J Intell Inf Syst (2012) 39:749–761
DOI 10.1007/s10844-012-0210-y

Extracting semantic relations to enrich domain ontologies

Minxin Shen · Duen-Ren Liu · Yu-Siang Huang

Received: 17 October 2011 / Revised: 17 May 2012 / Accepted: 17 May 2012 /
Published online: 9 June 2012

Abstract Domain ontologies facilitate the organization, sharing and reuse of domain knowledge, and enable various vertical domain applications to operate successfully. Most methods for automatically constructing ontologies focus on taxonomic relations, such as is-kind-of and is-part-of relations. However, much of the domain-specific semantics is ignored. This work proposes a semi-unsupervised approach for extracting semantic relations from domain-specific text documents. The approach effectively utilizes text mining and existing taxonomic relations in domain ontologies to discover candidate keywords that can represent semantic relations. A preliminary experiment on the natural science domain (Taiwan K9 education) indicates that the proposed method yields valuable recommendations. This work enriches domain ontologies by adding distilled semantics.

Keywords Ontology learning · Relation extraction · Semantic relation · Text mining

1 Introduction

Ontologies capture domain knowledge using structured and relational representations. This effective organization of knowledge has led to the wide application of onto-

M. Shen (✉)
Industrial Technology Research Institute, 195, Sec. 4, Chung Hsing Rd., Chutung,
Hsinchu 31040, Taiwan, Republic of China
e-mail: mshen.tw@gmail.com

M. Shen · D.-R. Liu · Y.-S. Huang
Institute of Information Management, National Chiao Tung University,
1001 University Road, Hsinchu 300, Taiwan, Republic of China

D.-R. Liu
e-mail: dliu@mail.nctu.edu.tw

Y.-S. Huang
e-mail: ohoh.huang@gmail.com

logies to semantic computing applications, including semantic webs, expert systems, vertical searches, and others. Although such applications of ontologies are valuable, the construction of a domain ontology is labor-intensive and time-consuming. Rapid increases in knowledge further complicate this problem (Hepp 2007).

Various ontology-learning methods have been developed to automate the construction process (Maedche et al. 2002; Cimiano et al. 2005; Sumida et al. 2006). Most focus on taxonomic structures because hierarchical classification is extensively adopted in representing knowledge. The higher entities in ontologies are more general, while the lower entities are more particular. However, such taxonomic relations (is-kind-of and is-part-of relations) only partially capture the relevant knowledge. Semantic relations among ontological entities provide more domain-specific associations. For example, two sibling concepts, satellite and planet, may involve a semantic relation "revolution" in a K9 natural science ontology. Therefore, this work seeks to develop a method for extracting semantic ontological relations from unstructured text documents.

The automatic extraction of semantic relations has been investigated extensively. The process involves two tasks: detecting related concepts and labeling their relationships. Usually, verbs that connect two ontological concepts are regarded as candidate labels for semantic relations. Hence, most studies have focused on the co-occurrence of concept pairs and verbs (Kavalec et al. 2004; Maedche and Staab 2000; Schutz and Buitelaar 2005; Villaverde et al. 2009; Weichselbrauna et al. 2010). Hasegawa et al. (2004) further considered the context of verbs to improve extraction. However, these studies do not exploit taxonomic relations to discover semantic relations. Vertical domains usually include obtainable taxonomy knowledge. Additionally, domain ontologies are often constructed by extending domain semantics to existing general ontologies, such as the Suggested Upper Merged Ontology (SUMO) (SUMO 2011).

This work develops a novel text mining approach for discovering potential semantic relations to extend existing taxonomic hierarchies. The proposed approach first adopts the Chi-Square test of independence to determine whether two concepts are correlated. Sentences that contain correlated concepts are collected for subsequent relation extraction. Then, hierarchical clustering is applied to group sentences with similar contexts. Synonym dictionaries are employed to improve clustering results. Finally, a keyword weighting scheme that effectively utilizes taxonomic relations and semantic context clusters is proposed to rank the extracted candidate labels. A domain ontology for K9 natural science is adopted to illustrate and validate the proposed approach. Experimental results indicate the effectiveness of the proposed approach in discovering semantically related concepts.

The rest of this paper is organized as follows. Section 2 reviews the literature on extracting semantic relations. Section 3 presents the proposed framework for discovering semantic relations. Section 4 presents the experimental results concerning the K9 natural science domain. Finally, Section 5 draws conclusions.

2 Review of semantic relation extraction

Methods for extracting semantic relations can be classified as supervised and unsupervised. Supervised approaches depend on predefined relations and aim to identify which types of relations hold between a pair of entities. Various machine learning

algorithms have been applied for predicting relations, including Support Vector Machine (Zelenko et al. 2003; Che etal. 2005; Culotta and Sorensen 2004), Conditional Random Fields (Culotta et al. 2006), and Maximum Entropy (Zhang et al. 2006). However, supervised methods require predefined relations and annotated training data. Ontology construction seeks to find previously unknown relations, rather than known relations, for which supervised approaches are therefore ineffective. Unsupervised approaches do not require training data and so are useful for constructing ontologies.

Several recent studies have explored unsupervised approaches. Maedche et al. (Maedche et al. 2002; Maedche and Staab 2000) applied association mining to find relations between concepts. They consulted experts to specify labels of those relations. Kavalec et al. (Kavalec et al. 2004), Villaverde et al. (Villaverde et al. 2009), Weichselbrauna et al. (Weichselbrauna et al. 2010), and Serra and Girardi (Serra and Girardi 2011) further discovered associated concept pairs and verbs, and then employed the verbs to label semantic relations. J. Punuru and J. Chen (Punuru and Chen 2011) utilized the distributions of co-occurring concepts and verbs as significance measures for identifying verbs as semantic labels. Moreover, (Weichselbraun et al. 2009) using verb vectors and vector centroids to handle the circumstance that multiple verbs co-occur with candidate concept pairs. However, most of these works consider the scope of verbs and do not utilize other context words near the verbs and existing taxonomic relations.

Hasegawa et al. (2004) further considered context of verbs, meaning verbs and their co-occurring non-verbs. The contexts of concept pairs were clustered, and then representative words of the clusters were recommended as relation labels. However, Hasegawa et al. considered only seven types of entities defined in ACE (ACE 2005). Moreover, they did not utilize clustering results to analyze the discriminative power of candidate words. Additionally, the studies cited above do not utilize taxonomic relations to discover semantic relations. A novel keyword weighting scheme that jointly considers context words (semantic context) and taxonomic relations (structural context), is proposed in this work to increase the accuracy of recommendation. Generally, association rule mining and pointwise mutual information (PMI) are two main approaches cited above for quantifying the association strength between two ontology concepts. While the former represents one-way dependency, the latter denotes two-way associations, and is thus adopted as the baseline in this paper. The experiment results in Section 4 show that the proposed keyword weighting scheme for verb selection achieves improvement by utilizing the clustering-based discrimination measures.

Sánchez and Moreno (2008) proposed a PMI-based incremental learning approach. Verbs frequently and directly connected to given domain concepts are first discovered, and then new concepts related to those verbs are identified and inserted to the original ontology. The cascaded expansion step is repeated until stop criteria are met. In contrast, our work discovers unknown labels between given concept pairs from a closed-domain corpus. Moreover, the application of semantic context for discriminative comparisons of candidate verbs (TFICF stated in Section 3.4.1) and the utilization of structural context (semantic relations of child concepts stated in Section 3.4.2) are our main contributions. For a given concept pair, as the evaluation shows in Section 4, exploring its direct context (co-occurring words in the same sentence) and indirect context (the co-occurring words that also appears with child concepts) improves the accuracy of relation labeling.

3 Extracting ontological semantic relation

Figure 1 presents the proposed framework for relation extraction. *Domain ontology* in the figure is the target ontology that comprises concept entities and corresponding taxonomic relations. Additionally, the *Domain-specific corpus* is the source of extracted semantic relations. For example, in the experiment described in Section 4, the corpus may include K9 natural science-related lectures, reports, and text books. The goal of this work is to extend domain ontology by adding semantic relations extracted from the domain-specific corpus. The generated ontology is an abstraction of the domain corpus. That is, all relations in the ontology are supported by the corpus.

The extraction process has four steps: *Sentence selection* first extracts sentences from the domain corpus, and then segments them into ontological concepts or words with part-of-speech (POS) tags. *Relation detection* further filters out sentences that do not contain statistically correlated concept pairs. Then, *context clustering* groups retain sentences that contain the concept pairs. The generated clusters contain candidate words for describing relations between pairs of concepts. These words are weighed and ranked in the *keyword ranking* step. Finally, the ranked words are recommended as relation labels, which ontology engineers then validate. The details of the above four steps are as follows.

3.1 Sentence selection

Sentences are the context from which relation labels are extracted. The sentence selection step collects sentences that contain at least one pair of ontological concepts

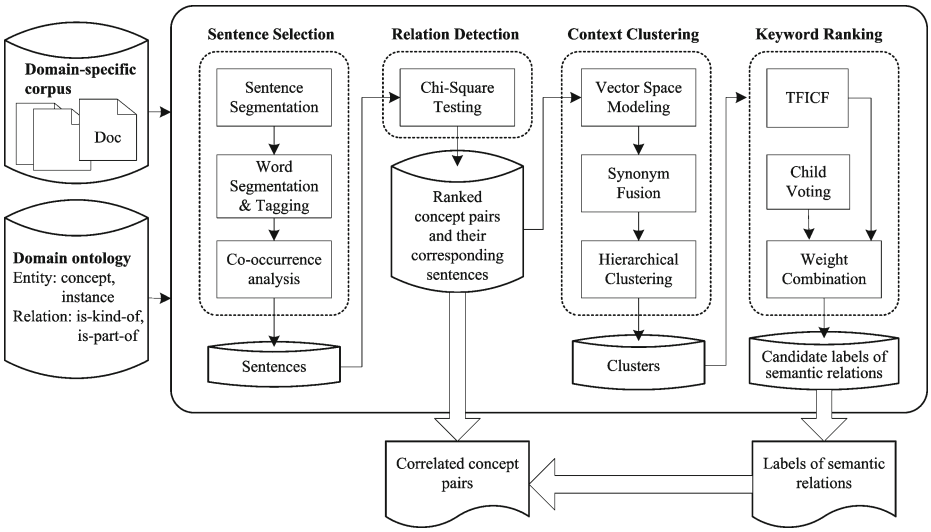

Fig. 1 Framework for extracting ontological semantic relations

and a verb. First, as shown in Fig. 1, all documents in the domain-related corpus are segmented into sentences using the punctuation marks, full stop, exclamation mark and question mark.

Then, these sentences are segmented into ontology concepts or words using a hybrid method. The maximum matching algorithm is employed to segment ontology entities (concept and instances), which are then labeled with the corresponding ontology concepts. For example, "Earth" is segmented out and tagged as "PLANET". For Chinese ontology learning, the Hidden Markov Model (HMM) algorithm, trained using the SINICA balanced corpus (Chen and Huang 1997), is employed to segment out Chinese words with part-of-speech (POS) tags from the other sentence fragments. For English ontology learning, words are segmented by blanks, and then tagged by a HMM POS tagger. This work argues that the labels of concept relations appear as verbs in sentences. Accordingly, co-occurrence analysis is applied to collect sentences that contain at least two ontology entities and at least one verb.

Table 1 exemplifies sentence selection. Following the sentence and word segmentation, words are annotated with part-of-speech tags (Verb or Noun) or ontology concepts. Finally, both s_1 and s_2 contain at least two ontology concepts and one verb, and are thus preserved. Notably, nouns, as well as verbs, are also retained as *context words*.

3.2 Relation detection

To prevent blind relation extraction from all sentences in the previous step, relation detection determines whether two concepts are related to each other and can thus be

Table 1 Example of sentence selection

Step	Results
Sentence segmentation	s_1: The distance between the Earth and the Sun is suitable to cause water to exist in liquid form, and thus promotes the survival of living things. s_2: Teachers provide instances, and students judge whether the reason is weathering or erosion. s_3: Arterial wall is thicker and very elastic.
Word segmentation and tagging	s_1: The distance/N between the Earth/PLANET and the Sun/STAR is/V suitable to cause/V water/N to exist in liquid form/N, and thus promotes/V the survival/N of living things/N. s_2: Teachers/TEACHER provide/V instances/N, and students/STUDENT judge/V whether the reason/N is/V weathering/N or erosion/N. s_3: Arterial wall/N is/V thicker and very elastic.
Co-occurrence analysis	s_1: The distance/N between the Earth/PLANET and the Sun/STAR is/V suitable to cause/V water/N to exist in liquid form/N, and thus promotes/V the survival/N of living things/N. s_2: Teachers/TEACHER provide/V instances/N, and students/STUDENT judge/V whether the reason/N is/V weathering/N or erosion/N. ~~s_3~~: Arterial wall/N is/V thicker and very elastic.

formulated as a Chi-Square test of independence. Equation (1) yields the Chi-Square statistic.

$$\chi^2 = \sum_{i,j} \frac{(O_{ij} - E_{ij})^2}{E_{ij}}, \tag{1}$$

where O_{ij}/E_{ij} denotes the observed/expected frequency of ontology concepts c_i and c_j.

As well as indicating whether two concepts are related, Chi-Square values also represent the degree of relatedness and are used to rank concept pairs. Larger Chi-Square values indicate stronger relatedness. Ontology engineers can annotate top-ranked concept pairs. In summary, the output of relation detection ranks correlated concept pairs and their corresponding sentences. Although some previous works have adopted pointwise mutual information as a co-occurrence indicator, Manning and Schütze (Manning and Schütze 1999) concluded that PMI is not a good measure of association between elements.

3.3 Context clustering

The sentences that provide the context of each discovered concept pair are transformed to feature vectors for clustering. The generated context clusters may represent semantic relations of the concept pair for further extraction, as stated in Section 3.4. Figure 1 shows the process of context clustering, the details of which are described below.

3.3.1 Context vector

The sentences that correspond to a concept pair are characterized by the union of context words, and transformed to context vectors. That is, each sentence is represented as a feature vector, as defined in Eq. (3).

For a given concept pair *cp*, sentence

$$s_i = (f_1, f_2, \ldots, f_j), \tag{2}$$

where $f_j = \begin{cases} 1, & \text{if word } w_j \text{ occures in sentence } s_i \\ 0, & \text{otherwise} \end{cases}$, $j = 1,2,\ldots,|W_{cp}|$,

$$s_i \in S_{cp},$$

where S_{cp} is the set of sentences that contain the concept pair *cp*, and $|S_{cp}|$ is the size of S_{cp},

$$w_j \in W_{cp} \text{ and } W_{cp} = \bigcup_i W_i, \tag{3}$$

where W_i is the set of context words in s_i, and $|W_{cp}|$ is the size of W_{cp}.

For example, the context words in the two sentences s_1 and s_2 in Table 2 are "instruct utilize, broadcast, process, present" and "use, teach, process, present", respectively. The union of these context words yields the features, i.e., (instruct utilize, broadcast, process, present, use, teach). Accordingly, the vector representations of these two sentences are (1, 1, 1, 1, 1, 0, 0) and (0, 0, 0, 1, 1, 1, 1), respectively.

Table 2 Example of context vector

Concept Pair: Teacher – Student							
s_1: Instruct utilize broadcast process present							
s_2: Use teach process present							
Before synonym fusion							
Features:	Instruct	Utilize	Broadcast	Process	Present	Use	Teach
s_1:	1	1	1	1	1	0	0
s_2:	0	0	0	1	1	1	1
After synonym fusion							
Features:	Instruct Teach	Utilize Use	Broadcast	Process	Present		
s_1:	1	1	1	1	1		
s_2:	1	1	0	1	1		

3.3.2 Synonym fusion

The words that are common to two sentences are too few because a sentence is far shorter than a document, so feature vectors are sparse. Moreover, the lack of repetition of the same words exacerbates the problem of sparseness. To address this issue, the synonym dictionaries, English WordNet and English-Chinese Bilingual WordNet (Academia-Sinica 2005), are adopted to merge synonyms to support the calculation of the semantic similarity among sentences.

Table 2 shows an example in which the two sentences indicate that the semantic relation "instruct" holds for the concept pair (teacher, student). The feature vector of these two sentences has seven dimensions if all of the words are considered. Table 3 reveals that "instruct" and "teach" belong to the same *synset* (synonym set), and so can be merged as shown in the lower part of Table 2. Synonym fusion reduces the number of feature dimensions to five. Moreover, Jaccard similarity between the two sentences is increased from 0.286 to 0.8. Notably, word sense disambiguation approaches are not adopted for polysemy resolution. Context words are merged if they appear in the same synset, since their co-occurring pair of concepts provides adequate semantic cues.

3.3.3 Sentence clustering

Sentence vectors of a particular concept pair are then clustered. Each cluster may represent a semantic relation of the concept pair. The hierarchical agglomerative clustering algorithm (HAC) is employed for context clustering. Distance measure is based on Jaccard similarity. Moreover, cluster-merging follows the complete linkage criterion; it stops when the inter-cluster distance is greater than a given threshold. Please refer to (Manning et al. 2008) for the details of the clustering algorithm.

Table 3 Example of synonym sets from the English WordNet synonym dictionary

Synonym set
teach, learn, **instruct**
use, **utilize**, utilise, apply, employ

3.4 Keyword ranking

The final step, as shown in Fig. 1, is to recommend appropriate keywords for the labeling of the semantic relations of concept pairs. These candidate keywords are derived from the context clusters, as discussed in Section 3.3 Two weighting schemes, *Term frequency and inverse cluster frequency* (TFICF), and *child voting* (CV), are developed to calculate the keyword weights. Finally, candidate keywords are ranked using linear combined weight scores to label the semantic relations.

3.4.1 Term frequency and inverse cluster frequency (TFICF)

The weights of the words in the clusters are estimated in terms of *term frequency* (TF) and *inverse cluster frequency* (ICF). TFICF is very similar to TFIDF (term frequency - inverse document frequency). TF is the occurrence of a word in a cluster. A word is important in a cluster if the word appears frequently in that cluster. ICF is the number of clusters in which the word occurs. If a word appears in numerous clusters, the discriminative power of the word is low. Accordingly, TFICF estimates both the relative importance of words in clusters and its discrimination among clusters. Formally, TFICF is calculated using Eq. (4), and normalized to [0..1] using Eq. (5).

TFICF of word w_i in cluster c_j is

$$w_{ij} = \left(\frac{TF_{ij}}{\max_i \left(TF_{ij} \right)} \right) \times \log \frac{N}{n_i}, \tag{4}$$

where TF_{ij} is the frequency of word w_i in cluster c_j, N is the number of clusters, and n_i is the number of clusters that contain word w_i.

$$w_{ij}^{tficf} = \frac{w_{ij}}{\max_j \left(w_{ij} \right)} \tag{5}$$

Conventional approaches including PMI and association rules tend to choose the verbs that repeatedly appear with target concept pairs. Although cluster analysis is applied to group similar contexts (Kavalec et al. 2004), frequently mentioned verbs still have a higher possibility of being selected than rare ones. However, popular verbs are sometimes meaningless because of generality. Therefore, this work further proposes ICF as a compromise factor to assign more weights to discriminative terms among context clusters.

3.4.2 Child voting

Taxonomic relations are also utilized to estimate keyword weights. This idea is motivated by inheritance. In a taxonomic hierarchy, child concepts inherit the attributes of a parent concept. Thus, semantic relations may also be inherited. For example, to determine the semantic relation between concept pair (E, F) in Fig. 2, the contexts of (E, L), (E, M), (F, J) and (F, K) are also considered. However, previous studies focus merely on co-occurrence analysis of verbs and target concept pairs, i.e., (E, F) (Kavalec et al. 2004; Hasegawa et al. 2004; Villaverde et al. 2009; Serra and Girardi 2011). Abstract and concrete concepts may be used interchangeably for reinterpreting the meaning of concepts, resulting in data sparseness. Exploring the

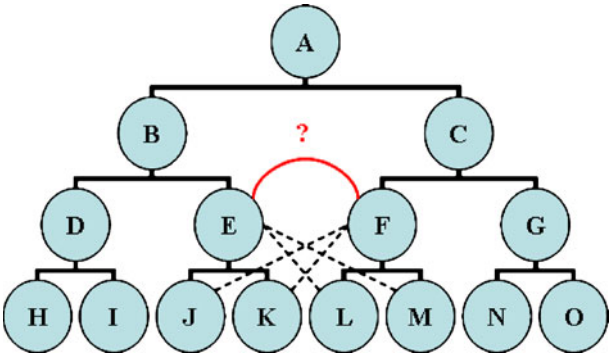

Fig. 2 Example of child voting scheme

contexts of (E, L), (E, M), (F, J), and (F, K) may provide indirect clues for finding appropriate relation labels.

A child voting scheme, given by Algorithm 1 below, is developed to refine the weights of context words. The main idea is to generate candidate words from the contexts of pairs of child concepts. These context words are candidate labels. Then, Eq. (6) is employed to estimate the child voting scores of the words.

$$w_i^{vote} = \frac{votei}{\max\limits_i (vote_i)}, \tag{6}$$

where $vote_i$ is the number of votes of word w_i.

Table 4 presents an example of the application of the algorithm. Candidate words are derived after context clustering for all pairs of child concepts of concepts E or F in Fig. 2. Then, the votes for candidate words are calculated and normalized using Eq. (6). For example, w_2 receives the most votes (4) and the voting score of w_2 and w_3 are 1 (3/3) and 0.67 (2/3), respectively.

Algorithm 1 (child voting scheme)

Input: concept pair $cp = (c_i, c_j)$
Output: voting scores of candidate words for labeling cp
begin
  for each concept c_i in c_p
    for each concept is c_{is} in subclass of c_i
      perform context clustering for (c_{is}, c_j), where $c_j \in cp$, $j \neq i$
      generate candidate words from context clusters
      calculate child voting scores of candidate words using Eq. (6)
  return candidate words and corresponding voting scores;
end

Table 4 Example of child voting results

Concept pair	Candidate words
(E,L)	w_1, w_2, w_3, w_4, w_5
(E,M)	$w_2, w_4, w_5, w_6, w_7, w_8$
(F,J)	$w_2, w_3, w_8, w_{10}, w_{11}, w_{12}$
(F,K)	w_4, w_8, w_9, w_{13}

3.4.3 Keyword weights

Word weights are finally determined as linear combinations of normalized TF-ICF and the child votes using Eq. (7). Candidate labels for semantic relations are ranked by weight.

$$w_{ij}^{weight} = \alpha \cdot w_{ij}^{tficf} + (1-\alpha) \cdot w_i^{vote},\ 0 \le \alpha \le 1 \tag{7}$$

4 Experiment on K9 natural science domain

The experimental domain is K9 natural science in Taiwan. Textbooks and teachers' manuals, provided by a leading publisher (Kang-Hsuan Educational Publishing Group, http://www.knsh.com.tw), are used as the experimental corpus. This corpus comprised 255 documents. An ontology is constructed manually from this corpus. The ontology consists of 272 concepts, 1336 instances, and taxonomic relations.

This experiment discovers previously unknown semantic relations, and does not have predefined answers. Accordingly, a natural science expert is consulted to evaluate the extraction results. For a concept pair, corresponding sentences extracted from the corpus are presented and the expert is consulted to select a verb as the correct label. Thus, the performance of semantic relation extraction is evaluated in terms of *accuracy*, as defined in Eq. (8). Its denominator is the number of concept pairs for evaluation and its numerator is the number of correctly predicted labels.

$$\text{accuracy} = \frac{\#\left(\text{correct predicted labels of concept pairs}\right)}{\#\left(\text{concep pairs}\right)} \tag{8}$$

The Chi-square based relation detection suggests that 706 concept pairs are correlated at the 95% confidence level. The top 100 concept pairs when ranked by Chi-Square value are the evaluation targets. The labels of concept pairs recommended by TFICF and Child voting are evaluated against the manually labeled results. In addition, the labels of these 100 concept pairs recommended by the pointwise mutual information approach, as defined in Eq. (9), are also evaluated for comparison.

$$PMI\left(\left(c_1 \wedge c_2\right), w\right) = \frac{P\left(\left(c_1 \wedge c_2\right)|w\right)}{P\left(\left(c_1\right)|w\right) P\left(\left(c_2\right)|w\right)} \tag{9}$$

Figure 3 plots the effects of the different parts of the proposed method; the horizontal axis represents the adopted approaches and the vertical axis represents the corresponding accuracy. Herein, the similarity threshold to stop clustering is 0.25, and TFICF+CV is obtained using Eq. (7), with α as 0.7. The leftmost two bars show that PMI and TF produced similar results, perhaps because PMI is applied to the 100 pairs derived by chi-square analysis. Accordingly, both PMI and TF only employ word frequencies. Contrarily, TFICF utilizes results of context clustering to reduce the impact of frequent but minor verbs (e.g., use and utilize), and thus discovers informative verbs for labeling.

Besides, the child voting scheme properly leverages taxonomic context and thus alleviates the data sparseness problem. That is, for a concept pair, one of them may co-occur with the other's children under the circumstance such as paraphrasing or explaining the meaning of child concepts. Thus, child voting recovers the hidden term frequencies. However, as the rightmost bar shown in Fig. 3, the improvement

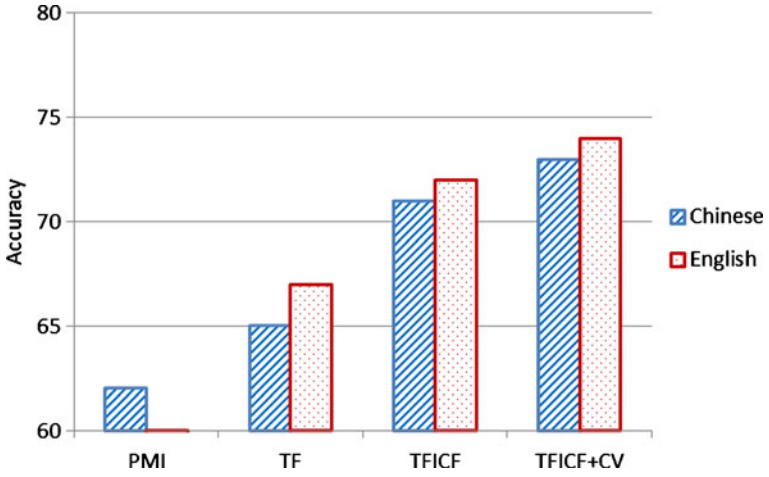

Fig. 3 Evaluation results

of accuracy is not very significant due to the fact that child voting is only applied to eight concept pairs in this experiment. Figure 4 shows the impact of child voting with different weighs (α). Initially, CV has no impact while $\alpha \geq 0.8$. For example, for the concept pair (Star, Season), TFICF+CV suggests a wrong label "identify" mainly according to ICF. Nevertheless, when $0.7 \geq \alpha \geq 0.5$, TFICF+CV advises "appear" as the label, which is consistent with the manual annotation. Note that child voting also has a negative effect when $\alpha \leq 0.4$ due to overweighting of term frequencies. For example, instead of choosing the correct verb "rotate" to label the pair (Star, Planet), TFICF+CV proposes "move" as the semantic label, since "move" appears repeatedly among child instances. In sum, the value of weight α is recommend among $0.6 \sim 0.8$ since TFICF is more general and directly measures the discriminative power of candidate labels, while child voting indirectly reflects term frequencies.

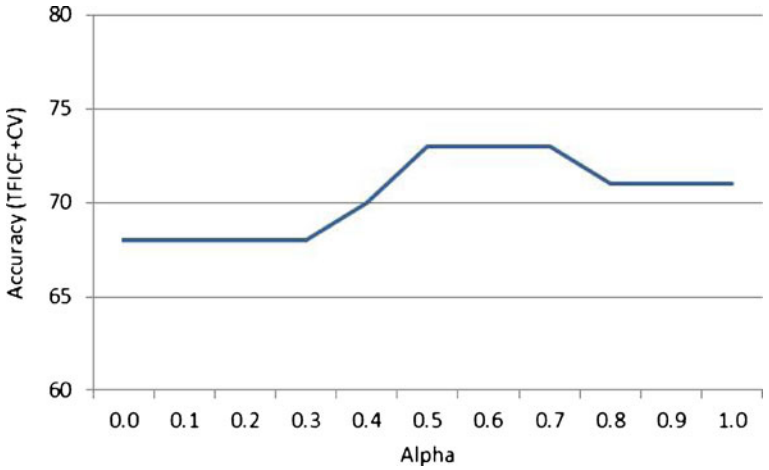

Fig. 4 Impact of alpha value in Eq. (7)

To validate the influence of language difference, the 100 concept pairs and their corresponding sentences are manually translated into English and the same experiments are conducted. Generally, the results are similar to those in Chinese. Errors in Chinese word segmentation and POS tagging are the main causes of performance differences. Since English WordNet has been translated into various languages, the proposed approaches can be easily adapted to different languages.

5 Conclusions and future work

Ontology construction is costly but domain knowledge changes rapidly, resulting in the failure of the ontology-based applications. This work develops a statistical approach for extracting semantic relations from text documents in order to assist in the construction process. First, the TFICF scheme discovers discriminative verbs among concept pairs, based on context clustering, and avoids overweighting of word frequencies that may result in the recommendation of frequent but insignificant verbs as semantic labels. Furthermore, the child voting scheme exploits taxonomic relations to estimate the frequencies of candidate verbs that are used interchangeably with parent and child concepts, which consequently reduce data sparseness. The preliminary experiment reveals that the combination of TFICF and the child voting scheme increases the accuracy of relation extraction. In summary, this work provides a novel keyword ranking mechanism for labeling semantic relations to enrich the taxonomic domain ontology. Although these findings are interesting, broad studies on method parameters for different text corpora should be further explored with a view to automating the learning process. Moreover, future investigations may apply deep semantic analysis, including, for example, semantic role labeling, to elucidate further interrelationships within the domain corpus, and thereby improve extraction performance.

Acknowledgements This research was partially supported by the National Science Council of the Taiwan under grant NSC 99-2410-H-009-034-MY3 and NSC 101-2811-H-009-002.

References

Academia-Sinica (2005). *The Academia Sinica Bilingual Wordnet (Sinica BOW)*. The Association for Computational Linguistics and Chinese Language Processing (ACLCLP).

ACE (2005). Automatic content extraction. http://projects.ldc.upenn.edu/ace/docs/.

Che, W., Liu, T., & Li, S. (2005). Automatic entity relation extraction. *Journal of Chinese Information Processing, 19*(2), 1–6.

Chen, K. J., & Huang, C. R. (1997). Academia sinica balanced corpus of modern Chinese. http://www.sinica.edu.tw/ftms-bin/kiwi1/mkiwi.sh.

Cimiano, P., Hotho, A., & Staab, S. (2005). Learning concept hierarchies from text corpora using formal concept analysis. *Journal of Artificial Intelligence Research, 24*, 305–339.

Culotta, A., McCallum, A., & Betz, J. (2006). Integrating probabilistic extraction models and data mining to discover relations and patterns in text. In *2006 North American Chapter of the Association for Computational Linguistics - Human Language Technologies (HLT-NAACL 2006), New York, 2006* (pp. 296–303).

Culotta, A., & Sorensen, J. (2004). Dependency tree kernels for relation extraction. In *Association of Computational Linguistics (ACL'04), Barcelona, Spain, 2004* (pp. 423–429).

Hasegawa, T., Sekine, S., & Grishman, R. (2004). Discovering relations among named entities from large corpora. In *Association of Computational Linguistics (ACL'04), Barcelona, Spain, 2004* (pp. 415–422).

Hepp, M. (2007). Possible ontologies: How reality constrains the development of relevant ontologies. *IEEE Internet Computing, 11*(1), 90–96.

Kavalec, M., Maedche, A., & Svatek, V. (2004). Discovery of lexical entries for non-taxonomic relations in ontology learning. In *The 30th conference on current trends in theory and practice of computer science, Merin, Czech Republic, January 24–30 2004* (Vol. LNCS 2932, pp. 249–256). Springer.

Maedche, A., Pekar, V., & Staab, S. (2002). Ontology learning part one: On discovering taxonomic relations from the Web. In *Web intelligence, New York, U.S.A., 2002* (pp. 301–322). Springer Verlag.

Maedche, A., & Staab, S. (2000). Discovering conceptual relations from text. In *The 14th European conference on artificial intelligence (ECAI 2000), Berlin, Germany, 2000* (pp. 321–325).

Manning, C., Raghavan, P., & Schütze, H. (2008). *Introduction to information retrieval*. Cambridge University Press.

Manning, C., & Schütze, H. (1999). *Foundations of statistical natural language processing*. MIT Press.

Punuru, J., & Chen, J. (2011). Learning non-taxonomical semantic relations from domain texts. *Journal of Intelligent Information Systems, 38*(1), 191–207.

Sánchez, D., & Moreno, A. (2008). Learning non-taxonomic relationships from Web documents for domain ontology construction. *Data and Knowledge Engineering, 64*(3), 600–623.

Schutz, A., & Buitelaar, P. (2005). RelExt: A tool for relation extraction from text in ontology extension. In *The 4th international semantic Web conference (ISWC), Galway, Ireland, 2005* (Vol. 3729, pp. 593–606).

Serra, I., & Girardi, R. (2011). Extracting non-taxonomic relationships of ontologies from texts. In *The 6th International conference SOCO 2011 soft computing models in industrial and environmental applications, Ostrava, Czech Republic, 2011* (Vol. 87, pp. 329–338). Springer

Sumida, A., Torisawa, K., & Shinzato, K. (2006). Concept-instance relation extraction from simple noun sequences using a full-text search engine. In *The ISWC 2006 workshop on Web content mining with human language technologies (WebConMine), Athens, GA, U.S.A.*

SUMO (2011). Suggested upper merged ontology. http://www.ontologyportal.org.

Villaverde, J., Persson, A., Godoy, D., & Amandi, A. (2009). Supporting the discovery and labeling of non-taxonomic relationships in ontology learning. *Expert Systems with Applications, 36*(7), 10288–10294.

Weichselbrauna, A., Wohlgenannta, G., & Scharl, A. (2010). Refining non-taxonomic relation labels with external structured data to support ontology learning. *Data & Knowledge Engineering, 69*(8), 763–778.

Weichselbraun, A., Wohlgenannt, G., Scharl, A., Granitzer, M., Neidhart, T., & Juffinger, A. (2009). Discovery and evaluation of non-taxonomic relations in domain ontologies. *International Journal of Metadata, Semantics and Ontologies, 4*(3), 212–222.

Zelenko, D., Aone, C., & Richardella, A. (2003). Kernel methods for relation extraction. *Journal of Machine Learning Research, 3*(6), 1083–1106.

Zhang, S., Wen, J., Wang, X., & Li, L. (2006). Automatic entity relation extraction based on maximum entropy. In *The 6th international conference on intelligent systems design and applications (ISDA'06), Jinan, China, 2006* (Vol. 1, pp. 540–544). IEEE Computer Society.