The Opinion Mining Based on Fuzzy Domain Sentiment Ontology Tree for Product Reviews

Xinhui Nie, Lizhen Liu, Hanshi Wang*, Wei Song
College of Information Engineering, Capital Normal University, Beijing, China
ldx4611@126.com

Jingli Lu
Grasslands Research Centre, New Zealand
Janny.jingli.lu@gmail.com

Abstract—**In this paper, we present a novel method that integrates domain sentiment knowledge into the analysis approach to deal with feature-level opinion mining By constructing a domain ontology called Fuzzy Domain Sentiment Ontology Tree (FDSOT), we then utilize the prior sentiment knowledge of our ontology to achieve significantly accuracy in sentiment classification. Particularly, the FDSOT is the conceptual model which represents the semantic relation between features and sentiment words. The evaluation is based on the Chinese product reviews collected from 360buy.com[1]. The experimental results demonstrate that our approach is able to automatically identify the domain-dependent polarity for a large subset of sentiment expression and effectively improve the performance of opinion mining.**

Index Terms—f**eature-sentiment pairs, product reviews ontology-based, sentiment classification**

I. Introduction

With the dramatic growth of e-commerce and the popularity of online merchants, these numbers of online product reviews become an opinion resource which is very useful for both potential customers and product manufacturers. This situation is also notable in Chinese web services. However, manually browsing a large number of consumer reviews posted to the Web may not be feasible, if not totally impossible. In addition, product reviews are usually fragmental and buried among the domain-specific knowledge. Therefore, research on automatic opinion mining of review texts has become a popular research topic at the cross-field of Web text mining, natural language processing and computational linguistics.

The understanding of these unstructured product reviews is not only important for humans, but also is critical for human-computer interaction. One resource that is much needed in solving the problem of emotional understanding is ontology. The ontology is a formal knowledge representation method to facilitate human and computer interactions and it can be expressed by using formal semantic markup languages such as RDF and OWL [1]. And in particular, domain ontology constructs the relationship among domain concepts and helps in determining the structure of knowledge [2].

As such, this paper will study the problem of opinion mining on product reviews through a novel method which is based on Fuzzy Domain Sentiment Ontology Tree (FDSOT). Our method performed at the feature-level to provide the in-depth sentiment analysis for target product features. We believe that labeling product reviews with product features and their corresponding opinions is an effective way to solve the domain-dependent sentiment problem.

The remainder of this paper is organized as follows. After reviewing the related work in section 2, we present the basis of ontology model and sentiment classification in section 3. Then we present the experimental results and discussion in section 4. Finally, we conclude by showing the utility of our method and describe the future plan in section 5.

II. Related Work

The early research of opinion mining (or sentiment analysis) is defined as the task of the sentiment classification at document-level [3]. However, for many opinion expressions such as twitter [4], micro blogs, and customer feedback reviews only judging the sentiment orientation is not sufficient. Therefore, increasingly more research has examined opinion mining at the sentence, phrase level [5-7] and more fine-grained feature-level [8-11] in recent years.

Feature-based opinion mining [12,13] (also called aspect-based sentiment analysis) is the research problem that focuses on the recognition of all sentiment expressions within a given document (e.g. a customer review) and the features to which they refer. The product reviews which people commented have many features (aspects) and different opinion about each feature. While many existing researches of feature-level opinion mining have been focused on English texts, and little study has been conducted on Chinese texts [14]. Due to the complexity of Chinese expression and the limited resources of Chinese sentiment analysis, our work which

* Supported by the Beijing Natural Science Foundation (No. 4133084) and the Beijing Key Disciplines of Computer Application Technology.

* Corresponding Author: Hanshi Wang

doi:10.4304/jsw.8.11.2682-2687

performs opinion mining for Chinese product reviews is a more challenging task. Our research approach is driven by WeiWei and Gulla [15], which proposed a HL-SOT approach to labeling product attributes and their associated sentiments in product reviews by a Hierarchical Learning (HL) process with a defined Sentiment Ontology Tree (SOT). However, the SOT is manually constructed. What attributes to be included in a product's SOT and how to structure these attributes in the SOT is an effort of human beings. The sizes and structures of SOT constructed by different individuals may vary. The key difference between our work and theirs is that we exploit a set of seeds which include product features and sentiment words to construct a FDSOT model and develop it through extracting feature-sentiment pairs. Lina Zhou and Pimwadee Chaovalit [16] proposed an ontology-supported polarity mining (OSPM) approach. Their approach aims to enhance polarity mining with ontology by providing detailed topic-specific information. OSPM is evaluated in the movie review domain using both supervised and unsupervised techniques. The key point of this study is polarity mining, not ontology construction. Lau et al [17] contributed to the development of a novel fuzzy domain ontology extraction method for adaptive e-learning.

III. Sentiment Analysis based on Fuzzy Domain Sentiment Ontology Tree

The opinion mining architecture based on Fuzzy Domain Sentiment Ontology Tree is illustrated in Fig 1. In the proposed architecture, opinion mining starts with collecting raw corpus from the Internet. Preprocessing cleans up the corpus such as stop word removal, Part-of-Speech (POS) tagging. The FDSOT is generated by the seed set and candidate feature-sentiment pairs extracted from review corpus. By matching with the FDSOT and analyzing the polarity of product features from all the reviews, a final presentation of sentiment analysis can be shown for users.

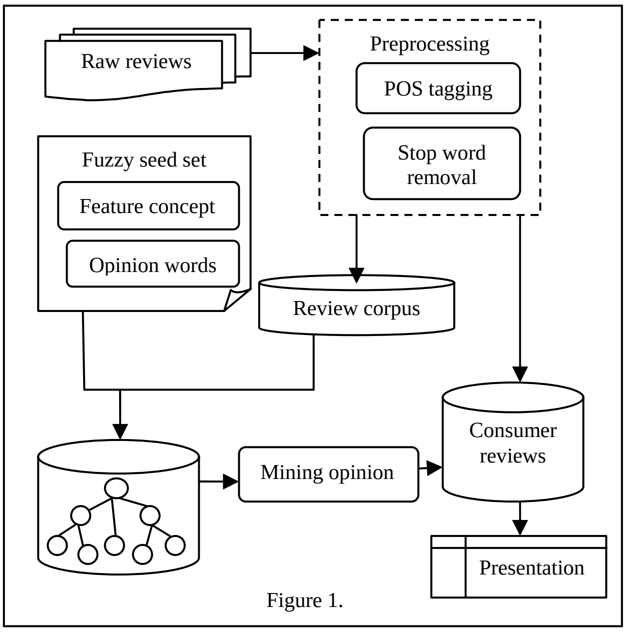

Figure 1. The architecture of opinion mining system

A. Model Construction

Ontology is generally considered as a formal specification of conceptualization which consists of concepts and their relationships [18]. The fuzzy domain sentiment ontology tree extraction method is developed based on fuzzy sets and semantic relations which offer the expressive power to capture the uncertainty presented in opinion mining. Our objective is to build the concept model which represents the semantic relation between product features and sentiment words. Note that product features (aspects) mean product components and attributes such as "外形(appearance)", "键盘(keyboard)", "做工(workmanship)". Sentiment words are sentiment-conveying terms which bear some sentiment orientation (positive or negative) such as "完美 (perfect)", "便宜(cheap)", "难看 (ugly)" and so on.

FDSOT model: it is expressed in a tree-hierarchy of concepts. The root node of FDSOT is product itself (e.g. mobile phone), each non-leaf child node of the root of the FDSOT represents a sub-feature belonging to its parent feature like "电池 (battery)". All the leaf nodes of FDSOT represent sentiment (positive or negative) nodes respectively associated with their parent nodes. This definition successfully describes the semantic relation between features and their associated sentiment.

***Fuzzy set*:** it consists of a synonyms set of sentiment words (positive and negative). For example, the sentiment node "便宜(cheap)", which has similar sentiment words like "廉价", "实惠".

Fuzzy relationship: it defines the semantic relation for product features and sentiment words in FDSOT. We use the score of sentiment orientation *SO* to describe the grade of sentiment categories (positive and negative).

B. Seeds Collecting and Corpus Preprocessing

Seeds collecting and corpus preprocessing are the preparation of FDSOT constructing. Our objective that collects seeds is to put them as the sentiment clue and extracted features and sentiment words from review corpus.

We collect some domain feature concepts by their frequency in reviews. Then we partly capture general sentiment words from synonym dictionary and sentiment lexicon of Hownet. The obligatory task of our method is pre-processing of review texts because the Chinese texts are different from English words which can be delimited by white space. Therefore, the preprocessing is done by word segmentation, POS tagging (by ICTCLAS) and stop word removal.

C. Extracting Features and Sentiment Words

After the seeds collecting and preprocessing step, we described our methods to extract product feature and sentiment words from online customer reviews. The goal of extraction is to get candidate feature words to be paired with their sentiment words. In Chinese, noun and noun phrase are regarded as the most informative patterns to represent the product features of a particular product. The Adjectives or Adverbs associated with the product features within a review sentence are extracted as the candidate sentiments. Traditional method exploits sentiment coherency within a sentence to extract sentiment candidate and then use a statistical method to determine whether a candidate is correct. This method can be unreliable when the occurrences of candidates are infrequent with small corpora. Therefore, we attempt to respectively extract frequent candidates and infrequent candidates.

As basic classification unit for our fine-grained sentiment analysis, we choose sub-sentence segments. In this work, sub-sentence means we segment product review into clauses by punctuation containing a comma. Then we choose the candidate noun and adjective (adverb) which are frequent co-occurrence (within a text window of 5 sizes) in a sub-sentence. The formula is shown as following:

$$MI\ (F_i, S_i) = \log \frac{P(F_i S_i)}{P(F_i) P(S_i)} \tag{1}$$

Where $MI(F_i, A_i)$ is a function to estimate the degree of association between product feature F_i .and its sentiment word S_i. $P(F_i, S_i)$ is the frequency of co-occurring terms. In our experiment, the threshold θ is 0.637.

As mentioned, the first stage takes the frequency-based view to discover frequent features and sentiment words. As a result, it is incapable of extracting infrequent candidates. We then utilize the seed set to learn new candidate terms. It is a measure to complement some weaknesses of the above method. The adjective, adverb or verb associated with the seed-features within a sub-sentence is extracted as a candidate sentiment words. It means that the adjective or adverb next to a seed-feature from both sides is extracted. And we extract candidate features with the similar method. The details are as following:

$$If\ \exists (NN, A^\circ = S\ , A) \rightarrow Pair = (NN, A^\circ), (NN, A) \tag{2}$$

$$If\ \exists (NN^\circ = F, NN, A) \rightarrow Pair\ = (NN, A), (NN^\circ, A) \tag{3}$$

Where *NN* represents noun, noun phrase, *A* represents adjective or adverb, NN° and A° respectively denotes seed features *F* and seed sentiment words *S* which we collected before. The *Pair* means the feature-sentiment pair (*F*, *S*). The basic idea is that we employ the part-of-speech and semantic relation to capture the infrequent candidate features and their opinion. If one clause has been identified with known features (sentiment words), the potential sentiment word (feature) will be paired with the term.

D. Computing Sentiment Orientation

After extracting pairs of product feature and sentiment word, we propose a method to predict the sentiment orientation for each feature-sentiment pair. The basic idea is that a positive product review is more likely to contain positive opinion pairs than a negative review does. Therefore, we may use the sentiment polarity label of a review text to infer the sentiment orientation of a product feature within the clause.

To calculate the sentiment orientation of a pair, at first we classify positive and negative polarity labels and then proposed a formula to predict the value of sentiment. The function is defined as follows:

$$SO(p) = P(pos \mid p) \times \log_2 \frac{P(pos \mid p)}{P(pos)} - P(neg \mid p) \times \log_2 \frac{P(neg \mid p)}{P(neg)} \tag{4}$$

Where $SO(p)$ represents the polarity score of an opinion pair *p*. The P_{pos} (P_{neg}) is the prior probability that a review is positive (negative) respectively. The $P(pos \mid p)$ ($P(neg \mid p)$) is the estimated conditional probability that a review is positive (negative) if it contains the opinion pair *p*. Their formulations are described by (5) and (6).

$$P(pos) = \frac{fre(R_{pos})}{fre(R)},\ P(neg) = \frac{fre(R_{neg})}{fre(R)} \tag{5}$$

$$P(pos \mid p) = \frac{fre(p_{pos})}{fre(p)},\ P(neg \mid p) = \frac{fre(p_{neg})}{fre(p)} \tag{6}$$

Where P_{pos} (P_{neg}) is the percentage of the frequency between the positive (negative) review R_{pos} and overall reviews *R*. The $P(pos \mid p)$ ($P(neg \mid p)$) denotes the rate of $fre(p_{pos})$ ($fre(p_{neg})$) and $fre(p)$. The $fre(p_{pos})$ is the number of positive reviews which contain the opinion pair *p*.

Based on the aforementioned information, we can get opinion pairs and their sentiment orientation. Finally, we can induce an entity model of FDSOT.

E. Sentiment Classification

Automatic sentiment analysis has been well studied with a variety of lexicon-based and machine learning based systems. In our method, we employed the FDSOT as the domain sentiment knowledge to mine the product features and their sentiment (positive or negative) in each sentence. The FDSOT which we have constructed is a conceptual model for the representation of specific-domain sentiment knowledge. The prior sentiment resource which is held by our ontology is utilized to analyze the target review.

Given a large number of product reviews about various detail features, we can identify the sentiments that are highly associated with a given product feature by matching the candidate feature and sentiment word with the common feature-opinion pairs in our FDSOT. As mentioned before, we choose the representative product features from each review and candidate sentiments which are close to the product features. The polarity score of feature-sentiment pair is calculated by finding out the closest matching result between candidate sentiment words in reviews and a fuzzy set of sentiment words in FDSOT. We then take the negation of sentiment into account within a text window. For example, if words such as "不 (no)", "没有(not)", "比不上 (far less)", and so on is found, the negation of the sentiment word is assumed. As result, the small negation dictionary is manually compiled by us. It consists of 27 negation words. Finally, we present the sentiment polarity of review sentence.

IV. Experiments and Discussions

The experiments are carried out on product reviews from 360buy.com. We crawled about 4000 reviews from different brands mobile phone because there are no authoritative Chinese review data sets. This testing corpus consists of 1000 product reviews. To conduct evaluations, we pre-construct an evaluation set. Each review of the testing reviews is manually annotated the product feature words, sentiment words and sentimental tendencies of reviews by our members.

To evaluate the method, we adopt precision, recall, and F-measure to measure the effectiveness of identifying feature-sentiment pairs and their sentiment polarity. The definitions of Precision, Recall and F-measure are as following:

$$\text{Precision} = \frac{\#\{identified(F,S) \cap labeling(F,S)\}}{\#\{identified(F,S)\}} \tag{6}$$

$$\text{Recall} = \frac{\#\{identified(F,S) \cap labeling(F,S)\}}{\#\{labeling(F,S)\}} \tag{7}$$

$$F-measure = \frac{2 \times \Pr ecision \times \mathrm{Re}\, call}{\Pr ecision + \mathrm{Re}\, call} \tag{8}$$

Where Precision is the fraction of the focuses detected that are considered feature-sentiment pairs. Recall is the fraction of the focuses detected that are relevant to the labeled pairs. The Precision, Recall and F-measure of sentiment polarity are similar with the above definition.

A. Experimental Result of FDSOT

In this section, we present and discuss the experimental result of the FDSOT construction. We first perform simple preprocessing on these reviews. The preprocessing procedures of reviews include Chinese word segment, part of speech tagging. Then we remove the irrelevant context from review texts. Analyzing the 3000 reviews, the total number of the clauses is 10208. After that we put the data set as inputs to build our FDSOT by our method as mentioned before. A snapshot of our ontology is depicted in Fig 2. Samples of fuzzy sets in our FDSOT are illustrated in Table I.

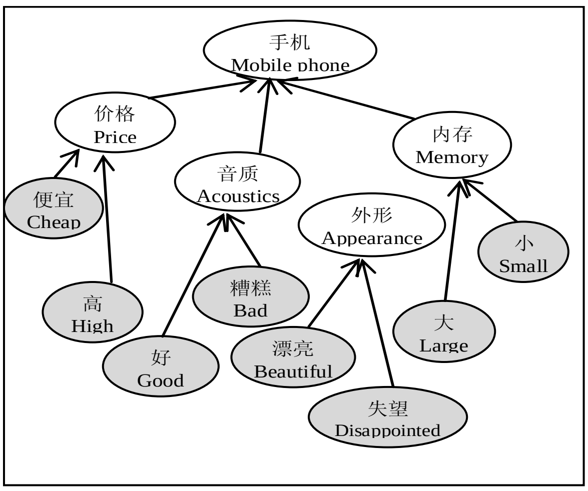

Figure 2. A snapshot of the FDSOT for mobile phone

TABLE I.
THE CONCEPT NODES AND FUZZY SETS OF FDSOT

Non-leaf Nodes	**Samples of fuzzy sets**
外观(appearance)	“外形, 外貌, 样式, 款式”
内存(memory)	“存储卡, 内存卡”
音质(acoustics)	“声音,音效,音色”
屏幕(screen)	“屏,显示屏,触摸屏”
功能(function)	“性能, 属性, 总体性”
价格(price)	“价格,价钱,价位”
Leaf nodes	**Samples of fuzzy sets**
简洁(briefness)	“简单, 简朴, 简约”
喜欢(love)	“喜爱, 钟爱, 给力”
便宜(cheap)	“便宜, 廉价, 物美价廉”
漂亮(beautiful)	“美丽, 美观, 好看”
差劲(bad)	“差,次品, 坏 ”
满意(satisfied)	“合意, 中意, 称心”

B. Evaluation of Opinion Mining with FDSOT

In the phase, our purpose is to examine the performance of the proposed method which performs the opinion mining based on FDSOT for the testing corpus. From Table II, we can see that opinion mining based on FDSOT achieved higher accuracy. The improvement of performance is mainly due to the increasing of sentiment knowledge in FDSOT. The experimental results are partly reported in Table III. Our method achieves the relatively higher F-measure score because of the automatic learning of sentiment contextual knowledge for opinion pairs. The knowledge was then applied to bootstrap opinion mining during the testing experiment. By the results in Table II, we can see that the result of identifying product features and sentiment words are relatively effective. According to our in-depth analysis, it was revealed that standard names of product features such as “外形(appearance)”, “价格(price)” and “屏幕(screen)” for mobile phones were frequently used in consumer reviews. Due to diversities of opinion expressions, some factor was scanted by us. For example, quite a number of product features were not extracted from the reviews if we used the “Verb Noun” pattern like “送货(deliver goods)” to represent product features. And it does not always have feature and its sentiment concurrently in each sentence.

TABLE II.
F-MEASURE, PRECISION AND RECALL OF EXPERIMENT

Opinion mining for Mobile phone			
Analysis Category	**Precision**	**Recall**	**F-Measure**
Feature-sentiment pairs	0.732	0.409	0.522
Sentiment polarity	0.587	0.625	0.605

TABLE III.
SENTIMENTAL TENDENCIES FOR FEATURES OF MOBILE PHONE

Features	**Sentiment**	**polarity**
手机(mobile phone)	不错 (good)	+0.729
屏幕 (Screen)	大 (big)	+0.738
散热 (heat dissipation)	好 (good)	+0.824
画面(picture)	清晰(clear)	+0.786
价格(price)	便宜(cheap)	+0.712
外形(appearance)	漂亮(beautiful)	+0.751
产品(product)	满意(satisfied)	+0.693
颜色(color)	不错(good)	+0.685
屏幕(screen)	死机 (crash)	-0.583
信号(signal)	糟糕(bad)	-0.809
运行速度(speed)	慢(slow)	-0.627
手感(feel)	差劲(poor)	-0.822
品牌(brand)	失望(disappointed)	-0.646
做工(workmanship)	不给力 (disappointed)	-0.791

V. CONCLUSION AND FUTURE WORK

This paper has introduced the FDSOT, a new model to represent domain-dependent sentiment knowledge by using the space of product features and sentiment words. This model can store sentiment polarity of opinion pairs with varying contextual information. It is clear from the experimental results presented that our ontology can be automatically constructed to facilitate opinion mining and accurately predict the polarities of sentiments. As a result, organizations can develop effective business strategies related to marketing, customer support, and product design functions in a timely fashion.

In our future work, there are many possibilities. First, we wish to extend the product features and sentiment words by adding other parts-of-speech like gerundial phrases conveying features and verbs bearing opinions. We would also like to further improve the classification accuracy and identify the implicit features by utilizing knowledge resource in our ontology expansion.

ACKNOWLEDGMENT

This work was supported by the Beijing Natural Science Foundation (No. 4133084) and the Beijing Key Disciplines of Computer Application Technology.

REFERENCES

[1] Jike Ge, Zushu Li, Taifu Li, “A Novel Chinese Domain Ontology Construction Method for Petroleum Exploration Information,” *Journal of Computers*, Vol 7, No 6 (2012), pp. 1445-1452, Jun 2012.

[2] Weifeng Pan, Shan Li, “Tag Ontology Automatic Building for Semantic Searching of Services: a Case Study on Mashup Services,” *Journal of Computers*, Vol 7, No 12 (2012), pp. 2979-2986, Dec 2012.

[3] Turney P D., “Thumbs up or thumbs down?: semantic orientation applied to unsupervised classification of reviews,” *Proceedings of the 40th Annual Meeting on Association for Computational Linguistics, Stroudsburg, PA, USA: Association for Computational Linguistics*, pp. 417–424, 2002.

[4] Zhao J, Dong L, Wu J, Xu K., “MoodLens: an emoticon-based sentiment analysis system for Chinese tweets,” *Proceedings of the 18th ACM SIGKDD international conference on Knowledge discovery and data mining*, ACM, pp. 1528-1531, 2012.

[5] T. Wilson, J. Wiebe, and P. Hoffmann, “Recognizing contextual polarity in phrase-level sentiment analysis,” *in Proceedings of the Conference on Human Language Technology and Empirical Methods in Natural Language Processing. Association for Computational Linguistics*, pp. 347–354, 2005.

[6] Xu, R., Wong, K.F., Lu, Q., Xia, Y., and Li, W. “Learning Knowledge from Relevant Webpage for Opinion Analysis”, *in Proceedings of the 2008 IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology, Sydney, Australia*, pp. 307–313. 2008.

[7] Agarwal, A., Biadsy, F., and Mckeown, F. “Contextual Phrase-Level Polarity Analysis Using Lexical Affect Scoring and Syntactic N-Grams,” *in Proceedings of the 12th Conference of the European Chapter of the Association for Computational Linguistics*, pp.24-32. 2009.

[8] Archak, N., Ghose, A., and Ipeirotis, P. “Show me the money!: deriving the pricing power of product features by mining consumer reviews”, *in Proceedings of the 13th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Jose, California* , pp. 56 - 65. 2007.

[9] Lei Zhang and Bing Liu, “Extracting Resource Terms for Sentiment Analysis,” *Proceedings of the 5th International Joint Conference on Natural Language Processing (IJCNLP-2011)*, pp.1171-1179, November 8-13, 2011.

[10] Ming Xie, Chanle Wu, Yunlu Zhang, “A New Intelligent Topic Extraction Model on Web,” *Journal of Computers*, Vol 6, No 3 (2011), pp. 466-473, Mar 2011.

[11] Veale T. A context-sensitive, multi-faceted model of lexico-conceptual affect. *Proceedings of the 50th Annual Meeting of the Association for Computational Linguistics: Short Papers-Volume 2, ACL*, pp. 75-79, 2012.

[12] Zhai, Z., Liu, B., Xu, H. and Jia, P. “Clustering Product Features for Opinion Mining”, *Proceedings of Fourth ACM International Conference on Web Search and Data Mining (WSDM-2011)*, pp.347-354, 2011.

[13] Qu L, Gemulla R, Weikum G., “A weakly supervised model for sentence-level semantic orientation analysis with multiple experts,” *Proceedings of the 2012 Joint*

Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning, ACL, pp. 149-59, 2012.
[14] Zhang H, Yu Z, Xu M, Shi Y., "An Improved Method to Building a Score Lexicon for Chinese Sentiment Analysis," *Semantics, Knowledge and Grids (SKG), 2012 Eighth International Conference on, IEEE*, pp. 241-244, 2012.
[15] WeiWei and Jon Atle Gulla., "Sentiment Learning on Product Reviews via Sentiment Ontology Tree," *Proceedings of the Association for Computational Linguistics (ACL)*, pp.404-413, 2010.
[16] Zhou ,L .and P. Chaovalit, "Ontology-supported Polarity Mining", *Journal of the American Society for Information Science and Technology*,59(1), 1-13, 2008
[17] Lau, R.Y.K., Song, D., Li, Y., Cheung, C.H., Hao, J.X. "Towards A Fuzzy Domain Ontology Extraction Method for Adaptive e-Learning", *IEEE Transactions on Knowledge and Data Engineering*, (21:6), pp. 800-813, 2009.
[18] Gruber, T.R. "A translation approach to portable ontology specifications", *Knowledge Acquisition*, (5:2), pp. 199–220, 1993.

Xinhui Nie is studying for a master's degree at Capital Normal University. Now, her research interest includes knowledge acquisition, Opinion Mining, Data Mining, and Computational Linguistics.

Prof. Lizhen Liu holds a PhD in Computer Application from the Beijing Institute of Technology, China. She is currently a Professor at the Capital Normal University. Her research interests include text mining, sentiment analysis, knowledge acquisition, and the design of Intelligent Tutoring Systems.

She has published in journals and conferences like Knowledge and Information Systems, International Journal of Information & Computational Science, Journal of Computers, IEEE World Congress on Intelligent Control and Automation, CSCWD and so on.

She served as PC of a number of international conferences, including IEEE Pervasive Computing and Application, International Symposium on IT in Medicine & Education, Computer Supported Cooperative Work in Design and so on. Prof. Lizhen, the director of intelligence science & technology department, is also a member of ACM, IEEE and CCF.

Dr. Hanshi Wang holds a PhD in Computer Application from the Beijing Institute of Technology, China. He is currently a lecturer at the Capital Normal University. His research interests focus on computational linguistics and natural language processing, especially unsupervised methods in the area.

He has published his important work on the famous journal of Computational Linguistics (CL), and other international conferences.

Dr. Wei Song holds a PhD in Computer Application from the Harbin Institute of Technology, China. He is currently a lecturer at the Capital Normal University. His research interests focus on Query Understanding, Sentiment Analysis, Personalized Search, Social Computing.

He has published his important work on the Proceedings of the ACM SIGIR Conference, International Conference on Computational Linguistics (COLING), Journal of Chinese Information Processing and so on.

Dr. Jingli Lu received a doctoral degree in Computer Science from Massey University, New Zealand. She is currently a Software Engineer at Agresearch Ltd, New Zealand. Her research interests include Data Mining and Machine Learning, Fuzzy Set Theory, and Speech Signal Processing. She has published several papers in highly-rankly international journals and conferences.