Ontology Mapping Constructing by means of Low Rank Distance Matrix Optimization

Caihua Qiu

Abstract—**Ontology mapping, based on similarity calculation, aims to find the similar concepts among different ontologies. In order to mathematically represent concepts, it is common to represent all information of a concept in a fixed dimensional vector. Therefore, the similarity calculation can be converted into a distance calculation between vectors, and the smaller the distance is, the larger the similarity will be. In this paper, the low-rank matrix learning strategy is used to obtain the corresponding ontology mapping strategy. The core idea of the method is to control the upper bound of the distance of the similar vertex pairs in the sample and the lower bound of the distance of dissimilar vertex pairs. At the same time, the rank of the matrix is integrated into the optimization conditions. The effectiveness of the proposed ontology trick is illustrated by the construction of ontology mapping on three ontology data.**

Index Terms—**ontology, similarity measure, distance computing, ontology mapping.**

I. Introduction

THE term ontology derives from philosophy and is used to describe the essential associations between things. After being introduced into the computer field, as a data structure representation model, ontology has been widely used in various fields of computer science. The research of ontology similarity calculation and ontology mapping algorithm has become one of the core contents in the field of knowledge representation. At the same time, as a conceptual structure, the ontology has been widely used in many fields such as biomedicine, geography, physics, and social sciences. For ontology mapping, the essence is to calculate the similarity between concepts from different ontology. Therefore, the ontology similarity calculation algorithm can become an ontology mapping algorithm by appropriate conversion.

In information retrieval, the ontology vertices represent concepts, and the edges represent the interrelationships between concepts. By determining the parameter C by the domain expert, the concept set corresponding to all the vertices B satisfying $Sim(A,B) > C$ is returned to the user as a query extension of the concept corresponding to the vertex A. For the ontology mapping, the graphs $G_1, G_2, \cdots, G_m$ correspond to the bodies $O_1, O_2, \cdots, O_m$, respectively. Each $A \in V(G_i)$, where $1 \leq i \leq m$, finds all sets of corresponding concepts of vertex B satisfying $Sim(A,B) > C$ in $G - G_i$ and returns to the user as a query extension of the concept corresponding to vertex A.

The ontology is used as a tool in various fields. For example, in the biological field, "GO" Ontology (http://www.geneontology.org) contains information on cellular components, molecular functions and biological processes, and contains about 23,700 terms, and more than 16 million in about 20 biological databases. The gene is annotated. Analysis of the ontology can help biologists understand the interrelated features of genes between different biological databases.

In recent years, ontology has been applied to various fields. Przydzial [1] applied the ontology to protein retrieval in pharmaceuticals. Koehler et al. [2] applied ontology to a database of characterization between molecules and diseases. Ivanovic and Budimac [3] reviewed the application of ontology in the medical field. Hristoskova [4] applied the ontology to the creation of a personal care system. Kabir et al. [5] used the ontology data structure to establish an effective social information management platform. Ma et al. [6] proposed an ontology model framework based on stable semantic retrieval. Li et al. [7] obtained a new ontology data representation model and applied it to the customer shopping system. Santodomingo et al. [8] proposed a matching system for expert knowledge in the ontology domain. Pizzuti et al. [9] innovated the food ontology and gave several practical applications of the ontology. Lasierra et al. [10] proposed that the ontology can be applied to the design of buildings and applied to the design and maintenance of the patient's home. Carlini and Makowski [11] applied the genetic GO ontology to the study of preferred cryptographic words in insect homology. Nicolai [12] elaborated on the theory of deflation and its ontology representation. Correae et al. [13] incorporated ontology methods into annotated scientific file systems based on modular technology and applied them in the field of drug and infectious disease control. Duran-Limon et al. [14] proposed an ontology-based product architecture derivation method. Chabot et al. [15] obtained a reconstruction and analysis method for the timetable of digital events based on ontology technology. Elbers and Taylor [16] gave gene workflow algorithms based on ontology technology and apply them to target region generation in targeted enrichment experiments. Rani et al. [17] obtained an ontology-based adaptive personalized learning system that was simultaneously implemented by cloud storage technology through software agents. Sangiacomo [18] studied the role of ontology in behavioral decision systems and obtained several results. Azevedo et al. [19] built a model for analyzing the resources and capabilities of enterprise architecture modeling under the ontology framework. Wimmer and Rada [20] established an ontology-based analysis and evaluation algorithm to determine the quality of information. Trokanas and Cecelja [21] established an ontology evaluation system in the field of chemical process systems engineering. Chhun et al. [22] established QoS ontology for service selection and reuse. The focus of this ontology is to evaluate the

Manuscript received December 07, 2019; revised April 04, 2020. This work was supported in part by General project of social science and technology development in Dongguan 2018 (no. 20185071541236) and Scientific Research Projects of Guangdong University of Science & Technology in 2018 (no. GKY-2017KYYB-6).

C. Qiu is with the Department of Computer Science, Guangdong University of Science and Technology, Dongguan, China, and Bulacan State University, Philippines. Email: qiucaihua_gk@sina.com.

quality of the service. Costa et al. [23] constructed the corresponding ontology algorithm for knowledge sharing and reuse in the field of architecture. Panov et al. [24] established the gene ontology OntoDT for the representation of data type knowledge. Kutikov et al. [25] established a urological label ontology study for the standardization of social media communication. Grandi [26] studied the multi-version ontology personalized dynamic class hierarchy management technology and introduced a storage scheme, which allows the representation and management of the temporal relational database and the evolution hierarchy of a multi-version ontology. Kontopoulos et al. [27] proposed an ontology-based decision support tool and applied it to the optimization of solar water heating systems. Hoyle and Brass [28] presented the theory of statistical mechanics in the process of annotating objects, the terminology of which is derived from the ontology domain. Solano et al. [29] proposed ontology technology for integrated processing and detection processes, and the ontology used is focused on the evaluation of resource capabilities. Aime and Charlet [30] applied social psychology knowledge to ontology engineering.

However, most of the current ontology engineering application algorithms are heuristics that design similarity calculation formula by experience. The disadvantages of these heuristic ontology similarity calculation model are: (1) A large number of parameters require experienced domain experts to decide in actual operation, and the quality of the domain experts can directly determine the success or failure of the algorithm. (2) It is not intuitive and requires a lot of calculations. It requires detailed analysis of the ontology structure and related features in advance, and requires intuitive experience to accurately reflect the key information of the ontology.

In addition, the defects of various ontology learning algorithms are obvious. For example, the ontology learning algorithm based on the sorting method has its obvious advantages, that is, it intuitively maps the entire ontology structure to the real number axis, and each concept corresponds to a real number on the axis. At the same time, because the algorithm only focuses on the relative sizes of different concepts under the ontology ranking function, it does not depend on label data and is very suitable for unsupervised learning.

II. Ontology setting

We assume that all information related to the ontology concept is encapsulated in a d-dimensional vector, which includes conceptual information and location information of the corresponding ontology vertices and neighborhood information in the ontology graph. For convenience of discussion, v is used to represent the ontology concept itself, the corresponding vertices on the ontology graph, and the corresponding d-dimensional vectors. In the following article, we no longer specifically point out that v is a vector, and it also represents a vertex in a particular statement. This slightly confusing writing does not affect the correct understanding of our article.

The ultimate goal of the ontology algorithm is to obtain the similarity between the vertices. Therefore, it can be understood that the goal of the ontology algorithm is to obtain a similarity matrix whose i-th row and j-th column elements correspond to the similarity between the vertex v_i and the vertex v_j. When the information of the corresponding vertices of the ontology concept is vectorized, their similarities can be considered from the perspective of geometric distance. That is, the smaller the geometric distance between the two vertices is, the greater the similarity will be. In this sense, the similarity matrix can be obtained by a distance matrix (the distance between the vertex v_i and the vertex v_j of the i-th row and the j-th column element of the matrix), or the degree of similarity of each pair of vertices can be directly measured by the distance matrix. The greater the distance is, the smaller the similarity becomes.

One type of existing method is to obtain the distance matrix by learning the Mahalanobis matrix, and thereby determine the degree of similarity between the vertices. Let $\mathbf{M} = \mathbf{P}^T\mathbf{P}$ the distance matrix derived for the transformation matrix P, and the distance between the vertices v_i and v_j in the ontology and $\mathbf{M}$ is calculated as follows:

$$d^2_{\mathbf{M}}(v_i, v_j) = (v_i - v_j)^T\mathbf{M}(v_i - v_j),$$

where $v_i = (v_i^1, \cdots, v_i^d) \in \mathbb{R}^d$, that is, the information corresponding to each vertex is represented by a d-dimensional vector. The role of the Mahalanobis matrix is essentially the extraction of data and the extraction of feature information. An important research content is how to get $\mathbf{M}$ better, and analyze the complexity, convergence order and statistical error of the algorithm.

For the ontology distance matrix learning, what we have to consider is how to get a good similarity matrix or distance matrix. As a good distance matrix, we measure it from the following two aspects: First, $\mathbf{M}$ should retain the structure of the ontology graph. That is to say, the distance between the similar vertices should be smaller than the distance of dissimilar vertices; secondly, after the vector information of the ontology is vectorized, the vector corresponding to each ontology vertex contains a large amount of information such as the name, instance, and neighborhood of the ontology vertex, but what is really relevant to an application area is some of its special components. Therefore, the good distance matrix $\mathbf{M}$ should be able to effectively remove noise during the dimensionality reduction process.

Recently, lager amount of learning tricks are applied in the ontology similarity computing and ontology mapping, and also several paper contribution in the theoretical analysis. Gao and Farahani [31] studied the generalization bounds and uniform bounds with convex ontology loss function in multi-dividing ontology setting. Gao et al. [32] raised partial multi-dividing ontology learning algorithm. Gao and Xu [33] presented the stability analysis of learning algorithms for ontology similarity computation. More ontology learning algorithms and analysis can be referred to [34], [35], [36], [37] and [38].

In this work, we present a new ontology learning algorithm for searching the similar concepts among different ontologies using the tricks of vector distance learning. The main ontology learning algorithm is based on the low rank matrix optimization method in which the smaple ontology vertex pairs are divided into two parts: similarity and dissimilarity; restrict conditions are designed in the different setting of ontology problem and finally we show the effectiveness of the new approach by comparing experimental data.

III. Ontology algortihm description using low rank matrix optimization

Assume that given an otnology data set with n vertices $V = \{v_i\}_{i=1}^n \subseteq \mathbb{R}^d$ and there are two sets of pairwise constraints connect with these ontology data vertices:

$$S = \{(i,j)|v_i \quad \text{and} \quad v_j \quad \text{are judged to be similar}\},$$
$$D = \{(i,j)|v_i \quad \text{and} \quad v_j \quad \text{are judged to be dissimilar}\}. \tag{1}$$

Here S is the set of constraints on ontology similar vertices, and D is the set of constraints on ontology dissimilar vertices. Each pairwise constraint (i,j) reveals that if two ontology vertices v_i and v_j are similar or dissimilar judged by filed experts. Notices that it is not necessary for all pair of vertices in V to be included in $S \cup D$.

Let $d(v_i, v_j)$ be the distance between two ontology vertices v_i and v_j, and $\mathbf{M} \in \mathbb{R}^{d\times d}$ be a symmetric metric matrix. With $\mathbf{M}$, the ontology distance function can be stated as follows:

$$d_{\mathbf{M}}(v_i, v_j) = \|v_i - v_j\|_{\mathbf{M}} = \sqrt{(v_i - v_j)^T \mathbf{M}(v_i - v_j)}. \tag{2}$$

Specifically, since the symmetric matrix $\mathbf{M}$ needs to be a valid metric, we assume that it satisfies the triangle inequality conditions and also a semidefinite positive matrix with $\mathbf{M} \succeq 0$. Usually, $\mathbf{M}$ is a Mahalanobis matrix with distance computing in the vector space $\mathbb{R}^d$ and the ontology distance function degenerates into standard Euclidean distance formula when $\mathbf{M}$ set to be identity matrix $\mathbf{I} \in \mathbb{R}^{d\times d}$. Originally, our goal was to learn the ontology distance function, but now by means of (2) the aim is transformed to get an optimal ontology symmetric distance matrix $\mathbf{M} \in \mathbb{R}^{d\times d}$ in light of a set of ontology data vertices X on a vector space $\mathbb{R}^d$ together with a set of constraints on ontology similar vertices S and a set of constraints on ontology dissimilar vertices D, which can be stated as the following ontology optimization model

$$\min_{\mathbf{M}\succeq 0} g(\mathbf{M}, V, S, D) \tag{3}$$

where $\mathbf{M}$ is a positive semidefinite matrix and g is a suitable ontology objective function defined over the given ontology vertices and constraints.

Now we discuss the principle for the ontology distance matrix learning, which directly affects ontology learning frameworks and algorithms. First principle is to minimize the distances among the ontology data vertices with ontology similar constraints and maximize the distances among the ontology data vertices with ontology dissimilar constraints. Using this principle, ontology objective function g can simply expressed as

$$g(\mathbf{M}) = \sum_{(i,j)\in S} \|v_i - v_j\|_{\mathbf{M}}^2 - \gamma \sum_{(i,j)\in D} \|v_i - v_j\|_{\mathbf{M}}^2, \tag{4}$$

where γ is a positive balance parameter. To get the ontology objective function g, we always assume that $\mathbf{M}$ is decomposable, i.e., there is a matrix $\mathbf{W} = (\mathbf{w}_1, \cdots, \mathbf{w}_p) \in \mathbb{R}^{d\times p}$ corresponding to a linear map: $\mathbf{W}^T : \mathbb{R}^d \to \mathbb{R}^p$ and satisfying $\mathbf{M} = \mathbf{W}\mathbf{W}^T$ and $\mathbf{w}_1, \cdots, \mathbf{w}_p$ is not linear dependent (equally, the rank of matrix $\mathbf{M}$ is p). In this way, we have

$$\begin{aligned} & \|v_i - v_j\|_{\mathbf{M}} = \sqrt{(v_i - v_j)^T \mathbf{M}(v_i - v_j)} \\ = & \sqrt{(v_i - v_j)^T \mathbf{W}\mathbf{W}^T (v_i - v_j)} = \|\mathbf{W}^T (v_i - v_j)\|. \end{aligned}$$

Through this transformation, the original $\mathbf{M}$ norm calculation is converted to the simplest Euclidean norm.

How to understand this transformation? We look at the following conversion. For two vectors (ontology vertices) v_1 and v_2, the most primitive method is to use cosine to represent similarity, i.e., $cos(v_1, v_2) = \frac{v_1^T v_2}{\|v_1\|\cdot\|v_2\|}$. In the setting that v_1 and v_2 are normalized to endow the standard L_2 norm, then the cosine similarity is equivalent to the standard Euclidean distance which can also be denoted as $d(v_1, v_2) = 2 - 2cos(v_1, v_2)$. The function of matrix $\mathbf{W}$ is like a projection map for vector denoted by $\vartheta(v) = \mathbf{W}^T v$, and the above transformation process can be re-described as

$$\begin{aligned} & d^2(\vartheta(v_1), \vartheta(v_2)) \\ = & d^2(\mathbf{W}^T v_1, \mathbf{W}^T v_2) \\ = & (\mathbf{W}^T v_1 - \mathbf{W}^T v_2)^T (\mathbf{W}^T v_1 - \mathbf{W}^T v_2) \\ = & d_{\mathbf{M}}^2(v_1, v_2) = (v_1 - v_2)^T \mathbf{W}\mathbf{W}^T (v_1 - v_2). \end{aligned}$$

When it comes to semi-supervised ontology learning setting which contains the unlabeled ontology data, we should assume an affinity matrix $\mathbf{A} \in \mathbb{R}^{n\times n}$ on V where each element A_{ij} reveals the measure of affinity between two ontology vertices v_i and v_j in the following way: the larger A_{ij} is, the large similarity betwwen v_i and v_j becomes. With the affinity matrix $\mathbf{A}$, set $\mathrm{T}(\cdot)$ is the trace operator, $\mathbf{D}$ is a diagonal matrix with $D_{ii} = \sum_{j=1}^n A_{ij}$ and $\mathbf{L} = \mathbf{D} - \mathbf{A}$ is the Laplacian matrix of ontology graph, then the function in (4) can be stated as

$$\begin{aligned} & g(\mathbf{M}, V, S, D) = \frac{1}{2}\sum_{i,j=1}^n \|v_i - v_j\|_{\mathbf{M}}^2 A_{ij} \\ = & \frac{1}{2}\sum_{i,j=1}^n \|\mathbf{W}^T (v_i - v_j)\|^2 A_{ij} \\ = & \sum_{r=1}^p \mathbf{w}_r^T \mathbf{V}(\mathbf{D} - \mathbf{A})\mathbf{V}^T \mathbf{w}_r = \sum_{r=1}^p \mathbf{w}_r^T \mathbf{V}\mathbf{L}\mathbf{V}^T \mathbf{w}_r \\ = & \mathrm{T}(\mathbf{W}^T \mathbf{V}\mathbf{L}\mathbf{V}^T \mathbf{W}) = \mathrm{T}(\mathbf{V}\mathbf{L}\mathbf{V}^T \mathbf{W}\mathbf{W}^T) \\ = & \mathrm{T}(\mathbf{V}\mathbf{L}\mathbf{V}^T \mathbf{M}), \end{aligned} \tag{5}$$

where $A_{ij} = 1$ if $(i,j) \in S$ and $A_{ij} = -1$ if $(i,j) \in D$ in the supervised ontology learning setting, while in the semi-supervised ontology setting, $A_{ij} = 1$ if v_i is among k nearest neighbors of v_j and vice versa. However, due to its obvious defects, we always need to improve its expression of $\mathbf{A}$.

We aim to obtain better affinity matrix $\mathbf{A}$ in semi-supervised ontology learning settings. By means of weak affinities followed by k-NN trick (denote $N(v_i)$ as the k nearest neighbors of vertex v_i), the strong affinities by the given pairwise ontology vertex constraints to the global wide of the ontology data. The neighborhood symmetric matrix $\mathbf{P} \in \mathbb{R}^{n\times n}$ for ontology data V is denoted as $P_{ij} = 1/k$ if $j \in N(v_i)$ and $P_{ij} = 0$ otherwise.

Let $\mathbf{A}^0 \in \mathbb{R}^{n\times n}$ be the initial affinity matrix with $A_{ij}^0 = 1$ for any $(i,j) \in S$, $A_{ij}^0 = -1$ for any $(i,j) \in D$, and $A_{ij}^0 = 0$ for any other cases. Let $t \in \mathbb{N}^+ \cup \{0\}$ be a counting parameter, $\mathbf{A}^t$ be an affinity matrix in the t-th iteration, $\mathbf{A}_i^t$ be the i-th row of $\mathbf{A}^t$, and $\lambda \in (0,1)$ be a balance parameter. Simply, the whole iteration processure can be stated as

$$\mathbf{A}_i^{t+1} = (1-\lambda)\mathbf{A}_i^0 + \lambda\sum_{j=1}^n P_{ij}\mathbf{A}_j^t, \tag{6}$$

and furthermore

$$\mathbf{A}^{t+1}=(1-\lambda)\mathbf{A}^{0}+\lambda\mathbf{P}\mathbf{A}^{t}. \quad (7)$$

Since all eigenvalues of $\mathbf{P}$ located between -1 and 1, and the limitation $\mathbf{A}^{*}=\lim_{t\to\infty}\mathbf{A}^{t}$ exists, we yield

$$\mathbf{A}^{*}=(1-\lambda)(\mathbf{I}-\lambda\mathbf{P})^{-1}\mathbf{A}^{0}. \quad (8)$$

The final ontology affinity matrix $\mathbf{A}$ is denoted by $\mathbf{A}_{ij}=\frac{\mathbf{A}^{*}_{ij}+\mathbf{A}^{*}_{ji}}{2}$ if $|\mathbf{A}^{*}_{ij}+\mathbf{A}^{*}_{ji}\geq c$ where parameter c meets $0<c<2$; and $\mathbf{A}_{ij}=0$ otherwise.

Let $\mathcal{S}^{d}_{+}$ be the subspace of positive semidefinite matrices in $\mathbb{R}^{d\times d}$, $\|\mathbf{M}\|_{0}$ be the number of nonzero components in $\mathbf{M}$, and $\|\mathbf{M}\|_{1}$ be the sum of absolute values of all components in $\mathbf{M}$. The inner product operator of matrix is defined as $<\mathbf{A},\mathbf{B}>=T(\mathbf{A}^{T}\mathbf{B})$. Let Σ be the empirical covariance matrix and ρ be a positive balance parameter. By means of log determinant measure, the corresponding ontology optimization problem can be stated as

$$\max_{\mathbf{M}\in\mathcal{S}^{d}_{+}}\log\det\mathbf{M}-<\Sigma,\mathbf{M}>-\rho\|\mathbf{M}\|_{0}, \quad (9)$$

where the last part $\rho\|\mathbf{M}\|_{0}$ is used to control the sparsity of $\mathbf{M}$. We can use $\|\|_{1}$ norm of matrix and then (9) becomes

$$\max_{\mathbf{M}\in\mathcal{S}^{d}_{+}}\log\det\mathbf{M}-<\Sigma,\mathbf{M}>-\rho\|\mathbf{M}\|_{1}, \quad (10)$$

which equals to

$$\min_{\mathbf{M}\in\mathcal{S}^{d}_{+}}-\log\det\mathbf{M}+<\Sigma,\mathbf{M}>+\rho\|\mathbf{M}\|_{1}. \quad (11)$$

Set $\Sigma=\mathbf{M}_{0}^{-1}+\rho_{1}\mathbf{V}\mathbf{L}\mathbf{V}^{-1}$ with a positive parameter ρ_{1}. Note that the sparsity part of (11) can be reduced to the condition part, which leads to a following max-min ontology learning problem

$$\min_{\mathbf{M}\in\mathcal{S}^{d}_{+}}\max_{\|\mathbf{W}\|_{\infty}\leq\rho}-\log\det\mathbf{M}+<\Sigma+\mathbf{W},\mathbf{M}>. \quad (12)$$

In light of exchaning max and min parts in (12), the dual problem of (11) can be expressed as

$$\max_{\|\mathbf{W}\|_{\infty}\leq\rho}\min_{\mathbf{M}\in\mathcal{S}^{d}_{+}}-\log\det\mathbf{M}+<\Sigma+\mathbf{W},\mathbf{M}>, \quad (13)$$

and it equals to

$$\begin{aligned}&\max-\log\det\Upsilon+d\\ &\|\Upsilon-\Sigma\|\leq\rho.\end{aligned} \quad (14)$$

Set $h(\mathbf{M})=\rho\|\mathbf{M}\|_{1}$ and

$$\chi(\mathbf{M})=-\log\det\mathbf{M}+<\Sigma,\mathbf{M}>. \quad (15)$$

Hence, ontology problem (11) can be re-written by

$$\min\chi(\mathbf{M})+h(\mathbf{M}). \quad (16)$$

Let κ be a smoothness parameter, $\|\cdot\|_{F}$ be the Frobenius norm and $h_{\kappa}(\mathbf{M})$ be the smooth approximation of $h(\mathbf{M})$ which is formulated as

$$h_{\kappa}(\mathbf{M})=\max_{\mathbf{W}:\|\mathbf{W}\|_{\infty}\leq\rho}\{<\mathbf{W},\mathbf{M}>-\frac{\kappa}{2}\|\mathbf{W}\|^{2}_{F}\}. \quad (17)$$

Therefore, ontology learning framewrok (16) can be represented as

$$\min\chi(\mathbf{M})+h_{\kappa}(\mathbf{M}). \quad (18)$$

The simple computing processes for ontology optimization (18) can be described as follows: at beginning set $\mathbf{M}=\Phi^{0}$ and $t=0$; do loop $\mathbf{M}^{t+1}=\arg\min_{\mathbf{M}}\chi(\mathbf{M})+h_{\kappa}(\Phi^{t})+<\bigtriangledown h_{\kappa}(\Phi^{t}),\mathbf{M}-\Phi^{t}>+\frac{\max\{L(\chi),L(h_{\kappa})\}}{2}\|\mathbf{M}-\Phi^{t}\|^{2}_{F}$, $\Phi^{t+1}=\arg\min_{\Phi}\chi(\mathbf{M}^{t+1})+<\bigtriangledown\chi(\mathbf{M}^{t+1}),\Phi-\mathbf{M}^{t+1}>+\frac{\max\{L(\chi),L(h_{\kappa})\}}{2}\|\Phi-\mathbf{M}^{t+1}\|^{2}_{F}+h_{\kappa}(\Phi)$, $t=t+1$ until $\mathbf{M}^{t}$ converges; and finally return $\mathbf{M}=\mathbf{M}^{t}$.

Let $\mathbf{M}_{0}$ be the initial matrix as defined before, and $D(\mathbf{M}|\mathbf{M}_{0})$ be a non-negative convex logarithmic determinant objective function. For instance, $D(\mathbf{M}|\mathbf{M}_{0})=T(\mathbf{M}\mathbf{M}_{0}^{-1})-\log|\mathbf{M}\mathbf{M}_{0}^{-1}|-d$. Using two positive parameters c_{1} and c_{2}, the ontology optimization model can be stated as

$$\begin{aligned}\min_{\mathbf{M}}\quad & D(\mathbf{M}|\mathbf{M}_{0})\\ \text{s.t.}\quad & d_{\mathbf{M}}(v_{i},v_{j})\leq c_{1}\quad(i,j)\in S\\ & d_{\mathbf{M}}(v_{i},v_{j})\geq c_{2}\quad(i,j)\in D.\end{aligned} \quad (19)$$

In the low rank Mahalanobis distance setting, the corresponding convert can be formulated as

$$\begin{aligned}&d^{2}_{\mathbf{I}+\mathbf{M}}(v_{1},v_{2})\\ =\;&(v_{1}-v_{2})^{T}(\mathbf{I}+\mathbf{M})(v_{1}-v_{2})\\ =\;&(v_{1}-v_{2})^{T}(v_{1}-v_{2})+(v_{1}-v_{2})^{T}\mathbf{M}(v_{1}-v_{2})\\ =\;&d^{2}(v_{1},v_{2})+d^{2}_{\mathbf{M}}(v_{1},v_{2}),\end{aligned}$$

where $\mathbf{M}$ is low rank. Since the rank of both $\mathbf{M}$ and $\mathbf{W}$ is p, the low-rank ontology distance learning framework can be stated as

$$\begin{aligned}\min_{\mathbf{M}}\quad & D(\mathbf{M}|\mathbf{W}\mathbf{W}^{T})\\ \text{s.t.}\quad & d_{\mathbf{M}}(v_{i},v_{j})\leq c_{1}\quad(i,j)\in S\\ & d_{\mathbf{M}}(v_{i},v_{j})\geq c_{2}\quad(i,j)\in D\\ & r(\mathbf{M})\leq p.\end{aligned} \quad (20)$$

Let $\Xi=\mathbf{W}(\mathbf{W}^{T}\mathbf{W})^{-\frac{1}{2}}$ be an orthogonal expression of the columns of $\mathbf{W}$. Hence, we have $\Xi^{T}\Xi=\mathbf{I}$ and $\Xi\Xi^{T}\mathbf{M}\Xi\Xi^{T}-\Xi\Xi^{T}\mathbf{M}_{0}\Xi\Xi^{T}=\Xi\Xi^{T}(\mathbf{M}-\mathbf{M}_{0})\Xi\Xi^{T}$. The ontology distance metric learning problem can be summaried as to learn a full rank Mahalanobis matrix with restrict $\mathbf{M}=\mathbf{I}+\Xi\Xi^{T}(\mathbf{M}-\mathbf{I})\Xi\Xi^{T}$ and this condition can be relaxed to more generalized $\mathbf{M}_{0}$ where $\mathbf{M}=\mathbf{I}+\Xi\Xi^{T}(\mathbf{M}-\mathbf{M}_{0})\Xi\Xi^{T}$. The corresponding ontology distance learning problem becomes

$$\begin{aligned}\min_{\mathbf{M}}\quad & D(\mathbf{M}|\mathbf{I})\\ \text{s.t.}\quad & d_{\mathbf{M}}(v_{i},v_{j})\leq c_{1}\quad(i,j)\in S\\ & d_{\mathbf{M}}(v_{i},v_{j})\geq c_{2}\quad(i,j)\in D\\ & \mathbf{M}=\mathbf{I}+\Xi\Xi^{T}(\mathbf{M}-\mathbf{I})\Xi\Xi^{T}.\end{aligned} \quad (21)$$

It is easy to check that

$$\begin{aligned}d_{\mathbf{M}}(v_{i},v_{j})&=d_{\mathbf{I}+\mathbf{M}'}(v_{i},v_{j})=d_{\mathbf{I}}(v_{i},v_{j})+d_{\mathbf{M}'}(v_{i},v_{j})\\ &=d_{\mathbf{I}}(v_{i},v_{j})-d_{\Xi\Xi^{T}}(v_{i},v_{j})+d_{\Xi\Xi^{T}\mathbf{M}\Xi\Xi^{T}}(v_{i},v_{j}),\end{aligned}$$

where $\mathbf{M}'$ is a low rank matrix with its rank at most p. Thus, set third marginal parameter $c_{ij}=d_{\mathbf{I}}(v_{i},v_{j})-d_{\Xi\Xi^{T}}(v_{i},v_{j})$, the ontology problem (21) can be re-stated as

$$\begin{aligned}\min_{\mathbf{M}}\quad & D(\mathbf{M}|\mathbf{I})\\ \text{s.t.}\quad & d_{\Xi\Xi^{T}\mathbf{M}\Xi\Xi^{T}}(v_{i},v_{j})\leq c_{1}-c_{ij}\quad(i,j)\in S\\ & d_{\Xi\Xi^{T}\mathbf{M}\Xi\Xi^{T}}(v_{i},v_{j})\geq c_{2}-c_{ij}\quad(i,j)\in D\\ & \mathbf{M}=\mathbf{I}+\Xi\Xi^{T}(\mathbf{M}-\mathbf{I})\Xi\Xi^{T}.\end{aligned} \quad (22)$$

The optimal solution of ontology probelm (23) meets $\mathbf{M}^* = \mathbf{I} + \Xi\Xi^T(\mathbf{M}^* - \mathbf{I})\Xi\Xi^T$, and it can be further rewritten as

$$\begin{array}{lll} \min\limits_{\mathbf{M}} & D(\mathbf{M}|\mathbf{I}_p) & \\ \text{s.t.} & d_{\mathbf{M}}(\Xi^T v_i, \Xi^T v_j) \le c_1 - c_{ij} & (i,j) \in S \\ & d_{\mathbf{M}}(\Xi^T v_i, \Xi^T v_j) \ge c_2 - c_{ij} & (i,j) \in D. \end{array} \quad (23)$$

The whole iterative process can be described as follows: given slack punish parameter ι, set $\Lambda = \mathbf{W}$, for any i and j, $\psi_{ij} = \omega = 0$; in each iterate, label $\tau = 1$ if similarity constrict and $\tau = -1$ if dissimilarity constrict, $d = (v_i, v_j)^T \Lambda^T \Lambda (v_i, v_j)$, $\zeta = \min\{\psi_{ij}, \frac{\tau\iota}{\iota+1} + \frac{1}{d} - \frac{1}{\omega_{ij}}\}$, $\varpi = \frac{\tau\zeta}{d\tau\zeta+1}$, $\psi_{ij} = \psi_{ij} - \zeta$, $\omega_{ij} = \frac{\omega_{ij}\iota}{\iota+\tau\zeta\omega_{ij}}$,$\Lambda = \Lambda + (\sqrt{\varpi+1}-1)(v_i - v_j)(v_i - v_j)^T\Lambda$; finally return $\mathbf{M} = \Lambda\Lambda^T$.

IV. Experiments

The main purpose of this part is to test the validity of our proposed distance computing based ontology learning algorithm with respect to ontology mapping structure. We will test on three databases: "University" ontologies, "Mathematics" ontologies and "Chemical Index" ontologies. Since in these three ontologies, the total number of vertices are small, so we use the ontology sample with small capacity, but it doesn't affect the validity of the test.

A. Experiment on "University" data

We use "University" ontologies O_1 and O_2 in the first experiment which is used to describe the basic organizational structure of the university. The structures of O_1 and O_2 are respectively presented in Figure 1 and Figure 2. We set the experiment with the aim to give ontology mapping between O_1 and O_2. We take $P@N$ precision ratio as a criterion to measure the quality of the experiment, and ontology algorithms in Gao et al. [40], [43] and [45] on "University" ontologies. Then we compare the precision ratios yielded from the three ontology learning tricks, and some results can be referred to Table 1.

We can see from the table 1 by comparing the average precision ratios obtained from the three algorithms and our newly proposed one that when $N = 1$, 3 or 5, the precision ratios by means of our new ontology mapping algorithms are much higher than that got from ontology learning algorithms proposed in Gao et al. [40], [43] and [45]. Generally speaking, the average ratio grows apparently as N increases, no matter in any algorithm here. When $N = 3$, our new algorithm shows most advantage over the others. As a result, our distance computing based ontology learning algorithms turn out to be better and more effective than those raised by Gao et al. [40], [43] and [45].

B. Experiment on "Mathematics" data

We use mathematical ontologies O_3 and O_4 (Figures 3 and 4 present the basic structures of O_3 and O_4, respectively) for our second experiment to test the availability of new proposed distance computing based algorithm with regard to ontology mapping constructing. These ontologies is used to describe the relationship between different research branches in different disciplines of mathematics. The aim is to compute the similarity-based ontology mapping between O_3 and O_4 using our proposed distance computing based algorithm. The $P@N$ criterion is also applied as criterion to test the experiment data. Ontology algorithms introduced in Gao and Zhu [39], Gao et al. [41], and Wu et al. [44] are also acted on mathematical ontologies, and the precision ratios are compared among four methods. Partial of experiment results can be referred to Table 2.

According to the experiment data presented in Table 2, it's easy to find that when N increases the average precision ratios show clear tendency to increase with it. Generally speaking, the advantage of our new algorithm is obvious by comparing the ratios. Even though the ratio from algorithm in Wu et al. [39] is the same as that from the new one when $N = 1$, when N becomes larger, the advantage of our new algorithm becomes more obvious. As a result,our distance computing based ontology approach performances much more efficient than ontology learning methods studied in Gao and Zhu [39], Gao et al. [41], and Wu et al. [44], especially when N is large enough.

C. Experiment on chemical index data

The "Chemical Index" ontologies O_5 and O_6 (the basic structures of these two ontologies can be referred to Figures 5 and 6, respectively) for the third experiment. The purpose is to infer similarity-based ontology mapping between O_5 and O_6 in lingt of our distance computing based algorithm. The $P@N$ criterion is also used to measure the equality of the data result. Ontology learning frameworks introduced in Gao and Zhu [39], Gao et al. [42], Wu et al. [44] are implemented on "Chemical Index" ontologies, and Table 3 presents the precision ratios deduced from these four ontology learning tricks.

Note that the Figure 5 and Figure 6 presented above only contain of partial vertices of ontologies. In fact, there are 68 concepts in chemical index ontology O_5, and 46 concepts in chemical index ontology O_6. The concepts contained in O_5 or O_6 but not displayed in Figure 5 and Figure 6 are: "singly vertex-weighted Wiener number", "multiplicative Wiener index", "terminal Wiener index", "generalized Harary index", "second atom bond connectivity index", $\cdots$, "fifth atombond connectivity index", "revised edge Szeged index", "GeneralCo-PI index", "Shultz polynomial", "zeroth-order general Randic index", "eccentric connectivity polynomial", "second geometric-arithmetic index", $\cdots$, "fifth geometric-arithmetic index", "first Zagreb polynomial",$\cdots$, "sixth Zagreb polynomial", etc.

Following the compared partial data manifested in Table 3, we see that when N increases the average precision ratios show clear tendency to increase with it. Generally speaking, the advantage of our new algorithm is obvious by comparing the ratios. Even though the ratio from algorithm in Wu et al. [39] is the same as that from the new one when $N = 1$, when N becomes larger, the advantage of our new algorithm becomes more obvious. As a result, our distance computing based ontology algorithm is more efficient than ontology learning technologies studied in Gao and Zhu [39], Gao et al. [42], Wu et al. [44], especially when N is becoming large.

V. Conclusions

As a structured representation tool, the role of the ontology has gradually emerged in various application fields, and

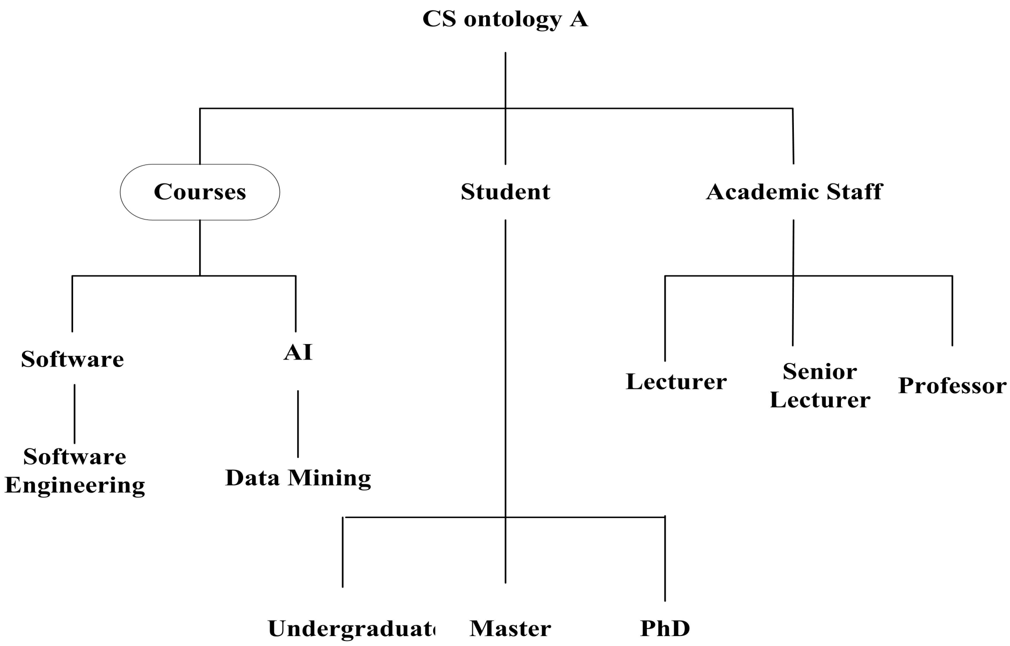

Fig. 1. "University" Ontology O_1

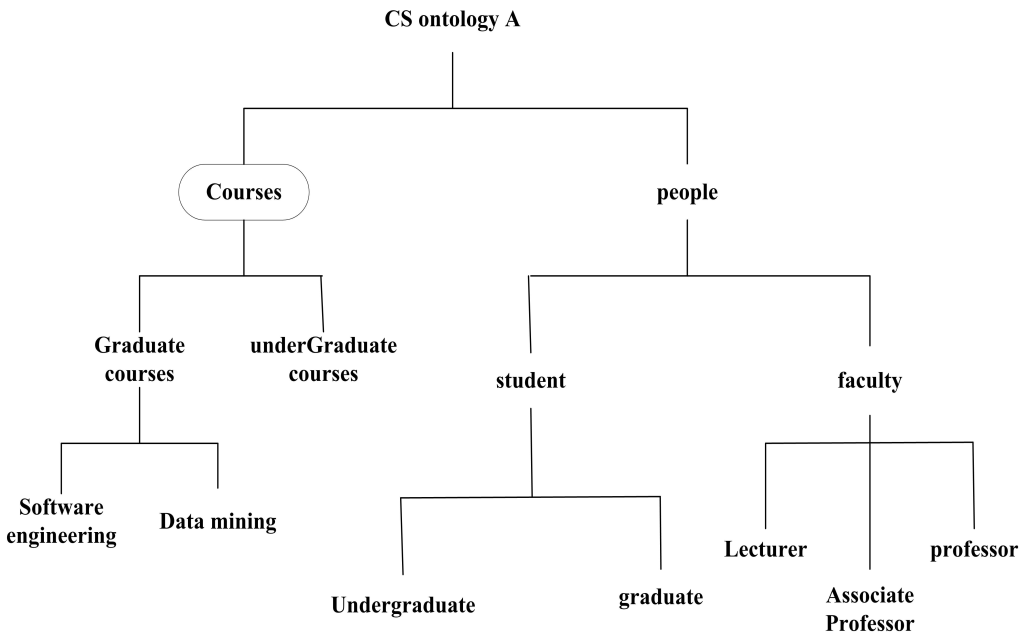

Fig. 2. "University" Ontology O_2

TABLE I
PRTIAL OF THE EXPERIMENT DATA FOR ONTOLOGY MAPPING ON "UNIVERSITY" ONTOLOGIES.

	P@1 average precision ratio	P@3 average precision ratio	P@5 average precision ratio
our ontology learning algorithm	0.5357	0.6429	0.7071
ontology algorithm in Gao et al. [40]	0.5000	0.5962	0.6857
ontology algorithm in Gao et al. [43]	0.4643	0.5714	0.6642
ontology algorithm in Gao et al. [45]	0.4286	0.5238	0.5929

its strong advantages have been highly praised by domain experts. Every year, new ontologies are constructed in all fields of natural science, and new ontology algorithms are constantly being created and expanded for the needs of various applications.

In this paper, we focus on the similarity based ontology

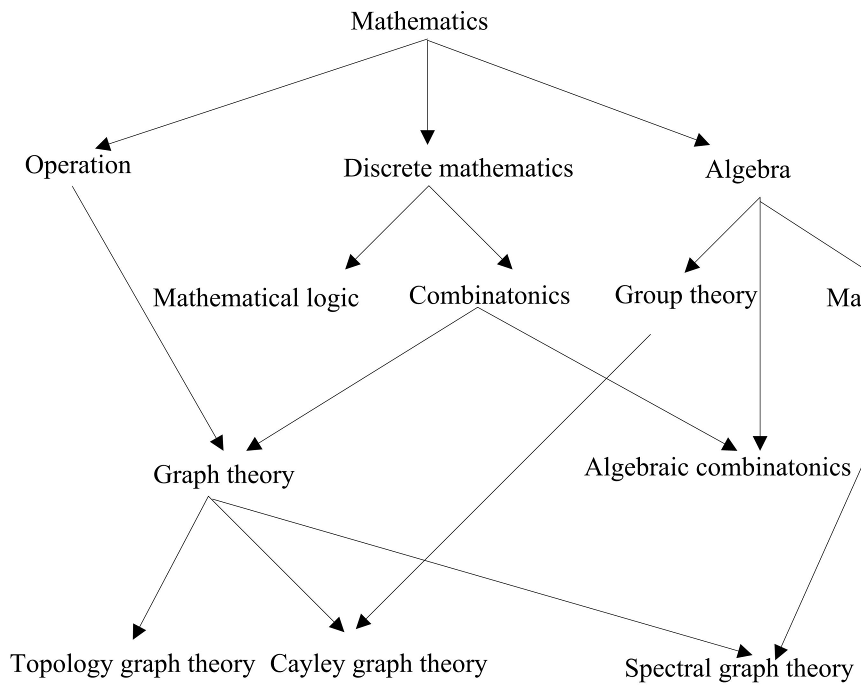

Fig. 3. “Mathematics” Ontology O_3

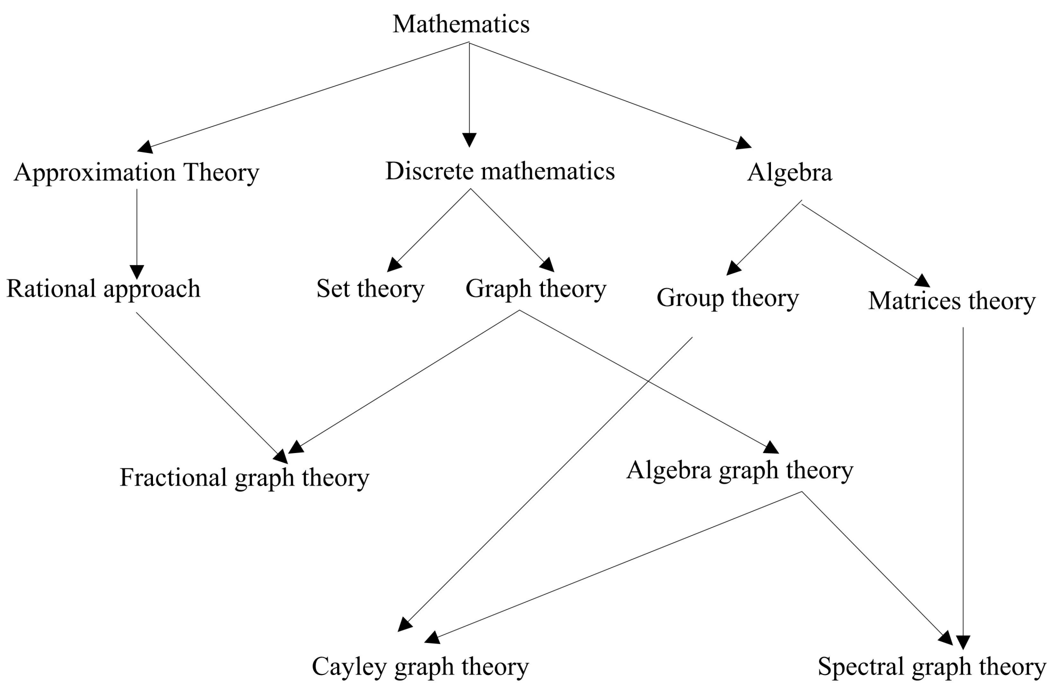

Fig. 4. “Mathematics” Ontology O_4

TABLE II
PRTIAL OF THE EXPERIMENT DATA FOR ONTOLOGY MAPPING ON “MATHEMATICS” ONTOLOGIES.

	P@1 average precision ratio	**P@3 average precision ratio**	**P@5 average precision ratio**
our ontology learning algorithm	0.3846	0.5128	0.6923
ontology algorithm in Gao and Zhu [39]	0.3077	0.4359	0.5615
ontology algorithm in Gao et al. [41]	0.3462	0.3974	0.5231
ontology algorithm in Wu et al. [44]	0.3846	0.5000	0.6769

mapping. The main idea of the designed algorithm is to determine the similarity between the two conceptual correspondence vectors by calculating the geometric distance between them. Considering the complexity of the obtained algorithm and the cost of ontology construction, our goal is to make the distance matrix $\mathbf{M}$ to be low-rank sparse

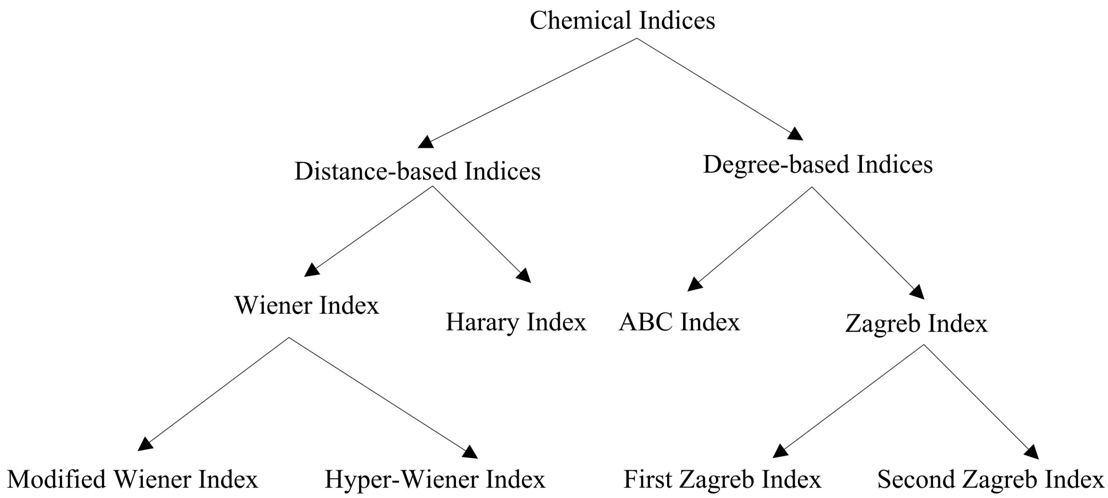

Fig. 5. "Chemical Index" Ontology O_5

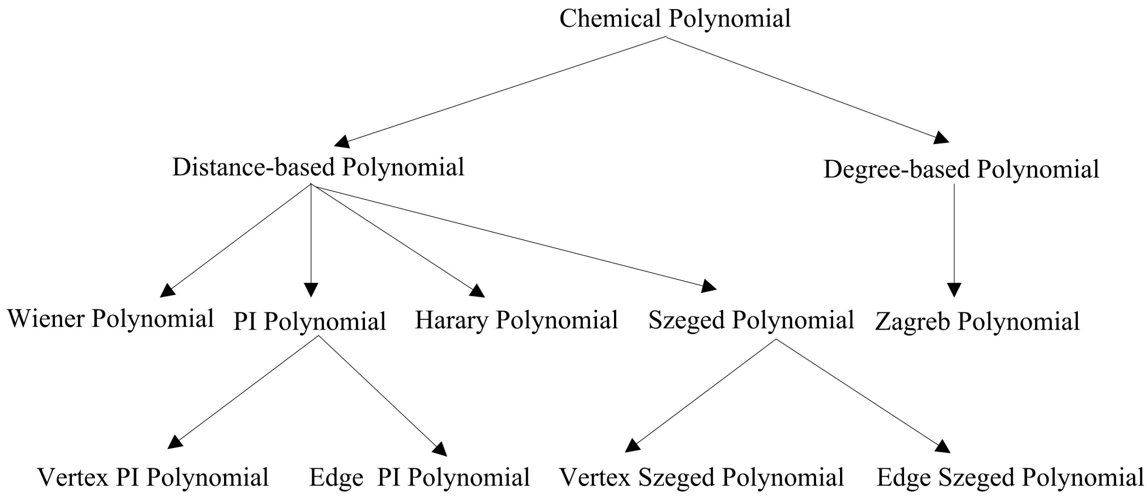

Fig. 6. "Chemical Index" Ontology O_6

TABLE III
PRTIAL OF THE EXPERIMENT DATA FOR ONTOLOGY MAPPING ON "CHEMICAL INDEX" ONTOLOGIES.

	P@1 average precision ratio	**P@3 average precision ratio**	**P@5 average precision ratio**
our ontology learning algorithm	0.4123	0.5263	0.6982
ontology algorithm in Gao and Zhu [39]	0.3247	0.4415	0.5667
ontology algorithm in Gao et al. [42]	0.3947	0.4678	0.5807
ontology algorithm in Wu et al. [44]	0.4123	0.5058	0.6754

matrix, and it is done by optimizing the constraints of the constraints in the model. Finally, the deduced ontology mapping algorithm is applied to "University" ontologies, "Mathematics" ontologies and "Chemical Index" ontologies, all of which have achieved high efficiency.

We believe that with the deepening of ontology applications and the advancement of related research in the future, more algorithms will be proposed and effectively applied to various ontology engineering fields. At the same time, we believe that the following can be used as a subject for further research:
(1) How to improve the existing ontology learning algorithm to make it meet the needs of large-scale computing and real-time computing requires further research.
(2) How to integrate ontology features into some well-known learning methods (for example: reinforcement learning, convolutional neural networks, graph neural networks, etc.) requires further research.
(3) We find that most of the current research on ontology is still focused on ontology inference, that is, in the process of ontology construction, using related methods of mathematical logic to logically model the ontology. The advantage of this is that the relationship between the entire ontology concepts is supported by a set of logic theories, which makes the ontology use mathematical logic as a tool in practical applications, and the derivation of the relationship between the concepts is more rigorous.

From the existing research results, mathematical logic and machine learning are two modeling tools of ontology, and there is no relationship between them. The question now is, if an ontology uses mathematical logic and machine learning to build computational models, can a bridge be established between these two models. That is, can the elements of ontology logic be integrated into the ontology learning process?

References

[1] J. M. Przydzial, B. Bhhatarai, and A. Koleti, "GPCR ontology: development and application of a G protein-coupled receptor pharmacology knowledge framework," *Bioinformatics*, vol. 29, pp. 3211–3219, 2013.

[2] S. Koehler, S. C. Doelken, and C. J. Mungall. "The human phenotype ontology project: linking molecular biology and disease through phenotype data," *Nucleic Acids Res.*, vol. 42, pp. 966–974, 2014.

[3] M. Ivanovic and Z. Budimac, "An overview of ontologies and data resources in medical domains," *Expert Syst. Appl.*, vol. 41, pp. 5158–15166, 2014.

[4] A. Hristoskova, V. Sakkalis, and G. Zacharioudakis, "Ontology-driven monitoring of patient's vital signs enabling personalized medical detection and alert," *Sensors*, vol. 14, pp. 1598–1628, 2014.

[5] M. A. Kabir, J. Han, and J. Yu, "User-centric social context information management: an ontology-based approach and platform," *Pers. Ubiquit. Comput.*, vol. 18, pp. 1061–1083, 2014.

[6] Y. L. Ma, L. Liu, K. Lu, et al. "A graph derivation based approach for measuring and comparing structural semantics of ontologies," *IEEE T. Knowl. Data En.*, vol. 26, pp. 1039–1052, 2014.

[7] Z. Li, H. S. Guo, Y. S. Yuan, et al. "Ontology representation of online shopping customers knowledge in enterprise information," *Appl. Mech. Mater.*, vol. 483, pp. 603–606, 2014.

[8] R. Santodomingo, S. Rohjans, M. Uslar, et al. "Ontology matching system for future energy smart grids," *Eng. Appl. Artif. Intel.*, vol. 32, pp. 242–257, 2014.

[9] T. Pizzuti, G. Mirabelli, M. A. Sanz-Bobi, et al. "Food Track & Trace ontology for helping the food traceability control," *J. Food Eng.*, vol. 120, pp. 17–30, 2014.

[10] N. Lasierra, A. Alesanco, and J. Garcia, "Designing an architecture for monitoring patients at home: Ontologies and web services for clinical and technical management integration," *IEEE J. Biomed. Health Inform.*, vol. 18, pp. 896-906, 2014.

[11] D. B. Carlini and M. Makowski, "Codon bias and gene ontology in holometabolous and hemimetabolous insects," *J. Exp. Zool. Part B*, vol. 324, pp. 686–698, 2015.

[12] C. Nicolai, "Deflationary truth and the ontology of expressions," *Synthese*, vol. 192, pp. 4031–4055, 2015.

[13] C. G. P. Correae, A. M. de Carvalho Moura, and C. M. Claudia, "A multi-ontology approach to annotate scientific documents based on a modularization technique," *J. Biomed. Inform.*, vol. 58, pp. 208–219, 2015.

[14] H. A. Duran-Limon, C. A. Garcia-Rios, F. E. Castillo-Barrera, et al. "An ontology-based product architecture derivation approach," *IEEE T. Software Eng.*, vol. 41, pp. 1153–1168, 2015.

[15] Y. Chabot, A. Bertaux, C. Nicolle, et al. "An ontology-based approach for the reconstruction and analysis of digital incidents timelines," *Digit. Invest.*, vol. 15, pp. 83–100, 2015.

[16] J. P. Elbers and S. S. Taylor, "GO2TR: a gene ontology-based workflow to generate target regions for target enrichment experiments," *Conserv. Genet. Res.*, vol. 7, pp. 851–857, 2015.

[17] M. Rani, R. Nayak, and O. P. Vyas, "An ontology-based adaptive personalized e-learning system, assisted by software agents on cloud storage," *Knowl. Based Syst.*, vol. 90, pp. 33–48, 2015.

[18] A. Sangiacomo, "The ontology of determination: from descartes to spinoza," *Sciencein Context*, vol. 28, pp. 515–543, 2015.

[19] C. L. B. Azevedo, M. E. Iacob, J. P. A. Almeida, et al. "Modeling resources and capabilities in enterprise architecture: a well-founded ontology-based proposal for ArchiMate," *Inform. Sci.*, vol. 54, pp. 235–262, 2015.

[20] H. Wimmer and R. Rada, "Good versus bad knowledge: ontology guided evolutionary algorithms," *Expert Syst. Appl.*, vol. 42, pp. 8039–8051, 2015.

[21] N. Trokanas and F. Cecelja, "Ontology evaluation for reuse in the domain of process systems engineering," *Comput. Chem. Eng.*, vol. 85, pp. 177–187, 2016.

[22] S. Chhun, N. Moalla, and Y. Ouzrout, "QoS ontology for service selection and reuse," *J. Intell. Manuf.*, vol. 27, pp. 187–199, 2016.

[23] R. Costa, C. Lima, J. Sarraipa, et al. "Facilitating knowledge sharing and reuse in building and construction domain: an ontology-based approach," *J. Intell. Manuf.*, vol. 27, pp. 263–282, 2016.

[24] P. Panov, L. N. Soldatova, and S. Dzeroski, "Generic ontology of datatypes," *Inform. Sci.*, vol. 329, pp. 900–920, 2016.

[25] A. Kutikov, H. H. Woo, and J. W. Catto, "Urology tag ontology project: standardizing social media communication descriptors," *Eur. Urol.*, vol. 69, pp. 183–185, 2016.

[26] F. Grandi, "Dynamic class hierarchy management for multi-version ontology-based personalization," *J. Comput. Syst. Sci.*, vol. 82, pp. 69–90, 2016.

[27] E. Kontopoulos, G. Martinopoulos, D. Lazarou, et al. "An ontology-based decision support tool for optimizing domestic solar hot water system selection," *J. Clean. Prod.*, vol. 112, pp. 4636–4646, 2016.

[28] D. C. Hoyle, A. Brass, "Statistical mechanics of ontology based annotations," *Physica A*, vol. 442, pp. 284–299, 2016.

[29] L. Solano, F. Romero, and P. Rosado, "An ontology for integrated machining and inspection process planning focusing on resource capabilities," *Int. J. Comp. Integ. M.*, vol. 29, pp. 1–15, 2016.

[30] X. Aime and J. Charlet, "Social psychology insights into ontology engineering," *Future Gener. Comp. Sy.*, vol. 54, pp. 348–351, 2016.

[31] W. Gao and M. R. Farahani, "Generalization bounds and uniform bounds for multi-dividing ontology algorithms with convex ontology loss function," *Comput. J.*, vol. 60, pp. 1289–1299, 2017.

[32] W. Gao, J. L. G. Guirao, B. Basavanagoud, et al. "Partial multi-dividing ontology learning algorithm," *Inform. Sci.*, vol. 467, pp. 35–58, 2018.

[33] W. Gao and T. W. Xu, "Stability analysis of learning algorithms for ontology similarity computation," *Abstr. Appl. Anal.*, vol. 2013, http://dx.doi.org/10.1155/2013/174802, 2013.

[34] L. L. Zhu, Y. Pan, M. K. Jamil, and W. Gao. "Boosting based ontology sparse vector computation approach," Eng. Lett., vol. 25, no. 4, pp. 406–415, 2017.

[35] M. H. Lan, J. Xu, and W. Gao. "Ontology feature extraction via vector learning algorithm and applied to similarity measuring and ontology mapping," IAENG Int. J. Comp. Sci., vol. 43, no. 1, pp. 10–19, 2016.

[36] L. L. Zhu, W. G. Tao, X. Z. Min, and W. Gao. "Theoretical characteristics of ontology learning algorithm in multi-dividing setting," IAENG Int. J. Comp. Sci., vol. 43, no. 2, pp. 184–191, 2016.

[37] L. L. Zhu, Y. Pan, M. R. Farahani, and W. Gao "Theoretical characteristics on scoring function in multi-dividing setting," IAENG Int. J. Appl. Math., vol. 47, no. 1, pp. 28–36, 2017.

[38] M. X. Chen, B. X. Liu, D. S. Zeng, and W. Gao, "A framework for ontology–driven similarity measuring using vector learning tricks," Eng. Lett., vol. 27, no. 3, pp. 549-558, 2019.

[39] W. Gao and L. L. Zhu, "Gradient learning algorithms for ontology computing," *Comput. Intell. Neur.*, vol. 2014, http://dx.doi.org/10.1155/2014/438291, 2014.

[40] W. Gao, L. L. Zhu, and K. Y. Wang, "Ontology sparse vector learning algorithm for ontology similarity measuring and ontology mapping via ADAL technology," *Int. J. Bifurcat. Chaos*, vol. 25, DOI: 10.1142/S0218127415400349, 2015.

[41] W. Gao, L. L. Zhu, and K. Y. Wang, "Ranking based ontology scheming using eigenpair computation," *J. Intell. Fuzzy Syst.*, vol. 31, pp. 2411–2419, 2016.

[42] W. Gao, Y. Guo, and K. Y. Wang, "Ontology algorithm using singular value decomposition and applied in multidisciplinary," *Cluster Comput.*, vol. 19, pp. 2201–2210, 2016.

[43] W. Gao, M. R. Farahani, A. Aslam, et al. "Distance learning techniques for ontology similarity measuring and ontology mapping," *Cluster Comput.*, vol. 20, pp. 959–968, 2017.

[44] J. Z. Wu, X. Yu, and W. Gao, "Disequilibrium multi dividing ontology learning algorithm," *Commun. Stat.-Theor. M.*, vol. 46, pp. 8925–8942, 2017.

[45] W. Gao, A. Q. Baig, H. Ali, et al. "Margin based ontology sparse vector learning algorithm and applied in biology science," *Saudi J. Biol. Sci.*, vol. 24, pp. 132–138, 2017.

Caihua Qiu is with the Department of Computer Science, Guangdong University of Science and Technology, Dongguan, China, and Bulacan State University, Philippines.