

## ONTOLOGY SIMILARITY COMPUTING BASED ON STOCHASTIC PRIMAL DUAL COORDINATE TECHNIQUE

GUOSHUN LIU<sup>1</sup>, ZHIYANG JIA, WEI GAO

**ABSTRACT.** With the extensive application of ontology in the fields of information retrieval and artificial intelligence, the ontology-based conceptual similarity calculation becomes a hot topic in ontology research. The essence of ontology learning is to obtain the ontology function through the learning of ontology samples, so as to map the vertices in each ontology graph into real numbers, and finally determine the similarity between corresponding concepts by the difference between real numbers. The essence of ontology mapping is to calculate concepts from different ontologies. In this paper, we introduce new ontology similarity computing in view of stochastic primal dual coordinate method, and two experiments show the effectiveness of our proposed ontology algorithm.

Mathematics Subject Classification: 60J20, 65L20.

*Key words and phrases:* ontology; similarity measuring; ontology mapping; machine learning.

### 1. Introduction

Ontologies play an increasingly important role in the fields of software engineering, artificial intelligence, information retrieval, and Web services. According to the degree of domain dependence, the ontology can be divided into Top Ontology, Domain Ontology, Task Ontology and Application Ontology. As a kind of description framework and standard specification of knowledge, it is the representation of related concepts and the relationship between concepts. It selects general terms and concepts in the fields described, and uses ontology language

---

Received 06 April 2018. Revised 25 August 2018.

<sup>1</sup> Corresponding Author

© 2018 Guoshun Liu, Zhiyang Jia, Wei Gao. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

to define and describe. Standard universal unambiguous domain ontology. Domain ontology is a tool for knowledge exchange and sharing between designers and systems. It facilitates the processing of structured knowledge and supports semantic operations. It has been widely used in knowledge management and data mining. The introduction of ontology concept promotes the integration of traditional information retrieval mechanism and semantic content, which can effectively improve the search efficiency and search accuracy of knowledge. Several algorithms and related results on ontology can refer to Wang *et al.* [1], Cheng *et al.* [2], Peng *et al.* [3], Zhang *et al.* [4], Zhang *et al.* [5], Zhang *et al.* [6], Xue and Pan [7], Zhong *et al.* [8], Jiang *et al.* [9], and Liang [10].

With the advent of the era of big data, the ontology needs to deal with not only text concepts, but also images, audio and video. Therefore, machine learning algorithms have been widely used in ontology similarity calculations in recent years. A classic method is to learn the ontology function through ontology sample set, this ontology function maps the whole ontology graph to a one-dimensional real number, and each ontology vertex corresponds to a real number. The similarity of the concepts is judged by comparing the distances between the real numbers. Since the information of each ontology vertex is expressed as a  $p$ -dimensional vector, the ontology function can be regarded as a map  $f : \mathbb{R}^p \rightarrow \mathbb{R}$ . In this paper, we present an ontology learning algorithm for ontology similarity measuring and ontology mapping by means of stochastic primal dual coordinate trick.

## 2. Setting

First, we use a  $p$ -dimensional vector to express all the information of a vertex  $v$  (corresponding to a concept). For convenience, we use  $v$  to denote vertex, vector and concept as well. Assume that  $\{v_i\}_{i=1}^n$  is the ontology sample set,  $n$  is the capacity of ontology sample,  $\{l_i\}_{i=1}^n$  are the ontology loss function,  $\beta$  is the ontology sparse vector, and  $g$  is the punish function. In this way, to learn an optimal ontology function can be transform to learn an ontology sparse vector  $\beta$ , and generally the optimization model can be formulated as

$$\min_{\beta \in \mathbb{R}^p} P_\lambda(\beta) = \frac{1}{n} \sum_{i \in \{1, \dots, n\}} l_i(v_i^T \beta) + \lambda g(\beta), \quad (1)$$

where  $\lambda$  is a positive balance parameter. For example, in the setting of mixture norm punishment, the punish function can be expressed as  $g(\beta) = \|\beta\|_1 + \|\beta\|_2^2$ . Let  $\mathbf{V} \in \mathbb{R}^{n \times p}$  be the ontology information matrix which is denoted as  $\mathbf{V} = [v_1, \dots, v_n]^T$ , and its  $j$ -th column is denoted as  $\mathbf{V}_{:j}$ . Let  $\{l_i^*\}_{i=1}^n$  be the convex conjugate function of  $\{l_i\}_{i=1}^n$ , and  $g^*$  be the convex conjugate function of  $g$ . Then, the dual problem of ontology problem (1) can be written as

$$\max_{\alpha} D_\lambda(\alpha) = -\frac{1}{n} \sum_{i \in \{1, \dots, n\}} l_i^*(-\alpha_i) - \lambda g^*\left(\frac{\mathbf{V}^T \alpha}{\lambda n}\right). \quad (2)$$

Assume that ontology loss functions  $l_i$  are  $\frac{1}{\gamma}$ -smooth. Using the standard stochastic primal-dual coordinate method for regularized empirical risk minimization, we can get the following implementation process (see Zhang and Xiao [11] for more details). First, input mini-batch size  $a$ ,  $\beta^0$ ,  $\alpha^0$  as initial solutions,  $\bar{\beta}^0 \leftarrow \beta^0$ ,  $\bar{\alpha}^0 \leftarrow \alpha^0$ . Then, for  $t = 1, 2, \dots$  to converged do: Generate random index from sampling probability  $p$  (in Zhang and Xiao [11],  $p_i = \frac{1}{2n} + \frac{\|v_i\|_2}{2\sum_{k \in \{1, \dots, n\}} \|v_k\|_2}$ )  $a$  times with replacement (denote  $K$  as the set of random indices); set  $\sigma_i$  and  $\tau$  as parameters depend on  $p$ , and  $\theta = \max\{1 - 2\tau\lambda, 1 - (\max_{i \in \{1, \dots, n\}} \frac{1}{ap_i} + \frac{n}{2a\sigma_i\gamma})^{-1}\}$ . Then update the dual and the primal coordinates  $\alpha_i^{t+1} = \operatorname{argmax}_{\mathbf{w} \in \mathbb{R}} \{-\mathbf{w} < v_i, \bar{\beta}^t > -l_i^*(-\mathbf{w}) - \frac{np_i}{2\sigma_i}(\mathbf{w} - \alpha_i^t)^2\}$  if  $i \in K$  (otherwise,  $\alpha_i^{t+1} = \alpha_i^t$ ),  $\bar{\alpha}_i^{t+1} = \alpha_i^t + \frac{(\alpha_i^{t+1} - \alpha_i^t)}{ap_i n}$  for  $i \in \{1, \dots, n\}$ ,  $\beta^{t+1} = \operatorname{argmin}_{\beta \in \mathbb{R}^p} \{\lambda g(\beta) - \frac{1}{n} \beta^T \mathbf{V}^T \bar{\alpha}^{t+1} + \frac{\|\beta - \beta^t\|_2^2}{2\tau}\}$ , and  $\bar{\beta}^{t+1} = \beta^{t+1} + \theta(\beta^{t+1} - \beta^t)$ .

### 3. Ontology learning algorithm using stochastic primal dual coordinate trick

Now, in this section, we introduce the trick of ontology optimization algorithm in terms of stochastic primal dual coordinate techniques. Set dual optimality violation  $\kappa_i^t$  for  $i \in \{1, \dots, n\}$  and each iteration  $t$  as

$$\kappa_i^t = -\alpha_i^t - \nabla l_i(v_i^T \beta^t). \quad (3)$$

Hence, the sampling probability is set as (if  $\kappa_i \neq 0$ )

$$p_i = \frac{1}{|\{i \in \{1, \dots, n\}, \kappa_i^t \neq 0\}|}, \quad (4)$$

and  $p_i = 0$  otherwise. Then, the above stochastic primal dual coordinate based ontology implementation can be modified as follows. Input mini-batch size  $a$ ,  $\beta^0$ , and  $\alpha^0$ ; for  $t = 1, 2, \dots$  to converged do:  $\alpha_i^t \leftarrow \operatorname{update}(\alpha_i^{t-1})$  (for  $i \in \{1, \dots, n\}$ ),  $\beta^t \leftarrow \operatorname{update}(\beta^{t-1})$ , determine  $\kappa^t$  and  $p$  using (4),  $(\hat{\beta}^0, \hat{\alpha}^0) \leftarrow (\beta^t, \alpha^t)$ , after the inner loop do  $(\beta^t, \alpha^t) \leftarrow (\hat{\beta}^u, \hat{\alpha}^u)$  if  $P_\lambda(\hat{\beta}^u) - D_\lambda(\hat{\alpha}^u) < P_\lambda(\beta^t) - D_\lambda(\alpha^t)$ ; for inner loop  $u = 1, \dots, \lceil \frac{n}{a} \rceil$ , randomly pick up a subset indices  $K$ ,  $\hat{\alpha}_i^u \leftarrow \operatorname{update}(\hat{\alpha}_i^{u-1})$  if  $i \in K$ , and  $\hat{\beta}^u \leftarrow \operatorname{update}(\hat{\beta}^{u-1})$ .

By setting  $\psi_j^t = |\beta_j^t - \nabla g_j^*(\frac{\mathbf{V}_j^T \alpha^t}{n\lambda})|$ , the the above stochastic primal dual coordinate based ontology implementation can be further modified as follows. Input mini-batch size  $a$ ,  $\beta^0$ , and  $\alpha^0$ ; for  $t = 1, 2, \dots$  to converged do:  $\alpha_i^t \leftarrow \operatorname{update}(\alpha_i^{t-1})$  (for  $i \in \{1, \dots, n\}$ ),  $\beta^t \leftarrow \operatorname{update}(\beta^{t-1})$ , determine  $\kappa^t$ ,  $\psi_j^t$  and  $p$  using (4),  $(\hat{\beta}^0, \hat{\alpha}^0) \leftarrow (\beta^t, \alpha^t)$ , after the inner loop do  $(\beta^t, \alpha^t) \leftarrow (\hat{\beta}^u, \hat{\alpha}^u)$  if  $P_\lambda(\hat{\beta}^u) - D_\lambda(\hat{\alpha}^u) < P_\lambda(\beta^t) - D_\lambda(\alpha^t)$ ; for inner loop  $u = 1, \dots, \lceil \frac{n}{a} \rceil$ , randomly pick up a subset indices  $K$ ,  $\hat{\alpha}_i^u \leftarrow \operatorname{update}(\hat{\alpha}_i^{u-1})$  if  $i \in K$ , and  $\hat{\beta}_j^u \leftarrow \operatorname{update}(\hat{\beta}_j^{u-1})$  for  $j \in \{m \in \{1, \dots, p\}, \psi_m^t \neq 0\}$ .

At last, one remark is that if we randomly pick up a subset  $M_h$  follows the probability vector  $q$ , then the iterative process presented in last section can be

re-written as follows. First, input dual mini-batch size  $a$ , primal mini-batch size  $b$ ,  $\beta^0$ ,  $\alpha^0$  as initial solutions,  $\bar{\beta}^0 \leftarrow \beta^0$ ,  $\bar{\alpha}^0 \leftarrow \alpha^0$ . Partition the indices of primal variables into  $p \bmod b$  mini-batches whose the number of elements are  $b + 1$  and  $(\lfloor \frac{p}{b} \rfloor - p \bmod b)$  mini-batches whose the number of elements are  $b$ . For  $t = 1, 2, \dots$  to converged do: Generate random index from sampling probability  $p$   $a$  times with replacement (denote  $K$  as the set of random indices); randomly pick up a subset  $M_h$  follows the probability vector  $q$ ; update the dual and the primal coordinates  $\alpha_i^{t+1} = \operatorname{argmax}_{\mathbf{w} \in \mathbb{R}} \{-\mathbf{w} \langle v_i, \bar{\beta}^t \rangle - l_i^*(-\mathbf{w}) - \frac{np_i}{2\sigma_i}(\mathbf{w} - \alpha_i^t)^2\}$  if  $i \in K$  (otherwise,  $\alpha_i^{t+1} = \alpha_i^t$ ),  $\bar{\alpha}_i^{t+1} = \alpha_i^t + \frac{(\alpha_i^{t+1} - \alpha_i^t)}{ap_i n}$  for  $i \in \{1, \dots, n\}$ ,  $\beta_j^{t+1} = \operatorname{argmin}_{x \in \mathbb{R}} \{\lambda g(x) - \frac{x}{n} \langle \mathbf{V} : j, \bar{\alpha}^{t+1} \rangle + \frac{qM_h p}{2\tau |M_h|} (x - \beta_j^t)^2\}$  if  $j \in M_h$  (otherwise,  $\beta_j^{t+1} = \beta_j^t$ ), and  $\bar{\beta}^{t+1} = \beta^{t+1} + \theta(\beta^{t+1} - \beta^t)$ .

#### 4. Experiments

Four simulation experiments relating the ontology similarity measure and ontology mapping are designed below. In order to adjacent to the setting of ontology algorithm, a  $p$ -dimensional vector is applied to express the information of each vertex. The information includes the name, instance, attribute and structure of a vertex. Here the instance of vertex is used to express the set of the reachable vertex in the directed ontology graph. The effectiveness of main ontology algorithm is verified in the following two experiments.

**4.1. Ontology similarity measure experiment on plant data.** In the first experiment, “PO” ontology built in <http://www.plantontology.org>. was adopted to test the efficiency of the proposed new algorithm for ontology similarity measuring. The basic structure of “PO” is shown in Figure 1.  $P@N$  standard was also used for this experiment. Furthermore, ontology methods discussed in [12], [13] and [14] for the “PO” ontology were considered as well. The accuracies of these three algorithms were calculated to compare with that obtained through the algorithm proposed in the paper. Table 1 presents part of the data.

We first give the closest  $N$  concepts for every vertex on the ontology graph by experts in plant field, and then we obtain the first  $N$  concepts for every vertex on ontology graph by the our algorithm, and compute the precision ratio.

	$P@3$ average precision ratio	$P@5$ average precision ratio	$P@10$ average precision ratio
Algorithm in our paper	0.4892	0.5739	0.7103
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

Tab. 1. The experiment results of ontology similarity measure

When  $N = 3, 5$ , or  $10$ , the precision ratio obtained by the proposed algorithm is higher than those determined by the algorithms in [12], [13] and [14]. In

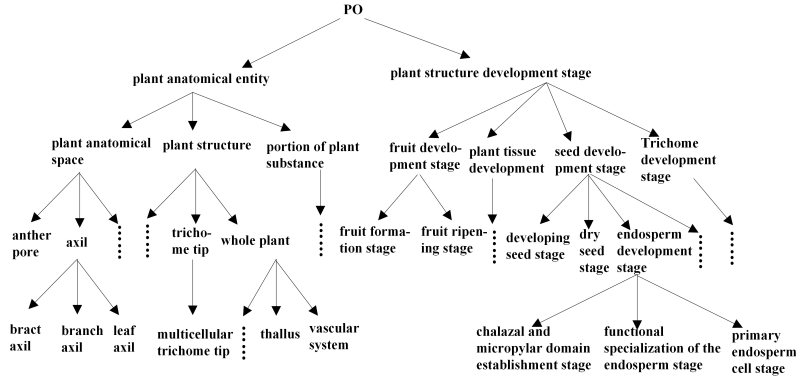


FIGURE 1. The structure of “PO” ontology.

particular, the precision ratios of those algorithms increase significantly as  $N$  increases. It can be concluded that the algorithm described in our paper is superior to those proposed in [12], [13] and [14].

**4.2. Ontology mapping experiment on humanoid robotics data.** Humanoid robotics ontologies  $O_2$  and  $O_3$  which were applied for our second experiment. Figure 2 and Figure 3 exhibit the structures of  $O_2$  and  $O_3$ , respectively. This experiment is to determine ontology mapping between  $O_2$  and  $O_3$  by our algorithm. Likewise, the authors applied  $P@N$  criterion to measure the equality of the experiment. The ontology algorithms used in [15], [13] and [14] to humanoid robotics ontologies were employed. Then the precision ratios of these three methods were compared with that of the proposed algorithm. The experimental results are demonstrated in Table 2.

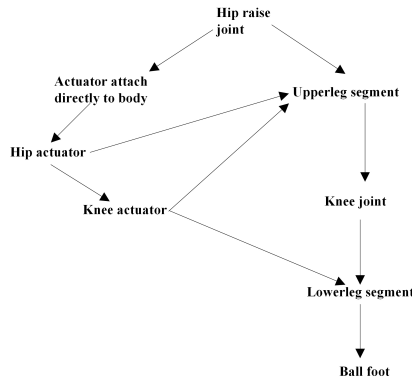
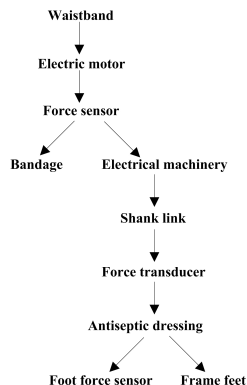


FIGURE 2. “humanoid robotics” ontology  $O_2$ .

FIGURE 3. “humanoid robotics” ontology  $O_3$ .

	$P@1$ average precision ratio	$P@3$ average precision ratio	$P@5$ average precision ratio
Algorithm in our paper	0.2778	0.5000	0.5667
Algorithm in [15]	0.2778	0.4815	0.5444
Algorithm in [13]	0.2222	0.4074	0.4889
Algorithm in [14]	0.2778	0.4630	0.5333

Tab. 2. The experiment results of ontology mapping

The experimental results in Table 2 reveal that our algorithm performs more efficiently than those described in [15], [13] and [14], particularly when  $N$  is large enough.

### Conflict of Interests

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

### REFERENCES

1. Wang, M., Wang, J., Guo, L., & Harn, L. (2018). Inverted XML Access Control Model Based on Ontology Semantic Dependency. *CMC-COMPUTERS MATERIALS & CON-TINUA*, 55(3), 465-482. <https://doi.org/10.3970/cmc.2018.02568>
2. Cheng, X., Xiao, X., & Chou, K. C. (2018). pLoc-mGneg: Predict subcellular localization of Gram-negative bacterial proteins by deep gene ontology learning via general PseAAC. *Genomics*, 110(4), 231-239. <https://doi.org/10.1016/j.ygeno.2017.10.002>
3. Peng, Y., Jiang, Y., & Radivojac, P. (2018). Enumerating consistent sub-graphs of directed acyclic graphs: an insight into biomedical ontologies. *Bioinformatics*, 34(13), 313- 322. <https://doi.org/10.1093/bioinformatics/bty268>
4. Zhang, C., Zheng, W., Freddolino, P. L., & Zhang, Y. (2018). MetaGO: Predicting Gene Ontology of Non-homologous Proteins Through Low-Resolution Protein Structure Prediction and ProteinProtein Network Mapping. *Journal of molecular biology*, 430(15), 2256-2265. <https://doi.org/10.1016/j.jmb.2018.03.004>

5. Zhang, H., Guo, Y., Li, Q., George, T. J., Shenkman, E., Modave, F., & Bian, J. (2018). An ontology-guided semantic data integration framework to support integrative data analysis of cancer survival. *BMC medical informatics and decision making*, 18(2), 41. <https://doi.org/10.1186/s12911-018-0636-4>.
6. Zhang, X., Zhang, Z., Wang, H., Meng, M., & Pan, D. (2018). Metallic materials ontology population from LOD based on conditional random field. *Computers in Industry*, 99, 140-155. <https://doi.org/10.1016/j.compind.2018.03.032>
7. Xue, X., & Pan, J. S. (2018). A Compact Co-Evolutionary Algorithm for sensor ontology meta-matching. *Knowledge and Information Systems*, 56(2), 335-353. <https://doi.org/10.1007/s10115-017-1101-x>
8. Zhong, B., Gan, C., Luo, H., & Xing, X. (2018). Ontology-based framework for building environmental monitoring and compliance checking under BIM environment. *Building and Environment*, 141, 127-142. <https://doi.org/10.1016/j.buildenv.2018.05.046>
9. Jiang, S., Wang, N., & Wu, J. (2018). Combining BIM and Ontology to Facilitate Intelligent Green Building Evaluation. *Journal of Computing in Civil Engineering*, 32(5), 04018039. [https://doi.org/10.1061/\(ASCE\)CP.1943-5487.0000786](https://doi.org/10.1061/(ASCE)CP.1943-5487.0000786)
10. Liang, J. S. (2018). An ontology-oriented knowledge methodology for process planning in additive layer manufacturing. *Robotics and Computer-Integrated Manufacturing*, 53, 28-44. <https://doi.org/10.1016/j.rcim.2018.03.003>
11. Zhang, Y., & Xiao, L. (2017). Stochastic primal-dual coordinate method for regularized empirical risk minimization. *The Journal of Machine Learning Research*, 18(1), 2939-2980.
12. Wang, Y. Y., Gao, W., Zhang, Y. G., & Gao, Y. (2010, November). Ontology similarity computation use ranking learning method. In *The 3rd International Conference on Computational Intelligence and Industrial Application* (pp. 20-22).
13. Huang, X., Xu, T., Gao, W., & Jia, Z. (2011). Ontology similarity measure and ontology mapping via fast ranking method. *International Journal of Applied Physics and Mathematics*, 1(1), 54-59.
14. Gao, W., & Liang, L. (2012). Ontology similarity measure by optimizing NDCG measure and application in physics education. In *Future Communication, Computing, Control and Management* (pp. 415-421). Springer, Berlin, Heidelberg. [https://doi.org/10.1007/978-3-642-27314-8\\_56](https://doi.org/10.1007/978-3-642-27314-8_56)
15. Gao, W., & Lan, M. H. (2011). Ontology mapping algorithm based on ranking learning method. *Microelectronics and Computer*, 28(9), 59-61.

#### **Guoshun Liu**

School of Mathematics and Physics, Jiangsu University of Technology, Changzhou 213001, China.

e-mail: [liuguoshun@jsut.edu.cn](mailto:liuguoshun@jsut.edu.cn)

#### **Zhiyang Jia**

Tourism and culture college, Yunnan University, Lijiang 674100, China.

e-mail: [jdyjazz@gmail.com](mailto:jdyjazz@gmail.com)

#### **Wei Gao**

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

e-mail: [gaoweiyunnan.edu.cn](mailto:gaoweiyunnan.edu.cn)