A Knowledge-Constrained Role-Based Access Control model for protecting patient privacy in hospital information systems


It is advisable to refer to the publisher’s version if you intend to cite from the work. See Guidance on citing.

To link to this article DOI: http://dx.doi.org/10.1109/JBHI.2017.2696573

Publisher: IEEE

All outputs in CentAUR are protected by Intellectual Property Rights law, including copyright law. Copyright and IPR is retained by the creators or other copyright holders. Terms and conditions for use of this material are defined in the End User Agreement.
www.reading.ac.uk/centaur

CentAUR

Central Archive at the University of Reading

Reading's research outputs online
A Knowledge-Constrained Access Control Model for Protecting Patient Privacy in Hospital Information Systems

Runtong Zhang, Senior Member, IEEE, Donghua Chen, Xiaopu Shang, Member, IEEE, Xiaomin Zhu, and Kecheng Liu

Abstract—Current access control mechanisms of the hospital information system can hardly identify the real access intention of system users. A relaxed access control increases the risk of compromise of patient privacy. To reduce unnecessary access of patient information by hospital staff, this paper proposes a Knowledge-Constrained Role-Based Access Control (KC-RBAC) model in which a variety of medical domain knowledge is considered in access control. Based on the proposed Purpose Tree and knowledge-involved algorithms, the model can dynamically define the boundary of access to the patient information according to the context, which helps protect patient privacy by controlling access. Compared with the Role-Based Access Control model, KC-RBAC can effectively protect patient information according to the results of the experiments.

Index Terms—Privacy, knowledge, access control, information systems, hospital

I. INTRODUCTION

The hospital information systems, including Hospital Information System (HIS) [1], Laboratory Information Management System (LIS) [2], Picture Archiving and Communication System (PACS) [3], Radiology Information System (RIS) [4] and Electronic Medical Record System (EMRS) [5], have been adopted in many hospitals to meet the needs of different users. Data in these systems which are associated with the personal information of the patients are crucial resources for the users. For example, physicians can carry out diagnosis and treatment for patients according to their personal information and medical history; and medical researchers can review and study the cases. As a consequence, to make the information and data accessible to appropriate users is essential for the design of such information systems. However, the information in medical and health data in the hospital information systems is extremely personal and private; therefore access to which should be strictly controlled to protect patient privacy and avoid information leak. Protecting the information in such systems is an important issue. To protect patient privacy, from the legal and technical levels, many studies [6–7] have been conducted in the area of information security. However, the existing access control models have not given full consideration on domain knowledge such as the hospital processes and clinical pathways. Following a preliminary version of this work [8], this paper studies the privacy-protection mechanism for information systems in the hospital and puts forward a Knowledge-Constrained Role-Based Access Control (KC-RBAC) model, based on the Role-Based Access Control (RBAC) model.

The rest of this paper is organized as follows. Section II reviews related RBAC work and patient privacy. Section III describes our research motivation and the formal model description of KC-RBAC, where Knowledge Module (KM) and Purpose Tree (PT) are introduced. Section IV gives the three algorithms to identify real user purposes, determine user permissions, and protect patient information using medical knowledge. Experiments have been conducted to verify the performance of KC-RBAC model and are discussed in Section V. This paper ends with conclusions and perspectives.

II. RELATED WORK

A widely adopted model, the RBAC model, is proposed to simplify the access management of information systems [9]. In RBAC, system users are distinguished by roles according to the access policies, which enables the system to assign a new system user with the corresponding access rights automatically rather than case by case manually. Hung et al. [10] propose a privacy-based framework to tackle the need in e-Healthcare services to protect health information. The access purpose is also considered in the design of access control mechanism, which allows privacy officers to specify what data should not be used for certain purposes [11]. With the use of big data, the knowledge from the large volume of medical data can help improve data management with security and performance constraints [12]. The personal health record system offers new opportunities for personalized healthcare management because patients normally worry about their personal information being used inappropriately [13]. Faced the rapid growth of users and information, Hsu et al. [14] propose the role-based access control model to deal with authorization in the healthcare systems.
systems. To capture domain knowledge in a formal language, such as ontology modeling and semantic modeling, to enable automatic performance of access control, is necessary [15]. Gritzalis et al. [16] conduct a study on risk assessment to protect patient privacy in a shared care platform for the treatment of patients suffering from beta thalassemia. Rostad et al. [17] believe that the main problem facing the busy clinicians is to avoid being exposed with both irrelevant and relevant information at the same time. Hence, most studies focusing on the privacy protection in hospital information systems are based on the RBAC.

Another research mainstream focuses on the improvement of application systems using semantic approaches and medical domain knowledge [18, 19]. In the personal health record and health knowledge sharing system design, Lee et al. [20] use the Integral Healthcare Enterprise-Cross Enterprise Document Sharing (IHE-XDS) and the W3C Web Ontology Language (OWL) to maintain the Personal Health Records (PHRs) and collate the useful health Web resources related to the personal diseases. Making better use of existing data in the electronic health records to identify eligible subjects can improve efficiency and quality of medical care [21]. Discovering the knowledge from the medical records may support medical personnel in making clinical decisions and also help improve personalized medicine and care [22]. Considering the difficulties of ensuring the security and appropriate use of patient health information contained in medical records, an explanation-based auditing system (EBAS) [23] was proposed to distinguish the clinical or operational reasons and present metrics to determine which hospital employees are responsible for treating a given diagnosis.

It is also widely argued that, the knowledge in medical domains should be seriously considered when using RBAC model in HIS. Some knowledge like the relationship between different diseases is given in some international coding standards, such as ICD-10 (International Classification of Diseases 10th edition) and SNOMED CT (Systematized Nomenclature of Medicine-Clinical Terms), and other biomedical domain ontologies. Beimel et al. [24] present the SitBAC knowledge framework to infer new knowledge based on the incoming data access request from the realization process. For the pervasive healthcare systems, Li et al. [25] propose an authorization model that supports specifying and enforcing authorizations in a flexible and efficient way to conceptualize the data and explicitly express the relationship among concepts and instances. Considering that personal healthcare applications range over many disciplines, Blobel et al. [26] introduce the care paradigms and discuss the requirements to meet the business objectives. In order to acquire better accuracy and coverage, Gordona et al. [27] propose a cost-effective semi-automated method for generating a useful knowledge compendium with minimal reliance on domain experts. To some extent, these studies are to make use of medical domain ontology to improve the performance of RBAC.

A comparison of the main access control models reviewed above is given in Table 1 from the aspects of role, purpose, knowledge, confidentiality, integrity, and flexibility. For the patient privacy protection in hospital information systems, current access control models still need further improvement in reducing access rights by combining the medical domain knowledge and the existing RBAC model.

### III. A KNOWLEDGE-CONSTRAINED ACCESS CONTROL MODEL

#### A. Motivation and General Idea

RBAC and its relevant models can restrict access rights according to the role of the users, but this is based on their specialties. Theoretically, some employees may have rights to access all the medical records of the patients even if their jobs have no relationship with patient information at that time. This brings potential risks of patient privacy disclosure. Thusly, to make the users’ access right more accurate, we take the medical domain knowledge and context information into consideration to formulate a novel knowledge-constrained access control model—KC-RBAC.

Figure 1 gives an example of the traditional RBAC model, where $A_i (i = 1, 2, 3, ..., n)$ represents the $i$th independent access right. In this case, ROLE 1 possesses several access rights which are different from those of ROLE 2.

![Fig. 1. Difference of access rights between two different roles in the traditional RBAC model](image)

Following the case of Figure 1, taking knowledge and context into consideration, another case is given in Figure 2.

![Fig. 2. Difference of access rights for the same role when considering knowledge and context](image)

In this case, the access rights ($A_i$, $i=1,2,...,m$) of the same role (ROLE 2) are redefined as $a_j (j=1,2,...,n)$ by introducing the knowledge and context factors. It brings a better control on accessing data and information. Even for the same role, the acquired patient information is different because the knowledge has a dynamic effect on those user permissions. Moreover, this mechanism has no impact on the origin system business logic. The context in Figure 2 is usually associated with the real access purposes of system users. In the hospital process, user intention is recognized as the original purposes that the user intends to do something. It differs from the predefined user purposes of visiting the systems. In this paper, user intentions are usually obscure and hard to be predicted, whereas user purposes are very clear and predefined in the design of system. Figure 3 illustrates the application of KC-RBAC in real life.

<table>
<thead>
<tr>
<th>Models</th>
<th>RBAC</th>
<th>PE-RBAC</th>
<th>P-RBAC</th>
<th>OBAM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Role</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td>Purpose</td>
<td>NO</td>
<td>NO</td>
<td>YES</td>
<td>NO</td>
</tr>
<tr>
<td>Knowledge</td>
<td>NO</td>
<td>NO</td>
<td>NO</td>
<td>YES</td>
</tr>
<tr>
<td>Confidentiality</td>
<td>NO</td>
<td>YES</td>
<td>NO</td>
<td>NO</td>
</tr>
<tr>
<td>Integrity</td>
<td>YES</td>
<td>NO</td>
<td>NO</td>
<td>YES</td>
</tr>
<tr>
<td>Flexibility</td>
<td>NO</td>
<td>NO</td>
<td>NO</td>
<td>YES</td>
</tr>
</tbody>
</table>

### Table 1: COMPARISON OF THE EXISTING ACCESS CONTROL MODELS
The rest of the paper explores the models and algorithms to implement our motivation and idea.

B. Model Description

Based on the RBAC model, medical domain knowledge and user purpose are involved in the new model.

1) General Model

KC-RBAC is used to define the mapping relationships between roles and permissions by referring to the medical knowledge and the PT with the knowledge inferences and semantic analysis to obtain more precise control of permission. PT serves as a tree hierarchy that describes the purpose of the use of the hospital information systems. Though the role hierarchy in hospitals is static, permissions for a role, as aforementioned, are dynamic and determined by the PT of the user. Following the basic structure of the traditional RBAC model, our proposed KC-RBAC model is shown in Figure 4.

Different from the RBAC model, the knowledge module in Figure 4 is introduced to identify the real access purposes of the users. The access purpose in this paper is a concept that represents the real needs of the system users in a hospital. For example, if a doctor wants to offer medical advices to his or her patients, the system should provide the doctor only the necessary access rights to the relevant patient information. However, the key to identify what the doctor actually needs depends on the KC-RBAC model which analyzes the user intention based on the knowledge. The knowledge relevant to the scope of the patient information acquired includes the hospital processes, roles, as well as the domain knowledge in the medical field. Thus, the access control model can restrict the hospital employees’ access to other information which is unnecessary to the process of medical diagnosis and treatment. Such an approach will protect patient privacy as far as possible.

2) Knowledge Module

The Knowledge Module is the key component in KC-RBAC to determine the access rights of system users to protect patient privacy. Those basic knowledge, such as ICD-10 and SNOMED CT, provides the ability of knowledge inference in the KC-RBAC.

In RBAC, the role hierarchy is preliminary knowledge in medical information systems. Many different roles with different purposes exist in the hospital because of the different processes in the medical departments. Thus, as a key part of the knowledge-based system, the role hierarchy is very important. The medical institutions normally have many employees, so the rights of the role in the system are hard to distinguish even by using contexts. To give strict protection for patient privacy in information systems in the hospital, the permissions of the role hierarchy in Figure 4 should be divided into two types, namely, public and private permissions. The logic of the knowledge module is shown as Figure 5. Figure 5 shows that the key part of the Knowledge Module is the knowledge reasoning engine. It provides the user with patient information that a user actually needs based on the calculation of medical knowledge. Except for the aforementioned basic knowledge, the main knowledge sources in medical information systems are the knowledge contained in the patient’s medical records corresponding to the knowledge bases such as the SNOMED CT; the domain query language of which is SNOMED CT Expression Constraint Language (SNOMED ECL) to support the search for knowledge in massive knowledge. The number of medical records in the hospital increases rapidly, and the invisible domain knowledge in those records are valuable. Thus the medical records are useful to be reused in our proposed model.

3) Purpose Tree Module

The Purpose Tree (PT) Module is another key component that illustrates the user purposes of visiting the system in a hierarchical structure. In the different processes of the hospital, physicians and other hospital staffs may request for different information of the patient in HIS, which we call the motivation of visiting as visiting purpose. In order to describe the visiting purpose, PT is introduced in this model. Table II gives a typical PT of HIS. The Level-1 column suggests several general purposes that may appear in the HIS. Based on this, more specific system access purposes are introduced in Level-2 and Level-3. The relationships of these purposes can be graphitized in a tree hierarchy with three levels.
Definition II (PT):

- **P**, **PT**, **UI**, and **OU** (purpose, Purpose Tree, user intention, and outside user, respectively)
-  
  \[ P_u = f_d(UI, PT) = \{p | p \in PT, u \text{ is } p\text{'s owner}, UI = f_u(R, OU, and PM) \} \]
- \( f_u \) is a function to obtain what purpose they want based on the determination of **UI** and **PT**.
-  
  \[ P_u \] is the result of the total calculation of **PT** for user **u**.

Definition III (Knowledge Ontology):

- **AT**, **DE**, **RE**, and **P** (attributes, description, relationship, and purpose, respectively)
- **K** = \( \{AT, DE, RE, and P\} \) \( P \in PT \) is the last part of **P**, which is the key part to associate knowledge and purpose in this paper. The other parts are traditional relationship of domain ontology.
-  
  \[ K_u = K_r \cap K_{PM} \cap K_p, K_r \]
- is a set of knowledge constrained by ROLES, **K_{PM}** is a set of knowledge constrained by permission, and **K_p** is all knowledge with purpose **p**.

In Definition II, **UI** indicates the set of purposes such as hospital employees. **OU** represents those people who want to protect their privacy as much as possible (such as the patients). Here, we define a function \( f_u \) as

\[
f_p(UI, PT) = \{p | \text{diff}(p, iu) > \epsilon\}
\]

\[ p, ui \in PT, UI, 0 \leq \epsilon \leq 1 \]

\[ \text{diff}(x, y) = \begin{cases} e = \text{sim}(x, y) & x = y \\ \epsilon & x \neq y \end{cases} \]

In Equation (1), the result of \( f_u \) is defined as a set in which the elements are determined by function \( \text{diff}(p, iu) \) as in Equation (2). The function \( \text{diff} \) aims to calculate the similarity between user purpose and user intention. The function \( \text{sim}(x, y) \) calculates the similarity between \( x \) and \( y \). All the results of the calculation of function \( \text{diff} \) should be treated as a set of purposes that the users will own. The **UI** can be calculated by

\[
UI = \{ui | ui \in OU, ui \in R \cap PM\}
\]

Thusly, according to the Definition II, the purpose set of specific users can be calculated. As aforementioned, the ontology of medical domain knowledge, such as SNOMED CT, can help the system find the most knowledgeable relationships in patient information for better privacy protection in Definition III.

### IV. Algorithms and Experimentation

The formal model description above outlines the basic components of KC-RBAC and their relationships. Three methods in the KC-RBAC are proposed: Purpose Identification Algorithm (PRA), Permission Determination Algorithm (PDA), and Knowledge-involved Information Filtration Algorithm (KIFA).

#### A. Purpose Identification Algorithm (PRA)

PRA is an algorithm to identify the actual purposes of visiting the hospital information systems based on the intentions before using the systems. Algorithm I gives the step-by-step description.

**Algorithm I PRA**

**Input:** Purpose Tree **PT**, User Intention **UI**

**Output:** Purpose Set **UI**

**Step 1:**

1. For each role **r** in **PT**, calculate its similarity with user intention **ui** using function **sim(x, y)**.
   
   \[
   \text{sim}(x, y) = \frac{\sum_{i \in x} \sum_{j \in y} \text{diff}(x, y)}{|x| |y|}
   \]

2. If **sim(x, y)** is greater than a predefined threshold, **x** is considered a relevant purpose for **ui**.

**Step 2:**

1. Collect all relevant purposes and form the set **UI**.
To evaluate the performance of Algorithm I, it is necessary to compare the difference between a set of predicted purposes and a set of original purposes by assigning weights to the purposes and knowledge. The weight of a purpose is an integer, ranging from 1 to 10. A larger value means a more important purpose. The difference can be calculated by Equation (2). Additionally, the aggregate weight of UI (user intentions) and UP (recognized purposes) can be calculated as

\[ W_i = a \sum_{p \in I} W_i^p + b \sum_{p \in P} W_i^p \]  \hspace{1cm} (4)

\[ W_p = a \sum_{i \in I} W_i^p + b \sum_{i \in P} W_i^p - c \sum_{i \in P} W_i^p \]  \hspace{1cm} (5)

where \( W_i \) stands for the weight of UI, \( W_p \) stands for the weight of the user intention, and \( W_p \) means the weight of knowledge associated with the purpose \( p \). The constant \( a \) and \( b \) are set as 0.2 and 0.8 respectively. In Equation (5), \( W_p \) stands for the weight value of UP, and \( W_p \) stands for the knowledge associated with purpose \( p \). If there is a strong relationship between \( W_i \) and \( W_p \), \( W_p \) as a coordinator, is able to enhance its effect in the expecting results. The constant \( c \) is set as 0.2. Thusly, a weighted PT is built to quantify the model.

B. Permission Determination Algorithm (PDA)

Algorithm I returns a subset of the PT, which represents the actual purposes of a system user. Given that the system modules in the hospital are associated with specific access rights for a role, the following algorithm (Algorithm II) is presented to summarize all the permissions based on their actual purposes determined by Algorithm I.

Algorithm II PDA

Input: List<Permission> PM, Role R, PurposeTree PT, List<Permission> UM

Output: PM

1) Let \( p_1, p_2, ..., p_n \) as the purpose Tree calculated by FindCurrentUserPurpose() in Algorithm I
2) Let \( U \) as the current user in HIS
3) Let \( r_1, r_2, ..., r_n \) as elements of Role R
4) Let \( PT \) as Total Purpose Tree
5) Let \( pm_1, pm_2, ..., pm_n \) as elements of List<Permission>
6) Let \( ump_1, ump_2, ..., ump_n \) as elements of List<Permission> that \( r \) has
7) Let \( PM \) as List<Permission> that user have
8) for each \( pm_i \) in \( PM \) where \( i=1,2,...,n \) do
9) \[ iif(pm_i \in ump_{rIsRoleOf(U)}) \]
10) for each \( p_j \) in \( P \) where \( j=1,2,...,m \) do
11) \[ if(pm_j \rightarrow PM) \]
12) \[ PM.push(pm_j) \]
13) \[ } \]
14) \[ } \]
15) return \( PM \)

Algorithm II enables the KC-RBAC model to cluster those access rights based on the result of Algorithm I because the subset of PT may contain the duplicate access rights generated in Algorithm I.

C. Knowledge-Involved Information Filtration Algorithm

The outputs of Algorithm I consist of the purposes and their associated medical standard knowledge including ICD-10 and SNOMED CT, which contains massive medical concepts and relationships. Moreover, the purposes are associated with specific permissions (access rights) in Algorithm II. To integrate their relationships, Algorithm III, also called Knowledge-Involved Information Filtration Algorithm (KIFA), is presented below.

Algorithm III KIFA

Input: List<Permission> PM, Knowledge K, Roles R, Permission PM, Information I

Output: \( I \) as user information

1) Let \( k_1, k_2, k_3, ..., k_n \) as elements of Knowledge in libraries (K)
2) Let \( r_1, r_2, r_3, ..., r_n \) as elements of Knowledge that matches the purposes (K)
3) Let \( i_1, i_2, ..., i_n \) as elements of necessary information(I)
4) Let \( I_1, I_2, ..., I_n \) as elements of total information(I)
5) for each \( k_i \) where \( i=1,2,...,n \) do
6) \[ if((k_i.attribute(P) eq p_1) \]
7) \[ iif(I_i \in (R \cup PM)) \]
8) \[ r.push(k_i) \]
9) \[ } \]
10) \[ } \]
11) \[ } \]
12) \[ for each \( i \), where \( i=0,1,2,...,n \) do \]
13) \[ if(\{k_i \ hasKnowledge(k_i) \}) \]
14) \[ i.push(I_i) \]
15) \[ } \]
16) \[ return I \]

By using medical domain knowledge-driven libraries, based on the actual purposes of requesting of information, the KC-RBAC model can finally extract the necessary data from the system when hospital employees request patient information.

V. RESULTS AND DISCUSSION

Based on the proposed model and algorithms, this section summarizes the results of our experiments. More importantly, we report how the model works in the existing hospital information systems in real cases.

A. Case Studies of KC-RBAC in the Hospital

To evaluate the performance of KC-RBAC, we established a prototype hospital information system with real medical records of cancer patients (8,654 copies of records) to simulate the real process of querying the patient information in a hospital. These data contain four kinds of illness information, including esophageal cancer, stomach cancer, kidney cancer, bowel cancer, and leukemia (2,033 records). Each record contains 27 data fields, 11 of which are patient privacy data fields.

Patient personal data, such as the names, identification
numbers, addresses, and so on, are very sensitive for both patients and medical employees. The proposed PT, KM, and RH are introduced independently in this system. Two test indicators are used to evaluate the results of the experiments, including the number of medical records accessed (MRA) and the amount of medical knowledge (MK) as shown in Table III. In this simulation, we assume that each medical record is a 1*n matrix, each cell in the matrix can be seen as the basic unit of patient information. At the same time, each cell can be mapped to a set of medical domain knowledge. Thusly, the MRA is measured based on the number of medical records multiplied by the filtering ratio of the purpose, knowledge, role, and knowledge-role. And the MK is measured based on corresponding MRA multiplied by the number of concepts and rules of domain knowledge in each cell. These relationships and equations are listed in Table III.

A series of access control models are separately implemented upon the system to examine patient privacy protection performance of the different models. In order to illustrate how the model can protect patient privacy, two cases are given to discuss and evaluate the advantages of KC-RBAC.

### Table III

THE FORMAL EQUATIONS TO CALCULATE MRA AND MK OF THE DIFFERENT MODELS IN THE EXPERIMENTS

<table>
<thead>
<tr>
<th>Models</th>
<th>Components</th>
<th>MRA</th>
<th>MK</th>
</tr>
</thead>
<tbody>
<tr>
<td>Open-AC</td>
<td>N (all records)</td>
<td>MRA*K_p</td>
<td>MRA*K_p</td>
</tr>
<tr>
<td>RBAC</td>
<td>N* r(%)</td>
<td>MRA*K_p</td>
<td>MRA*K_p</td>
</tr>
<tr>
<td>P-RBAC</td>
<td>N* r(%)</td>
<td>MRA*K_p</td>
<td>MRA*K_p</td>
</tr>
<tr>
<td>K-RBAC</td>
<td>N* r(%)</td>
<td>MRA*K_p</td>
<td>MRA*K_p</td>
</tr>
<tr>
<td>KC-RBAC</td>
<td>N* r(%)</td>
<td>MRA*K_p</td>
<td>MRA*K_p</td>
</tr>
</tbody>
</table>

(N, number of medical records in the system; r(%), percentage of the role-based records; p(%), percentage of the purpose-based records; k(%), percentage of the knowledge-based records; K_p, K_r, and K_s, the purpose-, role-, and context-based knowledge)

**Case 1:** The matron (head nurse) specifies that a nurse is responsible for a patient, and the KC-RBAC-based system can automatically filter and provide information about the care but doesn’t disclose unnecessary information of the patient.

In Case 1, the nurses follow hospital policies to take care of patients, which means that they do not have to know all of the patient information. The KC-RBAC model in this case firstly determined the preliminary permission of the nurses based on their roles and context of hospital policies and associated the nurses to their patients. Thusly, the nurses know only the necessary patient information. Role, context of policies, and permissions are considered to protect patient privacy in the nursing process. When applying KC-RBAC model in the context of Case 1, Algorithm I can filter out the unreasonable user intentions and only give necessary access rights to user according to the using context. Figure 6 denotes the difference of patient information leaks between two identification algorithms by using the intention-based and purpose-based rules respectively. In the Intention-based Identification algorithm, we simulated the process of visting HIS without consideration on whether their visiting purposes are necessary for the users in some context (e.g. a nurse on duty tries to search all records for specific patients in the system). In contrast, for the Purpose-based Identification algorithm, the user intentions are double checked by Algorithm I and mapped to a set of predefined purposes in PT. Based on this, in the simulation, we generate a random array as the knowledge set, and then assume some mapping relationships between the generated knowledge and patient information (data cells in the 1*n matrix). Then we run ten rounds of comparative experiments. The simulation results are shown as Figure 6.

In Figure 6, the horizontal axis represents the number of rounds and the vertical axis represents the estimated values of patient information leaks in the simulation. The larger the value of vertical axis, the more information or access rights the user acquire. The values of the Intention-based Identification in vertical axis are higher than those of the Purpose-based Identification in each round of experiments. This indicates that only the Intention-based Identification is hard to reduce the information exposed to system users if no further procedures are adopted to identify whether the intentions are necessary.

**Case 2:** A doctor wants to do scientific research based on the electronic medical records in hospital information systems. Given that he/she may have a specific scientific research purpose, the KC-RBAC-based system automatically filters out all personal information and only displays purpose-related information.

In this case, the information that the doctor acquired in Case 2 contains not only the related medical records but also the knowledge associated with the related medical domains, which help the doctors do their research. The medical domain is involved in the response of the KC-RBAC model to provide doctors with more medical reference information. Given that the purpose of the doctor is scientific research, the patient information will be hidden, reducing the risk of privacy leaking. A series of experiments with several different access control models were simulated according to Table III. One experiment is to compare the difference of the MRA indicator in several models when querying the medical records. Another one is to compare the difference of the MK indicator in different models when querying medical records for research. The results of the experiments are shown in Figure 7.
because the patient information is widely existing in different sub-systems. All the access requests on patient information by hospital employees will be reorganized and modified by the KC-RBAC model to protect patient privacy. However, the KC-RBAC model is designed for the access control in those hospitals with complex processes and organizational structure, and as a consequence, it is not suitable for individual doctors.

The input of Knowledge Module in the KC-RBAC is dynamic, while the role hierarchy of RBAC model is static. For example, the adopted knowledge includes ICD-10 and SNOMED CT, the structure of which supports the expansion of medical knowledge in the future without changing our model. The user intentions for using the system are associated with the purpose hierarchy, and also mapped to the medical domain knowledge that includes the semantic maps of medical concepts. This flexible knowledge organization allows possible knowledge inference to our Knowledge Module.

VI. CONCLUSION

In a traditional RBAC model, one role is always mapped to a fixed set of access rights, which means the user with a specific role can access to corresponding modules of the systems in any context. This kind of access control mechanism lacks flexibility. In this paper, we propose a KC-RBAC model which has a more precise access control according to the user’s context. The identification of visiting purpose is realized based on the PT which indicates the relationships of purposes in different context. Compared with current RBAC model and its improvement models, KC-RBAC can effectively zoom out the scope of data and information that a system user can access. The reduction of patient information range accessed by the system user objectively enhances the privacy protection for patients. Based on results of the experiments, the proposed model is shown to be highly flexible to protect patient privacy in HIS compared with other access control mechanisms. Although the model is special for HIS, it also can be used for privacy protection in other information systems if corresponding context knowledge can be described and encoded in the knowledge module.

REFERENCES


