A Framework for Ontology-Driven Similarity Measuring Using Vector Learning Tricks

Mengxiang Chen, Beixiong Liu, Desheng Zeng and Wei Gao,

Abstract—**Ontology learning problem has raised much attention in semantic structure expression and information retrieval. As a powerful tool, ontology is evenly employed in various subjects, such as neuroscience, medical science, pharmacopedia, chemistry, education and other social science. Ontology similarity measuring plays a vital role in practical implementations since essential issues of ontology mapping are also similarity calculating. In ontology function learning, one learns a real valued score function that assigns scores to each ontology vertex which corresponds to a concept. Thus, the similarity between vertices is determined by means of the absolute value of difference between their corresponding scores. In this paper, we report the new optimization algorithms for obtaining ontology function in view of ontology sparse vector learning. The implementation of ontology algorithms is mainly based on iterative calculation in which we consider the whole matrix version of framework and ontology sparse vector are updated in each iterative. The data results obtained from four simulation experiments reveal that our newly proposed ontology approach has high efficiency and accuracy in biology and plant science with regard to ontology similarity measure, and humanoid robotics and education science with regard to ontology mapping.**

Index Terms—**ontology, similarity measure, ontology mapping, ontology sparse vector, iterative computation.**

I. Introduction

There was no a clear definition on ontology in computer science at the beginning. Since 1991, researchers have been constantly giving and modifying various definitions of ontology. Recently, The most well accepted definition is proposed by Borst in 1997, i.e., "the ontology is a displayed specification of the conceptualization of sharing". It points out several distinctive features of the ontology, namely that, the ontology is a shared, conceptual, explicit, formal specification. The specific meaning is described as follows:

- **Ontology is a conceptual model.** It implies that ontology is derived from the relevant concepts of phenomenon in reality, whose meaning is independent from the specific environmental state. It describes the relationship between concepts.
- **Explicit.** The referred concepts and the constraints on these concepts are clearly defined, ie there is no ambiguity.
- **Formalization.** It means that the defined ontology is machine readable, not a natural language. Formalization can accurately describe a concept. Computers help to get a more accurate understanding of formal content.
- **Sharing.** The ontology reflects the commonly recognized knowledge, especially the set of concepts recognized in the relevant field, which aimed at groups rather than individuals. It is vital to share this feature, which states the reason why ontology is used in various areas.

The ontology is generated for the communication between different individuals, such as sharing, interoperability and so on. It provides a clear consensus for this exchange and ontology can capture knowledge in related fields. By identifying vocabulary that is commonly recognized in the field (specific or general), and giving the relationship between these vocabulary and vocabulary from different levels, it provides a semantic support for communication between individuals. In the past, individual communication often lacked semantic support. The semantic information is not inherent in the concept itself, but is generated under the designer's design. Different designers may give different meanings to the same concept, and even the same designer may give different meanings to the same concept under different circumstances. The relationship between concepts is not clear in which different perceptions of concepts make it difficult for us to communicate smoothly in different fields. The incomplete understanding of the concept also makes the intelligence of computer processing greatly compromised. With the development of ontology technology, these problems are gradually overcome.

As a conceptual semantic model, ontology has become a useful tool in computer science and information technology, which has permeated in intelligence decision making, data integration, image process, knowledge management, and collaboration. Also, it has been widely applied in pharmacology science, biology science, GIS, medical science and social sciences (for instance, see Acharya et al. [1], Shahsavani et al. [2], Horne et al. [3], Sormaz and Sarkar [4], Jayawardhana and Gorsevski [5], Ledvinka et al. [6], Sacha et al. [7], Reyes-Alvarez et al. [8], Nadal et al. [9], and Oliva et al. [10]).

Ontology is modelled by a graph structure, directed or undirected. Each vertex in ontology graph corresponds to a concept and each (directed or undirected) edge expresses an owner-member relationship (or potential relations) between two concepts in its corresponding ontology. Let O be an ontology and G be an ontology (directed) graph corresponding to O. The aim of ontology engineering applications is to get the similarity computation criterion which is used to judge the similarities between ontology vertices. Thus, the relevance between ontology concepts is determined in view of the vertex similarity. Moreover, ontology mapping is used to deduce the ontology similarity between vertices from multi-

Manuscript received February 13, 2019; revised May 15 2019. This work was supported in part by the "Research on High Precision and Convenient Positioning Terminal System Based on Beidou Satellite" and Foshan Science and Technology Bureau Science and Technology Planning Projects 2017 (No. 2017AB004072).

M.X. Chen is with the Department of Mechanical and Electrical Engineering, Guangdong Polyechnic of Environmental Protection Engineering, Foshan 528216, China. e-mail: mengxiang_chen@126.com.

B.X. Liu is with the Department of Mechanical and Electrical Engineering, Guangdong Polyechnic of Environmental Protection Engineering, Foshan 528216, China.

D.S. Zeng is with Information Engeering InstituteGuangdong Innovative Technical CollegeDongguan 523960China.

W. Gao is with School of Information Science and Technology, Yunnan Normal University, Kunming 650500, China.

ontologies which is built as a bridge connecting different ontologies. Therefore, the problem of ontology mapping can also be summarized as ontology similarity measuring, and these can be implemented by the same ontology learning algorithm.

In view of its powerful performance and good functionality, ontology has been applied in various disciplines and receives very good effect. Ochieng and Kyanda [11] demonstrated the spectral partitioning of an ontology which can generate high quality partitions geared towards matching between two different ontologies. By means of ontology techniques, McGarry et al. [12] identified drugs with similar side-effects which are used in drug repositioning to apply existing drugs to different diseases or medical conditions, alleviating to a certain extent the time and cost expended in drug development. Deepak and Priyadarshini [13] proposed system classifies the ontologies using SVM and a Homonym LookUp directory. Benavides et al. [14] described a study of the use of knowledge models represented in ontologies for building Computer Aided Control Systems Design (CACSD) tools. Kumar and Thangamani [15] raised multi-ontology based points of interests and parallel fuzzy clustering algorithm for travel sequence recommendation with mobile communication on big social media. Wheeler et al. [16] implemented an ontology-based knowledge model to formally conceptualise relevant knowledge in hypertension clinical practice guidelines, behaviour change models and associated behaviour change strategies. Haendel et al. [17] described ontologies and their use in computational reasoning to support precise classification of patients for diagnosis, care management, and translational research. Zaletelj et al. [18] proposed an extensible foundational ontology for manufacturing-system modelling in which the formal definitions of the modelling environment itself enable the definition of the manufacturing system's elements. Gyrard et al. [19] considered four ontology catalogs that are relevant for IoT and smart cities, and demonstrated how can ontology catalogs be more effectively used to design and develop smart city applications. Alobaidi et al. [20] proposed novel automated ontology generation framework consists of five major modules which allowed mitigating the time-intensity to build ontologies and achieve machine interoperability.

Specially, different kinds of learning algorithm were introduced in the ontology function learning. Using these learning methods, each ontology vertex is mapped into a real number, and the similarity between concepts of ontology is determined by means the difference between their corresponding real numbers. In the learning setting, we should mathematicise the ontology information, i..e., for each vertex in ontology graph, all its information is enclosure in a vector. By slightly confusing the notations, we denote v by both the ontology vertex and its corresponding vector. Hence, this vector is mapped to a real number in terms of ontology function, therefore these ontology learning algorithms are kinds of dimensionality reduction techniques.

It has large number of effective ontology learning method raised and applied in ontology engineering, and several contributions presented the theoretical analysis of ontology learning algorithm from the perspective of statistical learning theory. Gao et al. [21] analysised the strong and weak stability of k-partite ranking based ontology learning algorithm. Gao and Zhu [22] manifested the gradient learning algorithms for ontology similairty computing and ontology mapping. Gao et al. [23] presented an ontology sparse vector learning algorithm for ontology similarity measuring and ontology mapping via ADAL technology. Gao et al. [24] proposed the ranking based ontology scheming in terms of eigenpair computation. Gao et al. [25] raised new ontology algorithmin light of singular value decomposition and applied in multidisciplinary. Gao et al. [26] determined the trick of ontology similarity measuring and ontology mapping using distance learning. Wu et al. [27] studied the ontology learning trick using disequilibrium multi dividing method. Gao and Farahani [28] researched the generalization bounds and uniform bounds for multi-dividing ontology algorithms with convex ontology loss function. Gao et al. [29] proposed partial multi-dividing ontology learning algorithm and obtained some theoretical results from statistical view. Gao and Xu [30] yielded the stability analysis of learning algorithms for ontology similarity computation. More related contexts can be referred to [31], [32], [33], [34]

In this paper, we present the new learning approach for ontology application. The key tricks of our ontology sparse iterative algorithm are concerning the designing of update rules in particular mathematical settings. The rest of this paper is arranged as follows: first, the notations and the setting of ontology sparse vector learning are introduced; then, the detailed ontology iterative algorithms for ontology sparse vector learning are presented in Section 3; finally, the proposed ontology algorithms are employed in plant science, physical education, biology science and humanoid robotics to verify the effectiveness of algorithms on similarity measuring and ontology mapping, respectively.

II. Setting

Let V be an ontology instance space. For any vertex $v \in V(G)$, all its related information is expressed by a p dimensional vector, i.e., $v = (v^1, \cdots, v^p)^T$. W.L.O.G., by slightly confusing the notation, we use v to denote both v and its corresponding vector. For the given real-valued ontology function f, the similarity between two vertices v_i and v_j is judged by $|f(v_i) - f(v_j)|$.

Here, the dimension p of vector is always large since it contains all the information of the corresponding concept, including attribute and the neighborhood structure in the ontology graph. For example, in biology ontology or chemical ontology, the information of all genes, molecular structure, chemical process and disease or medicinal may be contained in a vector. Furthermore, the structure of ontology graph becomes very complicated since its vertex number becomes large, and one typical instance is the GIS (Geographic Information System) ontology. These factors lead to the high calculation complexity of ontology similarity measuring and ontology mapping application. However, the similarity between the ontology vertices is only determined by a small number of vector components. For example, in the gene ontology, only a small number of diseased genes lead to a genetic disease, and we can ignore most of the other genes. Another example, in the application of GIS ontology, if an accident happens and causes casualties somewhere, we should find the nearest hospital without considering the shops, factories and schools nearby. That is to say, we only consider the neighborhood information which satisfies the specific needs on the ontology graph. From this point of view, large number of industrial and academic interests are attracted to study the sparse ontology learning algorithm.

Actually, one sparse ontology function is expressed by

$$f_{\mathbf{w}}(v)=\sum_{i=1}^{p}v^{i}w_{i}+\delta, \tag{1}$$

where $\mathbf{w}=(w_1,\cdots,w_p)^T$ is a sparse vector and δ is a noise term. The sparse vector $\mathbf{w}$ is used to shrink irrelevant component to zero. Thus, we should learn sparse vector $\mathbf{w}$ first to determine the ontology function f. Hence, by ignoring the noise term, the response vector $\mathbf{y}=(y_1,\cdots,y_n)^T$ can be expressed by

$$\mathbf{y}=\mathbf{V}\mathbf{w} \tag{2}$$

where $\mathbf{V}\in R^{n\times p}$ is a ontology information matrix, and for each ontology instance (v_i,y_i), we have $y_i=v_i^T\mathbf{w}=\sum_{j=1}^{p}v_i^jw_j$ where $v_i=(v_i^1,v_i^2,\cdots,v_i^p)^T$.

Here, we don't give review of ontology sparse vector learning algortihms, but give an example to show how to get the optimal solution. Let $S=\{(v_i,y_i)\}_{i=1}^n$ be the ontology tranining sample set with $v_i\in\mathbb{R}^p$ and $y_i\in\mathbb{R}$. We assume that labels are centered and ontology data are standardized, i.e., $\sum_{i=1}^n y_i=0$, $\sum_{i=1}^n v_{ij}=0$ and $\sum_{i=1}^n v_{ij}^2=1$. In this setting, one of ontology optimization framework can be expressed as

$$\min_{\mathbf{w}}\frac{1}{2}\sum_{i=1}^{n}(v_i^T\mathbf{w}-y_i)^2 \tag{3}$$

$$\text{s.t.}\quad \|\mathbf{w}\|_1+c_1\sum_{j<k}\max\{|w_j|,|w_k|\}\le c_2,$$

where $\|\mathbf{w}\|_1$ controls the sparsity of ontology sparse vector $\mathbf{w}$, c_1 and c_2 are two positive balance parameters. If combin the ontology optimization framework (3) with its restrict conditions, we obtain

$$\begin{aligned}F(\mathbf{w},\lambda_1,\lambda_2) &= \min_{\mathbf{w}}\frac{1}{2}\sum_{i=1}^{n}(v_i^T\mathbf{w}-y_i)^2+\lambda_1\|\mathbf{w}\|_1\\ &\quad +\lambda_2\sum_{i<j}\max\{|w_i|,|w_j|\},\end{aligned} \tag{4}$$

where λ_1 and λ_2 are positive balance parameters. Suppose $\mathbf{w}$ is the optimal solution of (4), $o_j\in\{1,\cdots,p\}$ is the order of $|w_j|$ among $\{|w_1|,|w_2|,\cdots,|w_p|\}$ such that $|w_{j_1}|\le|w_{j_2}|$ if $o(j_1)<o(j_2)$. The set $\mathcal{K}_k\subseteq\{1,\cdots,p\}$ meets the following two conditions: $|w_{j_1}|=|w_{j_2}|=\vartheta_k$ for any $j_1,j_2\in\mathcal{K}_k$ with $j_1\ne j_2$; $|w_j|\ne\vartheta_k$ if $j\in\{1,\cdots,p\}$ and $j\in\mathcal{K}_k$. It reveals that ϑ_k $(0\le\vartheta_1<\vartheta_2<\cdots<\vartheta_K)$ represents the common value of $|w_j|$ for set $\mathcal{K}_k$ and $\mathcal{K}_1\cup\mathcal{K}_2\cup\cdots\cup\mathcal{K}_K=\{1,\cdots,p\}$. Thus, the ontology optimization framework can be re-formulated as

$$\min_{\vartheta}\frac{1}{2}\sum_{i=1}^{n}(\tilde{v}_i^T\vartheta-y_i)^2+\sum_{k=1}^{K}\beta_k\vartheta_k \tag{5}$$

$$\text{s.t.}\quad 0\le\vartheta_1<\vartheta_2<\cdots<\vartheta_K,$$

where $\tilde{v}_i=\{\tilde{v}_{i1},\tilde{v}_{i2},\cdots,\tilde{v}_{iK}\}$, $\tilde{v}_{ik}=\sum_{j\in\mathcal{K}_k}\text{sign}(w_j)v_{ij}$ and $\beta_k=\sum_{j\in\mathcal{K}_k}(\lambda_1+(o(j)-1)\lambda_2)$.

By defining the active set by means of ϑ_k as $\mathcal{A}=\{k\in\{1,\cdots,K\}|\vartheta_k>0\}$ and $\overline{\mathcal{A}}=\{1,\cdots,K\}-\mathcal{A}$, the optimal conditions of (5) can be re-expressed as

$$\sum_{i=1}^{n}-\tilde{v}_{ik}(\tilde{v}_i^T\vartheta-y_i)+\beta_k=0,\quad \forall k\in\mathcal{A} \tag{6}$$

$$0<\vartheta_{2-|\overline{\mathcal{A}}|}<\cdots<\vartheta_{K-1}<\vartheta_K \tag{7}$$

$$\vartheta_{\overline{\mathcal{A}}}=0. \tag{8}$$

For two parameter λ_1 and λ_2, set $\Delta\lambda=\begin{pmatrix}\Delta\lambda_1\\ \Delta\lambda_2\end{pmatrix}$, $d=\begin{pmatrix}d_1\\ d_2\end{pmatrix}$ and $\Delta\lambda=d\Delta\eta$, where $\Delta\eta$ is a parameter used to control the adjustment qualities of two parameters. Let $\Delta\vartheta$ be the changes of parameter ϑ with direction ξ, $\Delta\eta_{max}$ be te maximum change of $\Delta\eta$, and ε be the accuracy parameter. Set $\tilde{\beta}_k=\sum_{j\in\mathcal{K}_k}(d_1+(o(j)-1)d_2)$ and we infer the following linear system for each $k\in\mathcal{A}$:

$$\sum_{i=1}^{n}\tilde{v}_{ik}\tilde{v}_{i\mathcal{A}}^T\Delta\vartheta_{\mathcal{A}}+\tilde{\beta}_k\Delta\eta=0. \tag{9}$$

Let $\tilde{B}$ be the $K\times K$ diagonal matrix with elements $\tilde{\beta}_k$, and $\tilde{V}$ be a $n\times K$ matrix whose i-th row is $\tilde{v}_i^T$. Hence, (9) can be expressed as

$$\tilde{V}_{*\mathcal{A}}^T\tilde{V}_{*\mathcal{A}}\Delta\vartheta_{\mathcal{A}}+\tilde{B}_{\mathcal{A}\mathcal{A}}\Delta\eta=0. \tag{10}$$

Let $H_{\mathcal{A}\mathcal{A}}=\tilde{V}_{*\mathcal{A}}^T\tilde{V}_{*\mathcal{A}}$ and $\xi_{\mathcal{A}}=\frac{\Delta\vartheta_{\mathcal{A}}}{\Delta\eta}$. Then (10) can be re-formulated by

$$H_{\mathcal{A}\mathcal{A}}\xi_{\mathcal{A}}=-\tilde{B}_{\mathcal{A}\mathcal{A}}. \tag{11}$$

Thereore, $\xi_{\mathcal{A}}$ (the direction of $\Delta\vartheta_{\mathcal{A}}$) can be obtained by solving (11).

The maximum adjustment $\Delta\eta_{max}$ should be determined after yielding the linear relationships of $\xi_{\mathcal{A}}$ in which we need to consider the three main classes of situations. If a certain ϑ_k in $\mathcal{A}$ equals to zero, then the maximal possible $\Delta\eta^{\mathcal{A}}$ can be calculated before a certain $\vartheta_k\in\mathcal{A}$ moves to $\mathcal{A}$ using the constraints $\vartheta_k+\xi_k\Delta\eta>0$ for any $k\in\mathcal{A}$ in the optimality conditions of ontology learning alorithm. If the pair of feature sets change their orders of ϑ_k, in light of (7), the optimality conditions of ontology learning algorithm rely on a fix orders of ϑ_k. Hence, the maximal possible $\Delta\eta^o$ can be determined in view of constraints $\vartheta_k+\xi_k\Delta\eta<\vartheta_{k+1}+\xi_{k+1}\Delta\eta$ before a pair of $\mathcal{K}_k$ and $\mathcal{K}_{k+1}$ change their orders. If the termination condition is satisfied, i.e., η reaches $\overline{\eta}$, then the maximal adjustment value before the ontology algorithm satisfies the termination condition is $\overline{\eta}-\eta$. Finally, the smallest of three values $\{\overline{\eta}-\eta,\Delta\eta^{\mathcal{A}},\Delta\eta^o\}$ constitutes the maximal adjustment value of $\Delta\eta_{max}$.

The solution $\mathbf{w}$ can be ensured as a ε-approximation solution with $F(\mathbf{w},\lambda_1,\lambda_2)-F(\mathbf{w}^*,\lambda_1,\lambda_2)\le\varepsilon$ by the duality gap $K(\mathbf{w},\lambda_1,\lambda_2)=F(\mathbf{w},\lambda_1,\lambda_2)-\tilde{F}(\alpha,\lambda_1,\lambda_2)\le\varepsilon$ due to the ontology optimization problem $F(\mathbf{w})$ in (4) is a convex problem, where $\mathbf{w}^*$ is an optimal solution of $F(\mathbf{w},\lambda_1,\lambda_2)$, α is the dual variable, and $\tilde{F}(\alpha,\lambda_1,\lambda_2)$ is the dual of $F(\mathbf{w},\lambda_1,\lambda_2)$. Specifically,

$$\tilde{F}(\alpha,\lambda_1,\lambda_2)=\max_{\alpha}-\frac{\alpha^T\alpha}{2}-\alpha^Ty \tag{12}$$

$$\text{s.t.}\quad \max_{\sum_{j=1}^{p}(\lambda_1+\lambda_2(o(j)-1))|w_j|=1}\alpha^T\mathbf{V}\mathbf{w}\le 1,$$

where $\mathbf{V} \in \mathbb{R}^{n\times p}$ is ontology information matrix and optimal α of $\tilde{F}(\alpha, \lambda_1, \lambda_2)$ can be analytically computed as $\alpha = (\mathbf{V}\mathbf{w} - \mathbf{y}) \min\{\frac{1}{r*(V^T(\mathbf{V}\mathbf{w}-\mathbf{y}))}, 1\}$ where $|\gamma_1| \leq |\gamma_2| \leq \cdots \leq |\gamma_p|$ and

$$r^*(\gamma) = \max_{j\in\{1,\cdots,p\}} \frac{\sum_{i=1}^{j} |\gamma_j|}{\sum_{i=1}^{j} \lambda_1 + (i-1)\lambda_2}.$$

On the how to compute the value of $K(\vartheta, \lambda_1, \lambda_2)$, we can get it using the following steps: given $\mathbf{w}$ or ϑ, λ_1 and λ_2; calculate $\gamma = V^T(\mathbf{V}\mathbf{w} - \mathbf{y})$ and decline order the γ_i; determine $r^*(\gamma)$; compute the optimal α of $\tilde{F}(\alpha, \lambda_1, \lambda_2)$; return the duality gap $K(\mathbf{w}, \lambda_1, \lambda_2) = F(\mathbf{w}, \lambda_1, \lambda_2) - \tilde{F}(\alpha, \lambda_1, \lambda_2)$ in terms of (12).

The whole processes can be described as follows: given direction number and accuracy parameter ε,an interval $[\eta_1, \eta_2]$ of η, determine the solustion ϑ and sets $\mathcal{K}_k$ for $\eta = \eta_1$; repeat the following actions until $\eta > \eta_2$ or $K(\vartheta, \lambda_1, \lambda_2) > \varepsilon$: calculate $\Delta\vartheta$ and $\Delta\eta$, update η, ϑ, λ_1, λ_2, $\mathcal{A}$ and ϑ_k; determine $K(\vartheta, \lambda_1, \lambda_2)$. Finally return a solution in $[\eta_1, \eta_2]$ with regard to λ_1 and λ_2.

III. Main Ontology Algorithms

In this section, we present the ontology sparse learning techniques, and thus apply it in the ontology engineering. Two main ontology algorithms for ontology similarity measuring and ontology mapping are manifested.

A. Greedy ontology algorithm and its optimization tricks

Now we consider the following ontology model

$$y_i = \mathbf{w}^T v_i + \delta_i \tag{13}$$

or an ontology logistic regression framework

$$p(y_i|v_i) = \frac{1}{1 + e^{-\mathbf{w}^T v_i - w_0}}, \tag{14}$$

where $\delta = (\delta_1, \cdots, \delta_p)^T \in \mathbb{R}^p$ is an ontology disturbance vector and $w_0 \in \mathbb{R}$ is an ontology disturbance term. A modified ontology linear model with pairwise interaction terms ($\Psi \in \mathbb{R}^{p\times p}$ is a weight matrix which is connected with pairwise interactions) can be formulated as

$$y_i = \mathbf{w}^T v_i + v_i^T \Psi v_i + \delta_i, \tag{15}$$

or ontology logistic version

$$p(y_i|v_i) = \frac{1}{1 + e^{-y_i(\mathbf{w}^T v_i + v_i^T \Psi v_i + w_0)}}. \tag{16}$$

The ontology loss function corresponding to this framework is denoted by

$$l(\mathbf{w}, \Psi) = \frac{1}{2} \sum_{i=1}^{n} \|y_i - \mathbf{w}^T v_i - v_i^T \Psi v_i\|_2^2, \tag{17}$$

or the logistic ontology loss

$$l_{log}(\mathbf{w}, \Psi, w_0) = \sum_{i=1}^{n} \log(1 + e^{-y_i(\mathbf{w}^T v_i + v_i^T \Psi v_i + w_0)}). \tag{18}$$

Assume that Ψ can be expressed as the tensor product of K rank one matrices for pairwise interactions, i.e., $\Psi = \sum_{k=1}^{K} \varrho_k \otimes \varrho_k$. Let $\widehat{\mathbf{w}}$ and $\widehat{\varrho}_k$ be the estimator of $\mathbf{w}$ and ϱ_k respectively. Then the ontology sparse vector is obtained below

$$\begin{aligned}\{\widehat{\mathbf{w}}, \widehat{\varrho}_k\} \quad = \quad & \arg\min_{\varrho_k, \mathbf{w}} l(\mathbf{w}, \Psi) + \lambda_{\mathbf{w}} \|\mathbf{w}\|_1 \\ & + \sum_{k=1}^{K} \lambda_{\varrho_k} \|\varrho_k\|_1,\end{aligned} \tag{19}$$

where $\lambda_{\mathbf{w}}$ and λ_{ϱ_k} for $k \in \{1, \cdots, K\}$ are positive balance parameters. In the logistic ontology case, the corresponding framework can be obtained replace ontology loss function in (19) by l_{log}. Set Q as objective function of (19) and t is a counting number in the iteration process.

Let $L(\varrho_k)$ be the ontology loss function of ontology framework (19) with regard to ϱ_k. The optimality condition for ontology framework (19) can be stated as $\bigtriangledown_j L(\varrho_k) + \lambda_{\varrho_k} \mathrm{sgn}(\varrho_{kj}) = 0$ if $\varrho_{kj} \neq 0$; otherwise $| \bigtriangledown_j L(\varrho_k)| \leq \lambda_{\varrho_k}$. Also, we can re-write the ontology optimality conditions for $\mathbf{w}$ using the similar fashion. Let

$$\bigtriangledown_j L(\varrho_k) = \frac{1}{2} \sum_i (-2v_i^j v_i^T \varrho_k)(y_i - \mathbf{w}^T v_i - v_i^T \Psi v_i).$$

Thus, the subgradient $\bigtriangledown_j^s f(\varrho_k)$ for each ϱ_{kj} is $\bigtriangledown_j L(\varrho_k) + \lambda_{\varrho_k} \mathrm{sgn}(\varrho_{kj})$ if $|\varrho_{kj}| > 0$; $\bigtriangledown_j L(\varrho_k) + \lambda_{\varrho_k}$ if $\varrho_{kj} = 0$ and $\bigtriangledown_j L(\varrho_k) < -\lambda_{\varrho_k}$; $\bigtriangledown_j L(\varrho_k) - \lambda_{\varrho_k}$ if $\varrho_{kj} = 0$ and $\bigtriangledown_j L(\varrho_k) > \lambda_{\varrho_k}$; and 0 if $-\lambda_{\varrho_k} \leq \bigtriangledown_j L(\varrho_k) \leq \lambda_{\varrho_k}$. And the subgradient with regard to $\mathbf{w}$ can be inferred using the similar way, and the differential of ontology loss function in ontology framework (19) with regard to $\mathbf{w}$ can be determined by

$$\bigtriangledown_j L(\mathbf{w}) = \frac{1}{2} \sum_i (-2v_i^j)(y_i - \mathbf{w}^T v_i - v_i^T \Psi v_i).$$

The projection operator theory can be used to optimal the ontology problem. Let $\mathcal{P}_{\mathcal{O}}$ and $\mathcal{P}_{\mathcal{S}}$ be two projection operators, $\Gamma_{\mathbf{w}}$ be the positive definite approximation of Hessian of quadratic approximation of ontology function $f(\mathbf{w})$, $\gamma_{\mathbf{w}}$ and γ_{ϱ_k} be the parameter for step sizes. Here $\mathcal{P}_{\mathcal{O}}$ projects the step into the orthant including ϱ_k and $\mathbf{w}$, and $\mathcal{P}_{\mathcal{S}}$ projects the Newton-like direction to ensure the descent direction. Hence, the update can be described as

$$\mathbf{w} \leftarrow \mathcal{P}_{\mathcal{O}}(\mathbf{w} - \gamma_{\mathbf{w}} \mathcal{P}_{\mathcal{S}}(\Gamma_{\mathbf{w}}^{-1} \bigtriangledown^s f(\mathbf{w}))),$$

$$\varrho_k \leftarrow \mathcal{P}_{\mathcal{O}}(\varrho_k - \gamma_{\varrho_k} \mathcal{P}_{\mathcal{S}}(\Gamma_{\varrho_k}^{-1} \bigtriangledown^s f(\varrho_k))).$$

More contexts on projection operator can be referred to Magnus and Brosens [35], Pluta and Russo [36], Rosenthal [37], Arias and Gonzalez [38], Somai et al. [39], Jorgensen et al. [40], Leble et al. [41], Bramati et al. [42], and Censor and Mansour [43].

Also, soft computing tricks can be well applied to get the ontology optimization solution. By setting S as the soft thresholding operator, the updated rule can be formulated as

$$\begin{aligned}\tilde{\mathbf{w}}_j(\lambda_{\mathbf{w}}) \quad \leftarrow \quad & S(\tilde{\mathbf{w}}_j(\lambda_{\mathbf{w}}) + \sum_{i=1}^{n} V_{ij}(y_i \\ & - \sum_{k\neq j} V_{jk} \tilde{\mathbf{w}}_k - \sum_{k} V_{ik} \Psi V_{ki}), \lambda_{\mathbf{w}}),\end{aligned}$$

$$\tilde{\varrho}_{kj}(\lambda_{\varrho_k}) \leftarrow S(\tilde{\varrho}_{kj}(\lambda_{\varrho_k}) + \sum_{i=1}^{n} V_{ij}(\sum_{k}\sum_{r=1}^{P} \varrho_{kr}V_{ir})[y_i - \sum_{k\neq j} V_{jk}\tilde{\mathbf{w}}_k - \sum_{k} V_{ik}\Psi V_{ki}], \lambda_{\varrho_k}).$$

A greedy ontology implementation process can be described as follows: set $\mathbf{w} = 0$, $K = 1$ and $\varrho_K = 1$. $\varrho_{K-1} > 0$ for $K \geq 2$; repeat actions $w_j^t = \arg\min_j Q((w_1^t, \cdots, w_{j-1}^t, w_{j+1}^{t-1}, \cdots, w_p^{t-1}), \varrho_k^{t-1})$ and $\varrho_{k,j}^t = \arg\min_j Q((\varrho_{k,1}^t, \cdots, \varrho_{k,j-1}^t, \varrho_{k,j+1}^{t-1}, \cdots, \varrho_{k,p}^{t-1}), \varrho_k^{t-1})$ untill convergence; $K = K+1$ and $\varrho_K = 1$; jump out of the condition loop and delete ϱ_K and ϱ_{K-1} from ϱ.

B. Ontology sparse vector learning in special disturbance setting

In the part, we propose the ontology sparse vector learning algorithm in the special mathematical setting. First we express the ontology model using the matrix version as follows:

$$\mathbf{y} = \mathbf{V}^T\mathbf{w} + \epsilon, \tag{20}$$

where $\mathbf{y} \in \mathbb{R}^n$, $\mathbf{V}$ is the ontology information matrix, $\mathbf{w} = (w_1, \cdots, w_p)^T \in \mathbb{R}^p$ is the ontology sparse vector and $\epsilon \in \mathbb{R}^n$ is a disturbance vector which meets $\epsilon \sim \mathcal{N}(0, \Pi)$. In some machine learning works, $\Pi^{-1} = (c_{ij})$ expresses the partial covariance sturcture of data.

Assume that there are m observation ontology sparse vectors $\mathbf{w}^i = (w_1^i, w_2^i, \cdots, w_p^i)^T$ and their corresponding estimators are $\mathbf{y}^i = (y_1^i, y_2^i, \cdots, y_p^i)^T$ (here $i \in \{1, \cdots, m\}$). The l_1 norm of matrix is denoted as $\|\Pi\|_1 = \sum_{i\leq j}|c_{ij}|$ and $\|\mathbf{V}\|_1 = \sum_{i,j}|v_{ij}|$. The ontology loss function in this setting can be formulated as

$$l_{m,\lambda}(\mathbf{V}, \Pi) = l_m(\mathbf{V}, \Pi) + \lambda_1\|\Pi^{-1}\|_1 + \lambda_2\|\mathbf{V}\|_1, \tag{21}$$

where ontology loss function $l_m(\mathbf{V}, \Pi)$ is used to measure the quality of observation and their estimators, and λ_1, λ_2 are balance parameters. For example, it can be detailed expressed as $l_m(\mathbf{V}, \Pi) = \mathrm{Trace}((\mathbf{Y} - \mathbf{VW})^T(\mathbf{Y} - \mathbf{VW})\Pi^{-1}) - \log\Pi^{-1}$, where $\mathbf{Y} = (\mathbf{y}^1, \cdots, \mathbf{y}^m) \in \mathbb{R}^{n\times m}$, $\mathbf{W} = (\mathbf{w}^1, \cdots, \mathbf{w}^m) \in \mathbb{R}^{p\times m}$, $\mathbf{V} \in \mathbb{R}^{n\times p}$.

Let $d(\mathbf{y}|\mathbf{w})$ and $d(\mathbf{y}|\mathbf{w}, \mathbf{V}, \Pi)$ be true conditional distribution function and parametetric distribution function, respectively. Then the expected value $\mathbb{E}[d(\mathbf{y}|\mathbf{w}, \mathbf{V}, \Pi)]$ can be approximated using the empirical version $\frac{1}{m}\sum_{i=1}^{m} d(\mathbf{y}^i|\mathbf{w}^i, \mathbf{V}, \Pi)$, and ontology loss function can be written in empirical version (set $\int d^2(\mathbf{y}|\mathbf{w}, \mathbf{V}, \Pi)d\mathbf{y} = \frac{1}{2^n\pi^{\frac{n}{2}}|\Pi|^{\frac{1}{2}}}$):

$$\tilde{l}_m(\mathbf{V}, \Pi) = \int d^2(\mathbf{y}|\mathbf{w}, \mathbf{V}, \Pi)d\mathbf{y} - \frac{2}{m}\sum_{i=1}^{m} d(\mathbf{y}^i|\mathbf{w}^i, \mathbf{V}, \Pi). \tag{22}$$

Furthermore,

$$\tilde{l}_{m,\lambda}(\mathbf{V}, \Pi) = \frac{|\Pi^{-1}|^{\frac{1}{2}}}{2^n\pi^{\frac{n}{2}}} - \frac{2}{m}\sum_{i=1}^{m} d(\mathbf{y}^i|\mathbf{w}^i, \mathbf{V}, \Pi) + \lambda_1\|\Pi^{-1}\|_1 + \lambda_2\|\mathbf{V}\|_1. \tag{23}$$

When it comes to logistic ontology loss, (22) can be re-written as

$$l_m(\mathbf{V}, \Pi) = -\log[\frac{1}{m}\sum_{i=1}^{m} e^{-\frac{1}{2}(\mathbf{y}^i - V^T\mathbf{w}^i)^T\Pi^{-1}(\mathbf{y}^i - V^T\mathbf{w}^i)}] - \frac{1}{2}\log|\Pi^{-1}|, \tag{24}$$

and additional, (23) can be re-written as

$$l_{m,\lambda}(\mathbf{V}, \Pi) = -\log[\frac{1}{m}\sum_{i=1}^{m} e^{-\frac{1}{2}(\mathbf{y}^i - V^T\mathbf{w}^i)^T\Pi^{-1}(\mathbf{y}^i - V^T\mathbf{w}^i)}] - \frac{1}{2}\log|\Pi^{-1}| + \lambda_1\|\Pi^{-1}\|_1 + \lambda_2\|\mathbf{V}\|_1. \tag{25}$$

Set $\Theta = (\mathbf{V}, \Pi^{-1})$ and $\rho_i(\Theta) = -\frac{(\mathbf{y}^i - V^T\mathbf{w}^i)^T\Pi^{-1}(\mathbf{y}^i - V^T\mathbf{w}^i)}{2}$. By setting Θ_0 as the initial value and C is a positive parameter which is not dependent on Θ, then the first order approximation of $\log\{\frac{\sum_{i=1}^{m} e^{\rho_i(\Theta)}}{m}\}$ with regrad to Θ can be stated by

$$\frac{1}{m}\sum_{i=1}^{m} \frac{me^{\rho_i(\Theta_0)}}{\sum_{i=1}^{m} e^{\rho_i(\Theta_0)}} \bigtriangledown \rho_i(\Theta_0)^T(\Theta - \Theta_0) + C.$$

The objection function (25) can be approximated by

$$- \log|\Pi^{-1}| + \frac{1}{m}\sum_{i=1}^{m} \frac{me^{\rho_i(\Theta_0)}}{\sum_{i=1}^{m} e^{\rho_i(\Theta_0)}}(\mathbf{y}^i - V^T\mathbf{w}^i)^T\Pi^{-1}(\mathbf{y}^i - V^T\mathbf{w}^i) + \lambda_1\|\Pi^{-1}\|_1 + \lambda_2\|\mathbf{V}\|_1. \tag{26}$$

Set $\Lambda = \Lambda(\Theta_0) = \frac{1}{m}\sum_{i=1}^{m} \frac{me^{\rho_i(\Theta_0)}}{\sum_{i=1}^{m} e^{\rho_i(\Theta_0)}}(\mathbf{y}^i - V^T\mathbf{w}^i)(\mathbf{y}^i - V^T\mathbf{w}^i)^T$, then (26) can be re-stated as

$$- \log|\Pi^{-1}| + \mathrm{Trace}(\Pi^{-1}\Lambda(\Theta_0)) + \lambda_1\|\Pi^{-1}\|_1 + \lambda_2\|\mathbf{V}\|_1. \tag{27}$$

The whole iterative processes can be described as follows: input Π_0^{-1}, ε_1 and ε_2 are precision parameters, $\mathbf{V}_0$ and all ontology informations; get Λ, $\widehat{\Pi}^{-1} = \arg\min\{-\log|\Pi^{-1}| + \mathrm{Trace}(\Pi^{-1}\Lambda(\Theta_0)) + \lambda_1\|\Pi^{-1}\|_1 + \lambda_2\|\mathbf{V}\|_1\}$ in terms of (27) and $\widehat{\mathbf{V}}$ by means of (26) and given Π_0^{-1}; if $\|\widehat{\Pi}^{-1} - \Pi_0^{-1}\|_F^2 \leq \varepsilon_1$ and $\|\widehat{\mathbf{V}} - \mathbf{V}_0\|_F^2 \leq \varepsilon_2$, then finish the procedures. Otherwise, set $\Pi_0^{-1} = \widehat{\Pi}^{-1}$ and $\mathbf{V}_0 = \widehat{\mathbf{V}}$ and return back to the last step.

We found that in a big data environment, data is generally sparse. Conversely, dense data can only appear in small data environments. Therefore, the above sparse vector algorithm is consistent with the ontology similarity calculation and dimensionality reduction in most big data environments.

IV. Experiments

In this section, we present four simulation experiments on ontology similarity measure and ontology mapping respectively to verity the effectiveness of proposed ontology algorithms. To make experiment result as precise as possible, our presented ontology algorithms are run in C++, using LAPACK and BLAS libraries for linear algebra computation.

We chose these ontology data sets as experimental objects, not only because they play a major role in their respective fields, but also serve as a standard ontology sample. They are constructed in accordance with the following principles.

• **Clarity and objectivity.** The ontology should give a clear and objective semantic definition of the defined terms in

natural language.
• **Completeness.** The definitions given are complete and they fully express the meaning of the terms described.
• **Consistency.** The inference derived from the term is compatible with the meaning of the term itself, and there is no contradiction.
• **Maximum monotonic scalability.** When adding general or special terms to the ontology, there is no need to modify the existing content.
• **Minimum commitment.** As few constraints as possible to present the modelled object.
• **Ontological distinction principle.** The classes in the ontology should be disjoint.
• **Diversification of hierarchies.** The multiple inheritance mechanisms are potential to be enhanced.
• Modularity to minimize the degree of coupling between the modules.
• **Minimizat ion of the semantic distance.** The semantic distance between the sibling concepts is minimized, and the concepts with similar meanings are abstracted as much as possible, which are represented by the same meta-language.
• **Standardization of names.** Use standard names whenever possible.

Here we need to be particularly reminded that in the specific algorithm program, the calculation between sparse vectors is different from the traditional vector calculation. Since most of the components are zero, binary search techniques can be used for sparse vector calculations and improve program efficiency.

A. Experiment on biology data

First, we employ our ontology algorithm in the field of biology. The “Go” ontology O_1 was constructed in http://www.geneontology.org (Figure 1 presents the basic structure of O_1) is a structured database which stores biological gene-related information. $P@N$ (Precision Ratio) is used as criterion to weight the equality of the result data. It implemented within several steps:
• first, we infer the closest N concepts for each vertex on the ontology graph by experts.
• then, we derive the N concepts with largest similarity for each vertex on ontology graph by our proposed algorithm.
• finally, we calculate the precision ratio for each vertex, and at last determine the precision ratio for whole ontology graph.

Also, ontology function training approaches introduced in Gao et al. [23], [25] and [44] are also acted on “Go” ontology. The precision ratios are compared from these four techniques, and partial experiment results can refer to Table 1.

From the precision ratio comparison presented in Table 1, we see that the precision ratio gotten from our ontology algorithm is much higher than the precision ratio obtained by ontology learning algorithms proposed in Gao et al. [23], [25] and [44] when N =3, 5, 10 or 20. Therefore, we can draw the conclusion that the ontology algorithm presented in our paper is superior to the ontology approaches described by Gao et al. [23], [25] and [44] in the biology gene “Go” ontology application.

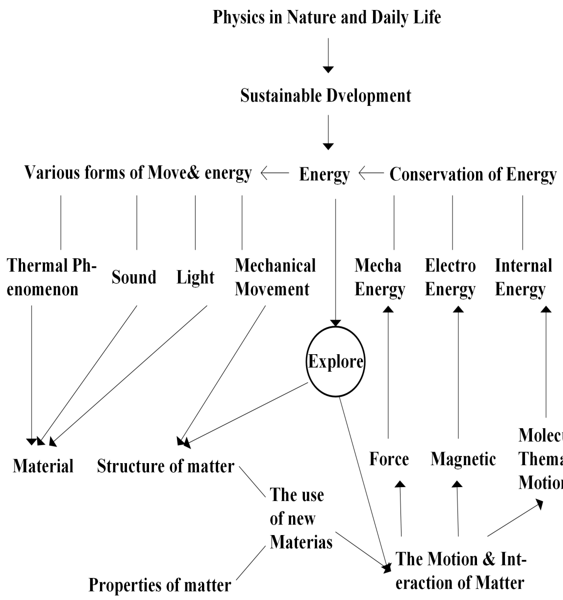

Fig. 2. “Physical Education” Ontology O_2

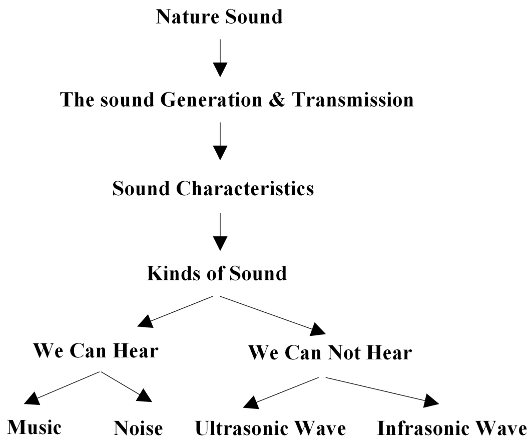

Fig. 3. “Physical Education” Ontology O_3

B. Experiment on physical education data

The physical education ontologies O_2 and O_3 (the structures of O_2 and O_3 can refer to Figure 2 and Figure 3, respectively) are used to test the serviceability of our algorithm for the ontology mapping construction which is purposed to compute the similarity between vertices from ontology O_2 and O_3. $P@N$ is also used as the criterion to measure the equality of the obtained results. Ontology learning algorithms described in Gao et al. [23], [24] and [25] are also act on “physical education” ontology. The comparison is made on precision ratios among these four learning approaches, and partial result can refer to Table 2.

According to the precision ratio comparison revealed in Table 2, we see that our proposed ontology learning our algorithm is more efficient than ontology learning techniques manifested in Gao et al. [23], [24] and [25].

C. Experiment on plant data

The “PO” ontology O_4 was constructed in http://www.plantontology.org (Figure 4 reveals the basic structure of O_4) which is used as a dictionary in plant

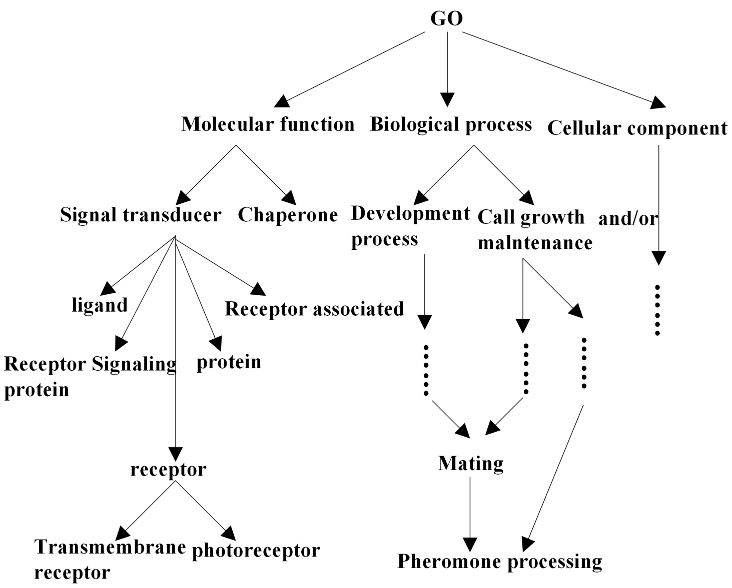

Fig. 1. "Go" ontology

TABLE I
PARTIAL OF SIMILARITY COMPUTING DATA ON "GO" ONTOLOGY.

	P@3 average precision ratio	P@5 average precision ratio	P@10 average precision ratio	P@20 average precision ratio
our ontology learning algorithm	0.5653	0.6971	0.8396	0.9472
ontology algorithm in Gao et al. [23]	0.5192	0.6569	0.8031	0.9192
ontology algorithm in Gao et al. [25]	0.5201	0.6663	0.8275	0.9417
ontology algorithm in Gao et al. [44]	0.5649	0.6827	0.8124	0.9371

TABLE II
PRTIAL OF THE EXPERIMENT DATA FOR ONTOLOGY MAPPING ON PHYSICAL EDUCATION ONTOLOGIES.

	P@1 average precision ratio	P@3 average precision ratio	P@5 average precision ratio
our ontology learning algorithm	0.6913	0.7957	0.9355
ontology algorithm in Gao et al. [23]	0.6774	0.7849	0.9032
ontology algorithm in Gao et al. [24]	0.6913	0.7634	0.8968
ontology algorithm in Gao et al. [25]	0.6774	0.7957	0.9290

science, and it is used to verify the efficiency of our ontology learning algorithm for similarity measuring application. We use $P@N$ standard again in this experiment. Furthermore, previously ontology learning tricks manifested in Gao et al. [22], [25] and [44] are also acted on the "PO" ontology. The precision ratios by these four algorithms are determined, respectively. Partial of the precision ratios are compared in Table 3.

When N =3, 5, or 10, the precision ratios yielded by ontology learning algorithm introduced in Gao et al. [22], [25] and [44] are lower than the precision ratio deduced in terms of our ontology learning technology. Clearly, our proposed ontology learning algorithm is superior to the method described by Gao et al. [22], [25] and [44] with respect to similarity computation in plant science field.

D. Experiment on humanoid robotics data

Two humanoid robotics ontologies O_5 and O_6 was introduced by Gao and Zhu [22] (the fundamental structures of O_5 and O_6 are showed in Figure 5 and Figure 6 respectively) which are used for our last experiment. In this experiment, we aim to obtain similarity between ontology vertices from O_5 and O_6, and thus infer the ontology mapping according to these similarities. Once again, $P@N$ criterion is employed to judge the equality of the experiment conclusion. Ontology learning tricks obtained in Gao et al. [22], [23] and [25] were also acted on humanoid robotics ontologies. The precision ratios which are gotten from four ontology optimization methods are compared and partially presented in Table 4.

According to the compared precision ratios manifested in Table 4, we draw the conclusion that our described ontology learning algorithm acts more efficiently than ontology learning approaches introduced in Gao et al. [22], [23] and [25]. These superiority is manifested to be more evident as N becomes larger.

V. CONCLUSIONS

In this article, new ontology computation techniques are introduced for ontology similarity measure and ontology mapping applications. The detailed optimization scheming relies on ontology sparse vector learning, and the target ontology sparse vector is obtained via iterative computation. Simulation precision ratios of the four experiments in the last section imply that our new ontology technology has high efficiency in biology, physics education, plant science and humanoid robotics. The new ontology algorithms raised

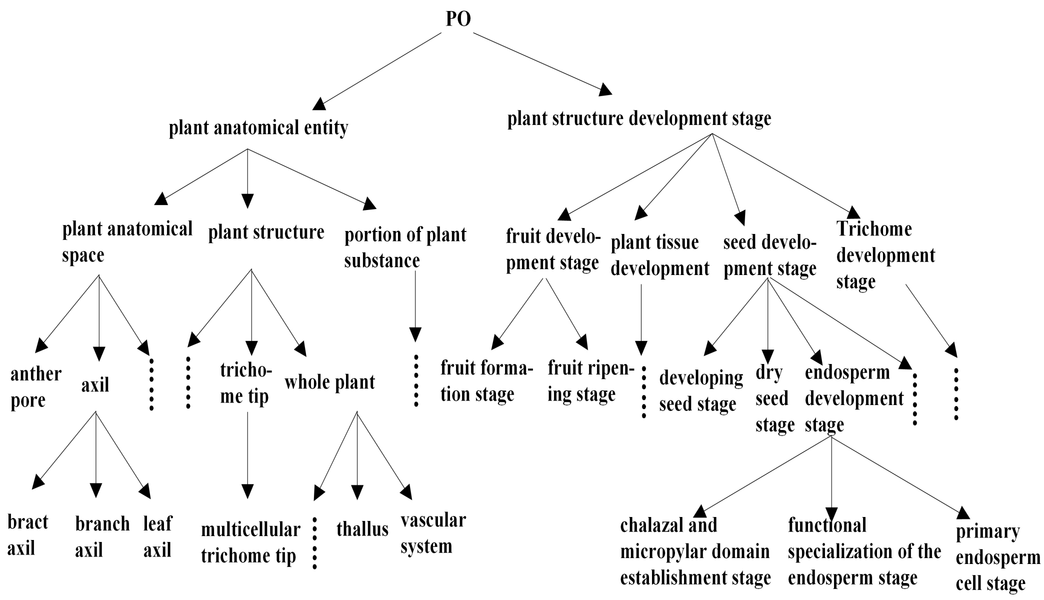

Fig. 4. “PO” ontology O_4

TABLE III
PARTIAL OF ONTOLOGY SIMILARITY COMPUTING DATA ON “PO” ONTOLOGY.

	P@3 average precision ratio	P@5 average precision ratio	P@10 average precision ratio	P@20 average precision ratio
our ontology learning algorithm	0.5371	0.6763	0.9037	0.9698
ontology algorithm in Gao et al. [22]	0.5042	0.6216	0.7853	0.9034
ontology algorithm in Gao et al. [25]	0.5081	0.6549	0.8104	0.9317
ontology algorithm in Gao et al. [44]	0.5360	0.6664	0.9004	0.9673

TABLE IV
PARTIAL ONTOLOGY MAPPING DATA ON HUMANOID ROBOTICS ONTOLOGIES.

	P@1 average precision ratio	P@3 average precision ratio	P@5 average precision ratio
our ontology learning algorithm	0.4444	0.6296	0.8444
ontology algorithm in Gao et al. [22]	0.4444	0.5185	0.6111
ontology algorithm in Gao et al. [23]	0.2778	0.6111	0.7889
ontology algorithm in Gao et al. [25]	0.4444	0.5370	0.8222

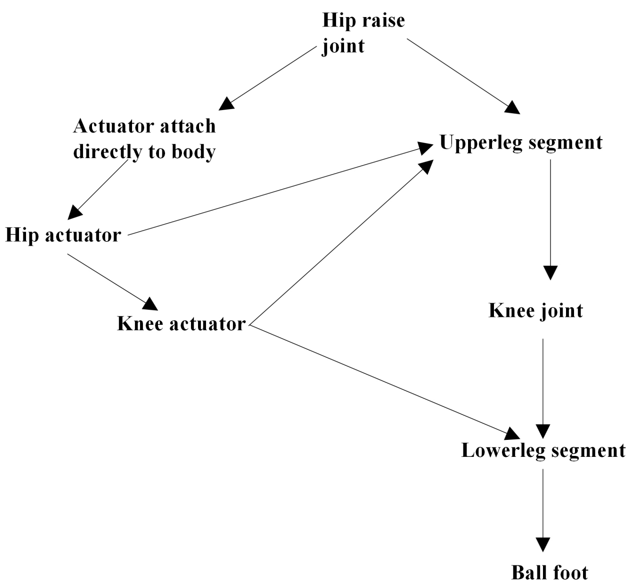

Fig. 5. “Humanoid Robotics” Ontology O_5

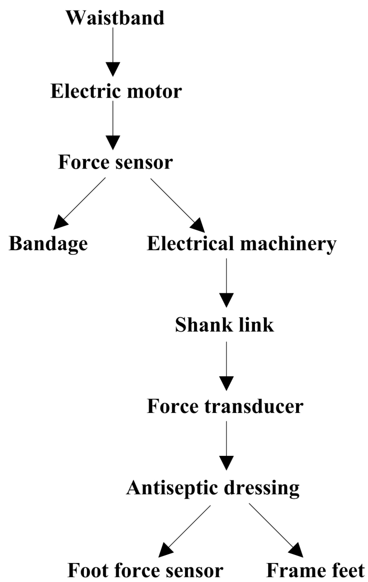

Fig. 6. “Humanoid Robotics” Ontology O_6

in this paper illustrate promising application prospects in multiple disciplines.

The following topics can be considered as the future ongoing research work.

• What are the statistical characteristics (for example, the computing complexity and convergence rate) of the given ontology learning algorithm?

• The ontologies of different application domains have their own different characteristics. Whether the algorithm proposed in this paper can exert better efficiency in other engineering fields waiting for further verification.

• Ontology sparse vector computes only one trick among learning techniques. Thus, it is natural to ask whether other machine learning techniques are suitable for ontology graph similarity applications.

VI. Acknowledgments

We thank all the reviewers for their constructive suggestions on how to improve the quality of this paper. This research is supported by (1) Research on High Precision and Convenient Positioning Terminal System Based on Beidou Satellite. (2) Foshan Science and Technology Bureau Science and Technology Planning Projects 2017 (No. 2017AB004072).

VII. Conflicts of Interest

The authors declare that there is no conflict of interests regarding the publication of this paper.

References

[1] S. Acharya, S. Saha, and P. Pradhan. "Novel symmetry-based gene-gene dissimilarity measures utilizing Gene Ontology: Application in gene clustering," *Gene*, vol. 679, pp. 341–351, 2018.

[2] S. Shahsavani, M. R. Shooshtarian, H. R. Tabatabaei, and M. Dehghani. "An ontology-based approach for modeling the heavy metals' temporal and spatial variations in groundwater resources by using fuzzy multi-criteria group decision-making models and geographical information system," *Desalin. Water Treat.*, vol. 136, pp. 259–267, 2018.

[3] C. Horne, R. Ritchie, and H. Griffiths. "Proposed ontology for cognitive radar systems," *IET Radar Sonar Nav.*, vol. 12, pp. 1363–1370, 2018.

[4] L. Sormaz and A. Sarkar. "SIMPM-Upper-level ontology for manufacturing process plan network generation," *Robot. Com.-Int. Manuf.*, vol. 55, pp. 183–198, 2019.

[5] U. K. Jayawardhana and P. V. Gorsevski. "An ontology-based framework for extracting spatio-temporal influenza data using Twitter," *Int. J. Digi. Earth*, vol. 12, pp. 2–24, 2019.

[6] M. Ledvinka, A. Lalis, and P. Kremen. "Toward data-driven safety: an ontology-based information system," *J. Aerospace Inform. Sys.*, vol. 16, pp. 22–36, 2019.

[7] D. Sacha, M. Kraus, D. A. Keim, and M. Chen. "VIS4ML: an ontology for visual analytics assisted machine learning," *IEEE T. Vis. Comput. Gr.*, vol. 25, pp. 385–395, 2019.

[8] L. Reyes-Alvarez, M. del Mar Roldan-Garcia, and J. F. Aldana-Montes. "Tool for materializing OWL ontologies in a column-oriented database," *Software Pract. Exper.*, vol. 49, pp. 100–119, 2019.

[9] S. Nadal, O. Romero, A. Abello, P. Vassiliadis, and S. Vansummeren. "An integration-oriented ontology to govern evolution in Big Data ecosystems," *Infrom. Syst.*, vol. 79, pp. 3–19, 2019.

[10] J. T. Oliva, H. D. Lee, N. Spolaor, W. S. R. Takaki, C. S. R. Coy, J. J. Fagundes, and F. C. Wu. "Computational system based on ontologies to automate the mapping process of medical reports into structured databases," *Expert Syst. Appl.*, **2019**, vol.115, pp. 37–56, 2019.

[11] P. Ochieng and S. Kyanda. "A K-way spectral partitioning of an ontology for ontology matching," *Distrib. Parallel Dat.*, vol. 36, pp. 643–673, 2018.

[12] K. McGarry, Y. Graham, S. McDonald, and A. Rashid. "RESKO: Repositioning drugs by using side effects and knowledge from ontologies," *Knowl.-Based Syst.*, vol. 160, pp. 34–48, 2018.

[13] G. Deepak and J. S. Priyadarshini. "Personalized and Enhanced Hybridized Semantic Algorithm for web image retrieval incorporating ontology classification, strategic query expansion, and content-based analysis." *Comput. Electr. Eng.*, vol. 72, pp. 14–25, 2018.

[14] C. Benavides, I. Garcia, H. Alaiz, and L. Quesada. "An ontology-based approach to knowledge representation for Computer-Aided Control System Design," *Data Knowl. Eng.*, vol. 118, pp. 107–125, 2018.

[15] N. S. Kumar and M. Thangamani. "Multi-ontology based points of interests (MO-POIS) and parallel fuzzy clustering (PFC) algorithm for travel sequence recommendation with mobile communication on big social media," *Wireless Pers. Commun.*, vol. 103, pp. 991–1010, 2018.

[16] T. S. Wheeler, T. M. Vallis, N. B. Giacomantonio, and S. R. Abidi. "Feasibility and usability of an ontology-based mobile intervention for patients with hypertension," *Int. J. Med. Inform.*, vol. 119, pp. 8–16, 2018.

[17] M. A. Haendel, C. G. Chute, and P. N. Robinson. "Classification, ontology, and precision medicine," *Int. J. Med. Inform.*, vol. 379, pp. 1452–1462, 2018.

[18] V. Zaletelj, R. Vrabic, E. Hozdic, and P. Butala. "A foundational ontology for the modelling of manufacturing systems," *Adv. Eng. Infrom.*, vol. 38, pp. 129–141, 2018.

[19] A. Gyrard, A. Zimmermann, and A. Sheth. "Building IoT-based applications for smart cities: How can ontology catalogs help?," *IEEE Internet Things J.*, vol. 5, pp. 3978–3990, 2018.

[20] M. Alobaidi, K. M. Malik, M. Hussain. "Automated ontology generation framework powered by linked biomedical ontologies for disease-drug domain," *Comput. Meth. Prog. Bio.*, vol. 165, pp. 117–128, 2018.

[21] W. Gao, Y. Gao, and Y. G. Zhang. "Strong and weak stability of k-partite ranking algorithm," *Information*, vol. 15, pp. 4585–4590, 2012.

[22] 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.

[23] 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.

[24] 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.

[25] 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.

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

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

[28] 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.

[29] W. Gao, J. L. G. Guirao, B. Basavanagoud, and J. Z. Wu. "Partial multi-dividing ontology learning algorithm," *Inform. Sciences*, vol. 467, pp. 35–58, 2018.

[30] 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.

[31] 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.

[32] 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.

[33] 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.

[34] 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.

[35] W. Magnus and F. Brosens. "Occupation numbers in a quantum canonical ensemble: a projection operator approach," *Physica A*, vol. 518, pp. 253–264, 2019.

[36] R. Pluta and B. Russo. "Projections on von Neumann algebras as limits of elementary operators," *Linear Algebra Appl.*, vol. 557, pp. 188–200, 2018.

[37] A. Rosenthal. "Algebraic determination of back-projection operators for optoacoustic tomography," *Bio. Opt. Eepr.*, vol. 9, pp. 5173–5193, 2018.

[38] M. L. Arias and M. C. Gonzalez. "Products of projections and self-adjoint operators," *Linear Algebra Appl.*, vol. 555, pp. 70–83, 2018.

[39] V. Somai, D. Legrady, and G. Tolnai. "Singular value decomposition analysis of back projection operator of maximum likelihood expectation maximization PET image reconstruction," *Radiol. Oncol.*, vol. 52, pp. 337–345, 2018.
[40] P. Jorgensen, E. Pearse, and F. Tian. "Unbounded operators in Hilbert space, duality rules, characteristic projections, and their applications," *Anal. AND Math. Phys.*, vol. 8, pp. 351–382, 2018.
[41] S. B. Leble, S. D. Vereshchagin, and I. S. Vereshchagina. "Diagnostics of atmospheric disturbances using the method of projection operators," *Russ. J. Phys. Chem. B*, vol. 12, pp. 554–561, 2018.
[42] R. Bramati, V. Casarino, and P. Ciatti. "On the norms of quaternionic harmonic projection operators," *C. R. Math.*, vol. 356, pp. 482–488, 2018.
[43] Y. Censor and R. Mansour. "Convergence analysis of processes with valiant projection operators in hilbert space," *J. Optimiz. Theory App.*, vol. 176, pp. 35–56, 2018.
[44] W. Gao, A. Q. Baig, H. Ali, W. Sajjad, and M. R. Farahani. "Margin based ontology sparse vector learning algorithm and applied in biology science," *Saudi J. Biol. Sci.*, vol. 24, pp. 132–138, 2017.

Mengxiang Chen obtained Master degree in Software Engineering, Huazhong University of Science and Technology, China in 2013. He worked in Department of Mechanical and Electrical Engineering, Guangdong Polyechnic of Environmental Protection Engineering. His research interests are Software Engineering and IoT applications, etc.

Beixiong Liu work in Department of Mechanical and Electrical Engineering, Guangdong Polyechnic of Environmental Protection Engineering.

Desheng Zeng work in Information Engeering Institute, Guangdong Innovative Technical College.

Wei Gao is a professor in School of Information Science and Technology, Yunnan Normal University. His research interests are graph theory and machine learning.