



Journal of Computational Mathematica

Journal homepage: www.shcpub.edu.in



Graph Laplacian Based Ontology Regularization Distance Framework for Ontology Similarity Measuring and Ontology Mapping

¹Linli Zhu, ²Yu Pan ³Mohammad Reza Farahani and ⁴Wei Gao

Received on 01 Mar 2017, Accepted on 08 April 2017

ABSTRACT. The core of a large number of ontology engineering applications is similarity measuring, and the essence of ontology mapping and its applications are also similarity computation. There are a great variety of learning techniques are designed for ontology similarity calculating in different engineering circumstances. In these learning settings, all the semantic information of concept is enclosed in the p dimensional vector. In this paper, the ontology similarity measuring is considered via geometric distance calculating between the vectors of ontology vertex corresponding to. We use regularization framework to solute the optimal ontology distance function and apply graph Laplacian to make full use of the unlabeled ontology information. Finally, two simulation experiments imply that our graph Laplacian based ontology regularization distance framework works well in plant and humanoid robotics ontology applications in which higher precision ratio results are obtained from our algorithm than previous ontology learning tricks.

Key Words: Ontology; Similarity computation; Ontology mapping; Graph Laplacian; Regularization framework

1 Introduction

For recent years, the researchers have witnessed ontology's plenty of applications in many fields, like image retrieval, information systems, industrial control, collaboration, as well as biomimetic robot. It's well known for the researchers that ontology, a tool of knowledge representation and storage, is an effective concept semantic computation model. Therefore, ontology also attracts researchers from different other disciplines such as GIS, education science and management, chemistry index calculating, neuroscience, pharmaceuticals and nanotechnology (for instance, see Kim et al., [1], Slota et al., [2], Azevedo et al., [3], Wimmer and Rada [4], Dececchi et al., [5],

¹Corresponding Author: E-mail: zhulinli@jsut.edu.cn (Linli Zhu).

²School of Computer Engineering, Jiangsu University of Technology, Changzhou, Jiangsu - 213001, China

³Department of Applied Mathematics, Iran University of Science and Technology, Narmak, Tehran - 16844, Iran

⁴School of Information Science and Technology, Yunnan Normal University, Kunming 650500, China

Santos [6], Santipantakis and Vouros [7], Euzenat [8], Menzel [9] and Takahashi and Kadobayashi [10]).

In specific, we use a (directed) graph to represent an ontology whose vertex is consistent with a concept. Moreover, the relationship between two concepts is represented by the edge on an ontology graph. It will be quite useful to get the best similarity function in order to calculate the semantic similarity between vertices (concepts), which could be the aims of ontology applications. Therefore, for the reason that it searched the highly similarity vertices among two or more ontology graphs, it's feasible for us to regard the nature of ontology mapping problem as ontology similarity measuring.

Recent years, ontology learning algorithms are attracted much attention from the researchers. Zhu and Gao [14] proposed a new optimization model for ontology similarity measurement and ontology mapping in multi-dividing setting by means of using partial AUC criterion. Gao et al., [16] further studied the ontology algorithm in multi-dividing setting. In his point of view, the empirical multi-dividing ontology framework can be presented as conditional linear statistical, and an approximation result is obtained from the projection method. Based on the assumptions of low noise, Gao et al., [18] obtained the upper bound and lower bound minimax learning rate. Additionally, other theoretical results can be found in Gao et al., [19].

In these ontology learning algorithms, the information of concept is expressed in a p dimensional vector. By slightly confusing the notation, we use v to both express the vertex in ontology graph and its corresponding vector in the following context. Such mathematical representations inspire us to consider whether the similarity between ontology vertex can be computed in terms of geometric distance of their corresponding vectors in high dimension space. The smaller distance between corresponding vectors are the more similarity there is between two ontology vertices and their expression concepts.

In this paper, for this purpose, we report a new ontology optimization framework for ontology similarity measure and ontology mapping by Laplacian based ontology regularized distance function learning. We discuss two points in designing the framework:

- How the empirical ontology data is used in the regularization learning framework?
- Can unlabeled ontology data participate in the learning procedure?

Then, in the next section, we will present the main algorithms are simultaneously give the above two questions positive answers. At last, two experiments are manifested to measure the effectiveness of our new ontology algorithm in plant science and humanoid robotics applications with respect to similarity measuring and ontology mapping, respectively.

2 Algorithm Description

In this section, we propose our main ontology algorithm which is used to determine the geometric distance of two vectors to represent two ontology vertices (i.e., two ontology concepts).

2.1 Using Empirical Ontology Data

Let $\Omega = \{v_1, v_2, \dots, v_n\} \subseteq \mathbb{R}^p$ be the ontology sample data in the p -dimensional vector space. We directly answer the first question raised in the former section. To use the empirical ontology data in the regularization learning framework, we take S and D as the set of similarity and

dissimilarity pair vertices, i.e.,

$$S = \{(v_i, v_j) | \text{vertex } v_i \text{ and } v_j \text{ are similarity}\},$$

$$D = \{(v_i, v_j) | \text{vertex } v_i \text{ and } v_j \text{ are dissimilarity}\}.$$

Note that $(v_i, v_j) \in S$ or D is determined by experts in certain special ontology applications.

Let $d(v_i, v_j)$ be the geometric distance between ontology vertices v_i and v_j , which reflects the similarity between their corresponding concepts. Our goal is to calculate the distance function and then measure the similarities between the ontology vertices. Set $\mathbf{W} \in \mathbb{R}^{p \times p}$ be the symmetric distance metric with its elements which can be stated as follows.

$$d_{\mathbf{W}}(v_i, v_j) = \|v_i - v_j\|_{\mathbf{W}} = \sqrt{(v_i - v_j)^T \mathbf{W} (v_i - v_j)} = \sqrt{\text{Tr}(\mathbf{W}(v_i - v_j)(v_i - v_j)^T)}. \quad (1)$$

where $\text{Tr}(\cdot)$ denotes the trace of matrix. It is easy to check that \mathbf{W} is a valid metric \Leftrightarrow it meets the triangle inequality and non-negativity $\Leftrightarrow \mathbf{W}$ is positive semi-definite (PSD) matrix which is represented as $\mathbf{W} \succeq 0$. Specifically, the matrix \mathbf{W} formulates a classes of Mahalanobis distances on \mathbb{R}^p . In particular, the distance function (1) returns back to the Euclidean distance if $\mathbf{W} = \mathbf{I}$. However, in the real ontology applications, we need to study an optimal distance function from the ontology data instead of using the standard Euclidean function. For this purpose, the formal distance function learning are introduced below.

2.2 Regularization Ontology Learning Model

Our ontology distance function learning problem aims to search a matrix $\mathbf{W} \in \mathbb{R}^{p \times p}$ from ontology data set $\Omega = \{v_1, v_2, \dots, v_n\} \subseteq \mathbb{R}^p$ and two sets S, D can be stated as an optimization framework presented below:

$$\min_{\mathbf{W} \succeq 0} f(\mathbf{W}, S, D, \Omega), \quad (2)$$

where \mathbf{W} is a PSD matrix and f is a certain ontology objective function defined over the fixed ontology data.

The core of solving the ontology distance function learning problem in formulation (2) is how to express a suitable ontology objective function f and then determine an efficient algorithm for solving the solution of optimization problem. The aim of this subsection is to discuss some rules for representing appropriate optimization about the ontology distance function learning problem.

One idea for distance function learning is to minimize the distances between the ontology vertices with similar pairs and maximize the distances between the ontology vertices with dissimilar pairs. We call it a min-max ontology learning technology through which we can formulate the ontology distance function learning problem as a convex optimization framework:

$$\begin{aligned} & \min_{\mathbf{W} \succeq 0} \sum_{(v_i, v_j) \in S} \|v_i - v_j\|_{\mathbf{W}}^2 & (3) \\ \text{s.t.} & \sum_{(v_i, v_j) \in D} \|v_i - v_j\|_{\mathbf{W}} \geq 1. \end{aligned}$$

This framework aims to search the metric \mathbf{W} by minimizing the sum of $\|v_i - v_j\|_{\mathbf{W}}^2$ among the similar ontology vertex pairs and also limit the sum of $\|v_i - v_j\|_{\mathbf{W}}$ among the dissimilar ontology vertex pairs not less than 1.

To improve the ontology distance function learning trick for special ontology engineering, we use the regularization techniques which are applied to strengthen the robustness performance and generalization of the distance function learning in practical ontology applications. As we know, the regularization learning framework has been employed in many learning tasks which can prevent the overfitting situations (see Lim and Hastie [20], Burke and Hoheisel [21], Tarres and Yao [22], El-Nabarawy and Abdelbar [23], and Mozaffari and Azad [24]).

In this paper, our regularization framework for ontology distance metric learning is presented as follows:

$$\min_{\mathbf{W} \succeq 0} g(\mathbf{W}) + \gamma_s L_s(S) + \gamma_d L_d(D), \tag{4}$$

where $g(\mathbf{W})$ is a regularizer related on the target matrix \mathbf{W} , and $L_s(\cdot)$ and $L_d(\cdot)$ are two classes of ontology loss function defined on the sets of similar and dissimilar ontology vertex pairs. γ_s and γ_d are two balance parameters which are used to tradeoff between similar and dissimilar ontology vertex pairs. For instance, these two ontology loss functions can be formulated as follows:

$$L_s(\cdot) = \sum_{(v_i, v_j) \in S} \|v_i - v_j\|_{\mathbf{W}}^2, \tag{5}$$

$$L_d(\cdot) = \sum_{(v_i, v_j) \in D} \|v_i - v_j\|_{\mathbf{W}}^2. \tag{6}$$

On how to choose a suitable regularizer, we use the following Frobenius norm based regularizer:

$$g(\mathbf{W}) = \|\mathbf{W}\|_F = \sqrt{\sum_{i,j=1}^p w_{ij}^2}, \tag{7}$$

where it can avoid too large elements appearing in the matrix \mathbf{W} but it's unable to use the unlabeled ontology information. To solve this defect, we consider the regularization framework which includes the unlabeled data information.

2.3 Using Unlabeled Ontology Data

Now, we focus on how to incorporate the unlabeled ontology data within the above regularization learning framework.

Let $N(v)$ be the set of the nearest neighbors of the ontology vertex v which are determined in view of standard Euclidean distance. For fixed \mathbf{W} with n ontology vertices, the weight matrix $\omega \in \mathbb{R}^{n \times n} = [\omega_{ij}]_{i,j=1}^n$ is determined by the following rule: $\omega_{ij} = 1$ if $v_i \in N(v_j)$ or $v_j \in N(v_i)$, and $\omega_{ij} = 0$ otherwise.

For matrix \mathbf{W} , let $\mathbf{U}^T : \mathbb{R}^p \rightarrow \mathbb{R}^m$ be a kind of corresponding linear mapping with $\mathbf{U} = [\mathbf{u}_1, \dots, \mathbf{u}_m] \in \mathbb{R}^{p \times m}$ such that it stratifies $\mathbf{W} = \mathbf{U}\mathbf{U}^T$ and thus:

$$d(v_i, v_j) = \|\mathbf{U}^T(v_i - v_j)\|^2 = (v_i - v_j)^T \mathbf{U}\mathbf{U}^T(v_i - v_j) = (v_i - v_j)^T \mathbf{W}(v_i - v_j). \tag{8}$$

Clearly, the unlabeled ontology data information is contained in weight matrix ω and we can

apply the Laplacian regularizer described as follows:

$$g(\mathbf{W}) = \frac{1}{2} \sum_{i,j=1}^n \|\mathbf{U}^T v_i - \mathbf{U}^T v_j\|^2 \omega_{ij} \tag{9}$$

$$= \sum_{k=1}^m \mathbf{u}_k^T \mathbf{V}(\mathbf{D} - \omega) \mathbf{V}^T \mathbf{u}_k \tag{10}$$

$$= \sum_{k=1}^m \mathbf{u}_k^T \mathbf{V} \mathbf{L} \mathbf{V}^T \mathbf{u}_k = \text{Tr}(\mathbf{U}^T \mathbf{V} \mathbf{L} \mathbf{V}^T \mathbf{U}) \tag{11}$$

$$= \text{Tr}(\mathbf{V} \mathbf{L} \mathbf{V}^T \mathbf{U} \mathbf{U}^T) = \text{Tr}(\mathbf{V} \mathbf{L} \mathbf{V}^T \mathbf{W}), \tag{12}$$

where $\mathbf{V} = (v_1, \dots, v_n)$ is ontology information matrix, \mathbf{D} is a diagonal matrix with $D_{ii} = \sum_j \omega_{ij}$, $\mathbf{L} = \mathbf{D} - \omega$ is Laplacian matrix defined on the ontology graph, and $\text{Tr}(\cdot)$ is denoted as the trace of matrix.

Now, the Laplacian based ontology regularized distance function learning can be denoted by

$$\min_{\mathbf{W} \geq 0} \text{Tr}(\mathbf{V} \mathbf{L} \mathbf{V}^T \mathbf{W}) + \gamma_s \sum_{(v_i, v_j) \in S} \|v_i - v_j\|_{\mathbf{W}}^2 + \gamma_d \sum_{(v_i, v_j) \in D} \|v_i - v_j\|_{\mathbf{W}}^2. \tag{13}$$

In the very special situation, the dissimilar factor $\gamma_d \rightarrow 0$, the ontology optimization problem (13) will be degenerated to the trivial solution $\mathbf{W} = \mathbf{0}$. To avoid this case, the ontology optimization problem (13) can be further revised to a new one:

$$\begin{aligned} \min_{\mathbf{W} \geq 0} \text{Tr}(\mathbf{V} \mathbf{L} \mathbf{V}^T \mathbf{W}) + \gamma_s \sum_{(v_i, v_j) \in S} \|v_i - v_j\|_{\mathbf{W}}^2 + \gamma_d \sum_{(v_i, v_j) \in D} \|v_i - v_j\|_{\mathbf{W}}^2, \\ \text{s.t. } \log\det(\mathbf{W}) \geq 0, \end{aligned} \tag{14}$$

where the restrictive condition $\log\det(\mathbf{W}) \geq 0$ implies that it can avoid the trivial solution $\mathbf{W} = \mathbf{0}$.

By simply calculating, the complexity of ontology learning algorithm described above is $O(n^2)$ where n is the size of ontology data.

3 Experiments

In this section, two simulation experiments relevance ontology similarity measure and ontology mapping are designed in the following section. We use a vector with p dimension to express each vertex's information to make it realizable to be close to the setting of ontology algorithm. By the way, the vector contains some information like the name, instance, attribute and structure of vertex.

3.1 Experiment on ontology similarity measuring

“PO” ontology, which can be referred to <http://www.plantontology.org> are chosen as a tool in the first experiment, and the basic structure of it is described in Figure 1. What's more, the quality of the experiment is measured by $P@N$ (see Craswell and Hawking [25] for more details).

With the high validity of our algorithm, it shows great prospect for promoting the development of plant ontology. It's well known for the researchers, the plant ontology are often used to represent plant anatomy (such as plant organ, plant cell, whole plant, portion of plant tissue, and vascular system, etc), morphology and stages of development for all plants(such as plant tissue development stage , leaf development stage , whole plant development stage, seed development stage, and sporophyte development stage, etc.). The plant ontology is pretty helpful to build a semantic framework for meaningful cross-species queries which cross the gene expression and phenotype data sets from plant genomics and genetics experiments, which embraces quite promising future.

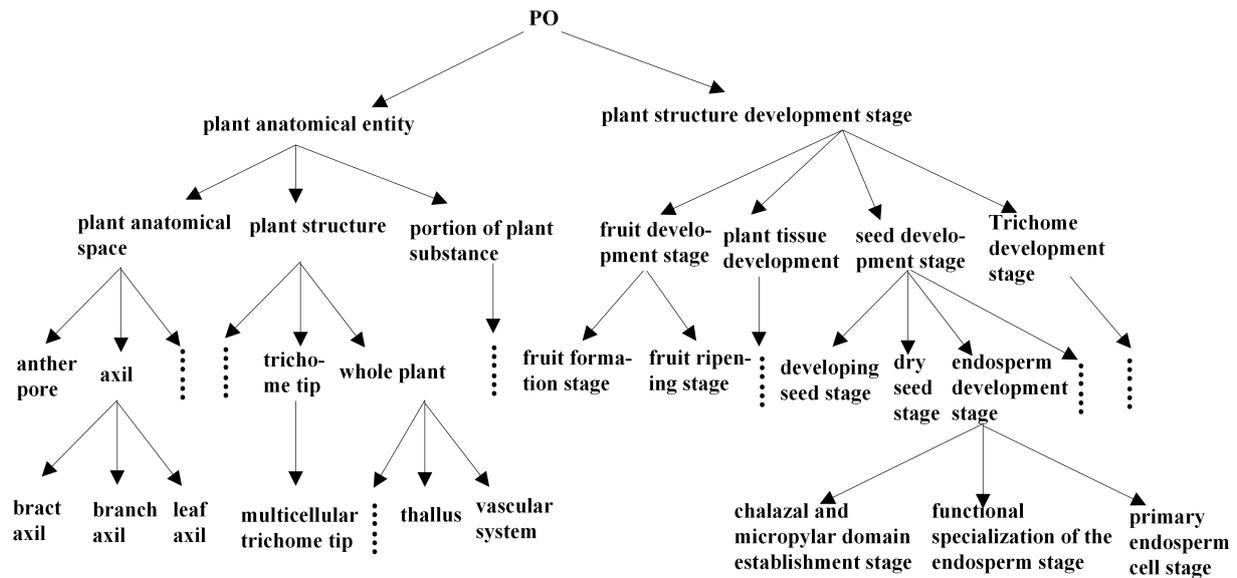


Fig. 1: The structure of “PO” ontology

In order to prove whether our algorithm is effective, we borrow the ontology algorithms raised in Lan et al., [11], Gao and Shi [12] and Gao et al., [13] to work on the ontology graph G . Then, based on $P@N$ criterion, we take the three ways to determine the precision ratio. Parts of the data can refer to Table 1.

Table 1: The experiment results of ontology similarity measure

	$P@3$ average precision ratio	$P@5$ average precision ratio	$P@10$ average precision ratio
Algorithm in our paper	0.4232	0.5293	0.6207
Algorithm in [11]	0.2837	0.3574	0.5425
Algorithm in [12]	0.3546	0.4383	0.6170
Algorithm in [13]	0.3191	0.3872	0.5723

From the Table 1, we can see clearly that when $N = 3, 5$ or 10 , the precision ratio obtained through our algorithm is higher than that obtained by algorithms raised in Lan et al., [11], Gao and Shi [12] and Gao et al., [13]. Moreover, as N is becoming larger, the differences become more obvious between them.

3.2 Experiment on ontology mapping

“Humanoid Robotics” ontologies O_2 and O_3 (the structure of O_2 and O_3 can refer to Gao and Zhu [17]) are chosen for our second experiment. Since the primary task here is to construct ontology mapping between O_2 and O_3 , $P@N$ precision ratio criterion is also taken as a measure to check the effectiveness of our algorithm. Then, we make the ontology algorithms introduced in Lan et al., [11], Gao and Shi [12] and Gao et al., [13] work on an ontology graph, get their precision ratio and compare them. Parts of data can refer to Table 2.

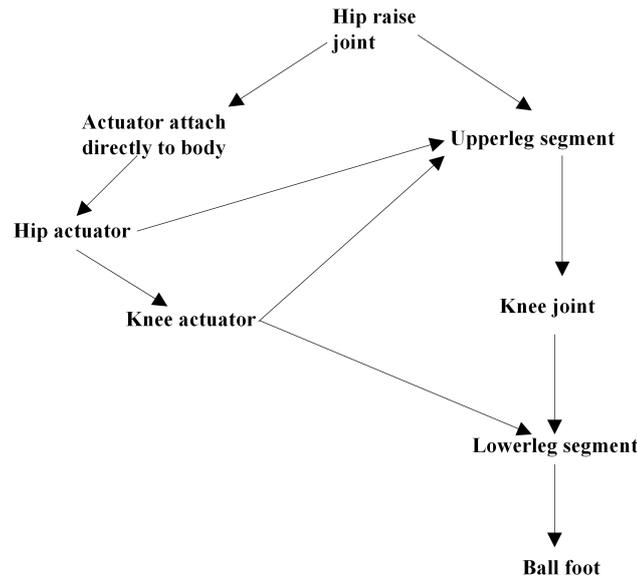


Fig. 2: “humanoid robotics” ontology O_2

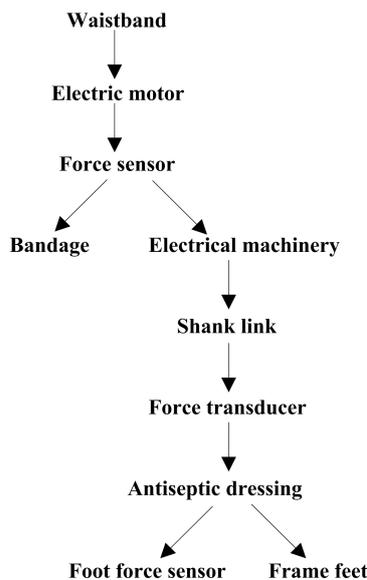


Fig. 3: “humanoid robotics” ontology O_3

From the Table 2, we can see clearly that when $N = 1, 3$ or 5 , the precision ratio obtained through our algorithm is higher than that obtained by algorithms raised in Lan et al., [11], Gao

Table 2: The experiment results of ontology mapping

	$P@1$ average precision ratio	$P@3$ average precision ratio	$P@5$ average precision ratio
Algorithm in our paper	0.1538	0.4074	0.6778
Algorithm in [11]	0.1538	0.3077	0.5231
Algorithm in [12]	0	0.2821	0.4923
Algorithm in [13]	0.0769	0.3077	0.5385

and Shi [12] and Gao et al., [13]. In this way, our algorithm has higher validity in constructing the ontology mapping between two humanoid robotics ontologies, which is pretty obvious.

4 Conclusions

As a validity model for data representation, management and computation, ontology has been employed in a large number of engineering applications and has proved to have high efficiency. In this article, we consider the ontology similarity measuring in view of distance calculating. The two ontology vertices have high similarity if the geometric distance between their corresponding vectors are small. The ontology data includes two special sets: S and D which are denoted as the set of similarity vertices pair and dissimilarity vertices pair determined by experts. The regularization framework is used for ontology learning and the information of unlabeled ontology data is fully participated in the learning process by graph Laplacian technology. Finally, two simulation experiments are manifested to verify the effectiveness of our new algorithm.

Acknowledgment

We thank the reviewers for their constructive comments in improving the quality of this paper. We wish to acknowledge the Open Research Fund by Jiangsu Key Laboratory of Recycling and Reuse Technology for Mechanical and Electronic Products(RRME-KF1612), the Industry-Academia Cooperation Innovation Fund Project of Jiangsu Province(BY2016030-06) and Six Talent Peaks Project in Jiangsu Province(2016-XYDXXJS-020).

References

- [1] H. H. Kim, S. Y. Lee, S. Y. Baik, J. H. Kim, MELLO: Medical lifelog ontology for data terms from self-tracking and lifelog devices, *International Journal of Medical Informatics*, 84(12) (2015) 1099-1110.
- [2] M. Slota, J. Leite, T. Swift, On updates of hybrid knowledge bases composed of ontologies and rules, *Artificial Intelligence*, 229 (2015) 33-104.
- [3] C. L. B. Azevedo, M. E. Iacob, J. P. A. Almeida, M. van Sinderen, L. F. Pires, G. Guizzardi, Modeling resources and capabilities in enterprise architecture: A well-founded ontology-based proposal for archiMate, *Information Systems*, 54(2015) 235-262.

- [4] H. Wimmer, R. Rada, Good versus bad knowledge: Ontology guided evolutionary algorithms, *Expert Systems with Applications*, 42(21) (2015) 8039-8051.
- [5] T. A. Dececchi, J. P. Balhoff, H. Lapp, P. M. Mabee, Toward synthesizing our knowledge of morphology: using ontologies and machine reasoning to extract presence/absence evolutionary phenotypes across studies, *Systematic Biology*, 64(6) (2015) 936-952.
- [6] G. C. Santos, Ontological emergence: how is that possible? towards a new relational ontology, *Foundations of Science*, 20(4) (2015) 429-446.
- [7] G. Santipantakis, G. A. Vouros, Distributed reasoning with coupled ontologies: the E-SHIQ representation framework, *Knowledge and Information Systems*, 45(2) (2015) 491-534.
- [8] J. Euzenat, Revision in networks of ontologies, *Artificial Intelligence*, 228 (2015) 195-216.
- [9] C. Menzel, Necessary beings: an essay on ontology, modality, & the relations between them, *Philosophia Mathematica*, 23(3) (2015) 407-428.
- [10] T. Takahashi, Y. Kadobayashi, Reference ontology for cybersecurity operational information, *Computer Journal*, 58(10) (2015) 2297-2312.
- [11] M. Lan, Y. J. Ren, J. Xu, W. Gao, Ontology similarity computation using k -partite ranking method, *Journal of Computer Applications*, 32(4) (2012) 1094-1096.
- [12] W. Gao, L. Shi, Ontology similarity measure algorithm with operational cost and application in biology science, *The BioTechnology : An Indian Journal*, 8(11) (2013) 1572-1577.
- [13] W. Gao, Y. Gao, L. Liang, Diffusion and harmonic analysis on hypergraph and application in ontology similarity measure and ontology mapping, *Journal of Chemical and Pharmaceutical Research*, 5(9) (2013) 592-598.
- [14] L. L. Zhu, W. Gao, Ontology similarity measuring and ontology mapping based on new optimization model in multi-dividing setting, *Journal of Computational Information Systems*, 11(1) (2015) 377-386.
- [15] W. Gao, T. Xu, Stability analysis for ontology similarity computation, *Abstract and Applied Analysis*, Vol. 2013, 2013, Article ID 174802, 9 pages, <http://dx.doi.org/10.1155/2013/174802>.
- [16] W. Gao, T. W. Xu, J. H. Gan, J. X. Zhou, Linear statistical analysis of multi-dividing ontology algorithm, *Journal of Information and Computational Science*, 11(1) (2014) 151-159.
- [17] W. Gao, L. L. Zhu, Gradient learning algorithms for ontology computing, *Computational Intelligence and Neuroscience*, Volume 2014, Article ID 438291, 12 pages.
- [18] W. Gao, Y. Gao, Y. G. Zhang, L. Liang, Minimax Learning Rate for Multi-dividing Ontology Algorithm, *Journal of Information and Computational Science*, 11(6) (2014) 1853-1860.
- [19] W. Gao, L. Yan, L. Liang, Piecewise function approximation and vertex partitioning schemes for multi-dividing ontology algorithm in AUC criterion setting (I), *International Journal of Computer Applications in Technology*, 50(3/4) (2014) 226-231.
- [20] H. Lim, T. Hastie, Learning interactions via hierarchical group-Lasso regularization, *Journal of Computational and Graphical Statistics*, 24(3) (2015) 627-654.
- [21] J. V. Burke, T. Hoheisel, Matrix support functionals for inverse problems, regularization, and learning, *SIAM Journal on Optimization*, 25(2) (2015) 1135-1159.
- [22] P. Tarres, Y. Yao, Online learning as stochastic approximation of regularization paths: optimality and almost-sure convergence, *IEEE Transactions on Information Theory*, 60(9) (2014) 5716-5735.
- [23] I. El-Nabarawy, A. M. Abdelbar, Advanced learning methods and exponent regularization applied to a high order neural network, *Neural Computing & Applications*, 25(3-4) (2014) 897-910.
- [24] A. Mozaffari, N. L. Azad, Optimally pruned extreme learning machine with ensemble of regularization techniques and negative correlation penalty applied to automotive engine coldstart hydrocarbon emission identification, *Neurocomputing*, 131 (2014) 143-156.
- [25] N. Craswell, D. Hawking, Overview of the TREC 2003 web track, *Proceedings of the Twelfth Text*