Fall 10-12-2012

Design and Implementation of Service-Oriented Expert System

Chi-Chun Lo

Mian-Yuan Yang

Dai-Rong Lee

Chi-Hua Chen

Ding-Yuan Cheng

Design and Implementation of Service-Oriented Expert System

Chi-Chun Lo, Mian-Yuan Yang, Dai-Rong Lee, Chi-Hua Chen*, Ding-Yuan Cheng
National Chiao Tung University, Taiwan, China
chihua0826.iim96g@nctu.edu.tw

Abstract: In recent years, the Internet technologies are well developed and the Internet is filled with all kinds of information. Since the data storage is increasingly distributed and data formats are more diverged, data collection and integration for providing value-added services have gradually become important topics. In this study, we propose the Service-Oriented Expert System (SOES) based on Service Component Architecture (SCA) which can make the services on different platforms turn into a common service component on the Internet, concatenate all the service components by combining with the Enterprise Service Bus (ESB), and use both expert rules and data mining techniques to perform the data classification. The SOES is applied to analyze the annual financial information derived from electronic industry in the Taiwan Economic Journal (TEJ) during 2006 to 2008 for discovering the financial crisis enterprises. The experiment results show that using expert rules and decision tree to find the financial crisis enterprise is higher performance.

Keywords: Service Component Architecture, Service-Oriented Expert System, Enterprise Service Bus, Data Mining, Decision Tree.

1. Introduction

Old companies and poorly-managed companies close down after financial turmoil. When emerging companies and competent companies attempt to acquire the poorly-managed companies or expand their own business, they usually confront the problem of integrating data from heterogeneous platforms. Government departments, enterprises, or academic institutions all need to deal with the data from heterogeneous platforms. Furthermore, different departments or different units within the same department may develop systems with heterogeneous languages, platforms, or databases at different times in different places. Due to the rapid increase of the systems on each independent platform, critical information of an organization is distributed between various platforms and databases which may use incompatible data schema and data formats. Besides, it is too risky to update the old systems, resulting in the so-called "islands of information" throughout the organization. However, more and more applications need to retrieve information from different information systems, making information sharing an increasing demand. Therefore, integrating the heterogeneous technologies (including services, systems, and databases) and data models of different systems without updating the old system massively has become an important issue.

This study aim to propose a flexible method that can quickly find people's problems. We intend to build a prediction system using Service Component Architecture (SCA) in order to integrate all systems on heterogeneous platforms and enable semi-automated inferring on the platform. Previous studies have shown that the integration of heterogeneous platform architecture can help to reduce the time cost, while the use of prediction system can save human resource of the specialists in the field. Therefore, in this paper we hope to integrate the framework, modifying and expanding existing functions to reduce time cost and expert labor cost more significantly. In

Proceedings of the Twelfth International Conference on Electronic Business, Xi'an, China, October 12-16, 2012, 36-41.

addition, past studies usually use statistical method or data mining techniques to implementing expert systems. However, the statistical assumptions are not fully realistic and it is possible to find many non-necessary rules if we only use data mining techniques, so we expect to discover a more effective rule by combining expert knowledge and data mining techniques. The goal of our study is to build web service components for the systems on different platforms, combine Enterprise Service Bus (ESB) to integrate all service components, and use both expert rules and data mining techniques to produce real-time searching results.

This paper is divided into five sections; the next section is literature review of financial prediction methods and systems. The third section interprets how to apply SCA to design a service-oriented expert system and what efforts we have done to improve the expert prediction model. The fourth section presents the implementation of service-oriented expert system, analyzes the executing speed among heterogeneous platforms, and compares the accuracy of this study with previous studies. The final section shows the conclusion and future research directions.

2. Literature Review

Financial crisis prediction has long become an important research topic in various fields, and the definition of failure varies among studies. Beaver was among the first to study bankruptcy prediction, he used financial statements data to predict corporate failure. Beaver defined bankruptcy, preferred stock dividends in arrears, and arrears of debt as the financial crisis (Beaver, 1966). Altman is the first scholar who uses financial ratios in bankruptcy prediction. Altman's definition of corporate failure is that the firm is bankrupt, taken-over, or reorganized according to law (Altman, 1968). Deakin defined failure as insolvency, bankruptcy, or liquidation of firms (Deakin, 1972). Booth thinks that a company fails when it is delisted from stock exchange (Booth, 1983). Laitinen identified three failure types: chronic, revenue financing, and acute (Laitinen, 1991).

Among the studies, researchers often construct the prediction model in three kinds of ways. The first way is using statistical methods such as univariate analysis, multivariate analysis, discriminant analysis, and logistic regression. Another way is to adopt data mining techniques, like neural networks, decision tree. The third way is to use both statistical methods and data mining techniques. In recent years, some scholars use intellectual capital and macroeconomic indexes as prediction variables, but financial ratios are still the most used variables.

Odom and Sharda construct their model using both neural networks and multivariate analysis (Odom and Sharda, 1990). Altman et al. use multivariate discriminant analysis and neural networks model respectively to analyze the bankruptcy risk of Italian companies, and discovered that the performance of neural networks model is better (Altman et al., 1994).

3. Service-Oriented Expert System Based on SCA

This study constructs a service-oriented expert system based on service component architecture (SCA), uses SCA to build common web service components for heterogeneous services, integrates the components using Enterprise Service Bus, and combines expert rules with data mining techniques to classify the data. We combine SCA with ESB to integrate heterogeneous service components, as shown in figure 1, components (1). We also improve the rule inference by using expert rules with data mining techniques, as shown in figure 1, components (2). The three parts of the system design principles are described below.

3.1 Service Components Architecture

This study employs SCA to make the heterogeneous services turns into common web components. The SCA Assembly Model consists of a series of artifacts, which can be assembled into any applications needed. In the following section, we will define and explain how to apply SCA Assembly Model to our research.

3.2 Enterprise Service Bus

The Enterprise Service Bus (ESB) uses XML, web service interfaces, standardized rules, and routing to concatenate the service components converted by SCA (as shown in figure 2). This study uses ESB to integrate and compose the metadata of heterogeneous service components, and utilizes the agent program developed under the concept of middleware to process data integration and searching.

The above diagram describes the framework of ESB. There are three services in the ServiceGroup, ServiceComposite A, ServiceComposite B, and ServiceComposite C, each of which has its corresponding XML language.

Heterogeneous platform integration requires the two important concepts SCA and ESB. When we input a lot of services into SCA, it will output WSDL for each service and a number of Composites of the services. Then we input the Composites into ESB and get Integrated Information as the final output (as shown in figure 3).

3.3 Inferring Rules

In addition to integrate heterogeneous platforms, we also improve the rule inference to enable semi-automated inference. First, we find expert rules among previous literatures, extract key factors from the rules, retrieve the data of key factors, classify the data using data mining techniques, and finally use JESS to verify the problem results.

Expert rule is the key factor in this study. We store expert rules in the knowledge base, pre-process a series of data conversion, use data mining method to identify the relation between rules, and classify the data according to the relations (as shown in Figure 5). This method can improve the accuracy of expert rule inference.

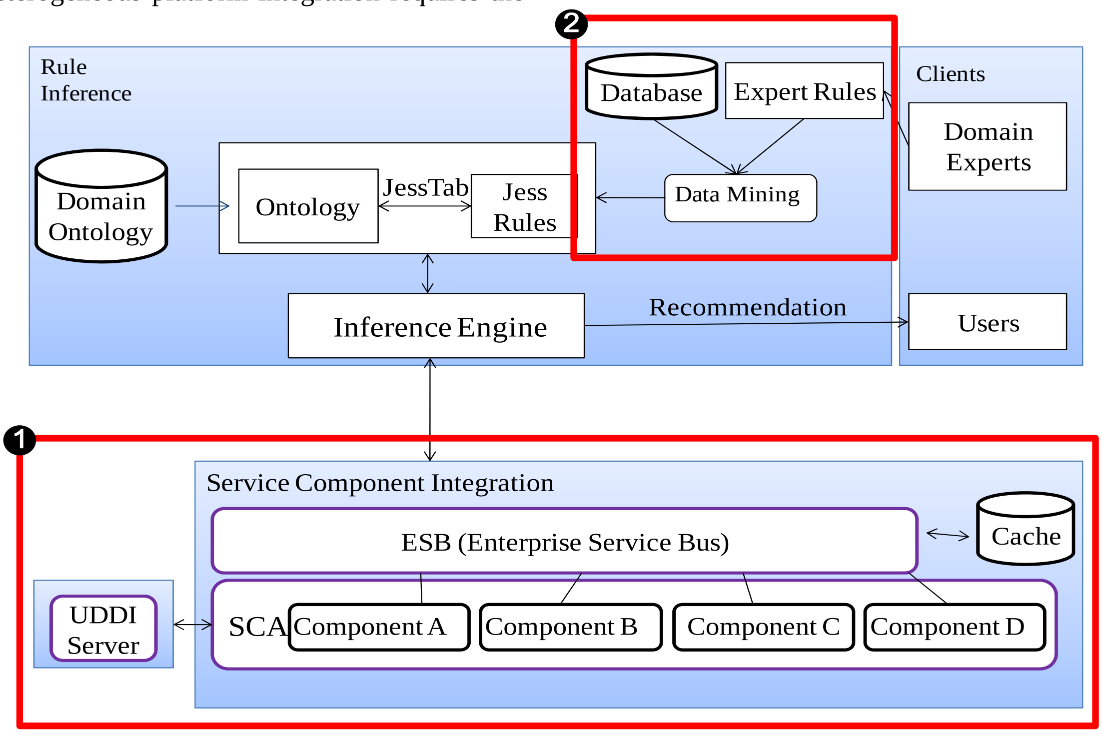

Fig. 1. System architecture

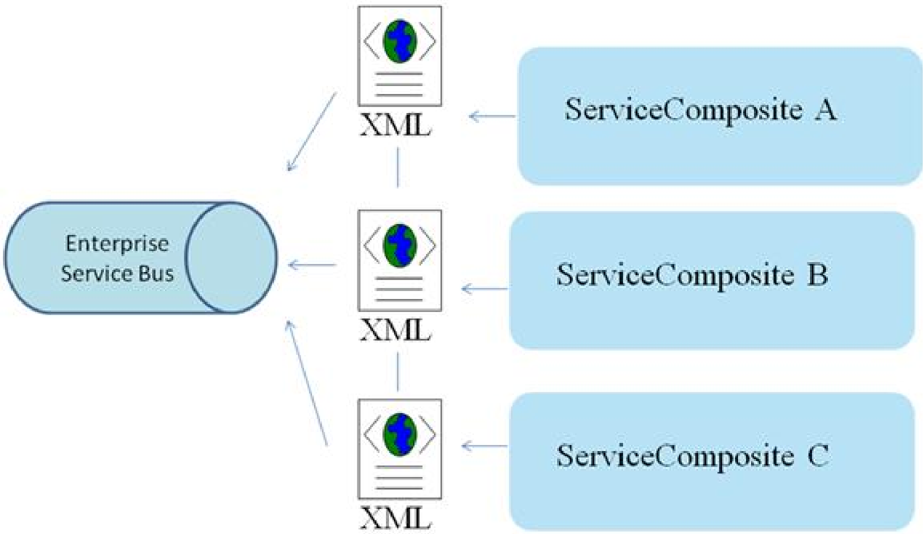

Fig. 2. Enterprise Service Bus

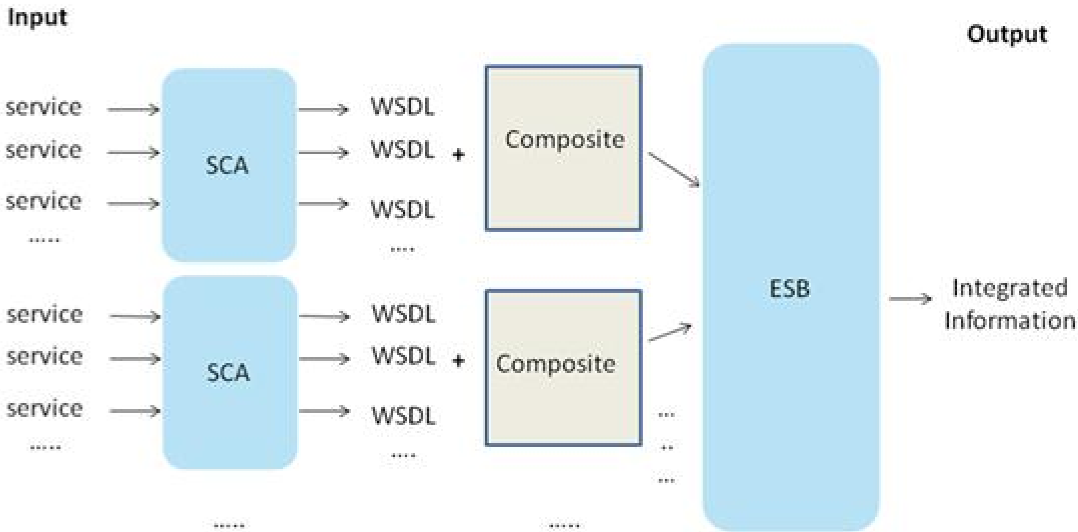

Fig. 3. SCA+ESB process

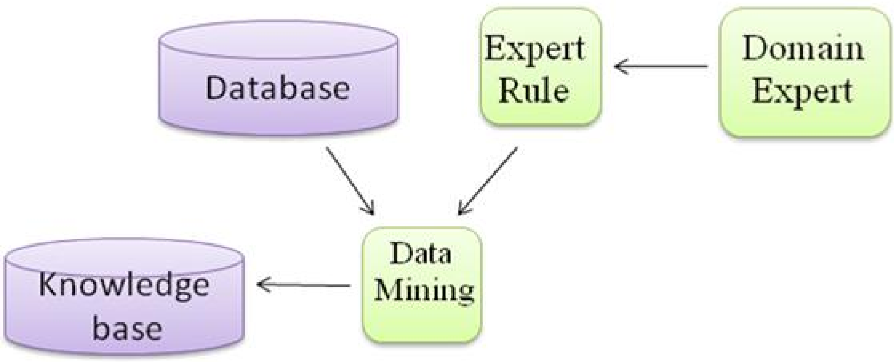

Fig. 4. Rule establishment

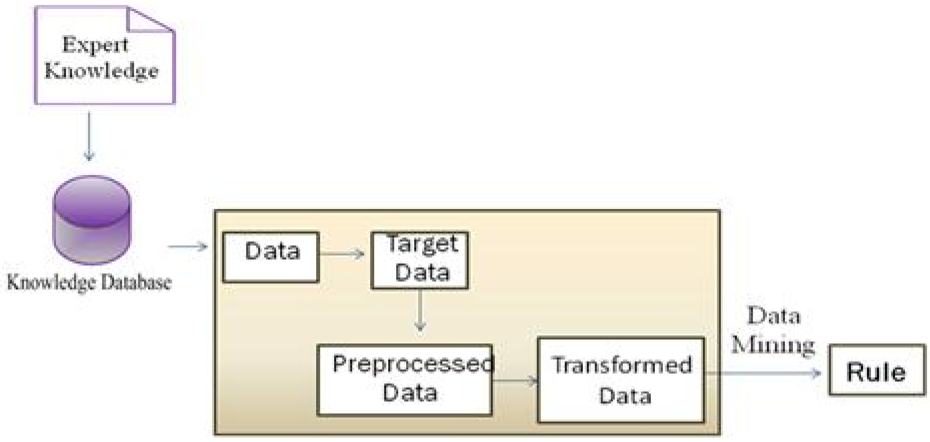

Fig. 5. Data mining process

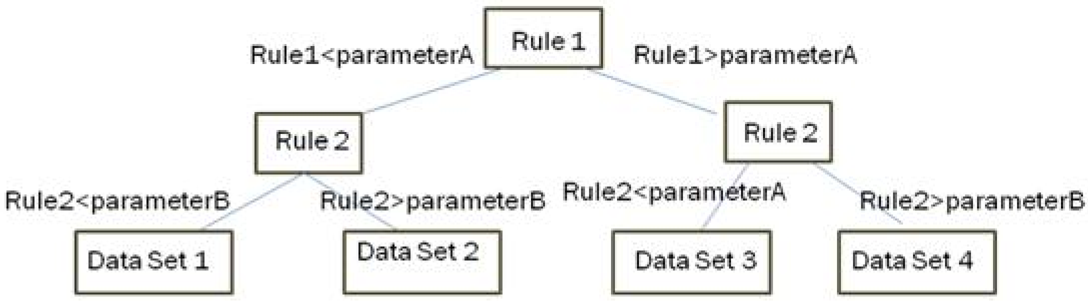

Fig. 6. Decision tree

In this study, we use decision tree method to identify the relations among different expert rules. Data is classified between the root node and child nodes and there would be various paths. As shown in figure 6, expert rules Rule1 and Rule2 are nodes of the decision tree. When we configure the variables between nodes, different relations are formed and new paths appear, thus produces different data sets (Data Set 1 - 4).

4. Comparison of System Efficiency

Many studies implement the integration of heterogeneous platform by combining SCA concept with BPEL method. In this section, we are going to compare the efficiency between two heterogeneous platforms: SCA concept with ESB method and SCA concept with BPEL method.

In this study, we use Apache Tuscany, Apache Synapse, and ActiveBPEL to implement SCA, ESB, and BPEL respectively.

First, we start from implementing two web services, Web Service A and Web Service B, under SCA environment and connect them with different financial databases, MySQL and SQL Server, respectively. Then we developed an application set up in both ESB and BPEL environments. The application can automatically determine which of Web Service A and Web Service B to call, and retrieve 25-pen data per round from Web Service A and Web Service B respectively. Table 1 shows the comparison of platform efficiency after 1, 10, 35, and 50 rounds.

Table 1. Comparison of platform efficiency (ms)

Method \ Rounds	1	10	35	50
SCA + ESB	3,391	22,991	91,105	110,333
SCA + BPEL	5,810	26,811	104,421	132,213

The experimental results prove that the efficiency of the combination of SCA and ESB is better than the combination of SCA and BPEL. The speed of data transferring of ESB is faster since the ESB can connect and convert data formats and has the distributed characteristics. When there are a lot of heterogeneous databases, it will become difficult to manage and maintain due to the large variety of data formats. But using ESB, we just have to transform the data into XML format and then convert the data formats through connected endpoints. Thus, the ESB improves not only the efficiency but the convenience of management.

5. Conclusion

This study builds an expert system based on SCA, uses expert rules to get key factors, classify data by data mining techniques, and then investigate by JESS. The architecture integrates heterogeneous platform and expert system allows user to discover the data they need much faster. The architecture can be applied into any field, in this study we use financial prediction system as an example, through the system architecture simulation, we can solve the heterogeneity problem among the financial services of each enterprise, quickly inferring the potential financial crisis enterprises, simplifying the regular data uploading procedure from enterprise to the Taiwan Stock Exchange Corporation (TWSE), and saving the cost for getting and integrating financial data. This study compares two types of heterogeneous platform implementation: composition of SCA and ESB, and composition of SCA and BPEL.

In this study, we broadly adopt semi-automated inference of expert knowledge, determine the rules using both expert knowledge and data mining techniques, and stored the rules into databases so that the rules can be reused. Nevertheless, the rules still need to be modified manually, and there is no algorithm to evaluate the level of rules. The algorithm can be studied in the future so that the inference of expert system can be automated.

Acknowledgments

The research is supported by the National Science Council of Taiwan under the grant Nos. NSC 100-2811-H-009-011 and NSC 100-2622-H-009-001-CC3.

References

[1] E.I. Altman, "Financial Ratios, Discriminant Analysis and the Prediction of Corporate Bankruptcy", Journal of Finance, Vol. 23, No. 4, pp.589-609, 1968.

[2] E.I. Altman, G.V. Marco, F. Varetto, "Corporate distress diagnosis: comparisons using linear discriminant analysis and neural networks", Journal of Banking and Finance, Vol.18, No. 3, pp. 505-529, 1994.

[3] W. Beaver, "Financial Ratios as Predictors of Failure", Journal of Accounting Research, Vol. 4, pp.72-102., 1966.

[4] E.B. Deakin, "A Discriminant Analysis of Predictors of Business Failure", Journal of Accounting Research, Vol. 10, No. 1, pp.167-179, 1972.

[5] E.K. Laitinen, "Financial Ratios and Different Failure Processes," Journal of Business Finance and Accounting, Vol.18, No. 5, pp. 649-673, 1991.

[6] M.D. Odom, R. Sharda, "A neural network model for bankruptcy prediction", in Proceedings of International Joint Conference on Neural Networks, Vol. 2, pp. 163- 168, San Diego, CA , USA, 1990.

[7] P.J. Booth, "Decomposition Measures and the Prediction of Financial Failure", Journal of Business Finance & Accounting, Vol. 10, No. 1,pp. 67-82,1983.