

# An Optimized Ontology Transfer Learning Method

Hong Tian<sup>1,3</sup>, Yunhong Li<sup>1</sup>, Hongbo Liu<sup>2,3</sup>, Ajith Abraham<sup>3</sup>

<sup>1</sup>School of Software, Dalian Jiaotong University, Dalian 116028, China

<sup>2</sup>School of Information, Dalian Maritime University, Dalian, 116026, China

<sup>3</sup>Machine Intelligence Research Labs - MIR Labs, USA

th@djtu.edu.cn, lhb@dlut.edu.cn, ajith.abraham@ieee.org

**Abstract**—Recently, more and more research is devoted for ontology in the semantic web domain. Firstly, a method for choosing the set of candidate similar concepts is presented based on ontology graphical structural features and data mining. Secondly, a calculation method of conceptual similarity is proposed based on the characteristics of the concept ontology and information content. Finally, the optimized ontology can be transferred into learning. Experimental results illustrate that this method is effective for computing the concept similarity and ontology can be transferred successfully to learn.

**Keywords**—ontology; data mining; concept similarity; transfer learning;

## I. INTRODUCTION

Currently, semantic web technology has become a key focus of the Internet applications. Semantic Internet and Semantic Share are the problems of concept matching and semantic matching in the ontology. Therefore, researching on ontology concept semantic matching is important. With the rapid development of ontology learning techniques, ontology research and development has become a hot spot in the semantic web research. Ontology has been widely used in computer science and information management. As an effective theory and method, and it is successfully applied in building a new intelligent search based information system.

This paper presents a transfer learning method by using optimized ontology transferred learning successfully and has played a positive role in the ontology-based non-structural or semi-structured text information learning.

## II. RELATED WORKS

### A. The concept of ontology

Ontology is the comprehensive explanation for the objective existent system, which core is the abstract nature of objective reality. Gruber [1] defined a typical ontology as  $O = (C, E, R, F, A)$ , where,  $C$  is the concept set,  $E$  is the set of the concept instance,  $R$  is the relationship set defined on the concept set,  $F$  is the sets of functions,  $A$  is the set of axioms.

### B. Related research

Before the ontology transfer learning algorithm is presented, we introduce the most relevant concepts:

- 1) Methods have been proposed based on comprehensive consideration of the data type attributes and object type attributes on language similarity, but its study was confined to the concept similarity calculation method, it should take into account other factors that may affect the final mapping results [2].
- 2) Edge weight is introduced and it is represented the semantic distance, the amount of information concept was introduced, the statistical characteristics of ontology database will be used in the calculation of similarity. At the same time, the depth of the concept, semantic coincidence degree and intensity will be considered in the calculation. But its calculation is in the same ontology [3].
- 3) The method is based on the concept instance, semantic similarity calculation method is improved. However, this method is suitable for tree structure ontology, in the actual generation of ontology, most of the relationships between ontology are multi-graph structure. In practice ontology, as the depth of concept node, density and relations between different concepts will affect the semantic similarity [4].

Therefore, other factors should be also considered in semantic distance calculation. By now, optimized ontology and its application for web services are still not enough studied.

### C. Description of the concept similarity

The two concepts are defined to be similar with each other when they have some common characteristics. Usually the similarity is numerical, value usually belongs to  $[0, 1]$ . It should be special treated when the similarity value is not in this range. There is not any common feature when the similarity value is 0. The two concepts are completely similar when the similarity value is 1. The two concepts have some common characteristics when the similarity value belongs to  $(0, 1)$  [5]. According to the semantic distance and impact factors have been identified and the density of concept is defined in the literature [6], [7], [8]. The following are definitions, which are foundations of similarity calculation and transfer learning methods:

*Definition 1 (Set of candidate similar concepts):* The set of candidate similar concepts are concept collection of two concept nodes with the semantic relating in different graph. Concept node is defined as  $C$ .

*Definition 2 (Path of concept):* In the same concept lattice, the path of concept is edge connecting any two nodes of concepts.  $E = [u, v]$  is defined as the path of concept, where  $u$  and  $v$  are concept nodes. The two concepts are also related.

*Definition 3 (Degree of concept):* In the concept lattice, each node has the path of concept, which is defined as  $degree(C)$ .

*Definition 4 (Density of side):*

$$E_{density} = 2 \times \frac{sum(E)_{C_1} + sum(E)_{C_2} - 1}{sum(E)_C} \quad (1)$$

Where  $sum(E)_C$  is the total numbers of side,  $sum(E)_{C_1}$  and  $sum(E)_{C_2}$  are the sides of node  $C_1$  and node  $C_2$ , respectively.

*Definition 5 (Fineness of concept):* In the concept lattice, the fineness is larger when concept node distribution is concentrated in some of the local context. Thereby, it expresses specific classification of the concept and the greater similarity of concept.

$$T(C) = \frac{(1 - \beta)E}{degree(C)} \quad (2)$$

$$\bar{E} = \frac{\sum_{i=0}^{n-1} e(i)}{n} \quad (3)$$

Where  $degree(C)$  express the degree of concept  $C$ ,  $\bar{E}$  is the average density of edge in concept map and  $\beta$  is regulator.

### III. TRANSFER LEARNING

There are two formal contexts  $(O_1, A_1, F_1)$  and  $(O_2, A_2, F_2)$ , where  $O_1$  and  $O_2$  are the sets of objects,  $A_1$  and  $A_2$  are the sets of attributes,  $F_1$  and  $F_2$  are the sets of relations,  $O(O_1, A_1, F_1)$  express the optimized ontology by formal concept analysis, and  $(O_2, A_2, F_2)$  express the concept lattice by the concept lattice algorithm [9].

#### A. Selecting the set of candidate similar concepts

Some practical factors should be taken into consider when selecting the set of candidate similar concepts. For example, the redundant or conflicting sets of candidate similar concept are easily obtained when concept lattice's scale is too large and the relationship is very complex. With the continued expansion of the candidate space, searching number will also be greatly increased, resulting in low searching efficiency, high complexity and other disadvantages. The different between the concepts and the time complexity can be identified. This paper proposes method as following:

There are ontology  $O(O_1, A_1, F_1)$  and concept lattice  $(O_2, A_2, F_2)$ , only a part of concepts in the concept lattice are similar to the concept in the ontology, the two concepts

will be calculated, the former comes from the concept lattice, and the latter comes from the set of candidate similar concepts. This method can greatly reduce the computational workload and working hours.

The set of candidate similar concepts are chosen as follows: the concept lattice is a graph-like structure and concept lattice can be generated as the composition of the graph by the concept. Each node contains two features, one is the objects feature  $O$  and the other is attributes feature  $A$ . According to Definition 3, the concept nodes are classified by the same concept degree in the concept lattice. Such as concept sets  $B_1, B_2, \dots, B_n$ , where  $B_1$  is a set of nodes, with concept degree 1. Then do the same operation in the ontology. The sets of the concept are obtained. Such as concept sets  $A_1, A_2, \dots, A_m$ . The concept from  $A_1, A_2, \dots, A_m$  is calculated with the other concept, which comes from  $B_1, B_2, \dots, B_n$ . When the elements in the set are so many, some elements can be extracted according to the actual situation. According to the algorithm analysis [10], the similarity threshold is setup as  $T = 0.3$ , this concept will be reserved in the set of candidate similar concepts when the concept similarity value is larger than this threshold. Then each concept in the  $A$  can calculate similarity value with the concept in the  $B$ , similarity formula detailed introduced below. This set is the set of candidate similar concepts.

#### B. Calculating the concept similarity

Considering some factors between the concept and according to the five definitions above, the improved conceptual similarity formula is as follows:

$$rela(A, B) = 1 - \sqrt{\frac{1}{M} \left( \frac{Q}{N_1} + \frac{Q}{N_2} \right)} \quad (4)$$

Where  $M = \frac{(1-\beta)\bar{E}}{N_1} \times \frac{(1-\beta)\bar{E}}{N_2}$ ,  $N_1 = degree(A)$ ,  $N_2 = degree(B)$ ,  $Q = (1 - \beta)\bar{E}$ .

The concept attribute similarity is calculated as follows:  $C_i = \{a_1, a_2, \dots, a_m\}$ , where  $\{a_1, a_2, \dots, a_m\}$  is the attribute set of  $C_i$ ,  $i$  is the number of concepts and  $m$  is the number of attributes in the ontology.  $C_j = \{b_1, b_2, \dots, b_n\}$ , where  $\{b_1, b_2, \dots, b_n\}$  is the attribute set of  $C_j$ ,  $j$  is the number of concepts and  $n$  is the number of attributes in the concept lattice. Then two concepts can be selected from  $C_i$  and  $C_j$  respectively, an attribute can be selected from the form concept, the attribute is calculated with all the attributes from the latter concept according to:

Equ.4, the similarity matrix can be obtained as below:

$$P_{ij} = \begin{bmatrix} p_{11} & p_{12} & \cdots & p_{1n} \\ p_{21} & p_{22} & \cdots & p_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ p_{m1} & p_{m2} & \cdots & p_{mn} \end{bmatrix}$$

Select the maximum value from each line in matrix  $P_{ij}$  and calculate the average. Then by selecting values from the maximum value, which are greater than the average value and comparing these values which are greater than or equal to the threshold (where, threshold is setup as  $T = 0.5$ ). The attribute is retained and the two concepts are similar. Then re-selecting the concept from  $C_i$  and the above process is repeated. The pseudo-code for the proposed transfer learning algorithm is illustrated in Algorithm 1.

- Step 1.** for all the concepts of optimization ontology and the similar concepts set in the concept lattice.
- Step 2.** for all selecting a concept is from the former node and finding out ones which from the latter nodes with the same degree node set respectively, and calculating the similarity value.
- Step 3.** When the similarity value  $\in (0,1)$ , holding the two nodes and identifying the two associated nodes sets respectively. Then finding the same degrees of two nodes from two sets and calculating the similarity value.
- Step 4.** When the two associated nodes have similarity value  $\in (0,1)$ , finishing transfer learning.
- Step 5.** Re-select the nodes to repeat the above process from all the concept nodes of optimization ontology.
- Step 6.** For all other set of candidates similar concepts, using of concept lattice reduction algorithm to eliminate redundant concept lattice.

**Algorithm 1:** Transfer learning algorithm.

#### IV. EXPERIMENT AND DISCUSSIONS

The government work report from 2000 to 2009 years as the experimental data in this paper. The report of the previous five years is optimized ontology, shown in Figure 1. The last five years is the prototype for the concept lattice, shown the figure Figure 2. Each circle presents a concept of ontology nodes, such as  $C_1, C_2, \dots, C_{10}$ , the object features and attributes are next to the concept node. Object attributes are the above braces and object features are the below braces, where i, ii, iii,  $\dots$  express from 2000 to 2009, each concept only give part of the attribute and the rest is omitted.

The experimental algorithm is explored by the Java language, the development platform is Eclipse and the experimental tests are carried out on a computer with CPU 2.99G, Memory 512M. According to Eq.4, using transfer learning algorithm and obtaining government reports ontology from 2005 to 2009 years is shown in Figure 2.

Many factors affect the time complexity. Such as, size of candidate space mapping, search frequency and the complex similarity measurement and so forth. The number of optimal ontology is  $m$ , the number of nodes in concept lattice is  $n$ , where  $m \ll n$ . Making use of the proposing method of conceptual similarity when do not change any circumstances,

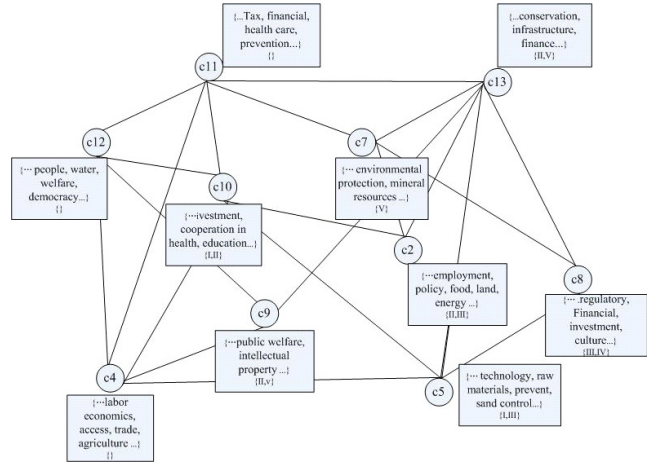


Figure 1. Government work report optimization of the prototype ontology from 2000 to 2004 years.

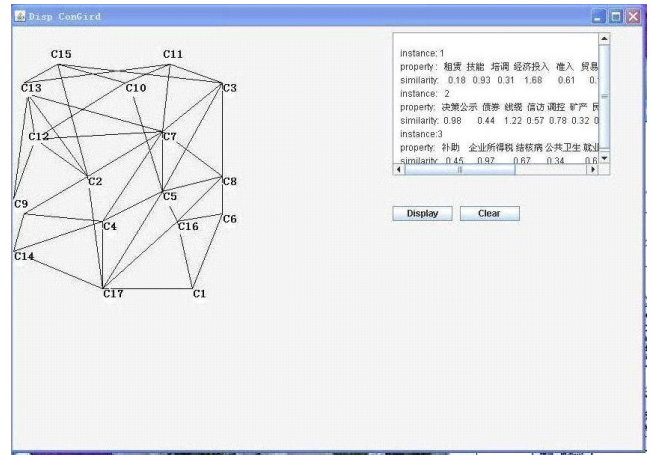


Figure 2. Government work transfer ontology.

the complexity is  $T_1 = O(n)$  in the worst case time. In the set of candidate similar concepts, the number of the concepts is  $s$  instead  $n$ . Therefore, the largest number of the concept comparison is  $m \times s$ , the maximum complexity of similarity computation is  $T_2 = O(m \log(2n))$ , the time complexity is  $T_3 = O(n \log n)$  when finishing transfer learning, Glue algorithm's time complexity is  $T = O(n^2)$ . The integration time complexity of this algorithm is

$$\begin{aligned} T &= O(n) + O(mn \log(2n)) + O(n \log n) \\ &= O(mn \log(2n)) \end{aligned}$$

Therefore, the time complexity of this algorithm is better than the previous results [11].

Figure 2 illustrates a transferring ontology and Figure 3 illustrates an optimized ontology which derived from the same data set. Finding out the nodes with the same degree from two figures and classifying the nodes. Firstly, comparing the nodes associated with this two nodes are whether the

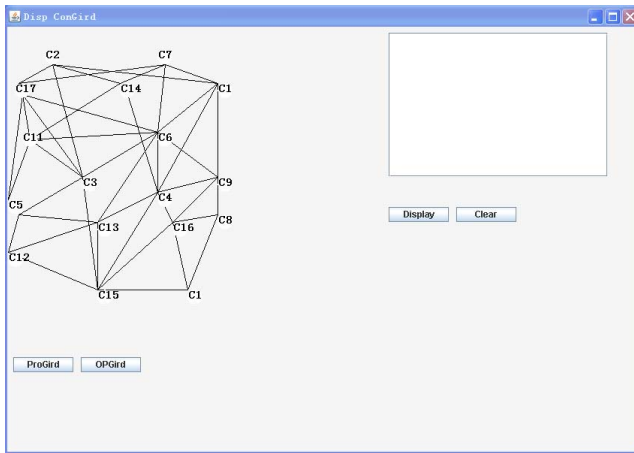


Figure 3. Government work optimized ontology.

same or not, the node is retained when they are the same nodes, otherwise is deleted. Secondly, comparing the two nodes which have the same degree whether has the same number of attributes or not, the node is abandoned when the attribute number of the former node is larger than that of the latter. Finally, identifying the same attributes number of the two nodes with the same similarity degree, then doing subtraction and holding the least absolute value. Using this method, all the concepts in Figure 3 nodes can be found out corresponding concept nodes in Figure 2. From the above two figures, optimized ontology can be transfer learning successfully and effect is obviously.

## V. CONCLUSION

This paper introduced an improved method of conceptual similarity based on ontology graphical structure, the characteristics of the concept and information content. The results shown that the time complexity of this algorithm was better than Glue and Ontology can be transferred learning successfully. Our further research direction is to study transfer learning ontology, and apply it to the ontology learning, so we can establish model of ontology transfer learning.

## ACKNOWLEDGMENT

This work is supported partly by the National Natural Science Foundation of China (No.60873054, 61074029), the Fundamental Research Funds for the Central Universities (2009QN043), Natural Science Foundation of Liaoning Province (No.20102014).

## REFERENCES

[1] T.R. Gruber. "A translation approach to portable ontology specifications". *Knowledge Acquisition*, 1993, 5(2), pp. 39–42.

- [2] G. Nie, X. Zuo, D. Chen. "Improved concept similarity computing approach in ontology mapping". *Computer Applications*, 2008, 28(6), pp. 1563–1565.
- [3] Z. Zhang, H. Zhao, Z. Zhang. "Concept similarity computation based on ontology". *Computer Engineering*, 2009, 35(7), pp. 17–19.
- [4] J. Wang, R. Li, Z. Li. "Research on method of concept semantic similarity based on ontology". *Computer Engineering*, 2007, 33(11), pp. 201–203.
- [5] Z. Gu, Y. Liu, G. Guo. "Improved method of conceptual similarity in ontology mapping". *Computer Engineering and Applications*, 2008, 44(8), pp. 67–69.
- [6] A. Budanitsky, G. Hirst. "Evaluating word net based measures of lexical semantic relatedness". *Computational Linguistics*, 2004, 1(1), pp. 1–49.
- [7] D. Xu, C. Zheng, K. Passi. "Semantic similarity based on the concept of SUMO". *Computer Applications*, 2006, 26(1), pp. 180–183.
- [8] H. Jiang. "Improvement of ontology semantic similarity computation method". *Computer Engineering and Applications*, 2008, 44(36), pp. 143–145.
- [9] R. Wille. "Restructuring lattice theory: an approach based on hierarchies of concepts". *Ordered Sets*, I. Rival (eds.), Reidel, Dordrecht, 1982, pp. 445–470.
- [10] Z. Nie. "Concept similarity analysis in ontology's automatic extraction". *Computer Engineering and Applications*, 2007, 43(26), pp. 159–163.
- [11] A. Doan, J. Madhavan, R. Dhamankar, P. Domingos, A. Halevy. "Learning to match ontologies on the semantic web". *VLDB Journal*, 2003, 12(4), pp. 303–319.