國 立 交 通 大 學

資訊管理研究所

碩 士 論 文

個 人 化 複 合 式 電 子 服 務 推 薦 之 實 作

Implementation of Personalized Recommendations for Composite E-services

研 究 生： 李 春 鋒

指導教授： 劉 敦 仁 博士

中 華 民 國 九 十 三 年 六 月

個 人 化 複 合 式 電 子 服 務 推 薦 之 實 作

Implementation of Personalized Recommendations for Composite E-services

研 究 生： 李 春 鋒 　　Student: Chun-Feng Lee

指導教授： 劉 敦 仁 　　Advisor: Duen-Ren Liu

國立交通大學

資訊管理研究所

碩士論文

A Thesis
Submitted to Institute of Information Management
College of Management
National Chiao Tung University
in Partial Fulfillment of the Requirements
for the Degree of
Master of Business Administration
in
Information Management
June 2004
Hsinchu, Taiwan, the Republic of China

中華民國九十三年六月

個人化複合式電子服務推薦之實作

研究生：李春鋒　　　　　　　　　　　　　指導教授：劉敦仁 博士

國立交通大學資訊管理研究所

摘要

隨著網路資源的蓬勃，電子商務市場產生劇烈的變動，強烈影響使用者的服務需求型態。然而企業所提供的單一電子服務已經無法滿足使用者需求，因此，企業開始與其他的電子服務提供者進行協同商務的整合，來發展複合式電子服務流程，以便提升本身在市場的競爭能力。複合式電子服務是多個基單一的電子服務所組成，其所帶來的附加價值遠高於單一的電子服務，再者，為了能夠在電子商務的環境中提供一對一行銷的服務，因此個人化推薦系統的建置也日趨重要。但僅管目前有許多平台紛紛提出複合式電子服務，但仍未見到針對複合式電子服務來進行個人化推薦之系統。

有鑑於此，本研究將運用合作式過濾方法以及資料探勘技術建置一個以複合式電子服務流程為主的個人化推薦系統，其目的在於將個人化推薦系統整合於複合式電子服務平台上，使客戶能更有效的選擇複合式電子服務來滿足其服務需求。

Implementation of Personalized Recommendations for Composite E-Services

Student：Chun-Feng Lee Advisor：Dr. Duen-Ren Liu

Institute of Information Management

National Chiao-Tung University

Abstract

With the explosive growth of information resources on the Internet, the electronic marketplace has changed dramatically, which strongly influences the customers' behaviour of demands. Individual e-services cannot fulfill customer's demands anymore. Therefore, recommender system are the solution that allow enterprises to develop one-to-one marketing strategies and provide adequate support to fulfill customers need to boost the success of online e-business. Enterprises need to provide various e-services and conduct composite e-services on the collaborative commerce environments, which aim to improve the market competitive advantages.

There are two objectives in this research. Firstly, we propose a useful personalized recommender system that combines collaborative filtering method and data mining techniques for generating composite e-services recommendations. Secondly, we provide an integrated knowledge map navigator with personalized recommendations in our proposed composite e-services platform. We expect not only to reduce the complexity and cost for the e-service providers online, but also to fulfill user's needs online.

誌謝

短短的兩年交大研究生涯在不知不覺中到了尾聲。首先，最要感謝我的指導教授-劉敦仁老師，這兩年對我的諄諄教誨，學生將永遠銘記在心。也要感謝口試委員陳恭老師、王朝煌老師、羅濟群老師所給予的寶貴意見，讓我的研究能更加完整。

感謝實驗室的學長學姐們，美玉、志坤、孟蓉、純和、美奐，謝謝你們給我在課業上以及生活上的諸多幫助，尤其要特別感謝雅月與怡瑾學姐在百忙中還費心給予我在論文寫作上的指導；感謝實驗室的同學們，嘉源、錦慧、昆學、瑞雲，有你們跟我一起為論文奮戰的感覺真好，期待我們的前途一片光明；感謝實驗室的學弟學妹們，桄瑋、韋孝、柏村、秋雯、明立，為我們實驗室注入新的活力與歡笑；此外，要感謝我們資管所最美麗的助理-淑惠，你有如大姐姐般的細心照顧我們這群資管所的小蘿蔔頭。

最後，要感謝我的父母和姐姐們的支持，有你們的鼓勵與關懷，才讓我得以盡心的去完成碩士學位。

Content

1. INTRODUCTION 1

1.1 Background and Motivations 1

1.2 Research Objectives 2

1.3 The Research Structure 3

2. RELATED WORKS 4

2.1 E-Service 4

2.1.1 E-Services and Composite E-Service Platform 4

2.1.2 Composite E-Service 5

2.2 Web Service 6

2.2.1 The Universal Description, Discovery and Integration (UDDI) 7

2.2.2 Web Service Description language (WSDL) 8

2.2.3 Simple Object Access Protocol (SOAP) 9

2.3 Topic Maps 10

2.3 Recommendation 10

2.4.1 Collaborative Filtering 11

2.4.2 Content Based Filtering 12

2.5 Data Mining 13

2.5.1 Introduction of Data Mining 13

2.5.2 Association Rules Mining 14

2.5.3 Process Mining 17

3. OVERVIEW OF COMPOSITE E-SERVICE PLATFORM 18

3. 1 Knowledge Map Building System 19

3.2 Personalized Recommendation System 20

4. PERSONALIZED RECOMMENDATION SYSTEM 22

4.1 User Feedback Module 24

4.2 User Preference Module 24

4.3 Flow Mining Module 25

4.3.1 Mining Frequent Ordering Sets.. 29

4.3.2 Mining Frequent Attributes Sets 30

4.3.3 Total Scores of Composite E-Service... 32

4.4 Personalized Recommendation Module 32

4.5 Middleware Module 33

5. SYSTEM IMPLEMENTATION AND DEMONSTRATION.. 36

5.1 Personalized Recommendation System Demonstration.. 36

5.1.1 Demonstration of Personalized Recommendation System 37

5.1.1.1 Feedback Page 38

5.1.1.2 Users Preference Similarity... 39

5.1.1.3 User's Unused E-Services Preference Prediction 42

5.1.1.4 Flow Ordering Score for Composite E-Service 45

5.1.1.5 Composite E-Service Recommendation Score.................................. 51

5.1.2 Perform Recommendations on the Composite E-Service Platform 54

5.1.2.1 User Identification.. 54

5.1.2.2 User Navigating Category Identification... 55

5.1.2.3 Matching Process between User and Recommendation List 55

6. CONCLUSION AND FUTURE WORKS 58

6.1 Conclusions 58

6.2 Future Works 59

REFERENCE.. 60

APPENDIX 64

List of Figure

Fig. 2.1 Composite Service Description Language diagram (CSDL).................................. 5

Fig. 2.2 UDDI、WSDL、SOAP 7

Fig. 2.3 Association Rule Mining Algorithm (1) 15

Fig. 2.3 Association Rule Mining Algorithm (2) 15

Fig. 2.3 Association Rule Mining Algorithm (3) 16

Fig. 2.4 Association Rule Mining Example... 16

Fig. 4.1 Personalized Recommendation System 23

Fig. 4.2 Infrastructure of Middleware Module... 34

Fig. 5.1 Starting Page of the Prototype System.. 37

Fig. 5.2 User E-Services Feedback Page 38

Fig. 5.3 Users Preference Similarity Page 40

Fig. 5.4 Comparison between Latest and Current User Preference Similarity 40

Fig. 5.5a Functional Description Link.. 41

Fig. 5.5b Functional Description for the User Preference Similarity 41

Fig. 5.6 User Unused E-Services Preference Prediction Page... 43

Fig. 5.7 E-service Preference Prediction Comparison 43

Fig. 5.8a Functional Description Link... 44

Fig. 5.8b Functional Description in User Preference Prediction Page............................... 44

Fig. 5.9 Composite E-Service Flow Ordering Calculation Page 46

Fig. 5.10 Comparison between Current and Latest Ordering Recommendation Score 46

Fig. 5.11a Functional Description Link... 47

Fig. 5.11b Functional Description in Composite E-Service Ordering Score Page............. 47

Fig. 5.12 Composite E-Service Attribute Calculation Page.. 49

Fig. 5.13 Comparison between Latest and Current Attribute Scores 49

Fig. 5.14a Functional Description Link 50

Fig. 5.14b Functional Description in Attributes Calculation Page 50

Fig. 5.15 Composite E-Service Recommendation Calculation Page 52

Fig. 5.16 Comparison between Current and Latest Recommendation Scores 52

Fig. 5.17a Functional Description Link 53

Fig. 5.17b Functional Description in Total Score Calculation Page 53

Fig. 5.18 User Login Page for Personalized Request 54

Fig. 5.19 Top 3 Personalized Composite E-Service Recommendations 56

Fig. 5.20 Top 3 Personalized Basic E-Service Recommendation 56

Fig. 5.21 Composite E-Service Flow Link 57

Fig. 5.22 Link from the Recommendation List Back to Composite E-Service Navigator 57

List of Table

Table 4.1 Example of Ordering List for Composite E-Service 27

Table 4.2 Example of Attributes List for E-Service 28

Table 4.3 Frequent Ordering Sets.. 29

Table 4.4 Frequent Attribute Sets 30

1. Introduction

At first, the research background and motivations are introduced. Then the objectives and the structures for this research are described in Sections 1.2 and 1.3, respectively.

1.1 Background and Motivations

With the explosive growth of Internet, the electronic marketplace has dramatically becomes the influential factor in the e-business transactions. Due to this change, enterprises provide various e-services and conduct composite e-services for collaborative commerce online to achieve e-market competitive advantages. In such complex collaborative commerce environments, online user is facing the difficulty of how to select the appropriate composite e-services for their needs. So it is essential to have an effective knowledge system which can manage and access the related information resources in the composite e-services environment.

A complete service normally consists of various basic services, so by providing individual e-service online will not satisfy customer's demand. Because of this, composite e-services have become the new solution for an enterprise to serve online user for their needs of a complete service. Enterprise provides e-services hoping to generate new revenue and strength and improve its competitiveness against their competitors online. Since each e-service providers might not use the same type of online system platform, it will be difficult for them to communicate and exchange information. Therefore, topic maps standard is the solution that provides a bridge between the domains of knowledge representation and information management for existing information resources.

In B2C e-commerce, many researches have proposed the composite e-service flow structures **[4][9]**, but very few researches consider about organize service providers' information resource through knowledge map and transform the existing information resource into valuable knowledge. Knowledge map combined with personalized recommendation approach can solve the bottleneck of semantic search from the website and assisting user with customized decision support.

1.2 Research Objectives

Based on the above motivations, we address the research objectives in two-folds, described as follows:

1. Implementing a well-known recommendation approach with data mining techniques for building a personalized recommender system under the composite e-services platform.
2. Implementing a prototype system that integrates knowledge maps navigator and personalized recommendations under the composite e-services platform.

1.3 The Research Structure

This research implements a personalized composite e-service recommendation system that aims to create value-added composite e-service flow to provide users with personalized recommendation support. While the online users are browsing the composite e-service information, the composite e-service platform will support user with the personalized recommendations, which assist users to make decision by selecting the suitable composite e-services. The rest of research is organized as follows, Section 2 introduces the related works, includes definition of composite e-services, collaborative recommendation, mining approach and topic map standards, etc. Section 3 describes the overview of the advanced composite e-service platform that includes knowledge map building system and personalized recommendation system. This research focuses on the personalized recommendation system. The section on knowledge maps building system will be introduced in another research. Section 4 describes the architecture and functionality of personalized recommendation system's architecture and functionality. Section 5 demonstrates the prototype system. The conclusion and future works are finally made in section 6.

2. Related works

The aim of this research is to implement a personalized recommender system under the composite e-services platform. The relate works of this research includes e-service definitions, web service standards, topic maps standards, recommendation approaches, and data mining techniques.

2.1 E-Service

2.1.1 E-Services and Composite E-Service Platform

E-services, a business concept developed by Hewlett Packard (HP), is the idea that the World Wide Web is moving beyond e-business and e-commerce into a new phase, where many business services can provide to the business or consumer using the Web. HP defines e-services as modular, nimble, electronic services that perform work, achieve tasks, or complete transactions **[3] [14]**. Almost any asset can be turned into an e-service and offered via the Internet to drive new revenue streams and create new efficiencies. R. Balakrishnan **[4]** defined an e-service as any application, written potentially in any language that conforms to a set of meta-data and can received and responds to messages as defined by its service specification.

2.1.2 Composite E-Service

Casati **[6] [7] [8] [9]** provides a definition, analysis and enactment of e-services: A composite e-service is a combination of several basic e-services and works similar to a workflow approach. Casati and other researchers have designed a composite service description language (CSDL) to describe the composition by means of a directed graph. The nodes in the CSDL represent the interactions between e-services while arcs define the execution of dependencies among services. Special nodes are provided to denote the starting and ending points of the composite service, or to route the execution flow as shown in Fig. 2.1.

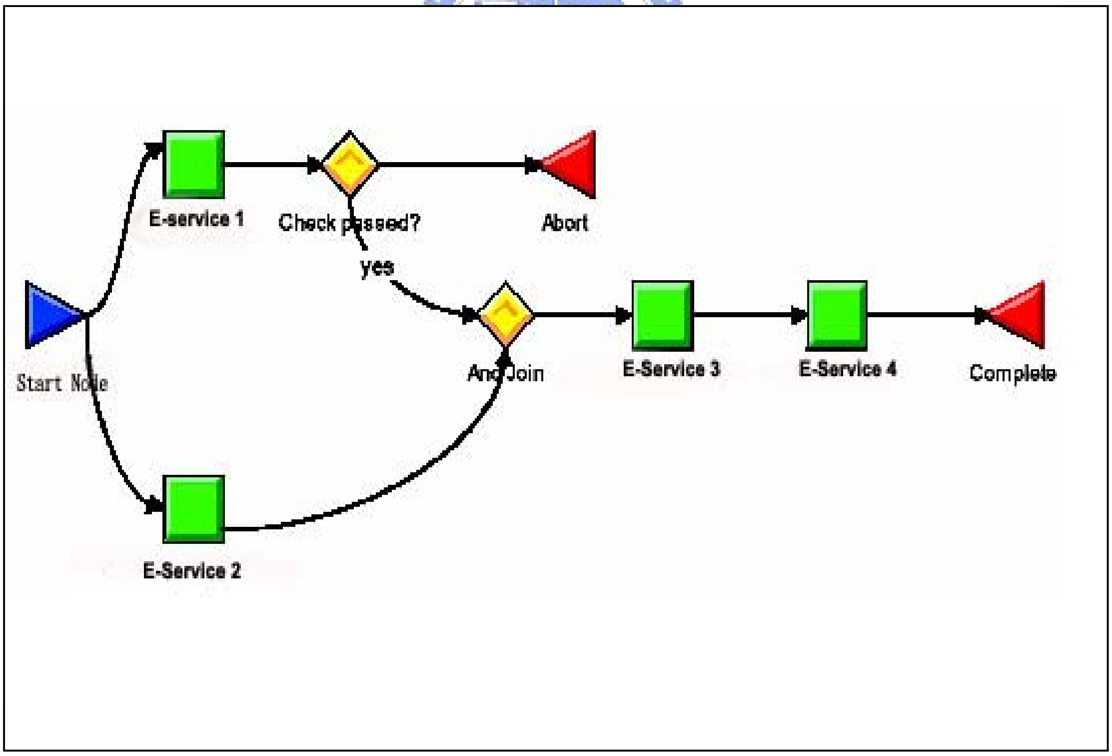


Fig. 2.1 Composite Service Description Language diagram (CSDL)

Piccinelli and Williams **[20]** developed a DySCo (Dynamic Service Composition) project. Their workflow model is a basis for a multi-layered composition framework for web services. Liu **[19]** performs a composite e-service platform with recommendation ability.

2.2 Web Service

> *Web services are a new breed of Web application. They are self-contained, self-describing, modular applications that can be published, located, and invoked across the Web. Web services perform functions, which can be anything from simple requests to complicated business processes. A sample Web service might provide stock quotes or process credit card transactions. Once a Web service is deployed, other applications (and other Web services) can discover and invoke the deployed service.*
>
> *IBM web service tutorial (www.ibm.com)*

The core of Web service standards mainly includes Simple Object Access Protocol (SOAP), Web Services Description Language (WSDL), and Universal Description, Discovery and Integration (UDDI). UDDI provides a mechanism for clients to find web services. WSDL defines services as collections of network endpoints or ports. SOAP is a message layout specification that defines a uniform way of passing XML-encoded data. Web services can significantly increase the online business's potential, by providing a way of automating e-service communication and discovery of e-services. The core of web service diagram is shown in Fig. 2.2

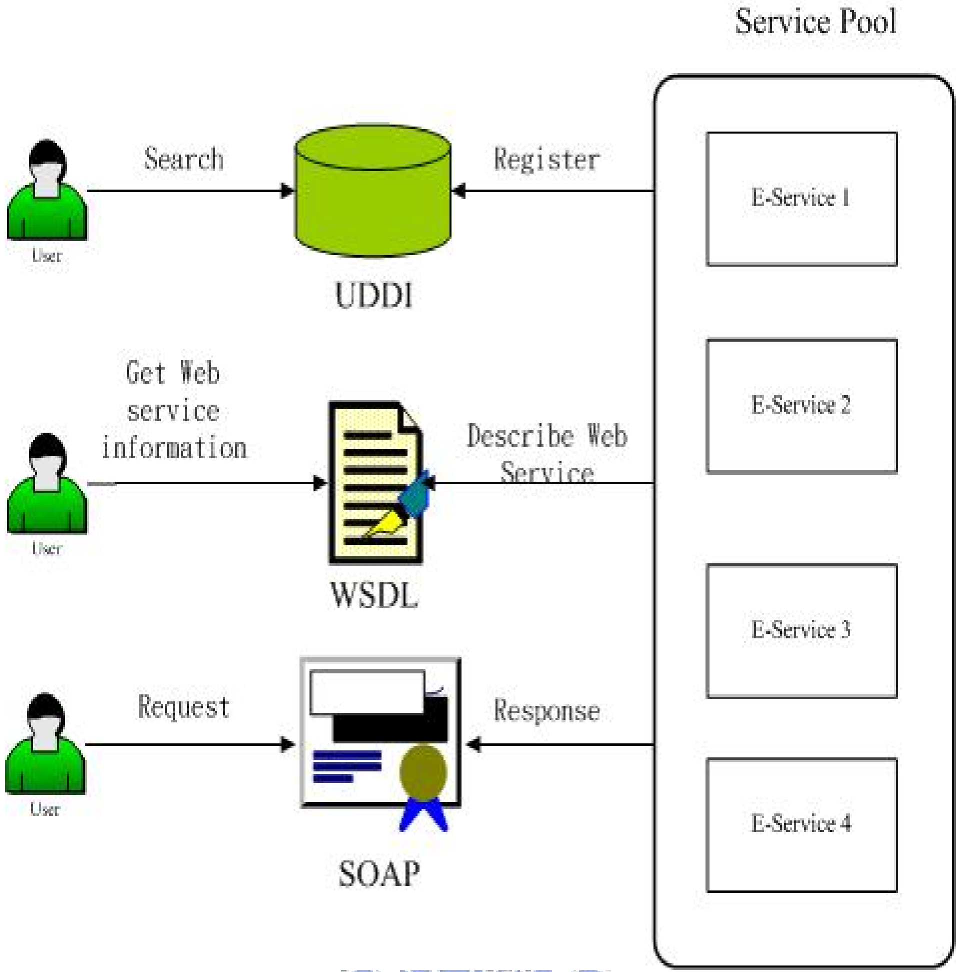


Fig. 2.2 UDDI、WSDL、SOAP

2.2.1 The Universal Description, Discovery and Integration (UDDI)

UDDI provides a mechanism for users to find web services **[11] [25]**. Using a UDDI interface, businesses can dynamically look up as well as discover web services provided by online business. A UDDI registry has two kinds of users: businesses and customers. Businesses are the one who want to publish the web service description while customers who want to obtain services descriptions of a certain kind and bind to them programmatically. UDDI itself is layered over SOAP and assumed that requests and responses are UDDI objects sent around as SOAP messages.

The UDDI consists of four levels of information, described as follows:

1. The top-level element is the *Business entity*: This provides information about the party who publishes services. This kind of information can be seen as the white pages of UDDI.
2. The second-level element is the *Business service*: This provides a description of the particular service and a list of categories that describe the service.
3. The third-level element is the *Binding templates*: This provides the more technical information about a web service.
4. The fourth-level element is the *Publish assertion*. This provides information about a relationship between two parties, asserted by one of both.

2.2.2 Web Service Description language (WSDL)

WSDL defines services as the collections of network endpoints or ports **[11] [27]**. In WSDL the abstract definition of endpoints and messages are separated from their concrete network deployment or data format bindings. This allows the reuse of abstract definitions of messages, which are abstract descriptions of the data being exchanged and the port types are represented as the abstract collections of operations. The concrete protocol and data format specifications for a particular port type constitute a binding. A port is defined by associating a network address with a binding; a collection of ports defines a service.

2.2.3 Simple Object Access Protocol (SOAP)

SOAP is a message layout specification that defines a uniform way of passing XML encoded data **[11] [24]**. It also defines a way to bind to HTTP as the underlying communication protocol for passing SOAP messages between two endpoints. SOAP is similar to techniques such as DCOM, RMI and CORBA in providing a simple, lightweight RPC-like mechanism. SOAP is basically a technology providing a very simple one-way as well as request/reply mechanism. A SOAP message contains three primary parts: an envelop, a header for adding application-specific features to SOAP message, and a body that contains information intended for the recipient.

2.3 Topic Maps

The Topic Maps is an international standard (ISO/IEC 13250) first published in January 2000 **[15]**. Topic Maps provide a bridge between the domains of knowledge representation and information management and link it to existing information resources. The basic concepts are shown as following:

- ***Topics*** are the main components in the topic maps. In a topic map, any given topic can be anything **[21]**; in the case of composite e-services, any e-service providers or e-service is a type of topic.

- ***Associations*** are the original information resource connected to the meaningful link that is specified among several topic names **[21]**.

- ***Occurrences*** are the information resources linked to the meaningful topic names. A topic may be linked to one or more information resources that are deemed to be relevant to the topic in some way. Such resources are called occurrences of the topic. **[21]**.

Topic Maps standardization can provide a clear structure in assisting an enterprise to organize knowledge from different information resources and building a knowledge-sharing environment for user to gain knowledge. The goal of Topic Maps visualization is to help users to locate relevant information quickly and explore the structure easily. There are two types of visualization: representation and navigation [12]. Our research in composite e-service platform is using navigation visualization interface as the interface.

2.4 Recommendation

2.4.1 Collaborative Filtering

Collaborative filtering approach identifies the relevant users who owns similar profiles and provides the data of their preferences. The basic idea of Collaborative Filtering algorithms is to provide e-service recommendations or predictions based on the opinions of other like-minded user. Users usually require input ratings about piece of information. These ratings are then used to compute user's correlation coefficients among existing users. The correlation coefficient is a measure of the similarity between two different users' preferences. The recommender system generates recommendations based on the predictions of user's preference. This approach doesn't consider any analysis of the item itself **[22]**. Example of system taking of this approach that achieved successfully result is GroupLens **[16]**. In recent recommendation approach, collaborative filtering often combines with other approach to increase the prediction for recommendations, such as content-based filtering and association rules **[17] [18] [20] [23]**.

Researchers have devised a number of collaborative filtering algorithms that can be divided into two main categories: Memory-based (user-based) and Model-based (item-based) algorithms **[5]**. We provide a detailed analysis of collaborative filtering based recommender system algorithms in following section.

1. ***Memory-based Collaborative Filtering Algorithms****.* Memory-based algorithms utilize the entire user-item database to generate a prediction. These systems employ statistical techniques to find a set of users; known as neighbours, that have a history of agreeing with the target user. Once a neighbourhood of users is formed, these systems use different algorithms to combine the preferences of neighbours to produce a prediction or top-n recommendation for the active user.

2. ***Model-based Collaborative Filtering Algorithms.*** Model-based collaborative filtering algorithms provide item recommendation by first develop a model of user ratings. Algorithms in this category take a probabilistic approach and envision the collaborative filtering process as computing the expected value of a user prediction, given his/her ratings on other items. The model building process is performed by different machine learning algorithms such as Bayesian network, clustering, and rule-based approaches.

2.4.2 Content Based Filtering

Content-based filtering collects information from the user and comparing its past behaviour and preference as the representation of contents of user profiles which express the interests of users. A user profile is built up by analyzing accumulated user rating content based on past purchasing behaviour. In this method, the techniques of information retrieval are used in content analysis.

2.5 Data Mining

2.5.1 Introduction of Data Mining

In the information age, because of the explosive growth of the data being collected in the databases, many enterprises require the data mining techniques tools to intelligently and automatically transform the processed data into useful information and knowledge to understand customer behaviour patterns, and improve both the quality of service and competitiveness.

Because of the above reason, data mining has become a research area with increasing importance. Data mining is also referred to as knowledge discovery in databases, which means a process of nontrivial extraction of implicit, previously unknown and potentially useful information (such as knowledge rules, constraints, regularities) from data in databases. **[10] [13]**

Several requirements and challenges of data mining are faced during the development of data mining techniques. **[10]**

1. Handling of different types of data.
2. Efficiency and scalability of data mining algorithms.
3. Usefulness, certainty and expressiveness of data mining results.
4. Expression of various kinds of data mining results.
5. Interactive mining knowledge at multiple abstraction levels.
6. Mining information from different sources of data.
7. Protection of privacy and data security.

Until now, there are many kinds of techniques that can be used to discover knowledge, such as association rules, classification rules, clustering, etc. In this research, the association rule mining technique is used to find frequent itemsets and obtain support value to make recommendations to users.

2.5.2 Association Rules Mining

Association rule algorithm identifies the correlations between items in transactional databases. Given a set of transactions, where each transaction is a set of items, an association rule is an expression *X* => *Y*, where *X* and *Y* are sets of items. The intuitive meaning of such a rule is that transactions in the database which contain the items in *X* tend to also contain the items in *Y*.
An example of such a rule might be that 80% of customers who purchase tires and auto accessories also buy some automotive services; here 80% is called the confidence of the rule. The support of the rule *X* => *Y* is the percentage of transactions that contain both X and Y.

1. **Support**: Percentage of transactions that contain both A & B:

   (A => B) = P (A *U* B)

2. **Confidence**: Percentage of transactions containing A, which also contain B:

   (A => B) = P (B | A)

The core of association rules mining is to find all rules that satisfy a user-specified minimum support and minimum confidence. Applications include cross-marketing, attached mailing, catalog design, loss-leader analysis, store layout, and customer segmentation based on buying patterns **[2]**.

The Apriori Algorithm is used in association rule mining and is shown in detail in Fig. 2.3.

```
Algorithm: Apriori. Find frequent itemsets using an iterative level-wise approach
           based on candidate generation.
Input: Database, D, of transactions; minimum support threshold, min_sup.
Output: L, frequent itemsets in D.
Method:

L1 = find_frequent_1-itemsets(D);
for (k=2; Lk-1 ≠ Ø; k++){
    Ck = apriori_gen(Lk-1 , min_sup);
    for each transaction t ∈ D {// scan D for counts
        Ct = subset(Ck, t); //get the subsets of t that are candidates
        for each candidate c ∈ Ct
            c.count++;
    }
    Lk = {c ∈ Ck | c.count ≥ min_sup}
}
return L= ∪k LK ;
```

Fig. 2.3 Association Rule Mining Algorithm (1)

```
procedure apriori_gen (Lk-1: frequent (k-1);
                       min_sup: minimum support threshold)

for each itemset l1 ∈ Lk-1
    for each itemset l2 ∈ Lk-1
        if (l1[1]=l2[1]) ∧ ..∧ (l1[k-2]=l2[k-2]) ∧ (l1[k-1] <l2[k-1]) then{
            c = l1 >< l2; //join step: generate candidates
            if has_infrequent_subset(c,Lk-1) then
                delete c; //prune step: remove unfruitful candidate
            else add c to Ck;
        }
return Ck;
```

Fig. 2.3 Association Rule Mining Algorithm (2)

```
procedure has_infrequent_subset (    c: candidate k-itemset;
                                     L_{k-1}: frequent (k-1)-itemsets    )
        // use prior knowledge
(1)     for each (k-1)-subset s of c
(2)         if s ∉ L_{k-1} then
(3)             return TRUE;
(4)     return FALSE;
```

Fig. 2.3 Association Rule Mining Algorithm (3)

Fig. 2.4 is an example of transaction database example; assume that the minimum support for the large itemset is 40%, applied to above association rule mining algorithm.

Database D

TID	Items
100	A C D
200	B C E
300	A B C E
400	B E

Scan D →

C_1

Itemset	Sup.
{A}	2
{B}	3
{C}	3
{D}	1
{E}	3

L_1

Itemset	Sup.
{A}	2
{B}	3
{C}	3
{E}	3

C_2

Itemset
{A B}
{A C}
{A E}
{B C}
{B E}
{C E}

Scan D →

C_2

Itemset	Sup.
{A B}	1
{A C}	2
{A E}	1
{B C}	2
{B E}	3
{C E}	2

L_2

Itemset	Sup.
{A C}	2
{B C}	2
{B E}	3
{C E}	2

C_3

Itemset
{B C E}

Scan D →

C_3

Itemset	Sup.
{B C E}	2

L_3

Itemset	Sup.
{B C E}	2

Fig. 2.4 Association Rule Mining Example

2.5.3 Process Mining

The goal of workflow management is to handle cases as efficient and effective as possible. A workflow process is designed to handle similar cases. Cases are handled by executing tasks in a specific order. The workflow process model specifies which tasks need to be executed and in what order. But workflow design is very complex and time consuming; therefore process mining is used to discovering the actual process models. Process mining starts by gathering information about the processes through workflow logs **[1] [28] [26].** Workflow log contains information about the workflow process as it is actually being executed. The goal of process mining is to reverse the process and collect data at runtime to support workflow design and analysis. The information collected at run-time can be used to derive a model explaining the events recorded.

In this research, process mining is used in the composite e-service flow schema, where association rule mining is used to find the item sets from composite e-service execution logs and generate into the recommendation list in the database.

3. Overview of Composite E-Service Platform

The composite e-service platform aims to add values to the composite e-services by providing users with the knowledge map navigator and personalized recommendation support. Our proposed composite e-service platform includes two main systems: Knowledge Map Building System and Personalized Recommendation System. Fig. 3.1 shows the architecture of the composite e-service platform.

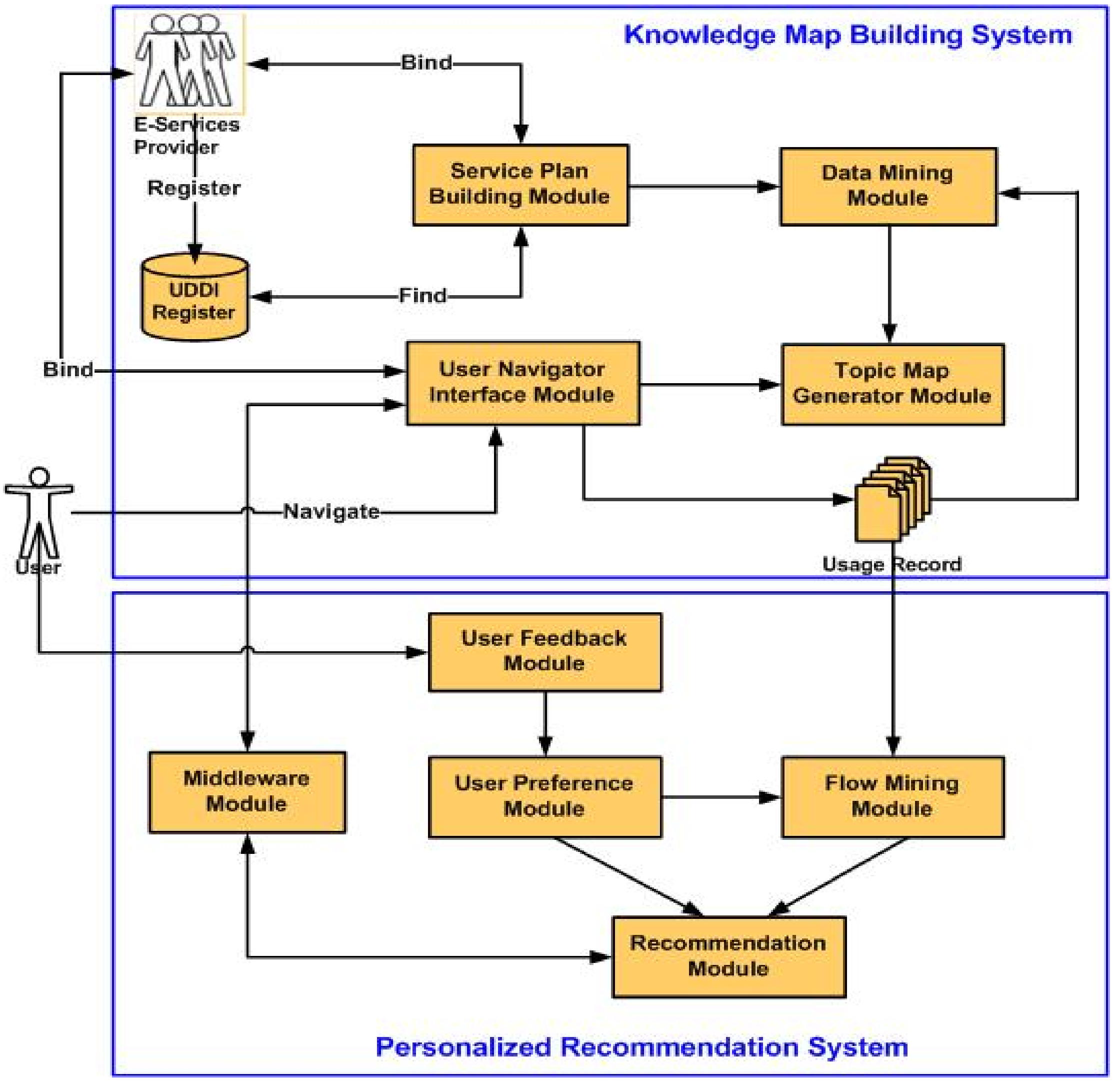


Fig. 3.1 The Architecture of Composite E-Service Platform.

3. 1 Knowledge Map Building System

The overall process of this system can be classified into two components: online and offline. Offline components consider the building and structuring of a knowledge maps environment through topic maps format. Once the knowledge map building system accomplishes its offline process, its online component considers the user friendliness for the knowledge maps navigator while searching for e-service online. The Knowledge map building offline components are covered in module one, two, and three; online component is covered in module four.

1. *Service plan building module*

   This phase is where various e-service providers provide their service information here and predefine some service plans by plan designers.

2. *Data mining module*

   Data mining approach is used here to find some useful knowledge patterns between composite e-services and their attributes.

3. *Topic Map Generator Module*

   XTM (XML topic maps) is used here to create knowledge map by arranging subjects and ontology relations due to knowledge patterns.

4. *User Navigating interface module*

   User can navigate the knowledge patterns for composite e-service by using a XTM based composite e-service navigator interface.

3.2 Personalized Recommendation System

The overall process of this system can also be classified into two main parts: online and offline. Offline components consider the clustering of users with their preference group and mining through preference group's usage records and in the end generate a personalized recommendation list. Online components of personalized recommendation system consist of two components: user feedback collection and a middleware which works closely with user navigating interface module from knowledge map building system. The personalized recommendation offline component is covered in modules one, two and three; online component is covered in modules four and five.

1. *User preference module*

   This phase uses collaborative filtering algorithm to find the similarity between users and to form the nearest neighbours of their preference similarity and make predictions of individual services based on the opinions of other like-minded users.

2. *Composite e-service flow mining module*

   This phase uses the association rule mining techniques and score approach to find composite e-service flow predictions based on user's usage records in the preference group.

3. *Personalized Composite E-Service Recommendation module*

   Top N approach is used here after generating complete composite e-service recommendations by gathering both basic service predictions and composite e-service flow predictions.

4. *User Feedback module*

This phase uses the collaborative filtering mining techniques to find user's preference group from customer database. It is very important to collect user's feedback about the preference for the different e-services after he/she finishes navigating the user navigating interface.

5. *Middleware module*

Middleware module is a bridge between the connections of both systems, while user searches the available service information in the composite e-service navigator; middleware module delivers some of personalized decision support recommendations for user to view.

4. Personalized Recommendation System

The personalized recommendation system implements two of the most popular approaches to generate personalized recommendation lists: collaborative filtering and data mining. By integrating collaborative filtering with data mining approach it lets the system to predict user's personal recommendations more intelligently than other earlier system, which only used data mining or collaborative filtering approach. The main objective of the personalized recommendation system is to analyze users' preferences on individual e-services and mining the scores for each composite e-service. The whole system consists of five modules and the detailed architecture system is shown in Fig. 4.1.

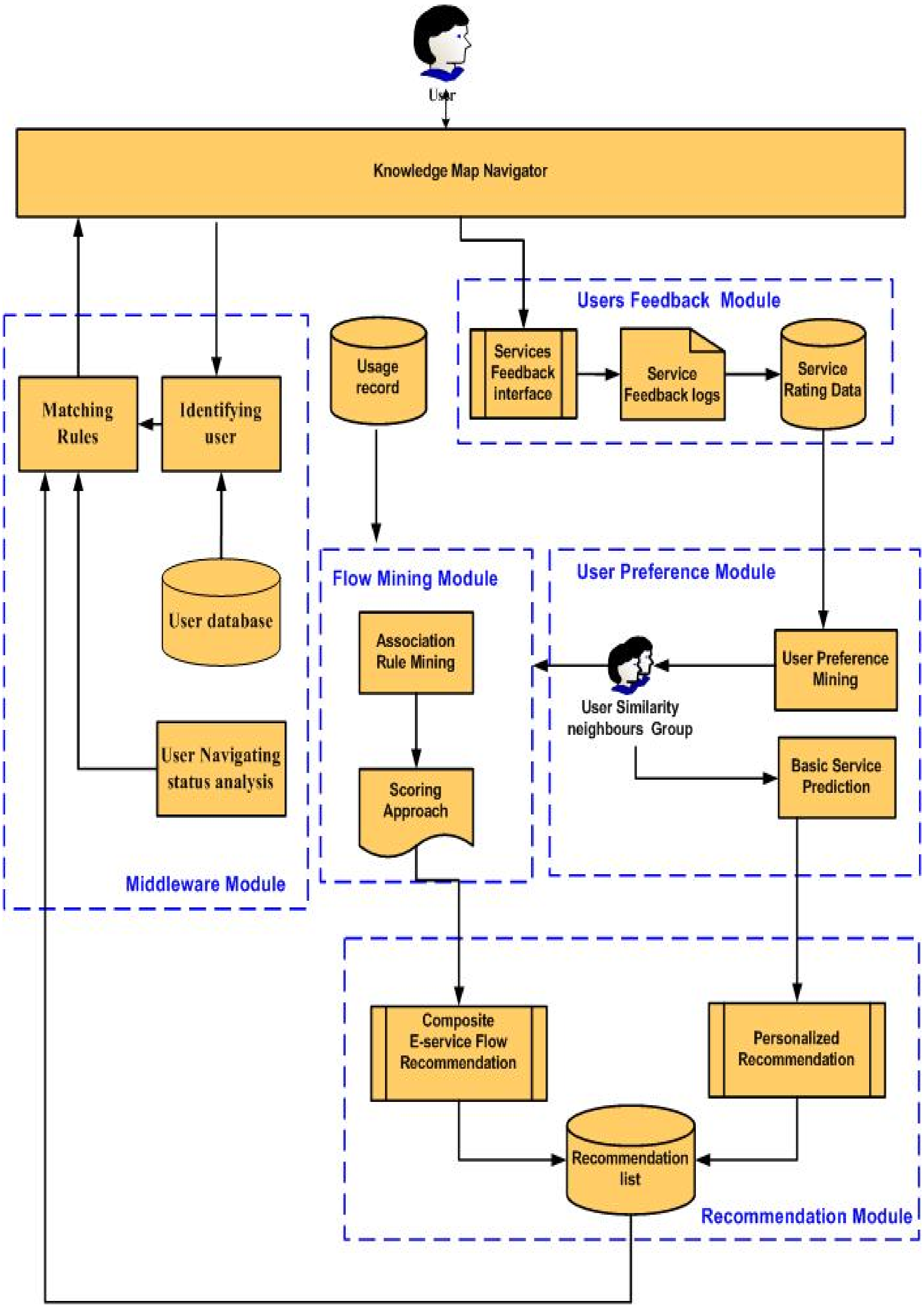


Fig. 4.1 Personalized Recommendation System

4.1 User Feedback Module

If a user uses any e-service that is provided in the composite e-service platform, we recommend the user to do an e-service evaluation feedback in our system. This explicit user data collects user rating on various e-services used by the users in the past.

4.2 User Preference Module

Pearson correlation coefficient is used to identify each user's preference similarity toward other users, and form the k-nearest neighbour for each user. It is calculated as in Equation 1.

$$w(u,i) = \frac{\sum_{j \in I_{ui}} \left(r_{u,j} - \overline{r_u}\right)\left(r_{i,j} - \overline{r_i}\right)}{\sqrt{\sum_{j \in I_{ui}} \left(r_{u,j} - \overline{r_u}\right)^2} \sqrt{\sum_{j \in I_{ui}} \left(r_{i.j} - \overline{r_i}\right)^2}} \qquad \textbf{(1)}$$

w(u,i) : The similarity between active user *u* (who we want to recommend it) and user *i*.

$r_{u,j}$, ($r_{i,j}$) : The rating of user *u*(*i*) for item *j*.

$\overline{r_u}, \overline{r_i}$: The average rating of user *u* (*i*)

I_{ui} : The set of items that were rated by both user *u* and user *i* .

This work does the above algorithm to calculate the similarity on e-service preference among all users, and then forms the nearest neighbours whose similarity value is greater than a specified threshold or select the nearest neighbours based on top N similarity value. Through the user's nearest neighbours, the system can predict user's preference on the individual e-services that user have not used before. The formula used to calculate the users' prediction score on unused e-services is shown in Equation 2.

$$p_{u,j} = \overline{r_u} + \frac{\sum_{i=1}^{n} w(u,i)\left(r_{j,i} - \overline{r_i}\right)}{\sum_{i=1}^{n} \left|w(u,i)\right|} \qquad (2)$$

$P_{u,j}$: Prediction for the active user *u* on item *j*.

$\overline{r_u}, \overline{r_i}$: The average rating of user *u* (*i*).

w(u,i) : The similarity between active user *u* and user *i*.

$r_{i,j}$: The rating of user *i* on item *j*

n : The number of user in neighborhood.

4.3 Flow Mining Module

Composite e-service is a composition of various individual e-services. This research uses a data mining approach to find frequent ordering scores and frequent attributes scores for every composite e-services flow schema stored in the database. To provide advanced recommendations of complete composite e-services, this system uses a scoring approach to recommend the top N composite e-services for customers.

The purpose of this module is the mining of the order between e-services and the attributes of basic e-services in the composite e-service. The mining results then form recommendation score for each composite e-service flow. Apriori algorithm is used here to mine the frequent orderings set between e-services and frequent attributes set of e-services.

Table 4.1 indicates the examples of composite e-service's flow definition for computer e-courses. CSID are the identifiers of flow schema definitions and CIS are the identifiers of instances of flow schemas. To clarify the steps for flow mining module, basic services are represented by capital letters. Flow schema definition is the graphic representation for composite e-services. Notably, the ordering between basic e-service A and B is transformed into (A, B). If the composite e-service exists a path flow from e-service A to e-service B, its means that e-service A precedes service B in the composite e-services. Therefore, in composite e-service CS05, we can derive the ordering list: {(B, D) (B, C) (B, E) (C, E)}.

Table 4.1 Example of Ordering List for Composite E-Service

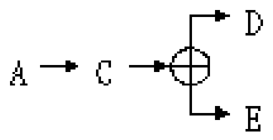
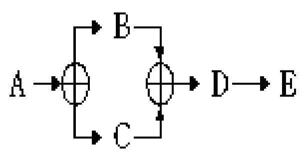
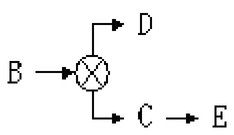
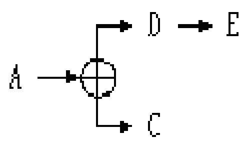
CSID	Flow Schema Definition	CIS	Instance Execution Logs	Ordering List
CS01		C001	ACD	(A, C) (A, D) (C, D)
		C002	ACE	(A, C) (A, E) (C, E)
CS02		C003	ABDE	(A, B) (A, D) (A, E) (B, D) (B, E) (D, E)
		C004	ACDE	(A, C) (A, D) (A, E) (C, D) (C, E) (D, E)
CS03	C → B → E	C007	CBE	(C, B) (C, E) (B, E)
CS04	A → B → G	C008	ABG	(A, B) (A, G) (B, G)
CS05		C008	BDCE	(B, D) (B, C) (B, E) (C, E)
CS06		C010	ADE	(A, D) (A, E) (D, E)
		C011	AC	(A, C)

Table 4.2 indicates the examples of e-service's attributes for computer e-courses. BSID are the identifiers of basic e-service in the composite e-service. To clarify the steps for attributes mining module, attributes is represented with small letters.

Table 4.2 Example of Attributes List for E-Service

BSID	Service Attributes (Database Computer Course)	Attributes List
BS01	Course Instructor= Mike, Course Level= Beginner, Course Location= Taipei, Course Provider=PC-School, Course Time schedule=Morning,	a1, b2, c3, d3, e2
BS02	Course Instructor= Nancy, Course Level= Advance, Course Location= Taipei, Course Provider= PC-School, Course Time schedule=Evening,	a2, b3, c3, d3, e1
BS03	Course Instructor= Mike, Course Level= Beginner, Course Location= TaiChung, Course Provider=Mitac, Course Time schedule= Evening,	a1, b2, c2, d1, e1

4.3.1 Mining Frequent Ordering Sets

The system can recommend the frequent ordering sets of each composite e-service flow. The following illustrates the mining of frequent ordering sets of flow schema. The frequent orderings of other composite e-services are derived similarly. Every element (X, Y) in the ordering list is a candidate item for deriving frequent ordering sets. The support of an ordering set is the ratio of instances that contain all orderings. Ordering sets with support values greater than the required minimum support values are called frequent ordering sets. We also use the Apriori Algorithm [2] to generate the frequent ordering sets that satisfy the required minimum support value. Suppose that the required minimum support value is 25%. Table 4.3 shows the example of frequent ordering sets and their corresponding support values for flow schema ACD.

Table 4.3 Frequent Ordering Sets

Frequent ordering sets	Support
(A,C)	55%
(A,D)	36%
(C,D)	21%

After finding the frequent ordering sets, the system computes the ordering score of composite e-services that include A, C, and D. The ordering score can be derived by the Equation 3.

$$\text{Order Score of composite e-service (CS)} = \sum_{\{x,y\} \in CS.ordering} \sup(\{x, y\}) \quad \textbf{(3)}$$

Where ((x, y) =CS ordering) means that the ordering (x, y) holds in the flow schema of the composite eservice. For example, in C001, the supports of (A, C), (A, D), and (C, D) are 55, 36, 21, respectively. Therefore, the ordering score of C001 = sup (A, C) + sup (A, D) + sup(C, D) = 55 + 36 + 21= 112. Ordering scores for other flow schema of the composite e-service are derived in the same way.

4.3.2 Mining Frequent Attributes Sets

According to these log data from table 4.2, we can use the Apriori algorithm **[2]** to discover the frequent attribute sets. Support of the attribute set is the ratio of those instances that contain all attributes in attribute set. Suppose that the required minimum support value is 25%, attribute sets with support values greater than minimum support value are called frequent attribute sets. Table 4.4 shows the example of the frequent attribute sets and their corresponding support values.

Table 4.4 Frequent Attribute Sets

Frequent Attribute Sets	**Support**
Course Instructor= Mike	55%
Course Level= Beginner	66%
Course Location= Taipei	36%
Course Provider=PC-School	21%
Course Time schedule=Morning	23%

The system gives each composite e-service a attribute score. In this example, we compute the attribute score of each flow schema of composite e-service that includes A, B, C, and D. To consider the constraints of customers, the system may not only select the flow schemas which include A, B, C, and D, but also select the ones whose attributes that meet the customer's constraints. The attribute score of a composite e-service is the summation of the attribute scores of its basic eservices, as in the Equation 4.

$$\textit{Attributes score of composite e-service} = \sum_{e \in CS} \sum_{p \in e.attribute} \sup(p) \quad \textbf{(4)}$$

Where *e* denotes a basic e-service in a composite eservice CS; *p* represents an attribute in e-service. Currently, the attribute scores only consider the support values of attribute of e-services. Each e-service has default attributes in a flow schema definition. For example, in C001, the supports of (Course Instructor= Mike), (Course Level= Beginner), (Course Location= Taipei), (Course Provider=PC-School), and (Course Time schedule=Morning) are 55, 66, 36, 21, 23 respectively. Therefore, the attribute score of C001 = sup(Course Instructor= Mike) + sup(Course Level= Beginner) + sup(Course Location= Taipei) + sup(Course Provider=PC-School) + sup(Course Time schedule=Morning) = 55 + 66 + 36 + 21 + 23 = 112. Attribute scores for other flow schemas of the composite e-service are derived in the same way.

4.3.3 Total Scores of Composite E-Service

A total score can be derived by summing up the frequent ordering score and the frequent attribute score. The formula is shown in Equation 5.

Total score = (frequent ordering score + frequent attributes score) (5)

The system ranks extracted composite e-service based on the total score. And the system recommends the top N flows of composite e-services to the users. The top three recommended composite e-services will be shown in the system.

4.4 Personalized Recommendation Module

To generate the complete personalized recommendation list for each user, this module generates recommendation list according to the recommendation scoring calculated from previous modules. Recommendations only take the top N in the recommendation list.

By combining the recommendations for composite e-service flows and basic services to form a complete personalized recommendations, users will not only getting recommendations on composite e-service flow definition, it also receives personalized basic service recommendations.

4.5 Middleware Module

In order to successfully implement the composite e-service platform environment, middleware module is an essential key. Middleware acts like a "glue" which connects the knowledge map building system and personalized recommender system together into a complete composite e-services platform.

While the user is browsing in the knowledge map navigator, middleware module delivers personalized decision support recommendations on composite e-services as to provide value added features to the user when they are using the composite e-services. While middleware module process been activated, user identification and service category identification is performed and then looks up the recommendation list in the database to extract a set of recommended composite e-service flows for the target user. It then transforms the selected set of recommendation list into a dynamic webpage (such as JSP or ASP etc) for user to view the knowledge maps navigator. The infrastructure of middleware module is shown in Fig 4.2.

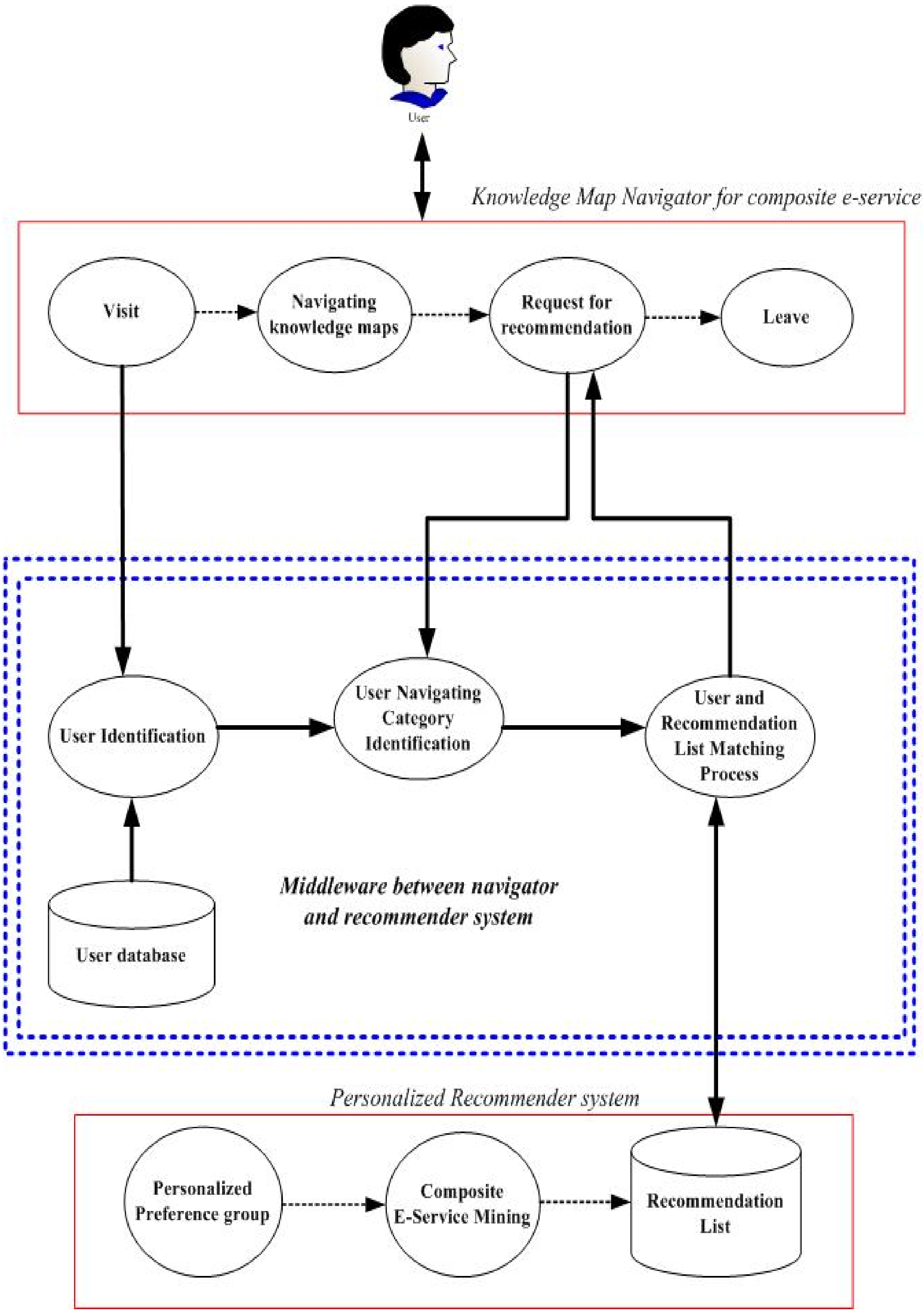


Fig. 4.2 Infrastructure of Middleware Module

Middleware module consists of three online processes as to deliver personalized recommendations to user while using the knowledge maps navigator:

1. User Identification

When user is using the composite e-service navigator, the user is required to login with the email and the password to identify user identity. When the middleware receives user's login information, it will confirm the information with the user database and save the information into the cookie if the user information matches with the user database's data.

2. User Navigating Category Identification

Before user gets a set of personalized recommendations for composite e-service, the system needs to identify which category in the e-service clusters that the user plans to view, so personalized recommender system can provide the correct type of recommendations according to the e-service cluster.

3. User and Recommendation List Matching Process

The personalized recommendations will be activated when both the user and the category of e-service cluster are identified. Matching process will match the recommendation database with user identification and the e-service category type. If matching process finds the match data in the database, it will extract a set of recommendations for both composite and basic e-services to user to view.

5. System Implementation and Demonstration

In order to illustrate how the proposed personalized recommendation system functions and works with the composite e-service platform, a prototype is developed using programming languages in JSP and SQL. The development tools include Borland J-Builder and Microsoft FrontPage. Web server is setup on Apache Tomcat 5 and database is on Microsoft SQL server 2000. Section 5.2.1 describes all the interfaces for the backend personalized recommendation generation, and section 5.2.2 introduces the interface for online personalized recommendations on composite e-service navigator.

5.1 Personalized Recommendation System Demonstration

This section is based on the system infrastructure described in chapter four. The system demonstration consists of two main parts: the personalized recommendations generation procedure which is shown in section 5.2.1 and the procedure of performing composite e-service recommendations to user on the composite e-service navigator which is shown in section 5.2.2. In this prototype system, online computer courses information is used as an example in the prototype system. The starting page of the prototype system is shown in Fig 5.1.

5.1.1 Demonstration of Personalized Recommendation System

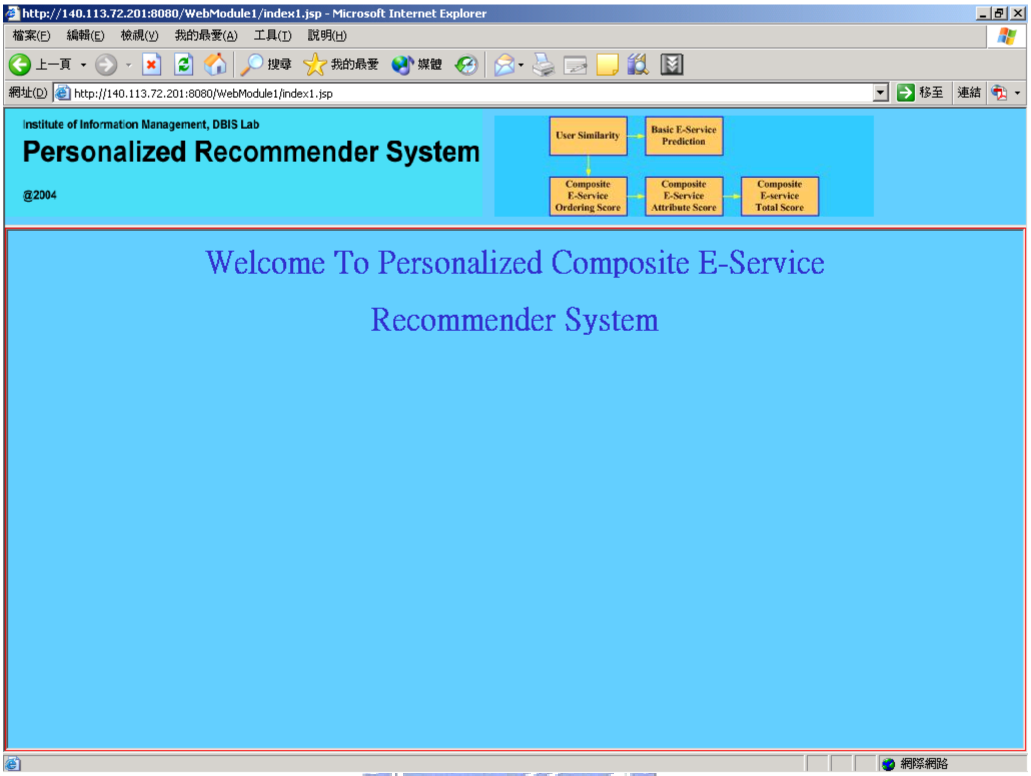


Fig. 5.1 Starting Page of the Prototype System

5.1.1.1 Feedback Page

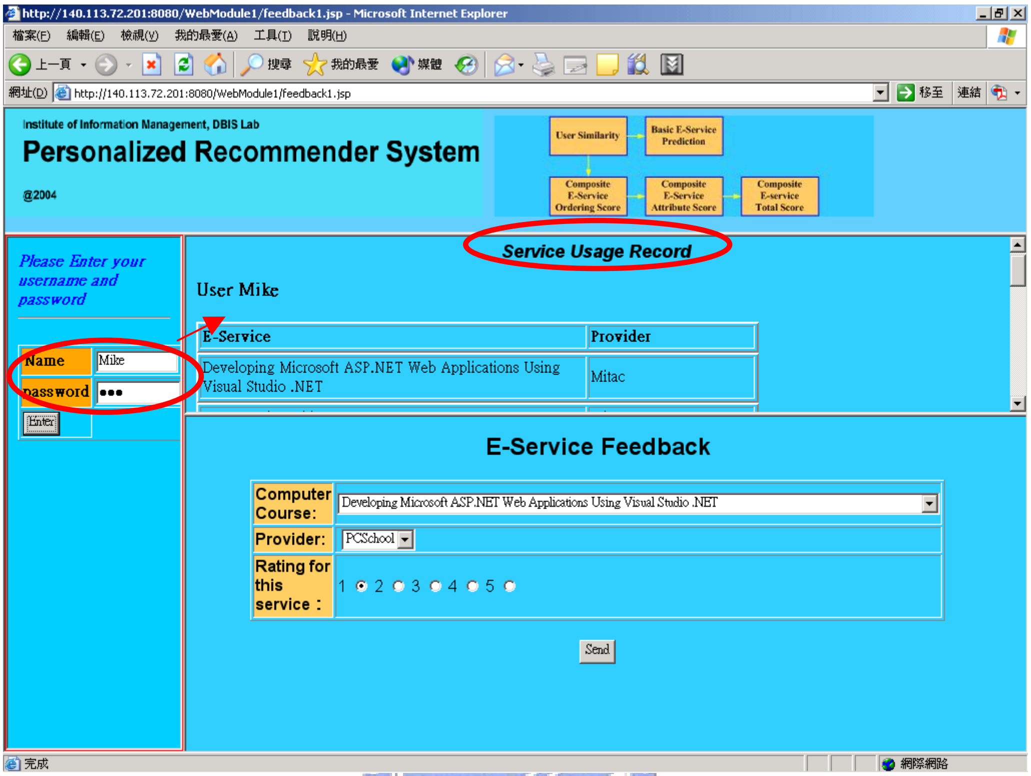


Fig. 5.2 User E-Services Feedback Page

In Fig. 5.2, before collecting feedback from users, it is required for users to login with their username and password to ensure the correct user feedback. After user has login into the feedback page, user's e-service usage records will be shown on the page to help users to memorize the e-service they have used before. In the prototype system, e-service name and provider of the e-service must be provided by the user and the each e-service rating value is from 1 to 5 and is shown from the following:

1 – Really bad e-service
2 – Bad e-service
3 – Average e-service
4 – Good e-service
5 – Excellent e-service

5.1.1.2 Users Preference Similarity

Click the hyperlink in the Fig 5.3 to enter the user preference's page. This page is used to find the similarity between all users in the database. The middle frame shows the last updated user similarity in the database and pressed the button circled in red to activate all users' similarity score shown in Fig. 5.3. And the results of renew similarity will be shown on the right side of the system, as shown in Fig.5.4. The similarity score is from 1 to -1. 1 means another user has similar preference behaviour with user and -1 mean opposite preference behaviour. While 0 means no correlation coefficient with user.

For Example, user Mike's rating value has been entered into the system for finding his preference group through the similarity score. From the result, user Lo score 0.52, user Amy's similarity is 0.51, user Tan's similarity is -0.27 and user Tsai's similarity is -0.18. From the result, we can predict that user Lo and Amy have very similar preference behaviour and will be cluster as Mike's preference group. User Tan's similarity is -0.18 means he has opposite preference behaviour against mike's preference

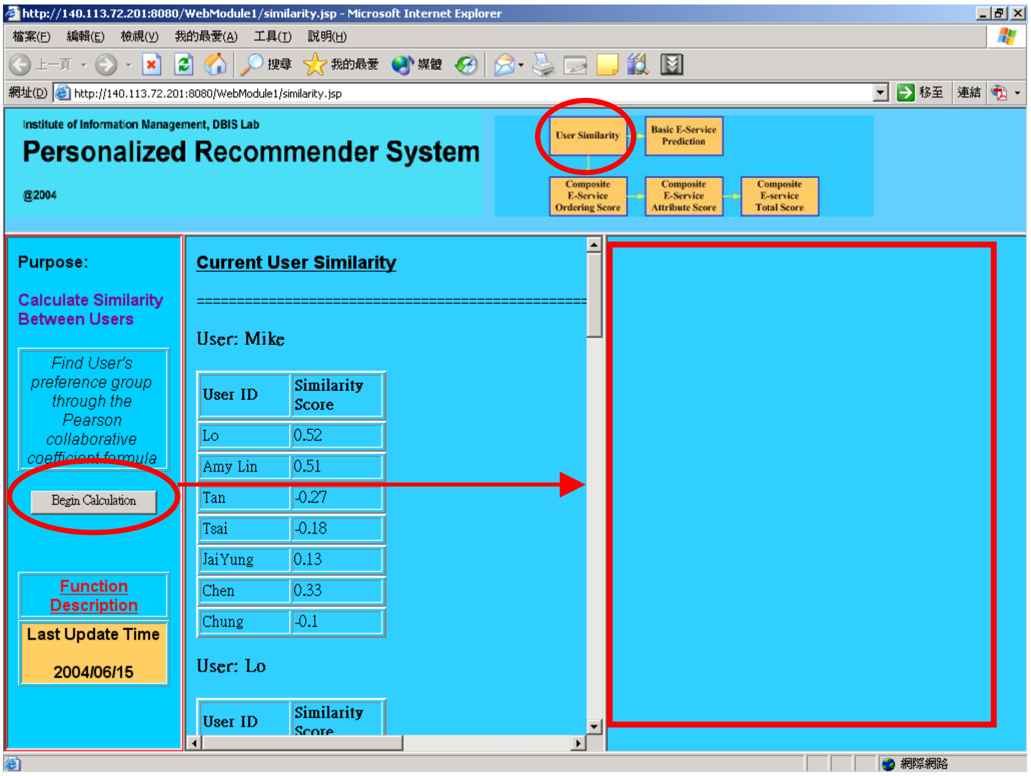


Fig. 5.3 Users Preference Similarity Page

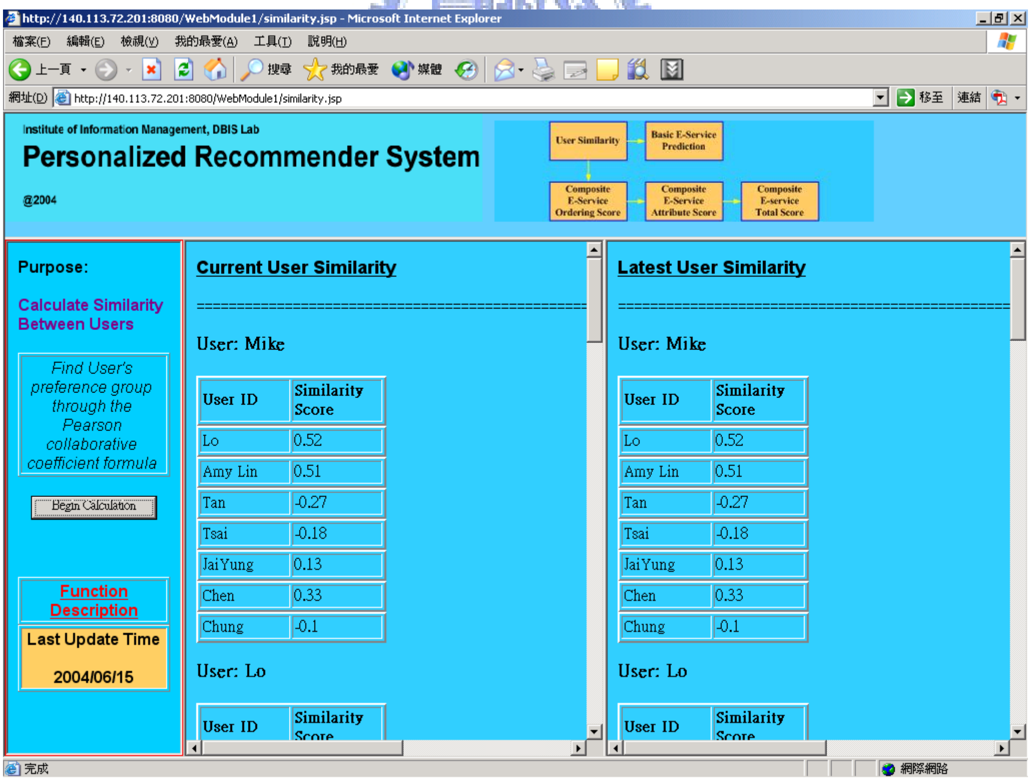


Fig. 5.4 Comparison between Latest and Current User Preference Similarity

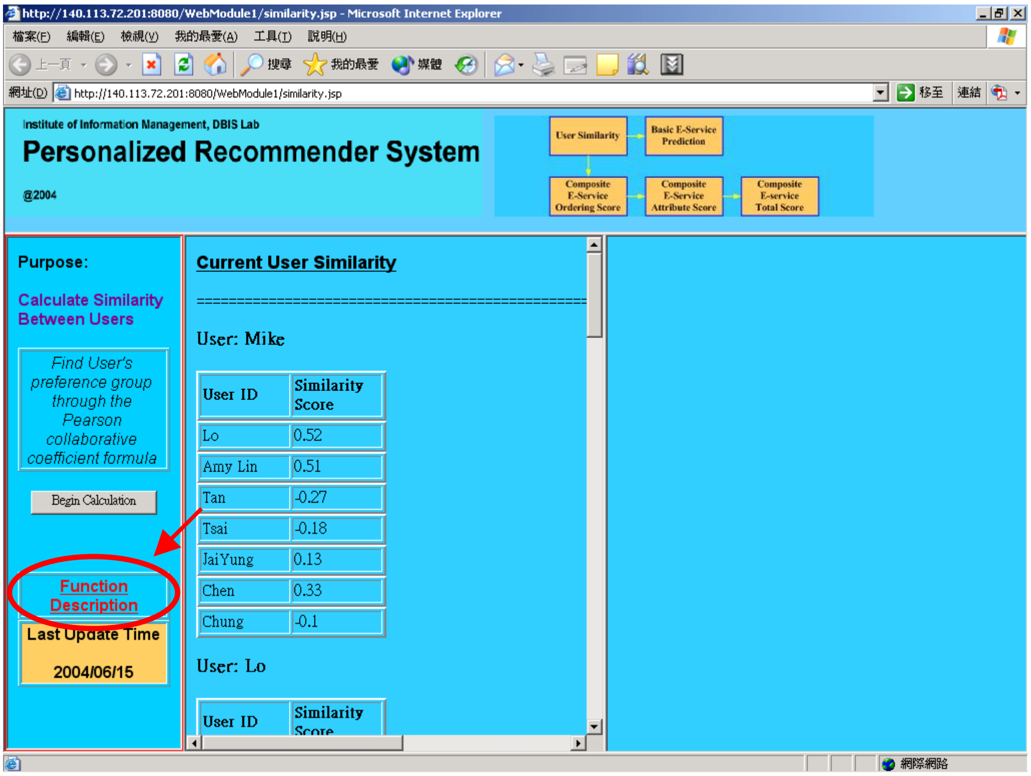


Fig. 5.5a Functional Description Link

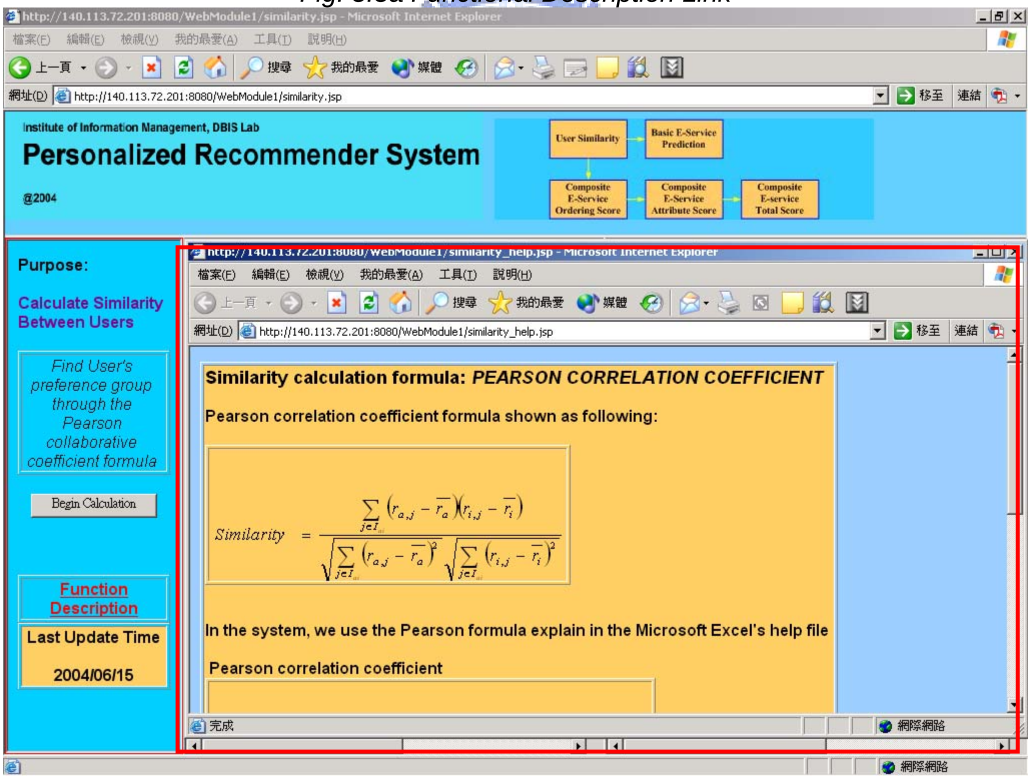


Fig. 5.5b Functional Description for the User Preference Similarity

Function description circled in Fig. 5.5 shows a help page which explains the formula used to calculate similarity between users.

5.1.1.3 User's Unused E-Services Preference Prediction

Through the similarity score, each user is clustered with his/her own preference group, the top 3 most correlated users are taken from preference group to predict the user's preference toward unused e-services. In Fig 5.6, the button on the left activates the new prediction calculation.

In Fig. 5.6, click the hyperlink on the top of the screen and enter into the basic service prediction page. In the middle frame, shows the last updated basic services prediction score for each user in the database. Press the button on the left to activate basic service score calculation per each basic service. And the results of renew service prediction score will be shown on the right side of the system screen as shown in Fig.5.7. The service prediction score is a score from 0 to 100. The higher the score means the higher interest for user to accept the e-service.

For example, the top 3 most correlated users in the user preference group for user Mike are Song, Chen and Tsai. And with the calculation it shows that user Mike will have a strong interest on e-service course of XML with a prediction score of 75. And this e-service can be recommended to user Mike as first recommendation for basic service.

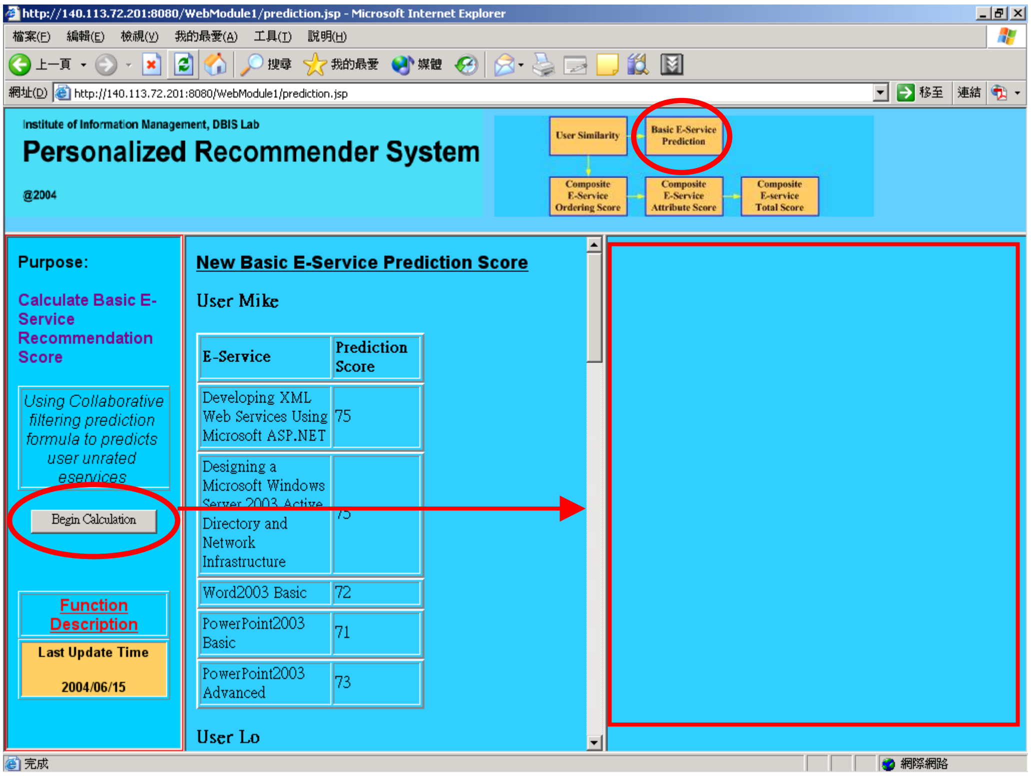


Fig. 5.6 User Unused E-Services Preference Prediction Page

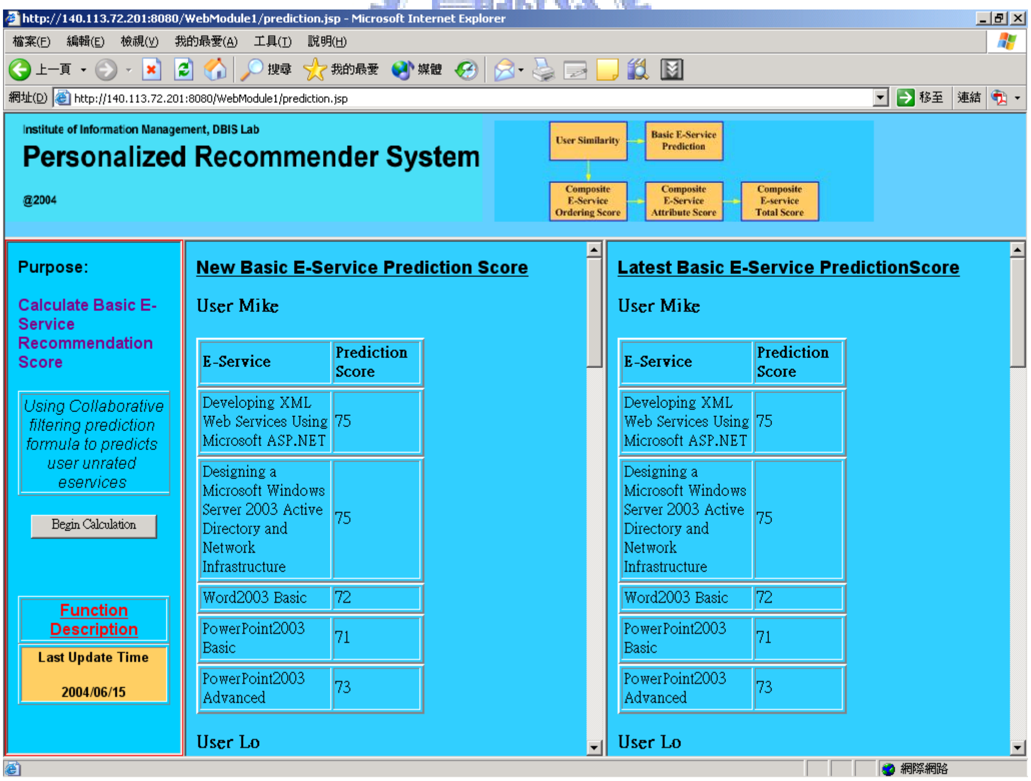


Fig. 5.7 E-service Preference Prediction Comparison

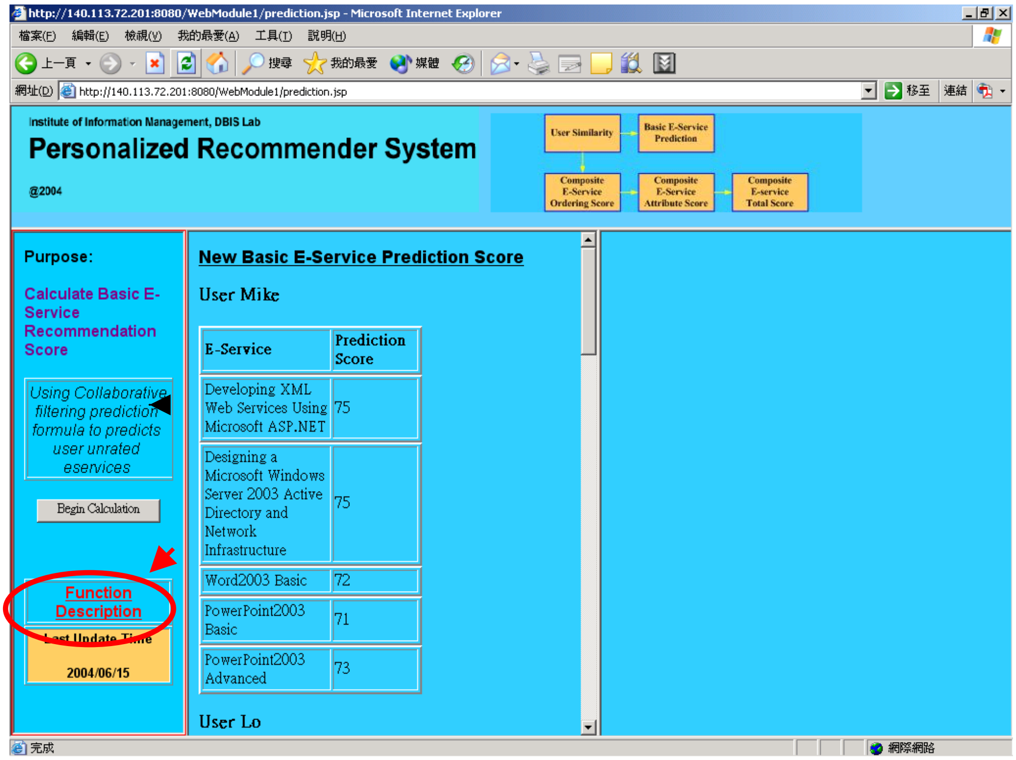


Fig. 5.8a Functional Description Link

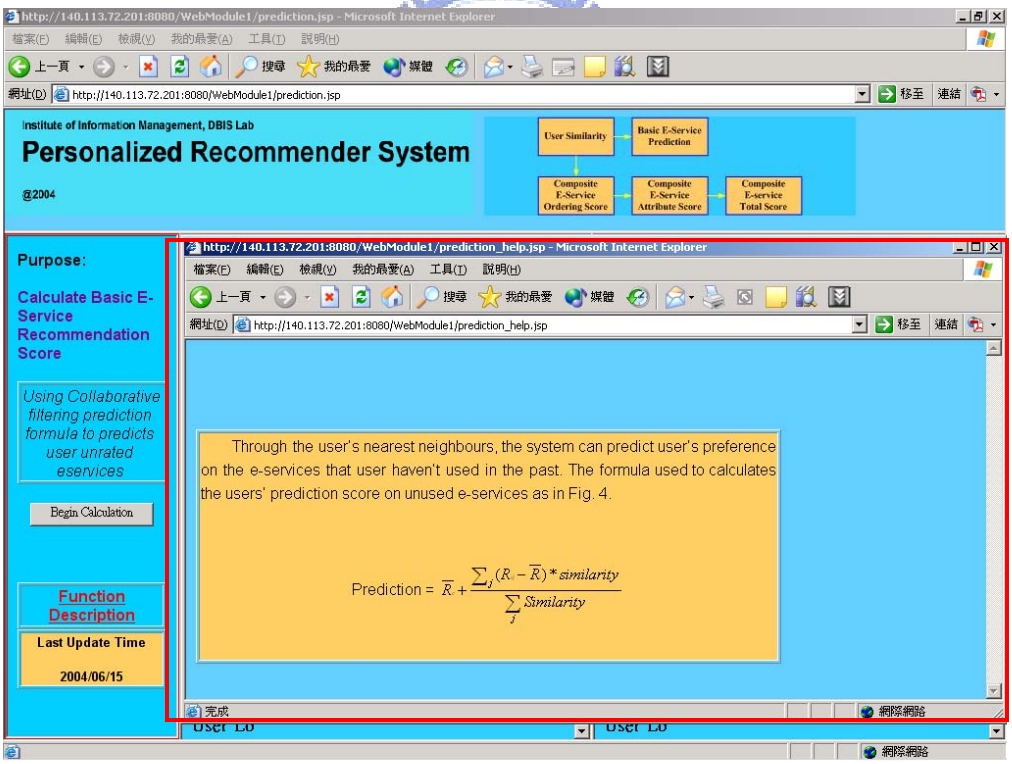


Fig. 5.8b Functional Description in User Preference Prediction Page

Function description in Fig. 5.8 shows a help page which explains the formula used to predict the e-service that is not used by the user before.

5.1.1.4 Flow Ordering Score for Composite E-Service

The composite e-service flow ordering scoring approach uses association rule mining to find the large frequent 1 support value as the basis for the flow ordering score calculation in each composite e-service. As mentioned in chapter four, e-services ordering are transformed as (A, B). For example, user Mike took a composite e-service flow:

VB => ASP => C#.

The flow ordering for the composite e-service consists ordering of (VB, ASP) and (ASP, C#). And two orderings are candidates in the association rule mining.

For example, the top 3 most correlated users in the user preference group for user Mike are Song, Chen and Tsai. By collecting their usage records for composite e-service, the association rule mining then finds the ordering set in the usage record and can gain support value for each ordering to score up all composite e-service flow in the database.

In Fig. 5.9, click the hyperlink on the top of the screen and enter into the ordering score calculation page. On the middle of screen, it shows the last updated composite e-service flow ordering scores for each user in the database. Press the button on the left to activate ordering calculation for each composite e-service. And the results of renew flow ordering score will be shown on the right side of the system screen as shown in Fig.5.10.

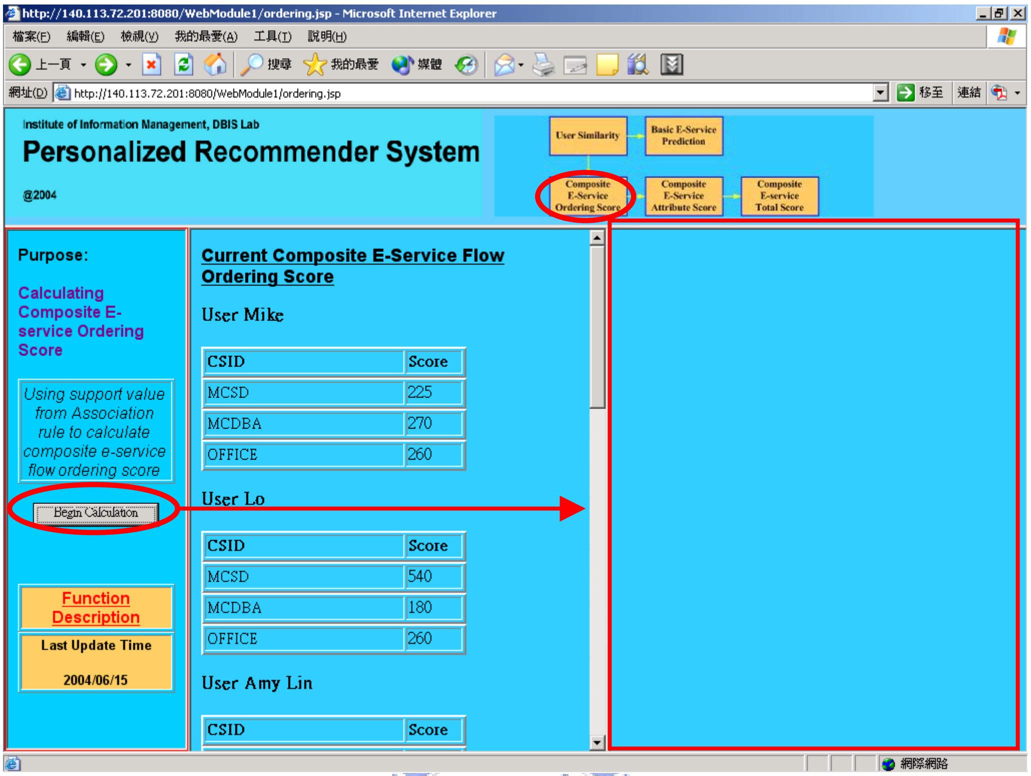


Fig. 5.9 Composite E-Service Flow Ordering Calculation Page

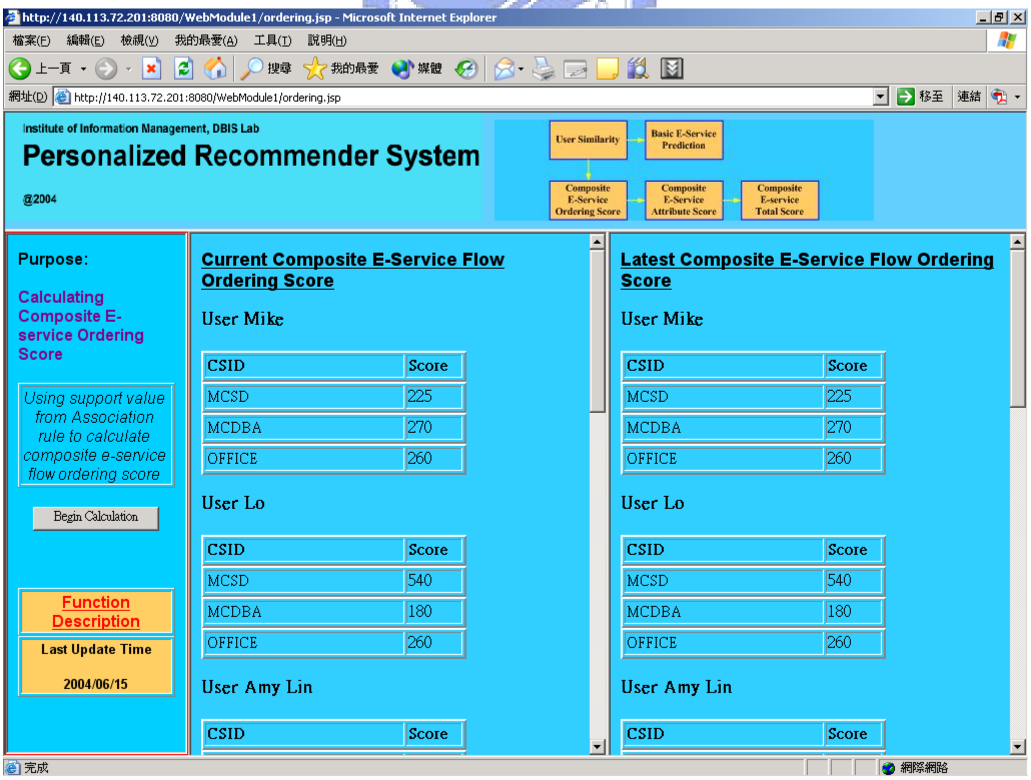


Fig. 5.10 Comparison between Current and Latest Ordering Recommendation Score

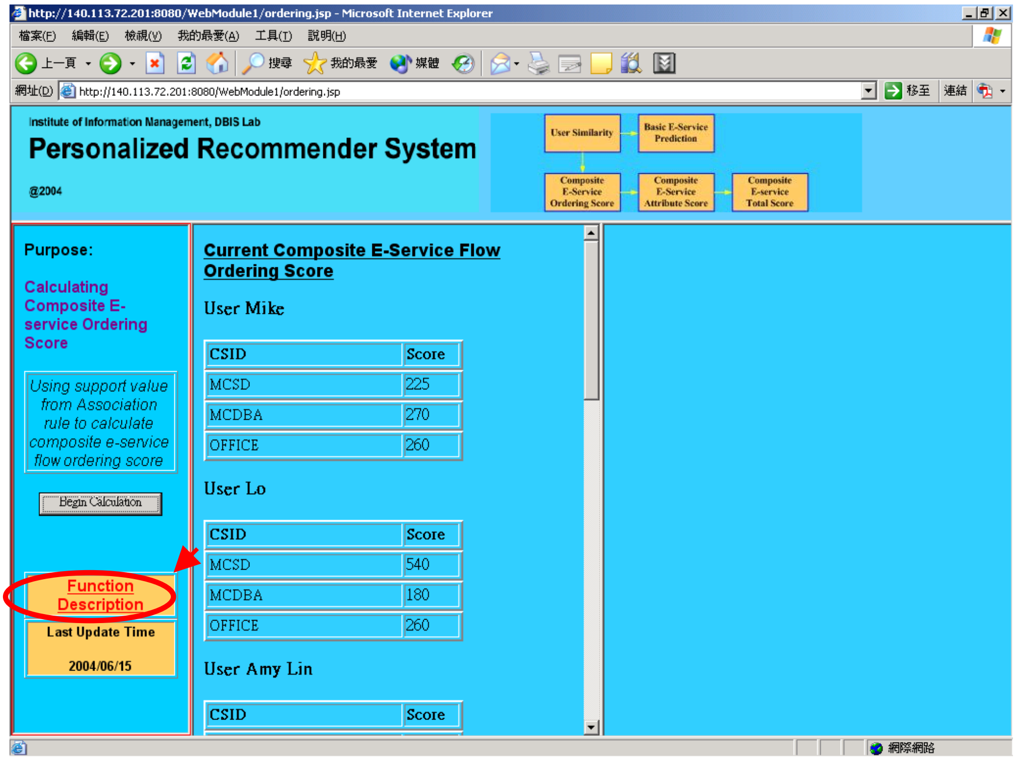


Fig. 5.11a Functional Description Link

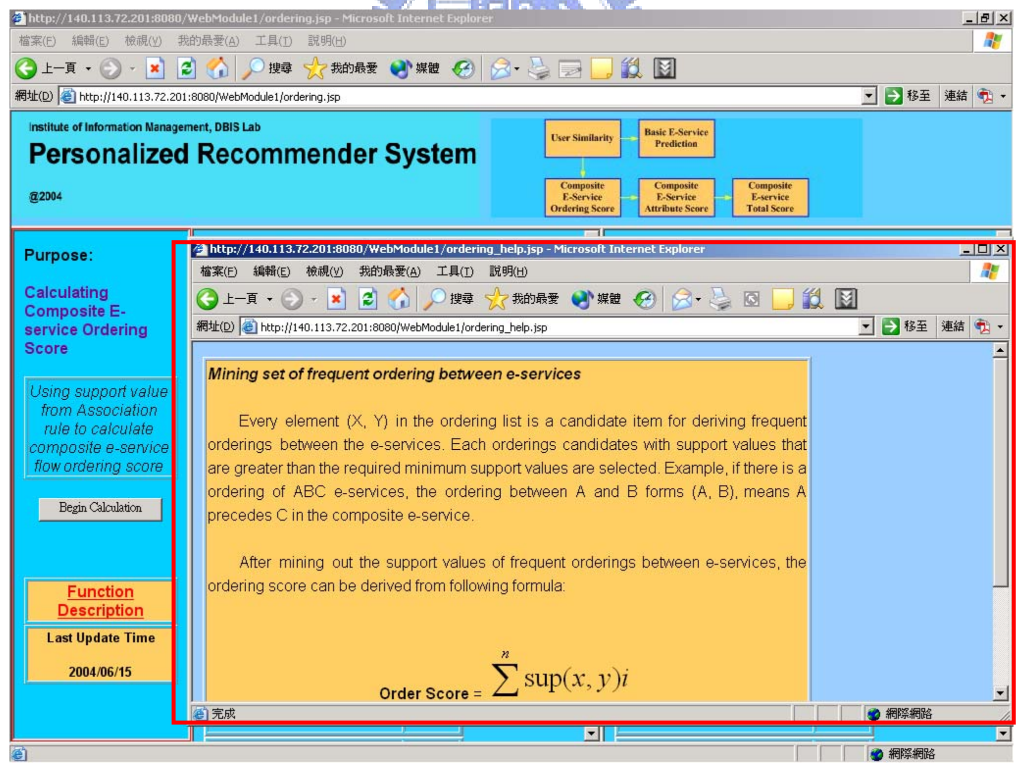


Fig. 5.11b Functional Description in Composite E-Service Ordering Score Page

Function description in Fig. 5.11 shows a help page which explains the formula used to calculate ordering score for each composite e-service.

5.1.1.5 Attributes Score for Composite E-Service

Each composite e-service consists of several basic e-services. Each basic service consists of its own attributes. In this prototype system, online e-learning computer courses are set as the example of basic e-service. Each computer course consists of five attribute types: time, instructor, provider, course level, and location. The composite e-service flow attributes scoring approach uses association rule mining to find the large frequent 1 support value for the attributes as the basis for the flow ordering score calculation in each composite e-service. These attributes are the candidates for the association rule mining. For example, user Mike took a composite e-service flow: *VB => ASP => C#.*

Each basic service consists of specific attributes:

VB: location=Taipei, time=Morning, Instructor=Nancy, provider=PC-school, level=beginner

ASP: location=Taipei, time=Afternoon, Instructor=Kit, provider=UMC, level=beginner

C#: location=Hsinchu, time=Evening, Instructor=Tom, provider=UMC, level=Advance

The association rules then mine the attributes set in the usage record and gain support value for each ordering to score up all composite e-service flow in the database.

In Fig. 5.12, click the hyperlink on the top of the screen and enter into the attribute score calculation page. In the middle of screen, it shows the last updated composite e-service attribute scores in the database. Press the button on the left to activate attribute score calculation for each composite e-service. And the results of renew attribute score will be shown on the right side of the system screen as shown in Fig. 5.13.

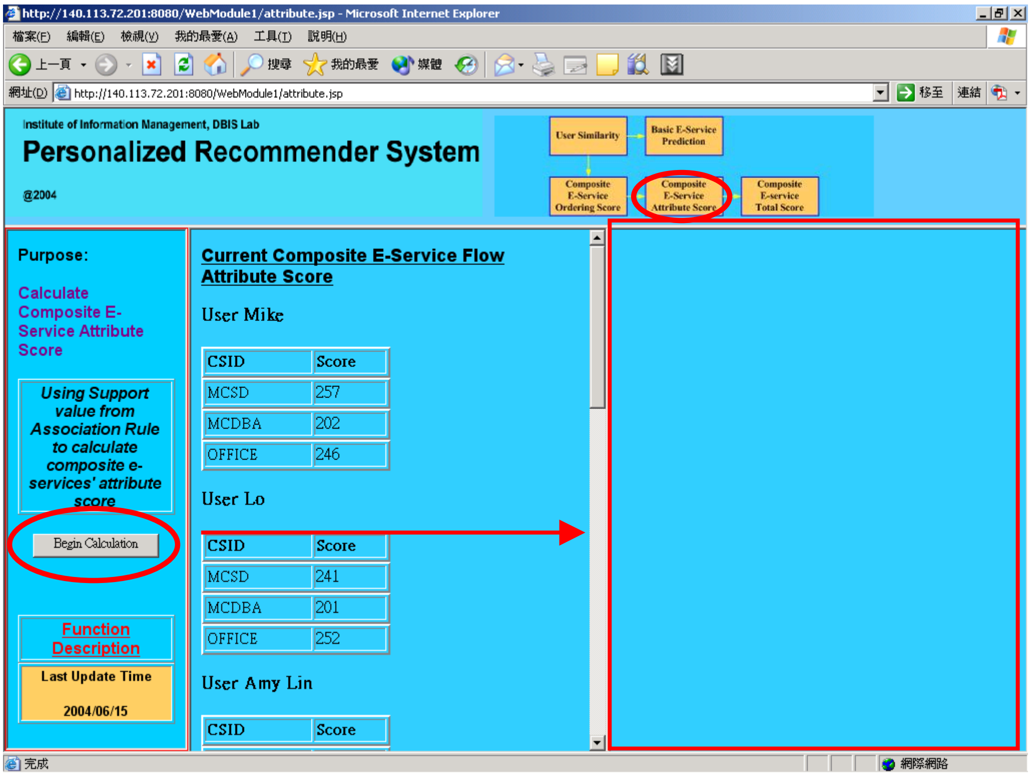


Fig. 5.12 Composite E-Service Attribute Calculation Page

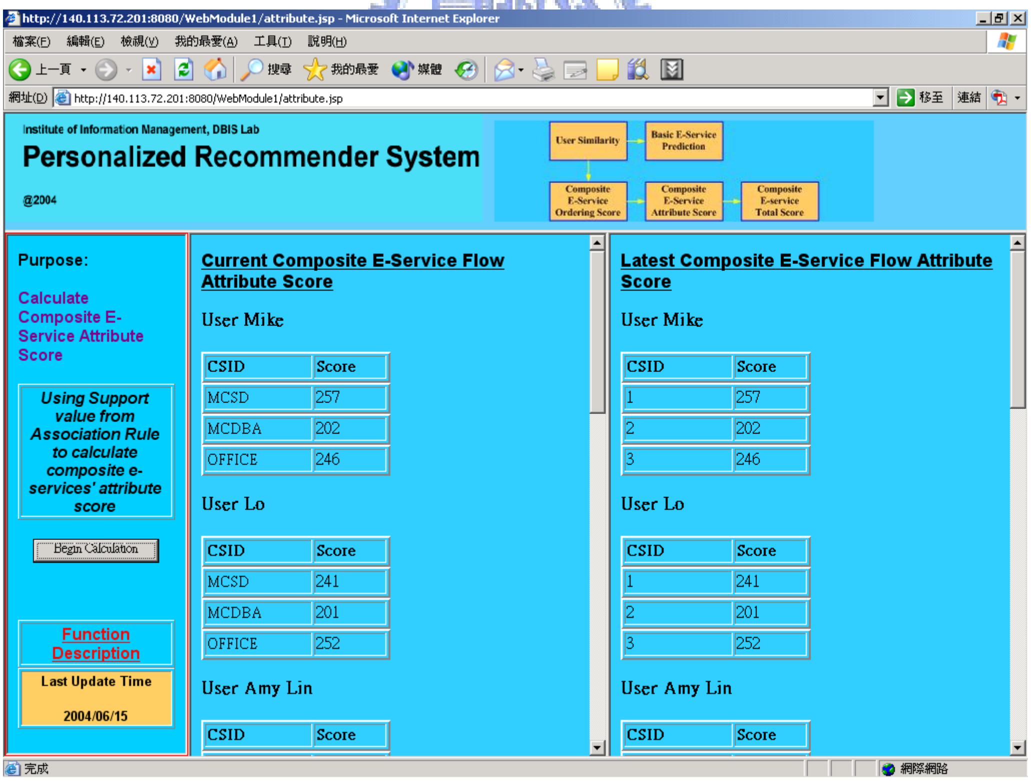


Fig. 5.13 Comparison between Latest and Current Attribute Scores

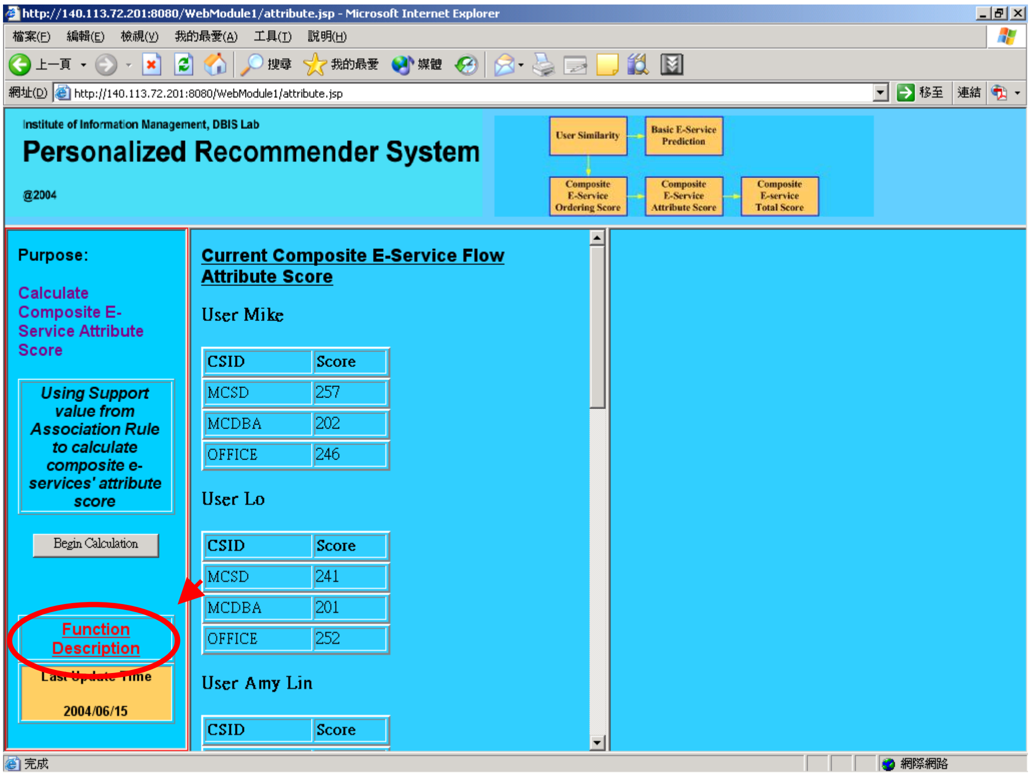


Fig. 5.14a Functional Description Link

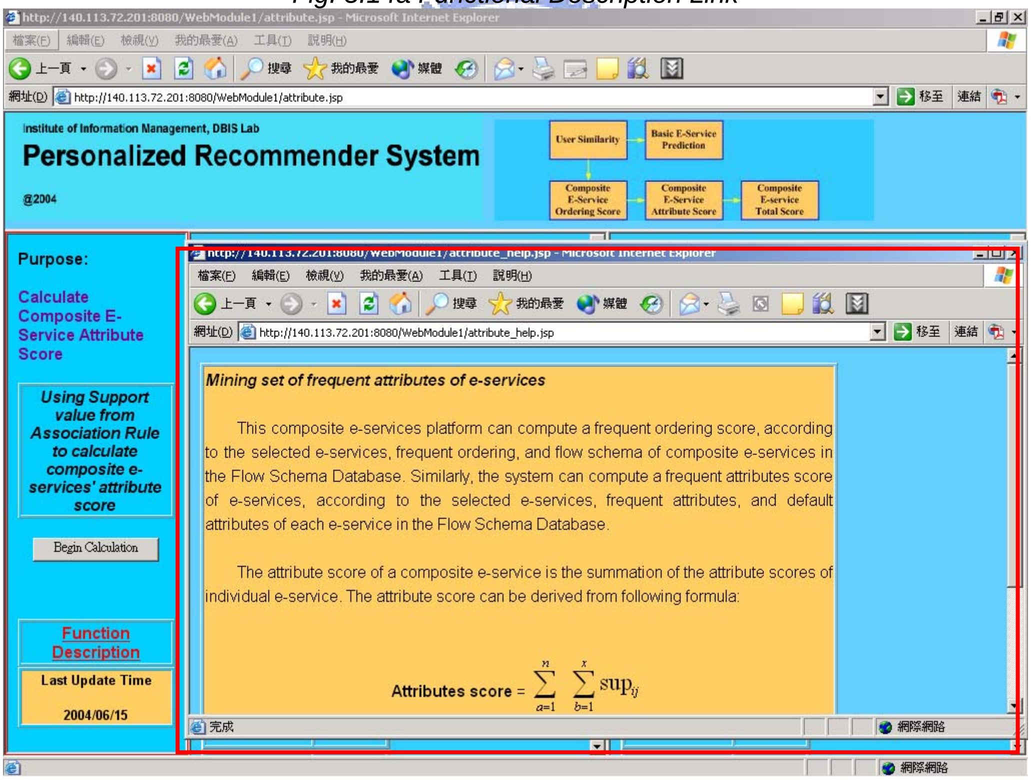


Fig. 5.14b Functional Description in Attributes Calculation Page

Function description in Fig. 5.14 shows a help page which explains the formula used to calculate attribute score for each composite e-service.

5.1.1.5 Composite E-Service Recommendation Score

Composite e-service's recommendation score is derived by summing up the frequent predicate score and the frequent ordering score. The system performs composite e-service flow recommendations to user by taking the top 3 of high recommendation scores in the composite e-service list.

For example, after both flow ordering score (125) and attributes score (250) for composite e-service (VB => ASP => C#) have been calculated. The recommendation score for this composite e-service flow is 125+250= 375. User Mike's recommendation score toward to this composite e-service is 375.

In Fig. 5.15, click the hyperlink on the top of the screen and enter the recommendation score calculation page. In the middle of screen, shows the last updated composite e-service recommendation scores for each composite e-service flow in the database. Press the button on the left to activate recommendation score calculation for each composite e-service. And the results of renew recommendation scores will be shown on the right side of the system screen as shown in Fig. 5.16.

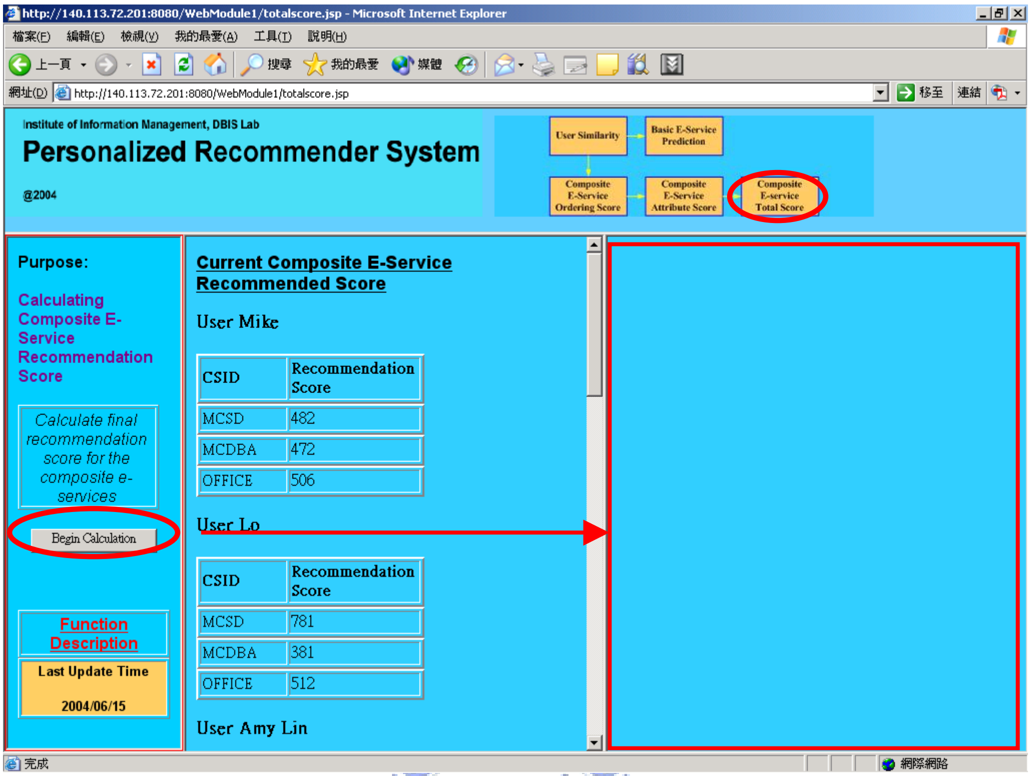


Fig. 5.15 Composite E-Service Recommendation Calculation Page

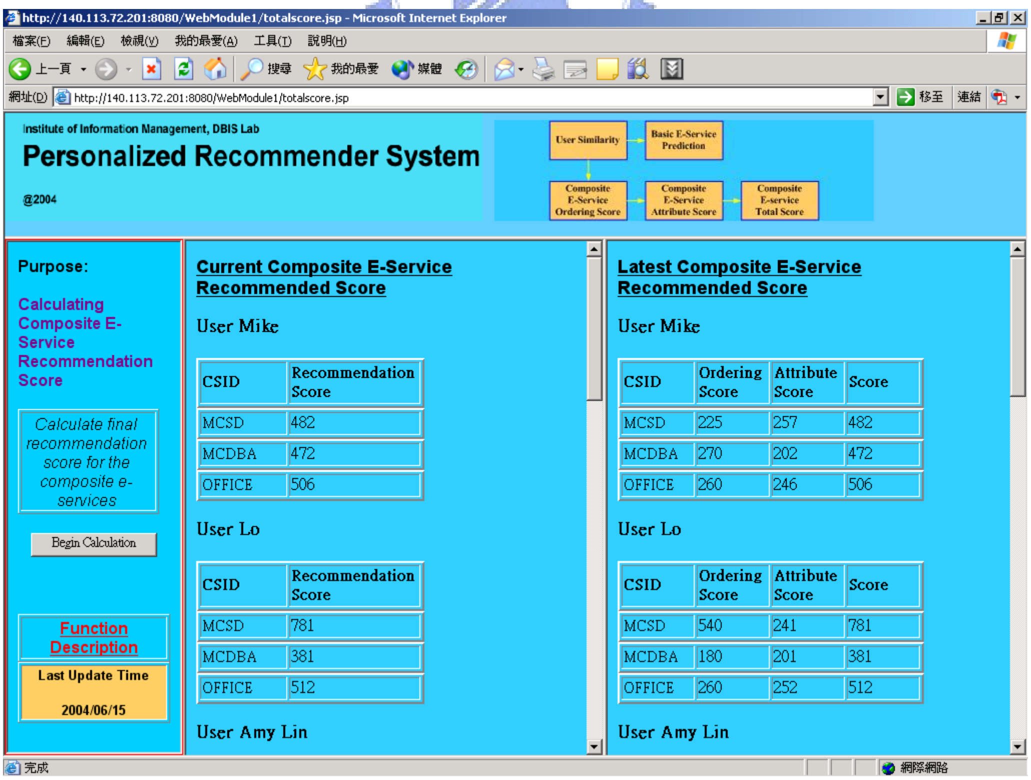


Fig. 5.16 Comparison between Current and Latest Recommendation Scores

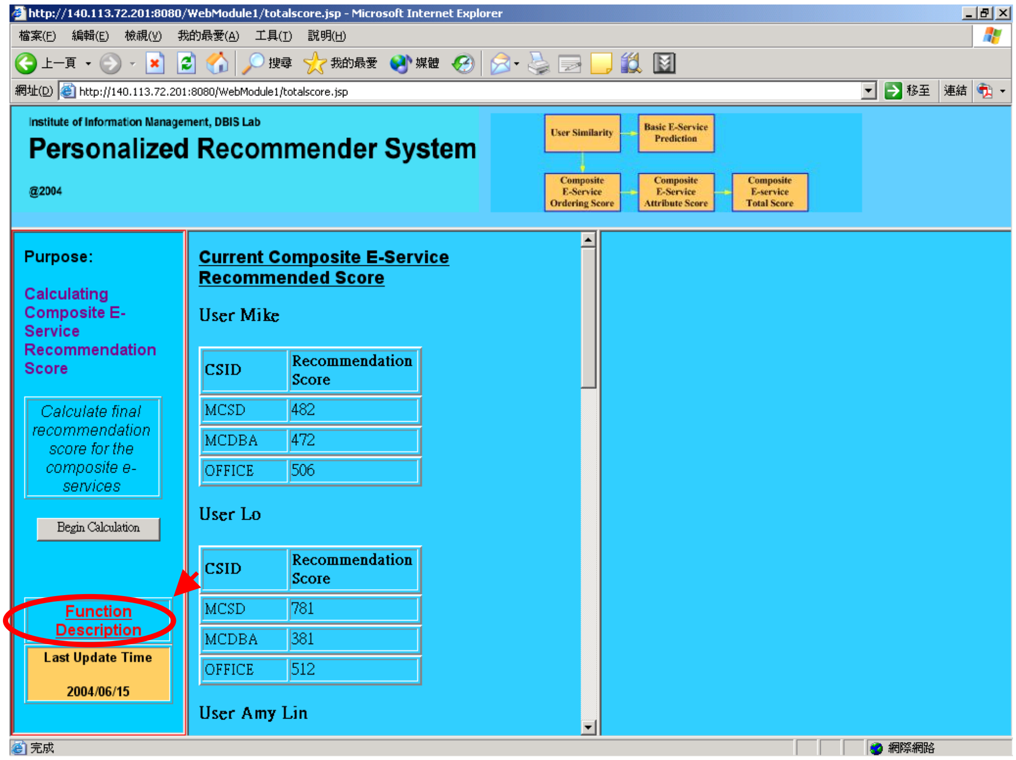


Fig. 5.17a Functional Description Link

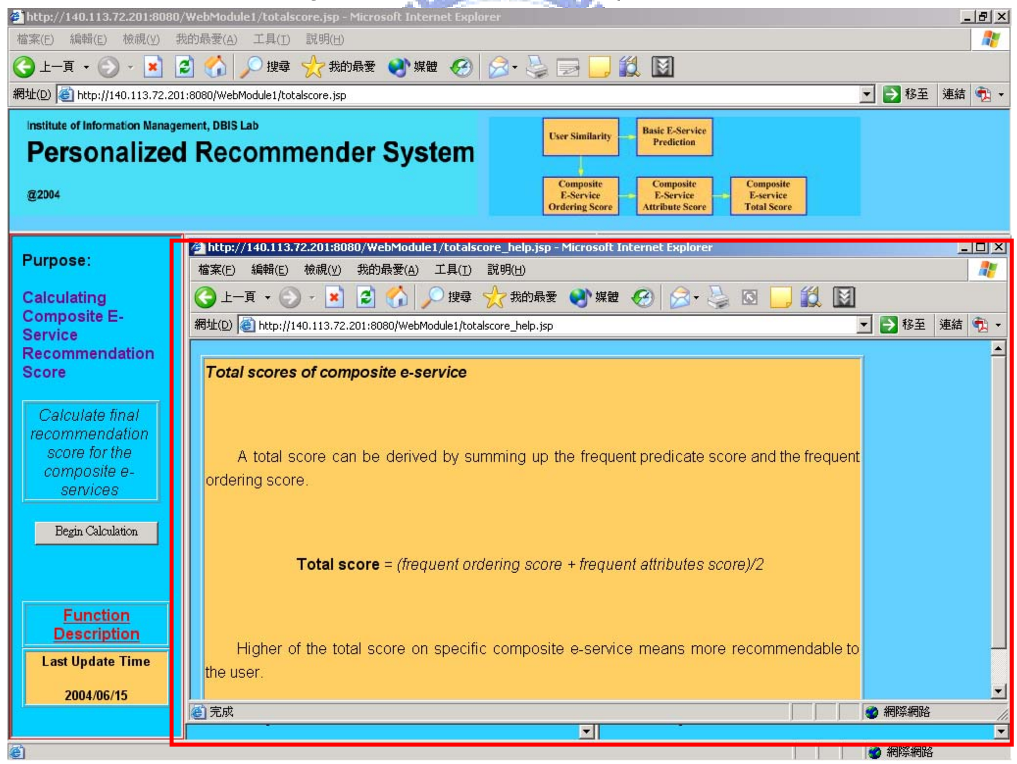


Fig. 5.17b Functional Description in Total Score Calculation Page

Function description in Fig. 5.17 shows a help page which explains the formula used to calculate recommendation score for composite e-service.

5.1.2 Perform Recommendations on the Composite E-Service Platform

In order to perform recommendations on the composite e-service navigator successfully, the middleware module acts as a bridge between the personalized recommender system and composite e-service navigator. The following is the demonstration for the middleware's three steps of the recommendations.

5.1.2.1 User Identification

When user is surfing in the composite e-service navigator, if he/she sends a request for a recommendation, the first step is to perform the login process. This step is used to identify user and save the information into cookie. Fig. 5.18 demonstrates the login process. User Mike has filled in his email address and the password to login to the composite e-service navigator.

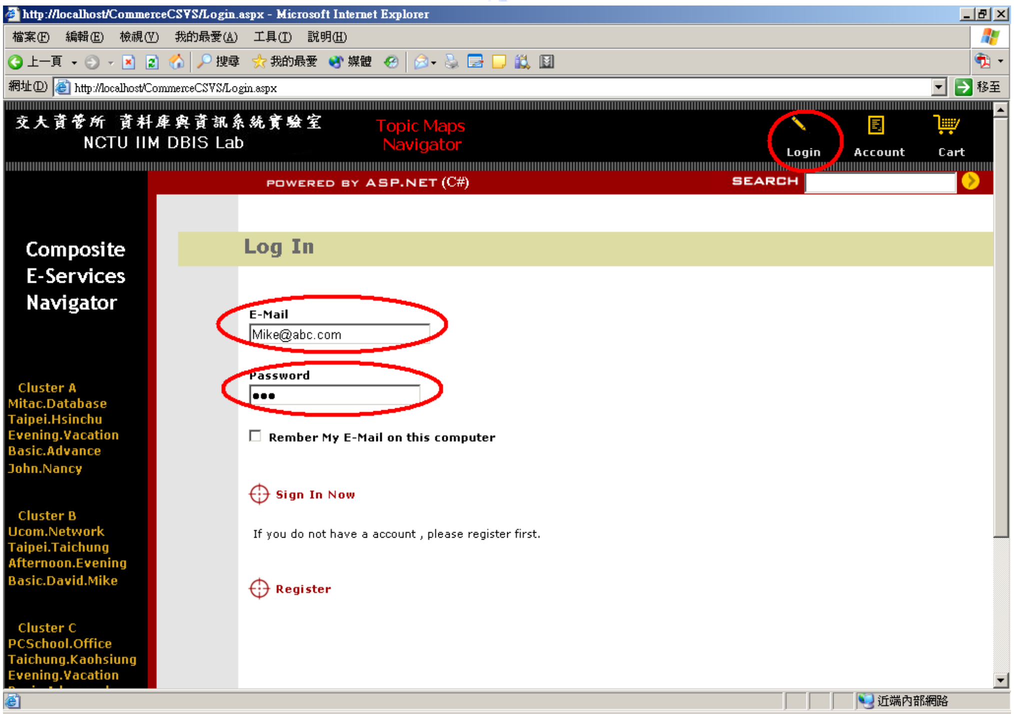


Fig. 5.18 User Login Page for Personalized Request

5.1.2.2 User Navigating Category Identification

When user Mike logs in the composite e-service navigator, before recommendations is performed, the system needs to identify the category of service cluster that the user plan to view. In Fig. 5.19, user Mike can choose one of the knowledge map clusters for the e-service to get his personalized recommendations and user Mike decides to click on a cluster type of e-service (e.g. cluster A) for his personalized recommendations.

5.1.2.3 Matching Process between User and Recommendation List

In Fig 5.20, after user Mike clicked on cluster A, a personalized recommendations list will pop up in the composite e-service navigator screen to suggest the kind of composite e-services that user Mike can take within this category. In the popup page of the recommendation list, the user can click on any e-services in the composite e-service flow and it will link user back to composite e-service navigator's e-service page for more details. Fig. 5.21 is the first composite e-service flow recommendation for user Mike to take. When user Mike clicked on the SQL course in the composite e-service flow, this will link user Mike back to the composite e-service navigator page where e-service (SQL 2000) is located, as shown in Fig. 5.22.

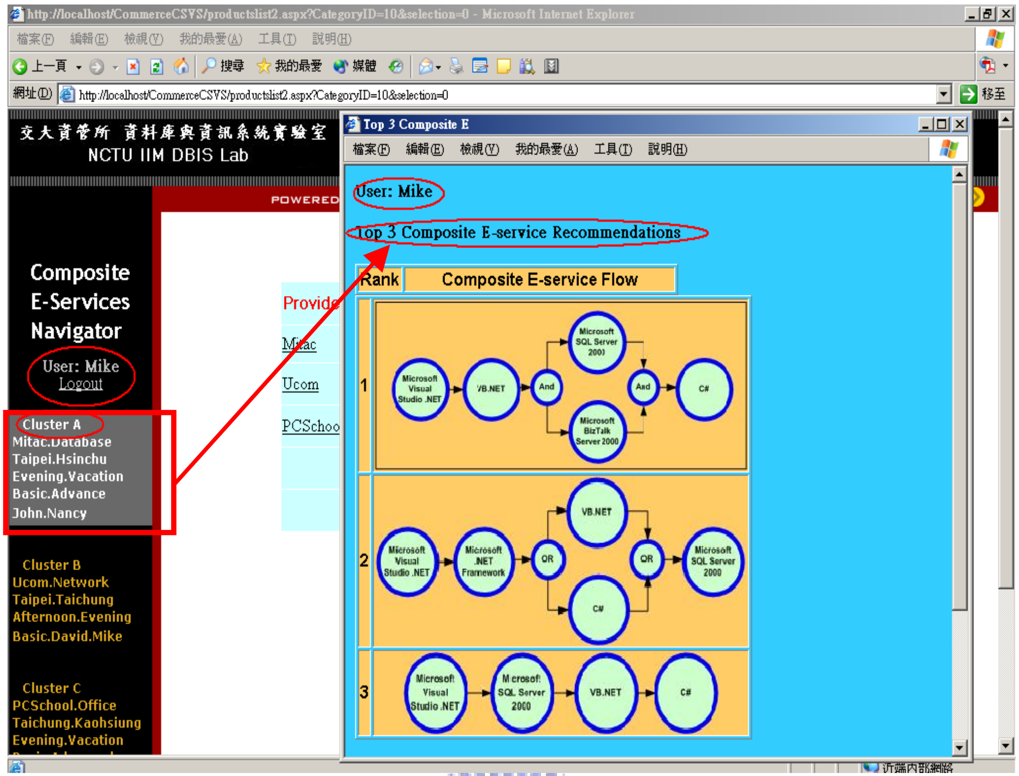


Fig. 5.19 Top 3 Personalized Composite E-Service Recommendations

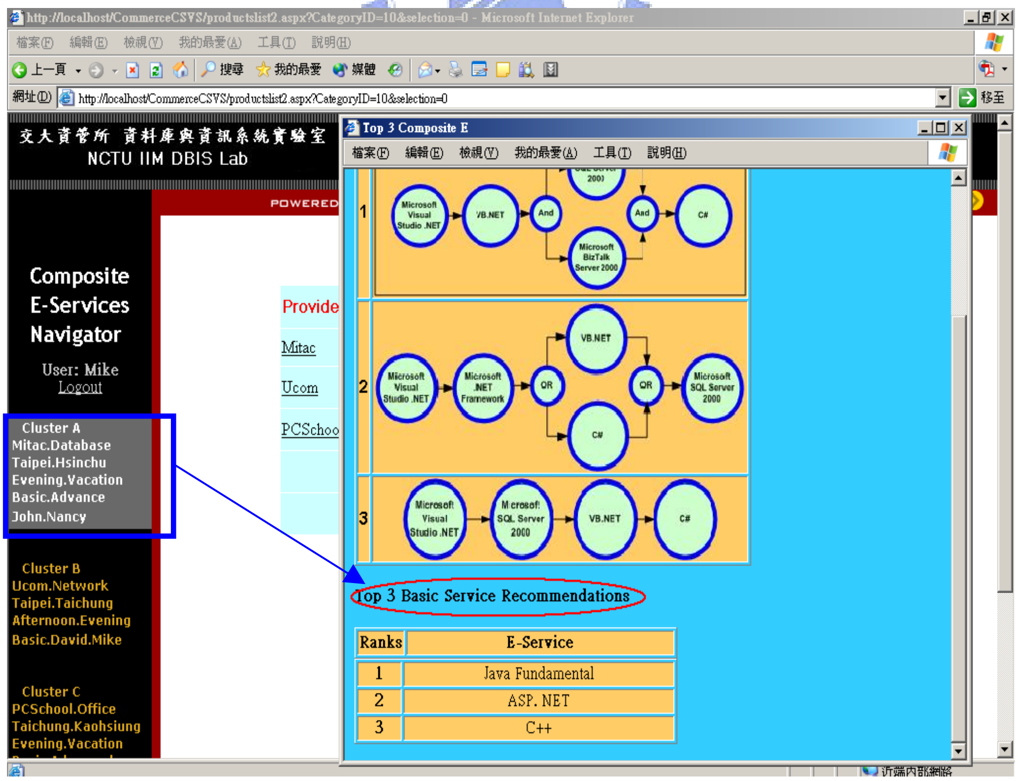


Fig. 5.20 Top 3 Personalized Basic E-Service Recommendation

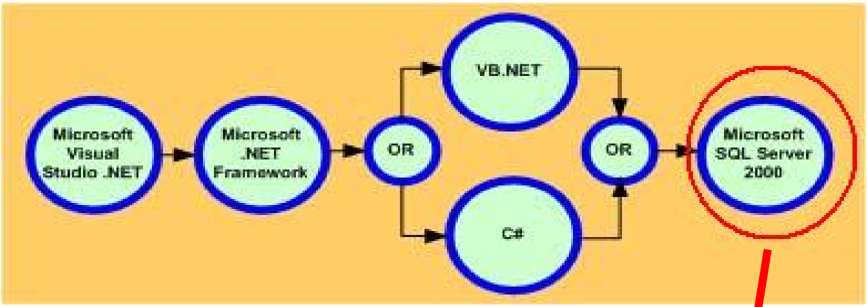


Fig. 5.21 Composite E-Service Flow Link

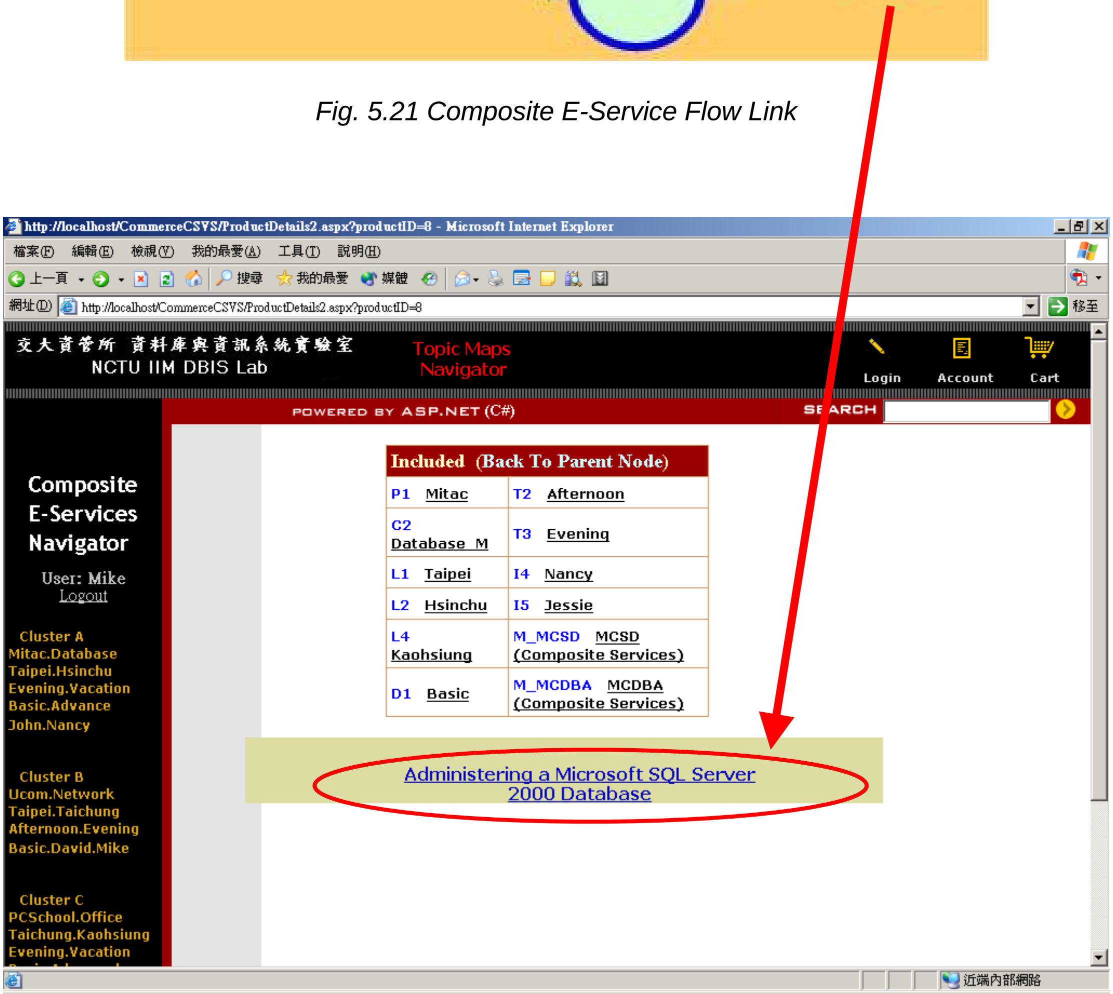


Fig. 5.22 Link from the Recommendation List Back to Composite E-Service Navigator

6. Conclusion and Future Works

6.1 Conclusions

With the rapid growth of electronic commerce, the electronic marketplace has dramatically changed, which strongly influences the customers' behaviour of demands. Providing individual e-service cannot fulfill user's various demands. So the composite e-services have become the new business solutions for enterprise to satisfy user's demands.

This research implements a composite e-service system based on knowledge map building and personalized recommendation approach. While users search the knowledge about composite e-service inside the composite e-service navigator, personalized recommendations for composite e-service is available to user for decision support to user. Our prototype system integrates collaborative filtering approach with data mining technique, which can effectively recommend personalized composite e-service to fulfill user's various demands. Therefore, this prototype system can assist e-service providers to increase the popularity of their e-services on the internet.

6.2 Future Works

Future works will be addressed in two directions:

1. This personalized recommendation system does not consider about first time users who have cold start problem. In future, the cold start problem should be explored to predict first time user with personalized recommendations.

2. This work implements a prototype of personalized recommendations for composite e-services. Further evaluation is needed to verify the effectiveness of the prototype system.

Reference

[1] R. Agrawal, D. Gunopulos, F. Leymann, “Mining Process Models from Workflow Logs”, Proc of the 6th International Conference on Extending Database Technology (EDBT'98), Valencia, Spain, March, 1998, pp. 23-27

[2] R. Agrawal, R. Srikant, “Mining Association Rules between Sets of Items in Large Databases”, Proc of ACM SIGMOD, May, 1993

[3] S. Akhil, M. Vijay, W Klaus, “Managing Next Generation E-Services”, HP Labs Technical Reports (http://www.hpl.hp.com/techreports/2000/HPL-2000-120.pdf)

[4] R. Balakrishnan, “A Services Framework Specification for Dynamic e-Services Interaction”, Proc. of the 4th International Enterprise Distributed Object Computing Conference, September 2000.

[5] J. S. Breese, D. Heckerman, and Kadie, C. (1998). Empirical Analysis of Predictive Algorithms for Collaborative Filtering. In Proceedings of the 14th Conference on Uncertainty in Artificial Intelligence, pp. 43-52.

[6] F. Casati, S. Ilnicki, L. J. Jin, V. Krishnamoorthy, and M. C. Shan, “Adaptive and Dynamic Service Composition in eFlow”, Proc. of the 12th International Conference of Advanced Information Systems Engineering, June 2000.

[7] F. Casati, S. Ilnicki, L. J. Jin, and M. C. Shan, “An Open, Flexible and Configurable System for Services Composition”, Proc. of the Second International Workshop on Advance Issues of E-Commerce and

Web-Based Information Systems, June 2000.

[8] F. Casati, M. C. Shan, "Definition, Execution, Analysis and Optimization of Composite e-Services", IEEE Data Engineering Bulletin, 24(1), March 2001.

[9] F. Casati, M. C. Shan, "Dynamic and Adaptive Composition of e-Services", Information Systems, 26(3), May 2001.

[10] M.S. Chen, J. S. Park and P. S. Yu, "Data Mining: An Overview from Database Perspective", IEEE Transactions on Knowledge and Data Engineering Vol. 8, No. 6, December 1996.

[11] D. Fensel, C. Bussler, "The Web Service Modeling Framework WSMF", Electronic Commerce Research and Applications, 1(2), 2002.

[12] B. L. Grand, M. Soto, "Visualization of the semantic Web: Topic Maps Visualization", Proceedings of the Sixth International Conference on Information Visualization 2002.

[13] J. Han, M. Kamber, Data Mining: Concepts and Techniques, Morgan Kaufmann Publixhers, USA, 2000

[14] Hewlett Packard (HP), http://www.hp.com/.

[15] International Organization for standardization (ISO), International Electro-technical Commission (IEC), "ISO/IEC 13250:2000, Topic Maps", Available http://www.iso.org.

[16] J. A. Konstan, B. N. Miller, D. Malts, J. L. Herlocker, L. R. Gordon, and J. Riedl, "GroupLens: Applying Collaborative Filtering to Usenet News", Communications of the ACM, 40(3), March 1997.

[17] C.H. Lee, Y. H. Kim, P. K. Rhee, “Web personalization expert with combining collaborative filtering and association rule mining technique”, Expert System with Application 21 (2001).

[18] Q. Li, B. M. Kim, “An Approach for combining Content-based and Collaborative filters”, IRAL2003, workshop of ACL2003, Sapporo, Japan.

[19] D-R. Liu, M. Shen, C-T. Liao, "Designing a composite e-service platform with recommendation function", Computer Standards & Interfaces 25 (2003)

[20] G. Piccinelli and S. L. Williams, “Workflow: A Language for Composing Web Services”, W.M.P. van der Aalst et al. (Eds.): BPM 2003, LNCS 2678, 2003.

[21] H.H. Rath, S. Pepper, “Topic Maps: Introduction and Allegro”, Proc of Markup Technology ’99 Conference, Pennsylvania, USA, December 7-9, 1999.

[22] P. Resnick, N. Iacovou, M. Suchak, P. Bergstrom, and J. Riedl, “GroupLens: An Open Architecture for Collaborative Filtering of Netnews”, Proc. of the Conference on Computer Supported Cooperative Work, October 1994.

[23] Badrul Sarwar, George Karypis, Joseph Konstan, and John Riedl. Item-Based Collaborative Filtering Recommendation Algorithms. *WWW10,* May 1-5, 2001, Hong Kong, ACM.

[24] Simple Object Access Protocol (SOAP) 1.2, http://www.w3.org/TR/soap/

[25] UDDI org, http://www.uddi.org/

[26] W.M.P. van der Aalst and A.J.M.M. Weijters, “Process Mining: A Research Agenda”, Computers in Industry, v.53 n.3, p.231-244, April 2004.

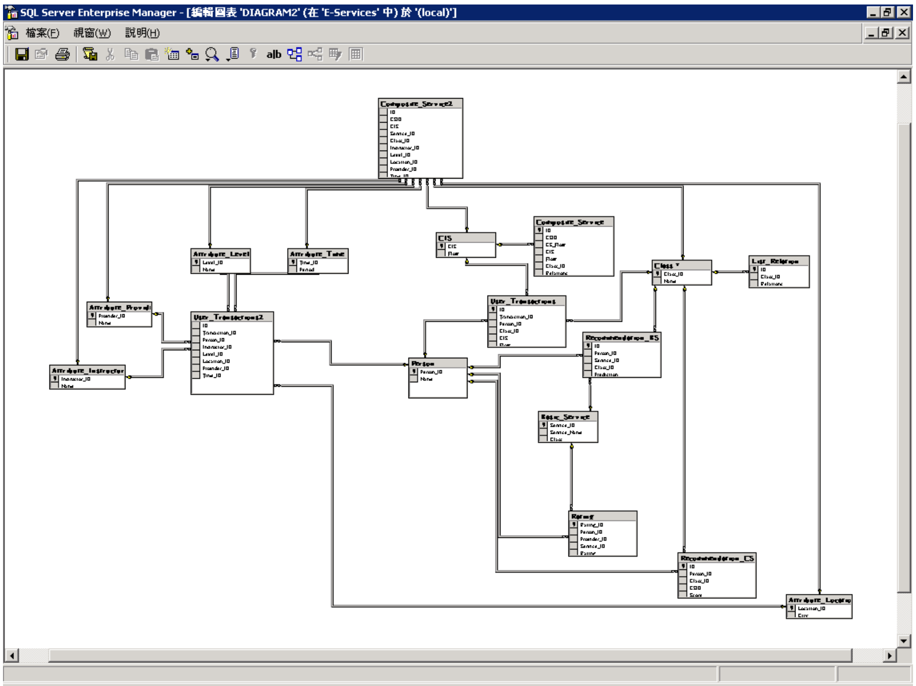
[27] Web Services Description Language (WSDL) 1.1, http://www.w3.org/TR/wsdl

[28] A.J.M.M. Weijters W.M.P van der Aalst, "Process Mining Discovering Workflow Models from Event-Based Data", Proc of the 13th Belgium-Netherlands Conference on Artificial Intelligence (BNAIC 2001), BNVKI, Maastricht, 2001, pp. 283-290

Appendix

Table Schema

This implement personalized recommendation system used the following table schema:



The whole table schema diagram for this system

Table Name: Attribute_Instructor **Purpose:** Store Composite E-Service Attributes information					
Column Name	***Type***	***Null***	***PK***	***FK***	***Explanation***
Instructor_ID	Char(10)	Y	●		Service instructor identify
Name	Char(10)	N			Service instructor name

Table Name: Attribute_Level **Purpose:** Store Composite E-Service Attributes information					
Column Name	***Type***	***Null***	***PK***	***FK***	***Explanation***
Level_ID	Char(10)	Y	●		Service difficulty identify
Name	Char(10)	N			Service difficulty level

Table Name: Attribute_Location **Purpose:** Store Composite E-Service Attributes information					
Column Name	***Type***	***Null***	***PK***	***FK***	***Explanation***
Location_ID	Char(10)	Y	●		Service location identify
City	Char(10)	N			Service location name

Table Name: Attribute_Provider **Purpose:** Store Composite E-Service Attributes information					
Column Name	***Type***	***Null***	***PK***	***FK***	***Explanation***
Provider_ID	Char(10)	Y	●		Service provider identify
Name	Char(10)	N			Service provider name

Table Name: Attribute_Time **Purpose:** Store Composite E-Service Attributes information					
Column Name	***Type***	***Null***	***PK***	***FK***	***Explanation***
Time_ID	Char(10)	Y	●		Time period identify
Period	Char(10)	N			Time period name

Table Name: Basic_Service

Purpose: Store information about type of basic service

Column Name	***Type***	***Null***	***PK***	***FK***	***Explanation***
Service_ID	Int(4)	Y	●		Basic Service identify
Service_Name	Char(10)	N			Basic service name
Class_ID	Int(4)	N		●	Service type identify

Table Name: CIS

Purpose: Store identifiers of instances of flow schema

Column Name	***Type***	***Null***	***PK***	***FK***	***Explanation***
CIS	Int(4)	Y	●		Instant flow schema identify
Flow	Char(10)	N			Flow schema order identify

Table Name: Class

Purpose: Store information of type of e-service area

Column Name	***Type***	***Null***	***PK***	***FK***	***Explanation***
Class_ID	Int(4)	Y	●		Service type identify
Name	Char(10)	N			Service type name

Table Name: Composite_Service **Purpose:** tore composite e-service's flow schema and it ordering relations					
Column Name	***Type***	***Null***	***PK***	***FK***	***Explanation***
ID	Int(4)	Y	●		Total list identify
CSID	Int(4)	N		●	Composite e-service identify
CS_Flow	Char(10)	N			Composite e-service flow schema
CIS	Int(4)	N		●	Instant flow schema identify
Flow	Char(10)	N			Flow schema order identify
Class_ID	Int(4)	N		●	Service type identify
Relations	Char(10)	N			Order sets of flow schema

Table Name: Composite_Service2 **Purpose:** Store composite e-service's flow schema and attributes					
Column Name	***Type***	***Null***	***PK***	***FK***	***Explanation***
ID	Int(4)	Y	●		Total list identify
CSID	Int(4)	N		●	Composite e-service identify
CIS	Int(4)	N		●	Instant flow schema identify
Service_ID	Int(4)	N		●	Basic Service identify
Class_ID	Int(4)	N		●	Service type identify
Instructor_ID	Char(10)	N		●	Service instructor identify
Level_ID	Char(10)	N		●	Service difficulty identify
Location_ID	Char(10)	N		●	Service location identify
Provider_ID	Char(10)	N		●	Service provider identify

Time_ID	Char(10)	N		●	Time period identify

Table Name: Person

Purpose: Store user person record

Column Name	***Type***	***Null***	***PK***	***FK***	***Explanation***
Person_ID	Int(4)	Y	●		User identify
Name	Char(10)	N			User name

Table Name: Rating

Purpose: Store user's rating record toward basic service

Column Name	*Type*	*Null*	*PK*	*FK*	*Explanation*
Rating_ID	**Int(4)**	**Y**	●		**Service rating identify**
Person_ID	**Int(4)**	**N**		●	**User identify**
Provider_ID	**Int(4)**	**N**		●	**Service provider identify**
Service_ID	**Int(4)**	**N**		●	**Basic Service identify**
Rating	**Int(4)**	**N**			**Service rating score**

Table Name: Recommendation_BS

Purpose: Store personalized recommendation list of basic service

Column Name	***Type***	***Null***	***PK***	***FK***	***Explanation***
ID	Int(4)	Y	●		Recommendation list identify
Person_ID	Int(4)	N		●	User identify
Service_ID	Int(4)	N		●	Basic Service identify
Class_ID	Int(4)	N		●	Service type identify
Predictiom	Int(4)	N			Service prediction score

Table Name: Recommendation_CS **Purpose:** Store personalized recommendation list of composite e-services					
Column Name	***Type***	***Null***	***PK***	***FK***	***Explanation***
ID	Int(4)	Y	●		Recommendation list identify
Person_ID	Int(4)	N		●	User identify
Class_ID	Int(4)	N		●	Service type identify
CSID	Int(4)	N		●	Composite e-service identify
Score	Int(4)	N			Service score

Table Name: User_Transactions **Purpose:** Store users' usage records about with flow schema					
Column Name	***Type***	***Null***	***PK***	***FK***	***Explanation***
ID	Int(4)	Y	●		Transaction total number
Transaction_ID	Int(4)	N		●	Transaction identify
Person_ID	Int(4)	N		●	User identify
Class_ID	Int(4)	N		●	Service type identify
CIS	Int(4)	N		●	Instant flow schema identify
Flow	Char(10)	N			Flow schema order identify

Table Name: User_Transactions2 **Purpose:** Store users' usage record about flow schema attributes					
Column Name	***Type***	***Null***	***PK***	***FK***	***Explanation***
ID	Int(4)	Y	●		Transaction total number
Transaction_ID	Int(4)	N		●	Transaction identify
Person_ID	Int(4)	N		●	User identify
Instructor_ID	Char(10)	N		●	Service instructor identify
Level_ID	Char(10)	N		●	Service difficulty identify
Location_ID	Char(10)	N		●	Service location identify
Provider_ID	Char(10)	N		●	Service provider identify
Time_ID	Char(10)	N		●	Time period identify