM²
¹Faculty of Engineering and Information Technology
²School of Finance and Economics
University of Technology, Sydney, Australia
jebrin@it.uts.edu.au, Mary-Anne.Williams@uts.edu.au, David.Goldbaum@uts.edu.au

Abstract: In the Service Web, customers’ feedback constitutes a substantial component of Web Service reputation and trustworthiness, which in turn impacts the service uptake by consumers in the future. This paper presents an approach to assess and predict reputation in service-oriented environments. For assessing a Web Service reputation, we define reputation key metrics to aggregate the feedback of different aspects of the ratings. In situations where rating feedback is not available, we propose a Feedback Forecasting Model (FFM), based on Expectation Disconfirmation Theory (EDT), to predict the reputation of a web service in dynamic settings. Then we introduce the semantic concept Reputation Aspect and show how to compute it efficiently. Finally we show how to integrate the Feedback Forecasting Model into Aspect-Based Reputation management framework.

To demonstrate the feasibility and effectiveness of our approach, we test the proposed model using our Service Selection Simulation Studio (4S). The empirical results included in this paper show the applicability and performance of the proposed Reputation Assessment based on the Customer Feedback Forecasting Model. We also show how our model is efficient, particularly in dynamic environments.

Keywords: Semantic web services; reputation assessment; automated feedback forecasting; reputation aspect.

I. INTRODUCTION

In the Service Web, Web services, and semantic technologies will emerge to create an environment where users and applications can search and compose services in an automatic and seamless manner [1]. The Service Web is expected to be a place where a huge number of web services will compete to offer a wide range of similar functionalities. It is expected that Web services will fully leverage the Semantic Web to outsource part of their functionality to other Web services [1]. Moreover, services from distributed locations can be composed to create new value-added composite services. In this case, some services may not have interacted before, while others may act maliciously to be selected. A key requirement is to provide trust mechanisms for quality access and retrieval of relevant Web services.

Web service reputation is regarded as a predictor of its future behavior [1]. It is a collective measure of the opinion of a community of users regarding their actual experience with the service [2]. It is computed as an aggregation of users’ feedbacks and reflects the reliability, trustworthiness and credibility of the service and its provider.

We extend the above definition to view different reputation aspects as components of the global reputation. Each reputation aspect is an aggregation of all feedbacks that address that aspect. Some aspects can address provider behavior, others address Web service behavior. While some aspects can be considered on the attribute level such as cost of web service, other aspects such as provider reliability or service duration.

In our proposed social recommender system [3], customers provide their queries for a specific service; they provide their preferences on different dimensions of functional, quality (QoS) and cost constraints. Each time customers request a service, they can set their preferences by assigning important weights for each criterion in advance. Then, the system will recommend services according to the preference setting. After completing interaction, customers rate provider’s performance of current interaction in terms of those criteria as the feedbacks of service satisfaction. Then feedback is used to update the reputation and trustworthiness of the service and its provider; which enhance the quality of recommendations and can be used to assess credibility of feedback and detect malicious parties who provide fake reports or feedback about the service and its provider.

Customer feedback constitutes a substantial component of Web service reputation and trustworthiness. Customer feedback can be either implicit or explicit depending on the way customer satisfaction is captured [3]. Although explicit methods capture user satisfaction more accurately, they are costly and often users do not cooperate.

The ‘feedback score’ is the most prominent indicator of a Web service’ reputation. At times, due to various reasons, a majority of the members may not be willing to engage in the rating process. In such situations of ratings scarcity [1] the accuracy of the reputation system may be compromised. In [4], Dellarocas and Wood (2008) investigate the consequences of non-random missing feedback on reputation scores. They find that dissatisfied consumers are more likely not to give feedback. In this paper we address the issue of ratings scarcity by proposing a Feedback Forecasting Model (FFM) as a crucial module of a feedback monitoring system and predicting the reputation of a given aspect based on feedback from available feedback data.
Recommender systems by their nature are centralized. In this paper we consider (1) Reputation Management is centralized and built in the recommender system. (2) Customers are either humans or agents (3) Web services using a set of standards such as WSDL, UDDI, and SOAP. Assessing credibility of raters and associated issues as free riding, fake identities, feedback ratings incentives, and reputation bootstrapping are outside the scope of this paper.

The rest of the paper is organized as follows. Section II outlines our motivations and contributions, Section III describes our terminology. Section IV presents Automated Feedback Monitoring System, and Section V outlines reputation system. The following sections present the evaluation of the proposed model, then related works. Finally, we conclude by summarizing our findings and our future plans for further work on the area.

II. MOTIVATIONS AND CONTRIBUTIONS

According to (Klein, Lambertz et al. 2009) eBay buyers comment on sellers 52.1% of the time. Ratings left by buyers were positive in 98.1% of the cases. From this ratio we can conclude that 47.9% of the buyers either fear revenge or they do not bother to give a seller rating. In such situations of ratings scarcity [1], the accuracy of the reputation system may be compromised. This ratio is critically important in providing reputation and trustworthiness; it motivates us to forecast customer feedback.

Furthermore, different customers might evaluate a web service from different QoS perspectives, such as response time, accuracy, execution time. For each attribute, they can set their preferences by assigning importance level (weight) for each criterion in advance, and then the customer develops a kind of impression \ satisfaction after using the service. The overall satisfaction depends on the aggregation of customer satisfaction in each attribute; which in turn develops overall customer feedback about the Web Service.

This implies that different customers have different interests, and feedback has multiple aspects associated with customer preferences and interests, which can play a role in deciding different reputation aspects of the Web service. Aspects vary in their nature such as provider aspect, cost aspect and network aspect, etc., which motivates us to model reputation on an aspect basis rather than on global reputation.

Another reputation management challenge is the problem that arises in all reputation management approaches which evaluate the reputation of the provider; if the provider provides two services with different capabilities for the same Web service community, it receives an average reputation over all web services provided. For example if provider (P) provides two services (S1 and S2) with varying qualities and reliabilities, it is clear that each service has a specific unique trustworthiness and consequently each service has a unique reputation. We argue that reputation management should be treated on Web service level where each service provided by the provider has its unique reputation.

Therefore there is a need for a rating mechanism that takes into account the factors discussed above to overcome these shortcomings and different reputation metrics such as: the time decay of the rating, the credibility of those providing rating, the reliability of the rating provided, and the availability of ratings.

The contributions of this paper are:

1. A Feedback Forecasting Model (FFM) that predicts consumer feedback about satisfaction with the Web Service based on the Expectation Disconfirmation Theory (EDT).
2. A Semantic Web Service Reputation framework based on reputation aspects of the Web service, considering service provider as one of these aspects.

III. FEEDBACK AND REPUTATION SEMANTICS

Different customers may evaluate a web service from different QoS aspects, such as reliability, availability, response time, accuracy, execution time and cost. For each aspect, the customer develops a sort of impression \ satisfaction. The overall satisfaction depends on the aggregation of customer satisfaction in each aspect, which in turn develops customer feedback about the Web Service. This emphasizes that feedback has multiple aspects associated with customer preferences and interests, which can play a role in deciding reputation of the Web service.

For example if Customer A has preferences \{Q1, Q5, Q3\} in a web service, while Customer B has preferences \{Q2, Q4, Q5\}; in current approaches, each customer feedback has one entry in the Reputation Management System. However, in such a situation; we believe that Customer A and Customer B feedbacks are worth only one entry in the Reputation Management System.

As shown in Figure 1, we propose a model to handle different aspects with different values in the Feedback Management System.

Let feedback = \{FuncFeedback, Non-FuncFeedback, costFeedback\}

If we consider a home loan service, any of the following aspects the customer may have interest in:

- FuncFeedback = RatingsOf{ Redraw Facility Avail(Y/N), Loan is Transferable (Y/N) }
- Non-FuncFeedback = RatingsOf{ reliability, availability, execution time, security }
- costFeedback= RatingsOf{ Interest Rate, Application Fee, Legal Fee, Settlement Fee, Monthly Service Fee }

Furthermore, we can consider the security aspect as:

- Security Aspect\{Accountability, Authentication, Authorization, Traceability, Auditability, Non-Repudiation, confidentiality, privacy, Encryption\}
Performance Aspect = \{\text{processing time, throughput, response time, latency}\}

Feedback Transaction can be captured either manually from the customer or from an automated feedback management system. If manual feedback is provided by a customer then automated feedback can be used to assess the credibility of customer feedback, alternatively it is used as a feedback entry in the feedback system.

A. Reputation Aspects

We argue that feedback ratings and consequently reputation are context specific particularly for selection and composition.

Even in the same context, there is a need to develop differentiated reputation evaluations for different aspects of a service [5]. For instance, a user might evaluate a web service from different QoS aspects, such as response time, accuracy, execution time. The global reputation depends on the combination of the reputations in each aspect. This emphasizes that reputation has multiple aspects, which can play a role in deciding whether a service is trustworthy for the purpose of selection and composition contexts.

For example eBay maintains different customer satisfaction categories \{aspects such as - Items as described, Communication, Postage time \} Delivery time and Shipment fees. Unlike eBay which maintains reputation on the provider level; our reputation management system maintains reputation on the Web Service level, so we can find the provider aspect appears in more than one web service reputation.

B. Reputation Representation

Reputation Aspect can be defined using a 3-tuple \((V, T, N)\) where \(V\) refers to reputation value \([0,1]\), \(T\) is the time stamp for the Aspect creation and \(N\) is the average number of ratings of all attributes in that aspect.

Global Reputation = \{set of \((1-N)\) of Reputation Aspects\}

Since user interests with some attributes decreases with time, (i.e. advances in technology ensures high quality over time), we believe our model has the following advantages over other reputation computational models: it provides the deemed required dynamism to join new attributes or ignore inactive attributes any time without the need to re-establish the reputation entry again. Furthermore, the model provides the required flexibility to provide reputation on specific Aspect about the web service which is important in service selection and composition.

IV. AUTOMATED FEEDBACK MONITORING SYSTEM

In this section we address the issue of the lack of feedback (ratings scarcity) by proposing a feedback forecasting model as module of feedback monitoring system.

The automation of the rating process has two folds: (1) provide a mean to assess credibility of human customers feedbacks, (2) provide forecasting of feedbacks on behalf of customers who are either afraid of other party revenge or they do not bother themselves with rating issue.

In order to automate the service rating process we elaborate a Feedback Forecasting Model based on customer satisfaction as shown in Figure 2. This model relies on the measurement of expected service utility and perceived service utility.

A. Feedback Forecasting Model

The model uses the customer's initial Web Service expectations, and then, after use of the Web Service, the model compares the perceived service performance to these initial service expectations. This comparison results in positive / negative disconfirmation, which, along with performance, influences customer satisfaction with the Web Service.
Service. Satisfaction with the Web Service can influence customer feedback about the service.

We assume that quality monitoring is achieved by a trusted system that produces credible QoS reports. According to [5] quality monitoring can either be supported by companies that manage service directories and are eager to control their services' quality, or by third party companies hired to achieve QoS monitoring tasks for them.

1) Feedback vs. Satisfaction
We base our model (as shown in Figure 3) on the expectancy-disconfirmation theory (EDT). This theory relates expectation, perceived quality and quality disconfirmation, i.e. the difference between the expected and perceived quality, to human being customer satisfaction.

![Feedback Forecasting Model](image)

Using EDT theory [6] and studies in the field of psychology and marketing, user satisfaction forecasting models have been derived, as shown in [2]. Inspired from Xiao and Boutaba’s (2007) work, we assimilate in our models have been derived, as shown in [2].

![Utility Function Computation](image)

There are two concepts of utility presented in this section: expected utility and perceived utility. Expected utility denotes a set of service related performance metrics that are measurable or observable. Together, they yield a single quantitative evaluation of utility.

2) Service Utility and Perceived Utility
As noted by Dabholkar, 1993 in [7], customer satisfaction and utility are not the same construct. Satisfaction is a customer’s subjective evaluation of the service performance, while utility is its objective measurable quantification. There are two concepts of utility presented in this section: expected utility and perceived utility. Expected utility is the objective measurable quantification of the service performance, while utility is its subjective evaluation. Together, they yield a single quantitative evaluation of utility.

3) Utility Function Computation:
The utility function $U_i$ from using service $S_i$ with (m) attributes as adopted from [8], is defined as:

$$U_i = \sum_{j=1}^{m} w_j * (Q^{(j)})^2 \text{ where } (1 \leq j \leq m)$$  

(1)

Where, $w_j$ is the normalized weight for attribute (j), such that: $\sum_{j=1}^{m} w_j = 1$

$Q^{(j)} = \begin{cases} \frac{q^{(j)} - \min^{(j)}}{\max^{(j)} - \min^{(j)}} & \text{if } \text{Maximize, } op = (\geq) \\ \frac{\max^{(j)} - q^{(j)}}{\max^{(j)} - \min^{(j)}} & \text{if } \text{Minimize, } op = (\leq) \\ 1 & \text{if } \max^{(j)} - \min^{(j)} = 0 \end{cases}$  

(2)

Max$^{(j)}$ and Min$^{(j)}$ are the maximum and minimum of attribute (j) from all available services in that category.

When there is a change in $q^{(j)}$ value, such that it does not belong to Min$^{(j)}$ and Max$^{(j)}$, then:

- $\min^{(j)} = q^{(j)}$ if $q^{(j)} < \min^{(j)}$ and $\text{Minimize, } op = (\leq)$
- $\max^{(j)} = q^{(j)}$ if $q^{(j)} > \max^{(j)}$ and $\text{Maximize, } op = (\geq)$

When we deal with one attribute only, then utility of that attribute: $U_i = Q^{(j)}$.  

B. Feedback Computation from Customer Satisfaction (CSAT)
Customer satisfaction can be modeled through the interaction between perceived utility and expected utility [7], expressed as a linear combination of a perception function and a disconfirmation function.

Following Xiao and Boutaba (2007), the satisfaction CSAT(s) of a customer with service (s) is defined as follows:

$$\text{CSAT}(s) = f_1(\text{Perceived Utility}(s)) + f_2(\text{Perceived Utility}(s) - \text{Expected Utility}(s))$$  

(3)

where $U_p(s)$ is the perceived measured utility, $U_e(s)$ is the expected utility, $f_1$ is the perception function which maps the perceived utility to user’s satisfaction, and $f_2$ is the disconfirmation function which reflects the subjectivity of user’s evaluation given its expectation as reference point.

By mapping FEEDBACK(s) to CSAT(s) in the equation 3, we obtain:

$$\text{FEEDBACK}(s) = f(\text{CSAT}(s))$$

![Perception Function](image)
Perception function \( f_1(x) \) is an increasing function defined in \([0, 1]\) and is bounded between \( f_1(0) = 0 \) and \( f_1(1) = 1 \). According to [7], as the utility increases, the customer becomes less sensitive to changes in utility. We believe that users are not much more sensitive to changes for very low utility values than they are for high utility values. In our model, we consider that \( f_1 \) varies slightly for utility values nearby \( 0 \) and for utility values nearby \( 1 \). Hence the concavity of \( f_1 \) changes at a particular utility value in the range \([0, 1]\). Like in [7], and with respect to the above requirements, \( f_1 \) can be defined as follows:

\[
f_1(x) = w_p x^{\mu_1}
\]

Where \( \mu_1 \geq 0, \mu_1 \leq 4w_p \) and \( 0 < w_p \leq 1 \). \( w_p \) and \( \mu_1 \) define the shape and concavity of the perception function \( f_1 \) respectively, as shown in Figure 4. The concavity of the perception function is the highest for values of \( \mu_1 = 4w_p \), (i.e. maximum positive disconfirmation). \( \mu_2 \) regulates the impact of negative disconfirmation on customer satisfaction. This model penalizes services that deliver low utility levels and reward services with high utility levels. Characteristics of the disconfirmation’ function is illustrated in Figure 5.

Customer satisfaction: By adding equations 4 and 5, we get Customer Satisfaction Function as shown in Figure-6. We observe that \( CASAT(s) \) is bounded between \( w_{dn} \) and \( w_p + w_{dp} \). In general, the choice of \( W \) parameters should follow the following rules [7]:

1. \( w_p \geq w_{dn} > w_{dp} \)
2. \( w_p, w_{dn}, w_{dp} \) Should be fixed for all customers of a service and
3. \( w_{dp} \) Should be small compared to \( W_p \) such that \(( w_p + w_{dp} = 1 \)

While \( \mu \) Parameters \( \mu_1 \geq 0, \mu_1 \leq 4w_p, \mu_2 \geq 0 \)

Disconfirmation Function:
The disconfirmation function \( f_2 \) accounts for the subjectivity of customer evaluation given a reference point (i.e. expectation). In [7], Tversky and Kahneman, 1973, found that “losses relative to a reference value looms larger than gains”. Grounded on this psychological theory, Anderson and Sullivan 1993, suggest that customer satisfaction is mildly increasing when perceived utility exceeds expectation and is significantly reduced when perceived utility falls below expectation. [7] formalize this interaction as a two piece increasing function as follows:

\[
f_2(x) = \begin{cases} 
  w_{dp} x & x \geq 0 \\
  w_{dn}(x + 1)^{\mu_2} - w_{dn} & x \leq 0 
\end{cases}
\]

Where \( x = U_p - U_e, \mu_2 \geq 0, \text{ and } w_{dp}, w_{dn} > 0 \)

We observe that the domain of \( f_2 \) is bounded between -1 and 1. The function is continuous (i.e. the two piecewise functions converge at \( x = 0 \)). The parameter \( w_{dp} \) controls the maximum value of \( f_2 \) (i.e. maximum positive disconfirmation) while \( w_{dn} \) controls the minimum value of \( f_2 \).
Special Case: When SLA is respected

When SLA is respected, i.e. the provider provided what was promised (advertised), then expected utility (Ue) = perceived utility (Up); consequently part 2 of equation (3), which calculates disconfirmation function becomes not applicable. In this situation \( W_p = 1 \) as shown in Figure-7.

Impact of \( \mu_1 \) Parameter on service domain

Variation in \( \mu_1 \) parameter can introduce a wide spectrum of Customer Satisfaction Models as shown in Figures (4, 7). \( \mu_1 \) Parameter is related to the cost of the service. For expensive services where the customer is very sensitive to service quality, \( \mu_1 \) should be highest where concavity down increases. While for cheap services \( \mu_1 \) should be lowest where concavity up increases. When \( \mu_1 = 1 \), then relation is linear.

In the next section we show how Feedback Forecasting Model used in the reputation management system.

V. REPUTATION MANAGEMENT SYSTEM

Reputation systems provide a way for building trustworthiness of Web Services. After a transaction is completed, customers provide ratings as feedback about their satisfaction with the Web Service performance. As shown in Figure 2, the reputation system collects ratings from customers who wish to provide feedback. When customers do not provide feedback the reputation system prompts the Feedback Forecasting module to provide its assessment of the Web Service performance. The reputation system continuously updates each Web Service reputation score as a function of the received ratings. Updated reputation scores are provided online for all the consumers (and Web Service providers) to see, and can be used by consumers to decide whether or not to use services from a particular provider.

Following [9], two fundamental aspects of centralized reputation systems are:
1. Centralized communication protocols that allow participants to provide ratings about transaction partners to the reputation system, as well as to obtain reputation scores of potential transaction partners from the reputation system.
2. A reputation computation approach to be used by the reputation system to derive reputation scores for each Web service based on received feedback ratings and possibly on other information such as Web service community opinions. In the following section we address computation model of Reputation Management System.

A. Reputation Computation Model

Different customers may provide their feedback about their satisfaction with the web service from different QoS aspects, such as reliability, availability, response time, accuracy, execution time and cost. For each aspect, the customer develops a sort of impression \( \text{satisfaction} \) in the range \([0,1]\). The overall satisfaction depends on the aggregation of customer satisfaction in each aspect; which in turn develops customer feedback about the Web Service.

For the \((j)\)th attribute of service \((S)\), customer feedback \([0,1]\), for specific transaction can be expressed as:

\[
\text{FEEDBACK}(S_j) = f(\text{CSAT}(S_j))
\]

where \((1 \leq j \leq m)\) and \((m \leq n)\), \((m)\) is the number of preferred QoS attributes of web service \((s)\) of customer \((i)\); \((n)\) is the total number of attributes of service \((S)\).

This emphasizes that feedback has multiple aspects associated with customer preferences and interests, which can play a role in deciding reputation of the Web service as viewed by that customer. Formally, the reputation aspect \( \text{j} \) attribute \((j)\) is denoted as: \( \text{REP}(S_j) \) in the range \([0,1]\), as viewed by customer \((i)\) after execution a transaction, and defined as:

\[
\text{REP}(S_j) = f(\text{FEEDBACK}(S_j))
\]

1) Reputation Computation Metrics

From [10], [1] we inspired the following reputation metrics that suits our centralized reputation system:

a) Customer Preferences:

Since each customer has a specific preference weight in her query for attribute \((j)\) denoted by \((W^j)\) in the range \([0,1]\), then Reputation of attribute \((j)\) is \( \text{REP}(S^j) \) as viewed by all customers, is the weighted average of all feedbacks from all customers \((N)\) who rated attribute \((j)\), and can be expressed as:

\[
\text{REP}(S^j) = \frac{\sum^N_{i=1} \text{FEEDBACK}(S^j)^i·W^j_i}{\sum^N_{i=1} W^j_i}
\]

Let us define weight average for attribute \((j)\) as:

\[
W^j_a = \frac{1}{N^j} \cdot \sum^N_{i=1} W^j_i
\]

Then Equation (8) can be rewritten as:

\[
\text{REP}(S^j) = \frac{1}{N^j} \cdot \frac{\sum^N_{i=1} \text{FEEDBACK}(S^j)^i·W^j_i}{W^j_a}
\]

b) Feedback Rating Credibility:

A reputation management system should weigh the ratings of highly credible raters more than raters with low credibility. The credibility of a service rater lies in the interval \([0,1]\) with 0 identifying a dishonest rater and 1 an honest one. The process involved in calculating rater’s credibility is out of the scope of this paper. Adopting Credibility of Rater from [1], Equation (8) can be rewritten as:

\[
\text{REP}(S^j) = N^j \cdot \frac{\sum^N_{i=1} \text{FEEDBACK}(S^j)^i·C^j_i·W^j_i}{\sum^N_{i=1} W^j_i·C^j_i}
\]

Note that: \((N^j)\) coefficient comes from aggregating feedback with two weighting parameters \((C^j_i'\text{and } W^j_i)\) at the same time.

c) Decay Factor:

Trust and reputation can increase or decrease with further experiences (interactions or observation), they also decay with time [5]. New experiences are more important than old.
ones; hence all the past reputation data may be of little or no importance. Adopting Decay factor from [11], given that all transactions are time stamped, equation (11) can be rewritten as:

\[ \text{REP}(S_i) = \frac{\sum_{t_{i+1}}^{t_i} \left( \frac{N}{\sum_{j=1}^{N} w_j} \cdot \frac{\sum_{j=1}^{N} \text{FEEDBACK}(s_i^j) \cdot C^j \cdot w_j}{\sum_{j=1}^{N} w_j} \right) \cdot f_d(t) }{\sum_{t_{i+1}}^{t_i} f_d(t)} \]  

(12)

The decay factor is given by:

\[ \lambda = e^{-\lambda_1 (t_2 - t_1)} \]

where each time period has its decay factor, \( \lambda_1 \in [0, 1] \), and \((t_2 - t_1)\) is the time interval difference between the present time and the time in which the rating was collected from the rater. It is clear that all feedbacks in a specific period have the same decay factor. For current period \((t_2 - t_1) = 0\), then \( \lambda = 1 \). For old feedbacks, when \( \lambda < \lambda_1 \) then rating is not considered as it is outdated. As shown in Figure 8, lower values of \( \lambda_1 \) give higher range of reference time.

Let us define rating credibility average for attribute (j) as:

\[ C_{a}^j = \frac{1}{N} \cdot \sum_{i=1}^{N} C_{a_i}^j \]  

(13)

Equation (12) can be rewritten as:

\[ \text{REP}(S_i) = \frac{\sum_{t_{i+1}}^{t_i} \left( \frac{N}{\sum_{j=1}^{N} w_j} \cdot \frac{\sum_{j=1}^{N} \text{FEEDBACK}(s_i^j) \cdot C^j \cdot w_j}{\sum_{j=1}^{N} w_j} \right) \cdot f_d(t) }{\sum_{t_{i+1}}^{t_i} f_d(t)} \]  

(14)

2) Reputation Aspect:

Reputation aspect is a group of Web Service related attributes aggregated to form a partial image of the Web Service Reputation.

Let an Aspect \((A1)\) be the aggregation of a number of Web Service attributes (k), each attribute associated with an average weight \( (W_a^j) \) and an average credibility \( (C^j_a) \),

where:

\[ n = \text{total number of the Web Service attributes}, (1 \leq k \leq n) \]

Then reputation Aspect \((A1)\) can be defined as:

\[ \text{REP}(S_i^{A1}) = \frac{\sum_{j=1}^{k} \text{REP}(s_i^j) \cdot W_a^j}{\sum_{j=1}^{k} W_a^j} \]  

(15)

3) Web Service Global Reputation:

Consequently, Global Reputation is the aggregation of all attributes \((n)\) of the web service, and can be defined as:

\[ \text{REP}(S_i) = \frac{\sum_{j=1}^{k} \text{REP}(s_i^j) \cdot W_a^j}{\sum_{j=1}^{k} W_a^j} \]  

(16)

Where \( (\text{REP}(S_i^j), W_a^j) \) are the \((j)\)th attribute reputation and average weight respectively, \((n)\) is total number of Web service attributes.

VI. SIMULATION AND EVALUATION

We extend our simulation tool “4S: Service Selection Simulation Studio” [12], which is developed under NetLogo platform\(^1\). The user interface as shown in Figure 9, we use it to analyze and evaluate the validity of our approach, in the following section we outline the testing environment and associated outcomes.

A. Simulation model:

Our model as shown in figure 6.1 is based on follow the leader principle [13] for social recommendations. It is composed of a fixed number of atomic services (50) with the same functional properties and varied in their (QoS) attributes; each atomic service maintains the following information: (service code, a list of QoS attribute codes and corresponding values). Each service\(^1\) information is hidden from others, and it is static during any simulation session.

Each simulation session is composed of a fixed set of rounds (500). In each round a fixed number of customers (100) enter their queries to the system. Each customer has the following information: (ID, a list of attribute QoS codes and corresponding values and weights). In addition, the system implements Reputation algorithm proposed in section (V.A) based on customers’ feedback which is predicted by the end of each transaction. Each service has an

\(^1\)http://ccl.northwestern.edu/netlogo/
initial reputation at the beginning of each session based on its capabilities.

Each simulation session starts by setting the services and customers with their corresponding information. Each round starts with a new set of customers with a new query. In each round every customer passes its information to the system, if the customer is a qualified leader, then the system performs the service selection process, selecting the best service from available services based on the utility perceived from that service and current reputation. If the customer is a follower, then the system selects randomly a friend from customer friends to follow.

By the end of each round, each customer provides feedback to the system about their satisfaction, and this feedback used to derive service reputation which has impact on service selection performance.

B. Simulation results:
The applicability and effectiveness of our approach is evaluated by observing the variance between the actual behavior and predicted reputation in the following cases:

1) Feedback forecasting Impact on lack of feedback
Analysis:
In this test, we simulate a scenario where 47.9% of the customers do not provide feedback (Klein, Lambertz et al. 2009). We calculated the associated reputation when feedback provided which is shown in the top of figure-6.2.1 and predict the reputation associated with the missing feedback. From this we conclude that forecasting feedback for missing feedback has significant impact on global reputation.

As shown in Figure 10, for a given service over 100 rounds, with 48% of the customers not providing feedback, reputation of the selected service dropped from (0.956 when feedback is provided to 0.451 when feedback is forecast to average 0.749 for both).

Conclusions:
1. Proposed Feedback Forecasting Model is a convenient and effective approach to predict service reputation.
2. Actual Web service reputation decrease about 20% from its value when we consider lack of feedbacks.

2) Attribute Based Reputation vs. Service Based Reputation Computation
In this test, we validate our argument that Web service reputation computation is more accurate when we base the computation on the attribute level, than using one feedback as one entry in the reputation computation, which used in other models such as [1].

For a given service over 100 rounds, with 100 customers in each round, each of them provide feedback about the service. Current approaches aggregate feedback to generate the reputation. Our computational model considers customer feedback as feedback for each attribute, then we aggregate this on the level of the attribute considering number of customers who are interested in that attribute. From varied attributes data we can calculate (Reputation aspect) or global web service reputation.

As shown in Figure 11, although 100 customers in each round provide feedback, average feedbacks per attribute is 45 with variations in the computed global reputation. The difference between the two values lies in the range (± 0.03) which has significant impact on service selection and composition.
Conclusions: Attribute Based Reputation Computation can be used to compute global reputation with more accuracy than Service Based Reputation Computation.

3) Reputation aspect Concept Validation

In this test, we validate our proposed argument that Reputation aspect differs from global reputation of the service.

Let ASPECT-A associated with attributes (3 & 7) of the web service.

Then Reputation aspect is the aggregation of feedbacks on these attributes only but not all attributes of the web service which is the case of global reputation.

As shown in Figure 12, Although 100 customers in each round provide feedback, average feedbacks per attribute is 43 with variations in the computed global reputation and the Reputation aspect. The difference between two values lies in the range (± 0.07) which has significant impact on service selection and composition.

Conclusions: Attribute Based Reputation Computation can be used to compute Reputation aspect, and there is a significance difference between Global Reputation and Reputation aspect.

4) Decay Factor impact on Initial Reputation

In this test, we validate our proposed algorithm to study the impact of the decay factor on initial reputation and outdated feedbacks.

For a given service over 100 rounds, in round 0 (i.e. when the service enters the market) it must have an appropriate initial reputation. According to decay factor, initial reputation and old feedback data decay over time, until they have no significant impact on the global reputation.

As shown in Figure 13, Initial reputation has significance indicated by (initRep%) in early rounds, while feedback reputation component grow exponentially and initial reputation component drop exponentially. In most current rounds the global reputation depends on the feedback.

Conclusions: Decay factor is reasonable approach in our algorithm.

VII. RELATED WORKS

Reputation assessment is an important aspect in Web service selection and composition. In order to obtain more reliable service reputation, several reputation models have been proposed in the literature. The spectrum of these reputation management solutions varies from “purely statistical” techniques to “heuristics-based” techniques. Bayesian systems [14], [15] and belief models such as [16] are the major examples of purely statistical techniques. Other approaches present heuristics-based solutions; for example, [1] present a hybrid solution defining key heuristics, and a statistical model (Hidden Markov Model, HMM-based) for reputation assessment. In the context of Web services’ reputation, interested readers are directed to good surveys about Web Services reputation such as [17], [18] and [19].

This paper defines reputation in a more reasonable way and supports measurable QoS constraints in Web service recommendation. Moreover, reputation is computed using objective consumer feedback, thus it avoids the problem of unfair rating.

Inspired by [10], [1] we developed reputation metrics that suits our centralized approach. Unlike previous works, every measurable QoS attribute is treated as an individual attribute in respect of reputation computation. Then from different attributes and based on the needs context (service selection or composition) a dynamic reputation aspect is concluded.

In this paper we address the lack of feedback or ratings scarcity as referred by [1], and how to forecast feedback in such situations. From Expectation Disconfirmation Theory (EDT) and from studies in the field of psychology and
marketing, user satisfaction forecasting models, like (Xiao and Boutaba 2007), [2] customer satisfaction have been derived. Inspired by Xiao and Boutaba’s (2007) work, we assimilate in our model customer expectation to the perceived quality.

[2] use EDT to forecast user satisfaction and feedback. Limam and Boutaba consider the SLA is respected. However, this approach fails to consider the dynamism of service behavior. We argue that SLA is not respected all the time because basic services may not guarantee the levels of service quality due to the influence of many factors (e.g. network transmission). We consider the general case is the disconfirmation and the special case as SLA respected.

Predicting Web service reputation behavior is a key issue in Web service provision. When the feedback series does not show any trend, it is hard to predict future service behavior. However, Smoothing-based forecasting techniques, like Moving Average, Weighted Moving Average, and Exponential Smoothing [2] can be used to predict near future behavior from past behaviors. Other more specific techniques like the Holt’s Linear Exponential and the Holt-Winters’ Forecasting are more suitable for long-term forecasting over data showing a trend and periodicity respectively.

VIII. CONCLUSION AND FUTURE WORK

Trust and reputation mechanisms have been used in many large open systems to solve the problem of selecting services/resources. A Web service system is a large open system too. Using trust and reputation mechanisms offers a promising way to solve the web service selection problem.

In this paper we proposed a feedback forecasting model to address lack of feedback (scarcity). Then we introduced the concept (Reputation aspect) from semantics perspective and showed how to compute it. Finally we integrated feedback forecasting model into Aspect-based reputation management framework.

We have evaluated our framework through simulations using our Service Selection Simulation Studio (4S). The results showed that the system succeeded in capturing service behaviors and in providing users with the best available choices. We proved feasibility of our Reputation computational approach by benchmarking it against other Reputation computation approaches. The results of the experiments included in this paper show the applicability and accuracy of the proposed Reputation Assessment based on Customer Feedback Forecasting Model.

In the future, we intend to build upon our proposed reputation management framework and propose a framework for QoS enabled reputation-based trustworthiness mechanism for web service selection and composition.

ACKNOWLEDGEMENT

Financial assistance was received from APA (Australian Postgraduate Award).

REFERENCES


