A Requirement-driven Approach for Information Integration in Regional Healthcare Environment

Xiuxiu Chen[a,b], Huiying Gao[b*], Kecheng Liu[c] and Hongqiao Yang[c]

[a] School of Management and Tourism, Fujian Agriculture and Forestry
University, China
[b] School of Management and Economics, Beijing Institute of Technology, China
[c] Informatics Research Centre, Henley Business School, University of Reading,
UK

Abstract

Information integration is an effective solution to support collaborations among various medical complexes in the regional healthcare environment. In this paper, an adaptive requirement-driven architecture for healthcare information integration based on service-oriented architecture is proposed. By adding business process layer and service layer, high business agility has been achieved. The function of each layer is discussed and the requirement-driven operational mechanism of this architecture is studied. Based on the requirement pattern modeling method, a case-based reasoning based requirement retrieval algorithm is designed in the requirement presentation layer to achieve the adaptive architecture and enable the information integration platform to have self-improvement and self-development capability. The capability of this approach has been demonstrated through an empirical case study in one of the largest hospitals in China and the outcome reveals that healthcare services quality can be improved significantly by the seamless integration of heterogeneous information in the different information systems.

Keywords: Information integration, Adaptive architecture, Requirement pattern, Case based reasoning, Web service, Norm

1. Introduction

Regional collaborative healthcare is an on-going effort to provide continuous cooperative healthcare services to patients, which plays a significant role in reducing medical errors, enhancing operating efficiency and improving the medical service quality [1]. Besides, many other organizations such as the government authorities need the medical information to support their decisions making. All of the above depend increasingly on the efficient management and sharing of different information sources. However, current medical information systems are separated from each other and established based on different platforms with several their own familiar design methods according to their individual requirements, resulting to amounts of "information islands" which construct barriers for maximal usage of medical sources. Therefore, an integrated information approach to provide seamless connectivity and sharing is urgent to be studied.

The main challenges in the information integration process are listed as follows.

(1) Massive healthcare data which typically exists in different forms (structured and unstructured) residing in heterogeneous and distributed healthcare systems spanning a wide range of architectures [2].

(2) Information is volatile and always changing with the business processes. Currently, how to invoke proper information according to the specific business process remains a challenge.

---

* Corresponding author: Huiying Gao, Associate professor. E-mail: huiying@bit.edu.cn.

(3) Information is required by different stakeholders in the regional healthcare ecosystem, *i.e.* patients, physicians, community healthcare centers and government authorities. Lack of data semantic interoperability solutions has been a major obstacle in the exchange of healthcare data between different stakeholders.

Therefore, this paper focuses on a requirement-driven approach to achieve the information integration in the current regional healthcare environment. The rest of the paper is organized as follows. In section 2, theoretical and practical background is briefly introduced. In section 3 the adaptive requirement-driven architecture of information integration is proposed and the main function of each layer is described respectively. In section 4, the requirement-driven mechanism including the modeling of requirement pattern and a CBR-based retrieval algorithm is analyzed. An empirical case study based on medical institutions in China is conducted and discussed in section 5. Finally section 6 displays our conclusion and outlook.

2. Related Theoretical and Practical Background

2.1. Service-Oriented Architecture (SOA)

As a new software architecture and design principle, SOA combines the benefits of both an in-house and a cloud environment solution to address the integration of variation business. It provides an ideal solution for a healthcare environment, which contains numerous heterogeneous data services and isolated, specialized workstations that are difficult to interface with one another [3]. Unlike the previous architectures, SOA adds a service layer between business layer and technology layer and takes full advantage of the basic blocks of "services" [4]. Specifically, it breaks down existing medical software systems into independent but interoperable web services, which are leveraged to create complex medical applications. Therefore, it is independent of any particular programming language, that is, a service written in one language should be able to interact with the other ones written in many other languages. Further, several evidences (e.g. [5]) suggest that healthcare information systems developed with the principles of the SOA can deliver an improvement of organization in scalability, sustainability and agility.

2.2. Case-Based Reasoning (CBR)

CBR is a cyclic and integrated process for a new problem solving by learning from a previous similar problem situation [6]. Usually, each previous problem is described as a problem feature set and solution set, then stored as a source case in the case base. A new target problem is solved by retrieving one or more source cases, reusing the cases directly, or revising the solution appropriately according to the differences between the source case and target problem, and retaining the new experience by incorporating it into the existing case base. In regional healthcare environment, it is also much more convenient to provide the right information to right users by retaining and utilizing similar past users' requirement cases than generating directly.

2.3. Related Information Integration Researches

Several efforts have been made to cover the integration methods of information from various heterogeneous systems among modern healthcare institutions in different aspects. Some of them identify standards or guidelines for the message exchanging in order to realize the collaboration of different healthcare systems, such as Integrating the Healthcare Enterprise (IHE), Health Level Seven (HL7), Digital Imaging and Communication in Medicine (DICOM). However, these standards

would still not take their expected roles unless there is a framework of model to invoke the right message at the right time to the right stakeholders. Other researches have focused on the data interoperability or functional integration. Such techniques include cloud health information systems technology architecture (CHISTAR), which has been applied on electronic health records (EHRs) to achieve semantic interoperability of massive healthcare data [7], and a unified SOA-based framework for healthcare information management to implement the functional integration with well-defined web services [8].

Due to the heterogeneity and diversity of healthcare systems, the variety of stakeholders and activities, as well as segmentation of healthcare services, three essential factors are necessary to fully integrated in order to make all kinds of medical information shared among the healthcare institutions without affecting the local systems: data, functions and workflow [9]. However, most of the current researches focus just on one or two aspects, let alone the cooperation among these three aspects. Based on the above studies, through integrating and extending part of our group's previous researches including a requirement-driven adaptive architecture (RDAA)[10] for information system management, capturing and analyzing the user's requirements with organizational semiotics methods [11], CBR-based knowledge reference and retrieval of the enterprise repository [12], an adaptive requirement-driven architecture based on SOA for healthcare information integration is proposed to satisfy the different users' changing needs and ensure the information integration from these three aspects.

3. Adaptive Requirement-Driven Architecture for Information Integration

In the IEEE 2000, STD 1471 architecture is the fundamental organization of a system embodied in its components, their relationships to each other and to the environment and the principles guiding its design and evolution. By separating the implementation of orchestration and services, the resulting system can easily be changed by orchestrating the services in another way or just using other services in a certain orchestration. In an ideal situation, the services do not need to change. Essentially based on the services identification, Figure 1 presents an overview of new adaptive service-oriented architecture aiming to realizing regional healthcare information integration according to the requirement-driven mechanism. Five main layers are distinguished: data sources layer, component layer, service layer (consist of application, business and orchestration services layers), business process layer and requirement presentation layer.

A. Data Sources Layer and Component Layer

Data sources layer is especially responsible for the storage of the data resources from different healthcare organizations, including all kinds of relationship/object databases (e.g. HIS, CIS, RIS, LIS), EPR and multimedia data stored in each hospital, healthcare data from the community stations, health insurance data and many other public service data from the government departments. It can also be the virtual database or the data warehouse which is built on the special entity database based on the medical information/document interchange standard (*i.e.* HL7, DICOM3, XDS).

Component layer has established the underlying basis used to describe specific technical capabilities which include all the applications developed in J2EE, Microsoft .NET, the CICS large-scale machinery as well as many other kinds of technology contained within healthcare organization's service logic. Common communication protocols make the communication between these components developed in different programming languages possible. As the existing application logic can be exposed as services through compiling the wrapped outer covering, business logic can be used by

different systems in heterogeneous environment. Therefore, the links between the components are built and broken when they respond to the upper service requests [7]. Due to the loose coupling of components, current medical information systems in different organizations can be integrated without any alterations and all the information can be reused to adapt to the evolving medical environment.
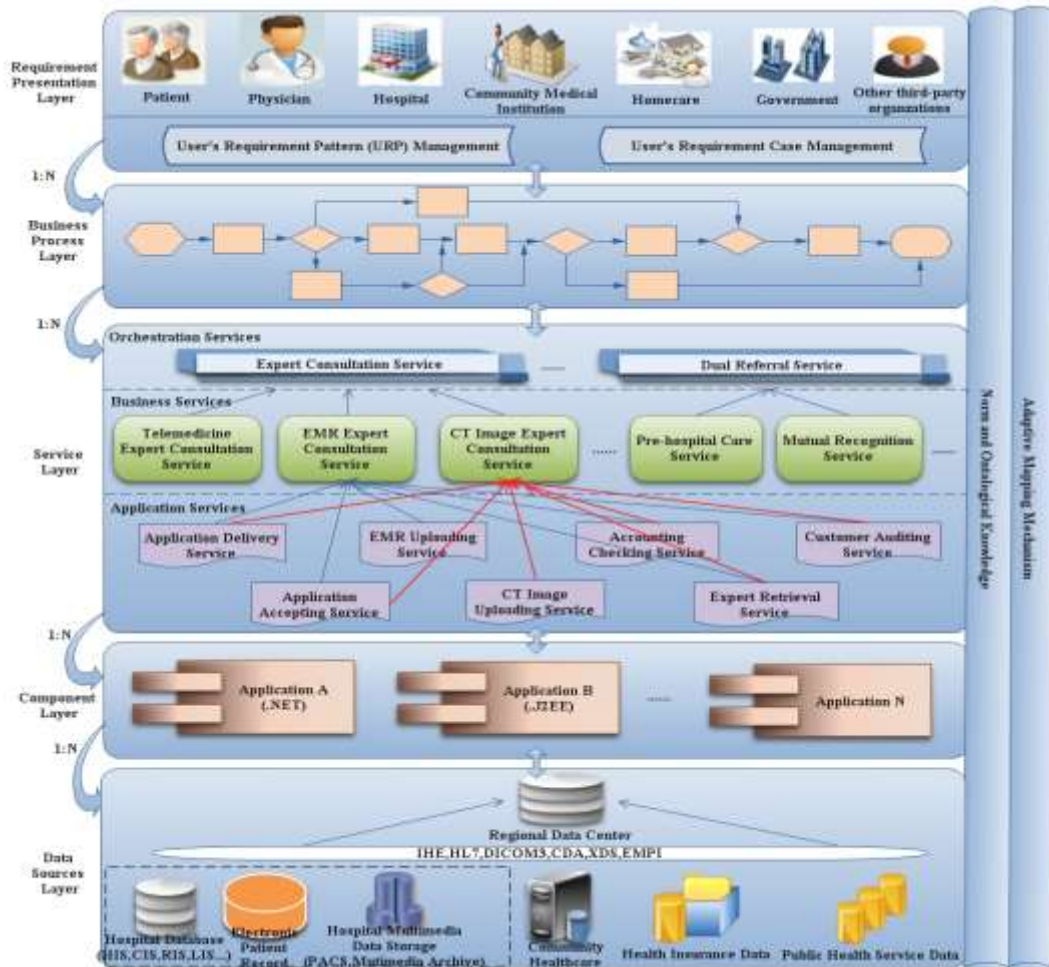

Figure 1. Requirement-Driven Architecture for Healthcare Information Integration

B. Service Layer

Service layer is composed of the particular business services race, reusable services and web services platform. The business services race is a group of services that aims at the special business areas and the corresponding data model. The reusable services refer to the sharing services in different business area. Web services platform applies an independent platform for business applications and technology platform environment for services definition and use. This layer can also be divided into three abstract layers-application services layer, business services layer and orchestration services layer [13].

Specifically, application services layer is composed of many function modules, such as CT image uploading service, expert retrieval service, application accepting service and so on. These primitive activities represent the smallest business logic that can be defined and carried out in the service-oriented environment. Services at this level could be called

application services, which aimed to provide reusable functions related to data processing in new or legacy applications environment.

Business services layer has introduced the services focusing on describing the business logic known as business services. Each of this kind service can be composed of the smaller granularity application services. For example, EMR expert consultation service may be composed of accounting checking, customer auditing, expert retrieval, application delivery, application accepting and EMR uploading service.

Orchestration services layer contains one or more processes, which is composed of business services and applications services according to the business rules and business logic embedded in the processes defined in the upper business process layer. By distinguishing numerous small services, a series of different orchestrations can be developed, which can also be reused as services. And it abstracts the business rules and services implementation of the logical sequence from the other services to promote their agility and reusability. It is also worth mentioning that the services in the lower layer can be reused for several times, e.g. application delivery service is both part of the EMR expert consultation service and part of CT image expert consultation service.

C. Business Process Layer

Business process layer provides numerous inter-departmental, end-to-end business processes and operational mechanisms. It undertakes the responsibilities of identification, modeling, development, deployment and management of business processes. Business processes can be got through the orchestration of the business service races in different areas and reusable services. It contacts with business service layer closely through "services". The description of the services provides a well-defined reasonable interface for business service race, so that business processes can transfer it easily without understanding business applications and details of technology platform. Registry of services in the business services layer ensures business process layer visit the required services dynamically in the right time.

D. Requirement Presentation Layer

Requirement presentation layer provides entrances of the regional cooperative healthcare information integration system for multiple stakeholders including patients, physicians, hospitals, community institutions, homes, governments and many other third-party organizations. The layer identifies and standardizes the requirements of users into query vectors, which will be used in the retrieval processes followed. It also establishes User's Requirement Case Base (URCB). User's requirement cases in high frequency are analyzed, classified and standardized into a series of User's Requirement Patterns (URPs). Main techniques in this layer will be elaborated in the section 4.

4. The Requirement-Driven Mechanism

The adaptive mechanism defined in the architecture can integrate the workflows, functions and data organically. In Figure 2, several artifacts in the application server are designed for the requirement presentation layer, business process layer and service layer respectively. "Pattern" is the atomic unit of information in these layers and follows the one-to-many relationship from top to bottom. That means a user pattern is often available by one or more business processes, a pattern of business process is made available by implementing multiple services, and so on. The searching and matching mechanism between adjacent layers are based on CBR and norms. Ontology and norm (including the social norm, organizational norm and the operational norm [14-15]) needed by each layer are classified and stored in the corresponding database. For the requirement presentation layer, URP acquisition engine is used to convert the user's requirements to patterns

through formalized representation and classification. The URP retrieval engine matches the user's new formalized requirement patterns with existing ones in the URP base. The integrated inference engine is responsible for generating the new patterns through the case, ontology and norm based inference with the help of the UR case base, ontology base and norm base, respectively. If the search and inference engine can find the similar cases, the business process invocation engine will call the adjacent business process layer to acquire processes in order to return the existing URP. Otherwise, search engine will discompose the sub queries, reorganize and composite the business processes and services. Therefore, this hierarchical operational mechanism enables the system to adapt to the evolvement of users' demands and environment with self-improvement and self-development capability.
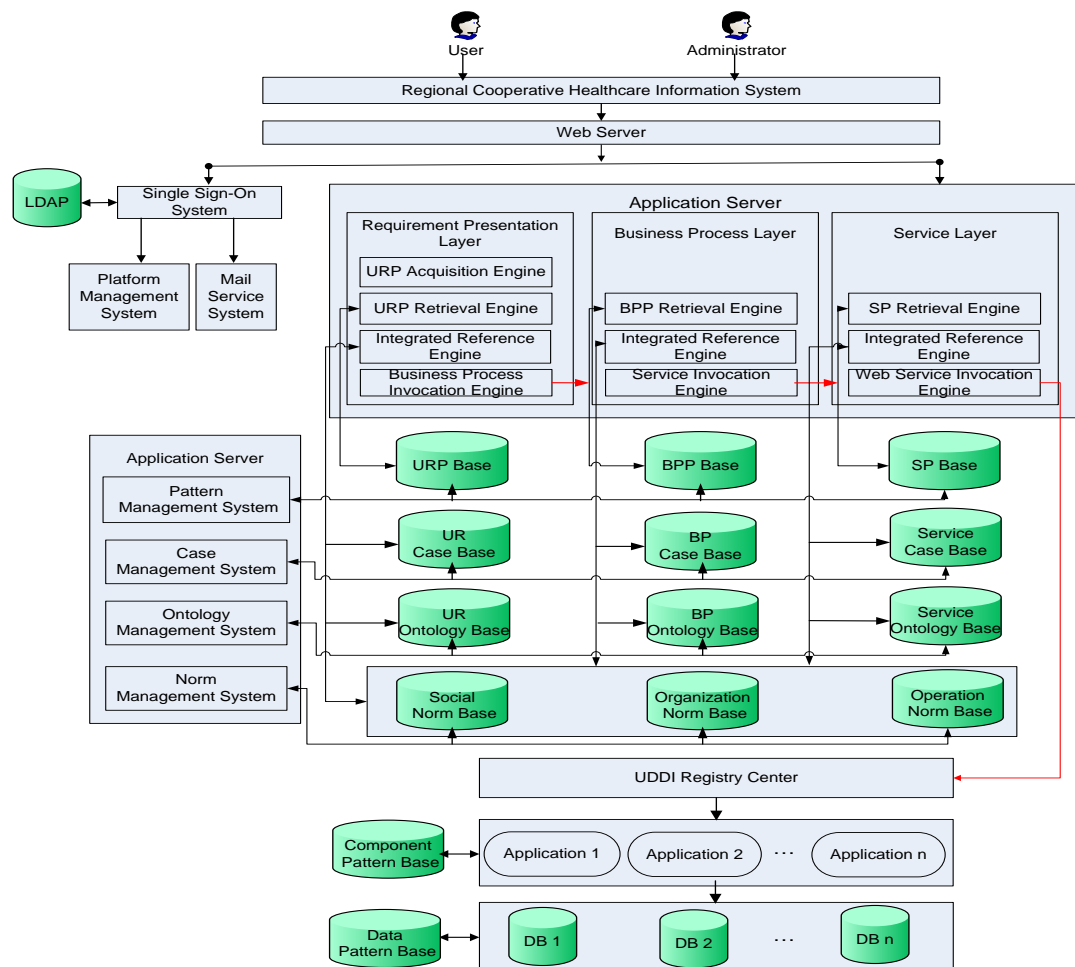

Figure 2. The Operational Mechanism of the Requirement-Driven Architecture

4.1. User's Requirement Pattern Modelling

URP can be defined as the abstract description of all the factors related to the system assignments and the relationships among them in order to achieve each stakeholder's own objectives on certain specific conditions. In our proposed architecture, UPR is expressed as $P = \{U, O, A, E, K, R, T\}$; $U = \{$Investors, Owners, Builders, Customers$\}$; $O = \{$Overall Objective, Business Objectives, IT Objectives$\}$; $A = \{$Core Assignments, Aided

Assignments}; $E = \{$IT Environment, Social Environments$\}$; $K = \{$Ontology, Norm$\}$; $R = \{R_1, R_2\}$, R_1 represents the relationships between the objectives and the assignments, while R_2 stands for the ones between the objectives and the environments. T means time.

As an important approach of requirement reuse, URP can make the process of locating and analyzing user's requirements more quickly and accurately. Two ways are mainly adopted in this paper to generate the URP. One is that knowledge engineers define the typical URP directly; the other is that the system carries out the integrated reasoning engine to create a new URP on condition when no URP in URP Base can meet the user's requirement. The 6-stage process of the latter way is shown in the Figure 3.
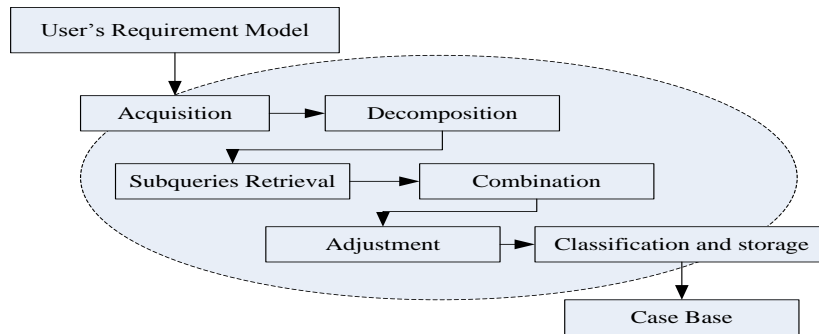

Figure 3. 6-Stage Lifecycle of URP Generation

Since the URPs are the results of analyzing users' requirements with the user's requirement model, their structures are also based on the pattern we built above. Figure 4 shows the CT image expert consultation URP presentation in URDL (User's Requirement Description Language) after the analysis.
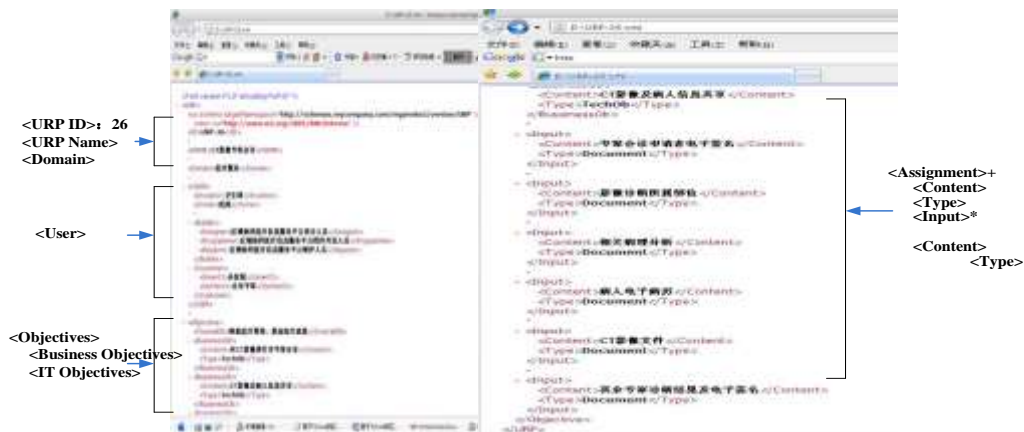

Figure 4. A Fragment of the CT Image Expert Consultation URP Presentation in URDL

The knowledge (K) in the requirement pattern includes ontology and norm. Ontology ensures the semantic interoperability of various kinds of information (workflows defined in the Business Process Layer, functions defined in the Service Layer and data stored in the Source Layer) as well as the accurate user requirement representation. Norm regulates the invoking and re-organization among these layers, which will be given an example in section 5.

4.2. User's Requirement Pattern Retrieval

The CBR-based retrieval strategy of URP provides support for getting right services and meeting user's requirement intelligently. Particularly, if the user finds similar user requirement patterns, the user requirement presentation layer will search the corresponding business processes according to the mapping information in the URP base. Then the business process layer will cooperate with the services layer to get the corresponding services. If there aren't similar patterns, the query vector should be decomposed into several independent sub queries and the matching is implemented on the UR case base. If the matching fails, the new user requirement can be standardized as a requirement case and saved in the UR case base. The CBR based matching algorithm is described as follows.

Suppose in an m dimension vector space, we have query vector $x_q = [x_q^1, x_q^2, ..., x_q^m]^T$ and case vector $x_d = [x_d^1, x_d^2, ..., x_d^m]^T$, the similarity between two vectors is generally measured as the cosine of the angle between them. Then,

(i) Input query demand and convert it into query vector x_q.

(ii) Compute the similarity between x_q and each user requirement case x_d by using (1).

$$Sim(x_q, x_d) = \frac{x_q \Box x_d}{|x_q| \cdot |x_d|} = \frac{\sum_{k=1}^{m} x_q^k \times x_d^k}{\sqrt{\sum_{k=1}^{m}(x_q^k)^2} \times \sqrt{\sum_{k=1}^{m}(x_d^k)^2}} \tag{1}$$

(iii) Return the user requirement cases as matching set G when $Sim(x_q, x_d) \geq \varepsilon$.

(iv) If $G = \varnothing$, then go to step (v), else stop.

(v) Decompose query vector x_q into sub query x_{q^i} ($i = 1, 2, ..., j, ..., n$), where the jth component of is 0, if $j \neq i$.

(vi) For $i = 1$ to n.

(vii) Compute the similarity between x_{q^i} with each user requirement case in case base.

(viii) Return matching set G_j.

(ix) End for.

5. Evaluation: an Empirical Case Study

In order to meet the various information requirements of stakeholders and realize healthcare information integration in the regional medical complexes for medical quality improvement and efficiency enhancement, we have established Regional Cooperative Healthcare Information System (RCHIS) implemented based on the proposed adaptive requirement-driven architecture.

A. Project Background

The Hospital of General Staff Department (HGSD) of the People's Liberation Army, one of the largest hospitals in China, possesses more than 50 subsystems that can be classified into 9 categories including physician workstation, nurse workstation, electronic medical record, clinical information, pharmacy management, decision support, OA, anesthesia management and audio-visual center. Each of these subsystems supports different aspects of daily activities in the hospital but most of them are separated. Besides, in recent years, it is urgent for HGSD to

realize information sharing and exchanging with other military hospitals, community medical institutions, sanitariums, General Military of Health and even the chronic patients' houses.

B. User's Requirement-driven Mechanism for Information Integration

Figure 5 shows the interface of URP retrieval. Users enter some key words, such as domain areas, main stakeholders, URP attributes and names, to search the proper URP from the case base.

Figure 5. Interface for URP Retrieval

If the existing URPs cannot meet the user's need, a new URP will be built according to the 6-stage process illustrated in the Figure 3. Now, if a URP which matches CT image expert consultation doesn't exist, a new URP can be generated based on the existing expert consultation pattern. Firstly, user gives the most simplified requirement field and description of problem, as shown in Figure 6. Then the requirement field and problem description will be matched to the field and objection of cases to find out the similar user requirement cases with the proposed retrieval algorithm. If the user requirement cases can be accepted, the system will carry out the reference engine; otherwise, further interaction is required and users will be asked to determine the target of system and the above procedure will be repeated until the case can be accepted or a new URP is built.
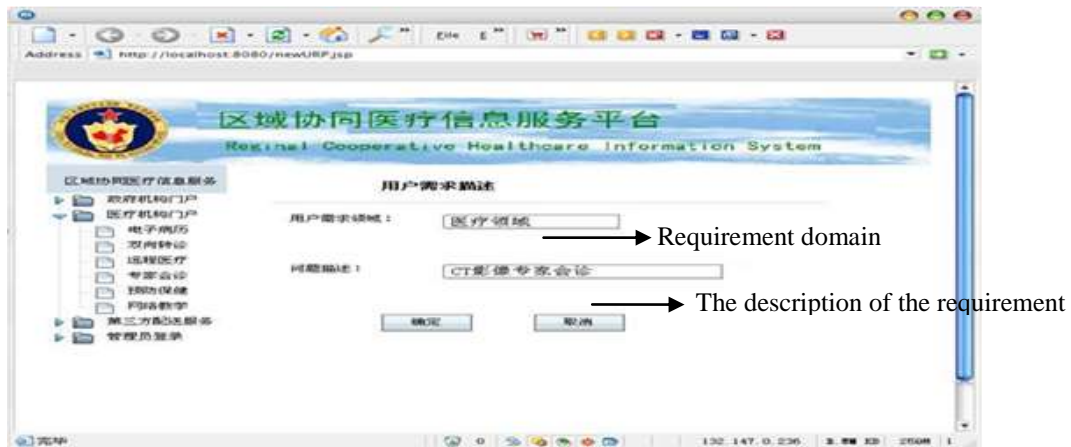

Figure 6. Interface for New URP Generation

Figure 7 shows the case of Electronic Health Record expert consultation can be accepted by users. After that, it will be revised with reference engine based on the ontology and norm knowledge in order to meet user requirement more accurately.
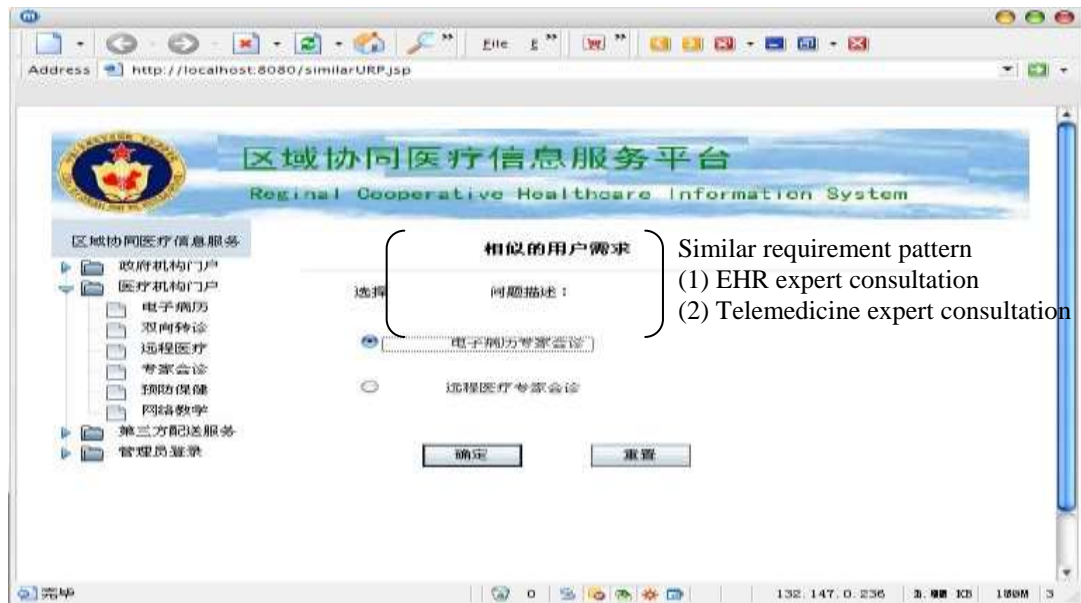

Figure 7. Similar URP Set

In the regional cooperative healthcare information integration platform implemented with our proposed architecture, we have established business process layer, service layer and data source layer for EHR expert consultation pattern and the three layers are available now. Then we can invoke the corresponding business process and lower services of the EHR expert consultation pattern to revise the existing norms of EHR expert consultation pattern to reorganize the required pattern. A good example can be:

N-customer2: Account Checking
Whenever < Old Customer>
IF <Pass the Account Checking>
Then <Choose Experts>and <Upload EHR>
Oblige To < Expert Consultation >
With revising and reusing, it can be updated as:
N-customer2: Account Checking
Whenever < Old Customer >
IF < Pass the Account Checking >
Then <Choose Experts> and <Upload EHR> and <Upload CT Images>
Oblige To < Expert Consultation >

When new requirement occurs such as requirement of CT image expert consultation, the system can revise the workflow of EHR expert consultation (shown in the Figure 8 (a)) to meet the requirement and eventually constitute a new application (Figure 8 (b)).
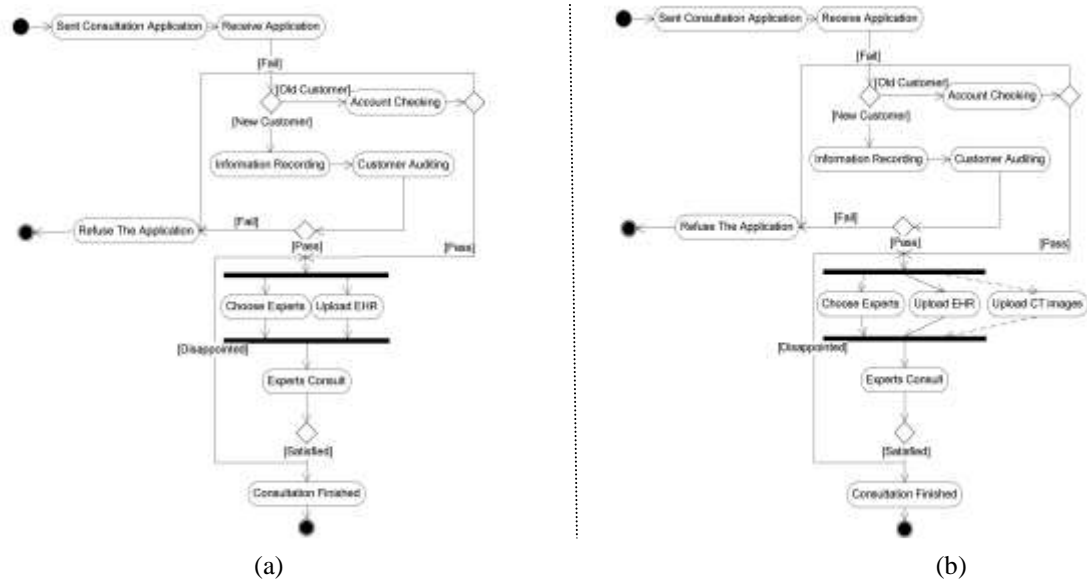
(a)   (b)

**Figure 8. Workflows of EHR Expert Consultation URP (A) and CT Image
Expert Consultation URP (B)**

6. Conclusions and Outlook

The information integration in regional healthcare environment is a hot issue at present and it plays a dominant role in improving the quality of medical services. A self-adaptive and service-oriented architecture for information integration is proposed in this paper. The function of each layer and the mapping mechanism among different layers are discussed. According to the description way of URP, a CBR based requirement retrieval algorithm is designed in the requirement presentation layer. If there are similar user requirement cases, the user requirement presentation layer will search the corresponding business processes according to the mapping information in the case base. Then the business process layer will cooperate with the service layer to get the corresponding services. If there aren't similar cases, the new user requirement should be standardized as a case and saved in the case base. It is an iterative process of problem solving, learning from the experiences and new problem solving. It enables the regional cooperative healthcare information integration platform to have self-adaptability, self-improvement and self-development, especially fitting for complex requirements, long-distance collaboration and heterogeneous structures of resource integration management.

Our approach is continuously being evaluated and improved with expanding user's requirements from different organizations. More investigations on the context environment including time and space features of the user's requirement should be conducted. Besides, the ontology reference mechanism should be further explored.

Acknowledgements

References

[1] L. Z. Ren, "Analysis on the Concept and System Architecture of Region Medial Coordination Information System", Information and communications technologies, **(2008)**, pp.6-10.

[2] E. Park and H.S. Nam, "A Service-Oriented Medical Framework for Fast and Adaptive Information Delivery in Mobile Environment", IEEE Transactions on Information Technology in Biomedicine, 13, 6 **(2009)**, pp.1049-1056.

[3] L. Constantinescu, J. Kim and D. Feng, "SparkMed: A Framework for Dynamic Integration of Multimedia Medical Data into Distributed m-Health Systems", IEEE Transactions on Information Technology in Biomedicine, vol. 16, no. 1, **(2012)**, pp.40-52.

[4] J. Haan, SOA and services identification. http://www.theenterprisearchitect.eu/archive/2007/04/26/soa_and_service_identification, **(2007)**.

[5] T. H. Yang, Y. S. Sun and F. Lai, "A scalable healthcare information system based on a service-oriented architecture", Journal of Medical Systems, vol. 35, no. 3, **(2011)**, pp. 391–407.

[6] A. Aamodt and E. Plaza, "Case-Based Reasoning: Foundational Issues, Methodological Variations and System Approaches", AI Communications, vol. 7, no. 1, **(1994)**, pp.39-59.

[7] A. Bahga and V. K. Madisetti, "A Cloud-based Approach for Interoperable Electronic Health Records (EHRs)", IEEE Journal of Biomedical and Health Informatics, vol. 17, no. 5, **(2013)**, pp.894-906.

[8] K. M. Anandkuma and C. Jayakumar, "A Unified Framework for Pervasive Healthcare Medical Computing Based on SOA", Advances in Computing and Information Technology, vol. 3, **(2013)**, pp. 319-327.

[9] I.E Azami, M.O.C. Malki and C. Tahon, "Integrating Hospital Information Systems in Healthcare Institutions: A Mediation Architecture", Journal of Medical Systems, vol. 36, no. 5, **(2012)**, pp. 3123-3134.

[10] R.C. Gan and G. X. Cui, "A Study on the Requirement-driven Adaptive Architecture of Information Systems", China Journal of Information Systems, vol. 2, no. 1, **(2008)**, pp. 66-74. (In Chinese)

[11] H.Q. Yang, K.C. Liu and W.Z. Li, "Adaptive requirement-driven architecture for integrated healthcare systems", Journal of Computers, vol. 5, no. 2, **(2010)**, pp.186-193.

[12] H. Y. Gao and X. X. Chen, "Ontology and CBR-based dynamic enterprise knowledge repository construction", Journal of Software, vol. 7, no. 6, **(2012)**, pp.1211-1218.

[13] T. Erl, "Service-Oriented Architecture Concepts, Technology and Design", China Machine Press, Beijing, **(2006)**. (In Chinese)

[14] K.C. Liu, "Semiotics in information systems engineering", Cambridge University Press, UK, **(2000)**.

[15] K.C. Liu and W.Z. Li, "Organisational Semiotics for Business Informatics", Routledge Press, UK **(2014)**.

Authors

Xiuxiu Chen, she is a lecturer in the School of Management and Tourism in Fujian Agriculture and Forestry University, China. She holds a PhD degree in Management Science and Engineering from Beijing Institute of Technology in 2015. She has ever been in the Informatics Research Centre, University of Reading for her Ph.D. research from 2013 to 2014. Her research focuses on healthcare informatics, information architecture and integration, knowledge based system and Organizational Semiotics. Her previous publications cover Organizational Semiotics in healthcare and system integration.

Huiying Gao, she is an Associate professor in the School of Management and Economics, Beijing Institute of Technology, China. She received her doctor degree in Management Science and Engineering from Beijing Institute of Technology in 2003. She has ever been in Technical University of Berlin, Germany to do her Ph.D. research work from 2002 to 2003. From 2008 to 2009 she has visited Karlsruhe Institute of Technology, Germany for half a year. Her current research interests include theory and method of information systems, content and knowledge management, semantic retrieval, intelligent information system *etc*.

Kecheng Liu, he is a fellow of the British Computer Society, full Professor of Applied Informatics, Director of the Informatics Research Centre and Head of School of Business Informatics, Systems and Accounting, within Henley Business School at the University of Reading, United Kingdom. He has published over 200 papers in the journals and conferences, and serves in editorial boards of several journals. He is one of the key figures in Organizational Semiotics. His research includes information systems, pervasive informatics and intelligent systems.

Hongqiao Yang, he is a co-operative supervisor of Informatics Research Centre, University of Reading, for PhD supervision in China. He conducts the postdoctoral research in computer science and technology in China. He is also the visiting scholar in Informatics Research Centre, University of Reading. His research interests are in the field of healthcare informatics, artificial intelligence, information system design and Organizational Semiotics.