

<http://ansinet.com/itj>

ITJ

ISSN 1812-5638

# INFORMATION TECHNOLOGY JOURNAL

**ANSI***net*

Asian Network for Scientific Information  
308 Lasani Town, Sargodha Road, Faisalabad - Pakistan

## An Attribute-based Sentiment Analysis System

<sup>1,2</sup>Hanxiao Shi, <sup>1</sup>Guodong Zhou and <sup>1</sup>Peide Qian

<sup>1</sup>School of Computer Science and Technology, Soochow University, Suzhou, Jiangsu, 215006, China

<sup>2</sup>School of Computer Science and Information Engineering, Zhejiang Gongshang University,  
Hangzhou, Zhejiang, 310018, China

---

**Abstract:** With the rapid development of E-Commerce, the number of customer reviews that a product receives grows rapidly. For a popular product, there are a large number of reviews. This makes it difficult for a potential customer to make an informed decision on purchasing the product, as well as for the manufacturer of the product to keep track and to manage customer opinions. Polarity analysis of reviews based on sentence or document level often can not show precise results, because of several attributes of the product included in one review sentence. This study uses semantic role labeling tool to realize shallow semantic analysis of reviews and propose a novel approach to focus the attribute-polarity pair and make sentiment orientation analysis. Finally the visualization of prototype system is implemented to help potential consumers make a whole understanding of all reviews. Experimental results on a real-world data set show that the system is feasible and effective.

**Key words:** Sentiment classification, opinion mining, semantic role labeling, domain-based dictionary

---

### INTRODUCTION

As the widespread use of computers and the high-speed development of the Internet, E-Commerce has already penetrated as a part of our daily life. Currently most information on the Internet exists in unstructured or semi-structured forms and presented the explosive growth. Besides, along with the development of e-commerce, the more appeared reviews not only help potential consumers to make decisions of the products in a certain extent, but also provide some good feedback for merchant. For instance, when a consumer plans to buy one digital camera, he will surf on the BBS or review sites to read the reviews from experienced consumers.

However, the reader's taste may differ from the reviewer's. For example, the reader may feel strongly about the image quality of digital camera, whereas many reviewers may focus on other aspects, such as the battery or the price. Thus, the reader is forced to wade through a large number of reviews looking for information about particular attributes of interest. Moreover, there is also some noise such as misleading articles often appearing in the first few pages, which would affect the comprehensiveness of browsers' information acquisition and correctness of their judgment. In addition, when it comes to the merchants, they also need to get the first-hand data of customer opinions from the reviews. Based on the opinions of each aspect of the product, the

merchants can grasp which products are satisfied by customers and which ones need improvements, etc. However, for a popular product, the number of reviews can be in hundreds or even thousands. This makes it difficult for a potential customer to read them to make an informed decision on whether to purchase the product. It also makes it difficult for the merchant of the product to keep track and to manage customer opinions. Nowadays, some websites have made quantized expressions for the Sentiment Orientation (SO) of their local review information, such as amazon.com, which has coarse-grained rating (5-star scale) for each review on its website and the 5-star is the best, while the 1-star is the worst, then giving the total rating. And ctrip.com has realized fine-grained rating for the hotel's each reviews in some extent, mainly aimed at four aspects as room hygiene, hotel service, surrounding environment and facilities, making 5-score rating, then provide the total scores.

In the past few years, many researchers transfer their interests from text classification (Subramanian and Ramraj, 2007; Umer and Khiyal, 2007) to sentiment analysis (Tuney, 2002; Pang *et al.*, 2002; Dave *et al.*, 2003; Hu and Liu, 2004a, b; Carenini *et al.*, 2006; Liu *et al.*, 2007; Baccianella *et al.*, 2009). Current researches mainly focused on proposing novel analyzing and processing technologies based on different domains, according to the large scale review data acquired from Internet, such as (1) Using Part-Of-Speech (POS) to tag the sentences, several

researchers summarized some rules to focus the object of opinion and sentiment items (Hu and Liu, 2004a; Baccianella *et al.*, 2009) and utilized the distribution rules of POS to extract the corresponding template for sentiment analysis (Liu and Seneff, 2009); (2) some scholars also started to make researches in sentiment orientation of different sentence structures (Narayanan *et al.*, 2009; Ku *et al.*, 2009); (3) Kim and Hovy (2006) used the technology of semantic role labeling (SRL), which was mainly utilized in news and public's opinion analysis, to help identify two main components of opinions: opinion holder and topic. Of course there were also some scholars aiming at the reviews in a special website and proposing special processing methods, so as to obtain the sentiment orientation of reviews (Danescu-Niculescu-Mizil *et al.*, 2009).

There have been three major ways for sentiment orientation analysis recently. The first one is based on the simple statistic (Turney, 2002; Nasukawa and Yi, 2003) mainly made simple statistics for orientation values to obtain the whole tendency of texts. This method is generally applied to the document-level sentiment analysis, such as Tsou *et al.* (2005) makes statistics for the sentiment orientation of news articles and measures the opinions of celebrities from the public through calculating the sentiment orientation of the words and comprehensively considers the spread, density and semantic intensity of the polarity elements. Although, the sentiment analysis based on the simple statistic belongs to coarse-grained orientation classification, because of its simple realization and not bad accuracy, it occupies a particular weight in the beginning of the orientation study.

The second one is based on machine learning, generating orientation classification model through the training of numerous labeled corpuses and then classifying the test texts using generated model. Pang *et al.* (2002) and Pang and Lee (2005) adopted the technology of standard bag-of-words and three machine learning methods (naive bayes, maximum entropy classifications and SVM) to make text orientation classification for the film reviews and, respectively compared them with the outcome of manual classification. The result of experiment shows that the method of SVM has the best effectiveness among several classifications. Chaovalit and Zhou (2005) also used the methods of machine learning and sentiment orientation to deeply mine the film reviews. Mullen and Collier (2004) proposed an approach to sentiment analysis which uses SVM to bring together diverse sources of potentially pertinent

information, including several favorability measures for phrases and adjectives and, where available, knowledge of the topic of the text. It indicates that the accuracy improves by joining the category feature of semantic orientation. Whitelaw *et al.* (2005) presented a method for sentiment classification based on extracting and analyzing appraisal groups which is represented as a set of attributes values in several task-independent semantic taxonomies. They used semi-automated methods to build a lexicon of appraisal adjectives and their modifiers and classified movie reviews using features based on these taxonomies combined with standard bag-of-words features and reported the accuracy of 90.2%.

The third one is based on the attribute-level sentiment orientation analysis. Comparative to the orientation analysis of the former two coarse-grained ones, this method belongs to fine-grained analysis. It realizes orientation analysis through the co-occurrence condition of attribute-sentiment pairs or through syntactic dependency analysis of attribute-sentiment pairs. There have been some researches on fine-grained analysis of sentiment orientation. Hu and Liu (2004b) used the co-occurrence between sentiment words and candidate attributes, on the basis of POS tagging, realizing the extraction from high-frequency attributes. They have already considered the dependence between attribute words and sentiment words when making the extraction of attribute-sentiment pairs. Miao *et al.* (2009) extracted reviews for each product based on a tuple including four elements (title, help, date, R-content), where title was the title of the review, help was the number of customers who find the review was helpful, date was the date when the review is commented, R-content was a set of sentences in customer reviews. They generated visual comparison between positive and negative evaluation of a particular feature which potential customers are interested in. They also proposed a novel ranking mechanism incorporating temporal opinion quality (TOQ) and relevance to meet customers' information need. Also, Popescu *et al.* (2005) presented an unsupervised information extraction system OPINE, which used relaxation labeling to find the semantic orientation of words in the context of given product features and sentences.

This study focuses on two aspects when making the fine-grained sentiment orientation analysis of reviews: (1) to identify the attributes of the product and (2) to judge whether the opinion of related attribute is positive or negative. We put forward to use SRL tool to make shallow semantic parsing for sentences and design a set of analysis method of fine-grained sentiment orientation

based on counting the probability of each semantic role for attributes or sentiment.

**SYSTEM INTRODUCTION**

The system consists of three parts: (1) review data pre-processing; (2) attribute-based fine-grained SO analysis and (3) user interface. The system architecture is shown in Fig. 1.

**Review data pre-processing:** The reviews are periodically crawled from the e-commerce websites such as taobao.com. These review documents are then filtered to remove HTML tags and select valuable items to store in database with structured data type. We represent each sentence of a customer review as a tuple including five elements [Product, Author, Review, Time, URL]. Product is the product name of commentary target; Author is the author of review; Review is the concrete review which has removed redundant tags; Time is the date-time when posting review; URL is the hyperlink of this review.

**Attribute-based fine-grained SO analysis:** Due to the existence that a semanteme has a variety of expressions and attribute or sentiment words have different meanings on different domains, we design a domain-based multi-dictionary which is divided as: basic polarity dictionary, special domain-based polarity dictionary, domain-based attribute dictionary, degree word dictionary and negative word dictionary. In each dictionary, it defines a variety of synonymous expression words for a semanteme. We make statistical analysis through the outcome of SRL with large amounts of data and summarize some rules to realize: (1) focusing the attributes of commentary object; (2) focusing sentiment words or phrase and (3) focusing the attribute-polarity pair and make SO analysis.

**User interface:** Users can visit and search the orientation outcome of target object by browsers. Assuming you want to search for sentiment orientation of one digital camera, you only need to input its name in the searching bar and click the search button. Then you can see all the reviews relative to this camera, in the mean time, you can also find the visualization of each attribute orientation based on these reviews, as shown in Fig. 2a. With a single glance of its visualization, general users even don't need to see the actual reviews because the visual illustration has already made a whole understanding of all reviews so that they can clearly see the strengths and weaknesses of each attribute. Users can also compare with the similar products and this system can give the result of visual

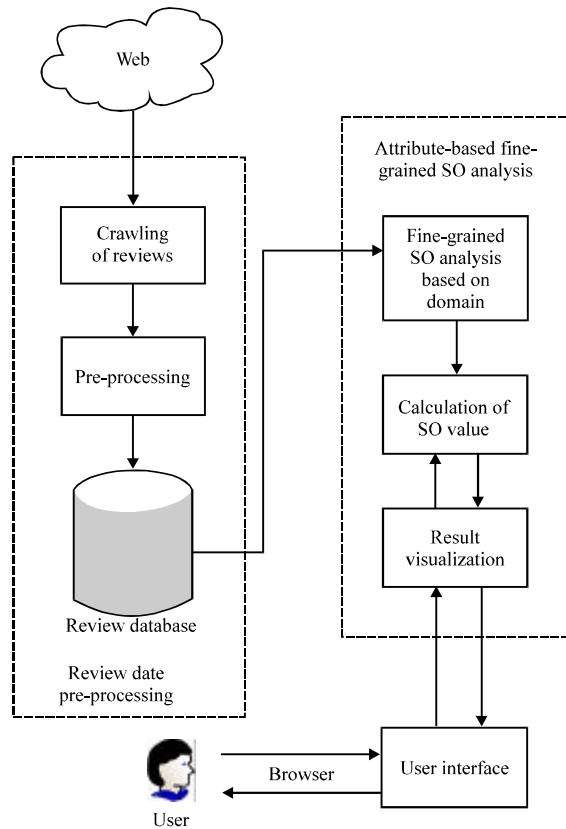


Fig. 1: The architecture of the SO analysis system

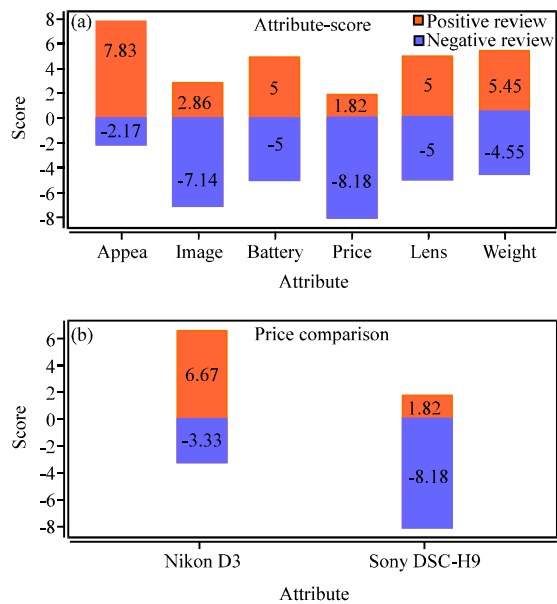


Fig. 2: Visual results based on attributes. (a) Visualization of each attributes orientation and b. Visual attributes comparison of opinions on two products

comparison based on attribute, as shown in Fig. 2b. Therefore, when users encounter difficulties in choosing two kinds of products, they can easily make the informed decision using our system function.

### FINE-GRAINED SO ANALYSIS

The previous methods of fine-grained SO analysis often use latent characteristic of words after POS tagging, e.g., adjective words are normally used to express sentiment and product attributes are usually nouns/noun phrases in review sentences. But these methods are lack of semantic understanding for whole sentences. We will utilize SRL to realize shallow semantic analysis of sentence and implement a fine-grained SO analysis system.

**Semantic role labeling:** The Semantic Role Labeling (SRL) is to assign syntactic constituents with semantic roles (arguments) of predicates (most frequently verbs) in sentences. A semantic role is the relationship that a syntactic constituent has with a predicate. Typical semantic arguments include Agent, Patient, Instrument, etc. and also adjunctive arguments indicating Locative, Temporal, Manner, Cause, etc. It can be used in lots of natural language processing application systems in which some kind of semantic interpretation is needed, such as question and answering, information extraction, machine translation, paraphrasing and so on.

The example is shown as follows:

相机很漂亮,但镜头太小,电池时间太短。(English: The camera is very beautiful, but the lens is too small and the time of battery is too short)

After SRL, it is show as follows:

[相机 Arg0][很 ARGM-ADV][漂亮 V],[但 ARGM-DIS] [镜头 Arg0] [太 ARGM-ADV] [小 V], [电池时间 Agr0] [太 ARGM-ADV] [短 V]

Obviously, the words "相机", "镜头" and "电池时间" belong to attribute words of digital camera and their semantic roles are all Arg0. In addition, the words "漂亮", "小" and "短" belong to SO words corresponding to the former three attributes respectively and their semantic roles are all V ("漂亮", "小" and "短" belong to predicative adjective, which is tagged by VA).

And the roles of these degree words "很" and "太" are ARGM-ADV.

From the result of this sentence, it is easy to see that SRL is helpful to the fine-grained SO analysis. We can make sampling analysis by large amounts of data, obtaining and summarizing some association rules to help us realize the fine-grained sentiment analysis better.

**The construction of domain-based multi-dictionary:** Judging the SO of sentences is always based on the method of sentiment words dictionary (Hu and Liu, 2004b), but it is usually lack of the consideration that there are many expressions for degree words and negative words relatively. So, we build dictionaries for them separately. In addition, with the sufficient consideration of the commonness and particularity of sentiment polarity in different domains, we conduct a dictionary classification: Basic polarity dictionary, Special domain-based polarity dictionary. Moreover we specially design domain dictionaries to store attribute sets of different domains. The lexical information acquisition mainly refers to HIT IR-Lab Tongyici Cilin and the sentiment lexicon of HowNet.

**Basic polarity dictionary:** A large part of the sentiment words/phrases have clear positive and negative polarity independent of domains. For instance, positive words: "高兴", "快乐", "漂亮", "健康", "先进", "美丽" (English: delighted, happy, pretty, healthy, advanced, beautiful); negative words: "错误", "浪费", "失误", "讨厌", "差劲", "难看" (English: wrong, wasteful, mistake, hateful, poor and troublesome). Whatever positive or negative the word is, it has relatively strengths or weaknesses. With the knowledge of linguistics, as well as the reference to polarity quantification design (Quan and Ren, 2009; Liu and Seneff, 2009), we give each polarity word/phrase a score, the value is from -1 to +1, for example, "漂亮" (beautiful) is +1, "难看" (ugly) is -1. We firstly establish baseline sentiment words assigned polarity score respectively and use different method (such as PMI (Turney, 2002; Turney and Littman, 2003), words similarity calculation based on HowNet (Zhu *et al.*, 2006) to calculate word similarity between new word/phrase and baseline sentiment words and then obtain ultimate similarity score for new word/phrase through weighted calculation among these methods. This method is also used for other dictionary.

**Special domain-based polarity dictionary:** The polarity of some sentiment words is relative with the domains of evaluation objects. In different domains, the polarity

expressed is totally different. Like the word "长" (long) is often described as the long time of battery usage in the domain of digital camera, which is a positive word, but is often described as taking long time, taking long way in other areas. Therefore we refer to the related resources and establish the corresponding special domain-based polarity dictionary which needs to be aimed at the different domains specially.

**Domain-based attribute dictionary:** According to the different domains, the attributes of evaluation objects generally have large differences. As digital products, we generally concern about the attributes such as appearance, battery, pixel, etc. However, the attributes of cars are usually focused on engine, appearance, oil consumption, etc.

To definite the product attributes, we first establish the corresponding standard attribute words using artificial filtering and revision and then using word similarity calculation algorithm described earlier, we expand this dictionary for each attribute word.

**Degree word dictionary:** When polarity words/phrases are modified by degree word in viewpoint sentences, the sentiment intensity of polarity words/phrases will be changed. Degree word acts on polarity words/phrases directly and also affects sentiment intensity of polarity words. In this study, we have collected degree words using language knowledge and some probability statistic. In order to quantify the results of sentiment orientation, we assign a score to each degree word, which value is from 0 to 1.

**Negative word dictionary:** When there is a negative word in a viewpoint sentence, the sentiment orientation of the sentence will change. Such as the sentence "这部手机不好看" (English: "this mobile-phone is not beautiful"), the word "好看" (beautiful) is originally positive polarity word, but because of the negative word "不" in front, the sentiment orientation of this sentence turns positive to negative.

Due to limitation of negative words, we compile them from relative corpuses and online documents and manually analyze and extract negative words/phrases.

**Fine-grained SO analysis of product attribute based on SRL:** Due to heavy workload for artificial judgment, we just analyzed the SRL outcome of 300 reviews and made the statistics for the probabilities which role is an attribute and which role has sentiment. For the outcome, please refer to Table 1.

Table 1: The probabilities which role is an attribute and which role has sentiment

Probability role	ARGM-TMP	A0	A1	V	
As attribute	0.04	0.76	0.14	0.06	
Probability role	ARGM-ADV	A1	A0	A2	V
As setiments	0.02	0.04	0.16	0.01	0.77

When designing the judgment rules for sentiment item determination in sentences, we first use polarity lexicon to judge the sentiment of role "V" and then to judge role "A1", "A0", "A2" sequentially. If these roles are not found with the emotional words/phrases, this sentence is considered an objective one and if these roles are found with the emotional words/phrases, corresponding role for attribute is sought for sequentially in current sentence according to probability distribution shown in Table 1. Note that the situation a sentiment corresponds to multiple attributes must be dealt with, such as the sentence that "相机的外形和价格都不错" (The figure and price of camera are all good), the sentiment word "不错(good)" corresponds to two attributes "外形" and "价格(price)". From the results of SRL, "相机的外形和价格" (the figure and price of camera) is Arg0.

According to the above methods, we consider the sentence as basic processing object and each review (could including several sentences) as processing unit and focus the attribute-sentiment pairs in one sentence by using the statistic rules of semantic role and finally calculate SO value of each attribute by analyzing all reviews. The calculation formulas of SO value are shown as follows:

$$SO_{fi} = \begin{cases} \beta S_{fi} + \alpha(1 - \beta)D_{fi} & \text{if no negative words} \\ \beta S_{fi} + \alpha(1 - \beta)D_{fi} & \text{if degree words lie between} \\ \text{negative words and sentiment words} & \\ \alpha(\beta S_{fi} + (1 - \beta)D_{fi}) & \text{if negative words lie between} \\ \text{deg ree words and sentiment words} & \end{cases} \quad (1)$$

where,  $S_{fi}$  denotes the orientation value of sentiment words/phrase related to attribute  $fi$  and  $D_{fi}$  denotes the value of degree words related to  $S_{fi}$ , the value of  $\beta$  is related with linguistic rule, in our system  $\beta$  is assigned to 0.6. If  $S_{fi}$  is a positive word,  $\alpha$  is assigned to 1, if  $S_{fi}$  is a negative word, then  $\alpha$  is assigned to -1.

$$SO_{fi} = \frac{\sum_{j=1}^n SO_{fi}}{\sum_{j=1}^n \text{count}(fi)} \quad (2)$$

(j is number of reviews)

Where  $SO_{fi}$  denotes the orientation value of attribute  $fi$  of current review processed and  $\overline{SO_{fi}}$  denotes the average value of all  $SO_{fi}$ . Using formula Eq. 1 and 2, we can easily calculate the tuple  $(SO_{f1}, SO_{f2}, \overline{SO_{f3}}, \dots, \overline{SO_{fn}})$ , which is sentiment orientation value of all attributes.

**EXPERIMENT RESULTS**

SRL tools are provided by NLP lab of Soochow University, which outperforms the state-of-the-art ones (Li *et al.*, 2009). We use Java to develop our demo system. The data of experiment are crawled from taobao.com (a famous Chinese C2C E-Commerce platform). We conducted our experiments using the customer reviews of two types of digital camera. Because of the too much work for artificial judgments, we, respectively extract 200 reviews of two types from reviews database which has been compiled previously. We arrange three people making artificial judgments for all 400 reviews and select the first six ones as object attributes from the domain-based attribute dictionary, such as "外形", "图像", "电池", "价格", "镜头" and "重量" (English: appearance, image, battery, price, lens and weight). We mainly implement attribute identification and calculation of SO value by using SRL tool. The orientation value is from -1 to +1, +1 is the highest positive and -1 is the largest negative. We allow the deviation threshold of orientation value of artificial judgment for 0.3 on the condition of correct polarity judgment (because the SO value judgment of each person can not be completely same, we allow the deviation in some extent). After their work is finished, some inconsistency of attribute identification and calculation of SO value among three people's results are found. The main reason lies in the mistake of manual judgments' and grammar errors of sentence itself. Through group discussion and modification, the consistency finally achieves 100% under the deviation threshold. All the attributes in each review are identified, calculating the corresponding average SO value, which is regarded as gold standard.

This system mainly realizes two function modules:

- Attribute-based fine-grained SO analysis of each product reviews and visual friendly display of results
- The multi-attribute visual comparison based on the reviews of similar products

In order to measure the efficiency of data processing, we mainly evaluate from two aspects:

**The effectiveness of attribute identification:** We analyze the effectiveness of attribute identification by comparing with POS-based method.

Table 2: The effectiveness comparison of attribute identification

Product name	SRL-based method		POS-based method	
	P	R	P	R
Digital camera 1	58.1	72.9	56.5	66.2
Digital camera 1	59.8	69.3	57.3	65.6
Average value	58.9	71.2	56.9	65.9

Table 2 shows that SRL-based method performed better than POS-based one. Because only noun/noun phrase (NP) is usually identified as attribute in POS-based method. However four roles are determined by probability in SRL-based method, meantime A0 even can be assigned to complex clause or different syntax structure, so SRL-based method can find more candidate attribute in corpus.

Because our approach is not sophisticated, there are some limitations: 1) there is pronoun resolution which needs to be processed in the reviews, needing to use natural language processing technology to improve efficiency; 2) the syntax of reviews is often incomplete, sometimes being lack of apparent attribute words, which need to strengthen the association rules; 3) the results of SRL have some errors, e.g. some incorrectness of Chinese word segmentation, which directly cause the decreasing of accuracy of attribute identification.

**Calculation for sentiment orientation of attributes:**

Firstly we analyze the identification effectiveness of sentiment words/phrases through recall and precision and then calculate the accuracy of attribute sentiment orientation. We allow the inaccuracy threshold of orientation value for 0.3 on the condition of correct polarity judgment.

Table 3 shows the evaluation results of the other two procedures: sentiment words/phrases identification and attribute orientation accuracy. According to statistical probability shown in Table 1, we judge relative role for sentiment words/phrases in sequence. The average precision of sentiment words/phrases identification is nearly 70%. The average recall of sentiment words/phrases identification is 72%. In the meantime, if correct identification of the attribute and sentiment words/phrases, the accuracy probability of orientation judgment is more than 90%.

We can see that our techniques are very promising, but we also note three main limitations of our system (1) sentiment words/phrases are not abundant, the sentiment lexicon needs to improve (2) the ways of expressing language diversify with some special expression methods such as enantiosis and irony, which need to be strengthened and (3) the expression method of sentiment orientation sometimes can realize through the comparison for different objects, such as "佳能相机的电池比它好" (the battery of Canon camera is better than it), the

Table 3: Results of sentiment words/phrases identification and attribute orientation accuracy

Objects of evaluation	Sentiment words/phrases identification		Attribute orientation accuracy
	Precision	Recall	
DC1	0.644	0.732	0.892
DC2	0.613	0.714	0.922
Average	0.629	0.723	0.907

current evaluation object is it (Sony). Due to the lack of special processing for comparative sentence, there is a deviation for attribute SO analysis.

### CONCLUSIONS

In this study, we use the SRL tool to make shallow semantic analysis of the sentences. Through analyzing the labeling rules of semantic roles, we also put forward a novel method to resolve the problem of fine-grained analysis of sentiment orientation. Experiment results prove our method is feasible.

In the future work, we plan to further improve the following things: (1) to incorporate the other NLP technologies, especially the anaphora resolution technology, to this system, in order to enhance the identification rate of attributes; (2) to improve the design of association rules to identify more attribute-sentiment pairs and (3) to try to handle some special language expression forms, such as enantiosis, irony, etc.

### ACKNOWLEDGMENTS

The authors want to express their special thanks to the SRL Tools provided by NLP lab of Soochow University. This study is supported by Project No.Y1090688 under Zhejiang Provincial Natural Science Foundation of China and Project No.2008C13082 under the Science and Technology Department of Zhejiang Province, China.

### REFERENCES

Baccianella, S., A. Esuli and F. Sebastiani, 2009. Multi-facet Rating of Product Reviews. Springer Press, France, pp: 461-472.

Carenini, G., R. Ng and A. Pauls, 2006. Multi-document summarization of evaluative text. Proceedings of the 11th EACL 2006, July 22-23, Association for Computational Linguistics Press, Sydney, Australia, pp: 305-312.

Chaovalit, P. and L. Zhou, 2005. Movie review mining: A comparison between supervised and unsupervised classification approaches. Proceedings of the 38th Hawaii International Conference on System Sciences, (HICSS'05), Hawaii, pp: 1-9.

Danescu-Niculescu-Mizil, C., G. Kossinets, J. Kleinberg and L. Lee, 2009. How opinions are received by online communities: a case study on amazon.com helpfulness votes. Proceedings of the WWW, April 20-24, ACM Press, Madrid, Spain, pp: 141-150.

Dave, K., S. Lawrence and D.M. Pennock, 2003. Mining the peanut gallery: opinion extraction and semantic classification of product reviews. Proceedings of the WWW, May 20-24, ACM Press, Budapest, Hungary, pp: 519-528.

Hu, M. and B. Liu, 2004a. Mining and summarizing customer reviews. Proceedings of the Tenth ACM SIGKDD International Conference on Knowledge Discovery and Data, Aug. 22-25, ACM Press, Washington, USA., pp: 168-177.

Hu, M. and B. Liu, 2004b. Mining opinion features in customer reviews. Proceedings of the 19th National Conference on American Association for Artificial Intelligence, July 25-29, AAAI Press, San Jose, California, pp: 755-760.

Kim, S.M. and E. Hovy, 2006. Extracting opinions, opinion holders and topics expressed in online news media text. Proceedings of the Workshop on Sentiment and Subjectivity in Text, July 22, Association for Computational Linguistics Press, Sydney, Australia, pp: 1-8.

Ku, L.W., T.H. Huang and H.H. Chen, 2009. Using morphological and syntactic structures for chinese opinion analysis. Proceedings of the Conference on Empirical Methods in Natural Language Processing, Aug. 6-7, Association for Computational Linguistics Press, Singapore, pp: 1260-1269.

Li, J.H., G.D. Zhou, H. Zhao, Q.M. Zhu and P.D. Qian, 2009. Improving nominal SRL in Chinese language with verbal SRL information and automatic predicate recognition. Proceedings of the Conference on Empirical Methods on Natural Language Processing, Aug. 6-7, Association for Computational Linguistics Press, Singapore, pp: 1280-1288.

Liu, J.J., Y.B. Cao, C.Y. Lin, Y.L. Huang and M. Zhou, 2007. Low-quality product review detection in opinion summarization. Proceedings of the Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning, June 28-30, Association for Computational Linguistics Press, Prague, Czech Republic, pp: 343-342.

Liu, J.J. and S. Seneff, 2009. Review sentiment scoring via a parse-and-paraphrase paradigm. Proceedings of the Conference on Empirical Methods on Natural Language Processing, Aug. 6-7, Association for Computational Linguistics Press, Singapore, pp: 161-169.



- Miao, Q., Q. Li and R. Dai, 2009. Amazing: A sentiment mining and retrieval system. *Expert Syst. Appl.*, 36: 7192-7198.
- Mullen, T. and N. Collier, 2004. Sentiment analysis using support vector machines with diverse information sources. *Proceedings of the Conference on Empirical Methods on Natural Language Processing*, July 25-26, Association for Computational Linguistics Press, Barcelona, Spain, pp: 412-418.
- Narayanan, R., B. Liu and A. Choudhary, 2009. Sentiment analysis of conditional sentences. *Proceedings of the Conference on Empirical Methods on Natural Language Processing*, Aug. 6-7, Association for Computational Linguistics Press, Singapore, pp 180-189.
- Nasukawa, T. and J. Yi, 2003. Sentiment analysis: Capturing favorability using natural language processing. *Proceedings of the 2nd International Conference on Knowledge Capture*, Oct. 28-31, ACM Press, Whistler, BC., pp: 70-77.
- Pang, B., L. Lee and S. Vaithyanathan, 2002. Thumbs up? Sentiment classification using machine learning techniques. *Proceedings of the Conference on Empirical Methods on Natural Language Processing*, July 6-7, Association for Computational Linguistics Press, Philadelphia, USA., pp: 79-86.
- Pang, B. and L. Lee, 2005. Seeing stars: Exploiting class relationships for sentiment categorization with respect to rating scales. *Proceedings of the ACL*, June 17, Association for Computational Linguistics Press, Michigan, pp: 115-124.
- Popescu, A.M., B. Nguyen and O. Etzioni, 2005. Extracting product features and opinions from reviews. *Proceedings of the HLT/EMNLP*, Oct. 6-8, Association for Computational Linguistics Press, Vancouver, BC., Canada, pp: 32-33.
- Quan, C.Q. and F.J. Ren, 2009. Construction of a blog emotion corpus for chinese emotional expression analysis. *Proceedings of the Conference on Empirical Methods on Natural Language Processing*, Aug. 6-7, Association for Computational Linguistics Press, Singapore, pp: 1446-1454.
- Subramanian, S.K. and N. Ramaraj, 2007. Automated classification of customer emails via association rule mining. *Inform. Technol. J.*, 6: 567-572.
- Tsou, B.K.Y., R.W.M. Yuen, O.Y. Kwong, T.B.Y. Lai and W.L. Wong, 2005. Polarity classification of celebrity coverage in the Chinese Press. *Proceeding of the International Conference on Intelligence Analysis*, May 2-6, McLean VA, pp: 1-6.
- Turney, P.D. and M.L. Littman, 2003. Measuring praise and criticism: Inference of semantic orientation from association. *Proc. ACM Trans. Inform. Syst.*, 21: 315-346.
- Turney, P.D., 2002. Thumbs up or thumbs down? Sentiment orientation applied to unsupervised classification of reviews. *Proceedings of the 40th Annual Meeting on Association for Computational Linguistics*, July 11-12, Philadelphia, USA., pp: 417-424.
- Umer, M.F. and M.S.H. Khiyal, 2007. Classification of textual documents using learning vector quantization. *Inform. Technol. J.*, 6: 154-159.
- Whitelaw, C., N. Garg and S. Argamon, 2005. Using appraisal groups for sentiment analysis. *Proceedings of the 14th ACM International Conference on Information and Knowledge Management*, Oct. 31-Nov. 5, ACM Press, pp: 625-631.
- Zhu, Y.L., J. Min, Y.Q. Zhou and X.J. Huang, 2006. Semantic orientation computing based on How Net. *J. Chinese Inform. Proc.*, 20: 14-20.