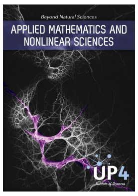
CrossMark

Applied Mathematics and Nonlinear Sciences 1(1) (2016) 159–174



Applied Mathematics and Nonlinear Sciences

http://journals.up4sciences.org

Ontology optimization tactics via distance calculating

Yun Gao[1†], Mohammad Reza Farahani[2], Wei Gao[3]

1. Department of Editorial, Yunnan Normal University, Kunming 650092, China
2. Department of Applied Mathematics, Iran University of Science and Technology, Narmak, Tehran 16844, Iran
3. School of Information Science and Technology, Yunnan Normal University, Kunming 650500, China

Submission Info

Communicated by F. Balibrea
Received 3th October 2015
Accepted 28th January 2016
Available online 29th January 2016

Abstract

In this article, we propose an ontology learning algorithm for ontology similarity measure and ontology mapping in view of distance function learning techniques. Using the distance computation formulation, all the pairs of ontology vertices are mapped into real numbers which express the distance of their corresponding vectors. The more distance between two vertices, the smaller similarity between their corresponding concepts. The stabilities of our learning algorithm are defined and several bounds are yielded via stability assumptions. The simulation experimental conclusions show that the new proposed ontology algorithm has high efficiency and accuracy in ontology similarity measure and ontology mapping in certain engineering applications.

Keywords and phrases: ontology, similarity measure, ontology mapping, distance computation, stability
2010 Mathematics Subject Classification: 60J20, 65L20

1 Introduction

Ontology originally comes from philosophy. It is used to describe the natural connection of things and their components' inherently hidden connections. Ontology is set up as a model for knowledge storage and representation in information and computer science. It has been extensively applied in different fields such as knowledge management, machine learning, information systems, image retrieval, information retrieval search extension,

†Corresponding author.
Email address: gaoyun@ynnu.edu.cn

ISSN 2444-8656 doi:10.21042/AMNS.2016.1.00012

collaboration and intelligent information integration. As a conceptually semantic model and an analysis tool, being quite effective, ontology has been favored by researchers from pharmacology science, biology science, medical science, geographic information system and social sciences since a few years ago (for instance, see Przydzial et al., [1], Koehler et al., [2], Ivanovic and Budimac [3], Hristoskova et al., [4], and Kabir [5]).

A simple graph is usually used by researchers to represent the structure of ontology. Every concept, objects and elements in ontology are made to correspond to a vertex. Each (directed or undirected) edge on an ontology graph represents a relationship (or potential link) between two concepts (objects or elements). Let O be an ontology and G be a simple graph corresponding to O. It can be attributed to getting. We use the similarity calculating function, the nature of ontology engineer application to compute the similarities between ontology vertices, which represent the intrinsic link between vertices in ontology graph. The ontology similarity measuring function is obtained by measuring the similarity between vertices from different ontologies. That is the goal of ontology mapping. The mapping serves as a bridge connecting different ontologies. Only through mapping, we gain a potential association between the objects or elements from different ontologies. The semi-positive score function $Sim : V \times V \to \mathbb{R}^+ \cup \{0\}$ maps each pair of vertices to a non-negative real number.

Several effective methods exist for getting efficient ontology similarity measure or ontology mapping algorithm in terms of ontology function. The ontology similarity calculation in terms of ranking learning technology was considered by Wang et al., [12]. The fast ontology algorithm in order to cut the time complexity for ontology application was raised by Huang et al., [13]. An ontology optimizing model in which the ontology function is determined by virtue of NDCG measure was presented by Gao and Liang [14], which is successfully applied in physics education. More ontology applications on various engineering can be refered to Gao et al., [11].

In this article, we determine a new ontology learning method by means of distance calculating. Moreover, we give a theoretical analysis for proposed ontology algorithm.

2 Algorithm Description

Let $\mathscr{S} = \{(v_i, v_j, y_{ij})\}_{i,j=1}^{N}$ be the ontology training data, where $v_i, v_j \in \mathbb{R}^p$ are ontology vectors and $y_{ij} = \pm 1$ (if v_i and v_j are similar, then $y_{ij} = 1$; otherwise, $y_{ij} = -1$. We also fixed m relevant source ontology training sets $\mathscr{S}_q = \{(v_{qi}, v_{qj}, y_{qij})\}_{i,j=1}^{N_q}$ $(q = 1, \cdots, m)$ if the number of target ontology training samples N is not large, and $v_{qi}, v_{qj} \in \mathbb{R}^p$ belong to the certain ontology feature space as v_i, v_j in this setting.

We aim to learn a distance function $d(v_i, v_j|\mathbf{W}) = (v_i - v_j)^T \mathbf{W}(v_i - v_j)$ which equals to learning a distance matrix $\mathbf{W}$, and the similarity or dissimilarity between a ontology vertex pair v_i and v_j is obtained by comparing $d(v_i, v_j|\mathbf{W})$ with a constant threshold parameter c. Specifically, our ontology optimization problem can be stated as

$$\arg\min \frac{1}{\binom{N}{2}} \sum_{i<j} g(y_{ij}[1 - \|v_i - v_j\|_{\mathbf{W}}^2]) + \frac{\eta}{2}\|\mathbf{W}\|_F^2 \tag{1}$$

$$\text{s.t.} \quad \sum_{i=1}^{m} \alpha_q = 1, \alpha_q \geq 0, q = 1, \cdots, m.$$

where $\|v_i - v_j\|_{\mathbf{W}}^2 = (v_i - v_j)^T \mathbf{W}(v_i - v_j)$, $g(z) = \max(0, b - z)$ is an ontology hinge loss function, $\|\mathbf{W}\|_F$ is the Frobenius norm of the metric $\mathbf{W}$ which is applied to control the model complexity, η is a balance parameter, and the constraint condition reveals that $\mathbf{W}$ is positive semi-definite.

The general version of ontology distance learning approach is formulated by

$$\arg\min \frac{1}{\binom{N}{2}} \sum_{i<j} L(v_i, v_j, y_{ij}) + \frac{\gamma_1}{2}\|\mathbf{W} - \mathbf{W}_D\|_F^2 + \frac{\gamma_2}{2}\|\alpha\|_2^2 + \gamma_3\|\theta\|_1, \tag{2}$$

$$\text{s.t.} \quad \sum_{i=1}^{m} \alpha_q = 1, \alpha_q \geq 0, q = 1, \cdots, m.$$

where $\mathbf{W}=\sum_{i=1}^{n}\theta_i\mathbf{u}_i\mathbf{u}_i^T$ and $\mathbf{W}_D=\sum_{q=1}^{m}\alpha_q\mathbf{W}_q$. Both $\|\alpha\|_2^2$ and $\|\theta\|_1$ are employed to control the complexity of model. In what follows, γ_1, γ_2 and γ_3 are all positive balance parameters.

Select $L(v_i,v_j,y_{ij})=g(y_{ij}[1-\|v_i-v_j\|_{\mathbf{W}}^2])$ and use the ontology hinge loss for g, that is to say, $g(z)=\max(0,b-z)$ and b is set to be 0. Thus, we deduce the following ontology optimization problem:

$$\arg\min\frac{1}{\binom{N}{2}}\sum_{i<j}g(y_{ij}[1-\|v_i-v_j\|_{\mathbf{W}}^2])+\frac{\gamma_{\mathbf{W}}}{2}\|\mathbf{W}-\mathbf{W}_D\|_F^2+\frac{\gamma_2}{2}\|\alpha\|_2^2+\gamma_3\|\theta\|_1, \tag{3}$$

$$\text{s.t.}\quad\sum_{i=1}^{n}\alpha_q=1,\alpha_q\geq 0,q=1,\cdots,m.$$

For short expressions, we use v_i, x_j and y_{ij} to denote v_k^1, v_k^2 and y_k with $k=1,\cdots,\binom{N}{2}=N'$. Let $\delta_k=v_k^1-v_k^2$ with $\|v_k^1-v_k^2\|_{\mathbf{W}}^2=\sum_{i=1}^{n}\theta_i\delta_k^T\mathbf{u}_i\mathbf{u}_i^T\delta_k=\delta^Tf_k$, $f_k=[f_k^1,\cdots,f_k^n]^T$ and $f_k^i=\delta_k^T\mathbf{u}_i\mathbf{u}_i^T\delta_k$. Therefore, the ontology problem (3) can be re-expressed as

$$\mathop{\mathbf{arg\,min}}_{\alpha,\theta}\frac{1}{N'}\sum_{k=1}^{N'}g(y_k[1-\theta^Tf_k)+\frac{\gamma_{\mathbf{W}}}{2}\|\mathbf{W}-\mathbf{W}_D\|_F^2+\frac{\gamma_2}{2}\|\alpha\|_2^2+\gamma_3\|\theta\|_1, \tag{4}$$

$$\text{s.t.}\quad\sum_{i=1}^{n}\alpha_q=1,\alpha_q\geq 0,q=1,\cdots,m.$$

The answer can be inferred by alternating between two sub ontology problems (minimization $\alpha=[\alpha_1,\cdots,\alpha_m]^T$ and $\theta=[\theta_1,\cdots,\theta_n]^T$ respectively) until its convergence.

Given α, the ontology optimization problem with respect to θ then it can be stated as

$$\arg\min_{\theta}F(\theta)=\Lambda(\theta)+\Omega(\theta) \tag{5}$$

where $\Lambda(\theta)=\frac{1}{N'}\sum_{k=1}^{N'}g(y_k[1-\theta^Tf_k)+\gamma_3\|\theta\|_1$, and $\Omega(\theta)=\frac{\gamma_{\mathbf{W}}}{2}\|\mathbf{W}-\mathbf{W}_D\|_F^2$. Since the ontology loss part $\Lambda(\theta)$ is non-differentiable, we should smooth the ontology loss and then solve (5) in terms of the gradient trick. Let $\Theta=\{x:0\leq x_k\leq 1,x\in\mathbb{R}^{N'}\}$ and σ be the smooth parameter. Then, the smoothed expression of the ontology hinge loss $g(f_k,y_k,\theta)=\max\{0,-y_k(1-\theta^Tf_k)\}$ can be formulated as

$$g_\sigma=\max_{x\in\Theta}x_k(-y_k(1-\theta^Tf_k))-\frac{\sigma}{2}\|f_k\|_\infty x_k^2, \tag{6}$$

where $\|f_k\|_\infty$ term is used as a normalization. In view of setting the objective ontology function of (6) to 0 and projecting x_k on Θ, we infer the following solution: $x_k=\text{median}\{\frac{-y_k(1-\theta^Tf_k)}{\sigma\|f_k\|_\infty},0,1\}$. Furthermore, the piece-wise approximation of g can be expressed as

$$g_\sigma=\begin{cases}0, & y_k(1-\theta^Tf_k)>0\\ -y_k(1-\theta^Tf_k)-\frac{\sigma}{2}\|f_k\|_\infty, & y_k(1-\theta^Tf_k)<-\sigma\|f_k\|_\infty\\ \frac{(y_k(1-\theta^Tf_k))^2}{2\sigma\|f_k\|_\infty}, & \text{Otherwise.}\end{cases} \tag{7}$$

By the computation and deduction, the gradient of the smoothed hinge ontology loss $g_\sigma(\theta)$ is

$$\frac{\partial g_\sigma(f_k,y_k,\theta)}{\partial\theta}=y_kf_kx_k. \tag{8}$$

Let $\mathbf{H}^\Lambda=[f_1,\cdots,f_{N'}]$ and $\mathbf{Y}=diag(y)$. We get $\frac{g_\sigma(\theta)}{\partial\theta}=\sum_k y_kf_kx_k=H^RYx$, and $L^g(\theta)\frac{N'}{\sigma}\max\frac{\|f_kf_k^T\|_2}{\|f_k\|_\infty}$ is the Lipschitz constant of $g_\sigma(\theta)$.

By setting $l(\theta) = \|\theta\|_1$, we infer the approximation of l with the smooth parameter σ' as

$$l'_\sigma = \begin{cases} -\theta_r - \frac{\sigma'}{2}, & \theta_r < -\sigma' \\ \theta_r - \frac{\sigma'}{2}, & \theta_r > \sigma' \\ \frac{\theta_r^2}{2\sigma'}, & \text{Otherwise.} \end{cases}$$

Furthermore, for each $x'_r = \text{median}\{\frac{\theta_r}{\sigma'}, -1, 1\}$, the gradient can be computed by $\frac{\partial \sum_{i=1}^{n} l_{\sigma'}(\theta_r)}{\partial\theta} = x'$ and the Lipschitz constant is denoted as $L^l(\theta) = \frac{1}{\sigma'}$.

Moreover, set $H_{st}^\Omega = \gamma_1 \text{Tr}((\mathbf{u}_s\mathbf{u}_s^T)(\mathbf{u}_t\mathbf{u}_t^T))$ and $f_r^\Omega = \gamma_1 \text{Tr}(\mathbf{W}_S^T(\mathbf{u}_t\mathbf{u}_t^T))$, we have $\frac{\partial\Omega(\theta)}{\partial\theta} = H^\Omega\theta - f^\Omega$, $\frac{\partial F_\sigma(\theta)}{\partial\theta} = \frac{1}{N'}H^R Yx + \gamma c x' + H^\Omega\theta - f^\Omega$ and $L_\sigma = \frac{1}{\sigma}\max_k \frac{\|f_k f_k^T\|_2}{\|f_k\|_\infty} + \frac{\gamma C}{\sigma'} + \|H^\Omega\|_2$ is the Lipschitz constant of $F(\theta)$.

Denote θ^t, y^t and z^t as the solutions in the t-th iteration round, and use $\widehat{\theta}$ as a guessed solution of θ. We obtain that L_σ is the Lipschitz constant of $F_\sigma(\theta)$ and the two attached ontology optimizations are stated as

$$\min_y < \triangledown F_\sigma(\theta^t), y - \theta^t > + \frac{L_\sigma}{2}\|y - \theta^t\|_2^2$$

and

$$\min_z \sum_{i=0}^{t} \frac{i+1}{2}[F_\sigma(\theta^i) + < \triangledown F_\sigma(\theta^i), y - \theta^i >] + \frac{L_\sigma}{2}\|y - \widehat{\theta}\|_2^2,$$

respectively. Set the gradients of the two objective ontology functions in the above two attached ontology problems to be zeros, we yield $y^t = \theta^t - \frac{1}{L_\sigma} \triangledown F_\sigma(\theta^t)$ and $z^t = \widehat{\theta} - \frac{1}{L_\sigma}\sum_{i=0}^{t}\frac{i+1}{2} \triangledown F_\sigma(\theta^i)$. Hence, we deduce $\theta^{t+1} = \frac{2}{t+3}z^t + \frac{t+1}{t+3}y^t$ and the stop criterion is given by $|F_\sigma(\theta^{t+1}) - F_\sigma(\theta^t)| < \varepsilon$.

Given θ, the optimization ontology problem on parameter α can be stated as

$$\arg\min_\alpha \frac{\gamma_1}{2}\|\mathbf{W} - \sum_{p=1}^{m}\alpha_p\mathbf{W}_p\|_F^2 + \frac{\gamma_2}{2}\|\alpha\|_2^2 \tag{9}$$

$$\text{s.t.} \quad \sum_{q=1}^{m}\alpha_q = 1, \alpha_q \geq 0, q = 1, \cdots, m.$$

And, the ontology problem (9) can be expressed in compact form which is stated by

$$\arg\min_\alpha \frac{1}{2}\alpha^T\mathbf{H}\alpha^T + \frac{\gamma_2}{2}\|\alpha\|_2^2 \tag{10}$$

$$\text{s.t.} \quad \sum_{q=1}^{m}\alpha_q = 1, \alpha_q \geq 0, q = 1, \cdots, m.$$

where $f = [f_1, \cdots, f_m]$ with $f_q = \gamma_1\text{Tr}(\mathbf{W}^T\mathbf{W}_q)$, and $\mathbf{H}$ is a symmetric positive semi-definite matrix such that $\mathbf{H}_{st} = \gamma_1\text{Tr}(\mathbf{W}_s^T\mathbf{W}_t)$. We only choose two elements α_i and α_j to update for each iteration. In order to meet the restraint $\sum_{q=1}^{m}\alpha_q = 1$, we get $\alpha_i^* + \alpha_j^* = \alpha_i + \alpha_j$, where α_i^* and α_j^* are the solutions of the current iteration. Then, according to (10) and set $\varepsilon_{ij} = (H_{ii} - H_{ij} - H_{ji} + H_{jj})\alpha_i - \sum_k(H_{ik} - H_{jk})\alpha_k$, we designed the updating rule as follows: $\alpha_i^* = \frac{\gamma_2(\alpha_i+\alpha_j)+(f_i-f_j)+\varepsilon_{ij}}{(H_{ii}-H_{ij}-H_{ji}+H_{jj})+2\gamma_2}$ and $\alpha_j^* = \alpha_i + \alpha_j - \alpha_i^*$. In case the obtained α_i^* and α_j^* don't meet the restraint $\alpha_q \geq 0$, we further set

$$\begin{cases} \alpha_i^* = 0, \alpha_j^* = \alpha_i + \alpha_j \text{ if } \gamma_2(\alpha_i + \alpha_j) + (h_i - h_j) + \varepsilon_{ij} \leq 0 \\ \alpha_j^* = 0, \alpha_i^* = \alpha_i + \alpha_j, \text{ if } \gamma_2(\alpha_i + \alpha_j) + (h_j - h_i) + \varepsilon_{ij} \leq 0. \end{cases}$$

The whole ontology algorithm is stated as follows:

Initialize: $\alpha^{(0)}$, $\theta^{(0)}$, $\gamma_2^{(0)}$ and $\gamma_3^{(0)}$. Set $t = 0$, construct $\mathbf{W}^0 = \sum_{r=0}^{n}\theta_r^{(0)}\mathbf{u}_r\mathbf{u}_r^T$ and $\mathbf{W}_S^0 = \sum_{q=1}^{m}\alpha_q^{(0)}\mathbf{W}_q$.

Iterate:
Optimize $\theta^{(t+1)} \leftarrow \arg\min_\theta \frac{1}{N'}\sum_{k=1}^{N'} g(y_k(1-\theta^T f_k)) + \frac{\gamma_1}{2}\|\mathbf{W}-\mathbf{W}_S^t\|_F^2 + \gamma_3^{(t)}\|\theta\|_t$
and update $\mathbf{W}^{(t+1)} = \sum_{r=1}^{n}\theta_r^{(t+1)}\mathbf{u}_r\mathbf{u}_r^T$;
Optimize $\alpha^{(t+1)} \leftarrow \arg\min_\alpha \frac{\gamma_1}{2}\|\mathbf{W}^{(t+1)}-\mathbf{W}_S\|_F^2 + \frac{\gamma_2^{(t)}}{2}\|\alpha\|_2^2$
and update $\mathbf{W}_S^{t+1} = \sum_{q=1}^{m}\alpha_q^{(t+1)}\mathbf{W}_q$;
Determine $\gamma_3^{t+1} = \frac{|\rho_C|(\frac{1}{N'}\sum_{k=1}^{N'} g(y_k(1-(\theta^{t+1})^T f_k)) + \frac{\gamma_1}{2}\|\mathbf{W}^{(t+1)}-\mathbf{W}_S^{(t)}\|_F^2)}{\|\theta^{(t+1)}\|_1}$;
Obtain $\gamma_2^{(t+1)} = \frac{|\rho_B|(\gamma_1\|\mathbf{W}^{(t+1)}-\mathbf{W}_S^{(t+1)}\|_F^2)}{\|\alpha^{(t+1)}\|_2^2}$;
$t \leftarrow t+1$.
Until convergence.

3 Stability Analysis

In this section, we give the theoretical analysis of our ontology algorithm via stability assumption.

3.1 Uniform stability

Definition 1. (Leave-One-Out) An ontology algorithm has uniform stability β_1 with respect to the ontology loss function l if the following holds

$$\forall s \in Z^m, \quad \forall i \in \{1,\cdots,m\}, \|l(f_s,\cdot) - l(f_{s^i},\cdot)\|_\infty \le \beta_1, \tag{11}$$

where Z is the ontology sample space, f_s is the ontology function determined by the ontology algorithm learning with the set of samples s, and $s^i = \{z_1,\cdots,z_{i-1},z_{i+1},\cdots,z_m\}$ denotes an ontology sample set with the i'-th element z_i deleted.

Definition 2. (Leave-Two-Out) An ontology algorithm has uniform stability β_2 with respect to the ontology loss function l if the following holds

$$\forall s \in Z^m, \quad \forall i \in \{1,\cdots,m\}, \|l(f_s,\cdot) - l(f_{s^{i,j}},\cdot)\|_\infty \le \beta_2, \tag{12}$$

where Z is the ontology sample space, f_s is the ontology function determined by the ontology algorithm learning with the set of samples s, and $s^{i,j}$ is the ontology sample set given from s by deleting two elements z_i and z_j.

For any convex and differentiable ontology function $F : \mathscr{F} \to \mathbb{R}$ as follows (here $\mathscr{F}$ denotes the Hilbert space): $\forall f,g \in \mathscr{F}, B_F(f||g) = F(f) - F(g) - \mathrm{Tr}(< f-g, \bigtriangledown F(g) >)$, we have $\partial F(f) = \{g \in \mathscr{F} | \forall f' \in \mathscr{F}, F(f') - F(f) \ge \mathrm{Tr}(< f'-f, \delta F(f) >)\}$. Let $\delta F(f)$ be any element of $\partial F(h)$. We infer $\forall f,f' \in \mathscr{F}, B_F(f'||f) = F(f') - F(f) - \mathrm{Tr}(< f'-f, \bigtriangledown F(f) >)$, $B_F(f'||f) \ge 0$ and $B_{P+Q} = B_P + B_Q$ for any convex ontology functions P and Q.

Lemma 1. *For any three distance metrics* $\mathbf{W}$ *and* $\mathbf{W}'$, *the following inequality established for any ontology sample* z_i *and* z_j

$$|V(\mathbf{W},z_i,z_j) - V(\mathbf{W}',z_i,z_j)| \le 4LM^2\|\mathbf{W}-\mathbf{W}'\|_F \tag{13}$$

Next, we describe the LOO and LTO stability of our algorithm.

Theorem 2. *Let* β_1 *and* β_2 *be the LOO and LTO stability of our ontology algorithm problem* (2). *Suppose that* $\|v\|_2 \le M$ *for any sample* v. *Then, we have*

$$\beta_1 \le \frac{32L^2M^4}{\gamma_1 N}, \quad \beta_2 \le \frac{64L^2M^4}{\gamma_1 N} \tag{14}$$

where L is the Lipschitz constant of the function g.

Proof. We only present the detailed proof of the first inequality, and the second one can be determined in the similar way. Let $F_{\mathscr{N}}(\theta) = P_{\mathscr{N}}(\theta) + Q(\theta)$, where $P_{\mathscr{N}}(\theta) = \frac{1}{\binom{N}{2}}\sum_{i<j} V(A, z_i, z_j)$ and $Q(\theta) = \frac{\gamma_1}{2}\|\mathbf{W} - \mathbf{W}_S\|_F^2 + \gamma_3\|\theta\|_1$. Clearly, both $P_{\mathscr{N}}(\theta$ and $Q(\theta)$ are convex. Suppose $\theta_{\mathscr{N}}$ and $\theta_{\mathscr{N}'}$ be the minimizers of $F_{\mathscr{N}}(\theta)$ and $F_{\mathscr{N}'}(\theta)$ respectively, where $\mathscr{N}'$ is the set of ontology examples that deletes $z_i \in \mathscr{N}$ from $\mathscr{N}$.

Note that

$$B_{F_{\mathscr{N}}}(\theta_{\mathscr{N}'}||\theta_{\mathscr{N}}) + B_{F_{\mathscr{N}'}}(\theta_{\mathscr{N}}||\theta_{\mathscr{N}'}) \geq B_Q(\theta_{\mathscr{N}'}||\theta_{\mathscr{N}}) + B_Q(\theta_{\mathscr{N}}||\theta_{\mathscr{N}'}).$$

Let $\triangle = \|\theta_{\mathscr{N}'}\|_1 - < \theta_{\mathscr{N}}, \mathrm{sgn}(\theta_{\mathscr{N}'}) > + \|\theta_{\mathscr{N}'}\|_1 - < \theta_{\mathscr{N}'}, \mathrm{sgn}(\theta_{\mathscr{N}}) > \geq 0$, $\mathrm{sgn}(\theta) = [\mathrm{sgn}(\theta_1), \cdots, \mathrm{sgn}(\theta_n)]^T$. Hence, we have $\frac{\partial Q(\theta_{\mathscr{N}})}{\partial \theta}$ where $\delta f(\theta)$ is the sub-gradient of $\|\theta\|_1$ and

$$B_Q(\theta_{\mathscr{N}'}||\theta_{\mathscr{N}}) + B_Q(\theta_{\mathscr{N}}||\theta_{\mathscr{N}'}) = \gamma_1\|\mathbf{W}_{\mathscr{N}'} - \mathbf{W}_{\mathscr{N}}\|_F^2 + \gamma_3\triangle.$$

We have $\partial F_{\mathscr{N}}(\theta_{\mathscr{N}}) = \partial F_{\mathscr{N}'}(\theta_{\mathscr{N}'}) = 0$ since $\theta_{\mathscr{N}}$ and $\theta_{\mathscr{N}'}$ are minimizers of $F_{\mathscr{N}}(\theta)$ and $F_{\mathscr{N}'}(\theta)$. Using Lemma 1, we obtain

$$\begin{aligned}
&\gamma_1\|\mathbf{W}_{\mathscr{N}'} - \mathbf{W}_{\mathscr{N}}\|_F^2 \leq B_{F_{\mathscr{N}}}(\theta_{\mathscr{N}'}||\theta_{\mathscr{N}}) + B_{F_{\mathscr{N}'}}(\theta_{\mathscr{N}}||\theta_{\mathscr{N}'})\\
&= F_{\mathscr{N}}(\theta_{\mathscr{N}'}) - F_{\mathscr{N}}(\theta_{\mathscr{N}}) - < \theta_{\mathscr{N}'} - \theta_{\mathscr{N}}, \partial F_{\mathscr{N}}(\theta_{\mathscr{N}}) > + F_{\mathscr{N}'}(\theta_{\mathscr{N}'}) - F_{\mathscr{N}'}(\theta_{\mathscr{N}'}) - < \theta_{\mathscr{N}} - \theta_{\mathscr{N}'}, \partial F_{\mathscr{N}'}(\theta'_{\mathscr{N}}) >\\
&= F_{\mathscr{N}}(\theta_{\mathscr{N}'}) - F_{\mathscr{N}}(\theta_{\mathscr{N}}) + F_{\mathscr{N}'}(\theta_{\mathscr{N}'}) - F_{\mathscr{N}'}(\theta_{\mathscr{N}'})\\
&= \frac{1}{\binom{N}{2}}\Big(\sum_{\mathscr{N}} V(\mathbf{W}_{\mathscr{N}'}, z_i, z_j) - \sum_{\mathscr{N}} V(\mathbf{W}_{\mathscr{N}}, z_i, z_j) + \sum_{\mathscr{N}'} V(\mathbf{W}_{\mathscr{N}}, z_{i'}, z_j) - \sum_{\mathscr{N}'} V(\mathbf{W}_{\mathscr{N}'}, z_{i'}, z_j)\Big)\\
&\leq \frac{1}{\binom{N}{2}}\Big(\sum_{\mathscr{N}} |V(\mathbf{W}_{\mathscr{N}'}, z_i, z_j) - V(\mathbf{W}_{\mathscr{N}}, z_i, z_j)| + \sum_{\mathscr{N}'} |V(\mathbf{W}_{\mathscr{N}}, z_{i'}, z_j) - V(\mathbf{W}_{\mathscr{N}'}, z_{i'}, z_j)|\Big)\\
&\leq \frac{8LM^2}{N}\|\mathbf{W}_{\mathscr{N}'} - \mathbf{W}_{\mathscr{N}}\|_F.
\end{aligned}$$

This implies that

$$\|\mathbf{W}_{\mathscr{N}} - \mathbf{W}_{\mathscr{N}'}\|_F \leq \frac{8LM^2}{\gamma_1 N}.$$

By virtue of $|V(\mathbf{W}_{\mathscr{N}}, z_i, z_j) - V(\mathbf{W}_{\mathscr{N}'}, z_i, z_j)| \leq 4LM^2\|\mathbf{W}_{\mathscr{N}} - \mathbf{W}_{\mathscr{N}'}\|_F$, we deduce

$$|V(\mathbf{W}_{\mathscr{N}}, z_i, z_j) - V(\mathbf{W}_{\mathscr{N}'}, z_i, z_j)| \leq \frac{32LM^2}{\gamma_1 N}.$$

Therefore, the expected result is obtained. □

Let $\mathscr{N}$ be the ontology sample set and $V(\mathbf{W}, z_i, z_j) = g(y_{ij}[1 - \|v_i - v_j\|_{\mathbf{W}}^2])$. In this sub-section, the empirical ontology risk and expected ontology risk are denoted by $R_{\mathscr{N}}(\mathbf{W}) = \frac{1}{\binom{N}{2}}\sum_{i<j} V(\mathbf{W}, z_i, z_j)$ and $R(\mathbf{W}) = \mathbb{E}(z_i, z_j)[V(\mathbf{W}, z_i, z_j)]$, respectively. We will determine the generalization bound $R(\mathbf{W}) - R_{\mathscr{N}}(\mathbf{W})$ in the next theorem. For this purpose, we should use the following McDiarmid inequality.

Theorem 3. [15] *Let $X_1, \cdots, X_N$ be independent random variables, each taking values in a set A. Let $\phi : A^N \to \mathbb{R}$ be such that for each $i \in \{1, \cdots, N\}$, there exists a constant $c_i > 0$ such that*

$$\sup_{x_1, \cdots, x_N \in A, x_i' \in A} |\phi(x_1, \cdots, x_N) - \phi(x_1, \cdots, x_{i-1}, x_i', x_{i+1}, \cdots, x_N)| \leq c_i. \tag{15}$$

Then for any $\varepsilon > 0$,

$$\mathbf{P}\{\phi(X_1, \cdots, X_N) - \mathbf{E}\{\phi(X_1, \cdots, X_N)\} \geq \varepsilon\} \leq e^{-2\varepsilon^2/\sum_{i=1}^N c_i^2}. \tag{16}$$

The generalization error bound via uniform stability is presented as follows.

Theorem 4. *Let $\mathscr{N}$ be a set of N randomly selected ontology samples and $\mathbf{W}_{\mathscr{N}}$ be the ontology distance matrix determined by (2). With probability at least $1-\delta$, we have*

$$R(\mathbf{W}_{\mathscr{N}})-R_{\mathscr{N}}(\mathbf{W}_{\mathscr{N}})\leq\frac{128L^2M^4}{\gamma_1 N}+\varsigma\sqrt{\frac{\ln\frac{1}{\delta}}{2N}},\tag{17}$$

where

$$\varsigma=\frac{128L^2M^4+4\gamma_1 g_{\mathbf{W}_s}+16\sqrt{2\gamma_1}LM^2\sqrt{g_{\mathbf{W}_s}+\gamma_3\|\gamma_3\|_1}}{\gamma_1}.$$

The method to proof Theorem 4 mainly followed by [16–19], we skip the detailed proof here.

3.2 Strong and weak stabilities

Naturally, the stability in uniform version is too restrictive for most learning algorithms, and only a small number of literatures presented that standard ontology learning algorithms met the uniform stability directly, most of these ontology learning algorithms were uncertain. Thus, we are inspired to consider the other "almost everywhere stability" beyond uniform stability in our ontology setting. We define strong and weak stabilities for our ontology framework which are also good measures to show how robust a ontology algorithm is. We assume $0<\delta_3,\delta_4<1$ in this subsection.

Definition 3. (Strong Stability) Let A be our ontology algorithm whose output on an ontology training sample Z is denoted by f_s, and let l be an ontology loss function. Let $\beta_3:\mathbb{N}\to\mathbb{R}$ and s^i be the ontology sample set which v_i is replaced by $v_i^{'}$. We say that ontology algorithm A has β_3 loss stable with respect to ontology loss l if for all $n\in\mathbb{N}$, $v_i^{'}\in V$, $i\in\{1,\cdots,n\}$, we have,

$$|l(f_s,\cdot)-l(f_{s^i},\cdot)|\leq\beta_3,\tag{18}$$

We say that the ontology algorithm A has strong loss stability β_3 if

$$\mathbb{P}\{A\quad\text{is}\quad\beta_3\quad\text{loss}\quad\text{stable}\quad\text{at}\quad s\}\geq 1-\delta.\tag{19}$$

Definition 4. (Weak Stability) Let A be our ontology algorithm whose output on an ontology training sample Z is denoted by f_s, and let l be an ontology loss function. Let $\beta_4:\mathbb{N}\to\mathbb{R}$. We say that our ontology algorithm A has weak loss stability β_4 if for all $n\in\mathbb{N}$, $i\in\{1,\cdots,n\}$, we have

$$\mathbb{P}\{|l(f_s,\cdot)-l(f_{s^i},\cdot)|\leq\beta_3\}\geq 1-\delta'.$$

We present the following lemma which is a fundamental for proving the results on strong and weak stability.

Lemma 5. (Kutin [22]) *Let $X_1,\cdots,X_N$ be independent random variables, each taking values in a set C. There is a "bad" subset $B\subseteq C$, where $\mathbb{P}(x_1,\cdots,x_N\in B)=\delta$. Let $\phi:C^N\to\mathbb{R}$ be such that for each $k\in\{1,\cdots,N\}$, there exists $b\geq c_k>0$ such that*

$$\sup_{x_1,\cdots,x_N\in C-B,x_k'\in C}|\phi(x_1,\cdots,x_N)-\phi(x_1,\cdots,x_{k-1},x_k^{'},x_{k+1},\cdots,x_N)|\leq c_k,$$

$$\sup_{x_1,\cdots,x_N\in C,x_k'\in C}|\phi(x_1,\cdots,x_N)-\phi(x_1,\cdots,x_{k-1},x_k^{'},x_{k+1},\cdots,x_N)|\leq b.$$

Then for any $\varepsilon>0$,

$$\mathbb{P}\{|\phi(X_1,\cdots,X_N)-\mathbb{E}\{\phi(X_1,\cdots,X_N)\}|\geq\varepsilon\}\leq 2(e^{-\varepsilon^2/8\sum_{i=1}^N c_i^2}+\frac{N^2b\delta}{\sum_{i=1}^N c_k}).$$

Lemma 6. (Kutin [22]) *Let $X_1,\cdots,X_N$ be independent random variables, each taking values in a set C. Let $\phi : C^N \to \mathbb{R}$ be such that for each $k \in \{1,\cdots,N\}$, there satisfies two condition inequalities in Lemma 1 by substituting $\frac{\lambda_k}{N}$ for c_k, and substituting e^{-KN} for δ. If $0 < \varepsilon \le \min_k T(b,\lambda_k,K)$, and $N \ge \max_k \Delta(b,\lambda_k,K,\varepsilon)$, then*

$$\mathbb{P}\{|\phi(X_1,\cdots,X_N) - \mathbb{E}\{\phi(X_1,\cdots,X_N)\}| \ge \varepsilon\} \le 4e^{-\varepsilon^2 N^2/40\sum_{i=1}^N \lambda_i^2}.$$

$$T(b,\lambda_k,K) = \min\{\frac{15\lambda_k}{2}, 4\lambda_k\sqrt{K}, \frac{\lambda_k^2 K}{b}\},$$

$$\Delta(b,\lambda_k,K,\varepsilon) = \max\{\frac{b}{\lambda_k}, \lambda_k\sqrt{40}, 3(\frac{24}{K}+3)\ln(\frac{24}{K}+3), \frac{1}{\varepsilon}\}.$$

The main result in this subsection is stated as follows.

Theorem 7. *Let A be our ontology algorithm whose output on an ontology training sample Z is denoted by f_s. Let l be an ontology loss function such that $0 \le l(f,\cdot) \le \Xi$ for all f and*

$$M^* = 2\beta_3 + \frac{4g\mathbf{w}_s}{N} + [\frac{8\sqrt{2}LM^2(\sqrt{g\mathbf{w}_s + \gamma_3(\|\theta_S\|_1 - \|\theta_{\mathcal{N}}\|_1)} + \sqrt{g\mathbf{w}_s + \gamma_3(\|\theta_S\|_1\|_1)})}{\sqrt{\gamma_1}N}].$$

1) Let β_3 such that our ontology algorithm A has strong loss stability (β_3,δ_1). Then for any $0 < \delta < 1$, with probability at least $1-\delta$, we have

$$R(\mathbf{W}_{\mathcal{N}}) - R_{\mathcal{N}}(\mathbf{W}_{\mathcal{N}}) \le \Xi + \sqrt{8N(M^*)^2 \ln\frac{2(M^*)^2}{8(M^*)^2 - 4N\Xi\delta_1}}.$$

2) Let β_4 such that our ontology algorithm A has weak loss stability (β_4,δ_2). And if

$$0 < \varepsilon \le \min\{\frac{15N(M^*)}{2}, 4NM^*\sqrt{\frac{\ln(1/\delta_1)}{N}}, \frac{N^2(M^*))^2\frac{\text{In}(1/\delta_2)}{N}}{2\Xi}\},$$

and

$$N \ge \max\{\frac{2\Xi}{N(M^*)}, N(M^*)\sqrt{40}, 3(\frac{24N}{\ln(1/\delta_2)}+3)\text{In}(\frac{24N}{\text{In}(1/\delta_2)}+3), \frac{1}{\varepsilon}\}.$$

Then, for any $0 < \delta < 1$, with probability at least $1-\delta$, we have

$$R(\mathbf{W}_{\mathcal{N}}) - R_{\mathcal{N}}(\mathbf{W}_{\mathcal{N}}) \le \Xi + \sqrt{\frac{40N(M^*)^2)^2\ln(\frac{4}{\delta})}{N}}.$$

The method to proof Theorem 7 is mainly followed by [20, 21], we skip the detailed proof here.

4 Experiments

In this section, we design five simulation experiments respectively concerning ontology measure and ontology mapping. In our experiment, we select the ontology loss function as the square loss. To make sure the accuracy of the comparison, we ran our algorithm in C++ through available LAPACK and BLAS libraries for linear algebra and operation computations. We implement five experiments on a double-core CPU with a memory of 8GB.

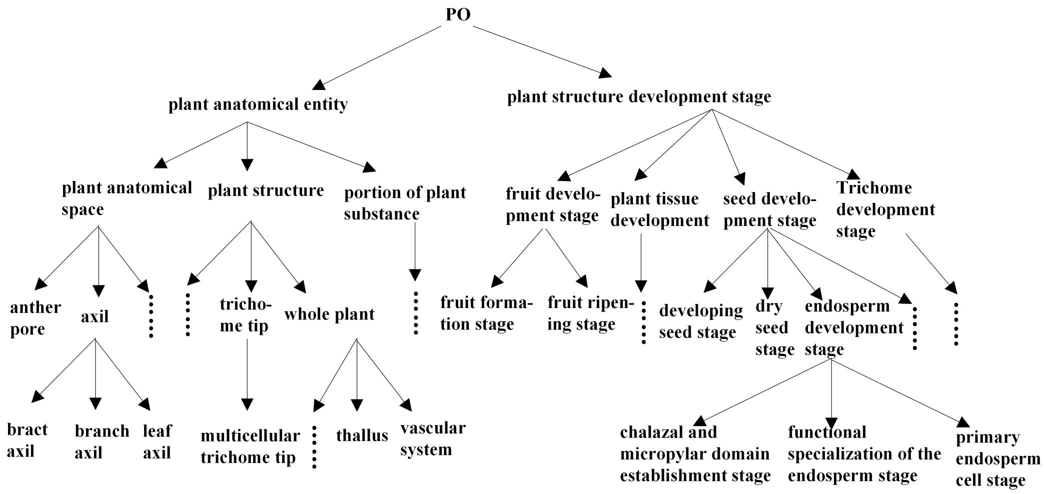


Fig. 1 The Structure of "PO" Ontology.

Table 1 Tab. 1.The Experiment Results of Ontology Similarity measure

	$P@3$ average precision ratio	$P@5$ average precision ratio	$P@10$ average precision ratio
Our Algorithm	0.5358	0.6517	0.8821
Algorithm in [12]	0.4549	0.5117	0.5859
Algorithm in [13]	0.4282	0.4849	0.5632
Algorithm in [14]	0.4831	0.5635	0.6871

4.1 Ontology similarity measure experiment on plant data

We use O_1, a plant "PO" ontology in the first experiment. It was constructed in www.plantontology.org. We use the structure of O_1 presented in Fig. 1. $P@N$ (Precision Ratio see Craswell and Hawking [5]) to measure the quality of the experiment data. At first, the closest N concepts for every vertex on the ontology graph in plant field was given by experts. Then we gain the first N concepts for every vertex on ontology graph by our algorithm, and compute the precision ratio.

Meanwhile, we apply ontology methods in [12], [13] and [14] to the "PO" ontology. Then after getting the average precision ratio by means of these three algorithms, the results with our algorithm are compared. Parts of the data can be referred to Table 1.

When N =3, 5 or 10, the precision ratio gained from our algorithms are a little bit higher than the precision ratio determined by algorithms proposed in [12], [13] and [14]. Furthermore, the precision ratios show it tends to increase apparently as N increases. As a result, our algorithms is proved to be better and more effective than those raised by [12], [13] and [14].

4.2 Ontology mapping experiment on humanoid robotics data

"Humanoid robotics" ontologies O_2 and O_3 are used in the second experiment. The structure of O_2 and O_3 are respectively presented in Fig. 2 and Fig. 3. The leg joint structure of bionic walking device for six-legged robot is presented by the ontology O_2. The exoskeleton frame of a robot with wearable and power assisted lower extremities is presented by the ontology O_3.

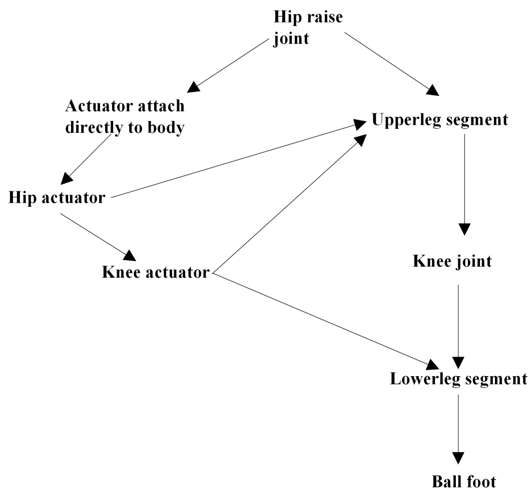


Fig. 2 ‘Humanoid Robotics” Ontology O_2.

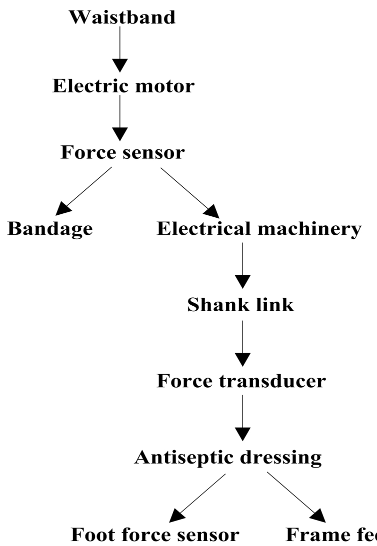


Fig. 3 “Humanoid Robotics” Ontology O_3.

Table 2 Tab. 2. The Experiment Results of Ontology Mapping

	$P@1$ average precision ratio	$P@3$ average precision ratio	$P@5$ average precision ratio
Our Algorithm	0.2778	0.5000	0.7667
Algorithm in [24]	0.2778	0.4815	0.5444
Algorithm in [13]	0.2222	0.4074	0.4889
Algorithm in [14]	0.2778	0.4630	0.5333

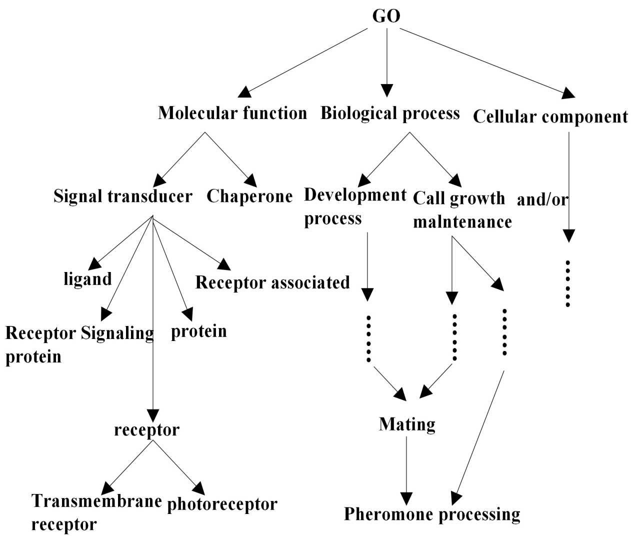


Fig. 4 The Structure of “GO” Ontology.

We set the experiment, aiming to get ontology mapping between O_2 and O_3. $P@N$ Precision Ratio is taken as a measure for the quality of experiment. After applying ontology algorithms in [24], [13] and [14] on “humanoid robotics” ontology and getting the average precision ratio, the precision ratios gained from these three methods are compared. Some results can refer to Table 2.

When $N = 1$, 3 or 5, the precision ratios gained from our new ontology algorithm are higher than the precision ratios determined by algorithms proposed in [24], [13] and [14]. Furthermore, the precision ratios show they tend to increase apparently as N increases. As a result, our algorithms shows much more efficiency than those raised by [24], [13] and [14].

4.3 Ontology similarity measure experiment on biology data

Gene “GO” ontology O_4 is used in the third experiment, which was constructed in the website http: //www. geneontology. We present the structure of O_4 in Figure 4. Again, $P@N$ is chosen as a measure for the quality of the experiment data. Then we apply the ontology methods in [13], [14] and [25] to the “GO” ontology. Then after getting the average precision ratio by means of these three algorithms, the results with our algorithm are compared. Parts of the data can be referred to Table 3.

When $N = 3$, 5 or 10, the precision ratios gained from our ontology algorithms are higher than the precision ratios determined by algorithms proposed in [13], [14] and [25]. Furthermore, the precision ratios show they tend to increase apparently as N increases. As a result, our algorithms turn out to have more effectiveness than those raised by [13], [14] and [25].

Table 3 Tab. 3. The Experiment Results of Ontology Similarity measure

	$P@3$ average precision ratio	$P@5$ average precision ratio	$P@10$ average precision ratio	$P@20$ average precision ratio
Our Algorithm	0.4987	0.6364	0.7602	0.8546
Algorithm in [13]	0.4638	0.5348	0.6234	0.7459
Algorithm in [14]	0.4356	0.4938	0.5647	0.7194
Algorithm in [25]	0.4213	0.5183	0.6019	0.7239

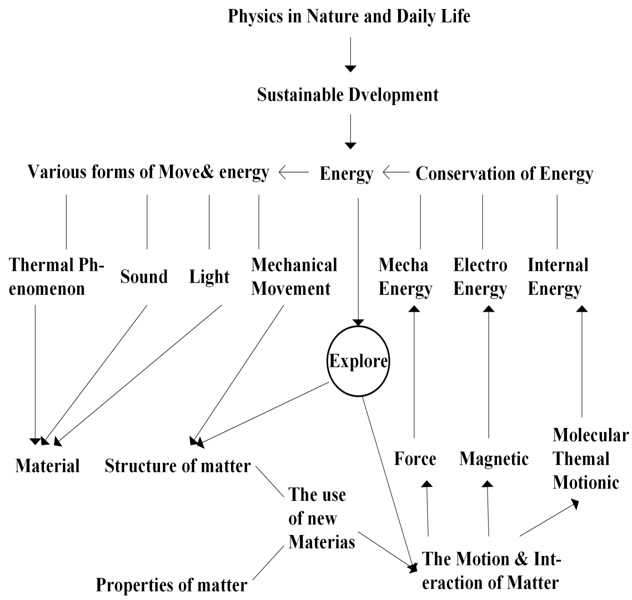


Fig. 5 "Physics Education" Ontology O_5.

4.4 Ontology mapping experiment on physics education data

"Physics education" ontologies O_5 and O_6 are used in the fourth experiment. We respectively present the structures of O_5 and O_6 in Fig. 5 and Fig. 6.

We set the experiment, aiming to give ontology mapping between O_5 and O_6. $P@N$ precision ratio is taken as a measure for the quality of the experiment. Ontology algorithms are applied in [13], [14] and [26] on "physics education" ontology. The precision ratio gotten from the three methods is compared. Some results can be referred to Table 4.

When $N = 1$, 3 or 5, the precision ratio in terms of our new ontology mapping algorithms are much higher than the precision ratio determined by algorithms proposed in [13], [14] and [26]. Furthermore, the precision ratios show they tend to increase apparently as N increases. As a result, our algorithms shows more effectiveness than those raised by [13], [14] and [26].

4.5 Ontology mapping experiment on university data

"University" ontologies O_7 and O_8 are applied in the last experiment. We present the structures of O_7 and O_8 in Fig. 7 and Fig. 8.

We set the experiment, aiming to give ontology mapping between O_7 and O_8. $P@N$ precision ratio is taken

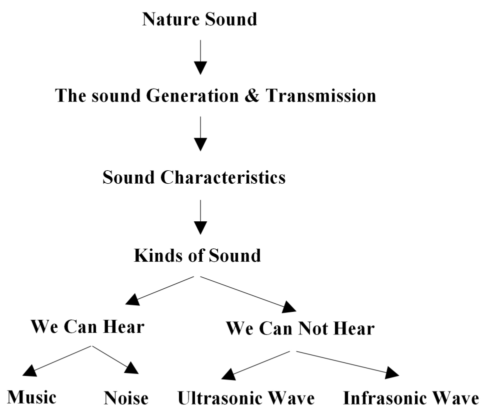


Fig. 6 "Physics Education" Ontology O_6.

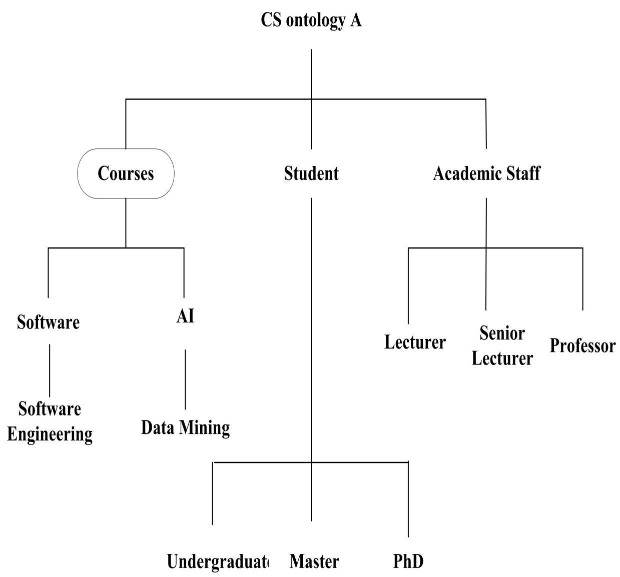


Fig. 7 "University" Ontology O_7.

Table 4 Tab. 4. The Experiment Results of Ontology Mapping

	$P@1$ average precision ratio	$P@3$ average precision ratio	$P@5$ average precision ratio
Our Algorithm	0.6913	0.7556	0.9161
Algorithm in [13]	0.6129	0.7312	0.7935
Algorithm in [14]	0.6913	0.7556	0.8452
Algorithm in [26]	0.6774	0.7742	0.8968

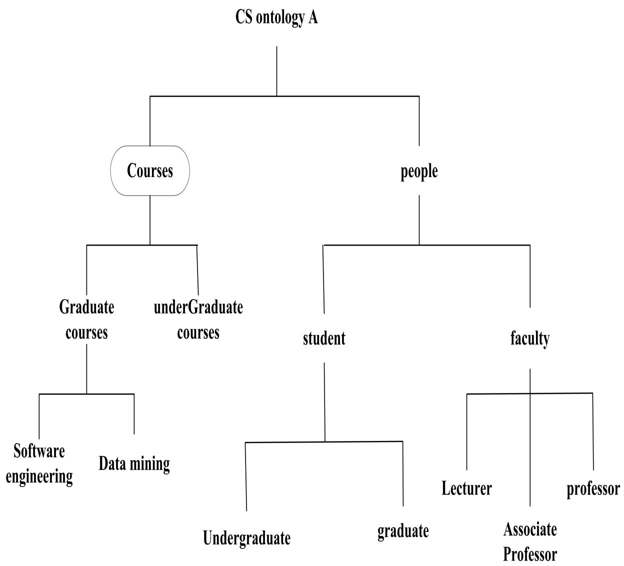


Fig. 8 "University" Ontology O_8.

Table 5 Tab. 5. The Experiment Results of Ontology Mapping

	$P@1$ average precision ratio	$P@3$ average precision ratio	$P@5$ average precision ratio
Our Algorithm	0.5714	0.6786	0.7714
Algorithm in [12]	0.5000	0.5952	0.6857
Algorithm in [13]	0.4286	0.5238	0.6071
Algorithm in [14]	0.5714	0.6429	0.6500

as a criterion to measure the quality of the experiment. Ontology algorithms are applied in [12], [13] and [14] on "University" ontology. The precision ratios gotten from the three methods are compared. Some results can be referred to Table 5.

When $N = 1$, 3 or 5, the precision ratios in terms of our new ontology mapping algorithms are much higher than the precision ratios determined by algorithms proposed in [12], [13] and [14]. Furthermore, the precision ratios show they tend to increase apparently as N increases. As a result, our algorithms turn out to have more effectiveness than those raised by [12], [13] and [14].

5 Conclusions

In this paper, the new ontology learning framework and its optimal approaches are manifested for ontology similarity calculating and ontology mapping. The new ontology algorithm is based on distance function learning tricks. Also, the stability analysis and generalized bounding computation of ontology learning algorithm are presented. Finally, simulation data in five experiments show that our new ontology learning algorithm has high efficiency in these engineering applications. The distance learning based ontology algorithm proposed in our paper illustrates the promising application prospects for multiple disciplines.

Acknowledgements

We thank the reviewers for their constructive comments in improving the quality of this paper. This work was supported in part by the Key Laboratory of Educational Informatization for Nationalities, Ministry of Education, NSFC (No.11401519), and the PhD initial funding of the third author.

References

[1] J.M. Przydzial, B. Bhhatarai, A. Koleti. (2013), GPCR ontology: development and application of a *G* protein-coupled receptor pharmacology knowledge framework, Bioinformatics, 29(24), 3211-3219. doi 10.1093/bioinformatics/btt565
[2] S. Koehler, S.C. Doelken, C.J. Mungall. (2014), The human phenotype ontology project: linking molecular biology and disease through phenotype data, Nucleic Acids Research, 42(D1), 966-974. doi 10.1093/nar/gkt1026
[3] M. Ivanovic, Z. Budimac. (2014), An overview of ontologies and data resources in medical domains, Expert Systems and Applications, 41(11), 5158-5166. doi 10.1016/j.eswa.2014.02.045
[4] A. Hristoskova, V. Sakkalis, G. Zacharioudakis. (2014), Ontology-driven monitoring of patient's vital signs enabling personalized medical detection and alert, Sensors, 14(1), 1598-1628. doi 10.3390/s140101598
[5] M.A. Kabir, J. Han, J. Yu. (2014), User-centric social context information management: an ontology-based approach and platform, Personal and Ubiquitous Computing, 18(5), 1061-1083. doi 10.1007/s00779-013-0720-9
[6] Y.L. Ma, L. Liu, K. Lu, B.H. Jin, X.J. Liu. (2014), A graph derivation based approach for measuring and comparing structural semantics of ontologies, IEEE Transactions on Knowledge and Data Engineering, 26(5), 1039-1052. doi 10.1109/TKDE.2013.120
[7] Z. Li, H.S. Guo, Y.S. Yuan, L.B. Sun. (2014), Ontology representation of online shopping customers knowledge in enterprise information, Applied Mechanics and Materials, 483, 603-606. doi 10.4028/www.scientific.net/AMM.483.603
[8] R. Santodomingo, S. Rohjans, M. Uslar, J.A. Rodriguez-Mondejar, M.A. Sanz-Bobi. (2014), Ontology matching system for future energy smart grids, Engineering Applications of Artificial Intelligence, 32, 242-257. doi 10.1016/j.engappai.2014.02.005
[9] T. Pizzuti, G. Mirabelli, M.A. Sanz-Bobi, F. Gomez-Gonzalez. (2014), Food track & trace ontology for helping the food traceability control, Journal of Food Engineering, 120(1), 17-30. doi 10.1016/j.jfoodeng.2013.07.017
[10] N. Lasierra, A. Alesanco, J. Garcia. (2014), Designing an architecture for monitoring patients at home: ontologies and web services for clinical and technical management integration, IEEE Journal of Biomedical and Health Informatics, 18(3), 896-906.
[11] W. Gao, L.L. Zhu, Y. Guo. (2015), Multi-dividing infinite push ontology algorithm, Engineering Letters, 23(3), 132-139.
[12] Y.Y. Wang, W. Gao, Y.G. Zhang, Y. Gao. (2010), Ontology similarity computation use ranking learning method, The 3rd International Conference on Computational Intelligence and Industrial Application, Wuhan, China, 2010: 20–22.
[13] X. Huang, T.W. Xu, W. Gao, Z.Y. Jia. (2011), Ontology similarity measure and ontology mapping via fast ranking method, International Journal of Applied Physics and Mathematics, 1(1), 54-59. doi 10.7763/IJAPM.2011.V1.11
[14] W. Gao, L. Liang. (2011), Ontology similarity measure by optimizing NDCG measure and application in physics education, Future Communication, Computing, Control and Management, 142, 415-421.
[15] C. McDiarmid. (1989), On the method of bounded differences, in Surveys in Combinatorics, Cambridge University Press, 1989, pp. 148-188.
[16] W. Gao, Y. G. Zhang, L. Liang, Y. M. Xia. (2010), Stability analysis for ranking algorithms, Proceedings 2010 IEEE International Conference on Information Theory and Information Security, Publisher: IEEE, 2010, pp. 973-976. doi 10.1109/ICITIS.2010.5689665
[17] Z. Y. Jia, W. Gao, X. G. He. (2011), Generalization bounds for ranking algorithm via query-level stabilities analysis,

ICCIC 2011, Part VI, CCIS 236, 2011, pp. 197-203.
[18] Y. Gao, W. Gao, Y. G. Zhang. (2011), Query-level stability of IRSVM for replacement case, Procedia Engineering, 15, 2150-2154. doi 10.1016/j.proeng.2011.08.402
[19] Y. Gao, W. Gao, Y. G. Zhang. (2011), Query-level stability of ranking SVM for replacement case, Procedia Engineering, 15, 2145-2149. doi 10.1016/j.proeng.2011.08.401
[20] W. Gao, Y. G. Zhang, Y. Gao, L. Liang, Y. M. Xia. (2013), Strong and weak stability of bipartite ranking algorithms, Proceedings of SPIE - The International Society for Optical Engineering, Doi: 10.1117/12.2030493.
[21] T. W. Xu, Y. G. Zhang, W. Gao. (2011), Generalization bounds for ordinal regression algorithms via strong and weak stability, Energy Procedia, 13, 3471-3478. doi 10.1016/j.egypro.2011.11.499
[22] S. Kutin. (2002), Extensions to McDiarmid's inequality when differences are bounded with high probability, Technical report, Department of Computer Science, The University of Chicago, 2002.
[23] N. Craswell, D. Hawking. (2003), Overview of the TREC 2003 web track, In proceedings of the Twelfth Text Retrieval Conference, Gaithersburg, Maryland, NIST Special Publication, 2003, pp. 78-92.
[24] W. Gao, M.H. Lan. (2011), Ontology mapping algorithm based on ranking learning method, Microelectronics and Computer, 28(9), 59-61.
[25] Y. Gao, W. Gao. (2012), Ontology similarity measure and ontology mapping via learning optimization similarity function, International Journal of Machine Learning and Computing, 2(2), 107-112. doi 10.7763/IJMLC.2012.V2.97
[26] W. Gao, Y. Gao, L. Liang. (2013), Diffusion and harmonic analysis on hypergraph and application in ontology similarity measure and ontology mapping, Journal of Chemical and Pharmaceutical Research, 5(9), 592-598.