Ontology-Based Knowledge Representation for Secure Self-Diagnosis in Patient-Centered Telehealth with Cloud Systems

Keke Gai¹, Meikang Qiu²*, Saravanan Jayaraman³, Lixin Tao⁴
¹ Department of Computer Science, Pace University, New York, NY, 10038, USA, kg7123w@pace.edu;
² Department of Computer Science, Pace University, New York, NY, 10038, USA, mqiu@pace.edu
³ Department of Computer Science, Pace University, New York, NY, 10038, USA, sj91845w@pace.edu
⁴ Department of Computer Science, Pace University, NY 10570, USA, ltao@pace.edu.

Abstract—The implementation of cloud computing in telehealth has enabled enormous benefits of improving healthcare services as well as active health approaches, such as Patient-Centered Telehealth (PCT). Applying cloud systems enables healthcare users to obtain medical information from multiple cloud-based platforms or sources. Various service deployments meet different customers’ needs. As one of the popular information sharing manners on clouds, Online Self-Diagnosis (OSD) has been widely implemented for recommending medicines or treatment plans in the varied medical fields. However, there is a great risk for users when executing OSD since lack of professional pharmacological knowledge. This paper concentrates on this issue and proposes an ontology-based approach representing the medical knowledge for alarming potential risks. The proposed mechanism is named Secure Ontology-based Self-diagnosis (SOS) Model that is designed to generate knowledge of medical treatments or incompatibility of medicines to avoid improper behaviors of diagnoses. Based on the proposed model, our algorithm uses structure-level technique applying semantic schemas, which is Propositional Satisfiability Matching Algorithm (PSMA). Our experiment has proved the feasibility of our proposed mechanism.

Index Terms—Ontology, knowledge presentation, self-diagnosis, cloud systems, patient-centered telehealth

I. INTRODUCTION

Currently, a broad implementation of cloud computing in tele-health domains has introduced enormous benefits to healthcare industry. Along with the development of cloud computing technologies, cloud-based tele-health services have become an efficient approach for physicians to diagnose or offer treatments. As one of the derivatives of applying clouds, Online Self-Diagnosis (OSD) is an implementation of Information-as-a-Service delivering users medical advice [1]. This new type of medical services provisions a variety of new methods for generating a Patient-Centered Telehealth (PCT) solution that enables patients to have involvements in the medical or healthcare diagnosis and treatment processes [2]. However, using this cloud-based solution also brought a few new potential risks to patients and other related parties. An approach identifying the hazards with representing the crucial knowledge is an emergent significant issue for achieving secure PCT.

Focusing on this new issue, this paper proposes a new cloud-based mechanism using ontology-based knowledge representation by defining new relations between diagnosis advice and medical incompatibility. The significance of this research is that being aware of the medical incompatibility relations can assist patients to reduce the risk of self-diagnosis. The main method used for achieving this goal is implementing ontology to identify the interrelationships between the relevant entities with adopting OSD systems.

This paper proposes a new model describing the new relations deriving from adopting cloud-based PCT systems. The model is named as Secure Ontology-based Self-diagnosis (SOS) model, which is designed to reduce the error levels by illustrating the nuances of diagnosis and mapping relations between symptoms and treatments. Our propose model uses Web Ontology Language (OWL) to formulate the relations between medical objects. Applying the proposed model can assist patients to reduce the repercussions caused by cloud-based self-diagnosis.

The implementation can also strengthen the role of the physicians instead of weakening doctors’ positions. The output of our proposed model will be a medical condition report that can assist patients to have an informative discussion with doctors before the final treatment plan is formed. Therefore, using the proposed model can be saturated with various medical services by offering the enhanced security levels and treatment efficiency.

The main algorithm in the proposed model is Propositional Satisfiability Matching Algorithm (PSMA), which is designed to execute ontology matching. In addition, the main contributions of this paper are:

1) Our proposed scheme can enable a cloud-based intelligent service offerings by providing medical knowledge
Section VII.

given in Section VI. Finally, we have our conclusions in Section V and experimental examination and proof are of our proposed model. Our main algorithms are exhibited the approach of implementing the proposed model. Next, in Section III, we give a motivational example to explain the approach of implementing the proposed model. Related works are summarized and presented in Section II. In Section IV represents the main concepts and the structure of our proposed model. Our main algorithms are exhibited in Section V and experimental examination and proof are given in Section VI. Finally, we have our conclusions in Section VII.

II. RELATED WORKS

As an emergent technique, cloud computing has been paid a high attention by researchers over years. Shifting the service model to a flexible and complexity-masked manner has changed the roles in the service delivery processes [3]. We summarize the recent research work of using ontology in clouds in this section.

Some recent researches have focused on creating an ontological cloud model to gain common access to the cloud services [4]. For example, mOSAIC was developed for a common ontology focusing on an open-source platform for identifying cloud services requests defined by cloud vendors [5]. Next, an ontology for search engine in cloud systems was proposed [6], in which the interrelationships between cloud concepts are defined by three similarity reasonings, including concept, object property, and datatype property. However, prior research did not concentrate on the ontological approaches in PCT systems.

Security issues were addressed by the prior research [7]–[10]. A security ontology model for cloud systems was created to prevent the systems from various threats [11]. Meanwhile, an ontological security-oriented service framework was developed to check the security compatibility, which was based on the security policies [12]. Nevertheless, most previous researches rarely addressed the security ontology when adopting ontology in cloud-based PCT systems.

Migrating on-premise services to the cloud-based deployments is a challenging issue for cloud vendors due to a group of problems, such as secure data migration and data validation in clouds [13]. Recent search also addressed this issue and one of the proposed solutions is using ontological approaches to monitor service deployments [14]. Nonetheless, the ontological approach focusing on dynamic service offerings has been rarely explored by the prior research.

Furthermore, different cloud service types enables various service implementations, such as Software-as-a-Service (SaaS), Infrastructure-as-a-Service (IaaS) [15], and Platform-as-a-Service (PaaS) [16]. The implementations of cloud computing can trigger various novel applications that can improve the existing business processes or create new services or products [17]. However, the relations between physicians and PCT have hardly been explored in the perspective of cloud service types.

Next, prior research has also addressed the optimizations of cloud computing for reaching green clouds [18], [19] or increasing working efficiency [20]. One research direction was using energy-aware approaches by optimizing the usage ratio to energy consumption on hardware side [21]–[23]. This approach can be achieved by increasing the performance of the computations in the clouds [24]–[26]. Another approach was optimizing voltage assignments to minimize the energy leakages [27]. The method of using loop scheduling optimization has been proved as an effective schema [28]. However, this research direction has limited concerns on user-centric issues.

Ontology has also been used in tele-health domains in recent years, which were explored by both researchers and practitioners [29], [30]. One direction is developing smart tele-health with using ontology. For instance, an ontology-based context-aware configurable system was proposed to gaining home-based continuous healthcare [31]. Another attempt was using ontological approach for telemedicin or emergency management in smart home [32], [33]. Various sensors were connected and deployed based on utilizing ontology-based approaches. Notwithstanding a lot of explorations in ontology, an investigation on creating new relations between physicians and PCT with using ontology in the clouds context was not done yet by the prior research.

In summary, this paper emphasizes the implementation of ontology in clouds and propose an approach of modeling new relations between patients and PCT with adopting ontology in cloud systems.

III. MOTIVATIONAL EXAMPLE

We describe a simple case of using our proposed model to explain the basic implementations in this section. The scenario illustrated in the motivational example is assisting patients to obtain a proper understanding of symptoms as well as the risks of treatments. The examined symptom is Acute Abdominal Pain (AAP) that is usually considered a common uncomfortable experience of stomach pain, such as gastritis. However, an arbitrary decision of self-diagnosis may result in missing the real pathogeny masquerading as the gastritis. In addition, the treatment plan based on the incorrect diagnosis will further put patients on a more dangerous position.

Our approach is using OWL to deliver integrated Web information with exact meanings of medical knowledge to patients and other users. The crucial part is identifying the relations between entities, which are represented by Classes in OWL. Fig. 1 represents an example of medical ontology showing a sharedAAPSytopOf ontology with using Resource Description Framework Schema (RDFS).
Our proposed model can identify the potential self-diagnosis by offering knowledge presentations. The relations and correspondences between entities are determined by processing Ontology Matching and the techniques are given in succeeding sections.

Listing 2: Example of OWL class construction combining different classes

IV. CONCEPTS AND THE PROPOSED MODEL

This section represents the operating principle of the proposed model as well as the main concepts in the model. Section IV-A represents the semantic techniques applied in the model, including theory and formula. Section IV-B describes the implementation of the proposed SOS model.

A. Semantic ontology matching

We use semantic Ontology Matching technique to generate new alignments from a group of ontologies. The matching process is an operation executing a function \( f \) that inputs an ontology set \( O \), an alignment \( A \), a parameter set \( P \), and a resource set \( R \) to output a new alignment \( A' \) corresponding the ontologies in \( O \) [34]. The mathematical expression can be formulated as follows.

\[
A' = f(O, A, P, R)
\]

\[
A' = \{ \forall i \in N, O_i \}
\]

In the formulation, we use a tuple to represent an ontology, \( O_i, E \) denotes the Classes set, \( I \) denotes the Individuals set, \( R \) denotes the Relations set, \( T \) denotes the Data Types set, and \( V \) denotes the Value set. An alignment refers to the interrelations between ontologies.

In order to output the new alignment, the critical component is creating the knowledge model for ontological semantics. The shared entities among different ontologies need to be corresponded. The process of corresponding ontologies is a Correspondence that is defined in Definition IV.1

**Definition IV.1** \( \exists \) a set of ontologies \( O \), a entity language set \( L \), a set of alignment relations \( R \), \( L = \{ L_i | L_1, L_2, ..., L_n \}, (i \in N) \); \( O = \{ O_i | O_1, O_2, ..., O_n \} \) \((i \in N)\); each \( O_i \) is associated with an entity language \( L_i \); \( L_i \in L \) \((i \in N)\), denoted as \( L_i(O_i) \). \( \exists \) two entities \( E_k \in O_k \subseteq O \) and \( E_j \in O_j \subseteq O \), an alignment relation set \( R_{k,j} \in R \), and \( r_{k,j} \in R_{k,j} \) between \( O_k \) and \( O_j \), denoted as \( R_{k,j} \Rightarrow C_{k,j} \subseteq \{O_k, O_j\} \), we define relation \( r_{k,j} \) holds both
ontologies $O_k$ and $O_j$. A Correspondence set is defined as $C = \{O, R\}, C = \{C_i|C_1, C_2, \ldots, C_n\}, (i \in N)$.

Definition IV.1 provides an operating principle asserting relations between ontologies. We mainly use a group of relations to interconnect different ontologies, such as owl:equivalentClass (=), rdfs:subClassOf (⊆), and owl:disjointWith (≠) [35]. Defining correspondences is a fundamental process for producing alignments. The definition of Alignments is formulated in Definition IV.2

**Definition IV.2.** Given an $O$ set, $O=\{O_1|O_2, \ldots, O_n\}$ ($i \in N$); an entity language set $L=\{L_1|L_2, \ldots, L_n\}$, ($i \in N$). $\exists C_i$ pairs two entities $E_j$ and $E_k$, $E_k \in O_k \subseteq O$ and $E_j \in O_j \subseteq O$, we define $A$ as an alignment over ontologies $C_k$ and $C_j$, $A \subseteq A$, $\forall C_i \in L_i(O_i)$, $\langle\{C_i\},=\rangle \in A$. We define the Total Alignment set $\Delta \cup \Delta$, if and only if $C_j=C_k \leftarrow (C_j, C',=) \in A \land (C_k, C',=) \in A$.

Definition IV.1 and IV.2 are the fundamentals of executing our proposed schema. Section IV-B describes the mechanism of applying these definitions.

**B. Secure Ontology-based Self-diagnosis (SOS) Model**

Our proposed model aims to provide online users with a secure approach for gaining secure information service offerings. Term Secure in SOS model refers to the security of the information itself, which can be also defined as minimizing the error rates and representing a knowledge presentation in a co-related manner.

SOS model uses Propositional Satisfiability (SAT) matching ontology, which is an efficient approach for semantic modeling [36]. The Axiom we used for SAT ontology matching is defined in Theorem IV.1.

**Theorem IV.1.** Given sets $A$ and $O$ with $C_i \in C$, and three random ontologies $O_m$, $O_n$, and $O_l$, $\{O_m, O_n, O_l\} \subseteq O$. $\exists$ random sets $r_{mn} \Rightarrow C_{mn}\equiv(O_m, O_n)$, $r_{nl} \Rightarrow C_{nl}\equiv(O_n, O_l)$. If $(O_m \equiv O_n)$, then it is true that $(C_{mn} \equiv C_{nl}) \land (\not\exists (C_{mn}, C_{nl},=) \in A)$. Denoted as $(\not\exists (C_{ml},=) \in A) \subseteq A$.

**Proof.** Given $A$, $\Delta$, and $O$ with $C_i \in C$, and $O_m$, $O_n$, and $O_l$, $\{O_m, O_n, O_l\} \subseteq O$. Assume there exists two relations $R_{mn} \Rightarrow C_{mn}\equiv(O_m, O_n)$, $R_{nl} \Rightarrow C_{nl}\equiv(O_n, O_l)$ and $(O_m \equiv O_n) \land (r_{mn} \equiv r_{nl})$. If it is true that $C_{mn} \neq C_{nl}$, then $\not\exists r_{mn} \in \Delta$ and $r_{nl} \in \Delta, \Delta \neq \Delta$. Conflict with one $\Delta$ fact. Proved.

Theorem IV.1 is used to formulate the ontology matching processes to find out all entities related to the target ontology with presenting the entities in a semantic means. Fig. 2 represents a manipulative process diagram of ontology matching that shows the output Alignment’ can be produced by an ontology matching with three main inputs, including Ontology $A$, Ontology $B$, and Alignment’. In the figure, State refers to the features of the input alignment located in the Before phase.

Parameters ($P$) and resources ($R$) are usually related to a few constraints, such as resources, languages, and properties. The constraints can be weighted for the purpose of the evaluation during the matching process. In After phase, a new Alignment’ is generated by which the relation types are determined. The output Alignment’ is a combination of a group of alignment candidates. The compliance measure can be accomplished by various approaches, which depends on the demands of the accuracy.

**V. ALGORITHMS**

This section describes the crucial algorithm, PSMA, used in our proposed model, which is designed for ontology match by using propositional techniques. The algorithm uses the Axioms given in Theorem IV.1. The inputs are a set of ontologies, $O$, and the corresponding alignments $\{A\}$. The selections of $P$ and $V$ need to be configured by the matching conditions and demands. The outputs are a set of new alignments, $A'$.

Considering the matching efficiency, we divide the target ontologies into a variety of groups based on the heterogeneous similarities of the entities in ontology. Using PSMA can separate inputs into different alignment candidates and converge the candidates at the end by which the new
alignment set is generated. The technique we use for solving the heterogeneous similarities is Weighting Product Model (WPM) [37]. Assume that there are two entities $E_j$ and $E_k$ in ontology $O$ and the number of comparing dimensions is $d$. We have $\delta(E_j, E_k) = \prod_{i=1}^{d} \delta_i(E_j, E_k)$ for $j, k \in N$. $w_i$ denotes for the weight of dimension $i$.

Algorithm VI.1 Propositional Satisfiability Matching Algorithm (PSMA)

Require: $\{O\}$, $\{A\}$, $P$, $R$

Ensure: $\{A^t\}$

1. Initialize datasets and input $\{O\}$, $\{A\}$, $P$, $R$
2. $A^T \leftarrow \emptyset$, $A^t_i \leftarrow \emptyset$
3. $\forall O_i \in \bigcap \exists P$ from $P$, $R$
4. /*Determine ontology ranks*/
5. Divide $\bigcap$ into a few groups of ontology sets $\bigcap_i^\bullet$
6. FOR $\forall \bigcap_i^\bullet \subseteq \bigcap$
7. Select candidate ontologies $O_a$, $O_b$
8. FOR $\forall O_i, O_j \in \bigcap_i^\bullet \{\{O_a, O_b\}, \{O_i, O_j\}\}$
9. Divide $\forall \forall \{O_i, O_j, O_a, O_b\}$ \{\{O_i, O_a\}, $O_b\}$ /*Partitioning*/
10. FOR $\forall (O_i, O_a), (O_j, O_b)$, till $d$
11. Composition using Theorem IV.1
12. Output alignment $A_t$, add $A_t$ to $A^T$
13. ENDFOR
14. Aggregate $A_T^T \Rightarrow A_T^\bigcap$, add it to $A^T$
15. $A^T_i \leftarrow \emptyset$
16. ENDFOR
17. ENDFOR
18. Aggregate $A^T \Rightarrow A^\bigcap$
19. $\{A^t\} \leftarrow A^\bigcap$ and RETURN $\{A^t\}$

Algorithm VI.1 represents the pseudocode of the proposed algorithm. The main phases of the algorithm include:

1) Input datasets, including ontology set, alignment set, parameter set, and resource.
2) Select parameters and resources and divide the ontology set into smaller sized ontology sets $\bigcap_i^\bullet$.
3) For each $\bigcap_i^\bullet$, select two candidate ontologies $O_a$, $O_b$. Use the candidate ontologies for semantic inference.
4) For each process of ontology matching, we divide the alignment as well as ontologies into different subsets, based on the deployed parameters.
5) Use the method defined in Theorem IV.1 and produce an alignment $A_t$ for a subset. Add $A_t$ to the temporary dataset $A^T$.
6) Aggregate all alignments in $A^T$ and produce $A^T$ and add it to a temporary dataset $A^T$.
7) Aggregate alignments in $A^T$ and produce output $\{A^t\}$.

VI. EXPERIMENT AND THE RESULTS

For the purpose of evaluating the experiment, we used compliance measures to examine the performance of compliance for the proposed schema. The performance of solving the heterogeneous similarities using WPM has been examined.

For processing these evaluations, we used a reference alignment $A_{ref}$. We configured four experimental settings that were associated with four correspondences, denoted as $\{C_1|C_{ref}, C_2|C_{ref}, C_3|C_{ref}, C_4|C_{ref}\} \subseteq A_{ref}$. The target evaluated alignments set was $A_e$, which was represented as $A_e = \{A_1|A_2, A_3, A_4\}$. Each each $A_1|A_2$, $\exists(C_1|C_{ref}, C_2|C_{ref}) \Rightarrow A_3$. Fig. 3 represents some experimental results concerning the heterogeneous similarities solvers.

According to the display of Fig. 3, the values of the bars determined the relations between correspondences. Each setting consisted of two bars showing the comparison results of $C_1$ and $C_2$. The figure depicts the following experimental results: $C_1^{12} \geq C_1^{11} \geq C_1^{1}$, $C_2^{22} \geq C_2^{2}$, $C_3^{22} \geq C_3^{3}$, and $C_4^{22} \geq C_4^{4}$. The outputs of the experiments matched the experimental configurations and the desired outcomes.

Moreover, three criteria were used in our experiment, including Precision, Recall, and F-Measure [38]. We defined the Precision as $P^e = |A_e \cap A_{ref}|/|A_e|$, Recall as $R^e = |A_e \cap A_{ref}|/|A_{ref}|$, and F-Measure as $FM = (P^e \times R^e)/((1 - \alpha) \times P^e + \alpha \times R^e)$, and $F.M$. Fig. 4 exhibits a group of experimental results about $P^e$, $R^e$, and $F.M$. The figure illustrated that our proposed schema had a high accuracy as
the smaller value of $FM$ is associated with a higher-level accuracy.

VII. CONCLUSIONS

This paper proposed a new schema modeling relations between medical knowledge entities using ontological approaches in cloud systems. The proposed solution was SOS model that secured the accuracy and scope of knowledge presentation in ontology-based OSD systems. The critical algorithm is PSMA that was designed to execute an efficient matching process. Our experimental research had proved the performance of the proposed model.

REFERENCES