Multiple Linear Regression Approach for Determining Quality of Web Service in Mobile Environment

Khaing Lae Lae Soe, and Thiri Haymar Kyaw

Abstract—Today, almost all services are available on the internet in the present day and we can just utilize these services by our portable devices anytime and anywhere. Consequently mobile devices are boosting in use and mobile applications are considered to be designed highly compatible for these devices. Network bandwidth, available memory, computation speed and operating system are significant factors for running mobile application efficiently. Therefore these factors are used as interesting contexts and then multiple linear regression is exploited to estimate quality of service returned from server. By using multiple linear regression, mobile application can get convenient service even if user does not know about the characteristics of his device and the application he uses. We use this model in CARM (Context Adaptable Reflective Middleware over Mobile environment). We demonstrate a service that can display townships in Yangon Division is developed using Apache Axis2. Map service can show three versions; full-color map, gray-color map and text only information. The appropriate version is predicted using multiple linear regression. The system can keep the balance of resource utilization in mobile devices and appropriate service for users. Dynamic change in mobile environment is solved using java dynamic proxy.

Keywords—Context, dynamic proxy, mobile environment, multiple linear regression, webservice

I. INTRODUCTION

The increasing use of mobile devices, such as mobile phones and personal digital assistants, has lead to many efforts of designers to choose flexible and convenient methods or techniques. These devices have scarce resources and exposed to variation of context. In order to provide reasonable computing environment, which approaches the best that currently available resources of contexts will allow, webservice should be context-awareness to be adapt to limited or fluctuating resources.

Context is the set of environmental states and settings that either determines an application’s behavior or in which an application event occurs and is interesting to the user [7]. They are significant factors for running mobile application efficiently. The contexts that can mostly affect mobile infrastructure are location, network, power, and storage and computation capacity.

Web Services are a self-describing, self-contained, modular application accessible over the web. It can provide heterogeneous middleware for mobile devices. What is more, it can increase interoperability, reusability, and composability of software systems. SOAP is lightweight, XML-based heterogeneous protocol that is easily configurable and extensible. Although SOAP is said to be bandwidth-consuming because of text-based protocol, other choices of message passing style such as JSON objects are available for mobile devices. Therefore, the marriage of webservces and mobile devices becomes popular technology combination.

This paper is concerned service selection for mobile computing via multiple linear regression. It is simple equation that is compatible for mobile devices. Although it causes overhead in training stage, it can provide faster estimation after achieving necessary relationships. The remainder of the paper is structured as follow: Section II describes related work. Section III is about multiple linear regression. Section IV demonstrates case-study, Map service. In Section V, we will discuss how multiple linear regression is used in forecasting suitable service. Section VI will finish this paper and discuss strengths and weakness.

II. RELATED WORK

Research area has been developed for mobile environment after boosting mobile devices. Reflective middleware is one option for dynamic adaptation for context changes that often occurs in that environment. In this section, we describe previous works of reflective middleware and some attempts for service adaptation.

The CARISMA project [2] is a context-aware middleware system for service adaptation in mobile situation. In this system, XML is used as “Application Profile” that encodes context-aware application-specific semantic information. The middleware proposes how context changes should be handled using policies. The service policy confliction is resolved by using a microeconomic approach that relies on a particular type of sealed-bid auction. The Chisel system [3] introduced a dynamic services adaptation framework which decomposes the
particular aspects of a service object into multiple possible behaviors. Whenever the context information changes, the service object will be adapted to use the different behaviors according to the adaptation policy.

MUSIC is a component-based planning framework that enables self-adaptation of mobile and ubiquitous applications in the presence of Service-Oriented Architectures (SOA). Although it is implemented as webservice approach to adaptation process, it focuses on changes in the service provider landscape in order to plug in interchangeably components and services providing the functionalities defined by the component framework [5]. It does not consider changes in client or consumer landscape.

All previous reflective middlewares acquire user’s preferences to achieve desired quality of applications. No reflective middleware solution exists that makes automatically adaptation of application according to available resources. However, FSAM, context-aware mobile computing middleware, formulate the service adaptation process by using fuzzy linguistic variables and membership degrees to define the context situations and the rules for adopting the policies of implementing a service [6].

On the other hand, logistic regression has been already used for grid computing [1] and we wishes to use this method because it can handle continuous input streams very well and current prediction is done by local functions which are using only a subset of the data. We use locally weighted learning to find out the mapping between input context and output quality. After that multiple linear regression is employed to predict suitable application version based on current contexts of client’s mobile device.

### III. MULTIPLE LINEAR REGRESSION

Multiple regression is an extension of simple linear regression. It is used when we want to predict the value of a variable based on the value of two or more other variables. The most valuable (and correct) use of regression is in making predictions. However, one point to keep in mind with regression analysis is that causal relationships among the variables cannot be determined. It means we can say that X "predicts" Y, we cannot say that X "causes" Y [8].

Multiple linear equation is as follow:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_v x_v + \epsilon$$  \hspace{1cm} (1)

where

- $y$ = an observed value of the response variable for a particular observation in the population
- $\beta_0$ = the constant term
- $\beta_j$ = the coefficient for the $j$th explanatory variable ($j = 1, 2, \ldots, v$)
- $x_j$ = a value of the $j$th explanatory variable for a particular observation ($j = 1, 2, \ldots, v$)
- $\epsilon$ = the residual for the particular observation in the population

### IV. CASE STUDY: MAP SERVICE

CARMM architecture allows the best possible quality of application based on available contexts. Firstly, the client must acquire the device’s contexts such as network bandwidth, processor usage, and so on. In context representation stage, acquired contexts are symbolized into specified format that agrees with the service provider. This agreement is written in XML file as shown in Fig. 1 and it is used in server-side to write configuration or policies of application.

On the server side, particular conditions of application like poor or good state of device are calculated after receiving context information. And then the decision making process takes place by seeing predefined configuration file. Subsequently, the output that emits from decision making is used to load related service class dynamically. Finally, the loaded class provides service to the users.

In case study application, the service that shows townships of Yangon Division in Myanmar is provided in the server side. We use Apache Axis2 as service provider and services are implemented using Java language. Mobile clients are developed for android platform. In map service, there are three versions to display map to the users: - full-colour map, gray-colour map and description texts only without photo.

Firstly, user can choose desired town in the list view as shown in figure 2(a). After selecting one of the towns, the client program requests map service from the server altogether with its current device’s state. The server responds the version convenient with the current context of the user’s mobile phone. The server makes decision which policy should be adopted by matching predefined rules and acquired contexts’ calculation.

In figure 2(b), the server give the result of good quality map in the time of client’s resources maximize. However, in figure 2(c), the facility of application is not much great because the client’s contexts degrade. Figure 2(d) is the poorest result returned from server because device exhausts one or more resources of device. The contexts to be considered are network speed, operating system (API level), processor speed and available memory. The users can view maps of 32 townships in Yangon Division of Myanmar. The returned Map can be viewed to zoom-level 3. Rules of which version should be available for which contexts should be negotiated between service provider and client.

```
<SERVICE Name="Version_1">
<CONTEXT id="good" policy="value<30000" />
<CONTEXT id="version_2">
<CONTEXT id="moderate" policy="30000<value<50000" />
<CONTEXT id="version_3">
<CONTEXT id="poor" value="value>50000" />
```

Fig. 1 Example XML Configuration File
V. LEARN RULES WITH MULTIPLE LINEAR REGRESSION

In Yangon Map Application, there are three services that can be provided from server. Therefore, three policies must be generated to provide suitable service.

If \( y \) (calculated result of contexts) is in the range of predefined policy 1, version 1 \((v_1)\) is sent to user. Let’s as say:

- \( p_1 \)= predefined policy of service 1
- \( p_2 \)= predefined policy of service 2
- \( p_3 \)= predefined policy of service 3

The procedure would be:

- If \((\min(p_1) < y < \max(p_1))\) \(v_1\) is sent to client’s device.
- Else If \((\min(p_2) < y < \max(p_2))\) \(v_2\) is sent to client’s device.
- Else \(v_3\) is sent to client’s device.

Rules of which version should be available for which contexts should be negotiated between service provider and client. In this place, rules are generated using multiple linear regression [4]. The contexts to be considered are network speed, operating system version, processor speed and available memory.

Multiple linear regression is used for its simplicity and competitive accuracy. It is exploited for estimating time to download map images. In this place, we use

\[
Y = \beta_0 + \beta_1 \cdot \text{bandwidth} + \beta_2 \cdot \text{os} + \beta_3 \cdot \text{cpuspeed} + \beta_4 \cdot \text{memory}
\]

(1)

A. Training Data Set

Android emulator API 8 to 18 is used to provide operating system parameters. In each emulator, various bandwidth (measured in KB/s) and available memory (measured in MB) is inputted to test the model. Depending on device model, processor speed may be varied. This is calculated using the idea of BogoMips in Linux kernel. BogoMips is unscientific measurement of CPU speed when it boots to calibrate an internal busy-loop. An oft-quoted definition of the term is "the number of million times per second a processor can do absolutely nothing". However, in this place, it is calculated when processor is in current state rather than the time of loading operating system. Dataset is shown in Fig. 3.

B. Using IBM SPSS Statistics

We use IBM SPSS Statistics to make calculation for multiple linear regression. Figure 4, 5, 6, 7 and 8 are screenshots of IBM SPSS output.

Dataset Summary can be seen below:

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>Minimum</th>
<th>Maximum</th>
<th>Mean</th>
<th>Std. Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>os</td>
<td>134</td>
<td>26</td>
<td>440</td>
<td>172.19</td>
<td>130.597</td>
</tr>
<tr>
<td>memory</td>
<td>134</td>
<td>10</td>
<td>205</td>
<td>62.2</td>
<td>47.141</td>
</tr>
<tr>
<td>bandwidth</td>
<td>134</td>
<td>233</td>
<td>1639</td>
<td>8591.81</td>
<td>4149.665</td>
</tr>
<tr>
<td>time</td>
<td>134</td>
<td>6847</td>
<td>66851</td>
<td>24056.37</td>
<td>14614.383</td>
</tr>
</tbody>
</table>

Fig. 3 Dataset for Map Service

Fig. 4 Descriptive Statistics
After calculating with this software to get coefficient values in fig. 5 and then substituting in (1), the time for downloading color map can be estimated as follow:

\[
\text{Time} = 46587.311 - (155.59 \times \text{bandwidth}) - (2537.006 \times \text{os}) - (0.565 \times \text{cpu speed}) + (13.329 \times \text{memory})
\] (2)

Finally, policies are generated using XML configuration file in Fig. 1.

**C. Model Fitness**

R Square value is 0.441; therefore, 44.1% of the variation in %time can be explained by this model as shown in Fig. 6. Fig 7 presents that the model, as a whole, is significant to the data. Fig 8 displays residual statistics of the model.

**VI. CONCLUSION**

In this paper, we propose CARMM middleware architecture allows the best possible quality of application on available resource utilization. The users or developers can provide customized rules and services using this architecture. Moreover, they can add new services and rules easily.

Multiple linear regression is used to generate rules for service adaptation. Using this model, clients do not need to know his or her device’s characteristics and how many services can provide from service provider. What is more, multiple linear regression provides some other desirable properties, including simplicity, capability of extrapolating, and fast processing. However, learning stage must be added to adapt changing contexts automatically in the system.

**ACKNOWLEDGMENT**

I would like to give my special thanks to my supervisor, Dr. Thiri Haymar Kyaw, Lecturer of our university, UTYCC, (University of Technology, Yatanarpone Cyber City) for her kindheartedly aids. I would like to express my deepest thanks to all my teachers and colleagues who lend a hand directly or indirectly during the laborious process of completing this work successfully. I also show appreciation to my parents for their mental supporting.

**REFERENCES**


