Partial Multi-dividing Ontology Learning Algorithm

Wei Gao[1]*, Juan Luis García Guirao[2], B. Basavanagoud[3], Jianzhang Wu[4]

1. School of Information Science and Technology, Yunnan Normal University, Kunming 650500, China
2. Departamento de Matemática Aplicaday Estadística, Universidad Politécnica de Cartagena, Hospital de Marina, 30203-Cartagena, Región de Murcia, Spain
3. Department of Mathematics, Karnatak University, Dharwad 580003, India
4. School of Computer Science and Engineer, Southeast University, Nanjing 210096, China

Abstract

As an effective data representation, storage, management, calculation and analysis model, ontology has been attracted more and more attention by scholars and been applied to various engineering disciplines. In the background of big data, the ontology is expected to increase the amount of data information, and the structure of its corresponding ontology graph has become complicated. Thus, it demands that the ontology algorithm should be more efficient than before. In the specific engineering application, the ontology algorithm is required to quickly find the semantic matching set of the concept and rank it back to the user according to their similarities. Therefore, using the learning trick to get the ontology algorithm has become a hot topic in the field of ontology research in recent years. In this paper, we present the partial multi-dividing ontology algorithm, and focus on the efficient approach for optimizing the partial multi-dividing ontology learning model. Furthermore, several theoretical results are given from a statistical learning theory perspective. Four experiments in different engineering fields are designed to show the serviceability of our partial multi-dividing algorithm from angles of ontology similarity measuring and ontology mapping building.

Keywords: ontology, similarity measuring, ontology mapping, multi-dividing setting, learning

1 Introduction

The concept of ontology originated from the study of the essential connection between things in the 1980s. And then it's introduced into the field of computer and information technology, and from the 90's it began to become the hot research filed of artificial intelligence. Because of its powerful semantic query and concept management ability, the ontology has been applied to other fields in the past 10 years. Now it has been permeated in nearly all disciplines, such as chemical science (Vijayasarathi and Sankar [1], and Banchetti-Robino [2]), pharmacology science (Baorto et al. [3] and Sarntivijai et al. [4]), biology science (Kohler et al. [5], Levine et al. [6], and Vishnu et al. [7]), psychology (Aime and Charlet [8] and Petrunia [9]), education system (Demartini et al. [10], Kruger-Ross [11], and Ochara [12]), geographic information system (GIS) (Vaccari et al. [13], Delgado et al. [14], and Tahmoorespur et al. [15]), medical science (Bertaud-Gounot et al. [16], Lousado et al. [17], and Kim et al. [18]), material science (Cuccia [19] and Ghibaudi and Cerruti [20]), and neuroscience (Bowden et al. [21] and Fumagalli [22]).

*Corresponding author. Tel.: +86 871 5516277; fax: +86 871 5516277. E-mail: gaowei@ynnu.edu.cn (W. Gao).

As a conceptual model, ontology storage and management the concept, has been widely concerned in the field of information retrieval. In light of ontology similarity calculation, we can effectively find the semantic similarity concept of the original retrieval concept, carry out the extended query in the retrieval, and return the result to the user. This trick can greatly improve the intelligence of the information retrieval. For example, if we retrieval the keyword "computer", the traditional way of search will return the computer-related information according to the degree of relevance from high to low and present them to the user. However, this retrieval is based on keyword matching, like a similar information contain word "laptop" can't be matched to "computer". But in fact the words "computer" and "laptop" share high semantics similarity. With the help of ontology for query expansion, it's found that the similarity between "laptop" and "computer" is very high. Thus, in order to find information related to the computer, we find laptop-related information, and then return back to the user according to the similarity. The advantage is that the retrieval of query information is intelligent and very comprehensive.

There are several advances in ontology semantic similarity computation. Rodriguez and Egenhofer [23] presented a method to compute semantic similarity which relaxes the demand of a single ontology. Steichen et al. [24] constructed a morphological abnormality ontology in breast pathology to assist inter-observer consensus, and it implemented position-based, content-based and mixed semantic similarity measures between concepts in this ontology. Al-Mubaid and Nguyen [25] proposed a ontology-structure-based trick for measuring semantic similarity across multiple ontologies. By means of human phenotype ontology, Kohler et al. [26] adapted semantic similarity metrics to compute phenotypic similarity between queries and hereditary diseases annotated. Batet et al. [27] studied a measure in view of the exploitation of the taxonomical structure of a biomedical ontology. Albacete et al. [28] gave proposal for computing a similarity function for each dimension of knowledge. Taha [29] presented techniques for determining the semantic relationships among GO terms. Taieb et al. [30] raised an ontology measure for quantifying the degree of the semantic similarity between concepts. Mazandu et al. [31] introduced ADaptable gene ontology semantic similarity-based Functional analysis. Lastra-Diaz et al. [32] presented a detailed companion reproducibility article of the trick and experiments proposed by former researchers.

Specifically, the framework of ontology can be expressed as a simple graph in which each concept (or element, or object) in ontology corresponds to a vertex in ontology graph and each edge on an ontology graph represents a potential link (or potential relationship) between two concepts. Let $G = (V(G), E(G))$ be a graph corresponding to the ontology O with vertex set $V(G)$ and edge set $E(G)$. In the engineering applications of ontology in various fields, the fundamental goal of ontology algorithm is to obtain the best ontology function which is applied to measure the similarities between ontology vertices in single ontology or multiple ontologies. The aim of ontology mapping is to get the high similarity vertices from different ontologies, i.e., to deduce the similarity between two or multiple ontologies, and it is used to build a bridge between different ontologies thus helps to yield a potential connection among the elements or objects from target ontologies.

At the beginning, the design of formula for ontology similarity measuring is heuristic based, i.e., the similarity formula is determined by the researchers according to the structural features of the ontology and the characteristics of the specific application domain. The shortcomings of this method are: 1) it relies on the participate of high-level field experts; 2) the similarity formula contains many man-made parameters; 3) it can't adapt to the dynamic changes in the ontology; 4) it has high complexity, and thus not suited in the specific application with big data background. In order to overcome these shortcomings, the machine learning techniques are gradually applied to the ontology algorithm. The specific idea is to get the optimal ontology function $f : V \rightarrow \mathbb{R}$ from the sample learning, which maps each vertex in ontology graph to a real number, and thus maps the whole ontology graph to the one dimension real axis (for multiple ontologies, we put all the graphs into one graph, each ontology is seen as a connected branch of the graph). Then the similarity between the ontology concepts is determined by the distance of their corresponding vertex on the real axis. It means, the similarity between vertices v_i and v_j is measured by $|f(v_i) - f(v_j)|$. The closer the distance is, the higher the similarity is, whereas the lower the distance is, the lower the similarity is. The advantage of this algorithm is that it doesn't depend on domain experts; the results are intuitive; the parameters set by man-made settings are greatly reduced; and most importantly, the computational complexity is greatly reduced because there is no pairwise similarity calculating.

There are several ontology learning algorithms and theoretical analysis results proposed in recent years. Gao et al. [33] studied the strong and weak stability of k-partite ranking based ontology algorithm. Gao and Xu [34] presented the uniform stability analysis of learning algorithms for ontology similarity computation. Gao and Zhu [35] raised gradient based ontology learning algorithm. Gao et al. [36] obtained the Ontology sparse vector learning algorithm using ADAL technology. Gao and Farahani [37] researched the generalization bounds and uniform bounds for multi-dividing ontology algorithms with convex ontology loss function. More related contexts can refer to Maire [38] and Ibrahim et al. [39].

Among these ontology learning algorithms, multi-dividing ontology algorithm is the most popular ontology learning approach in which all vertices in ontology graph or multi-ontology graph are divided into k parts (correspond to the k classes of rates). Assume that $f(v^a) > f(v^b)$ if v^a belongs to rate a and v^b belongs to rate b with $1 \leq a < b \leq k$. Gao and Farahani [37] and Wu et al. [40] presented some examples to show how multi-dividing ontology algorithm applied in the specific engineering applications, respectively. For ontology graph with tree or tree-likely structure, each kind of branch is corresponding to a rate in the dividing. Since most of ontology graphs have tree structure, multi-dividing ontology algorithm method is widely used in various of engineering filed like biology, medicine, chemistry, etc.

Although there have been several recent advances in developing algorithms for various settings of the multi-dividing ontology learning problem, the study of more available tricks and generalization properties of multi-dividing ontology learning algorithms has been largely limited to the special setting. It inspires us to explore more advanced techniques of ontology learning algorithm in multi-dividing setting and theoretical analysis from statistical learning theory.

In this paper, we present a partial multi-dividing ontology learning and study its statistical characteristics from a mathematical point of view. The rest of this paper is arranged as follows: firstly, we introduce the setting of multi-dividing ontology learning; secondly, the main algorithm is presented in Section 3; and at last, the effectiveness of proposed ontology learning algorithm is manifested by four experiments in various of engineering applications.

2 Setting

For our mathematical discussion and learning setting expression, for each vertex in the ontology graph, we use a p dimension vector to express all semantic information of its corresponding ontology concept. For convenience to express and discuss, we use v to denote the vertex v and its both corresponding vector in $\mathbb{R}^p$.

Let $V \subseteq \mathbb{R}^p$ ($p \in \mathbb{N}$) be a vertex space for ontology graph G, and the vertices in V are drawn independently and randomly according to certain unknown distribution $\mathcal{D}$. The target of ontology learning algorithms is to predict an ontology function $f : V \to \mathbb{R}$ in terms of ontology training set $S = \{v_1, \cdots, v_n\} \subseteq V$. As an ontology function, it assigns a real number to each vertex in ontology graph, and thus the similarities between two ontology concepts (assume v_i and v_j are corresponding to ontology vertices in ontology graph) are judged in view of the value of $|f(v_i) - f(v_j)|$. In the multi-dividing ontology setting, ontology vertices are divided into k parts (corresponding to k classes or k rates) and these k rates are deduced in light of the experts.

Formally, the learner is inferred to an ontology training set $S = (S^1, S^2, \cdots, S^k) \in V^{n_1} \times V^{n_2} \times \cdots \times V^{n_k}$ which consists of a sequence of ontology training samples $S_a = (v_1^a, \cdots, v_{n_a}^a) \in V^{n_a}$ ($1 \le a \le k$). By virtue of ontology sample S, a real-valued ontology function $f : V \to \mathbb{R}$ is learned which allocates the future S^a vertices larger value than S^b, where $a < b$. It means, a real-valued ontology function $f : V \to \mathbb{R}$ is predicted to assign a value to each vertex satisfies $f(v^a) > f(v^b)$ for any pair of (a, b) where $1 \le a < b \le k$. On the other hand, it can be regarded as a dimension reduction operator $f : \mathbb{R}^p \to \mathbb{R}$. In the linear ontology algorithm it aims to learn a linear ontology function $f : V \to \mathbb{R}$ denoted by $f(v) = \beta^\top v$ from a given ontology training set $S = (S^1, S^2, \cdots, S^k) \in V^{n_1} \times V^{n_2} \times \cdots \times V^{n_k}$, where $\beta \in \mathbb{R}^p$ is the linear ontology sparse vector.

Set $\mathcal{D}_a$ as the conditional distributions for each rate $1 \le a \le k$ and $n = \sum_{i=1}^k n_i$ as the total size of ontology sample set. In what follows, let $\mathbf{I}$ be the truth function satisfies $\mathbf{I}(\cdot) = 1$ if the assertion is correct and 0 otherwise. For each pair of (a, b) with $1 \le a < b \le k$, set false rates $\alpha_1^{a,b}$

and $\alpha_2^{a,b}$, where $0 \le \alpha_1^{a,b} < \alpha_2^{a,b} \le 1$. Symbols α_1 and α_2 are denoted as the abstract parameters in our setting which will be changed according to the difference of pair (a, b), i.e., $\alpha_1|_{a,b} = \alpha_1^{a,b}$ and $\alpha_2|_{a,b} = \alpha_2^{a,b}$. Sign function $\mathrm{sign}(x)$ is denoted as $\mathrm{sign}(x) = 1$ if $x > 0$, and -1 otherwise. For any $x \in \mathbb{R}$, $(x)_+ = \max\{x, 0\}$.

For an ontology function $f : V \to \mathbb{R}$ and threshold $t \in \mathbb{R}$, the true level rate of the binary classifier $\mathrm{sign}(f(x) - t)$ as the probability that it correctly classifies a random rate ontology vertices from D_a (here $1 \le a \le k$):

$$\mathrm{TR}_f^a(t) = \mathbb{P}_{v^a \sim D_a}[f(v^a) > t].$$

In the multi-dividing ontology (in short, it called MDO) setting, the model to measure the goodness of ontology function f can be formulated as

$$\mathrm{MDO}_f = \sum_{a=1}^{k-1} \sum_{b=a+1}^{k} \int_0^1 \mathrm{TR}_f^a((\mathrm{TR}_f^b)^{-1}(x))dx, \tag{2.1}$$

where $(\mathrm{TR}_f^b)^{-1}(x) = \inf\{t \in \mathbb{R} | \mathrm{TR}_f^b(t) \le x\}$. The evaluation model (2.1) equals to the area under the receiver operating characteristic curve criterion in the regress or classification setting with $k = 2$, and this is the motivation and rationality of (2.1).

Assume that there are no ties, then the multi-dividing ontology framework (2.1) can be expressed as

$$\mathrm{MDO}_f = \sum_{a=1}^{k-1} \sum_{b=a+1}^{k} \mathbb{P}_{(v^a, v^b) \sim \mathcal{D}_a \times \mathcal{D}_b}(f(v^a) > f(v^b)).$$

For each pair of (a, b) with $1 \le a < b \le k$, we are interested in the area under the curve between $\alpha_1^{a,b}$ and $\alpha_2^{a,b}$. We define the normalized partial MDO of ontology function f in the interval $[\alpha_1^{a,b}, \alpha_2^{a,b}]$ for each pair of (a, b) with $1 \le a < b \le k$ as

$$\mathrm{PMDO}_f = \sum_{a=1}^{k-1} \sum_{b=a+1}^{k} \frac{1}{\alpha_2^{a,b} - \alpha_1^{a,b}} \int_{\alpha_1^{a,b}}^{\alpha_2^{a,b}} \mathrm{TR}_f^a((\mathrm{TR}_f^b)^{-1}(x))dx.$$

Given an ontology sample set $S = (S^1, S^2, \cdots, S^k) \in V^{n_1} \times V^{n_2} \times \cdots \times V^{n_k}$, the empirical receiver operating characteristic curve corresponding to an ontology function $f : V \to \mathbb{R}$ is obtained by (here $a \in \{1, \cdots, k\}$)

$$\widehat{\mathrm{TR}}_f^a(t) = \frac{1}{n_a} \sum_{i=1}^{n_a} \mathbf{I}(f(v_i^a) > t).$$

Hence, the empirical version of MDO is given by

$$\widehat{\mathrm{MDO}}_f = \sum_{a=1}^{k-1} \sum_{b=a+1}^{k} \frac{1}{n_a n_b} \sum_{i=1}^{n_a} \sum_{j=1}^{n_b} \mathbf{I}(f(v_i^a) > f(v_j^b)). \tag{2.2}$$

For each pair of (a, b) with $1 \le a < b \le k$, by denoting $j_{\alpha_1}^{a,b} = \lfloor n_b \alpha_1^{a,b} \rfloor$ and $j_{\alpha_2}^{a,b} = \lceil n_b \alpha_2^{a,b} \rceil$, the normalized empirical partial multi-dividing ontology (PMDO) criterion of ontology function f in the interval $[\alpha_1^{a,b}, \alpha_2^{a,b}]$ can then be written as:

$$\widehat{\text{PMDO}}_f = \sum_{a=1}^{k-1}\sum_{b=a+1}^{k}\frac{1}{n_a(j_{\alpha_2}^{a,b} - j_{\alpha_1}^{a,b})}\sum_{i=1}^{n_a}\sum_{j=j_{\alpha_1}^{a,b}+1}^{j_{\alpha_2}^{a,b}}\mathbf{I}(f(v_i^a) > f(v_{(j)}^b)), \tag{2.3}$$

where $v_{(j)}^b$ denotes the ontology vertices in S^b ranked (in descending order of values) in the j-th position by ontology function f.

3 Describe of partial multi-dividing ontology algorithm

In this section, we consider ontology function $f : V \to \mathbb{R}$ denoted by $f(v) = \beta^\top v$ for some $\beta \in \mathbb{R}^p$. The contexts in this section is organized as follows: we first introduce the structural SVM based multi-dividing ontology framework with hinge ontology loss; then the partial multi-dividing ontology framework with hinge ontology loss is presented; next, we discuss the optimization methods for partial multi-dividing ontology framework based on structural SVM in interval $[0, \alpha_2^{a,b}]$ and $[0, \alpha_2^{a,b}]$, respectively in the third and fourth parts; finally, we compte the generalization bound for our ontology algorithm.

3.1 Multi-dividing ontology framework with structural SVM or hinge ontology loss

Given an ontology training set $S = (S^1, S^2, \cdots, S^k) \in V^{n_1} \times V^{n_2} \times \cdots \times V^{n_k}$, the aim here is to search an ontology function that gets maximum empirical MDO on ontology sample S, or equivalently minimizes the empirical multi-dividing ontology risk described as follows

$$\widehat{R}_{\text{MDO}}(\beta, S) = \sum_{a=1}^{k-1}\sum_{b=a+1}^{k}\frac{1}{n_a n_b}\sum_{i=1}^{n_a}\sum_{j=1}^{n_b}\mathbf{I}(\beta^\top v_i^a \le \beta^\top v_j^b). \tag{3.4}$$

In most situations, ontology optimization model (3.4) is difficult to determine its extreme value since the truth function $\mathbf{I}$ is non-differentiable. One trick to solve this problem is using hinge loss function to replace the truth function. For any ontology function f and vertices v^a and v^b from difficult rates, the pair-wise hinge ontology loss is defined as $(1 - (\beta^\top v^a - \beta^\top v^b))_+$ which is convex in β and an upper bound on $\mathbf{I}(\beta^\top v^a \le \beta^\top v^b)$. Thus, ontology optimization framework with pair-wise hinge ontology loss is presented as (which is convex in β)

$$\widehat{R}_{\text{MDO}}^{\text{hinge}}(\beta, S) = \sum_{a=1}^{k-1}\sum_{b=a+1}^{k}\frac{1}{n_a n_b}\sum_{i=1}^{n_a}\sum_{j=1}^{n_b}(1 - (\beta^\top v_i^a - \beta^\top v_j^b))_+. \tag{3.5}$$

For each pair of (a, b) with $1 \le a < b \le k$ and any ordering of the ontology training vertices, we represent the related ordering of the n_a ontology vertices in S^a and n_b ontology vertices in S^b by means of a matrix $\pi^{a,b} = [\pi^{a,b}]_{ij} \in \{0,1\}^{n_a \times n_b}$ as follows: $[\pi^{a,b}]_{ij} = 1$ if $f(v_i^a) \le f(v_j^b)$; $[\pi^{a,b}]_{ij} = 0$ if $f(v_i^a) > f(v_j^b)$. Note that for each pair of (a, b) with $1 \le a < b \le k$, not all $2^{n_a n_b}$ matrices in

$\{0,1\}^{n_a \times n_b}$ represent a valid relative ordering. Hence, for each pair of (a,b) with $1 \le a < b \le k$, set $\Pi^{a,b}_{n_a,n_b}$ as the set of all matrices in $\{0,1\}^{n_a \times n_b}$ that can express valid orderings, and obviously, for any $i \in \{1, \cdots, n_a\}$ and $j \in \{1, \cdots, n_b\}$ the correct relative ordering π^* has $(\pi^{a,b}_{ij})^* = 0$. In what follows, we always assume π as the abstract matrix in our setting, i.e., $\pi|_{a,b} = \pi^{a,b}$; similarly, Π is the matrix space in our setting with $\Pi|_{a,b} = \Pi^{a,b}_{n_a,n_b}$. We can regard π and Π as the symbols which are well defined in the various pair of (a,b).

Define the multi-dividing ontology loss of π with respect to π^* as

$$\Delta_{\mathrm{MDO}}(\pi^*, \pi) = \sum_{a=1}^{k-1} \sum_{b=a+1}^{k} \frac{1}{n_a n_b} \sum_{i=1}^{n_a} \sum_{j=1}^{n_b} \pi^{a,b}_{i,j}. \tag{3.6}$$

Clearly, for any π consistent with ontology function $f = \beta^\top v$, $\Delta_{\mathrm{MDO}}(\pi^*, \pi)$ evaluates to the multi-dividing ontology risk $\widehat{R}_{\mathrm{MDO}}(\beta, S)$ in ontology framework (3.4). Set a joint feature map between the input ontology training set and an output ordering matrix $\phi : (V^{n_1} \times \cdots \times V^{n_k}) \times \prod_{1 \le a < b \le k} \Pi^{a,b}_{n_a,n_b} \to \mathbb{R}^p$ as

$$\phi(S, \pi) = \sum_{a=1}^{k-1} \sum_{b=a+1}^{k} \frac{1}{n_a n_b} \sum_{i=1}^{n_a} \sum_{j=1}^{n_b} (1 - \pi^{a,b}_{ij})(v^a_i - v^b_j). \tag{3.7}$$

Moreover, for each pair of (a,b) with $1 \le a < b \le k$, we set $\phi^{a,b} : (V^{n_a} \times V^{n_b}) \times \Pi^{a,b}_{n_a,n_b} \to \mathbb{R}^p$ as

$$\phi^{a,b}(S^a \cup S^b, \pi^{a,b}) = \frac{1}{n_a n_b} \sum_{i=1}^{n_a} \sum_{j=1}^{n_b} (1 - \pi^{a,b}_{ij})(v^a_i - v^b_j).$$

The good selection of $\phi(S,\pi)$ satisfy that for any fixed $\beta \in \mathbb{R}$, maximizing $\beta^\top \phi(S,\pi)$ over π gets an ordering matrix consistent with the ontology function $f = \beta^\top v$, and thus the loss part evaluates to $\widehat{R}_{\mathrm{MDO}}(\beta, S)$. Now, the ontology problem of optimizing the MDO becomes to searching a $\beta \in \mathbb{R}^p$ in which the maximizer over π of $\beta^\top \phi(S,\pi)$ has the smallest multi-dividing ontology loss. We now express our ontology problem in light of the following structural SVM based version:

$$\widehat{R}^{\mathrm{struct}}_{\mathrm{MDO}}(\beta, S) = \max_{\pi} \{\Delta_{\mathrm{MDO}}(\pi^*, \pi) - (\beta^\top \phi(S, \pi^*) - \beta^\top \phi(S, \pi))\}. \tag{3.8}$$

Obviously, $\widehat{R}^{\mathrm{struct}}_{\mathrm{MDO}}(\beta, S)$ is convex in β. Furthermore, if let $\overline{\pi}$ be the maximizer of $\beta^\top \phi(S,\pi)$ over Π, we infer $\widehat{R}^{\mathrm{struct}}_{\mathrm{MDO}}(\beta, S) \ge \Delta_{\mathrm{MDO}}(\pi^*, \overline{\pi}) - (\beta^\top \phi(S, \pi^*) - \beta^\top \phi(S, \overline{\pi})) \ge \Delta_{\mathrm{MDO}}(\pi^*, \overline{\pi}) = \widehat{R}_{\mathrm{MDO}}(\beta, S)$.

Our first theoretical result stated below shows the equivalent of ontology optimization model by pair-wise hinge loss and structure SVM.

Theorem 1 *For any $\beta \in \mathbb{R}^p$ and training sample set $S = (S^1, S^2, \cdots, S^k) \in V^{n_1} \times V^{n_2} \times \cdots \times V^{n_k}$, we have $\widehat{R}^{\mathrm{struct}}_{\mathrm{MDO}}(\beta, S) = \widehat{R}^{\mathrm{hinge}}_{\mathrm{MDO}}(\beta, S)$.*

Proof of Theorem 1. The structural SVM based ontology loss can be simplified into a pair-wise form:

$$\widehat{R}^{\mathrm{struct}}_{\mathrm{MDO}}(\beta, S) = \max_{\pi} \{\Delta_{\mathrm{MDO}}(\pi^*, \pi) - (\beta^\top \phi(S, \pi^*) - \beta^\top \phi(S, \pi))\}$$

$$
\begin{aligned}
&= \max_{\pi}\{\sum_{a=1}^{k-1}\sum_{b=a+1}^{k}\frac{1}{n_a n_b}\sum_{i=1}^{n_a}\sum_{j=1}^{n_b}\pi_{ij}^{a,b}(1-\beta^{\top}(v_i^a-v_j^b))\}\\
&= \sum_{a=1}^{k-1}\sum_{b=a+1}^{k}\max_{\pi^{a,b}\in\{0,1\}^{n_a,n_b}}\{\frac{1}{n_a n_b}\sum_{i=1}^{n_a}\sum_{j=1}^{n_b}\pi_{ij}^{a,b}(1-\beta^{\top}(v_i^a-v_j^b))\}\\
&= \sum_{a=1}^{k-1}\sum_{b=a+1}^{k}\frac{1}{n_a n_b}\sum_{i=1}^{n_a}\sum_{j=1}^{n_b}\max_{\pi_{ij}^{a,b}\in\{0,1\}}\{\pi_{ij}^{a,b}(1-\beta^{\top}(v_i^a-v_j^b))\}\\
&= \sum_{a=1}^{k-1}\sum_{b=a+1}^{k}\frac{1}{n_a n_b}\sum_{i=1}^{n_a}\sum_{j=1}^{n_b}\overline{\pi}_{ij}^{a,b}(1-\beta^{\top}(v_i^a-v_j^b)),
\end{aligned}
$$

where $\overline{\pi}_{ij}^{a,b}=\mathbf{I}(\beta^{\top}(v_i^a-v_j^b)\le 1)$. Therefore, we deduce

$$
\begin{aligned}
\widehat{R}_{\mathrm{MDO}}^{\mathrm{struct}}(\beta,S) &= \sum_{a=1}^{k-1}\sum_{b=a+1}^{k}\frac{1}{n_a n_b}\sum_{i=1}^{n_a}\sum_{j=1}^{n_b}\overline{\pi}_{ij}^{a,b}(1-\beta^{\top}(v_i^a-v_j^b))\\
&= \sum_{a=1}^{k-1}\sum_{b=a+1}^{k}\frac{1}{n_a n_b}\sum_{i=1}^{n_a}\sum_{j=1}^{n_b}(1-\beta^{\top}(v_i^a-v_j^b))_{+}\\
&= \widehat{R}_{\mathrm{MDO}}^{\mathrm{hinge}}(\beta,S).
\end{aligned}
$$

Hence, we get the desired result. □

3.2 Partial multi-dividing ontology framework with hinge ontology loss

For a given ontology training set $S=(S^1,S^2,\cdots,S^k)\in V^{n_1}\times V^{n_2}\times\cdots\times V^{n_k}$, we aim to find an ontology function $f(v)=\beta^{\top}v$ that maximizes partial multi-dividing ontology risk in $[\alpha_1^{a,b},\alpha_2^{a,b}]$ (for each pair of (a,b) with $1\le a<b\le k$):

$$
\widehat{R}_{\mathrm{PMDO}(\alpha_1,\alpha_2)}(\beta,S)=\sum_{a=1}^{k-1}\sum_{b=a+1}^{k}\frac{1}{n_a(j_{\alpha_2^{a,b}}-j_{\alpha_1^{a,b}})}\sum_{i=1}^{n_a}\sum_{j=\alpha_1^{a,b}+1}^{\alpha_2^{a,b}}\mathbf{I}(\beta^{\top}v_i^a\le\beta^{\top}v_{(j)_\beta}^b), \qquad (3.9)
$$

where $v_{(j)_\beta}^b$ denotes the ontology vertices in S^b ranked (in descending order of values) in the j-th position by $f(v^b)=\beta^{\top}v^b$. Using hinge ontology loss function, the above loss (3.9) becomes

$$
\widehat{R}_{\mathrm{PMDO}(\alpha_1,\alpha_2)}^{\mathrm{hinge}}(\beta,S)=\sum_{a=1}^{k-1}\sum_{b=a+1}^{k}\frac{1}{n_a(j_{\alpha_2^{a,b}}-j_{\alpha_1^{a,b}})}\sum_{i=1}^{n_a}\sum_{j=\alpha_1^{a,b}+1}^{\alpha_2^{a,b}}(1-(\beta^{\top}v_i^a-\beta^{\top}v_{(j)_\beta}^b))_{+}. \qquad (3.10)
$$

When $\alpha_1^{a,b}=0$ and $j_{\alpha_1^{a,b}}=\lfloor n_b\alpha_1^{a,b}\rfloor=0$, the ontology risk (3.10) is manifested as

$$
\widehat{R}_{\mathrm{PMDO}(0,\alpha_2)}^{\mathrm{hinge}}(\beta,S)=\sum_{a=1}^{k-1}\sum_{b=a+1}^{k}\frac{1}{n_a j_{\alpha_2^{a,b}}}\sum_{i=1}^{n_a}\sum_{j=1}^{\alpha_2^{a,b}}(1-(\beta^{\top}v_i^a-\beta^{\top}v_{(j)_\beta}^b))_{+}. \qquad (3.11)
$$

Our next theoretical result stated below shows that the multi-dividing ontology loss depicted in (3.11) is convex with respect to β.

Theorem 2 *Let $\alpha_2^{a,b}>0$ for each pair of (a,b) with $1\le a<b\le k$. For any ontology training sample set $S=(S^1,S^2,\cdots,S^k)\in V^{n_1}\times V^{n_2}\times\cdots\times V^{n_k}$, the $\widehat{R}_{\mathrm{PMDO}(0,\alpha_2)}^{\mathrm{hinge}}(\beta,S)$ is convex in β.*

Proof of Theorem 2. Fix ontology sample $S=(S^1,S^2,\cdots,S^k)\in V^{n_1}\times V^{n_2}\times\cdots\times V^{n_k}$. Let $\beta_1,\beta_2\in\mathbb{R}^p$, $\lambda\in(0,1)$, and $\tilde{\beta}=\lambda\beta_1+(1-\lambda)\beta_2$. We aim to show that

$$\widehat{R}^{\text{hinge}}_{\text{PMDO}(0,\alpha_2)}(\tilde{\beta},S)\leq\lambda\widehat{R}^{\text{hinge}}_{\text{PMDO}(0,\alpha_2)}(\beta_1,S)+(1-\lambda)\widehat{R}^{\text{hinge}}_{\text{PMDO}(0,\alpha_2)}(\beta_2,S).$$

For each pair of (a,b) with $1\leq a<b\leq k$ and any vertex v^b ($b\in\{2,\cdots,k\}$), define $\Psi^{a,b}(\beta,S^a,v^b)=\frac{1}{n_a}\sum_{i=1}^{n_a}\mathbf{I}(\beta^\top v_i^a\leq\beta^\top v^b)$. In view of $\Psi^{a,b}(\beta,S^a,v^b)$ is convex in β and is monotonically increasing in the value $\beta^\top v^b$ assigned to v^b, we yield

$$\begin{aligned}
&\widehat{R}^{\text{hinge}}_{\text{PMDO}(0,\alpha_2)}(\tilde{\beta},S)=\sum_{a=1}^{k-1}\sum_{b=a+1}^{k}\frac{1}{j^{a,b}_{\alpha_2}}\sum_{j=1}^{j^{a,b}_{\alpha_2}}\Psi^{a,b}(\tilde{\beta},S^a,v^b_{(j)_{\tilde{\beta}}})\\
\leq&\sum_{a=1}^{k-1}\sum_{b=a+1}^{k}\{\lambda\frac{1}{j^{a,b}_{\alpha_2}}\sum_{j=1}^{j^{a,b}_{\alpha_2}}\Psi^{a,b}(\beta_1,S^a,v^b_{(j)_{\tilde{\beta}}})+(1-\lambda)\frac{1}{j^{a,b}_{\alpha_2}}\sum_{j=1}^{j^{a,b}_{\alpha_2}}\Psi^{a,b}(\beta_2,S^a,v^b_{(j)_{\tilde{\beta}}})\}\\
\leq&\sum_{a=1}^{k-1}\sum_{b=a+1}^{k}\{\lambda\frac{1}{j^{a,b}_{\alpha_2}}\sum_{j=1}^{j^{a,b}_{\alpha_2}}\Psi^{a,b}(\beta_1,S^a,v^b_{(j)_{\beta_1}})+(1-\lambda)\frac{1}{j^{a,b}_{\alpha_2}}\sum_{j=1}^{j^{a,b}_{\alpha_2}}\Psi^{a,b}(\beta_2,S^a,v^b_{(j)_{\beta_2}})\}\\
=&\lambda\widehat{R}^{\text{hinge}}_{\text{PMDO}(0,\alpha_2)}(\beta_1,S)+(1-\lambda)\widehat{R}^{\text{hinge}}_{\text{PMDO}(0,\alpha_2)}(\beta_2,S).
\end{aligned}$$

This completes the proof. □

It's still hard to solve the ontology optimization problem directly, although the multi-dividing ontology framework with hinge ontology loss and intervals $[0,\alpha_2^{a,b}]$ for each pair of (a,b) with $1\leq a<b\leq k$ being convex.

As a supplement, if we replace the loss term to the structural SVM for the multi-dividing ontology framework in (3.8) with $[\alpha_1^{a,b},\alpha_2^{a,b}]$ for each pair of (a,b) with $1\leq a<b\leq k$, then we obtain partial multi-dividing ontology framework as follows

$$\Delta_{\text{PMDO}}(\pi^*,\pi)=\sum_{a=1}^{k-1}\sum_{b=a+1}^{k}\frac{1}{n_a(j_{\alpha_2^{a,b}}-j_{\alpha_1^{a,b}})}\sum_{i=1}^{n_a}\sum_{j=1}^{n_b}\pi^{a,b}_{i,(j)_\pi},$$

where $(j)_\pi$ denotes the index of the j-th ranked ontology vertices in S^b consistent with π. Hence, the resulting ontology risk can be formulated by:

$$\widehat{R}^{\text{struct}}_{\text{PMDO}}(\beta,S)=\max_{\pi}\{\Delta_{\text{PMDO}}(\pi^*,\pi)-(\beta^\top\phi(S,\pi^*)-\beta^\top\phi(S,\pi))\}.\tag{3.12}$$

3.3 Partial multi-dividing ontology framework based on structural SVM trick in intervals $[0,\alpha_2^{a,b}]$

In this part, we consider the ontology approach for optimizing the partial multi-dividing ontology framework in intervals $[0,\alpha_2^{a,b}]$ (pair (a,b) satisfy $1\leq a<b\leq k$):

$$\widehat{R}_{\text{PMDO}(0,\alpha_2)}(\beta,S)=\sum_{a=1}^{k-1}\sum_{b=a+1}^{k}\frac{1}{n_a j_{\alpha_2^{a,b}}}\sum_{i=1}^{n_a}\sum_{j=1}^{\alpha_2^{a,b}}\mathbf{I}(\beta^\top v_i^a\leq\beta^\top v^b_{(j)_\beta}).\tag{3.13}$$

We use the techniques of structure SVM here and also provide an algorithm to estate the positions related to value of $\alpha_2^{a,b}$.

For each pair of (a,b) with $1 \le a < b \le k$ and any subset of negatives $\Xi^{a,b} \subseteq S^b$, let $\widehat{R}^{a,b}_{\rm MDO}(\beta, S^a, \Xi^{a,b})$ be the multi-dividing ontology risk of ontology function $f = \beta^\top v$ evaluated on an ontology sample containing all the S^a and the subset of Ξ. Then we obtain the following equivalent result which implies that S^b can reduce to a subset of size j_{α_2} for for each pair of (a,b) with $1 \le a < b \le k$.

Theorem 3 *For any $\beta \in \mathbb{R}^p$ and ontology training sample set $S = (S^1, S^2, \cdots, S^k) \in V^{n_1} \times V^{n_2} \times \cdots \times V^{n_k}$, we infer*

$$\begin{aligned}\widehat{R}_{{\rm PMDO}(0,\alpha_2)}(\beta, S) &= \sum_{a=1}^{k-1}\sum_{b=a+1}^{k} \max_{\Xi^{a,b}\subseteq S^b, |\Xi^{a,b}|=j^{a,b}_{\alpha_2}} \frac{1}{n_a j_{\alpha_2^{a,b}}} \sum_{i=1}^{n_a}\sum_{j=1}^{v^b\in\Xi^{a,b}} \mathbf{I}(\beta^\top v_i^a \le \beta^\top v^b) \\ &= \sum_{a=1}^{k-1}\sum_{b=a+1}^{k} \max_{\Xi^{a,b}\subseteq S^b, |\Xi^{a,b}|=j^{a,b}_{\alpha_2}} \widehat{R}^{a,b}_{\rm MDO}(\beta, S^a, \Xi^{a,b}). \qquad (3.14)\end{aligned}$$

Proof of Theorem 3. For each pair of (a,b) with $1 \le a < b \le k$ and any v^b, set $\Psi^{a,b}(\beta, S^a, v^b) = \frac{1}{n_a}\sum_{i=1}^{n_a} \mathbf{I}(\beta^\top v_i^a \le \beta^\top v^b)$. Notice that $\Psi^{a,b}(\beta, S^a, v^b)$ is monotonically increasing in the value $\beta^\top v^b$ corresponds to v^b. Thus $\widehat{R}^{a,b}_{\rm MDO}(\beta, S^a, \Xi^{a,b}) = \frac{1}{j^{a,b}_{\alpha_2}}\sum_{v^b\in\Xi^{a,b}} \Psi^{a,b}(\beta, S^a, v^b)$ takes the largest value if $\Xi^{a,b}$ contains the top valued $j^{a,b}_{\alpha_2}$ vertices in S^b, and by definition (3.13), this largest value is equal to the partial multi-dividing ontology risk of the fixed ontology function in the interval $[0, \alpha_2^{a,b}]$ for each pair of (a,b) with $1 \le a < b \le k$. □

For each pair of (a,b) with $1 \le a < b \le k$, S^a and any subset $\Xi^{a,b} = \{\xi_1^{a,b}, \cdots, \xi^{a,b}_{j_{\alpha_2^{a,b}}}\} \subseteq S^b$, let $\pi^{a,b} \in \{0,1\}^{n_a \times j_{\alpha_2^{a,b}}}$ be the truncated ordering matrices defined as $\pi^{a,b}_{ij} = 1$ if $f(v_i^a) \le f(\xi_j^{a,b})$ and $\pi^{a,b}_{ij} = 0$ if $f(v_i^a) > f(\xi_j^{a,b})$, where $i \in \{1, \cdots, n_a\}$ and $j \in \{1, \cdots, j_{\alpha_2^{a,b}}\}$.

Let $\Pi^{a,b}$ be the collection of all valid orderings for each pair of (a,b) with $1 \le a < b \le k$, and the correct ordering is stated as $(\pi^{a,b})^* = \mathbf{0}^{n_a \times j_{\alpha_2^{a,b}}}$. The joint feature map for n_a vertices from S^a and $j_{\alpha_2^{a,b}}$ vertices from S^b is redefined as $\phi^{a,b} : V^{n_a} \times V^{j_{\alpha_2^{a,b}}} \to \mathbb{R}^p$. In this way, the convex upper bound on the inner multi-dividing ontology part ($\widehat{R}^{a,b}_{\rm MDO}(\beta, S^a, \Xi^{a,b})$ for each pair of (a,b) with $1 \le a < b \le k$) in (3.14) can be expressed as

$$\max_{\pi^{a,b}\in\Pi_{n_a, j_{\alpha_2^{a,b}}}} \{\Delta_{\rm MDO}((\pi^{a,b})^*, \pi^{a,b}) - \beta^\top(\phi((S^a, \Xi^{a,b}), (\pi^{a,b})^*) - \phi((S^a, \Xi^{a,b}), \pi^{a,b}))\}.$$

By replacing the $\widehat{R}^{a,b}_{\rm MDO}(\beta, S^a, \Xi^{a,b})$ in (3.14) with the above expression, we infer

$$\begin{aligned}\widehat{R}^{\rm tight}_{{\rm PMDO}(0,\alpha_2)}(\beta, S) &= \sum_{a=1}^{k-1}\sum_{b=a+1}^{k} \max_{\Xi\subseteq S^b, |\Xi^{a,b}|=j_{\alpha_2^{a,b}}} \max_{\pi^{a,b}\in\Pi_{n_a, j_{\alpha_2^{a,b}}}} \{\Delta_{\rm MDO}((\pi^{a,b})^*, \pi^{a,b}) \qquad (3.15) \\ &\quad -\beta^\top(\phi((S^a, \Xi^{a,b}), (\pi^{a,b})^*) - \phi((S^a, \Xi^{a,b}), \pi^{a,b}))\}.\end{aligned}$$

where the $\pi_{ij}^{a,b}$'s indices over all vertices in S^a, and over vertices in the corresponding subset $\Xi^{a,b}$ in the outer argmax part.

By means of hinge ontology loss, the multi-dividing ontology risk can be simplified as

$$\widehat{R}^{\text{tight}}_{\text{PMDO}(0,\alpha_2^{a,b})}(\beta,S)=\sum_{a=1}^{k-1}\sum_{b=a+1}^{k}\max_{\Xi\subseteq S^b,|\Xi^{a,b}|=j_{\alpha_2^{a,b}}}\frac{1}{n_a j_{\alpha_2^{a,b}}}\sum_{i=1}^{n_a}\sum_{v^b\in\Xi^{a,b}}(1-(\beta^\top v_i^a-\beta^\top v^b))_+. \quad (3.16)$$

For each pair of (a,b) with $1\le a<b\le k$, let $\overline{\Xi}^{a,b}=\{\overline{\xi}_1^{a,b},\cdots,\overline{\xi}_{j_{\alpha_2^{a,b}}}^{a,b}\}$ be the set of vertices in S^b ranked in the top $j_{\alpha_2^{a,b}}$ positions (in descending order of values) by $f=\beta^\top v$. Hence, the largest objective value of each pair of (a,b) in (3.16) is obtained at $\overline{\Xi}^{a,b}$.

Define $\Psi^{a,b}(\beta,S^a,v^b)=\frac{1}{n_a}\sum_{i=1}^{n_a}(1-(\beta^\top v_i^a-\beta^\top v^b))_+$ for each pair of (a,b) with $1\le a<b\le k$ and any v^b. For any $\beta\in\mathbb{R}^p$ and ontology training sample set $S=(S^1,S^2,\cdots,S^k)\in V^{n_1}\times V^{n_2}\times\cdots\times V^{n_k}$, we yield $\widehat{R}^{\text{hinge}}_{\text{PMDO}(0,\alpha_2^{a,b})}(\beta,S)=\widehat{R}^{\text{tight}}_{\text{PMDO}(0,\alpha_2^{a,b})}(\beta,S)$.

Now, we present the algorithm to determine the tight set $\overline{\Xi}^{a,b}$ and truncated ordering matrices $\overline{\pi}^{a,b}$. The specific ontology optimization problem can be written as:

$$\sum_{a=1}^{k-1}\sum_{b=a+1}^{k}\max_{\Xi\subseteq S^b,|\Xi^{a,b}|=j_{\alpha_2^{a,b}},\pi\in\Pi_{n_a,j_{\alpha_2^{a,b}}}}\{\Delta_{\text{MDO}}((\pi^{a,b})^*,\pi^{a,b})-\beta^\top(\phi((S^a,\Xi^{a,b}),(\pi^{a,b})^*)-\phi((S^a,\Xi^{a,b}),\pi^{a,b}))\}.$$

From what we have discussed above, the above inner argmax is attained at the top $j_{\alpha_2}^{a,b}$ vertices $\overline{\Xi}^{a,b}=\{\overline{\xi}_1^{a,b},\cdots,\overline{\xi}_{j_{\alpha_2^{a,b}}}^{a,b}\}\in S^b$ for each pair of (a,b) with $1\le a<b\le k$ in terms of β. Hence, the rest task is to determine the optimal ordering matrix in $\Pi^{n_a,j_{\alpha_2^{a,b}}}$ with given $\overline{\Xi}^{a,b}$, and the ontology optimization problem thus be decomposed effectively. Specifically, the ontology optimization problem with given subset $\overline{\Xi}^{a,b}$ for each pair of (a,b) with $1\le a<b\le k$ can be represented as

$$\text{argmin}_\pi\sum_{a=1}^{k-1}\sum_{b=a+1}^{k}\frac{1}{n_a j_{\alpha_2^{a,b}}}\sum_{i=1}^{n_a}\sum_{j=1}^{\alpha_2^{a,b}}\pi_{ij}^{a,b}(1-\beta^\top(v_i^a-\overline{\xi}_j^{a,b})). \quad (3.17)$$

We consider solving a relaxed form of ontology optimization problem (3.17) over all matrices in $\{0,1\}^{n_a\times j_{\alpha_2^{a,b}}}$, and the optimal ordering matrix is obtained by $\overline{\pi}_{ij}^{a,b}=\mathbf{I}(\beta^\top v_i^a-\beta^\top\overline{\xi}_j^{a,b}\le 1)$ for each pair of (a,b) with $1\le a<b\le k$. Furthermore, $\overline{\pi}$ is also a solution to the original unrelaxed ontology optimization problem (3.17) for given $\overline{\Xi}$, and then $(\overline{\Xi},\overline{\pi})$ provides us the desired conclusion.

Algorithm 1. Determine $(\overline{\Xi},\overline{\pi})$ for partial multi-dividing ontology problem in interval $[0,\alpha_2^{a,b}]$ for each pair:

Step 1: Inputs: $S=(S^1,S^2,\cdots,S^k)\in V^{n_1}\times V^{n_2}\times\cdots\times V^{n_k}$, $\alpha_2^{a,b}$ for each pair of (a,b) with $1\le a<b\le k$, β

Step 2: Set $\overline{\Xi}^{a,b}=\{\overline{\xi}_1^{a,b},\cdots,\overline{\xi}_{j_{\alpha_2^{a,b}}}^{a,b}\}$ as the collection of vertices in S^b (for each pair (a,b)) ranked in the top $j_{\alpha_2^{a,b}}$ positions (in descending order of values) by $f=\beta^\top v$

Step 3: $\overline{\pi}_{ij}^{a,b}=\mathbf{I}(\beta^\top v_i^a-\beta^\top\overline{\xi}_j^{a,b}\le 1)$ where $i\in\{1,\cdots,n_a\}$, $j\in\{1,\cdots,j_{\alpha_2^{a,b}}\}$

Step 4: Output: $(\overline{\Xi}^{a,b},\overline{\pi}^{a,b})$

Clearly, computational time for each pair of (a,b) with $1\le a<b\le k$ is $O(n_a j_{\alpha_2^{a,b}}+n_b\log(n_b))$. Thus, the complexity of Algorithm 1 is $O(\sum_{a=1}^{k-1}\sum_{b=a+1}^{k}(n_a j_{\alpha_2^{a,b}}+n_b\log(n_b)))$.

3.4 Partial multi-dividing ontology framework based on structural SVM trick in intervals $[\alpha_1^{a,b}, \alpha_2^{a,b}]$

Recall that the structural SVM based partial multi-dividing ontology risk in $[\alpha_1^{a,b}, \alpha_2^{a,b}]$ is given by:

$$\widehat{R}_{\mathrm{PMDO}(\alpha_1,\alpha_2)}(\beta, S) = \sum_{a=1}^{k-1}\sum_{b=a+1}^{k}\frac{1}{n_a(j_{\alpha_2^{a,b}} - j_{\alpha_2^{a,b}})}\sum_{i=1}^{n_a}\sum_{j=\alpha_1^{a,b}+1}^{\alpha_2^{a,b}}\mathbf{I}(\beta^\top v_i^a \le \beta^\top v_{(j)_\beta}^b). \quad (3.18)$$

Similarly as Theorem 3, the partial multi-dividing ontology risk in interval $[\alpha_1^{a,b}, \alpha_2^{a,b}]$ (for each pair of (a, b) with $1 \le a < b \le k$) can be expressed as a maximum of a certain term over subsets of S^b with $j_{\alpha_2^{a,b}}$ vertices.

Theorem 4 *For any $\beta \in \mathbb{R}^p$ and ontology training sample set $S = (S^1, S^2, \cdots, S^k) \in V^{n_1} \times V^{n_2} \times \cdots \times V^{n_k}$,*

$$\widehat{R}_{\mathrm{PMDO}(\alpha_1,\alpha_2)}(\beta, S) = \sum_{a=1}^{k-1}\sum_{b=a+1}^{k}\max_{\Xi^{a,b}\subseteq S^b, |\Xi^{a,b}|=j_{\alpha_2^{a,b}}}\tilde{R}(\beta, S^a, \Xi^{a,b}),$$

where for any $\Xi^{a,b} = \{\xi_1^{a,b}, \cdots, \xi_{j_{\alpha_2^{a,b}}}^{a,b}\} \subseteq S^b$ satisfy $\beta^\top\xi_1^{a,b} \ge \cdots \ge \beta^\top\xi_{j_{\alpha_2^{a,b}}}^{a,b}$,

$$\tilde{R}(\beta, S^a, \Xi^{a,b}) = \frac{1}{n_a(j_{\alpha_2^{a,b}} - j_{\alpha_2^{a,b}})}\sum_{i=1}^{n_a}\sum_{j=\alpha_1^{a,b}+1}^{\alpha_2^{a,b}}\mathbf{I}(\beta^\top v_i^a \le \beta^\top\xi_j^{a,b}).$$

Proof of Theorem 4. Similar to depicted in Theorem 3, for each pair of (a, b) with $1 \le a < b \le k$ and any v^b, define $\Psi^{a,b}(\beta, S^a, v^b) = \frac{1}{n_a}\sum_{i=1}^{n_a}\mathbf{I}(\beta^\top v_i^a \le \beta^\top v^b)$. Then $\tilde{R}(\beta, S^a, \Xi^{a,b})$ evaluates to the average value of such quantity at the bottom ranked $j_{\alpha_2^{a,b}} - j_{\alpha_1^{a,b}}$ vertices in $\Xi^{a,b}$ by β, and $\Psi^{a,b}(\beta, S^a, v^b)$ is monotonically increasing in the value of $\beta^\top v^b$. Therefore, $\tilde{R}(\beta, S^a, \Xi^{a,b})$ takes the largest value if $\Xi^{a,b}$ includes vertices in the top $j_{\alpha_2^{a,b}}$ positions in S^b which ranked by β. In terms of (3.18), this largest value is equal to the partial multi-dividing ontology risk of the ontology function in $[\alpha_1^{a,b}, \alpha_2^{a,b}]$ for each pair of (a, b) with $1 \le a < b \le k$. □

Next, we discuss the convex upper bound on $\tilde{R}$, and thus get a convex optimization model in the partial multi-dividing setting with interval $[\alpha_1^{a,b}, \alpha_2^{a,b}]$ for each pair of (a, b) with $1 \le a < b \le k$. Set the loss of the truncated ordering matrices as follows:

$$\Delta_{\mathrm{PMDO}}^{\mathrm{tr}}(\pi^*, \pi) = \sum_{a=1}^{k-1}\sum_{b=a+1}^{k}\frac{1}{n_a(j_{\alpha_2^{a,b}} - j_{\alpha_2^{a,b}})}\sum_{i=1}^{n_a}\sum_{j=\alpha_1^{a,b}+1}^{\alpha_2^{a,b}}\pi_{i,(j)_{\pi^{a,b}}}^{a,b}. \quad (3.19)$$

Then, we deduce the convex upper bound on $\tilde{R}$:

$$\sum_{a=1}^{k-1}\sum_{b=a+1}^{k}\max_{\pi\in\Pi_{n_a, j_{\alpha_2^{a,b}}}}\{\Delta_{\mathrm{PMDO}}^{\mathrm{tr}}((\pi^{a,b})^*, \pi^{a,b}) - \beta^\top(\phi((S^a, \Xi^{a,b}), (\pi^{a,b})^*) - \phi((S^a, \Xi^{a,b}), \pi^{a,b}))\}.$$

By replacing $\tilde{R}$ in the partial multi-dividing ontology risk part in Theorem 4 with (3.19), we infer the following upper bounding multi-dividing ontology risk:

$$\begin{aligned}\widehat{R}^{\text{tight}}_{\text{PMDO}(\alpha_1,\alpha_2)}(\beta,S) \quad = \quad & \sum_{a=1}^{k-1}\sum_{b=a+1}^{k}\max_{\Xi^{a,b}\subseteq S^b,|\Xi^{a,b}|=j_{\alpha_2^{a,b}}}\max_{\pi\in\Pi_{n_a,j_{\alpha_2^{a,b}}}}\{\Delta^{\text{tr}}_{\text{PMDO}}((\pi^{a,b})^*,\pi^{a,b}) \\ & -\beta^\top(\phi((S^a,\Xi^{a,b}),(\pi^{a,b})^*)-\phi((S^a,\Xi^{a,b}),\pi^{a,b}))\}. \qquad (3.20)\end{aligned}$$

The risk (3.20) is the largest value of convex functions in β, and thus is convex in β. It reveals that (3.20) is attained by the top $j_{\alpha_2^{a,b}}$ vertices in S^b according to β for each pair of (a,b) with $1\le a<b\le k$.

For any $\Xi^{a,b}=\{\xi_1^{a,b},\cdots,\xi^{a,b}_{j_{\alpha_2^{a,b}}}\}\subseteq S^b$, we assume that $\beta^\top\xi_1^{a,b}\ge\cdots\beta^\top\xi^{a,b}_{j_{\alpha_2^{a,b}}}$. In terms of expanding the objective in multi-dividing ontology risk (3.20), we yield

$$\begin{aligned}&\sum_{a=1}^{k-1}\sum_{b=a+1}^{k}\max_{\Xi^{a,b}=\{\xi_1^{a,b},\cdots,\xi^{a,b}_{j_{\alpha_2^{a,b}}}\}\subseteq S^b}\max_{\pi\in\Pi_{n_a,j_{\alpha_2^{a,b}}}}\frac{1}{n_a(j_{\alpha_2^{a,b}}-j_{\alpha_1^{a,b}})}\sum_{i=1}^{n_a}[\sum_{j=j_{\alpha_1^{a,b}}+1}^{j_{\alpha_2^{a,b}}}\pi^{a,b}_{i(j)_{\pi^{a,b}}} \\ &\qquad -\sum_{j=1}^{j_{\alpha_2^{a,b}}}\pi^{a,b}_{ij}\beta^\top v_i^a+\sum_{j=1}^{j_{\alpha_2^{a,b}}}\pi^{a,b}_{ij}\beta^\top\xi_j^{a,b}].\end{aligned}$$

By setting $q_j^{a,b}=\sum_{i=1}^{n_a}\pi_{ij}^{a,b}$, the above expression is equivalent to

$$\begin{aligned}&\sum_{a=1}^{k-1}\sum_{b=a+1}^{k}\frac{1}{n_a(j_{\alpha_2^{a,b}}-j_{\alpha_1^{a,b}})}\max_{\pi\in\Pi^{n_a,j_{\alpha_2^{a,b}}}}[\sum_{i=1}^{n_a}[\sum_{j=j_{\alpha_1^{a,b}}+1}^{j_{\alpha_2^{a,b}}}\pi^{a,b}_{i(j)_{\pi^{a,b}}}-\sum_{j=1}^{j_{\alpha_2^{a,b}}}\pi^{a,b}_{ij}\beta^\top v_i^a] \\ &\qquad +\max_{\Xi^{a,b}=\{\xi_1^{a,b},\cdots,\xi^{a,b}_{j_{\alpha_2^{a,b}}}\}\subseteq S^b}\sum_{j=1}^{j_{\alpha_2^{a,b}}}q_j^{a,b}\beta^\top\xi_j^{a,b}].\end{aligned}$$

Observe that the only term above which relies on $\Xi^{a,b}$ is the third term, and this term gets the largest value if $\Xi^{a,b}$ contains the $j_{\alpha_2^{a,b}}$ vertices with the highest values by β. This implies the fact that: let $\overline{\Xi}^{a,b}=\{\overline{\xi}_1^{a,b},\cdots,\overline{\xi}^{a,b}_{j_{\alpha_2^{a,b}}}\}\subseteq S^b$ be the set of vertices in the top $j_{\alpha_2^{a,b}}$ positions (in descending order of values) by $f=\beta^\top v$, then the maximum value of multi-dividing ontology risk (3.20) is attained at $\overline{\Xi}^{a,b}$ for each pair of (a,b) with $1\le a<b\le k$.

As the last part of this subsection, we characterize the $\widehat{R}^{\text{tight}}_{\text{PMDO}(\alpha_1,\alpha_2)}$ which is described in the following theorem.

Theorem 5 *Let $0<\alpha_1^{a,b}<\alpha_2^{a,b}\le 1$ for each pair of (a,b) with $1\le a<b\le k$. Let $\widehat{R}^{\text{hinge}}_{\text{PMDO}(\alpha_1,\alpha_2)}$ be the multi-dividing ontology risk with hinge ontology loss in (3.10). Define*

$$\widehat{R}^{\text{hinge},+}_{\text{PMDO}(\alpha_1,\alpha_2)}(\beta,S)=\sum_{a=1}^{k-1}\sum_{b=a+1}^{k}\frac{1}{n_a(j_{\alpha_2^{a,b}}-j_{\alpha_1^{a,b}})}\sum_{i:\beta^\top v_i^a<\beta^\top v^b_{(j_{\alpha_1^{a,b}})_\beta}}\sum_{j=1}^{j_{\alpha_1^{a,b}}}(1-\beta^\top(v_i^a-v^b_{(j)_\beta}))_+,$$

$$\zeta_{[0,\alpha_1]}(\beta)=\sum_{a=1}^{k-1}\sum_{b=a+1}^{k}\frac{1}{n_a(j_{\alpha_2^{a,b}}-j_{\alpha_1^{a,b}})}\sum_{i=1}^{n_a}\sum_{j=1}^{j_{\alpha_1^{a,b}}}(-\beta^\top(v_i^a-v_{(j)_\beta}^b))_+,$$

and

$$\zeta^+_{[0,\alpha_1]}(\beta)=\sum_{a=1}^{k-1}\sum_{b=a+1}^{k}\frac{1}{n_a(j_{\alpha_2^{a,b}}-j_{\alpha_1^{a,b}})}\sum_{i:\beta^\top v_i^a<\beta^\top v^b_{(j_{\alpha_1^{a,b}})_\beta}}\sum_{j=1}^{j_{\alpha_1^{a,b}}}(-\beta^\top(v_i^a-v_{(j)_\beta}^b))_+.$$

Then for any ontology sample set $S=(S^1,S^2,\cdots,S^k)\in V^{n_1}\times V^{n_2}\times\cdots\times V^{n_k}$ *and* $\beta\in\mathbb{R}^p$*, we have*

$$\widehat{R}^{\text{hinge},+}_{\text{PMDO}(\alpha_1,\alpha_2)}(\beta,S)+\zeta^+_{[0,\alpha_1]}(\beta)\le\widehat{R}^{\text{tight}}_{\text{PMDO}(\alpha_1,\alpha_2)}(\beta,S)\le\widehat{R}^{\text{hinge}}_{\text{PMDO}(\alpha_1,\alpha_2)}(\beta,S)+\zeta_{[0,\alpha_1]}(\beta),$$

Moreover, if $[\beta^\top v_i^a-\beta^\top v_j^b]\ge 1$ *for each pair of* (a,b) *with* $1\le a<b\le k$ *and any* $i\in\{1,\cdots,n_a\}$*,* $j\in\{1,\cdots,n_b\}$*, then we obtain*

$$\widehat{R}^{\text{tight}}_{\text{PMDO}(\alpha_1,\alpha_2)}(\beta,S)=\widehat{R}^{\text{hinge}}_{\text{PMDO}(\alpha_1,\alpha_2)}(\beta,S)+\zeta_{[0,\alpha_1]}(\beta).$$

Proof of Theorem 5. For each pair of (a,b) with $1\le a<b\le k$, set

$$\Pi^{a,b,\beta}_{n_a,j_{\alpha_2^{a,b}}}=\{\pi^{a,b}\in\Pi_{n_a,j_{\alpha_2^{a,b}}}|\forall i,j_1<j_2:\pi^{a,b}_{i,(j_1)_\beta}\ge\pi^{a,b}_{i,(j_2)_\beta}\},$$

where as before $(j)_\beta$ denotes the j-th position ranked vertices in S^a or equivalently in $\overline{\Xi}^{a,b}$ when the vertices are ranked in descending order by $f=\beta^\top v$. The structural SVM based multi-dividing ontology risk in interval $[\alpha_1,\alpha_2]$ (for each pair of (a,b) with $1\le a<b\le k$) can be deduced in the following simplified version:

$$\begin{aligned}
&\widehat{R}^{\text{tight}}_{\text{PMDO}(\alpha_1,\alpha_2)}(\beta,S)\\
=\;&\sum_{a=1}^{k-1}\sum_{b=a+1}^{k}\max_{\pi^{a,b}\in\Pi^{a,b,\beta}_{n_a,j_{\alpha_2^{a,b}}}}\frac{1}{n_a(j_{\alpha_2^{a,b}}-j_{\alpha_1^{a,b}})}\sum_{i=1}^{n_a}[-\sum_{j=1}^{j_{\alpha_1^{a,b}}}\pi^{a,b}_{i(j)_\beta}\beta^\top(v_i^a-\overline{\xi}_j^{a,b})\\
&+\sum_{j=j_{\alpha_1^{a,b}}+1}^{j_{\alpha_2^{a,b}}}\pi^{a,b}_{i(j)_\beta}(1-\beta^\top(v_i^a-\overline{\xi}_j^{a,b}))]\\
=\;&\sum_{a=1}^{k-1}\sum_{b=a+1}^{k}\max_{\pi^{a,b}\in\Pi^{a,b,\beta}_{n_a,j_{\alpha_2^{a,b}}}}\frac{1}{n_a(j_{\alpha_2^{a,b}}-j_{\alpha_1^{a,b}})}\sum_{i=1}^{n_a}[-\sum_{j=1}^{j_{\alpha_1^{a,b}}}\pi^{a,b}_{i(j)_\beta}\beta^\top(v_i^a-v^b_{(j)_\beta})\\
&+\sum_{j=j_{\alpha_1^{a,b}}+1}^{j_{\alpha_2^{a,b}}}\pi^{a,b}_{i(j)_\beta}(1-\beta^\top(v_i^a-v^b_{(j)_\beta}))]\\
\le\;&\sum_{a=1}^{k-1}\sum_{b=a+1}^{k}\frac{1}{n_a(j_{\alpha_2^{a,b}}-j_{\alpha_1^{a,b}})}\sum_{i=1}^{n_a}[-\sum_{j=1}^{j_{\alpha_1^{a,b}}}\max_{\pi^{a,b}_{i(j)_\beta}\in\{0,1\}}\pi^{a,b}_{i(j)_\beta}\beta^\top(v_i^a-v^b_{(j)_\beta})
\end{aligned}$$

$$+\sum_{j=j_{\alpha_1^{a,b}}+1}^{j_{\alpha_2^{a,b}}}\max_{\pi^{a,b}_{i(j)_\beta}\in\{0,1\}}\pi^{a,b}_{i(j)_\beta}(1-\beta^\top(v_i^a-v^b_{(j)_\beta}))]$$

$$=\sum_{a=1}^{k-1}\sum_{b=a+1}^{k}\frac{1}{n_a(j_{\alpha_2^{a,b}}-j_{\alpha_1^{a,b}})}\sum_{i=1}^{n_a}[\sum_{j=1}^{j_{\alpha_1^{a,b}}}(-\beta^\top(v_i^a-v^b_{(j)_\beta}))_+ +\sum_{j=j_{\alpha_1^{a,b}}+1}^{j_{\alpha_2^{a,b}}}(1-\beta^\top(v_i^a-v^b_{(j)_\beta}))_+].$$

For each pair of (a,b) with $1\le a<b\le k$, define $\widehat{\pi}^{a,b}\in\{0,1\}^{n_a\times j_{\alpha_2^{a,b}}}$ as follows: for each v_i^a such that $\beta^\top v_i^a\le\beta^\top v^b_{(j_{\alpha_1^{a,b}})_\beta}$, we have $\widehat{\pi}^{a,b}_{i(j)_\beta}=1$ if $j\in\{1,\cdots,j_{\alpha_1^{a,b}}\}$ and $\widehat{\pi}^{a,b}_{i(j)_\beta}=\mathbf{I}(\beta^\top v_i^a-\beta^\top v^b_{(j)_\beta}\le 1)$ otherwise; for each v_i^a such that $\beta^\top v_i^a>\beta^\top v^b_{(j_{\alpha_1^{a,b}})_\beta}$, we have $\widehat{\pi}^{a,b}_{i(j)_\beta}=\mathbf{I}(\beta^\top v_i^a-\beta^\top v^b_{(j)_\beta}\le 0)$ if $j\in\{1,\cdots,j_{\alpha_1^{a,b}}-1\}$ and $\widehat{\pi}^{a,b}_{i(j)_\beta}=0$ otherwise.

Obviously, $\widehat{\pi}^{a,b}$ is a valid ordering matrix in $\Pi^{a,b,\beta}_{n_a,j_{\alpha_2^{a,b}}}$. Since for each pair of (a,b) with $1\le a<b\le k$ and for i satisfies $\beta^\top v_i^a\le\beta^\top v^b_{(j_{\alpha_1^{a,b}})_\beta}$ and $j\in\{1,\cdots,j_{\alpha_1^{a,b}}\}$, we get $-\beta^\top(v_i^a-v^b_{(j)_\beta})=-\beta^\top v_i^a+\beta^\top v^b_{(j)_\beta}\ge-\beta^\top v_i^a+\beta^\top v^b_{(j_{\alpha_1^{a,b}})_\beta}\ge 0$, which means $-\beta^\top(v_i^a-v^b_{(j)_\beta})=(-\beta^\top(v_i^a-v^b_{(j)_\beta}))_+$. Hence, by the definition of $\widehat{\pi}^{a,b}$, we get

$$\begin{aligned}
&\widehat{R}^{\text{tight}}_{\text{PMDO}(\alpha_1,\alpha_2)}(\beta,S)\\
\ge&\sum_{a=1}^{k-1}\sum_{b=a+1}^{k}\frac{1}{n_a(j_{\alpha_2^{a,b}}-j_{\alpha_1^{a,b}})}\sum_{i=1}^{n_a}[\sum_{j=1}^{j_{\alpha_1^{a,b}}}\widehat{\pi}^{a,b}_{i(j)_\beta}(-\beta^\top(v_i^a-v^b_{(j)_\beta}))+\sum_{j=j_{\alpha_1^{a,b}}+1}^{j_{\alpha_2^{a,b}}}\widehat{\pi}^{a,b}_{i(j)_\beta}(1-\beta^\top(v_i^a-v^b_{(j)_\beta}))]\\
=&\sum_{a=1}^{k-1}\sum_{b=a+1}^{k}\{\frac{1}{n_a(j_{\alpha_2^{a,b}}-j_{\alpha_1^{a,b}})}\sum_{i:\beta^\top v_i^a\le\beta^\top v^b_{(j_{\alpha_1^{a,b}})_\beta}}[\sum_{j=1}^{j_{\alpha_1^{a,b}}}(1)(-\beta^\top(v_i^a-v^b_{(j)_\beta}))\\
&+\sum_{j=j_{\alpha_1^{a,b}}+1}^{j_{\alpha_2^{a,b}}}(1-\beta^\top(v_i^a-v^b_{(j)_\beta}))_+]\\
&+\frac{1}{n_a(j_{\alpha_2^{a,b}}-j_{\alpha_1^{a,b}})}\sum_{i:\beta^\top v_i^a>\beta^\top v^b_{(j_{\alpha_1^{a,b}})_\beta}}[\sum_{j=1}^{j_{\alpha_1^{a,b}}}(-\beta^\top(v_i^a-v^b_{(j)_\beta}))_+ +0]\}\\
=&\sum_{a=1}^{k-1}\sum_{b=a+1}^{k}\{\frac{1}{n_a(j_{\alpha_2^{a,b}}-j_{\alpha_1^{a,b}})}\sum_{i:\beta^\top v_i^a\le\beta^\top v^b_{(j_{\alpha_1^{a,b}})_\beta}}[\sum_{j=1}^{j_{\alpha_1^{a,b}}}(-\beta^\top(v_i^a-v^b_{(j)_\beta}))_+\\
&+\sum_{j=j_{\alpha_1^{a,b}}+1}^{j_{\alpha_2^{a,b}}}(1-\beta^\top(v_i^a-v^b_{(j)_\beta}))_+]\\
&+\frac{1}{n_a(j_{\alpha_2^{a,b}}-j_{\alpha_1^{a,b}})}\sum_{i:\beta^\top v_i^a>\beta^\top v^b_{(j_{\alpha_1^{a,b}})_\beta}}\sum_{j=1}^{j_{\alpha_1^{a,b}}}(-\beta^\top(v_i^a-v^b_{(j)_\beta}))_+\}.
\end{aligned}\tag{3.21}$$

Since the last term is not less than zero, the desired bound is obtained.

At last, we present that the upper bound on the multi-dividing ontology risk establishes with equality if $|\beta^\top v_i^a - \beta^\top v_j^b| \ge 1$ for any $i \in \{1, \cdots, n_a\}$ and $j \in \{1, \cdots, n_b\}$. In this case, if $\beta^\top(v_i^a - v^b_{(j)_\beta})$ is positive for some $i \in \{1, \cdots, n_a\}$ and $j \in \{1, \cdots, n_b\}$, then it is also the situation $\beta^\top(v_i^a - v^b_{(j)_\beta}) \ge 1$. This implies, for each pair of (a, b) with $1 \le a < b \le k$, we obtain

$$\frac{1}{n_a(j_{\alpha_2^{a,b}} - j_{\alpha_1^{a,b}})} \sum_{i:\beta^\top v_i^a > \beta^\top v^b_{(j_{\alpha_1^{a,b}})_\beta}} \sum_{j=j_{\alpha_1^{a,b}}+1}^{j_{\alpha_2^{a,b}}} (1 - \beta^\top(v_i^a - v^b_{(j)_\beta}))_+ = 0.$$

Combining this to (3.21), we have

$$\begin{aligned}\widehat{R}^{\text{tight}}_{\text{PMDO}(\alpha_1,\alpha_2)}(\beta, S) \quad \ge \quad & \sum_{a=1}^{k-1} \sum_{b=a+1}^{k} \{\frac{1}{n_a(j_{\alpha_2^{a,b}} - j_{\alpha_1^{a,b}})} \sum_{i=1}^{n_a} [\sum_{j=1}^{j_{\alpha_1^{a,b}}} (-\beta^\top(v_i^a - v^b_{(j)_\beta}))_+ \\ & + \sum_{j=j_{\alpha_1^{a,b}}+1}^{j_{\alpha_2^{a,b}}} (1 - \beta^\top(v_i^a - v^b_{(j)_\beta}))_+].\end{aligned}$$

Therefore, we finish the proof of Theorem 5. □

3.5 Generalization bound for partial multi-dividing ontology learning algorithm

In this part, we consider the generalization properties of partial multi-dividing ontology learning algorithm, and the uniform convergence generalization bound is derived. By this way, we establish a good training performance generalization performance by virtue of partial multi-dividing ontology learning algorithm. For each pair of (a, b) with $1 \le a < b \le k$, let $\Theta^{a,b}_{\alpha_1^{a,b},\alpha_2^{a,b}}(f, v^b)$ be the indicator function which is 1 if $\mathbb{P}_{\tilde{v}^b \sim \mathcal{D}_b}(f(\tilde{v}^b) > f(v^b)) \in [\alpha_1^{a,b}, \alpha_2^{a,b}]$ and is 0 otherwise; $\widehat{\Theta}^{a,b}_{\alpha_1^{a,b},\alpha_2^{a,b}}(f, v_j^b)$ be an indicator function which is 1 if v_j^b in positions from $j_{\alpha_1^{a,b}} + 1$ to $j_{\alpha_2^{a,b}}$ in the ranking of all vertices in S^b by ontology function f. For an ontology function $f : \mathbb{R}^p \to \mathbb{R}$, define the expectation risk version as

$$R_{\text{PMDO}(\alpha_1,\alpha_2)}[f, \mathcal{D}] = \sum_{a=1}^{k-1} \sum_{b=a+1}^{k} \frac{1}{\alpha_2^{a,b} - \alpha_1^{a,b}} \mathbb{E}_{v^a \sim \mathcal{D}_a, v^b \sim \mathcal{D}_b}[\mathbf{I}(f(v^a) \le f(v^b)) \Theta^{a,b}_{\alpha_1^{a,b},\alpha_2^{a,b}}(f, v^b)], \quad (3.22)$$

and for an ontology sample set $S = (S^1, S^2, \cdots, S^k) \in V^{n_1} \times V^{n_2} \times \cdots \times V^{n_k}$, its corresponding empirical version is

$$\widehat{R}_{\text{PMDO}(\alpha_1,\alpha_2)}[f, S] = \sum_{a=1}^{k-1} \sum_{b=a+1}^{k} \frac{1}{n_a(j_{\alpha_2^{a,b}} - j_{\alpha_1^{a,b}})} \sum_{i=1}^{n_a} \sum_{j=1}^{n_b} [\mathbf{I}(f(v_i^a) \le f(v_i^b)) \widehat{\Theta}^{a,b}_{\alpha_1^{a,b},\alpha_2^{a,b}}(f, v_j^b)]. \quad (3.23)$$

Let $\mathcal{F}$ be the ontology function space, and the capacity of $\mathcal{F}$ will be measured in view of the VC dimension of the kind of classifiers yielded from ontology functions in the special class: $\Upsilon_{\mathcal{F}} = \{\text{sign} \circ (f - t) | f \in \mathcal{F}, t \in \mathbb{R}\}$. As the last theorem in our paper, we state below the uniform convergence bound for partial multi-dividing ontology algorithm.

Theorem 6 *Let $\mathcal{F}$ be a class of real-valued ontology functions on V, and $\Upsilon_{\mathcal{F}} = \{\text{sign} \circ (f - t) | f \in \mathcal{F}, t \in \mathbb{R}\}$. For all pairs of (a, b) with $1 \le a < b \le k$, let $n_{amax} = \max\{n_1, \cdots, n_{k-1}\}$, $n_{bmax} = \max\{n_2, \cdots, n_k\}$, $(\alpha_1^{\max}, \alpha_2^{\max}) = \max\{\alpha_2^{a,b} - \alpha_1^{a,b} | 1 \le a < b \le k\}$. Let $\delta > 0$. Then with probability at least $1-\delta$ over ontology sample $S = (S^1, S^2, \cdots, S^k) \in V^{n_1} \times V^{n_2} \times \cdots \times V^{n_k}$ from $\mathcal{D}_1 \times \cdots \times \mathcal{D}_k$, we have for all $f \in \mathcal{F}$,*

$$\begin{aligned} R_{\text{PMDO}(\alpha_1,\alpha_2)}[f, \mathcal{D}] \quad \le \quad & \widehat{R}_{\text{PMDO}(\alpha_1,\alpha_2)}[f, S] \\ & + C\left(\sqrt{\frac{p\ln(n_{amax}) + \ln(\frac{k(k-1)}{2\delta})}{n_{amax}}} + \frac{1}{\alpha_2^{\max} - \alpha_1^{\max}} \sqrt{\frac{p\ln(n_{bmax}) + \ln(\frac{k(k-1)}{2\delta})}{n_{bmax}}}\right), \end{aligned}$$

where p is the VC dimension of $\Upsilon_{\mathcal{F}}$, and positive parameter C is distribution-independent.

Proof of Theorem 6. Assume that f has no ties and $n_b\alpha_2^{a,b}$ is an integer. For each pair of (a, b) with $1 \le a < b \le k$, we define

$$t^{a,b}_{\mathcal{D}_b, f, \alpha_2^{a,b}} = \arg\inf_{t \in \mathbb{R}} \{t \in \mathbb{R} | \mathbb{P}_{v^b \sim \mathcal{D}_b}[f(v^b) > t] = \alpha_2^{a,b}\},$$

and its empirical version as

$$\widehat{t}^{a,b}_{S^b, f, \alpha_2^{a,b}} = \arg\min_{t \in \mathbb{R}} \{t \in \mathbb{R} | \frac{1}{n_b} \sum_{j=1}^{n_b} \mathbf{I}(f(v_j^b) > t) \ge \alpha_2^{a,b}\}.$$

Clearly, $\mathbb{E}_{v^b \sim \mathcal{D}_b}[\mathbf{I}(f(v^b) > t^{a,b}_{\mathcal{D}_b, f, \alpha_2^{a,b}})] = \alpha_2^{a,b}$. Since ontology function f has no ties, and $\widehat{t}^{a,b}_{S^b, f, \alpha_2^{a,b}}$ is the threshold on f which $n_b\alpha_2^{a,b}$ vertices in S^b are ranked by f, we infer $\sum_{j=1}^{n_b} \mathbf{I}(f(v_j^b) > \widehat{t}^{a,b}_{S^b, f, \alpha_2^{a,b}}) = n_b\alpha_2^{a,b}$.

For any ontology function $f : V \to \mathbb{R}$, the partial multi-dividing ontology risk in intervals $[0, \alpha_2^{a,b}]$ (pair (a, b) satisfy $1 \le a < b \le k$) can be stated as

$$R_{\alpha_2}[f, \mathcal{D}] = \sum_{a=1}^{k-1} \sum_{b=a+1}^{k} \frac{1}{\alpha_2^{a,b}} \mathbb{E}_{v^a \sim \mathcal{D}_a, v^b \sim \mathcal{D}_b}[\mathbf{I}(f(v^a) \le f(v^b), f(v^b) > t^{a,b}_{\mathcal{D}_b, f, \alpha_2^{a,b}})],$$

and its empirical version on ontology sample set S can be formulated by

$$\widehat{R}_{\alpha_2}[f, S] = \sum_{a=1}^{k-1} \sum_{b=a+1}^{k} \frac{1}{n_a n_b \alpha_2^{a,b}} \sum_{i=1}^{n_a} \sum_{j=1}^{n_b} \mathbf{I}(f(v_i^a) \le f(v_j^b), f(v_j^b) > \widehat{t}^{a,b}_{S^b, f, \alpha_2^{a,b}}).$$

For $1 \le a \le k$ and fixed any $\epsilon > 0$, using McDiarmid inequality, we get

$$\mathbb{P}_{S^a \sim \mathcal{D}_a^{n_a}}\left(\bigcup_{f \in \mathcal{F}} \bigcup_{t \in \mathbb{R}} \{|\frac{1}{n_a} \sum_{i=1}^{n_a} \mathbf{I}(f(v_i^a) \le t) - \mathbb{E}_{v^a \sim \mathcal{D}_a}[\mathbf{I}(f(v^a) \le t)]| \ge \epsilon\}\right) \le C^a n_a^p e^{-2n_a\epsilon^2} \tag{3.24}$$

where p is the VC dimension of $\Upsilon_{\mathcal{F}}$, and positive parameter C^a is distribution-independent.

For each pair of (a, b) with $1 \le a < b \le k$, an ontology function f and vertex v^b, define $l^{a,b}(f, v^b) = \mathbb{E}_{v^a \sim \mathcal{D}_a}[\mathbf{I}(f(v^a) \le f(v^b))]$. Moreover, we set

$$\tilde{R}^{a,b}_{\alpha_2^{a,b}}[f, \mathcal{D}_a \cup \mathcal{D}_b, S^b] = \frac{1}{n_b\alpha_2^{a,b}} \sum_{j=1}^{n_b} l^{a,b}(f, v_j^b)\mathbf{I}(f(v_j^b) > t^{a,b}_{\mathcal{D}_b, f, \alpha_2^{a,b}}),$$

$$\overline{R}^{a,b}_{\alpha_2^{a,b}}[f,\mathcal{D}_a,S^b]=\frac{1}{n_b\alpha_2^{a,b}}\sum_{j=1}^{n_b}l^{a,b}(f,v_j^b)\mathbf{I}(f(v_j^b)>\widehat{t}^{a,b}_{S^b,f,\alpha_2^{a,b}}).$$

Hence, for any $f\in\mathcal{F}$, we obtain

$$\begin{aligned}R_{\alpha_2}[f,\mathcal{D}]-\widehat{R}_{\alpha_2}[f,S] &= (R_{\alpha_2}[f,\mathcal{D}]-\sum_{a=1}^{k-1}\sum_{b=a+1}^{k}\tilde{R}^{a,b}_{\alpha_2^{a,b}}[f,\mathcal{D}_a\cup\mathcal{D}_b,S^b])\\ &\quad+(\sum_{a=1}^{k-1}\sum_{b=a+1}^{k}\tilde{R}^{a,b}_{\alpha_2^{a,b}}[f,\mathcal{D}_a\cup\mathcal{D}_b,S^b]-\sum_{a=1}^{k-1}\sum_{b=a+1}^{k}\overline{R}^{a,b}_{\alpha_2^{a,b}}[f,\mathcal{D}_a,S^b])\\ &\quad+(\sum_{a=1}^{k-1}\sum_{b=a+1}^{k}\overline{R}^{a,b}_{\alpha_2^{a,b}}[f,\mathcal{D}_a,S^b]-\widehat{R}_{\alpha_2}[f,S]).\end{aligned}$$

Thus for any $\epsilon>0$, we infer

$$\begin{aligned}&\mathbb{P}_{S=(S^1,S^2,\cdots,S^k)\in V^{n_1}\times V^{n_2}\times\cdots\times V^{n_k}}(\sup_{f\in\mathcal{F}}\{R_{\alpha_2}[f,\mathcal{D}]-\widehat{R}_{\alpha_2}[f,S]\geq\epsilon\})\\ \leq\ &\sum_{a=1}^{k-1}\sum_{b=a+1}^{k}\mathbb{P}_{S^b\sim\mathcal{D}_b^{n_b}}(\bigcup_{f\in\mathcal{F}}\{\frac{1}{\alpha_2^{a,b}}\mathbb{E}_{v^a\sim\mathcal{D}_a,v^b\sim\mathcal{D}_b}[\mathbf{I}(f(v^a)\leq f(v^b),f(v^b)>t^{a,b}_{\mathcal{D}_b,f,\alpha_2^{a,b}})]\\ &-\tilde{R}^{a,b}_{\alpha_2^{a,b}}[f,\mathcal{D}_a\cup\mathcal{D}_b,S^b]\geq\frac{\epsilon}{3}\})\\ &+\sum_{a=1}^{k-1}\sum_{b=a+1}^{k}\mathbb{P}_{S^b\sim\mathcal{D}_b^{n_b}}(\bigcup_{f\in\mathcal{F}}\{\tilde{R}^{a,b}_{\alpha_2^{a,b}}[f,\mathcal{D}_a\cup\mathcal{D}_b,S^b]-\overline{R}^{a,b}_{\alpha_2^{a,b}}[f,\mathcal{D}_a,S^b]\geq\frac{\epsilon}{3}\})\\ &+\sum_{a=1}^{k-1}\sum_{b=a+1}^{k}\mathbb{P}_{S^a\times S^b\sim\mathcal{D}_a^{n_a}\times\mathcal{D}_b^{n_b}}(\bigcup_{f\in\mathcal{F}}\{\overline{R}^{a,b}_{\alpha_2^{a,b}}[f,\mathcal{D}_a,S^b]\\ &-\frac{1}{n_an_b\alpha_2^{a,b}}\sum_{i=1}^{n_a}\sum_{j=1}^{n_b}\mathbf{I}(f(v_i^a)\leq f(v_j^b),f(v_j^b)>\widehat{t}^{a,b}_{S^b,f,\alpha_2^{a,b}})\geq\frac{\epsilon}{3}\}).\end{aligned}$$

Let N_1, N_2 and N_3 be the first, second and third terms of right side of the above inequality. We need to bound N_1, N_2 and N_3 separately. For the first part, according to the assumption that ontology function f has no ties, we have

$$\begin{aligned}&R_{\alpha_2}[f,\mathcal{D}]-\sum_{a=1}^{k-1}\sum_{b=a+1}^{k}\tilde{R}^{a,b}_{\alpha_2^{a,b}}[f,\mathcal{D}_a\cup\mathcal{D}_b,S^b]\\ =\ &\sum_{a=1}^{k-1}\sum_{b=a+1}^{k}\{\frac{1}{\alpha_2^{a,b}}\mathbb{E}_{v^b}[l^{a,b}(f,v^b)f(v^b)>t^{a,b}_{\mathcal{D}_b,f,\alpha_2^{a,b}})]-\frac{1}{n_b\alpha_2^{a,b}}\sum_{j=1}^{n_b}l^{a,b}(f,v_j^b)\mathbf{I}(f(v_j^b)>t^{a,b}_{\mathcal{D}_b,f,\alpha_2^{a,b}})\}\\ =\ &\sum_{a=1}^{k-1}\sum_{b=a+1}^{k}\{\frac{1}{\alpha_2^{a,b}}\mathbb{E}_{v^a}[\mathbb{E}_{v^b}[\mathbf{I}(f(v^a)\leq f(v^b))\mathbf{I}(f(v^b)>t^{a,b}_{\mathcal{D}_b,f,\alpha_2^{a,b}})]\\ &-\frac{1}{n_b}\sum_{j=1}^{n_b}\mathbf{I}(f(v^a)\leq f(v_j^b))\mathbf{I}(f(v_j^b)>t^{a,b}_{\mathcal{D}_b,f,\alpha_2^{a,b}})]\}\\ =\ &\sum_{a=1}^{k-1}\sum_{b=a+1}^{k}\{\frac{1}{\alpha_2^{a,b}}\mathbb{E}_{v^a}[\mathbb{E}_{v^b}[\mathbf{I}(f(v^b)>\max\{f(v^a),t^{a,b}_{\mathcal{D}_b,f,\alpha_2^{a,b}}\})]\\ &-\frac{1}{n_b}\sum_{j=1}^{n_b}\mathbf{I}(f(v_j^b)>\max\{f(v^a),t^{a,b}_{\mathcal{D}_b,f,\alpha_2^{a,b}}\})]\}\end{aligned}$$

$$
\begin{aligned}
&\leq \sum_{a=1}^{k-1}\sum_{b=a+1}^{k}\{\frac{1}{\alpha_2^{a,b}}\sup_{v^a\in V}[\mathbb{E}_{v^b}[\mathbf{I}(f(v^b)>\max\{f(v^a),t^{a,b}_{\mathcal{D}_b,f,\alpha_2^{a,b}}\})]\\
&\quad -\frac{1}{n_b}\sum_{j=1}^{n_b}\mathbf{I}(f(v_j^b)>\max\{f(v^a),t^{a,b}_{\mathcal{D}_b,f,\alpha_2^{a,b}}\})]\}\\
&\leq \sum_{a=1}^{k-1}\sum_{b=a+1}^{k}\sup_{t\in\mathbb{R}}|\frac{1}{\alpha_2^{a,b}}\mathbb{E}_{v^b}(\mathbf{I}(f(v^b)>t))-\frac{1}{n_b}\sum_{j=1}^{n_b}(\mathbf{I}(f(v_j^b)>t)])|,
\end{aligned}
$$

Thus, using (3.24), we yield

$$
\begin{aligned}
N_1 &= \sum_{a=1}^{k-1}\sum_{b=a+1}^{k}\mathbb{P}_{S^b\sim\mathcal{D}_b^{n_b}}(\bigcup_{f\in\mathcal{F}}\{\frac{1}{\alpha_2^{a,b}}\mathbb{E}_{v^a\sim\mathcal{D}_a,v^b\sim\mathcal{D}_b}[\mathbf{I}(f(v^a)\leq f(v^b),f(v^b)>t^{a,b}_{\mathcal{D}_b,f,\alpha_2^{a,b}})]\\
&\quad -\tilde{R}^{a,b}_{\alpha_2^{a,b}}[f,\mathcal{D}_a\cup\mathcal{D}_b,S^b]\geq\frac{\epsilon}{3}\})\\
&\leq \sum_{a=1}^{k-1}\sum_{b=a+1}^{k}\mathbb{P}_{S^b\sim\mathcal{D}_b^{n_b}}(\bigcup_{f\in\mathcal{F}}\{\sup_{t\in\mathbb{R}}|\mathbb{E}_{v^b}(\mathbf{I}(f(v^b)>t))-\frac{1}{n_b}\sum_{j=1}^{n_b}(\mathbf{I}(f(v_j^b)>t)])|\geq\frac{\alpha_2^{a,b}\epsilon}{3}\})\\
&= \sum_{a=1}^{k-1}\sum_{b=a+1}^{k}\mathbb{P}_{S^b\sim\mathcal{D}_b^{n_b}}(\bigcup_{f\in\mathcal{F}}\bigcup_{t\in\mathbb{R}}\{|\mathbb{E}_{v^b}(\mathbf{I}(f(v^b)>t))-\frac{1}{n_b}\sum_{j=1}^{n_b}(\mathbf{I}(f(v_j^b)>t)])|\geq\frac{\alpha_2^{a,b}\epsilon}{3}\})\\
&\leq \sum_{a=1}^{k-1}\sum_{b=a+1}^{k}C^b n_b^p e^{-\frac{2n_b(\alpha_2^{a,b})^2\epsilon^2}{9}}.
\end{aligned}
$$

For the second term, for each pair of (a,b) with $1\leq a<b\leq k$, note that if $t^{a,b}_{\mathcal{D}_b,f,\alpha_2^{a,b}}\leq\widehat{t}^{a,b}_{S^b,f,\alpha_2^{a,b}}$, then $\mathbf{I}(f(v^b)>t^{a,b}_{\mathcal{D}_b,f,\alpha_2^{a,b}})-\mathbf{I}(f(v^b)>\widehat{t}^{a,b}_{S^b,f,\alpha_2^{a,b}})\geq 0$ for any $v^b\in V$, and if $t^{a,b}_{\mathcal{D}_b,f,\alpha_2^{a,b}}>\widehat{t}^{a,b}_{S^b,f,\alpha_2^{a,b}}$, then $\mathbf{I}(f(v^b)>t^{a,b}_{\mathcal{D}_b,f,\alpha_2^{a,b}})-\mathbf{I}(f(v^b)>\widehat{t}^{a,b}_{S^b,f,\alpha_2^{a,b}})\leq 0$ for any $v^b\in V$. Since one of these two cases will always hold, and l is bounded by ($0\leq l^{a,b}(f,v^b)\leq 1$ for any $v^b\in V$), combining with the definition of $\widehat{t}^{a,b}_{S^b,f,\alpha_2^{a,b}}$ and $t^{a,b}_{\mathcal{D}_b,f,\alpha_2^{a,b}}$, and (3.24), we deduce

$$
\begin{aligned}
N_2 &= \sum_{a=1}^{k-1}\sum_{b=a+1}^{k}\mathbb{P}_{S^b\sim\mathcal{D}_b^{n_b}}(\bigcup_{f\in\mathcal{F}}\{\tilde{R}^{a,b}_{\alpha_2^{a,b}}[f,\mathcal{D}_a\cup\mathcal{D}_b,S^b]-\overline{R}^{a,b}_{\alpha_2^{a,b}}[f,\mathcal{D}_a,S^b]\geq\frac{\epsilon}{3}\})\\
&= \sum_{a=1}^{k-1}\sum_{b=a+1}^{k}\mathbb{P}_{S^b\sim\mathcal{D}_b^{n_b}}(\bigcup_{f\in\mathcal{F}}\{\frac{1}{n_b\alpha_2^{a,b}}\sum_{j=1}^{n_b}l^{a,b}(f,v_j^b)\mathbf{I}(f(v_j^b)>t^{a,b}_{\mathcal{D}_b,f,\alpha_2^{a,b}})\\
&\quad -\frac{1}{n_b\alpha_2^{a,b}}\sum_{j=1}^{n_b}l^{a,b}(f,v_j^b)\mathbf{I}(f(v_j^b)>\widehat{t}^{a,b}_{S^b,f,\alpha_2^{a,b}})\geq\frac{\epsilon}{3}\})\\
&\leq \sum_{a=1}^{k-1}\sum_{b=a+1}^{k}\mathbb{P}_{S^b\sim\mathcal{D}_b^{n_b}}(\bigcup_{f\in\mathcal{F}}\{|\frac{1}{n_b\alpha_2^{a,b}}\sum_{j=1}^{n_b}l^{a,b}(f,v_j^b)[\mathbf{I}(f(v_j^b)>t^{a,b}_{\mathcal{D}_b,f,\alpha_2^{a,b}})\\
&\quad -\mathbf{I}(f(v_j^b)>\widehat{t}^{a,b}_{S^b,f,\alpha_2^{a,b}})|\geq\frac{\epsilon}{3}\})\\
&\leq \sum_{a=1}^{k-1}\sum_{b=a+1}^{k}\mathbb{P}_{S^b\sim\mathcal{D}_b^{n_b}}(\bigcup_{f\in\mathcal{F}}\{|\frac{1}{n_b\alpha_2^{a,b}}\sum_{j=1}^{n_b}[\mathbf{I}(f(v_j^b)>t^{a,b}_{\mathcal{D}_b,f,\alpha_2^{a,b}})\\
&\quad -\mathbf{I}(f(v_j^b)>\widehat{t}^{a,b}_{S^b,f,\alpha_2^{a,b}})|\geq\frac{\epsilon}{3}\})
\end{aligned}
$$

$$
\begin{aligned}
&= \sum_{a=1}^{k-1}\sum_{b=a+1}^{k} \mathbb{P}_{S^b\sim\mathcal{D}_b^{n_b}}(\bigcup_{f\in\mathcal{F}}\{|\frac{1}{n_b}\sum_{j=1}^{n_b}\mathbf{I}(f(v_j^b) > t^{a,b}_{\mathcal{D}_b,f,\alpha_2^{a,b}}) \\
&\quad -\frac{1}{n_b}\sum_{j=1}^{n_b}\mathbf{I}(f(v_j^b) > \widehat{t}^{a,b}_{S^b,f,\alpha_2^{a,b}})| \geq \frac{\alpha_2^{a,b}\epsilon}{3}\}) \\
&= \sum_{a=1}^{k-1}\sum_{b=a+1}^{k} \mathbb{P}_{S^b\sim\mathcal{D}_b^{n_b}}(\bigcup_{f\in\mathcal{F}}\{|\frac{1}{n_b}\sum_{j=1}^{n_b}\mathbf{I}(f(v_j^b) > t^{a,b}_{\mathcal{D}_b,f,\alpha_2^{a,b}}) - \alpha_2^{a,b}| \geq \frac{\alpha_2^{a,b}\epsilon}{3}\}) \\
&= \sum_{a=1}^{k-1}\sum_{b=a+1}^{k} \mathbb{P}_{S^b\sim\mathcal{D}_b^{n_b}}(\bigcup_{f\in\mathcal{F}}\{|\frac{1}{n_b}\sum_{j=1}^{n_b}\mathbf{I}(f(v_j^b) > t^{a,b}_{\mathcal{D}_b,f,\alpha_2^{a,b}}) \\
&\quad -\mathbb{E}_{v^b}[\mathbf{I}(f(v^b) > t^{a,b}_{\mathcal{D}_b,f,\alpha_2^{a,b}})]| \geq \frac{\alpha_2^{a,b}\epsilon}{3}\}) \\
&\leq \sum_{a=1}^{k-1}\sum_{b=a+1}^{k} \mathbb{P}_{S^b\sim\mathcal{D}_b^{n_b}}(\bigcup_{f\in\mathcal{F}}\bigcup_{t\in\mathbb{R}}\{|\frac{1}{n_b}\sum_{j=1}^{n_b}\mathbf{I}(f(v_j^b) > t) - \mathbb{E}_{v^b}[\mathbf{I}(f(v^b) > t)]| \geq \frac{\alpha_2^{a,b}\epsilon}{3}\}) \\
&\leq \sum_{a=1}^{k-1}\sum_{b=a+1}^{k} C^b n_b^p e^{-\frac{2n_b(\alpha_2^{a,b})^2\epsilon^2}{9}}.
\end{aligned}
$$

Now, we bound the term N_3.

$$
\begin{aligned}
&\sum_{a=1}^{k-1}\sum_{b=a+1}^{k} \overline{R}^{a,b}_{\alpha_2^{a,b}}[f,\mathcal{D}_a,S^b] - \widehat{R}_{\alpha_2}[f,S] \\
&= \sum_{a=1}^{k-1}\sum_{b=a+1}^{k}\{\frac{1}{n_b\alpha_2^{a,b}}\sum_{j=1}^{n_b}\mathbb{E}_{v^a}[\mathbf{I}(f(v^a)\leq f(v_j^b), f(v_j^b) > \widehat{t}^{a,b}_{S^b,f,\alpha_2^{a,b}})\}] \\
&\quad -\frac{1}{n_a n_b\alpha_2^{a,b}}\sum_{i=1}^{n_a}\sum_{j=1}^{n_b}\mathbf{I}(f(v_i^a)\leq f(v_j^b), f(v_j^b) > \widehat{t}^{a,b}_{S^b,f,\alpha_2^{a,b}})\} \\
&= \sum_{a=1}^{k-1}\sum_{b=a+1}^{k}\{\frac{1}{n_b\alpha_2^{a,b}}\sum_{j=1}^{n_b}\mathbf{I}(f(v_j^b) > \widehat{t}^{a,b}_{S^b,f,\alpha_2^{a,b}})[\mathbb{E}_{v^a}\mathbf{I}(f(v^a)\leq f(v_j^b)) - \frac{1}{n_a}\sum_{i=1}^{n_a}\mathbf{I}(f(v_i^a)\leq f(v_j^b))]\} \\
&\leq \sum_{a=1}^{k-1}\sum_{b=a+1}^{k}\{\frac{1}{n_b\alpha_2^{a,b}}\sum_{j=1}^{n_b}\mathbf{I}(f(v_j^b) > \widehat{t}^{a,b}_{S^b,f,\alpha_2^{a,b}})\sup_{t\in\mathbb{R}}|\mathbb{E}_{v^a}\mathbf{I}(f(v^a)\leq t) - \frac{1}{n_a}\sum_{i=1}^{n_a}\mathbf{I}(f(v_i^a)\leq t)|\} \\
&= \sum_{a=1}^{k-1}\sum_{b=a+1}^{k}\{\frac{1}{n_b\alpha_2^{a,b}}(n_b\alpha_2^{a,b})\sup_{t\in\mathbb{R}}|\mathbb{E}_{v^a}\mathbf{I}(f(v^a)\leq t) - \frac{1}{n_a}\sum_{i=1}^{n_a}\mathbf{I}(f(v_i^a)\leq t)|\} \\
&= \sum_{a=1}^{k-1}\sum_{b=a+1}^{k}\{\sup_{t\in\mathbb{R}}|\mathbb{E}_{v^a}\mathbf{I}(f(v^a)\leq t) - \frac{1}{n_a}\sum_{i=1}^{n_a}\mathbf{I}(f(v_i^a)\leq t)|\},
\end{aligned}
$$

and thus in light of (3.24), we derive

$$
\begin{aligned}
N_3 &= \sum_{a=1}^{k-1}\sum_{b=a+1}^{k}\mathbb{P}_{S^a\times S^b\sim\mathcal{D}_a^{n_a}\times\mathcal{D}_b^{n_b}}(\bigcup_{f\in\mathcal{F}}\{\overline{R}^{a,b}_{\alpha_2^{a,b}}[f,\mathcal{D}_a,S^b] \\
&\quad -\frac{1}{n_a n_b\alpha_2^{a,b}}\sum_{i=1}^{n_a}\sum_{j=1}^{n_b}\mathbf{I}(f(v_i^a)\leq f(v_j^b), f(v_j^b) > \widehat{t}^{a,b}_{S^b,f,\alpha_2^{a,b}}) \geq \frac{\epsilon}{3}\}) \\
&= \sum_{a=1}^{k-1}\sum_{b=a+1}^{k}\mathbb{E}_{S^b}[\mathbb{P}_{S^a|S^b}(\bigcup_{f\in\mathcal{F}}\{\overline{R}^{a,b}_{\alpha_2^{a,b}}[f,\mathcal{D}_a,S^b]
\end{aligned}
$$

$$-\frac{1}{n_a n_b \alpha_2^{a,b}}\sum_{i=1}^{n_a}\sum_{j=1}^{n_b}\mathbf{I}(f(v_i^a)\le f(v_j^b), f(v_j^b)>\widehat{t}^{a,b}_{S^b,f,\alpha_2^{a,b}})\ge\frac{\epsilon}{3}\})]$$

$$\le \sum_{a=1}^{k-1}\sum_{b=a+1}^{k}\mathbb{E}_{S^b}[\mathbb{P}_{S^a|S^b}(\bigcup_{f\in\mathcal{F}}\{\sup_{t\in\mathbb{R}}|\mathbb{E}_{v^a}[\mathbf{I}(f(v^a)\le t)]]-\frac{1}{n_a}\sum_{i=1}^{n_a}\mathbf{I}(f(v_i^a)\le t)|\ge\frac{\epsilon}{3}\})]$$

$$\le \sum_{a=1}^{k-1}\sum_{b=a+1}^{k}\mathbb{E}_{S^b}[\mathbb{P}_{S^a|S^b}(\bigcup_{f\in\mathcal{F}}\bigcup_{t\in\mathbb{R}}\{|\mathbb{E}_{v^a}[\mathbf{I}(f(v^a)\le t)]]-\frac{1}{n_a}\sum_{i=1}^{n_a}\mathbf{I}(f(v_i^a)\le t)|\ge\frac{\epsilon}{3}\})]$$

$$\le \sum_{a=1}^{k-1}\sum_{b=a+1}^{k}C^a n_a^p e^{-\frac{2n_a\epsilon^2}{9}}.$$

Combining the bounds we show above in three cases, the final result followed by setting the right-hand side equal to δ and solving it for ϵ. □

4 Experiments

To implement our algorithm with mathematical learning setting, for each vertex in each ontology in the experiments, we use fix dimensional vector to express vertex's semantic and construct information. It stated that all the information of the concept include it's name, attribute, instance and structure of vertex in the ontology graph is packaged in its corresponding vector. In this section, four experiments are designed and presented to measure the effectiveness of our partial multi-dividing ontology learning algorithm manifested in the former sections. We use C++ to run our main algorithm, and available LAPACK and BLAS libraries for linear calculating are used here. All four experiments are implemented on a 32G memory multi-core CPU.

4.1 Ontology similarity measure experiment on plant data

In the field of plant science, "PO" ontology O_1 (see http: //www.plantontology.org for more details, and its basic structure is extracted, draw and depicted in Figure 1) is applied to present plant morphology and anatomy, and development stages for nearly all types of plants. The aim of construction "PO" ontology is to establish a semantic framework for worthy cross-species queries across gene expressions and phenotype data collections followed from genetics experiments and plant genomics. Indeed, "PO" ontology can be regarded as a dictionary which helps plant scientists to search the concept and understand the potential connection between plant species, biochemical processes and the surrounding climate and environment impact. We use "PO" ontology to test the productiveness of our partial multi-dividing ontology algorithm with respect to ontology similarity measuring. Since there are two branches "plant structure development stage" and "plant anatomical entity" in its ontology graph, we set $k=2$. We use $P@N$ (Precision Ratio, see Craswell and Hawking [44] for more details on this index) to evaluate the equality of the experiment results. The specific procedures can be stated as follows: firstly, with the help of experts, the most similarity N vertices for each vertex are listed; secondly, in terms of our algorithm, other most similarity N for each

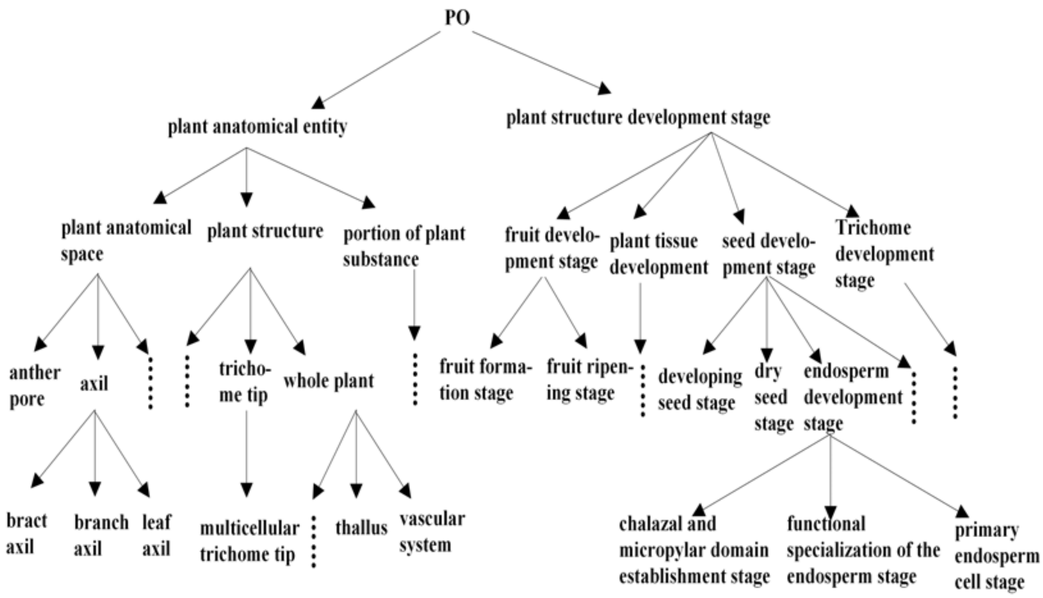

Figure 1: The Structure of "PO" Ontology O_1.

vertex are obtained; thirdly, the $P@N$ precision ratios for all vertices are inferred; lastly, the average precision ratio for the whole ontology graph is finally calculated. Furthermore, ontology learning technologies proposed in Gao and Zhu [35], Gao et al. [45] and [46] are also acted on the "PO" ontology and their corresponding precision ratios are calculated respectively. Several parts of the compared data can be referred to Table 1.

	$P@3$ average precision ratio	$P@5$ average precision ratio	$P@10$ average precision ratio	$P@20$ average precision ratio
PMDO Algorithm in our paper	0.5383	0.6903	0.9156	0.9705
Algorithm in Gao and Zhu [35]	0.5042	0.6216	0.7853	0.9034
Algorithm in Gao et al. [45]	0.4921	6152	0.8113	0.9174
Algorithm in Gao et al. [46]	0.5360	0.6664	0.9004	0.9673

Table 1. The experiment results of ontology similarity measure

In light of the compared data presented in Table 1, we see that the precision ratio using our partial multi-dividing ontology algorithm is clearly higher than the precision ratio determined by algorithms in Gao and Zhu [35], Gao et al. [45] and [46] when N=3, 5, 10 or 20. Hence, we conclude that our partial multi-dividing ontology algorithm discussed in our paper is superior to the learning techniques that Gao and Zhu [35], Gao et al. [45] and [46] raised.

4.2 Ontology mapping experiment on mathematical data

Mathematical ontologies O_2 and O_3 (Figure 2 and Figure 3 present the basic structures of O_2 and O_3, respectively) are employed in our second experiment to test the effectiveness of our new proposed partial multi-dividing algorithm with regard to building ontology mapping between them.

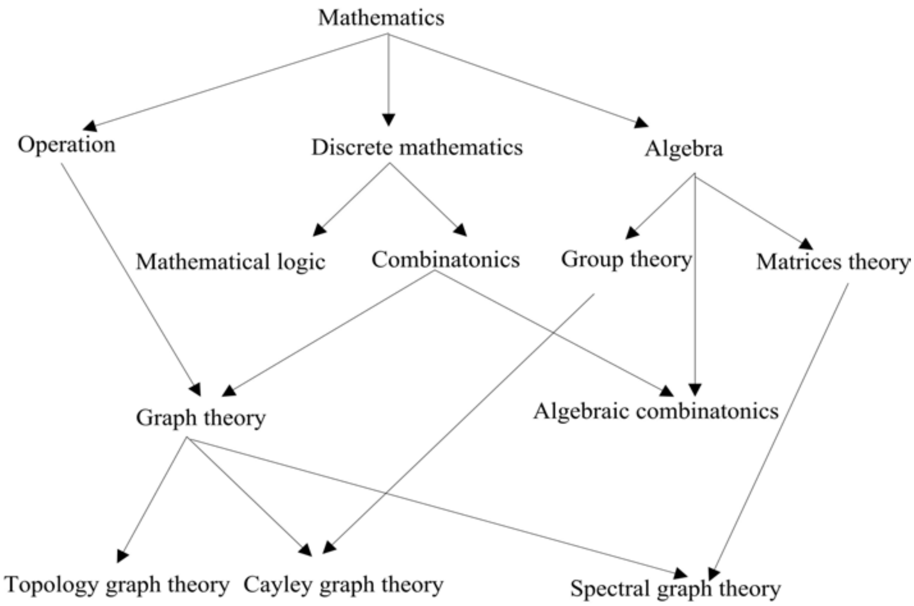

Figure 2: "mathematical" ontology O_2

This experiment aims to determine the similarity based ontology mapping between O_2 and O_3 using our ontology learning algorithm. In view of analyzing of ontology graph structures, we set $k = 3$. The $P@N$ criterion is applied as well as criterion to measure the quality of experiment data. Moreover, ontology learning algorithms introduced in Gao and Zhu [35], Wu et al. [40] and Gao et al. [47] are also implemented on mathematical ontologies, and the precision ratios are compared using these methods. Parts of compared experiment data can be referred to Table 2.

	P@1 average precision ratio	P@3 average precision ratio	P@5 average precision ratio
PMDO Algorithm in our paper	0.4231	0.5128	0.7154
Algorithm in Gao and Zhu [35]	0.3077	0.4359	0.5615
Algorithm in Wu et al. [40]	0.3846	0.5000	0.6769
Algorithm in Gao et al. [47]	0.3462	0.3974	0.5231

Table 2. The experiment results of ontology mapping

The experiment compared results for $N = 1, 3, 5$ presented in Table 2 reveals that our partial multi-dividing ontology algorithm performances much more efficient than ontology learning algorithms propsed in Gao and Zhu [35], Wu et al. [40] and Gao et al. [47] especially as N increases.

4.3 Ontology similarity measure experiment on biology data

In our third experiment, we aim to utilize the "GO" ontology O_4 (see http: //www.geneontology.org for more details. Its basic structure is extracted and showed in Figure 4) to test the practicability of our partial multi-dividing ontology learning algorithm on biology engineering application. As a collaborative effort to deal with the need for consistent expressions of gene products across databases,

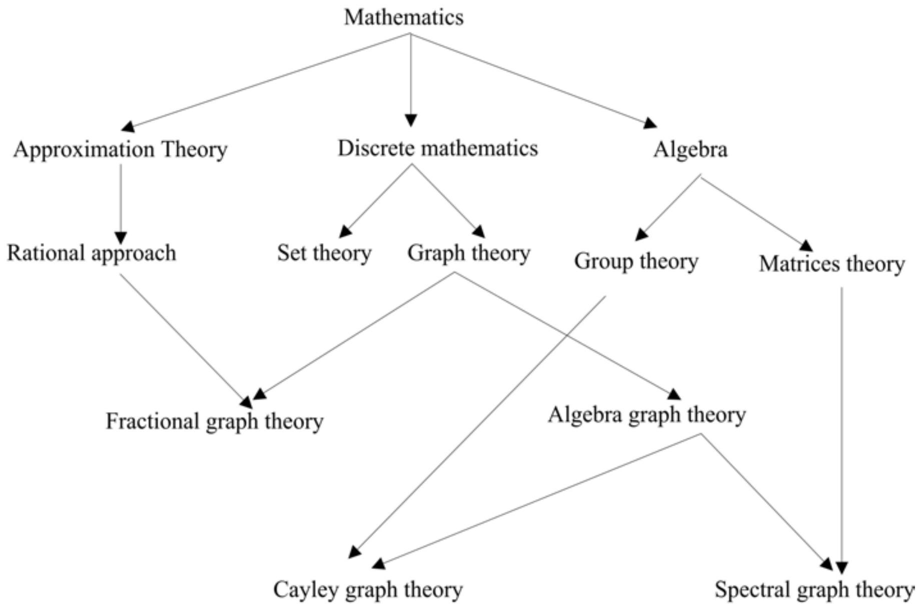

Figure 3: "mathematical" ontology O_3

"GO" ontology was founded in 1998. Worked as an encyclopedia, it concludes a great number of databases, and contains various of the world's major repositories for animal, plant and microbial genomes. Since it contains three main branches: "Molecular function", "Biological process" and "Cellular Component", we set $k = 3$. Ontology learning approaches discussed in Gao and Zhu [35], Gao et al. [45] and [46] are implemented on "GO" ontology, and the $P@N$ accuracy rates by these algorithms are determined. Parts of compared the precision ratio data inferred form these these techniques are presented in Table 3.

	$P@3$ average precision ratio	$P@5$ average precision ratio	$P@10$ average precision ratio	$P@20$ average precision ratio
PMDO Algorithm in our paper	0.5670	0.6985	0.8305	0.9671
Algorithm in Gao and Zhu [35]	0.4988	0.6142	0.7478	0.9101
Algorithm in Gao et al. [45]	0.4792	0.5409	0.6672	0.8449
Algorithm in Gao et al. [46]	0.5649	0.6827	0.8124	0.9371

Table 3. The experiment results of ontology similarity measure

By comparison of experiment results for N=3, 5, 10 or 20 described in Table 3, we see that the partial multi-dividing ontology algorithm depicted in this article is superior to the approaches proposed by Gao and Zhu [35], Gao et al. [45] and [46]. Therefore, our new introduced partial multi-dividing ontology algorithm has more productivity on similarity measuring in biology data applications.

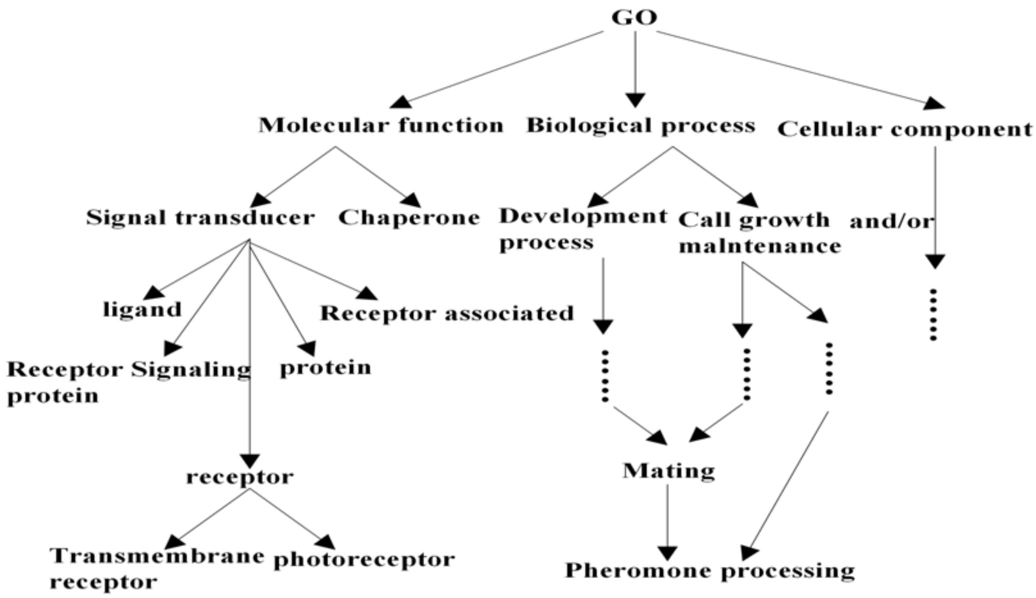

Figure 4: The Structure of "GO" Ontology O_4.

4.4 Ontology mapping experiment on chemical index data

The "Chemical Index" ontologies O_5 and O_6 (the basic structures of O_5 and O_6 are extracted and presented in Figure 5 and Figure 6, respectively) for our last experiment. The figures only present partial vertices of two ontologies, and in fact O_5 and O_6 contain 68 concepts and 46 concepts, respectively. There are many concepts of O_5 and O_6 not displayed in Figure 5 and Figure 6 such as "singly vertex-weighted Wiener number", "multiplicative Wiener index", "terminal Wiener index", "generalized Harary index","second atom bond connectivity index", $\cdots$, "General Co-PI index","fifth atom bond connectivity index", "revised edge Szeged index", "eccentric connectivity polynomial","Shultz polynomial", "second geometric-arithmetic index", "zeroth-order general Randic index", $\cdots$, "Zagreb polynomial", "fifth geometric-arithmetic index", $\cdots$, "sixth Zagreb polynomial", etc.

In chemical graph theory, topological polynomials are defined closely related to topological index. For example, for each vertex (which express an atom) v in molecular graph, the eccentricity of v is denoted by $ec(v) = \max\{u \in V(G)|d(u,v)\}$, then the sixth Zagreb index and the sixth Zagreb polynomial are formulated as

$$Zg_6(G) = \sum_{uv \in E(G)} ec(u)ec(v)$$

and

$$Zg_6(G,x) = \sum_{uv \in E(G)} x^{ec(u)ec(v)},$$

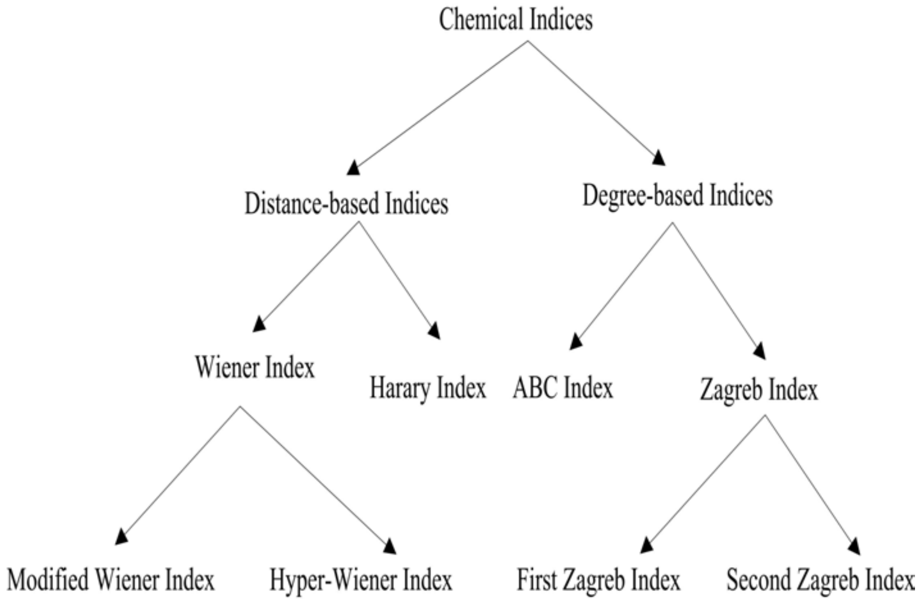

Figure 5: “Chemical Index”ontology O_5

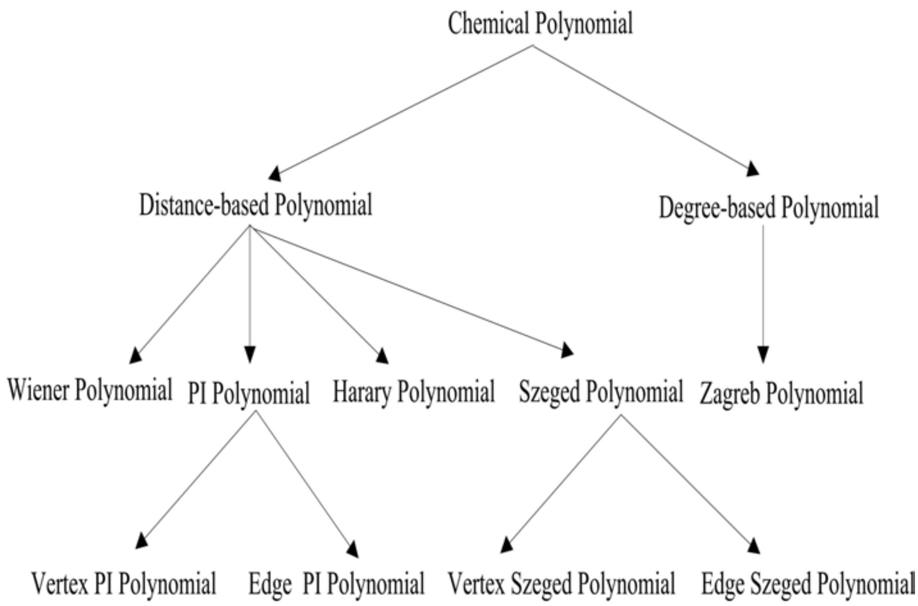

Figure 6: “Chemical Index” ontology O_6

respectively; the fourth Zagreb index and the fourth Zagreb polynomial are formulated as

$$Zg_4(G) = \sum_{uv \in E(G)} (ec(u) + ec(v))$$

and

$$Zg_4(G,x) = \sum_{uv \in E(G)} x^{ec(u)+ec(v)},$$

respectively. From this point of view, it is necessary to build the connection between two ontologies. It will help the researchers in theoretical chemistry and mathematical filed to search the related index and polynomial, and find their potential relationship.

The aim of this experiment is to yield the similarity based ontology mapping between ontologies O_5 and O_6 in terms of ontology learning algorithm raised in this paper. By virtue of ontology structure analysis, we set $k = 3$. Again, the $P@N$ criterion is applied to study the equality of the experiment result. In order to compare the result data, ontology optimization learning frameworks introduced in Gao and Zhu [35], Wu et al. [40], and Gao et al. [48] are implemented on chemical index ontologies, and the precision ratios obtained from these ontology learning algorithms are manifested in Table 4.

	$P@1$ average precision ratio	$P@3$ average precision ratio	$P@5$ average precision ratio
PMDO Algorithm in our paper	0.4386	0.5263	0.7035
Algorithm in Gao and Zhu [35]	0.3247	0.4415	0.5667
Algorithm in Wu et al. [40]	0.4123	0.5058	0.6754
Algorithm in Gao et al. [48]	0.3947	0.4678	0.5807

Table 4. The experiment results of ontology mapping

In terms of compared data depicted in Table 4, we ensure that our partial multi-dividing ontology algorithm is much more efficient than ontology learning tricks described in Gao and Zhu [35], Wu et al. [40] and Gao et al. [48] especially as N becoming large.

5 Conclusion

In recent years, since most ontology structure can be expressed as a tree or analogous to tree, multi-dividing ontology learning becomes a hot topic in ontology research in which all concepts are divided into k parts corresponding to k rates according to the branches of ontology tree, and the rank among these k parts are determined by domain experts. There are several advances both in theoretical and engineering applications in multi-dividing ontology setting, and proved to be in high multidisciplinary special applications.

In our article, we present the partial ontology learning algorithm in multi-dividing setting for both ontology similarity computation and ontology mapping. Several theoretical results on this special learning setting are studied, and the optimal real-valued ontology score function is then yielded. Using this ontology function, each ontology vertex (corresponding to an ontology concept) is mapped into a real number, and the similarity between two vertices v_i and v_j is judged in view of $|f(v_i) - f(v_j)|$. Four experiments are designed to check the practicability and productiveness of our new ontology learning algorithm, and the compared results verify the efficiency of new multi-dividing ontology algorithm.

Conflict of Interests

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

Acknowledgment

We thank the reviewers for their constructive comments in improving the quality of this paper. This work has been partially supported by MINECO grant number MTM2014-51891-P and Fundación Séneca de la Región de Murcia grant number 19219/PI/14 and National Science Foundation of China grant number 11401519.

References

[1] D. Vijayasarathi and P. Sankar, Encoding of Fundamental Chemical Entities of Organic Reactivity Interest using chemical ontology and XML, Journal of Molecular Graphics & Modelling, vol. 61, pp. 30-43, 2015.

[2] M. P. Banchetti-Robino, Van Helmont's hybrid ontology and its influence on the chemical interpretation of spirit and ferment, Foundations of Chemistry, vol. 18, no. 2, pp. 103-112, 2016.

[3] D. Baorto, L. Li, and J. J. Cimino, Practical experience with the maintenance and auditing of a large medical ontology, Journal of Biomedical Informatics, vol. 42, no. 3, pp. 494-503, 2009.

[4] S. Sarntivijai, S. Zhang, D. G. Jagannathan, S. Zaman, K. K. Burkhart, G. S. Omenn, Y. Q. He, B. D. Athey, and D. R. Abernethy, Linking medDRA (R)-coded clinical phenotypes to biological mechanisms by the ontology of adverse events: a pilot study on tyrosine kinase inhibitors, Drug Safety, vol. 39, no. 7, pp. 697-707, 2016.

[5] S. Kohler, S. C. Doelken, C. J. Mungall, S. Bauer, H. V. Firth, I. Bailleul-Forestier, G. C. M. Black, D. L. Brown, M. Brudno,J. Campbell, et al., The Human Phenotype Ontology project: linking molecular biology and disease through phenotype data, Nucleic Acids Research, vol. 42, no. D1, pp. D966-D974, 2014.

[6] D. M. Levine, N. K. Dutta, J. Eckels, C. Scanga, C. Stein, S. Mehra, D. Kaushal, P. C. Karakousis, and H. Salamon, A tuberculosis ontology for host systems biology, Tuberculosis, vol. 95, no. 5, pp. 570-574, 2015.

[7] U. S. Vishnu, J. Sankarasubramanian, P. Gunasekaran, J. Sridhar, and J. Rajendhran, Omics of brucella: species-specific sRNA-mediated gene ontology regulatory networks identified by

computational biology, OMICS-A Journal of Integrative Biology, vol. 20, no. 6, pp. 375-385, 2016.

[8] X. Aime and j. Charlet, Social psychology insights into ontology engineering, Future Generation Computer Systems-The International Journal of Escience, vol. 54, pp. 348-351, 2016.

[9] O. E. Petrunia, Dimensionally ontology V. Frankl as the conceptual basis for interdisciplinary synthesis of Biomedicine, psychology and computing, Biomedicine Radioengineering, vol. 4, pp. 58-59, 2015.

[10] G. Demartini, I. Enchev, J. Gapany, and P. Cudre-Mauroux, The bowlogna ontology: fostering open curricula and agile knowledge bases for Europe's higher education landscape, Semantic Web, vol. 4, no. 1, pp. 53-63, 2013.

[11] M. Kruger-Ross, Raising the question of being in education by way of heidegger's phenomenological ontology, Indo-Pacific Journal of Phenomenology, vol. 15, no. 2, pp. 1-12, 2015.

[12] N. M. Ochara, Towards a regional ontology of management education in africa: a complexity leadership theory perspective, Acta Commercii, vol. 17, no. 1, pp. 1-8, 2017.

[13] L. Vaccari, P. Shvaiko, J. Pane, P. Besana, and M. Marchese, An evaluation of ontology matching in geo-service applications, Geoinformatica, vol. 16, no. 1, pp. 31-66, 2012.

[14] F. Delgado, M. M. Martinez-Gonzalez, and J. Finat, An evaluation of ontology matching techniques on geospatial ontologies, International Journal of Geographical Information Science, vol. 27, no. 12, pp. 2279-2301, 2013.

[15] M. Tahmoorespur, S. Jelokhani-Niaraki, Z. Minuchehr, and M. R. Nassiri, An ontology-based GIS for genomic data management of rumen microbes, Genomics & Informatics, vol. 13, no. 1, pp. 7-14, 2015.

[16] V. Bertaud-Gounot, R. Duvauferrier, and A. Burgun, Ontology and medical diagnosis, Informatics for Health & Social Care, vol. 37, no. 2, pp. 51-61, 2012.

[17] J. P. Lousado, V. Rodrigues, N. Costa, R. Carvalho, C. Albuquerque, and V. Franco, Medical ontology for treatment of clinical data from children and youth, Atencion Primaria, vol. 45, no. 2, pp. 223-227, 2013.

[18] H. H. Kim, S. Y. Lee, S. Y. Baik, and J. H. Kim, MELLO: medical lifelog ontology for data terms from self-tracking and lifelog devices, International Journal of Medical Informatics, vol. 84, no. 12, pp. 1099-1110, 2015.

[19] E. Cuccia, Aquinas's ontology of the material world. change: hylomorphism and material objects, Scripta Mediaevalia, vol. 9, no. 1, pp. 82-84, 2016.

[20] E. Ghibaudi and L. Cerruti, Chemical substance, material, product, goods, waste: a changing ontology, Foundations of Chemistry, vol. 19, no. 2, pp. 97-123, 2017.

[21] D. M. Bowden, E. Song, J. Kosheleva, and M. F. Dubach, NeuroNames: an ontology for the brainInfo portal to neuroscience on the web, Neuroinformatics, vol. 10, no. 1, pp. 97-114, 2012.

[22] R. Fumagalli, Choice models and realistic ontologies: three challenges to neuro-psychological modellers, European Journal for Philosophy of Science, vol. 6, no. 1, pp. 145-164, 2016.

[23] M. A. Rodriguez and M. J. Egenhofer, Determining semantic similarity among entity classes from different ontologies, IEEE Transactions on Knowledge and Data Engineering, vol. 15, no. 2, pp. 442-456, 2003.

[24] O. Steichen, C. D. Le Bozec, M. Thieu, E. Zapletal, and M. C. Jaulent, Computation of semantic similarity within an ontology of breast pathology to assist inter-observer consensus, Computers in Biology and Medicine, vol. 36, no. 7-8, pp. 768-788, 2006.

[25] H. Al-Mubaid and H. A. Nguyen, Measuring semantic similarity between biomedical concepts within multiple ontologies, IEEE Transactions on Systems and Cybernetics Part C-Applications and Reviews, vol. 39, no. 4, pp. 389-398, 2009.

[26] S. Kohler, M. H. Schulz, P. Krawitz, S. Bauer, S. Dolken, C. E. Ott, C. Mundlos, D. Horn, S. Mundlos, and P. N. Robinson, Clinical diagnostics in human genetics with semantic similarity searches in ontologies, American Journal of Human Genetics, vol. 85, no. 4, pp. 457-464, 2009.

[27] M. Batet, D. Sanchez, and A. Valls, An ontology-based measure to compute semantic similarity in biomedicine, Journal of Biomedical Informatics, vol. 44, no. 1, pp. 118-125, 2011.

[28] E. Albacete, J. Calle, E. Castro, and D. Cuadra, Semantic similarity measures applied to an ontology for human-like interaction, Journal of Artifical Intelligence Research, vol. 44, pp. 397-421, 2012.

[29] K. Taha, Determining the semantic similarities among gene ontology Terms, IEEE Journal of Biomedical and Health Informatics, vol. 17, no. 3, pp. 512-525, 2013.

[30] M. A. H. Taieb, M. Ben Aouicha, and A. Ben Hamadou, Ontology-based approach for measuring semantic similarity, Engineering Applications of Artifical Intelligence, vol. 36, pp. 238-261, 2014.

[31] G. K. Mazandu, E. R. Chimusa, M. Mbiyavanga, and N. J. Mulder, A-DaGO-Fun: an adaptable gene ontology semantic similarity-based functional analysis tool, Bioinformatics, vol. 32, no. 3, pp. 477-479, 2016.

[32] J. J. Lastra-Diaz, A. Garcia-Serrano, M. Batet, M. Fernandez, and F. Chirigati, HESML: A scalable ontology-based semantic similarity measures library with a set of reproducible experiments and a replication dataset, Information Systems, vol. 66, pp. 97-118, 2017.

[33] W. Gao, Y. Gao, and Y. G. Zhang, Strong and weak stability of k-partite ranking algorithms, Information-An International Interdisciplinary Journal, vol. 15, no. 11A, pp. 4585-4590, 2012.

[34] W. Gao and T. W. Xu, Stability analysis of learning algorithms for ontology similarity computation, Abstract and Applied Analysis, 2013, Article ID 174802, 9 pages, http://dx.doi.org/10.1155/2013/174802.

[35] W. Gao and L. L. Zhu, Gradient learning algorithms for ontology computing, Computational Intelligence and Neuroscience, Volume 2014, Article ID 438291, 12 pages, http://dx.doi.org/10.1155/2014/438291.

[36] W. Gao, L. L. Zhu, and K. Y. Wang, Ontology sparse vector learning algorithm for ontology similarity measuring and ontology mapping via ADAL technology, International Journal of Bifurcation and Chaos, vol. 25, no. 14, 1540034 (12 pages), DOI: 10.1142/S0218127415400349, 2015.

[37] W. Gao and M. R. Farahani, Generalization bounds and uniform bounds for multi-dividing ontology algorithms with convex ontology loss function, The Computer Journal, DOI:10.1093/comjnl/bxw107.

[38] L. Maire, Large solutions for cooperative logistic systems: existence and uniqueness in star-shaped domains, Applied Mathematics and Nonlinear Sciences, vol. 2, no. 2, pp. 249-258, 2017.

[39] E. Ibrahim, S. Y. Atwa, and M. K. Ammar, The effect of two-temperature on thermoelastic medium with diffusion due to three phase-lag model, Applied Mathematics and Nonlinear Sciences, vol. 2, no. 2, pp. 259-270, 2017.

[40] J. Z. Wu, X. Yu, and W. Gao, Disequilibrium multi-dividing ontology learning algorithm, Communications in Statistics-Theory and Methods, http://dx.doi.org/10.1080/03610926.2016.1197254.

[41] F. Cucker and D. X. Zhou, Learning Theory: an Approximation Theory Viewpoint. Cambridge: Cambridge University Press, 2007.

[42] D. X. Zhou, On grouping effect of elastic net, Statistics & Probability Letters, vol. 83, no. 9, pp. 2108-2112, 2013.

[43] S. Smale and D. X. Zhou, Geometry on probability spaces, Constructive Approximation, vol. 30, no. 3, pp. 311-323, 2009.

[44] N. Craswell and D. Hawking, Overview of the TREC 2003 Web Track, In proceedings of the Twelfth Text Retrieval Conference, Gaithersburg, Maryland, NIST Special Publication, 2003, pp. 78-92.

[45] W. Gao, M. R. Farahani, A. Aslam, and S. Hosamani, Distance learning techniques for ontology similarity measuring and ontology mapping, Cluster Computing-The Journal of Networks, Software Tools and Applications, vol. 20, no. 2, pp. 959-968, 2017.

[46] W. Gao, A. Q. Baig, H. Ali, W. Sajjad, and M. R. Farahani, Margin based ontology sparse vector learning algorithm and applied in biology science, Saudi Journal of Biological Sciences, vol. 24, no. 1, pp. 132-138, 2017.

[47] W. Gao, L. L. Zhu, and K. Y. Wang, Ranking based ontology scheming using eigenpair computation, Journal of Intelligent & Fuzzy Systems, vol. 31, no. 4, pp. 2411-2419, 2016.

[48] W. Gao, Y. Guo, and K. Y. Wang, Ontology algorithm using singular value decomposition and applied in multidisciplinary, Cluster Computing-The Journal of Networks Software Tools and Applications, vol. 19, no. 4, pp. 2201-2210, 2016.