A FOUR-LEVEL LINEAR DISCRIMINANT ANALYSIS BASED SERVICE SELECTION IN THE CLOUD ENVIRONMENT

S. Bharath Bhushan¹, Pradeep Reddy C.H.¹*

¹ School of Information Technology and Engineering, VIT University Vellore-632014, Tamil Nadu, India

(Received: March 2016 / Revised: June 2016 / Accepted: July 2016)

ABSTRACT

The cloud is an outstanding platform to deal with functionally equivalent services which are exponentially increasing day-by-day. The selection of services to meet the client requirements is a subtle task. The services can be selected by ranking all the candidate services using their network and non-network Quality-of-Service (QoS) parameters, which is formulated as a NP hard optimization problem. In this paper, we proposed a linear discriminant analysis (LDA) based a four level matching model for service selection based on QoS parameters, which includes description matching of a service, matchmaking phase, LDA-based QoS matching and ranking. The LDA-service selection agent is deployed on each cloud to classify services into classes and rank the services based on the aggregate QoS value of each service. Finally, the test results show the efficiency in service selection with minimal discovery overhead, significant reduction in the computation time and the number of candidate services to be considered.

Keywords: Cloud computing; Linear Discriminant Analysis; Quality of Service; Ranking; Web service

1. INTRODUCTION

The merits of cloud powers a significant increase in number of similar services and service providers. The main cloud services: software, platform and infrastructure are provided as services on the basis of pay per use. So, selection of service provider and service selection plays a crucial role in their business activities. As the services are self-contained, loosely coupled processes deployed over a standard middleware platform can be described, published, discovered and invoked over a network. The challenge associated with cloud is in selecting the optimum required services, which are provided by different service providers with different QoS.

The selection of atomic service from a large number of similar services with a different quality of service is a multi-criteria decision problem Rajeswari et al. (2014). In this paper, we considered QoS parameters, such as response time, throughput, availability, successability and price Zhou et al. (2013). We proposed a four-level matching model to select a service from an optimal number of candidate services based on QoS criteria. The objective of our work is to achieve an efficiency in service selection with minimal discovery overhead, significant reduction in the number of candidate numbers to be considered and computation time.

Normally, the similarity computation is performed between a service request and published service based on which service is selected. In this process, the service is filtered based on service request
and then it assesses the QoS parameters of a service by selecting a criteria to obtain the overall score of filtered services. In order to reduce the number of candidate services and discovery overhead, the services are clustered based on QoS values by using linear discriminant analysis. None of the work has been carried out in LDA on service selection. The LDA is used for service discovery by reducing the dimension of service data and ignores the inequality of local data points of a similar class using matrix representation and calculates the overall score of QoS for each service, then ranks the services with the highest QoS values Izenman (2013). The intention of linear discriminant analysis is to detect a class to which services should belong to the closest mean. If the classes pose equal covariance and the predictor variables are subject to multivariate normal distribution, then LDA works more efficiently than any other of the discriminant analysis methods. In order to filter the candidate web services, we applied the term frequency-inverse document frequency (TFIDF) algorithm to compute the similarity degree between the service request and services published.

The contribution of this research is as follows:
1. The similar services are grouped into classes according to their QoS parameters
2. It will reduce discovery overhead
3. It reduces the number of candidate services
4. A significant reduction in computation time
5. It will ensure near optimal solution by selecting the best service from selected class

The rest of the paper is organized as follows: Related work about service selection based on QoS parameters in Section. 2. Section. 3 focuses on our proposed four level matching method and followed by an illustrative example. Experimental results and discussions are presented in Section. 4. Finally, we concluded our work in Section. 5.

2. RELATED WORK

Service selection has been a very important issue for service composition for years. This is because the improper selection of a service can affect the overall QoS of a composite service and this leads to user dissatisfaction. Researchers have adopted different approaches to select the best service from possible similar services.

In Zeng et al. (2004) they implemented a QoS based service selection through QoS ratings from the service requestor without including the context. Arasi et al. (2016) constructed a discriminant analysis model using a successability percentage of services. In Batra and Bawa (2011) they proposed a method to categorize the services into a set of already defined categories using principal component analysis. In Cardoso (2006) they computed the similarity between the service request and services published based on absolute distance between them. In Skoutas et al. (2007), the QoS requirement was detailed and structured into different classes, like security-related QoS, price-related QoS and run time-related QoS. In Papaioannou et al. (2006) they described QoS parameters which are based on QoS ontology models. In Tsesmetzis et al. (2006), they proposed a novel three-dimensional QoS model for web service discovery with a guaranteed QoS and distribution mechanism.

In Rajendran and Balasubramanies (2009), the services are selected by an agent-based web service discovery framework, which satisfies client preferences. Kalepu et al. (2004) proposed a QoS based web service selection and ranking by evaluating the reputation of a service in which it blindly considers the QoS values produced by users. Tian et al. (2004) presented a broker-based web service selection model that enables service selection based on QoS constraints. In D’Mello and Ananthanarayana (2008), also Aruna and Aramudhan, (2016) they explained the
web service selection mechanism, which ranks the candidate services based on prospective levels of satisfaction of requests.

Shao-chang Li et al. (2010) proposed a heuristic algorithm for selecting services based on QoS parameters, concerning the degree of user satisfaction. Almulla et al. (2011) and Bhushan and Pradeep (2016) presented a ranking mechanism by using a fuzzy constraint satisfaction problem, in which QoS criteria are considered. Ardagna and Pernici, (2007) proposed a service composition method based on a linear programming model, but it is difficult to solve the complex problem.

In Benatallah et al. (2002), it is based on request parameters that the service selection is performed by considering the past and current execution history. Another set of authors Lim et al. (2011) and Vergin Raja Sarobin et al., (2016a or 2016b) proposed a web service selection within a community by calculating scores for other slave web services based on QoS parameters by a master web service.

3. PROPOSED FOUR-LEVEL MATCHING MODEL FOR SERVICE SELECTION

In this paper, we propose a four-level matching model for service selection based on QoS parameters, which includes description matching of a service, matchmaking process, LDA-based QoS matching and ranking method. The flowchart of the proposed method is shown in Figure 1.

![Figure 1 Flow chart of the proposed model](image-url)
3.1. Description Matching of a Service

To find all similar services from multiple cloud domains, we employed term frequency (TF) - inverse document frequency (IDF) and cosine similarity to compute similarity scores of all the services with respect to the service request. Finally, the service set is filtered by applying thresholds.

The steps that are involved in description matching of a service is detailed below:

3.1.1. Term Frequency (TF)

TF measure will list down all the services based on service request. The web service description language (WSDL) will be a different size for each service, so we can adopt normalization to normalize the service description size.

3.1.2. Inverse Document Frequency (IDF)

In order to find similar services that are matched with a service request is our primary concern, but in Step 1 the terms involved in service request matching have equal importance. But in reality some terms have a minute capacity to decide relevant services and other terms will find more relevant services.

So, we need to weigh up and weigh down the terms for retrieving more relevant services from multiple clouds as shown in Equation 1:

\[
IDF_{ei}(term_i) = 1 + \log_2 \left( \frac{\text{total number of services in a cloud}}{\text{number of services with service requester term in it}} \right)
\]  

Aggregated the term count for all clouds by using Equation 2 shown below:

\[
term_i = IDF_{c1} + IDF_{c2} + \ldots + IDF_{cn}
\]  

Calculate IDF for each term in the service request from all clouds as shown in Equation 3:

\[
\begin{align*}
Term_1 &= IDF_{c1}(term_1) = 1 + \log_2(1) \\
Term_2 &= IDF_{c1}(term_2) = 1 + \log_2(1) \\
&\vdots \\
Term_n &= IDF_{c1}(term_n) = 1 + \log_2(1)
\end{align*}
\]  

3.1.3. Perform TF *IDF

Multiply the normalized term frequency with its inverse service frequency on each service in every cloud.

3.1.4. Vector Space Model – Cosine Similarity

A vector is derived for each service, and the set of services from multiple clouds is viewed as a set of vectors in a vector space.

To find the similarity between any two services is given by the Equations 4 and 5:

\[
\text{Cosine similarity } (S1,S2) = \text{Dot product } \left( \frac{(S1,S2)}{||S1|| \ast ||S2||} \right)
\]  

\[
\text{Dot Product } (S1,S2) = s_1[0] \ast s_2[0] + s_1[1] \ast s_2[1] + \cdots + s_1[n] \ast s_2[n]
\]  

\[
||S1|| = \sqrt{s_1[0]^2 + s_1[1]^2 + \cdots + s_1[n]^2}
\]  

\[
||S2|| = \sqrt{s_2[0]^2 + s_2[1]^2 + \cdots + s_2[n]^2}
\]
where $d_1,d_2$ are service description documents. Based on the similarity score of all relevant services from multiple clouds, these will be reduced by applying a threshold filter.

### 3.2. Matchmaking Phase

The matchmaking phase selects the candidate cloud service providers based on the service request, which will be input for LDA based QoS matching. The cloud service provider must satisfy the following constraint.

#### 3.2.1. Dimensions matching

The cloud service must have all QoS values which are in the service request. The cloud service QoS dimensions must be a superset of the service requestor QoS dimension. Equation 6 given below finds out whether the service meets the service requestor dimensions.

$$
M(sp, sr) = \frac{|\{s | s \in sp \land s \in sr\}|}{|\{s | s \in sr\}|}
$$

$$
M(sp, sr) = 1, \quad M(sp, sr) = 1
$$

$$
0, \quad M(sp, sr) < 1
$$

where $sp, sr$ is a service provider and a service requester.

### 3.3. Normalization

The values of various QoS parameters are different to performing numeric matching. So, the QoS parameters need to be quantified to have a uniform distribution. The QoS parameters are classified into positive and negative criteria based on their impact shown in the classification function.

The positive parameters are those with an increase in attribute value when the objective is increased and with such an increase in attribute value the objective function is going to decrease those which are negative criteria.

The positive parameters and negative parameters are normalized by Equations 7 amd 8 below:

$$
N_{ij}^+ = \frac{a_{ij} - q_{ij}^{\min}}{q_{ij}^{\max} - q_{ij}^{\min}} \quad \text{if } q_{ij}^{\max} - q_{ij}^{\min} \neq 0
$$

$$
N_{ij}^+ = 1 \quad \text{if } q_{ij}^{\max} - q_{ij}^{\min} = 0
$$

where $N_{ij}$ is the normalized values of $j^{th}$ parameter of the $i^{th}$ service.

$$
N_{ij}^- = \frac{q_{ij}^{\min} - q_{ij}^{\min}}{q_{ij}^{\max} - q_{ij}^{\min}} \quad \text{if } q_{ij}^{\max} - q_{ij}^{\min} \neq 0
$$

$$
N_{ij}^- = 1 \quad \text{if } q_{ij}^{\max} - q_{ij}^{\min} = 0
$$

where $N_{ij}$ is the maximum value of the $j^{th}$ column of the QoS matrix, $q_{ij}^{\max} = \max(q_{ij}), 1 \leq i \leq n$ and $N_{ij}$ is the minimum value of the $j^{th}$ column of the QoS matrix, $q_{ij}^{\min} = \min(q_{ij}), 1 \leq i \leq n$. $q_{ij}^{\min}$ denotes minimum QoS value and $q_{ij}^{\max}$ represents the maximum QoS value. After normalization, all the QoS values lie between a [0,1] interval.

### 3.4. Classification Function

The classification function can be used to determine to which class each service most likely belongs. The classification function allows us to compute a classification score for each service in each class, by applying Equation 9.
\[ S_i = c_i + w_{i1} \cdot x_1 + w_{i2} \cdot x_2 + \cdots + w_{im} \cdot x_m \]  \hspace{1cm} (9)

In this formula, the subscript \( i \) denotes the respective class, the subscripts \( 1, 2, \cdots, m \) denotes the \( m \) parameter, \( c_i \) is constant for the \( i^{th} \) class, \( w_{ij} \) is the weight for the \( j^{th} \) parameter in the computation of the classification score for the \( i^{th} \) class. \( x_j \) is the observed value for the respective service for the \( j^{th} \) parameter. \( S_i \) is the resultant classification score.

The services are classified into groups, which are represented as matrices consisting of similar services in a row and their QoS criteria are represented in columns as shown below in Equation 10.

\[
S = \begin{bmatrix}
S_{11} & S_{12} & \cdots & S_{1m} \\
S_{21} & S_{22} & \cdots & S_{2m} \\
\vdots & \vdots & \ddots & \vdots \\
S_{n1} & S_{n2} & \cdots & S_{nm}
\end{bmatrix}
\]  \hspace{1cm} (10)

The 4L-LDA-SC compute the service score is based on service requester weights as shown in the below column vector indicated in Equation 11:

\[
W = \begin{bmatrix}
w_1 \\
w_2 \\
\vdots \\
w_m
\end{bmatrix}
\]  \hspace{1cm} (11)

where \( w_1, w_2, \ldots, w_m \) are preferred weights of each QoS parameter.

### 3.5. Linear Discriminant Analysis

Linear discriminant analysis is used to classify the services into groups based on the service request. The proposed method classifies the services into two groups, which is a 2-class problem. The goal is to identify the best class with a good between class separability and to find the best service in the selected class with a good within class separability. The steps for performing linear discriminant analysis for service selection is illustrated below.

#### 3.5.1. Compute Mean

Compute Mean is calculated for two classes from the dataset. For two classes, compute the mean by using below Equation 12.

\[
\mu_1 = \frac{1}{N_1} \sum_{x \in c_1} x \\
\mu_2 = \frac{1}{N_2} \sum_{x \in c_2} x
\]  \hspace{1cm} (12)

where \( N_1, N_2 \) is the total number of services in class \( c_1 \) and \( c_2 \).

#### 3.5.2. Compute Covariance Matrix

For two classes, compute the covariance matrix by using Equation 13 below:

\[
S_1 = \sum_{x \in c_1} (x - \mu_1)(x - \mu_1)^T \\
S_2 = \sum_{x \in c_2} (x - \mu_2)(x - \mu_2)^T
\]  \hspace{1cm} (13)

where \( \mu_1, \mu_2 \) is the mean of two classes.
3.5.3. Compute the scatter matrices
Now we compute within-class and the between class scatter matrix, which helps in finding the best class and service with a good between class separability and within class separability.

3.5.3.1. Within-class scatter matrix \( (S_w) \)
The within-class scatter matrix \( S_w \) is computed by Equation 14 as follows:
\[
S_w = S_1 + S_2
\]  
(14)
where \( S_1, S_2 \) is the covariance matrix of two classes.

3.5.3.2. Between-class scatter matrix \( (S_B) \)
The between-class scatter matrix \( S_B \) is computed by Equation 15:
\[
S_B = (\mu_1 - \mu_2)(\mu_1 - \mu_2)^T
\]  
(15)
where \( \mu_1, \mu_2 \) is the mean of two classes.

3.5.4. Compute the eigenvectors and eigenvalues
Compute the eigenvectors and corresponding eigenvalues for the scatter matrices by using Equation 16:
\[
S_e^{-1}S_B = \lambda_e
\]  
(16)
The eigenvector corresponding to the smallest eigenvalue leads to bad separability between the two classes and the eigenvector that corresponds to the highest eigenvalue leads to good separability between the two classes. To have a better discriminant class scatter matrix value, it should be large and within a class scatter matrix, it should be small.

3.5.5. Compute score
The Mahalanobis Distance (P.C. Mahalanobis, 1936) is used to find the best service from the selected class. The service with the highest score will be selected as the best service. Equation 17 below is used to calculate the Mahalanobis Distance.
\[
d_m(x, y) = \sqrt{(x - m)^T S^{-1} (x - m)}
\]  
(17)
where \( x \) is the vector of service data, \( m \) is the vector of mean values, \( S^{-1} \) is the inverse covariance matrix, \( T \) is the indicates a vector should be transposed.

4. RESULTS AND DISCUSSION
In this paper, an LDA-based four-level matching model is proposed to rank similar services from multiple clouds based on QoS preferences. All experiments were implemented in MATLAB using an HP Pavilion dv6 laptop with 2.10 GHz Intel core processor and 2 GB RAM. The description and dimension matching is performed in the first and second stage as shown in Figure 1. Table 1 and Table 2 show the normalized QoS information of Class 1 and Class 2.

To compute between class scatter matrix and within class scatter matrix, we use the Equations 14 and 15, the resultant values are shown in Table 3 and Table 4. The \( S_B \) variable helps in identifying the best class and \( S_w \) variable finds the best service in the selected values with good within class separability.
A Four-level Linear Discriminant Analysis Based Service Selection in the Cloud Environment

Table 1 The normalized QoS information about Class 1 services after the description and matchmaking process

<table>
<thead>
<tr>
<th>Services</th>
<th>Response Time</th>
<th>Throughput</th>
<th>Availability</th>
<th>Successability</th>
<th>Price</th>
</tr>
</thead>
<tbody>
<tr>
<td>S2</td>
<td>0.1112</td>
<td>0.2816</td>
<td>0.5554</td>
<td>0.2008</td>
<td>0.3915</td>
</tr>
<tr>
<td>S3</td>
<td>0.1772</td>
<td>0.3392</td>
<td>0.2981</td>
<td>0.3534</td>
<td>0.3437</td>
</tr>
<tr>
<td>S6</td>
<td>0.2890</td>
<td>0.3392</td>
<td>0.4778</td>
<td>0.3333</td>
<td>0.3537</td>
</tr>
<tr>
<td>S8</td>
<td>0.1729</td>
<td>0.3392</td>
<td>0.1103</td>
<td>0.3333</td>
<td>0.3187</td>
</tr>
<tr>
<td>S9</td>
<td>0.2599</td>
<td>0.3392</td>
<td>0.3716</td>
<td>0.3212</td>
<td>0.3706</td>
</tr>
<tr>
<td>S11</td>
<td>0.3717</td>
<td>0.3392</td>
<td>0.1511</td>
<td>0.3855</td>
<td>0.3300</td>
</tr>
<tr>
<td>S14</td>
<td>0.3412</td>
<td>0.3392</td>
<td>0.1184</td>
<td>0.3614</td>
<td>0.3228</td>
</tr>
<tr>
<td>S16</td>
<td>0.5879</td>
<td>0.3392</td>
<td>0.3736</td>
<td>0.3333</td>
<td>0.2846</td>
</tr>
<tr>
<td>S18</td>
<td>0.4183</td>
<td>0.3392</td>
<td>0.2185</td>
<td>0.3453</td>
<td>0.2657</td>
</tr>
</tbody>
</table>

Table 2 The normalized QoS information about Class 2 services after the description and matchmaking process

<table>
<thead>
<tr>
<th>Services</th>
<th>Response Time</th>
<th>Throughput</th>
<th>Availability</th>
<th>Successability</th>
<th>Price</th>
</tr>
</thead>
<tbody>
<tr>
<td>S1</td>
<td>0.3311</td>
<td>0.3768</td>
<td>0.1166</td>
<td>0.2615</td>
<td>0.3833</td>
</tr>
<tr>
<td>S4</td>
<td>0.2902</td>
<td>0.3312</td>
<td>0.1747</td>
<td>0.2660</td>
<td>0.3450</td>
</tr>
<tr>
<td>S5</td>
<td>0.3720</td>
<td>0.3602</td>
<td>0.3011</td>
<td>0.3534</td>
<td>0.3437</td>
</tr>
<tr>
<td>S7</td>
<td>0.3720</td>
<td>0.3105</td>
<td>0.1709</td>
<td>0.3534</td>
<td>0.3437</td>
</tr>
<tr>
<td>S10</td>
<td>0.2902</td>
<td>0.3478</td>
<td>0.2837</td>
<td>0.4102</td>
<td>0.3450</td>
</tr>
<tr>
<td>S12</td>
<td>0.3720</td>
<td>0.3395</td>
<td>0.2722</td>
<td>0.4102</td>
<td>0.3450</td>
</tr>
<tr>
<td>S13</td>
<td>0.2902</td>
<td>0.2857</td>
<td>0.3687</td>
<td>0.4102</td>
<td>0.3450</td>
</tr>
<tr>
<td>S15</td>
<td>0.3720</td>
<td>0.3602</td>
<td>0.6455</td>
<td>0.4102</td>
<td>0.3450</td>
</tr>
<tr>
<td>S17</td>
<td>0.2902</td>
<td>0.2732</td>
<td>0.3589</td>
<td>0.4102</td>
<td>0.3450</td>
</tr>
</tbody>
</table>

Table 3 within Service Class $S_w$

<table>
<thead>
<tr>
<th>QoS</th>
<th>Response Time</th>
<th>Throughput</th>
<th>Availability</th>
<th>Successability</th>
<th>Price</th>
</tr>
</thead>
<tbody>
<tr>
<td>Response Time</td>
<td>10.6751</td>
<td>-17.6040</td>
<td>-0.5253</td>
<td>-1.8683</td>
<td>19.757</td>
</tr>
<tr>
<td>Throughput</td>
<td>-17.6040</td>
<td>133.3829</td>
<td>9.7370</td>
<td>21.5365</td>
<td>-33.834</td>
</tr>
<tr>
<td>Availability</td>
<td>-0.5253</td>
<td>9.7370</td>
<td>3.9565</td>
<td>4.5731</td>
<td>-0.3719</td>
</tr>
<tr>
<td>Successability</td>
<td>-1.8683</td>
<td>21.5365</td>
<td>4.5731</td>
<td>16.1673</td>
<td>-0.2224</td>
</tr>
<tr>
<td>Price</td>
<td>19.7579</td>
<td>33.8346</td>
<td>-0.3719</td>
<td>-0.2224</td>
<td>88.882</td>
</tr>
</tbody>
</table>

Table 4 Service between Class $S_b$

<table>
<thead>
<tr>
<th>QoS</th>
<th>Response Time</th>
<th>Throughput</th>
<th>Availability</th>
<th>Successability</th>
<th>Price</th>
</tr>
</thead>
<tbody>
<tr>
<td>Response Time</td>
<td>0.0443</td>
<td>-0.0019</td>
<td>0.0031</td>
<td>-0.0161</td>
<td>0.0015</td>
</tr>
<tr>
<td>Throughput</td>
<td>-0.1098</td>
<td>0.0046</td>
<td>-0.0077</td>
<td>0.0400</td>
<td>-0.0037</td>
</tr>
<tr>
<td>Availability</td>
<td>-0.0081</td>
<td>0.0003</td>
<td>-0.0006</td>
<td>0.0030</td>
<td>-0.0003</td>
</tr>
<tr>
<td>Successability</td>
<td>-0.0291</td>
<td>0.0012</td>
<td>-0.0020</td>
<td>0.0106</td>
<td>-0.0010</td>
</tr>
<tr>
<td>Price</td>
<td>0.0844</td>
<td>-0.0036</td>
<td>0.0059</td>
<td>-0.0307</td>
<td>0.0029</td>
</tr>
</tbody>
</table>

The optimum class selection is shown in Table 5, where $C_2$ is the best class with a good score. From Table 6, S15 is the best service with 0.4740 aggregate score and S9 is the worst service with the least score.
Table 5 optimum class selection

<table>
<thead>
<tr>
<th>Class</th>
<th>Score</th>
<th>Rank</th>
</tr>
</thead>
<tbody>
<tr>
<td>C1</td>
<td>0.29816</td>
<td>2</td>
</tr>
<tr>
<td>C2</td>
<td>0.64124</td>
<td>1</td>
</tr>
</tbody>
</table>

Table 6 Overall score and ranks of services in C2

<table>
<thead>
<tr>
<th>Sorted Services</th>
<th>Score</th>
<th>Rank</th>
</tr>
</thead>
<tbody>
<tr>
<td>S15</td>
<td>0.4740</td>
<td>1</td>
</tr>
<tr>
<td>S5</td>
<td>0.3582</td>
<td>2</td>
</tr>
<tr>
<td>S10</td>
<td>0.3449</td>
<td>3</td>
</tr>
<tr>
<td>S1</td>
<td>0.3316</td>
<td>4</td>
</tr>
<tr>
<td>S4</td>
<td>0.2986</td>
<td>5</td>
</tr>
<tr>
<td>S12</td>
<td>0.2978</td>
<td>6</td>
</tr>
<tr>
<td>S13</td>
<td>0.2943</td>
<td>7</td>
</tr>
<tr>
<td>S17</td>
<td>0.2819</td>
<td>8</td>
</tr>
<tr>
<td>S7</td>
<td>0.2724</td>
<td>9</td>
</tr>
</tbody>
</table>

In Figure 2, we compared our proposed Four-Level-Linear Discriminant Analysis Based Service Selection (4L-LDA-SC) with QoS Based-Web Services Selection Method (WSSM-Q), QoS Aware Service Selection Based on Clustering (SCB-QC) methods with respect to the number of candidate services required for execution of tasks. The existing WSSM-Q method selects services based on QoS constraints by evaluating all the candidate services without filtering, which shows an impact on computation time and optimality. The SCD-QC adopted a clustering technique to reduce the number of candidate services required for the evaluation process, but still it fails to attain an optimum number of candidate services. The 4L-LDA-SC dominates the other two methods with significant reduction in the number of candidate services required.
Figure 3 shows the importance of candidate service filtering at each level with respect to the variation of tasks. The candidate services at Level 3 will be considered for evaluation to find the best service.

In Figure 4, the proposed 4L-LDA-SC attains good computation time with respect to the variation of tasks from 20 to 100. The reduction in the number of candidate services has shown influence on computation time. The proposed method filters the candidate services by applying cosine similarity and dimension matching.

The quality of the proposed method is evaluated by comparing the overall classification function value of the selected services with an overall classification function value of the optimal selection obtained by the proposed method. Figure 5 shows the optimality value of different methods with respect to the variation of the number of candidate per class. The proposed method achieves the best result with more than 96% optimality ratio on average.

5. CONCLUSION

This paper proposes a four-level linear discriminant analysis based service selection (4L-LDA-SC) method to select a better service. The description and dimension matching of service filters the candidate services, which is input to LDA-based QoS matching. The candidate services are divided into two classes from which best class and service is selected. The proposed method achieves a reduced number of candidate services for evaluation and significant reduction in computation time over differing tasks. The 4L-LDA-SC method finds best service with more than 96% optimality ratio on average. In future work, we will use a neural network for classification and prediction of services based on user preferences.

6. REFERENCES


