Abstract

With the proliferation of Web services as a business solution to enterprise application integration, ranking and selecting the best Web services among the providers become an important factor in the success of the business solution. Quality of Service (QoS) determines the quality and usability of a service including its price, performance, reliability, integrity, accessibility, availability, interoperability, and security. Given a set of QoS attributes from a variety of sources, it is a challenge to sort through all of them and be able to get the best services that meet QoS requirement. In this paper, we describe a novel method by which Web services can be ranked and selected automatically based on a number of observed QoS parameters and feedback responses learned from prior knowledge. This new approach treats the observed Web services QoS attributes and target Web services relationship, represented by a matrix, as a statistical problem. Using Singular Value Decomposition (SVD) technique, and an user assisted weighting system, implicit higher order correlations among Web services and their associated QoS attributes are extracted and used to estimate the selection of recommended Web services.

1. Introduction

Web services enable software to be delivered and paid for as dynamic flow of services in contrast to fixed and pre-packaged products. It also enables the emergence of a new class of software developers and businesses that focus on delivering software as services via Web services. Web services greatly reduce the barrier to provide software and at the same time significantly increase its potential customer base. Such new paradigm in providing software functionalities, installation and upgrades alleviates businesses from the costly, complex and tedious tasks of traditional software integration, thus allowing them to focus on their primary businesses.

With the widespread availability of Web services, quality of service (QoS) [1] becomes a significant factor in selecting the best services among the providers, it constitutes the most important differentiating point for a set of Web services which provide similar functions. QoS covers a wide range of attributes that match the needs of service requestors with those of the service provider's. QoS generally refers to non-functional properties of Web services such as price, performance, reliability, integrity, availability, accessibility, interoperability, security, etc. [2] All these attributes are more “abstract” compared with functional requirements and are often subject to interpretation. A detailed description of these properties and the methods used to collect their ratings is beyond the scope of this paper. In this paper, we assume that these QoS attributes are collected by various reputable agencies and they are domain specific and readily available.

Traditional approaches utilize QoS policies [3] or some weighting system to sort through a pile of attributes in the hope of finding the ones that best fit the QoS requirements. This simple approach is acceptable if the number of attributes and the number of available services are relatively small. More importantly, not all QoS attributes that the Web services users require are available, and this leads to a problem in finding the best Web services given a QoS attribute requirement that may not be available in all related Web services. In addition, due to the more “abstract” nature of QoS attributes, a rule-based model [4] of getting the best Web services may be too restrictive and not be able to provide a comprehensive set of alternatives. In this paper, we propose and describe a new approach to facilitate the automatic selection and composition of Web services based on QoS by using Singular Value Decomposition technique (SVD) [5]
together with an adaptive weighting system that can react to newly available Web services and changing market conditions. We also introduce the concept of Quality Matrix for Web services as a model for capturing QoS knowledge for each of the Web services being monitored. SVD is a classical statistical method and is widely used in latent semantic analysis for information retrieval and image recognition [6]. Its use in autonomic system [7] for autonomic policy selection has been investigated recently. However, its application in ranking and selecting Web services has not been seriously explored. Our proposed Web services ranking and selection system enables dynamic interaction between the available QoS metrics and the weighting system with its weight entries being updated to enhance correlation among the data. We further introduce the concept of “Web Service Quality index”, a user defined weighting factor assigned to each attribute in the QoS metric to represent its relative importance.

In the rest of this paper, section 2 describes the system design and architecture. Section 3 describes the Quality Matrix for Web services. In section 4, we use an example to illustrate the creation of a QoS metrics and Web Services matrix and space by SVD technique. We then describe the Web services selection and composition process, and given a new set of attributes, how old data is mined to provide a set of possible Web services, and new knowledge is acquired to enable adaptation in section 5. In section 6, we describe the result. Section 7 discusses possible future work and concludes.

2. System Overview

![System Overview Diagram](image)

Fig.1 shows the overview of the proposed system. It includes a QoS metrics and Web services (QW) repository for storing QoS attributes and Web services patterns. A QW matrix module transforms an initial QoS attributes and Web services set obtained from the QW repository into an QW matrix, which is then decomposed by SVD transformation to form an n-dimensional QW space wherein QoS attributes and closely associated Web services are placed near one another. Web services which are closely associated with the required QoS attributes in this space are selected. Users can accept, reject or modify the recommended Web services. Each episode of successful Web service selection triggers the system to update the QW repository by putting a positive or negative weight on certain QW patterns.

3. Quality Matrix for Web Services

The Quality Matrix is used to capture the collection of QoS attributes associated with the set of target Web services. This matrix represents the Web services and their QoS metrics as mathematical objects, which captures the relationship among the Web services and their QoS metrics in a simple, compact and extendable form. Such representation enables fast and easy computation as there are numerous numerical analysis tools [8,9] and algorithms [10,11] readily available and therefore eliminates the need to develop the tools needed to support the knowledge model. It is extensible, adding and/or removing columns and rows does not alter the original data structure and meanings of the entries. It is also a compact way to capture knowledge since it preserves the intrinsic or latent relationship among the entries [12]. In our QW matrix, each cell contains a numerical index which represents a weighted rating of a QoS attribute of a Web service. The higher the number, the better the Web service with respect to the attribute. Since different rating agencies provide different rating systems and formats for each of the QoS attributes, such as either “positive”, “negative” or “neutral” as in Ebay, an user review index of 1 to 5 as in Amazon. We must convert these different rating representations into a single representation for processing. For each cell in the matrix, we therefore define a Quality Index (QI) which is a weighted average of values from selected rating agencies for a particular attribute:

\[ QI = W_1 Q_1 + W_2 Q_2 + \ldots + W_n Q_n / n \] (1)

Where
Qn = normalized, common quality index between 0 and 1.

Wn = weight assigned to the specific agency

Wu = weight assigned by the user or automatically based on historical data for the rating of a specific attribute

With M available Web services and N attributes in a QoS metric, our Quality and Web Services (QW) Matrix is an MxN matrix which can be readily decomposed by SVD technique. The decomposition enables the original MxN matrix to be approximated by the first k eigenvalues of the diagonal matrix (one of the 3 resulting decomposed matrices) of the original MxN matrix, resulting in a compressed representation of the original data. Using the compressed data, clusters of Web services with similar qualities can be identified and ranked accordingly.

4. Creation of QoS Metrics and Web Services Space

For ease of illustration and simplicity, a simulated set of data from earlier work on policy and document search was used as sample QoS metrics and Web services data set as shown in Figure 2.

E1 = Price
E2 = Availability
E3 = Response Time
E4 = Security
E5 = Accessibility
E6 = Problem Handling
E7 = Failure Rate
E8 = Integration Effort
E9 = Billing Accuracy

A1 = Web service from Provider 1
A2 = Web service from Provider 2
A3 = Web service from Provider 3
A4 = Web service from Provider 4
A5 = Web service from Provider 5

<table>
<thead>
<tr>
<th></th>
<th>A1</th>
<th>A2</th>
<th>A3</th>
<th>A4</th>
<th>A5</th>
</tr>
</thead>
<tbody>
<tr>
<td>E1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>E2</td>
<td>1</td>
<td></td>
<td></td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>E3</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>E4</td>
<td></td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>E5</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>E6</td>
<td>1</td>
<td></td>
<td></td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>E7</td>
<td></td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>E8</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>E9</td>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td></td>
</tr>
</tbody>
</table>

This dataset consists of m quality attributes (Em) and n Web services (An), where m=9 and n=5. The m quality attributes are entered as rows and the n Web services are entered as columns in the MxN matrix R. The entries in the QW matrix are simply QI of a particular QoS attributes for a specific Web service. For simplicity, QI = 1 (QI ranges from 0 to 1) is used; empty cell indicates the QoS rating for that particular attribute is not available for the associated Web service.

R is decomposed into three matrices [14, 15] by SVD as in Equation (1),

\[ R = E S A' \]  

E and A’ are the QW matrices of left and right singular vectors as shown in Figures 3 and 4, respectively. Both of them have orthogonal columns. As shown in Figure 5, S is the diagonal matrix of singular values ordered in decreasing magnitude. These special matrices are the result of a breakdown of the original QoS metrics and Web services relationships into linearly independent QoS metrics and Web services components. So each QoS metric and Web service is represented by its vector. As shown in Figures 5, the values for many of these singular values can be ignored as they become relatively small (weak correlation). Usually, only the first few largest singular values are needed and the rest deleted. Thus, a reduced model which approximately equals to the original QW model with fewer dimensions can be built. This process, in essence, captures the major relationships among QoS attributes and Web service while ignoring the minor ones by treating them as noise.
Figure 4. A’ Matrix

<table>
<thead>
<tr>
<th>2.47</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1.84</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.62</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.12</td>
<td>1.2</td>
<td>0.76</td>
</tr>
</tbody>
</table>

Figure 5. S Matrix

In a two dimensional model where k = 2 as shown in the shaded elements in Figures 3, 4 and 5, all the QoS attribute to QoS attribute, Web service to Web service, and QoS attribute to Web service similarities are now approximated by the first two largest singular values of S. As a result, the row vectors of the reduced matrices (shaded columns of the E matrix in Figure 3 and A’ matrix in Figure 4) are taken as coordinates of points representing QoS attribute and Web service in a two-dimensional space as shown in Figure 6 where QoS attributes are represented as *diamonds* and Web services as *squares*. The dot product or cosine between two vectors representing any two components corresponds to their estimated similarity.

5. Selecting and Composing Web Service Based on a New set of QoS Attributes

When the system receives a QoS attribute set which matches exactly one or more of the existing attribute set in the QW repository, the system simply retrieves its corresponding Web services directly from the QW repository. However, when a new set of QoS attributes is required with no exact match in the QW repository (no new QoS attributes), the Web services selection mechanism in the QW space is invoked. Using the required QoS attribute set, a pseudo-Web service vector is constructed as the weighted sum of its constituent attribute vectors. With appropriate rescaling of the axes, this amounts to placing the pseudo-Web service at the centroid of its corresponding QoS attribute points. Then, this pseudo-Web service is compared against all existing Web services by calculating the cosine between the pseudo-Web service vector and the existing Web service vectors as a similarity metric. Those Web services with the highest cosines (the nearest vectors) to the pseudo-Web service are selected. The resulting Web services are ranked according to their closeness to the pseudo-Web service vector and merged to form the recommended Web services set. Clearly, the choice of the threshold cosine value plays a significant role in the number and the accuracy of the Web services selected. The common practice is to use a small cosine value to enable a broader search space initially, and reduce the search space gradually as more data is accumulated to maximize accuracy.

When a new required QoS attribute set includes at least one new QoS attribute, the system excludes the new QoS attribute and only uses the existing QoS attributes to form the pseudo-Web service and select the recommended Web service as described in the previous paragraph. But this recommended Web service will be examined by the user. The user, at his discretion, can accept, modify this recommended Web service. Upon acceptance of this recommended Web service, the new Web service is recorded in the QW action repository and triggers the system to re-construct a new QW space which includes the new QoS attribute for subsequent uses. As a result, new knowledge is acquired and captured by the new QW matrix.

Using the SVD based Web services ranking and selecting system, we can expect a collection of Web services, with each of its member Web services ranked and selected individually, constitutes the best available set for a solution. This implies that a solution consists of multiple Web services can in fact achieves the best matched QoS requirement.

6. Results

With $k = 2$ and a threshold cosine value of 0.7, the following examples illustrate the
operation of the system under different conditions:

1) A QoS attribute set consists of E2 and E3, (user is interested solely in Response Time and Availability) a direct match of A2 (Web service from Provider 2) is found and the system simply returns A2 as the recommended Web service. It is also interesting to note that relevant Web services can be retrieved depending on their proximity to the pseudo Web services formed by E2 and E3, despite the fact that an exact match is found. This is useful in the case that the user needs to consider a broader spectrum of recommended Web services.

2) An observed QoS attribute set consists of E4 and E5 (user is interested in Security and Accessibility). A search indicates that there are no matching Web services in the current repository. A pseudo-Web service is constructed from E4 and E5, represented as point q in Figure 6 which is the centroid of vector E4 and E5. A2, A3 (Web service from Provider 2 and 3) are selected as they are within the dotted cone with a cosine value of 0.7 from q.

3) An observed QoS attribute set consists of E6, E9 (user is interested in Problem Handling Capability and Security) and a new QoS attribute E10 (Provider Reputation, not in the current QoS attribute list). The system uses E6 and E9 to form the pseudo Web service represented as point f in Figure 6. Using a cosine value of 0.7 from f, A5 (Web service from Provider 5) is selected as the recommended Web service. If user accepts this recommended Web service, it will trigger the recreation of QW matrix and new knowledge is acquired. This is interpreted as even the new QoS attribute is not available from the OoS list, its computation is based on the other 2 available attributes. The new QoS attribute will be included in the QoS attribute list and the relationship with other attributes and Web services will be established as user feedback is obtained.

7. Discussion and Conclusion

This paper introduces a system for Web service ranking and selection using a statistical approach with SVD. It provides a compact and efficient knowledge representation mechanism to represent QoS attributes and Web services relationship by using matrix and dimension reduction technique to extract meaningful relationship among QoS attributes and the corresponding Web services. It enables the selection of Web services without exact match of the required QoS attributes. Compared to other statistical approaches, the advantages of using SVD technique includes: 1) SVD is an automatic algorithm which does not require a formal knowledge model. 2) It can be used in any applications without customization. 3) It guarantees the best fit. 4) Increasing and decreasing the precision of the fit is relatively straightforward (e.g. by adjusting the threshold cosine value). 5) Theories of SVD and its variants have been well established. 6) Its uses in numerous different applications prove its usefulness 7) Tools for SVD transformation and analysis are readily available. However, for large matrix, it is computational intensive and updating the system [15] to reflect new data in a large matrix is a challenge. In addition, the automatically generated correlations among QoS attributes and Web services are sometimes not obvious to humans, especially when the orthogonal factors are large.

Initial simulated dataset yields promising results with the number of orthogonal factors k used in the reduced model chosen to be two to represent a 2-dimensional space. However, the representation of a conceptual space for any large QoS metrics and Web services collection usually requires a fairly large number of orthogonal factors. Finding a balance point between accuracy and coverage is a challenge as well as an art. In addition, the choice of the cosine value in locating recommended Web services relative to the pseudo Web service in the QW space is also an important factor as this parameter determines the level of accuracy of the selected Web services. The relationships among the number of QoS attributes, Web services, factor k and cosine values are interesting topics for research.

8. References