Ontology Similarity Measure and Ontology Mapping via Learning Optimization Similarity Function

Yun Gao and Wei Gao

***Abstract*—Ontology similarity calculation and ontology mapping are important research topics in information retrieval. By learning optimization similarity function, we propose the new algorithm for ontology similarity measure and ontology mapping. The stability of ontology algorithms is studied by adopting a strategy which adjusts the sample set by deleting one or two element from it. Relationship between uniform loss stability and uniform score stability is investigated. A sufficient condition for uniform score stability is given. The result of our work shows that if for any $(v,v') \in V\times V$, a kernel function $K((v,v'), (v,v'))$ has a limited upper bound, then the ontology algorithm which minimizes the regularization empirical *l*-error will have good uniform score stability. Also, two experiments results show that the proposed algorithm has high accuracy and efficiency for similarity calculation and ontology mapping.**

***Index Terms*—Ontology, similarity calculation, ontology mapping.**

I. Introduction

The generalized information retrieval refers to the process of organizing and storing the information by certain means, and in accordance with the needs of users to find out related information. The narrow information retrieval refers to the process of only finding out the needed information in information set, and it is equivalent to the information search which is often said. Information set refers to an organized information collection, it can be all the database records, also may be all literatures collected in the library. Of course, it can also be all types of information set released through Internet. Text information retrieval is to find the appropriate number of documents subset related to the requests and inquires in a large number of documents set.

Ontology abstracts certain application field of the real world into a set of concepts and relationships of concepts. Integrating the ontology into the technology of text information retrieval not only inherit the advantages of information retrieval but also overcome the limitations that concepts information retrieval cannot deal with the relationships of the concepts. As the ontology has the ability to express concept semantics through the relationship between concepts, portray the intrinsic link between concepts, and excavate those hidden and not clear concepts and information. So, it can better meet user requirements in the recall and precision aspects, and realize the retrieval intelligent. At the same time, ontology-based retrieval methods are more in line with the of human thought, it can overcome the shortcomings of the information redundancy or information missing caused by the traditional information retrieval methods, and the query results can be more reasonable. Now, ontology similarity computation is widely used in medical science biology science [see 1] and social science [see 2]. As ontology used in information retrieval [see 3], every vertex can be regard as a concept of ontology, measure the similarity of vertices using the information of ontology graph.

Let graphs $G_1,G_2,\ldots,G_k$ corresponding to ontologies $O_1,O_2,\ldots, O_k$, respectively, and $G=G_1+G_2+\ldots+G_k$. For every vertex $v\in V(G_i)$, where $1\le i\le k$, the goal of ontology mapping is finding similarity vertices from G-G_i. So, the ontology mapping problem is also ontology similarity measure problem. Choose the parameter $M\in[0,1]$, let A,B are two concepts on ontology and Sim(A,B)>M, then return B as retrieval expand when search concept A. So, the ontology mapping problem is also ontology similarity measure problem. Thus, the key trick for ontology similarity measure and ontology mapping is to find the best similarity function f: $V\times V \to \mathbb{R}^{+} \cup \{0\}$, which maps each pair of vertices to a non-negative real number. More details can be seen in [4-16].

The main contribution of our paper is proposed a new ontology similarity measure method and ontology mapping using the learning method. The organization of this paper is as follows: we describe the algorithm in Section II, and theory analysis in Section III. Two experiments are obtained in Section IV which shows that the new algorithm have high quality.

II. The new Algorithm To Learn Optimization Similarity Function

The input space is called the instance space, and all the instances in input space are drawn randomly and independently based on some (unknown) distribution. The labels in the finite label set are totally ordered and let y_i is label for some (v_i, v_i'), where v_i and v_i' are vertices in ontology graph or in multi-ontology graph. Without loss of generality, the training examples take labels Y=[0,M] for some M > 0. Then, a training set S can be constructed as $S=\{(v_1, v_1', y_1),\cdots, (v_n, v_n', y_n)\}$ of size n in $V\times V\times Y$, the new ontology algorithms try to find a mapping function f: $V\times V \to R^{+} \cup \{0\}$, which is can assign an accurate label to future instances.

Manuscript received March 7, 2012, revised March 28, 2012. The work is supported by National Natural Science Foundation of China [60903131].

Yun Gao is with the Department of Editorial, Yunnan Normal University, Kunming, Yunnan, China. (E-mail: gy64gy@sina.com).

Wei Gao is with the Department of Information, Yunnan Normal University, Kunming, Yunnan, China. Also he is PhD student in the Department of Mathematics, Soochow University, Suzhou, Jiangsu, China (E-mail: gaowei@ynnu.edu.cn).

Let l be loss function which is used to punish the inconsistent situation of wrong prediction. One common used loss function as absolute loss $l(f, v_i, v_i^{'}, y_i)=\left| f(v_i, v_i^{'}) - y_i \right|$.

The quality of a prediction rule f: $V\times V \rightarrow R^+ \cup \{0\}$ is measured by its expected error with respect to D:

$$R_l(f) = \mathbf{E}_{(v_i, v_i^{'}, y_i)\sim D}[l(f, v_i, v_i^{'}, y_i)]. \quad (1)$$

However, it cannot be estimated directly since distribution D unknown. Instead, we use empirical error to measure the algorithm:

$$\widehat{R}_l(f;S) = \frac{1}{n}\sum_{i=1}^{n} l(f, v_i, v_i^{'}, y_i). \quad (2)$$

Nevertheless, the similarity function given by the minimum $\widehat{R}_l(f;S)$ usually has bad smoothness. It is general to add a regularization term to solve this problem, and use the regularized empirical l-error with regularization parameter λ>0 to measure f:

$$\widehat{R}_l^{\lambda}(f;S) = \widehat{R}_l(f;S) + \lambda N(f), \quad (3)$$

where F is a class of real-valued functions on X, and N is a regularization function: $F \rightarrow R^+ \cup \{0\}$. Given a training sample S, a ranking algorithm outputs a ranking function $f_S \in F$ which satisfies:

$$f_S = \arg\min_{f\in F} \widehat{R}_l^{\lambda}(f;S). \quad (4)$$

III. Stability Analysis

An algorithm is called stable at a training set S if any change of a single point in S yields only a small change in the output. It is natural to consider that a good ranking algorithm should have good stability, that is, for a wild change of samples, the ranking function doesn't change too much. Some analyses of the stability of ranking algorithms are given in [17]-[36].

[36] studied the stability of ontology algorithm by replacing one element in sample set S with another outside of S. Also, [36] obtained some conclusions relatively on such stabilities. In this paper, we consider two modifications on training sets: 1) $S^{\backslash i'}$ is the set formed by removing the i' th element from S; 2) $S^{\backslash i',j'}$ formed by removing the i'th and j'th elements from S. Denote

$$\widehat{R}_l^{\backslash i'}(f;S^{\backslash i'}) = \frac{1}{n}\sum_{i\neq i'} l(f, v_i, v_i^{'}, y_i),$$

$$\widehat{R}_l^{\backslash i',j'}(f;S^{\backslash i',j'}) = \frac{1}{n}\sum_{i\neq i',j'} l(f, v_i, v_i^{'}, y_i).$$

They are called "leave-one-out" error and "leave-two-out" error, respectively. Note $\widehat{R}_l^{\lambda,\backslash i'}(f;S^{\backslash i'}) = \widehat{R}_l^{\backslash i'}(f;S^{\backslash i'}) + \lambda N(f)$ and $\widehat{R}_l^{\lambda,\backslash i',j'}(f;S^{\backslash i',j'}) = \widehat{R}_l^{\backslash i',j'}(f;S^{\backslash i',j'}) + \lambda N(f)$.

A. Some Definitions

Definition 1 (Uniform loss stability for leave one out). Let A be a ontology algorithm who's output on a training sample S denoted by f_S, and let l be a loss function. Let $\beta_1(n): N \rightarrow R$. We say that A has $\beta_1(n)$ loss stable with respect to l at S if for all $n\in N$, $S\in (V\times V\times Y)^n$ and $i'\in \{1,\cdots,n\}$ we have for all $(v,v',y) \in V\times V\times Y$,

$$\left| l(f_S, v, v', y) - l(f_{S^{\backslash i'}}, v, v', y)\right| \leq \beta_1(n).$$

Definition 2 (Uniform score stability for leave one out). Let A be a ontology algorithm whose output on a training sample S denoted by f_S. Let $\mu_1(n): N \rightarrow R$. We say that A has score stability $\mu_1(n)$ at $S\in (V\times V\times Y)^n$ if for all $n\in$ □ , and $i'\in \{1,\cdots,n\}$, we have for all $(v,v',y) \in V\times V\times Y$,

$$\left| f_S(v,v') - f_{S^{\backslash i'}}(v,v')\right| \leq \mu_1(n).$$

Definition 3 (Uniform loss stability for leave two out). Let A be a ontology algorithm whose output on a training sample S denoted by f_S, and let l be a loss function. Let $\beta_2(n): N \rightarrow R$. We say that A has $\beta_2(n)$ loss stable with respect to l at S if for all $n\in N$, $S\in (V\times V\times Y)^n$, and $i',j'\in \{1,\cdots,n\}$, $i'\neq j'$, we have for all $(v,v',y) \in V\times V\times Y$,

$$\left| l(f_S, v, v', y) - l(f_{S^{\backslash i',j'}}, v, v', y)\right| \leq \beta_2(n).$$

Definition 4 (Uniform score stability for leave two out). Let A be a ontology algorithm whose output on a training sample S denoted by f_S. Let $\mu_2(n): N \rightarrow R$. We say that A has score stability $\mu_2(n)$ at $S\in (V\times V\times Y)^n$ if for all $n\in$ □ , and $i',j'\in \{1,\cdots,n\}$, $i'\neq j'$, we have for all $(v,v',y) \in V\times V\times Y$,

$$\left| f_S(v,v') - f_{S^{\backslash i',j'}}(v,v')\right| \leq \mu_2(n).$$

Our error analysis provides a learning rate of algorithm when the loss function is σ-admissible.

Definition 5. Let l be a loss function, σ>0, and F be a class of real-valued functions on $V\times V$. We say that l is σ-admissibility with respect to F if for any f_1, $f_2\in F$ and $(v,v') \in V\times V$, we have

$$\left| l(f_1, v, v', y) - l(f_2, v, v', y)\right| \leq \sigma \left| f_1,(v,v') - f_2,(v,v')\right|.$$

B. Main Results and Proof

Lemma 1. Let l be a ontology loss such that $l(f; v, v', y)$ is convex in f. Let F be a convex class of real-valued functions on $V\times V$, and let $\sigma > 0$ be such that l is σ-admissible with respect to F. Let $\sigma > 0$, and let $N : F \rightarrow R^+ \cup \{0\}$ be a functional defined on F such that for sample set $S=\{(v_1, v_1^{'}, y_1), \cdots, (v_n, v_n^{'}, y_n)\}\in (V\times V\times Y)^n$, the regularized empirical l-error $\widehat{R}_l^{\lambda}(f;S)$ has a minimum (not necessarily unique) in F. Let A be a ontology algorithm for ontology defined by (1), and let $(v_i, v_i^{'}, y_i) \in S$, $i'\in \{1,\cdots,n\}$, $j'\in \{1,\cdots,n\}$, $i'\neq j'$. For brevity denote $f \equiv f_{S,\lambda}$, $f_i \equiv f_{S^{\backslash i'},\lambda}$, $f_{i,j} \equiv f_{S^{\backslash i',j'},\lambda}$, $\Delta_1 f = (f_i - f)$ and $\Delta_2 f = (f_{i,j} - f)$.

Then for any $t \in [0, 1]$,

$N(f)-N(f+t\Delta_1 f)+N(f_i)-N(f_i - t\Delta_1 f)$

$$\le \frac{t\sigma}{\lambda n}\sum_{i\neq i'}(|\Delta_1 f(v_i,v_i^{'})|+|\Delta_1 f(v_{i'},v_{i'}^{'})|),$$

$N(f)-N(f+t\Delta_2 f)+N(f_{i,j})-N(f_{i,j} - t\Delta_2 f)$

$$\le \frac{t\sigma}{\lambda n}\sum_{i\neq i', i\neq j'}(|\Delta_1 f(v_i,v_i^{'})|+|\Delta_1 f(v_{i'},v_{i'}^{'})|+|\Delta_1 f(v_{j'},v_{j'}^{'})|).$$

Proof. Recall that a convex function g verity:

$\forall x, y, \forall t \in [0,1]$ $g(x+t(y-x))-g(x) \le t(g(y)-g(x))$.

Since $l(f; v, v', y)$ is convex in f, then $\hat{R}_l^{\backslash i'}(f;S^{\backslash i'})$ and $\hat{R}_l^{\backslash i',j'}(f;S^{\backslash i',j'})$ are convex in f. By the convexity of $\hat{R}_l^{\backslash i'}(f;S^{\backslash i'})$, we know that for any $t \in [0, 1]$,

$$\hat{R}_l^{\backslash i'}(f+t\Delta_1 f;S^{\backslash i'}) - \hat{R}_l^{\backslash i'}(f;S^{\backslash i'}) \le t(\hat{R}_l^{\backslash i'}(f_i;S^{\backslash i'}) - \hat{R}_l^{\backslash i'}(f;S^{\backslash i'})) \quad (2)$$

and also (interchanging the roles of f and f_i)

$$\hat{R}_l^{\backslash i'}(f_i+t\Delta_1 f;S^{\backslash i'}) - \hat{R}_l^{\backslash i'}(f_i;S^{\backslash i'}) \le t(\hat{R}_l^{\backslash i'}(f;S^{\backslash i'}) - \hat{R}_l^{\backslash i'}(f_i;S^{\backslash i'})) \quad (3)$$

Adding Eqs (2) and (3), we get

$$\hat{R}_l^{\backslash i'}(f+t\Delta_1 f;S^{\backslash i'}) - \hat{R}_l^{\backslash i'}(f;S^{\backslash i'}) + \hat{R}_l^{\backslash i'}(f_i+t\Delta_1 f;S^{\backslash i'}) - \hat{R}_l^{\backslash i'}(f_i;S^{\backslash i'}) \le 0. \quad (4)$$

Now, since F is convex, then $(f+t\Delta_1 f) \in F$ and $(f_i+t\Delta_1 f) \in F$. Since f minimizes $\hat{R}_l^{\lambda}(f;S)$ in F and f_i minimizes $\hat{R}_l^{\lambda,\backslash i'}(f;S^{\backslash i'})$ in F, then

$$\hat{R}_l^{\lambda}(f;S) - \hat{R}_l^{\lambda}(f+t\Delta_1 f;S) \le 0, \quad (5)$$

$$\hat{R}_l^{\lambda,\backslash i'}(f_i;S^{\backslash i'}) - \hat{R}_l^{\lambda,\backslash i'}(f_i+t\Delta_1 f;S^{\backslash i'}) \le 0. \quad (6)$$

Adding Eqs.(4),(5),(6) we get

$\lambda(N(f)-N(f+t\Delta_1 f)+N(f_i)-N(f_i - t\Delta_1 f))$

$$\le \frac{t\sigma}{\lambda n}\sum_{i\neq i'}(|\Delta_1 f(v_i,v_i^{'})|+|\Delta_1 f(v_{i'},v_{i'}^{'})|).$$

Similarly, by the convexity of $\hat{R}_l^{\backslash i',j'}(f;S^{\backslash i',j'})$, we can get

$N(f)-N(f+t\Delta_2 f)+N(f_{i,j})-N(f_{i,j} - t\Delta_2 f)$

$$\le \frac{t\sigma}{\lambda n}\sum_{i\neq i', i\neq j'}(|\Delta_1 f(v_i,v_i^{'})|+|\Delta_1 f(v_{i'},v_{i'}^{'})|+|\Delta_1 f(v_{j'},v_{j'}^{'})|).$$

Thus, we complete the proof. □

Let F be a reproducing kernel Hilbert space (RKHS) of real-valued functions on $V\times V$, with kernel K: $(V\times V) \times (V\times V) \to \mathbb{R}$. Then the reproducing property of F gives that for all $f\in F$ and $(v,v')\in V\times V$,

$$|f(v,v')| = |\langle f, K_{(v,v')}\rangle_K|, \quad (7)$$

where $K_{(v,v')}: (V\times V) \to \mathbb{R}$ is defined as

$$K_{(v,v')}(v_1,v_1^{'}) = K((v,v'),(v_1,v_1^{'})),$$

and $\langle \cdot,\cdot \rangle_K$ denotes the RKHS inner product in F. It is easy to know that for all $f\in F$ and $(v,v')\in V\times V$,

$$f(v,v') \le \|f\|_K \|K_{(v,v')}\|_K \le \|f\|_K \sqrt{K((v,v'),(v,v'))} \quad (8)$$

where $\|\cdot\|_K$ denotes the RKHS norm in F. We consider ontology algorithms that perform regularization in the RKHS F using the squared norm in F as a regularizer. Specifically, let N: $F \to \mathbb{R}^+ \cup \{0\}$ be the regularizer defined by

$$N(f) = \|f\|_K^2.$$

Our result in this section shows that if the kernel K is a such $K((v,v'),(v,v'))$ which bounded for all $(v,v')\in V\times V$, then a ontology algorithm that minimizes an appropriate regularized error over F, with regularizer N as defined above, has good uniform score stability.

Theorem 1. Let F be an RKHS with kernel K such that for all $(v,v')\in V\times V$, $K((v,v'),(v,v')) \le \kappa^2$. Let l be a loss function such that $l(f;v,v',y)$ is convex in f and l is σ-admissible with respect to F. Let $\sigma>0$, and let A be a ontology algorithm that, given a training sample S, outputs a similarity function $f_S\in F$ that satisfies (4). Then A has uniform score stability $\mu_1(n)$ and $\mu_2(n)$, where for all $n\in N$,

$$\mu_1(n) = \frac{2(n-1)\sigma\kappa^2}{\lambda n}, \ \mu_2(n) = \frac{2(n-2)\sigma\kappa^2}{\lambda n}.$$

Proof. Let $n\in N$, $S=\{(v_1, v_1^{'}, y_1), \cdots, (v_n, v_n^{'}, y_n)\} \in (V\times V\times Y)^n$; $i', j' \in \{1, \cdots, n\}$, $i' \ne j'$. For the "leave-oneout" part of Theorem 2, applying Lemma 1 with t=1/2, by $N(f) = \|f\|_K^2$, we get (using the notation of Lemma 1) that

$$\|f\|_K^2 - \left\|f+\frac{1}{2}\Delta_1 f\right\|_K^2 + \|f_i\|_K^2 - \left\|f_i-\frac{1}{2}\Delta_1 f\right\|_K^2 \le \frac{\sigma}{2\lambda n}\sum_{i\neq i'}(|\Delta_1 f(v_i,v_i^{'})|+|\Delta_1 f(v_{i'},v_{i'}^{'})|). \quad (9)$$

By F is a vector space, $\Delta_1 f \in F$, $f+\frac{1}{2}\Delta_1 f \in F$, and $f_i-\frac{1}{2}\Delta_1 f \in F$. So, $\left\|f+\frac{1}{2}\Delta_1 f\right\|_K$ and $\left\|f_i-\frac{1}{2}\Delta_1 f\right\|_K$ are well defined. It is easy to see that

$$\|f\|_K^2 - \left\|f+\frac{1}{2}\Delta_1 f\right\|_K^2 + \|f_i\|_K^2 - \left\|f_i-\frac{1}{2}\Delta_1 f\right\|_K^2 = \frac{1}{2}\|\Delta_1 f\|_K^2$$

Combined with Eq.(9), this gives

$$\frac{1}{2}\|\Delta_1 f\|_K^2 \le \frac{\sigma}{2\lambda n}\sum_{i\neq i'}(|\Delta_1 f(v_i,v_i^{'})|+|\Delta_1 f(v_{i'},v_{i'}^{'})|).$$

Since $\Delta_1 f \in F$, we get

$$\frac{1}{2}\|\Delta_1 f\|_K^2 \le \frac{\sigma}{2\lambda n}\|\Delta_1 f\|_K \times \sum_{i\neq i'}(\sqrt{K((v_i,v_i^{'}),(v_i,v_i^{'}))} + \sqrt{K((v_{i'},v_{i'}^{'}),(v_{i'},v_{i'}^{'}))}).$$

Which gives

$$\|\Delta_1 f\|_K \le \frac{2(n-1)\sigma\kappa}{\lambda n}.$$

Thus, for all $(v,v')\in V\times V$,

$$\left| f_S(v,v') - f_{S^{\setminus i}}(v,v') \right| \le \frac{2(n-1)\sigma\kappa^2}{\lambda n}.$$

The proof of "leave-two-out" part is similar to that of "leave-one-out" part. □

The following simple lemma show the relationship between uniform loss stability and uniform score stability:

Lemma 2. Let *F* be a class of real-valued functions on $V\times V$, and let *A* be a ontology algorithm for ontology that given a sample set *S*, outputs a similarity function $f_S\in F$. If *A* has uniform score stability $\mu_1(n)$, $\mu_2(n)$ and *l* is a loss function that is σ-admissible with respect to *F*, then *A* has uniform loss stability $\beta_1(n)$ and $\beta_2(n)$ with respect to *l*, where for all $n\in N$,

$$\beta_1(n) = \sigma\ \mu_1(n),\ \beta_2(n) = \sigma\ \mu_2(n).$$

Using Theorem 1 and Lemma 2, we can immediate get the following corollary:

Corollary 1. Under the conditions of Theorem 1 and Lemma 3, A has uniform loss stability $\beta_1(n)$ and $\beta_2(n)$ with respect to *l*, where for all $n\in N$,

$$\beta_1(n) = \frac{2(n-1)\sigma^2\kappa^2}{\lambda n},\ \beta_2(n) = \frac{2(n-2)\sigma^2\kappa^2}{\lambda n}.$$

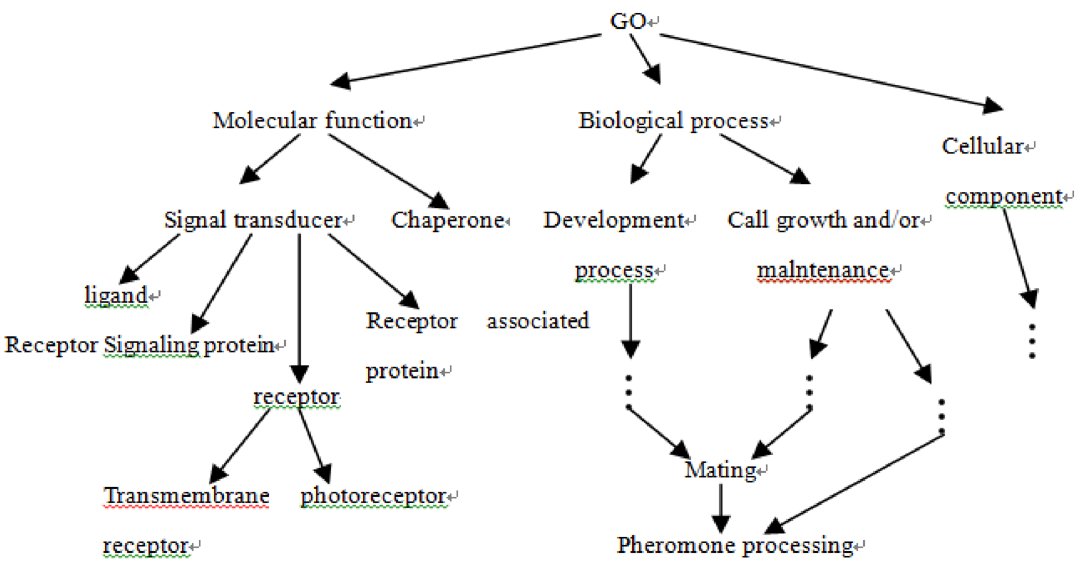

Fig. 1. "GO" ontology

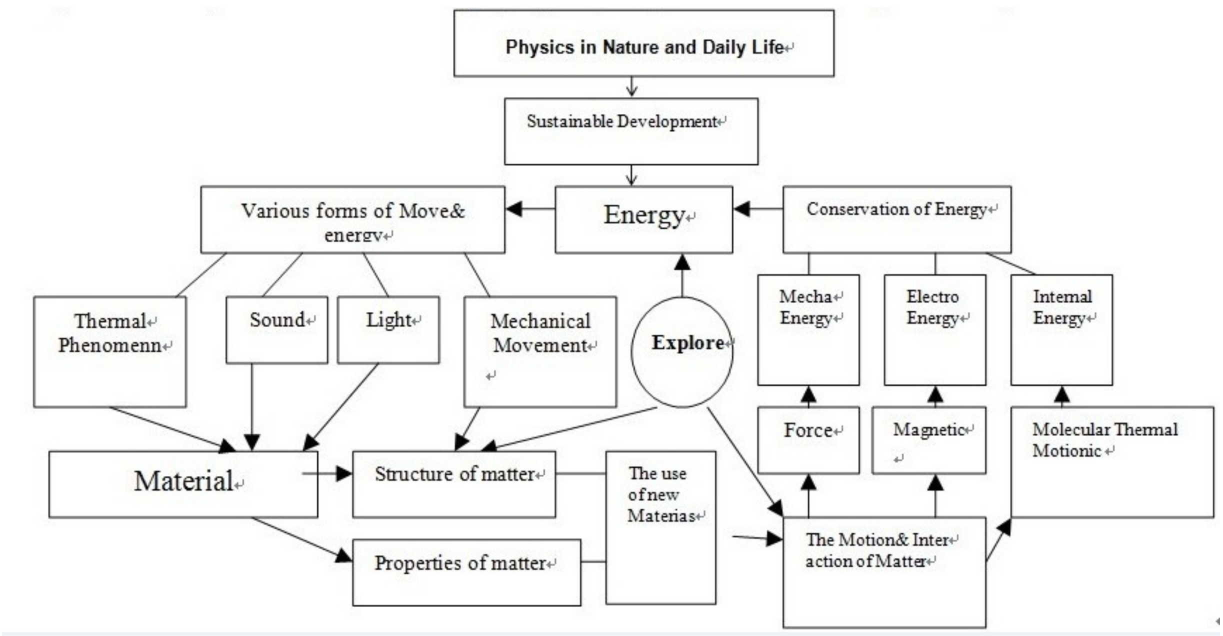

Fig.2. Physics education Ontology O_1

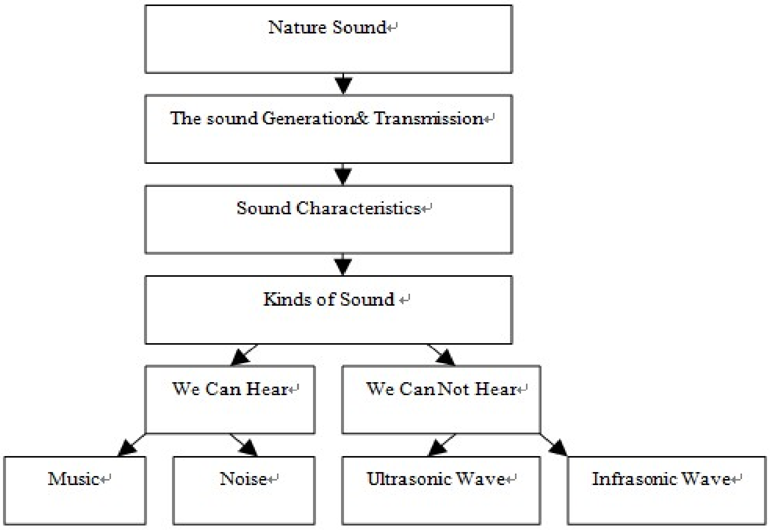

Fig.3. Physics education Ontology O_2

IV. EXPERIMENTS

To connect ontology to this ontology algorithm, we should use a vector to express the vertex of information. This vector contains the information of name, instance, attribute and structure of vertex, where the instance of vertex is the set of its reachable vertex in the directed ontology graph.

We use absolute loss $l(f, v_i, v_i^{'}, y_i)=\left| f(v_i, v_i^{'}) - y_i \right|$ as loss function and $N(f) = \| f \|_K^2$ as regularizer. Regularization parameter λ is determined by expert.

The first experiment concerns ontology similarity measurement is described as follows. In this experiment, we use computer ontology O_1 which was constructed in [37], Fig. 1 show O_1. We use *P@ N* (Precision Ratiosee [9]) to measure the equality of the experiment. First, the expert gives the first *N* concepts for every vertex on the ontology graph, and then we obtain the first *N* concepts for every vertex on ontology graph by the algorithm and compute the precision ratio.

The experiment shows that, *P@*1 Precision Ratio is 55. 73%, *P@*3 Precision Ratio is 68. 36%, *P@*5 Precision Ratio is 79. 78%. Thus the proposed algorithm has high efficiency for ontology measure.

For the second experiment, we use another Physics Education Ontologies O_2 and O_3 which constructed in [11], as Fig. 2 shows O_2 and Fig. 3 shows O_3. The goal of this experiment is given ontology mapping between O_2 and O_3. We also use *P@ N* Precision Ratio to measure the equality of experiment.

The experiment shows that, *P@*1 Precision Ratio is 53. 42%, *P@*3 Precision Ratio is 65. 79%, *P@*5 Precision Ratio is 78. 36%. Thus the proposed algorithm has high efficiency for ontology mapping.

V. CONCLUSION

In this paper, we give a new algorithm for measuring the ontology similarity and ontology mapping via learning optimization similarity function. The stability analysis concerns on "leave one out" and "leave two out" are given. Also, the new algorithms have high quality according to the experiments above.

ACKNOWLEDGMENT

First we thank the reviewers for their constructive comments in improving the quality of this paper. We also would like to thank the anonymous referees for providing us with constructive comments and suggestions.

REFERENCES

[1] P. Lambrix, and A. Edberg, "Evaluation of ontology tools in bioinformatics", *In: Paci.c Symposium on Biocomputing*, 2003, pp. 529-600 .

[2] A. Bouzeghoub, and A. Elbyed, Ontology mapping for web-based educational systems interoperability, *Interoperability in Business Information Systems*, vol.1, pp. 73-84, 2006.

[3] X. Su, and J.A. Gulla, "Semantic enrichment for ontology mapping", in Proc.The 9th International Conference to Information Systems (NLDB). 2004, pp.217-228.

[4] B. Hu, S. Dasmahapatra, P. Lewis, and N. Shadbolt, "Ontology-based medical image annotation with description logics", 15th IEEE International Conference on Tools with Artificial Intelligence, 2003, pp. 77-82.

[5] S. Liu , L. Chia, and S. Chan, "Ontology for naturescene image retrieval", In on the move to meaningful Internet systems, CoopIS, DOA, and ODBASE, 2004, pp. 1050-1061.

[6] P. Mork, A. Philip, and Bernstein, "Adapting a generic match algorithm to align ontologies of human anatomy", In: 20th International Conf. on Data Engineering. IEEE, Los Alamitos. 2004 , pp.787-790.

[7] J. Opitz, B. Parsia, and U. Sattler, "Using Ontologies for Medical Image Retrieval- An Experiment", Proceedings 5th Int. Workshop on OWL: Experiences and Directions, 2009, pp. 23-24.

[8] I. Spasic, S. Ananiadou, J. McNaught, and A. Kumar, "Text mining and ontologies in biomedicine: making sense of raw text", Briefings in bioinformatics, vol. 6, pp. 239-251, 2006.

[9] H. Wang, S. Liu, and L. Chia, "Does ontology help in image retrieval?-A comparison between keyword, text ontology and multi-modality ontology approaches", MM06, October, Santa Barbara, California, USA. ACM, 2006, pp. 109-112.

[10] X. Huang, T. Xu, W. Gao, and Zhiyang Jia, "Ontology Similarity Measure and Ontology Mapping Via Fast Ranking Method", International Journal of Applied Physics and Mathematics, Vol. 1, No. 1, pp.54-59, July 2011.

[11] W.Gao and L. Liang. "Ontology Similarity Measure by Optimizing NDCG Measure and Application in Physics Education", Future Communication, Computing, Control and Management, LNEE 142, pp. 415–421.
[12] X. Huang, T. Xu, W. Gao, and Zhiyang Jia, "A Fast Algorithm for Ontology Similarity Measure and Ontology Mapping", Proceedings of 2011 4th IEEE international conference on computer science and Information technology, June ,2011, Chengdu, China, pp.567-570.
[13] X. Huang, T.i Xu, W. Gao and Shu Gong, "Ontology Similarity Measure and Ontology Mapping Using Half Transductive Ranking", Proceedings of 2011 4th IEEE international conference on computer science and Information technology, June ,2011, Chengdu, China, pp.571-574.
[14] M. Lan, J.Xu, and W. Gao, "Ontology Mapping Algorithm Based on Primal RankRLS", 2011 3rd international comference on E-business and information system security, may, 2011, Wuhan, China, IEEE Press, Vol 2. pp.788-790.
[15] M. Lan, J. Xu, and Wei Gao, "Ontology Similarity Measure Based on Preference Graphs", 2011 3rd international comference on E-business and information system security, may, 2011, Wuhan, China, IEEE Press, Vol 2. 667-670.
[16] Y. Wang, W. Gao, Y. Zhang, and Y. Gao, "Ontology Similarity Computation Use Ranking Learning Method, Intelligent computation and industrial application", June, 2011, Hong Kong, China, Publisher: International industrial electronic center. Vol. III, pp.20-23.
[17] O. Bousquet, and A. Elisseeff, "Stability and generalization," Journal of Machine Learning Research, vol 2 , Mar.2002, pp. 499-526.
[18] S. Kutin, and P. Niyogi, "Almost-everywhere algorithmic stability and generalization error," In Proceedings of the 18th Conference on Uncertainty in Artificial Intelligence, 2002.
[19] S. Kutin, and P. Niyogi, "The interaction of stability and weakness in AdaBoost," Technical Report TR-2001-30, Computer Science Department, University of Chicago , 2001.
[20] S. Agarwal, and P. Niyogi, "Stability and generalization of bipartite ranking algorithms," In proceedings of the 18th Annual Conference on Learning Theory, 2005.
[21] S. Agarwal, and P. Niyogi, "Generalization bounds for ranking algorithms via algorithmic stability," Journal of Machine Learning Research, vol 10 , December.2009, pp. 441-474.
[22] S. Agarwal, "Generalization Bounds for Some Ordinal Regression Algorithms", ALT 2008, LNAI 5254, 2008, pp. 7-21.
[23] S. Agarwal, "Learning to rank on graphs", *Mach Learn,* vol. 81, pp. 333-357, 2010.
[24] O. Bousquet, and A. Elisseeff, "Stability and generalization", *Journal of Machine Learning Research*, vol. 2, pp. 499-526, 2002.
[25] S. Kutin, and P. Niyogi, "Almost-everywhere algorithmic stability and generalization error", In Proceedings of the 18th Conference on Uncertainty in Artificial Intelligence, 2002.
[26] X. He, W. Gao, and Z. Jia, "Generalization Bounds of Ranking via Query-Level Stability I", ICCIC 2011, Part VI, CCIS 236, pp. 188–196, 2011.
[27] Z. Jia, W. Gao, and X. He. "Generalization Bounds for Ranking Algorithm via Query-Level Stabilities Analysis", ICCIC 2011, Part VI, CCIS 236, pp. 197–203, 2011.
[28] Y. Gao, W. Gao, and Y. Zhang, "Query-Level Stability of IRSVM for Replacement Case", *Procedia Engineering*, vol, 15, pp. 2150 – 2154, 2011.
[29] Y. Gao, W. Gao, and Y. Zhang, "Query-Level Stability of Ranking SVM for Replacement Case", *Procedia Engineering*, vol.15, pp. 2145 – 2149, 2011.
[30] W. Gao, and Y. Zhang, "Generalization Bounds for Certain Class of Ranking Algorithm", *Advanced Materials Research,* Vol. 267, pp 456-461, 2011.
[31] W. Gao, Y. Zhang, and L. Liang, "Stability Analysis for Local Transductive Regression Algorithms", *Advanced Materials Research*, Vol. 267, pp 438-443, 2011.
[32] W. Gao, Y. Zhang, L. Liang, and Y. Xia, "Stability Analysis for Ranking Algorithms", Proceedings 2010 IEEE International Conference on Information Theory and Information Security, Beijing, China, Dec, 2010, pp.973-976.
[33] W. Gao, Y. Zhang, Y. Gao, L. Liang, and Youming Xia, "Strong and Weak Stability of Bipartite Ranking Algorithms", 2011 International Conference on Engineering and Information Management,April, 2011,Chengdu, China, IEEE press, pp.303-307.
[34] T. Xu, J.Gan, Y. Zhang, and W. Gao, "Generalization Bounds for Ranking Algorithms via Almost Everywhere Stability", 2011 3rd IEEE international conference on computer research and development, March, 2011, Shanghai, China, IEEE press, Vol, 04, pp.500-503.
[35] T. Xu, Y. Zhang, and W. Gao, "Generalization Bounds for Ordinal Regression Algorithms via Strong and Weak Stability", 2011 3rd International conference on Machine learning and computing, Singapore, Feb, 2011, V0l. 2, pp. 143-147.
[36] W. Gao, Y. Zhang, "Analysis for learning a similarity function with Ontology applications", under review.
[37] http: //www. geneontology. org.
[38] N. Craswelln, D. Hawking, "Overview of the TREC 2003 web track", Proceedings of the Twelfth Text Retrieval Conference. Gaithersburg, Maryland, NIST Special Publication, 2003, pp.78-92.

Yun Gao was born in Xingping, Yunnan Province, China, on Dec.10, 1964. Studied in the department of physics of Yunnan Normal University from 1980 to 1984 and got bachelor degree on physics. And then, he worked in the Journal of Yunnan Normal University as an editor and physics lecturer till now. During this years, always engaged in studying the computational physics and other related areas, such as condensed matter physics and computer science, have published more than 30 scientific papers in home and abroad.

Now, as the vice editor in chief of Yunnan Normal University and a researcher, his interests are mainly in: computational physics and computing method.

Wei Gao is a Ph.D. student at the Department of Mathematics, Soochow University, China. His research fields are graph theory, statistical learning theory, information retrieval, and artificial intelligence. He got two bachelor degrees which are computer science from Zhejiang Industrial University and mathematics education from Zhejiang Normal University. Then, he enrolled in department of computer science and information technology, Yunnan Normal University, and got a master degree there in 2009.