

SOAP Optimization using Multi Agents

4.1 Introduction

The previous chapter discussed about the genetic algorithms and multi agent technologies. It also highlighted the advantages of using GA as a hybrid method. Then it has also discussed about the possible models of using multi agent systems on data clustering process. It is always possible to do any work efficiently when we first organize the work very well. This is applicable to messaging also. If we first analyze the message well and then organize it properly then it is possible to do an efficient message transfer by utilizing resources optimally. This chapter describes the approach to optimize SOAP communication using multi agent negotiation methodology. This chapter will discuss about the users of the proposed system, inputs and output of the system and finally process and the features of the system. Further in detail it will discuss the approach taken to identify XML characteristics of the SOAP message, then approach of finding common repeating content patterns of the message and then how those patterns has been used to optimize the SOAP message.

4.2 Users

Any web-service application can be defined as users of this system. Ultimately the software developers who develop those web services will use this tool as a plug-in.

4.3 Input to the system

At the sender side of the web-service communication the web-service SOAP message that is to be transferred through the network is taken as the input to the system. But at the receiver side of the communication, the optimized message which has been transferred through the network is the input.

4.4 Output of the system

At the message sender side, system will output the optimized message which is the reduced version of the input SOAP message. Then this optimized message will be transferred through the network. At the receiver side of the message system will output the original version of the SOAP message.

4.5 Process

The process of SOAP communication optimization can be divided in to two parts. The identification of XML types and generation of basic templates based on these types will be discussed on SOAP message optimizing process. The second part of the process is further identification of static content inside the data section of the XML document and it will be discussed in the clustering process.

4.5.1 SOAP message optimizing process

If the SOAP message also organized properly then it is possible to increase the web-service performance drastically. Since the SOAP message is based on the XML, analyzing and organizing of the SOAP message should be done using the XML characteristics. Previous researches done on XML compression also confirm that it get better compression when the XML document is compressed using the XML characteristics. When it works for compression algorithms then it should work for message analyzing and organizing also. Therefore as the first step of optimization it should find the XML characteristics of the SOAP message. To find the XML characteristics it first identifies the basic elements that the XML document has been generated from. The basic components of the XML document can be identified by generating the schema of the XML document. After generating the schema document, each complex type on the schema document will be considered as the basic component of the document. Then it traverses through the XML document to find out the instance of those basic components and process them together. Because of the above approach all related data will be processed together and it will be more convenient to analyze the message by doing so.

Then it identify the common repeating content of the SOAP message as the static content of the document and classify rest of the content as the dynamic content. This

classification process has to be done carefully. The static content has to be identified so that the processing overhead at the time of organizing the message will be minimum as well as the storage overhead of those static content also will be minimum. It also must classify more content into static as possible.

After identifying the static content of the SOAP message it is possible to reorganize the SOAP message so that only the dynamic content will be transferred to the other party. The static content will be stored on both side of the communication and use as the pre-defined knowledge of the message. Only an identifier of the static pattern will be transferred with the message so that it will be possible to identify the correct static content pattern for the message at the receiver end.

4.5.2 Clustering process

It is the straightforward method to classify XML tags as the static content of the message. But this will not be enough because there can be repeating content or repeating patterns of content on values of XML nodes so that by deleting them the length of the message can be reduced drastically. It is possible to identify common content patterns of the text if common texts are clustered together. There are several ways of clustering data such as data mining techniques. But for those it is required to have a pre-knowledge of the data being clustered. And also it is required the user to select the variables that is to be used in clustering. But if it is used Multi Agent Negotiation method for clustering then it is possible to do the clustering automatically. It is never possible to predict the data content of a SOAP message on the web-service. Once the new data is entered to the system, those data will also be available on web-service messages. There can be upgrades to the web-service and therefore there can be changes to the SOAP message structure. Therefore the environment is dynamic.

Clustering of data to identify the common content patterns can be done in different ways. It is possible to create more number of clusters so that each cluster contains small number of members on it and therefore the common pattern of those members are more specific to those texts. That means larger part of the text is classified into static content. But there is more number of content templates generated on the system and therefore processing overhead and the storage overhead of the system is high. It is

also possible to cluster data into smaller number of clusters. In that way each cluster will contain large number of members in the cluster and therefore common text pattern generated from the cluster will be generic. Then less amount of content will be classified as the static content of the message and the large amount will be considered as dynamic content. By this way the data reduction of the SOAP message will be less and therefore the optimization process of SOAP message will not be done properly. There are many alternative solutions in between these extremes that can be used as the classification of static/dynamic content. The clustering mechanism must be capable of selecting a rational solution that will be capable of optimizing the SOAP message for its best.

When considering the above scenarios the approach that is suitable to find the optimal static/dynamic content patterns of the SOAP messages is to use Multi Agent System to cluster the text and then generate the static content pattern using those clusters. Two main types of Agent will be used in clustering. Those are cluster agent (the resource agent) and the text agent (the request agent). Cluster agents and text agent will negotiate to perform better clusters. For control the number of clusters at the start of the clustering process it will generate only few number of clusters (either two or three and the number is decided depending on the number of texts available for clustering). But for control the quality of the output of the cluster it will add some rules to the negotiation algorithms so that most suitable text for the current cluster pattern will be accepted in to the cluster as members. By this method it will identify the most suitable texts to be in the cluster and it will increase the quality of the output of the cluster output.

After identification of the static patterns of a particular SOAP message, during the communication it is possible to use this pattern as a pre-knowledge of the message. For that both parties engaged in the communication must have this knowledge. Therefore the identified static content pattern on one side of the communication is distributed to other side also. After the pattern is available on the other side also it is ready to use that knowledge in the communication. Because both parties know the same thing when a particular message is being transferred the redundant parts of the message (parts that both parties know) can be deleted from the message and then transfer. But there should be an indication of the message that such a part which was

originally on the message is now deleted. Therefore a unique key is generated per each static content pattern identified by the system and that key will be inserted into the message where static content is deleted. When such a message is received at the receiver end it will use that key and re-insert the corresponding static data into the message. By this process it will re-generate the original SOAP message.

4.6 Features

The system will automatically build the static content templates of the SOAP messages being transferred. It will also adapt to the changes of the SOAP message by improving the template itself. The system will automatically distribute the generated templates among the other connected nodes of the system. It will also guarantee of the consistence of the knowledge of all nodes of the system by allowing nodes to request the knowledge of new templates at the startup of the node. The template generation is done in optimal way so that it will generalized to cover more messages as possible and also it will contain more text as static content.

4.7 Summary



University of Moratuwa, Sri Lanka.
Electronic Theses & Dissertations

In this chapter it has discussed as to how the SOAP message communication can be optimized using the multi agent systems. It presented the approach with reference to users, input, output, process and the features of the system. The process has been described two major areas namely SOAP message optimization process and the clustering process. The multi agent negotiation methodology has been applied for both those areas in different context. The next chapter will discuss about the analysis and design phase of this project.