2009 IEEE/WIC/ACM International Joint Conference on Web Intelligence and Intelligent Agent Technology

An Adaptive Web Services Selection Method Based on the QoS Prediction Mechanism

Mu Li, JinPeng Huai, HuiPeng Guo
School of Computer Science and Engineering, Beihang University, Beijing, China
limu@act.buaa.edu.cn, huaijp@buaa.edu.cn
guohp@act.buaa.edu.cn

Abstract

In recent years, many QoS-based web service selection methods have been proposed. However, as QoS changes dynamically, the atomic services of a composite web service could be replaced with other ones that have better quality. The performance of a composite web service will be decreased if this replacement happens frequently in runtime. Predicting the change of QoS accurately in select phase can effectively reduce this web services "thrash". In this paper, we propose a web service selection algorithm GFS (Goodness-Fit Selection algorithm) based on QoS prediction mechanism in dynamic environments. We use structural equation to model the QoS measurement of web services. By taking the advantage of the prediction mechanism of structural equation model, we can quantitatively predict the change of quality of service dynamically. Optimal web service is selected based on the predicted results. Simulation results show that in dynamic environments, GFS provides higher selection accuracy than previous selection methods.

1. Introduction

Web services based on SOAP、WSDL、UDDI[1] and BPEL[2] have advantages in building business process with standard, loosely-coupled, reusable, flexible and effective nature. With the great increase of available services on the Web [2], automatic services selection has become a key problem in Web service research area. A lot of researches in the area focus on the importance of non-functional properties, also known as QoS (Quality of Service). These non-functional properties, like response time and availability are important for the success of software applications.

QoS is the overall quality score of a Web service and can be calculated by $\sum_{i=1}^{n} a_i v_i$ [3], in which, $v_i (1 \le i \le n)$ is the value of the ith QoS attribute of Web servive, and the relative $a_i (1 \le i \le n)$ is the user concern on it. From this expression, we can see that the value of QoS attributes and user concerns on them are two factors that determine the QoS of a Web service. Also, we can see that the QoS of Web services changes dynamically.

First of all $v_i (1 \le i \le n)$ is dynamic because of the uncertainty caused by the distributed, heterogeneous, autonomic, and dynamic features of web services. The services in a business process are distributed over the Web. They are provided by different organizations and run on different platforms. Any incident can affect the QoS of the composite web services, Service users are not trustful on the QoS data provided by the services providers. What's more the runtime environment is full of changes that come from dynamic publication, disablement, delete and update of individual services, the variety of services QoS, providers' cooperation relationships and network conditions.

On the other hand, the user concern on the QoS attributes changes as the user's focus changes. For example, cost may be of high priority in users' consideration in the initial period, so the weight (concern) of cost should higher than others in QoS calculation. But requirements of safety, reliability are growing gradually, the relative weights are increasing. Vice versa, the QoS attributes will be improved if they receive more concern. This complex interaction between user concerns and QoS attributes makes the selection of web services even harder. Literature [3-10] address the dynamic change of QoS, but to the best of our knowledge, there is no successful work for quantitive prediction of concern change.

So how to capture the QoS change trends of Web services and user concerns, and how to perform adaptive service selection based on these changes are two challenges for dynamic service selection.

In this paper, we address these problems and propose to use SEM (Structural Equation Modeling) to solve the multivariable equation group of QoS attributes and user concern on them, and then predict and validate their variables. This calculation can reduce the error in prediction due to information missing. For example the services provider may not publish service quality attribute or its concern. We extend SEM to predict QoS of web services.

Our approach is, first of all, analyzing the change of individual QoS attribute based on accumulated historical data and predicting its value in the coming time period. For SEM can provide a format to express the relationship of each variables, we can model QoS measurement based on SEM. This QoS measurement model can express the

978-0-7695-3801-3/09 $26.00 © 2009 IEEE
DOI 10.1109/WI-IAT.2009.363

relationship of QoS attributes and the user concerns, for efficiently analyzing the change of user concerns on every attributes and predicting the future value of the concerns. Based on this work, we use λ^2 examinations based on goodness-fit index, adjust it to examine the accuracy of prediction results, and define the concept of goodness-of-fit index for optimal services selection when the prediction results are close to each other. Our prediction mechanism can effectively overcome the difficulties in services selection in dynamic environment and decrease the "thrash" caused by adjusting services when their QoS descend.

The contributions of this paper include:

(1) Propose a modeling method based on SEM for QoS measurement of composite web services which can measure the changes of QoS accurately.

(2) Propose a quantitative prediction method of quality of web service by using the prediction mechanism of SEM in dynamic environment.

(3) Design a selection algorithm GFS (Goodness-Fit Selection algorithm) for optimal web services selection which can perform adaptive service selection.

The rest of this paper is organized as follows. Section 2 explains the motivation of our work by a simple scenario example. Section 3 briefly introduces SEM and constructs SEM to web services. Section 4 presents the web services selection method GFS based on the prediction mechanism of SEM. Section 5 shows the performance evaluation and comparison of different algorithms. Related works in this area is introduced in Section 6. The paper is concluded in Section 7.

2. Motivation Scenario

An example scenario is described in this section to explicitly express the effort of prediction technology for services selection and concerns for computing quality of services. To increase the QoS of synthesis services, fault-tolerance is used so that every activity has a number of candidate services. Let's assume that an activity, AirReservation, provides ticket booking and it has a set of candidate web services WS (including three services, ws_1, ws_2, and ws_3).

To ensure the customer satisfaction in using synthesis services, service with optimal QoS value should be selected in execution. Supposing, each of web services includes two quality dimensions, price and success rate for convenience. The formula of QoS then can be expressed as QoS = a_1*price + a_2*succRate, a_1 and a_2 are the user concern on price and success rate. Before the calculation, values of QoS attributes, price and success rate, need to be standardized by using the formalization method in the paper [3]. Obviously, quality of services will change if the change of price or success rate happens.

Now, we try to illustrate the impact on quality of service with the change of QoS attribute. Supposing, at time t_0, QoS attributes of WS (including ws_1, ws_2 and ws_3) are Price={0.5, 0, 1} and successRate={1, 0.5, 0} (after standardized). The user concern of price and success rate is {0.4, 0.6} and the vector {0.8, 0.3, 0.4} are value of QoS about WS. So the result of service selection is ws_1 because its QoS has the maximal value at t_0. But if the user adjusts the concern on price and success rate to {0.8, 0.2} at time t_1 (after t_0), the QoS values of WS are changed to {0.6, 0.1, 0.8}. The result of optimal selection is ws_3. From this example we can see the user concerns have great impact on the selection result.

Using above approach can select the execution service in candidate services at current time, but if the service are selected at t_0 and the composition service are executed at t_1, we do not have overall result to judge upon. So we need an approach to predict the trend of QoS to accurately select the optimal service at the execution time.

3. Forecast Modeling Based on Structural Equation

3.1. Overview of Structural Equation Modeling

Structural Equation Modeling is a multivariate statistical technique addressed by Bock and Bargmann (1969) for the first time. SEM has been successful applied in several scientific fields, like biological sciences, economics, and social sciences, etc. SEM can be used to clearly analyze the impact of individual variable on the entireness and the relationship between each individual variable. SEM integrates path analysis, confirmatory factor analysis and statistical testing methods in general and can analyze the causal relationship between variables, combined with the advantage of factor analysis and path analysis. SEM can analyze the relationship of equation groups, especially to the groups having causal relationship. So researchers can analyze and validate multiple set of variables with causal relationship by using SEM.

SEM can be used to solve the multivariable equation group of QoS attributes and user concern on them and predict and validate its variables. This calculation can also reduce the error in prediction due to information missing. For example the services provider may not publish quality attribute or its concern.

The process of calculating SEM is, first of all, through the manifest variables matrix, which is constituted by QoS and its multi-dimensions attributes, computing a covariance matrix S. The matrix S is the basis of SEM analysis. secondly, setting validation parameters (e.g., concerns of price) in SEM and estimating parameter matrix $\sum(\theta)$ from matrix *S*. If the theoretical assumption model and the estimated one from the sample demonstrate well fitness, the model is accepted. If good fitness does not exist, the model has to be changed. Iterating above process, researchers will finally reach a well fitted model [11, 12]. and finally predicting the change of $\sum(\theta)$ using the approach in [13].

By orthogonal transformation of covariance matrix which is computed to use manifest variables matrix, paper[13] raises the general issue about forecasting model. Assuming that there are samples at time point 1 to *T* and the sample covariance matrixes for each moment are calculated. As long as we can acquire the "sample" covariance matrix at moment *T+l*, then according to the system of structural equation we can derive covariance matrix at moment *T+l* of system's latent variables, then we can make conclusion on the system construction and causal relation of system's components.

3.2. Construction of Web Service SEM

In this approach, the selection of the web services that will execute an assigned task of a synthesis services is done at the last possible moment, without taking into account the other tasks involved in the synthesis services (candidate web services). When a task is actually executed, the system collects information and predicts QoS of each web service that can execute this task. After a task is actually executed, a quality vector is computed for each candidate web service. Based on these quality vectors, the system predicts change of quality of services by using SEM predicting method and selects one of the candidate Web service at last. In addition, because this selection process is based on concerns (weight) assigned by the user to each QoS attributes, the change of concerns value is considered in the approach too.

To illustrate the construction of web services SEM we use vector $X = \{X^1, X^2, X^3, X^4\}$ to indicate the four QoS attributes, X^1 = price, X^2 = reputation, X^3 = duration, X^4 = available. The vector *Y* indicates the value of the quality of services. The structural equation model is composed of *X* and *Y*. So, for each service, a model is constructed as shown Fig-1. In the model, *X*, *Y* are the manifest variables. η and $\xi = \{\xi_1, \xi_2, \xi_3\}$ are introduced as the latent variables of model. $\delta = \{\delta_1, \delta_2, \delta_3, \delta_4\}$ represents the error set generated in measuring quality targets *X*, $\delta_1, \delta_2, \delta_3, \delta_4$ are the errors that occur in measuring price, reputation, availability and duration respectively. ε represents error in the measurement of QoS target whose impact can also be neglected as δ.

ξ_1 abstracts QoS data pronounced by service supplier. It targets things like service execution prices advertised by service providers. ξ_2 is the QoS attribute aggregated by users' evaluations and feedback, such as reputation. The third kind is ξ_3. It denotes metrics like duration, success rate and available etc. The last latent variable is η, whose role is to establish contact between the former three latent variables and QoS.

$\Lambda_x = \{\lambda_1, \lambda_2, \lambda_3, \lambda_4\}$ describes the relationship between *X* and ξ. $\Lambda_y = \{\lambda_5\}$ describes the relationship between *Y* and ξ. Thus we construct a measurement model of web services by Λ_x, Λ_y, *X* and *Y*, and express it as formula 1.

$$\begin{cases} X = \Lambda_x \xi + \delta \\ Y = \Lambda_y \eta + \varepsilon \end{cases} \tag{1}$$

$\Gamma = \{\gamma_1, \gamma_2, \gamma_3\}$ describes the relationship between Λ_x and Λ_y. With Γ, Λ_x and Λ_y, we construct the structural model, and express it in formula 2.

$$\eta = \Gamma \xi + \zeta \tag{2}$$

According to definition of SEM in [13], we can construct a SEM of web service, with formula 1 and 2, and express it as in formula 3.

$$\begin{cases} \eta = \Gamma \xi + \zeta \\ \mathrm{X} = \Lambda_x \xi + \delta \\ Y = \Lambda_y \eta + \varepsilon \end{cases} \tag{3}$$

So, for a task, there is a set of candidate services $ws_j = \{ws_1, ws_2, \ldots, ws_n\}$ that can be used to execute this task. Each candidate service has *X* and *Y*. By merging the quality vector and QoS of all these candidate web services, a matrix $Q = [X_i, Y_i] = Q_{i,j}$, ($1 \leqslant i \leqslant n$, $1 \leqslant j \leqslant 6$) is built, in which each row Q_j corresponds to a web services ws_j while each column corresponds to QoS and quality dimension.

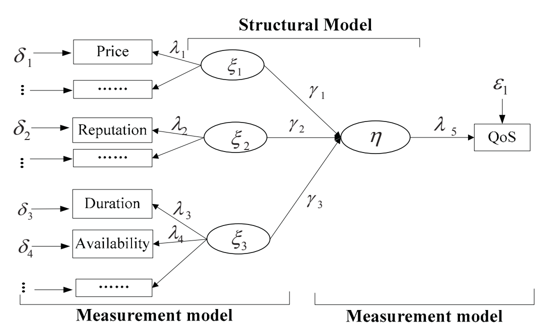

Fig- 1: Graphical illustration of SEM

Based on Q, we can calculate covariance matrix $\sum^t(\theta)$ with formula 1 and 4, and $\sum^t(\theta)$ is expressed as in formula 5.

$$\Sigma = \begin{pmatrix} Var(Y) & Cov(Y,X) \\ Cov(X,Y) & Var(X) \end{pmatrix} \quad (4)$$

$$\Sigma^t(\theta) = \begin{pmatrix} \Lambda_y Var(\eta)\Lambda'_y & \Lambda_y Cov(\xi)\Lambda'_x \\ \Lambda_y Cov(\xi)\Lambda'_x & \Lambda_x Var(\xi)\Lambda'_x \end{pmatrix} \quad (5)$$

Obviously above model is good for estimate parameters at any time point.

In our SEM, θ, which is calculated by $\sum^t(\theta)$, denotes the users' concerns. For example, we can achieve concern of price, reputation, duration and available with $\lambda_1 \times \lambda_5 \times \gamma_1$, $\lambda_2 \times \lambda_5 \times \gamma_2$, $\lambda_3 \times \lambda_5 \times \gamma 3$ and $\lambda_4 \times \lambda_5 \times \gamma_3$.

To note, we use static samples(i.e. example, we get QoS matrix in time t_0) only in this chapter, but the concerns of QoS attributes change in the dynamic scenario of network environment is more meaning full, and predicting concerns in coming moments with SEM is the key point in web services selection. We will discuss these problems in coming section 4.1.

4. Structural Equation for the Quality of Web Services

In this section, SEM is firstly shown that it is applicable to calculate the concerns of Web Service. Based on SEM, we analyze relationships of QoS, user concerns and quality attributes. After that we describe forecast algorithm in detail. We show the prediction of user concerns by our forecast algorithm mainly. The discussion on the evolution of quality attributes is briefly based on [3].

4.1. SEMSS: a Services Forecast algorithm

According to the structural equation model given in section 3.2, we have $Q^t = [X_i, Y_i]^t = (Q_{i,j}^t, 1 \leqslant i \leqslant n, 1 \leqslant j \leqslant 6)$ as the QoS matrix of candidate web services of task tk_j at time t_0. and S^t is calculated from Q^t, and express it as $S^t = G^t\Lambda^t(G^t)'$ at time t, in which G^t is the unit orthogonal matrix, and its column vector is the unit orthogonal vector of variance matrix S^t at time t. Λ^t is the eigenvalue matrix at time t, its diagonal elements are the eigenvalue, the other elements are 0. Eigenvalue λ_i^{T+l} at time $T+l$ can be forecasted if there is regression equation for aigenvalues $\hat{\lambda}^t{}_i = f_i(t)$, $i = 1,2,\cdots,p$. Unit orthogonal transformation matrix can be seen as the rotation in multi-dimensional space, so each element $G^t{}_{ij}$ of orthogonal matrix G^t denotes an angle $-\frac{\pi}{2} \le \theta^t{}_{ij} \le \frac{\pi}{2}$. After making a regression equation $\theta^t{}_{ij} = f_{ij}(t) + \varepsilon^t{}_{ij}$, for each time series $\theta^t{}_{ij}$, $1 \le i < j \le p$, $t = 1,2,\cdots,T$, we can forecast at time $T+l$.

By using the Maximum Likelihood Method （ML） we have the fitting function FML as follows:

$$FML() = (1/2)n[tr(\hat{R}^{T+l}(\Sigma(\theta)^{T+l})^{-1}) + \ln\left|\Sigma(\theta)^{T+l}\right| - \ln\left|\hat{R}^{T+l}\right| - (p+q)]$$

The user concern $a^{t+l}{}_i$ at the future time $T+l$ is predicted by the iteration of FML. And based on [6,7], the QoS attributes v^{t+l} can be computed. The QoS value of ws_j at the future time $T+l$ is calculated by $\sum_{i=1}^{p} a_i^{t+l}(v^i{}_j)^{t+l}$ (subscript j for the service number, super-script i for attribute number. Superscript $T+l$ denote future time.).

Following is the web services selection algorithm based on SEM: SEMSS（Structural Equation Modeling Services Selection Algorithm）

Algorithm: *SEMSS*

Input: $[X^j\ Y^j]^t$, s_j: Calculated Predict Results and Selection

Output: *A(s)*

1. $[X^j\ Y^j]^t$ **to** S^t;//$[X^j\ Y^j]$ is measurement sample in time t
2. $S^t = U^t T^t (U^t)'$
3. $U^t \to U^{t+l}$;
4. $T^t \to T^{t+l}$;
5. $\hat{U}^{T+l}\hat{T}^{T+l}(\hat{G}^{T+l})' \to S^{T+l}$;
6. **calculated** $\widehat{\Sigma(\theta)}^{T+l}$ **of** S^{T+l};
7. **evaluate** $\hat{\theta}$ with *FML*;// $\hat{\theta}$ is concerns about QoS **attributes and** θ **include** $\{\lambda_1,\lambda_2,\lambda_3,\lambda_4,\lambda_5,\gamma_1,\gamma_2,\gamma_3\}$
8. **prediction**(X_j^{t+l}) ; // X_i is exogenous manifest variable, that is properties of QoS in time $T+l$ using methods in [6,7]
9. **compute** *P(s_j)* **with** X_j^{t+l} **and** $\hat{\theta}_j$;// *P(s_j)* is prediction of QoS in $T+l$;
10. **selection max of** *P(s_j)* **to** *A(s)*;
11. **output the** *A(s)* and **exit**;

4.2. GFS Algorithm: an improved web services selection algorithm based on SEMSS

Forecasting future QoS will help users chose the best service. Future QoS may forecast by SEMSS approach and it is easy to choose the best service if there exist big different quality among services. How do people choose if the difference varies at a time period?

The service selection is reasonable if following condition occur:

(1) During a past period of time, the forecasting value fits the measured sample.

(2) Difference of forecasted value is big and clear enough, so correctness of choice is guaranteed.

We determine condition by using formula, where GFI is the metric to examine the match degree of sample covariance matrix S and the covariance matrix gotten by structural equation model. If sample covariance matrix S equals to predicted covariance matrix then GFI is set to 1:

$$GFI = 1 - \frac{FML\left[S, \Sigma(\hat{\theta})\right]}{FML\left[S, \Sigma(0)\right]} \tag{6}$$

Where $FML[S,\Sigma(\theta)]$ is the fitting function between the sample covariance and the covariance got from the model. $FML[S,\Sigma(0)]$ is the fitting function between sample covariance and null model, i.e. a model where manifest variables are independent each other, Obviously GFI is varies with time. If forecast value has a good fit with sample value in the passed period of time we then may think condition-1 is satisfied.

$CFI = \frac{1}{L}\int_T^{T+L}(\sum_{i=1}^{p-1} a_i^t X_i^t)dt$ is the criterion in condition-2. It is mean QoS within the future period [T, $T+l$].

The input of GFS (Goodness-Fit Selection) includes data table $[X^j\ Y^j]^t$ in time point t, redundant services sequence S_j, services duration $T+l$. The output is the selected service. Steps are:

(1) Predict covariance matrix $\Sigma(\theta)^{T+l}$ at time point $T+l$. Calculate θ if it matches rule t.

(2) Calculate the goodness-of-fit index and adjusted goodness-of-fit index of $\Sigma(\theta)^{T+l}$.

(3) Calculate the GFI of each service by the sorting result.

(4) Choose the optimal service, i.e., with largest GFI, at the time point $T+l$.

The major time complexity of GFS arises from the calculation in the prediction on SEM. If an activity consists of m available services, each of the services has a $n \times n$ QoS data table, the worst time complexity is $O(m \times n \times 2)$. Small time complexity indicates better performance.

Algorithm: *GFS*

Input: $[X^j\ Y^j]^t$, s_j: Calculated Predict Results and Selection

Output: *A(s)*

1. $[X^j\ Y^j]^t$ **to** S^t**;**//$[X^j\ Y^j]$ is measurement sample in time t
2. **transfer** S^t **to** S^{t+l};
3. **evaluate** $\hat{\theta}$ with $\hat{R}_j^{T+l}$ of $\hat{\Sigma}(\hat{\theta})^{T+l}$;// $\hat{\theta}$ include $\{\lambda_1,\lambda_2,\lambda_3,\lambda_4,\lambda_5,\gamma_1,\gamma_2,\gamma_3\}$
4. **prediction**(X_j^{t+l}) ; // X_i is exogenous manifest variable, that is properties of QoS in time T+l using methods in [6,7]
5. **compute** $P(s_j)$ with X_j^{t+l} and $\hat{\theta}_j$;// $P(s_j)$ is prediction of QoS in $T+l$;
6. **if** GFI $\notin$ [0.9, 1]
7. **then** remove *service* in s_j; // don't compute with SEM;
8. **else**
9. **arrange** s_j **order** by GFI **sorted** by $P(s_j)$;
10. **compute** GCI of s_j **in** $P(s_j)$;
11. **end if**
12. **select** max of $P(s_j)$ to *A(s)*;
13. **output** A(s) and **exit**;

5. Performance Analysis and Evaluation

5.1. Simulation Setting

To evaluate the effectiveness of our selection algorithm, simulations are conducted to analyze the predicted results of GFS and compare them to the actual data for validation. We develop and deploy web services with SOArWare and queries randomly. The SOArWare is one of production of our key project, service-oriented software product line. QoS data is retrieved by using method of [3, 6] and each web service has information including name, release, type, location, price, credit, duration, availability and success ratio, etc. The run time of web services is set to 100 days. In this period of time the attributes of QoS are monitored daily.

The sample matrix of web services is defined in section 3.2. In the following, we give a measurement parameter for examination accuracy of SEMSS and GFS algorithm.

5.2. Measurement Parameters

Error ration ω is a simulation parameter set for data analysis and defined by the following expression:

$$\omega = \frac{\left|Q_p - Q_r\right|}{Q_r} \tag{7}$$

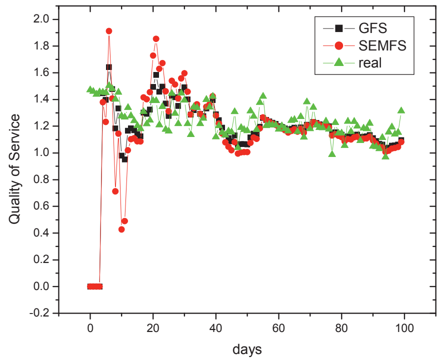

Fig-2 Prediction test in static environment

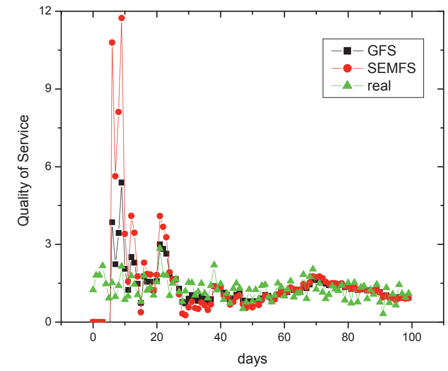

Fig-4 Prediction test in dynamic environment

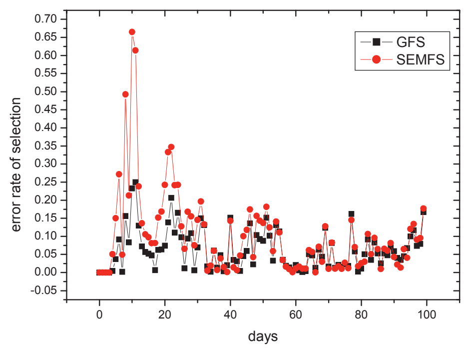

Fig-3 error rate of GFS and SEMSS in static environment

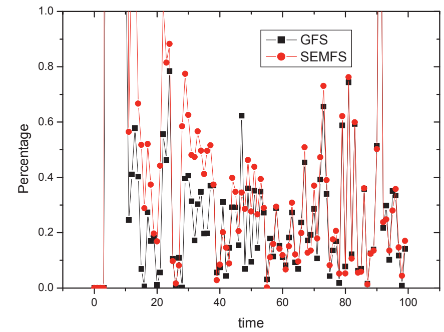

Fig-5 error rate of GFS and SEMSS in dynamic environment

In the expression, Q_r denotes the actual value of QoS obtained from the measurement in evolvement of the services, and Q_p denotes the predicted value gained through the SEMSS and GFS algorithm based on SEM on the simulation nodes. The difference ratio between the predicted value and actual value can be calculated by the expression. In our simulation, lower error ratio means higher accuracy of prediction, which proves the effectiveness of the selection algorithm.

5.3. Results Analysis

The first simulation is to analyze the difference between the QoS prediction mechanism and the actual execution results of the web services, under the assumption that the weights of QoS attributes are fixed or updated with a run time of 100 days and four attributes, namely the price, credit, availability and ratio of successful execution. The prediction of the service status of the coming second and fourth day starts on the day 5 and is compared to the actual measured data. As shown in Fig-2, in the first 30 days where the accumulated QoS data of the services is not enough, the difference between the predicted value and the actual value is rather big, while after that the difference is reducing gradually and becoming stable. As shown in Fig-4, similar to the first simulation, in the first 25 days where the accumulated QoS data of the services is not enough, the difference between the predicted value and the actual value is rather big, while after that the difference is reducing gradually and becoming stable.

The result in Fig-2 show that prediction algorithm based on the SEM in *T+l* closed real measurement QoS with accumulation of data in static environment. The result in Fig-3, through goodness-fit index, the error rate of GFS algorithm is better than SEMSS in prediction selection.

The second simulation assumes that the weights of QoS attributes change in the different time period. For example in the initial phase users may prefer services lower price, but as the use frequency and service quality increase, the weight of price decrease and the weight of execution success increase. That is the second simulation

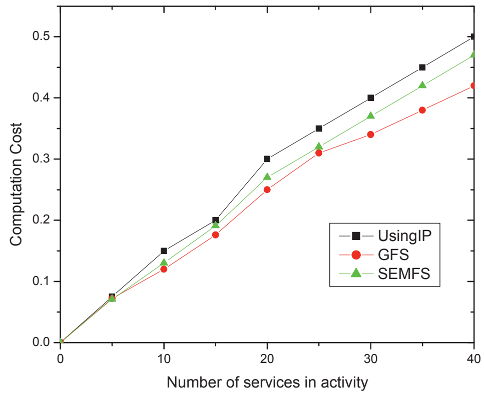

Fig-6 Experimental results (computation cost) in a dynamic environment, varying the number of services per task.

is to validate that is the second simulation is to validate the difference between the QoS prediction mechanism and The prediction of the service status of the coming second and fourth day starts on the day 5 and is compared to the actual measured data. Fig-4 shows that the impact of measurement error decreases as the data accumulated. Fig-5 shows that the impact of measurement error is slim in dynamic environment though.

The result in Fig-4 show that predicted and actual value in the first 30 days is caused by the lack of historical data that can be referenced to. But when the data is accumulated to a certain amount, the predicted value is inclined to the actual value stably.

As shown in Fig-5, GFS shows good adaptability in dynamic environment and the error ratio is acceptable and negligible because though goodness-fit index in selection.

The third experiment is to analyze computation cost between prediction selection algorithm and IP (Integer Programming). The results in Fig-6 show that GFS and SEMSS have insignificant difference in computation cost. In third experiment, we increase number of candidate services from 5 to 40, and record of computation cost in difference web services selection algorithm (IP, GFS and SEMSS) in dynamic environment.

6. Related Work

To achieve optimal selection result, many efforts have been made. Related work mainly involves service composition, dynamic selection, dynamic predicting and adaptation technology, monitoring and recovery technology, evaluation, replication technology and so forth.

Dynamic service composition method is proposed by L. Zeng and H. Sun [3,4]. In the definition period only required functions are defined. Component services are bound and instantiated at runtime. Composite service communicates with service registry dynamically according to the pre-established strategies. These strategies include choice based on QoS [14], or based on semantic [15] and so on. These methods improve the flexibility and adaptability in selection. But they fail to solve dynamically searching. The multiple remote interactions needed and the efficiency will be low. Because the existing service registries cannot guarantee the authenticity of data and the state of registered services, the quality of services can not be guaranteed.

H.P. Guo [6] proposed a method for adaptive maintenance of composite services in dynamic environment. The method use modeling the control process as a Markov decision process (MDP), and then predict service state with algorithm based on Kalman-Filter. L.S. Shao [7] proposed predictive methods based on making similarity mining and prediction from consumers' experiences. Their approach can make significant improvement on the effectiveness of QoS prediction for web services. However, they fail to support the adaptive concerns of attributes in dynamically environment.

S. Guinea [17] presented a self-healing method through service state monitoring and recovery of fails. V. Issarny and F. Tartanoglu [18] proposed a method to achieve fault tolerance and dependable composite service for forward error recovery and based on the concept of cooperative atomic action and web service composition action. But they unconsidered monitoring and recovery efficiency, costs and rewards, and also they do not assess the effect of the monitoring and recovery in quantitative forms.

In addition, WT. Balke and W. Matthias [19] proposed service usage patterns and enhance service discovery. Z. Maamar, S.K. Mostefaoui [20] presented a model that highlights the resource on which the web service is performed for the context of web service interactions. These researches focus on providing a mechanism to formalize service consumers' preferences, history, scenario and resource of service provider. The work bases on the pre-setup semantic and ontology models that need much manual work.

7. Conclusion and Future Work

To implement the optimal of web services selection, we analyze the change of individual QoS attribute based on accumulated historical data and predict its value in the coming time period. For SEM can provide a format to express the relationship of each variables, in this paper we model QoS measurement model using SEM to express the relationship of QoS attribute and users' concern, so that we can analyze the change of users' concerns on every attributes and predict the future value of the concerns. Based on these work, we use goodness-of-fit

λ^2 examination, goodness-of-fit index and adjusted goodness-of-fit index to examine the accuracy of prediction results and define the concept of closeness-of-fit index for optimal services selection when the prediction results are close to each other. Our prediction mechanism can effectively overcome the difficulties in services selection in dynamic environment and decrease the "thrash" caused by adjusting services when their QoS descend.

The major contributions of our paper provide, first of all, a modeling method based on structural equation model for QoS measurement of composite web services, and then a quantitative prediction method of web services QoS by using the prediction mechanism of structural equation model in dynamic environment. At last, we design a selection algorithm GFS for optimal web services selection.

Although significant achievements have been made so far, there are several interesting directions in our future work. We will investigate effect of extended SEM, e.g., to define parameters more specific to Web services in more flexible way. On the other hand, we will consider to more abundant composition model other than supporting the general mode of sequence, parallel, choice and iteration, e.g., more complex relationships, such as compensation and transaction.

8. Acknowledgement

This research was supported in part by China 863 High-tech Programme (SN:2007AA010301) and (SN: 2009AA01Z419).

9. References

[1] F. Curbera, M.J. Duftler, R. Khalaf, W. Nagy, N. Mukhi, and S.Weerawarana. IEEE Internet Computing: Spotlight - Unraveling the Web Services Web: An Introduction to SOAP, WSDL, and UDDI. IEEE Distributed Systems Online 3(4): (2002)

[2] Y. Liang, Y. Yan, and H. Liang. Composing Business Processes with Partial Observable Problem Space in Web Services Environments. In Web Services. International Conference on, pages 541 - 548, 2006.

[3] L. Zeng, B. Benatallah, A. H.H.Ngu, M. Dumas, J. Kalagnanam, and H. Chang. QoS-Aware Middleware for Web Services Composition. IEEE TRANSACTION ON SOFTWARE ENGINEERING, VOL 30, NO. 5. MAY 2004

[4] H. Sun, X. Wang, B. Zhou, and P. Zou. Research and Implementation of Dynamic Web Services Composition, APPT 2003, LNCS 2834, Springer-Verlag Berlin Heidelberg, pp.457–466, 2003.

[5] D. Mennie, and B. Pagurek. An architecture to support dynamic composition of service components. Proceedings of the WCOP. Sophia Antipolis, France, 2000.

[6] H.P. Guo, J.P. Huai, H. Li, T. Deng, Y. Li, and Z.X. Du: ANGEL: Optimal Configuration for High Available Service Composition. ICWS 2007. pp:280-287, 2007.

[7] L.S. Shao, J. Zhang, Y. Wei, J.F. Zhao, B. Xie, and H.Mei. Personalized QoS Prediction forWeb Services via Collaborative Filtering. ICWS 2007, pp.439-446, 2007.

[8] H. Harney, and P. Doshi. Speeding up Adaptation of Web Service Compositions Using Expiration Times. WWW (2007), Pages: 1023 – 1032, 2007.

[9] R. Jurca,, W. Binder, and B. Faltings. Reliable QoS Monitoring Based on Client Feedback. WWW (2007), Pages: 1003 - 1012 , 2007.

[10] K.Q. Wu, D.J. Lilja, and H. W.Bai. The Applicability of Adaptive Control Theory to QoS Design: Limitations and Solutions. IPDPS'05, ipdps,pp.272b, 2005.

[11] Ji.N. Li. Introduction of structural equation model. Hefei: Anhui University Press, 2004

[12] Z.G. Guo. Social statistic analysis methods-SPSS software application [M]. Beijing: China Renmin Universtiy Press. 2004

[13] H.W. Wang, and Y. Zhang. Forecast modeling for structural equation model. Journal of Beijing University of Aeronautics and Astronautics vol33, No.4. 2007

[14] L.Z. Zeng, H. Lei, and H. Chang. Monitoring the QoS for Web Services. ICSOC, pp. 132-144, 2007.

[15] S. Ran. A Model for Web Services Discovery with Qos. ACM SIGecom Exchanges. V4, I 1, pp. 1-10. 2003

[16] D. Wu, B.J. Parsia, E. Sirin, J. Hendler, and D. Nau. Automating DAML-S Web services composition using SHOP2. Proceedings of ISWS(2003), pp. 195-210, 2003.

[17] S. Guinea. Self-healing web service compostions. proceedings of the ICSE., pp. 655-655, 2005

[18] V. Issarny, F. Tartanoglu, A. Romanovsky, and N. Levy, . Coordinated forward error recovery for composite web services. Proceedings of the SRDS, pp. 167-176, 2003

[19] WT. Balke and W. Matthias. Towards Personalized Selection of Web Services. WWW, 2003.

[20] Z. Maamar, S.K. Mostefaoui, and Q.H. Mahmoud. Context for Personalized Web Services. In System Sciences. Proceedings of the 38th Annual Hawaii International Conference on, pp. 166b– 166b, 2005.