Feature-Level Sentiment Analysis Based on Rules and Fine-Grained Domain Ontology†

Wei Wei*, Yi-Ping Liu** and Lei-Ru Wei***

*Zhongnan University of Economics and Law, Big Data Institute, School of Statistics and Mathematics, Wuhan, China, < vivin@zuel.edu.cn >
**Wuhan University, School of information management,
Wuhan, China,
< 2856867730 @qq.com>
***Zhengzhou University of Light Industry, School of Economics and Management,
Zhengzhou, China, <weileiru83@163.com>

Wei Wei currently works as a teacher at the department of Big Data Institute, Zhongnan University of Economics and Law in China. She obtained her PhD degree from the School of Information Management, Wuhan University and attained her master's and bachelor's degrees in management information system. She does research in knowledge organization and intelligence system. Her research interests also include text mining.

Yi-Ping Liu received the master degree in management science and engineering from the School of Information Management, Wuhan University in 2017. She got her bachelor's degree at Shandong University. She is now working as a new media analyst and translator in Beijing. Her research interests lie in natural language processing and sentiment analysis.

Lei-Ru Wei is a lecture at the department of Economics and Management, Zhengzhou University of Light Industry in China. She holds a PhD degree in School of Information Management from Wuhan University. She is currently working on sentiment analysis and data mining. Her research interests also include knowledge organization.

Wei, Wei, Yi-Ping Liu and Lei-Ru Wei. 2020. "Feature-level Sentiment Analysis Based on Rules and Fine-Grained Domain Ontology." *Knowledge Organization* 47(2): 105-121. 44 references. DOI:10.5771/0943-7444-2020-2-105.

Abstract: Mining product reviews and sentiment analysis are of great significance, whether for academic research purposes or optimizing business strategies. We propose a feature-level sentiment analysis framework based on rules parsing and fine-grained domain ontology for Chinese reviews. Fine-grained ontology is used to describe synonymous expressions of product features, which are reflected in word changes in online reviews. First, a semiautomatic construction method is developed by using Word2Vec for fine-grained ontology. Then, feature-level sentiment analysis that combines rules parsing and the fine-grained domain ontology is conducted to extract explicit and implicit features from product reviews. Finally, the domain sentiment dictionary and context sentiment dictionary are established to identify sentiment polarities for the extracted feature-sentiment combinations. An experiment is conducted on the basis of product reviews crawled from Chinese e-commerce websites. The results demonstrate the effectiveness of our approach.

Received: 17 July 2019; Revised: 13 December 2019; Accepted: 24 January 2020

Keywords: sentiment analysis, feature-level, words, ontology extraction

† This work is supported in part by the Fundamental Research Funds for the Central Universities (grant nos. 412/31511941203).

1.0 Introduction

People's lifestyles have changed in the era of mobile internet. Almost nothing is separable from the internet, including food, clothing, and transportation. Everyone is both a recipient and a potential provider of information. For example, users can access required material or information services via the internet, such as online shopping, movie or music downloads, social interactions, and information browsing. Moreover, users can express their opinions and experiences through the internet, such as product reviews on e-commerce websites, movie reviews on film websites, and news reviews on social networking websites. Because of the large number of users, review data are increasing exponentially. These visible data are only the tip of iceberg, and a large part of the value is hidden at a deeper level. Mining and utilizing these reviews are of great significance

By analyzing the sentiment of user reviews, website operators or merchants can analyze the pros and cons of a product, speculate on users' preferences to develop a reasonable marketing strategy, and propose plans to improve the product's reputation and profitability. Producers of literary works such as movies (Kumar et al. 2019) and the tourism industry (Afzaal et al. 2019) can also understand popular trends and users' perceptions for the works. These reviews, which reflect users' opinions and sentimental tendencies, are increasingly of great value.

Review mining and sentiment analysis have become hot topics in both academia and industry, with many scholars having contributed excellent research. However, compared with abundant English resources, Chinese corpora for sentiment analysis are relatively limited (Chen and Huang 2019). Due to the complexity of Chinese text and nonstandard expressions in web reviews, some issues and difficulties in research on sentiment analysis yet remain. For example, nonstandard punctuation, unreasonable grammatical structure, and typos are abundant, but the current dependency parser can identify only standardized sentence elements. In feature-level sentiment analysis, most studies fail to notice the contextual specificity of sentiment words, and some sentiment words are often specific to only a certain product feature. In many studies that combine ontologies for feature extraction, traditional domain ontologies are based on mostly standardized professional terminology. Reviews frequently feature a large number of nonprofessional vocabularies and colloquial expressions that are not present in the ontologies and, therefore, ignored, which reduces the accuracy of sentiment analysis. In addition, the automated or semiautomatic construction of ontologies has always been a difficult aspect of research.

To address these issues, this study proposes a domain ontology construction and sentiment analysis method based on review mining. Fine-grained domain ontology for review mining is proposed to solve the issue of different types of synonymous or irregular descriptions of product features in Chinese reviews. We propose a domain ontology construction method based on Word2Vec. With assistance from machine learning to sort out fine-grained description words of product features, semiautomatic construction of product ontology is realized. Feature-level product review sentiment analysis involves two key steps in our study: feature extraction and sentiment classification. For feature extraction, rules parsing and domain ontology are used to extract features as explicit or implicit. For sentiment classification, we construct a domain sentiment dictionary and a context sentiment dictionary to overcome the defects of the existing general sentiment dictionary.

The remainder of this paper is organized as follows: Section 2 reviews previous research on review sentiment analysis, product feature extraction, and product review domain ontology. Section 3 explains the proposed approach and gives a detailed description of domain ontology and sentiment analysis. Section 4 presents the experiments and discusses the results. Section 5 concludes the paper and provides directions for future work.

2.0 Literature review

2.1 Feature-level sentiment analysis

Systematic research on sentiment analysis began with Turney's work in 2002. He used unsupervised machine learning algorithms to classify reviews into thumbs up and thumbs down (Peter 2002). In later research, sentiment analysis was divided into three levels: chapter-level sentiment analysis, sentence-level sentiment analysis, and feature-level sentiment analysis.

The main task of chapter-level sentiment analysis is to conduct sentiment classification of an entire document. Research ideas on this process can be divided into two types: those based on sentiment knowledge and those based on machine learning (Arruda et al. 2017). Sentence-level sentiment analysis divides an entire document into sentences and uses individual sentences as the object of the sentiment analysis. Review texts usually include two types of statements: subjective sentences and objective sentences. Sentences containing the users' sentiment tendencies are considered to be subjective sentences, while the objective sentences contain the users' descriptions of a certain target without emotional sentiment. As research progressed, people gradually discovered that many objective sentences describe a certain fact but still have sentiment tendencies. Therefore, the task gradually evolved from the recognition of subjective sentences to the recognition of sentiment sentences.

Feature-level sentiment analysis is also known as aspect-level analysis and is a fine-grained model of sentiment analysis that deals with determining the opinion or sentiment tendencies intended by social media users about a specific feature (aspect) of a product, service, or other entity (Medhat et al. 2014). This type of analysis usually includes identification of the opinion holders, extraction of the evaluation objects, extraction of sentiment words, and extraction of sentiment evaluation units. The opinion holder is the initiator of the sentiment opinion. The evaluation object is the target of the sentiment word in sentences, which may be the product feature. This technique will be detailed in Section 2.2. Sentiment words are the words that contain sentimental information in sentences, mostly adjectives and verbs. The sentiment evaluation is extracted as a mutual auxiliary unit rather than an evaluation object and a sentiment word separately. This is important because sentiment words do not always indicate sentiment tendencies when they appear alone, and the same sentiment word may have different sentiment tendencies when applied to different evaluation objects. According to Hu (Hu and Liu 2004; Toqir and Yu 2016), a feature-level sentiment analysis task can be divided into three main subtasks: feature extraction, sentiment lexicon analysis, and opinion summarization.

2.2 Product feature extraction

Features are the objects described by sentiment words in product reviews. For example, in the review "The screen of iPhone X is very large," the product feature is "screen." To develop and evaluate sentiment analysis at the feature level, feature extraction is a crucial process that can be either explicit or implicit. The feature is considered explicit if it is mentioned explicitly in the review sentences; otherwise it is considered implicit (Hu and Liu 2004).

Methods for the extraction of explicit features can be divided into two types: rule-based methods and machine learning-based methods. Li et al. (2010) improved a feature-mining method based on English reviews and applied it to Chinese reviews; a series of rules were proposed to define noun phrases, and association rules were used to mine Chinese product features. Wouter et al. (2014) proposed a method of matching the syntactic dependency path among different words in a sentence. To identify product features and their opinion words, ten handcrafted dependency paths were defined. The superiority of this method is that it requires only a small seed set, whereas other classifiers require a large trained corpus (Schouten and Frasincar 2016). The application of machine learning in explicit feature extraction can be divided into sequence models and topic models. The main principle of the sequence model is that a sentence is a grammatical relationship that connects words, so the feature extraction is regarded as a sequence tag task. The main research methods of sequence models are hidden Markov model (Wei 2009), conditional random field (Tang 2019), and maximum entropy model (Huang and Sun 2017), while topic models include latent Dirichlet allocation (LDA) (Cui et al. 2018) and probabilistic latent semantic analysis (PLSA) (Zhou 2016). Since LDA is designed to operate on the document level, employing it for much finer-grained feature-level sentiment analysis is not straightforward. Tang et al. (2019) proposed a joint aspect based sentiment topic model that extracted multi-grained aspects and opinions through the simultaneous modeling of aspects, opinions, sentiment polarities, and granularities by means of supervised learning, while using maximum entropy to improve performance.

The extraction of implicit features is of great help for sentiment analysis and can greatly improve the recall rate. Implicit feature extraction techniques can be classified into three approaches, namely, unsupervised, semi-supervised, and supervised (Mohammad et al. 2018). Because unsupervised methods do not require data annotation for implicit features or any sort of training, they are the most frequently used methods for feature extraction in previous research works. Commonly used methods for unsupervised implicit feature extraction include dependency parsing (Zainuddin et al. 2016), association rule mining (Mankar and Ingle 2015), ontology (Lazhar and Yamina, 2017), topic modeling (Rana et al. 2018), co-occurrence (Prasojo et al. 2015), and rule-based (Wan et al. 2018). Liao et al. (2019) focused on the recognition of fact-implied implicit sentiment at the sentence level. A multi-level semantic fusion method was proposed to learn the features. Semi-supervised implicit feature extraction utilizes both labeled and unlabeled data to extract implicit features from the corpus or require little training. Xu et al. (2015) extracted implicit features using both support vector machine (SVM) and explicit topic models. Semi-supervised methods are still not sufficiently explored compared with other types of methods. The supervised methods require labeled data and cannot be generalized easily. Schouten and Frasincar (2014) labeled the dataset with implicit features and computed the co-occurrence score between the labeled implicit features and other dictionaries. Hajar and Mohammed (2016) used a hybrid approach of the labeled corpus, WordNet, and Naive Bayes classifier for implicit feature extraction.

2.3 Product review domain ontology

Domain ontology is a professional ontology that describes the concepts and the relationship between concepts in a specific domain. The product review domain ontology is an ontology built on the reviews of a certain product, which represents the concepts, attributes, and relationships of a product domain.

Some scholars have applied ontological approaches to sentiment analysis. Yin et al. (2013) established a review mining model to identify feature-sentiment combinations based on domain ontology. Tang et al. (2016) constructed product feature ontology to classify feature words and then identified implicit features by calculating the collocation weights between sentiment words and product features. Santosh et al. (2016) presented an ontology to improve the performance of LDA. They used the ontology to identify appropriate features after clustering and showed that the accuracy of the feature extraction greatly improved. Sophie et al. (2018) focused on semantic enrichment by employing ontology features in determining the sentiment value of a given pair of review and feature. Chen et al. (2018) designed a text analytics framework to assess secondhand sellers' reputations and developed a feature extraction method that combined the results of domain ontology and topic modeling to extract topical features. Farman et al. (2019) proposed an ontology, LDA-based and word embedding approach for sentiment classification using ontology-generated topics and features.

According to Schouten and Frasincar (2016), because semantic methods naturally combine common sense knowledge with domain knowledge, ontologies are being used to improve feature detection. Combining concept-centric semantic methods with the power of machine learning will give rise to algorithms that can reason with language and concepts at a whole new level. However, the ontologies mentioned above are less suitable for the diverse and flexible online language expressions of web users in regard to review mining or social media analysis and basically constructed manually. We will improve upon this aspect.

3.0 Proposed approach

The overall architecture of the proposed approach is divided into two parts: fine-grained domain ontology construction based on Word2Vec and sentiment analysis based on rules parsing and domain ontology. These two parts are respectively shown in the upper half and lower half of Figure 1 and will be discussed in detail in Section 3.1 and Section 3.2.

3.1 Fine-grained domain ontology construction based on Word2vec

As the importance of review mining increases, the use of existing ontologies with standard terminology for knowledge representation and reasoning becomes less suitable

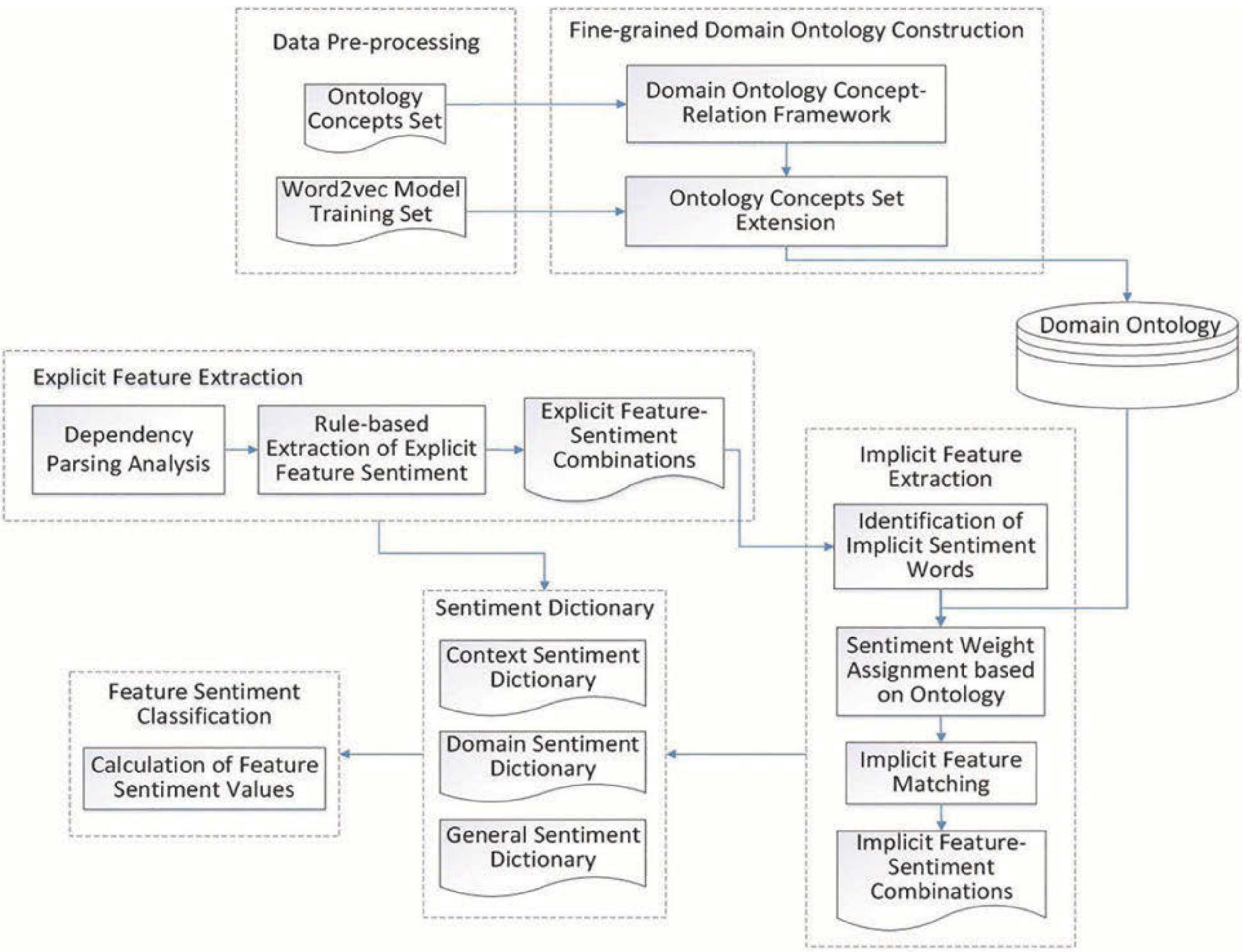

Figure 1. Overall architecture for domain ontology construction and sentiment analysis.

for the diverse and irregular expressions of social media users.

The architecture for domain ontology construction and sentiment analysis richness and complexity of Chinese ideograms makes these expressions more flexible and changeable. For example, the same entity or attribute may be represented in a variety of ways, and the inclusion of internet slang and buzzwords makes these expressions richer. A great breadth of informal vocabularies and spoken language is contained in web reviews, and different users may use different manners to describe the same product feature. According to word changes in user reviews, the domain ontology is no longer limited to professional vocabularies. For example, "film" in the review "This film is great" is synonymous with the concepts "movie," "motion picture," or "cinema." Each concept can have its own synonym set. In studies of review mining based on lexicons or ontologies, these words that do not appear in the dictionaries or ontologies are often missed and ignored. As a result, the effect of review mining is reduced. The fine-grained domain ontology (FDO) for review mining proposed in this paper can solve this problem. FDO is used to describe synonymous expressions of product features, which are reflected word changes in reviews. The ontology constructed is designed for the diverse and flexible online language expressions of social media users in regard to review mining or social media analysis.

Based on deep learning, we propose a domain ontology construction method based on Word2Vec. Through machine learning assisting in sorting fine-grained description words of product features, semi-automatic construction of product ontology is realized. The architecture for building a FDO based on Word2Vec is shown in the upper part of Figure 1. The process is divided into three main modules, namely, data pre-processing, construction of a domain ontology concept-relation framework, and an ontology concept words set extension based on Word2Vec. Taking the phone product as an example, we build a fine-grained phone product domain ontology (phone FDO) based on review mining.

The phone FDO adopts a semiautomatic construction method that includes manual construction of the domain ontology concept-relation framework, an automatic extension of the domain ontology concept words set, and the sentiment assignment of the concept words. The purpose of automatic extension of the domain ontology concept words set is to automatically form synonymous relationships in the ontology. The sentiment assignment of the concept words is an extension of sentiment description for ontology concepts, which will be described in Section 3.2.2.

3.1.1 Data pre-processing

This module collects the required corpus data from the internet and processes the data for cleaning, noise reduction, and word segmentation. This section includes two data sets: the ontology concepts set and the word vector model training set.

The ontology concepts set is used to construct the concept-relation framework of the phone domain ontology. The construction of product domain ontology requires some authoritative expertise that can represent the concepts of the product and the relationships between concepts. Domain concepts and relationships between concepts can be extracted from authoritative and specialized data, such as HowNet's Chinese structural information base, product parameter descriptions of e-commerce websites, and professional portals.

The word vector model training data set is used to provide data support for the Word2Vec tool. The training mode used by the Word2Vec model is the Skip-Gram in which we can obtain the context or similar words associated with an entered word. The Skip-Gram requires a large corpus for model training, so we must collect sufficient review data. Because the training of the Word2Vec model is based on words, word segmentation must be applied to the training corpus. To facilitate a unified language style between the training corpus and experimental data, the same type of mobile phone review data are selected as the Word2Vec training data.

3.1.2 Domain ontology concept-relation framework

This module is used to construct a concept-relation framework for the phone product ontology. The relationships between concepts in the phone product ontology mainly include synonymous relationships and subordinate relationships. For example, a synonymous relationship exists between "mobile phone" and "phone," and an overall-partial relationship exists between "phone" and "screen." The "screen" and "resolution" is the upper and lower position relations associated with the attribute.

The construction process of the concept-relation framework for phone domain ontology includes the seed concept words, the upper and lower positions, and the relationships between concept words. Phone product parameters and product manual data collected via an electronic product portal are summarized and extracted to form the seed concept words, which contain more specialized classification and descriptions of parameters, functions, and components in the phone field. Then, we refer to HowNet's conceptual subordinate relationship document, which includes the subordinate relationships between entity class, attribute class, and instance class, and the upper and lower position relationships between these classes. After these two steps, the concept-relation framework of the phone domain ontology is obtained.

3.1.3 Ontology concept words set extension

This module utilizes the Word2Vec tool to train the word vector model and then extends the seed concept words set. Word2Vec is a deep learning-based tool developed by Google and is an effective auxiliary for the semiautomatic construction of domain ontology.

An iterative algorithm is adopted to obtain an extended-words set of the seed concept words set. Figure 2 presents the process, and the details are as follows:

1) Initialize the input vocabularies by the seed concept words set obtained in Section 3.1.2.
2) Call the Word2Vec word vector model. Generate the concept words candidate set by setting a similarity threshold to obtain words with high similarities as the input vocabularies.
3) Iteratively input the difference sets between the output vocabularies and input vocabularies into the word vector model to include words larger than the similarity threshold and obtain the concept words candidate set.
4) Set the termination condition for iteration to end the algorithm.

The concept words candidate set extracted by the iterative algorithm should be further filtered. The Domain Membership Degree (Yu and Dang 2009) is used to analyze each candidate words. The basic idea of this method is that if a candidate word has a higher probability of appearing in the foreground corpus than in the background corpus and is evenly distributed in the foreground corpus, then the word is an ontology concept word in this field. Then, the words are arranged in descending order of Domain Membership Degree. According to the scale of ontology and the popularity of words, the words that best reflect product properties are identified as fine-grained ontology concept words. Finally, the resulting concept words set are written as the extended words of the seed concepts into the ontology conceptual framework.

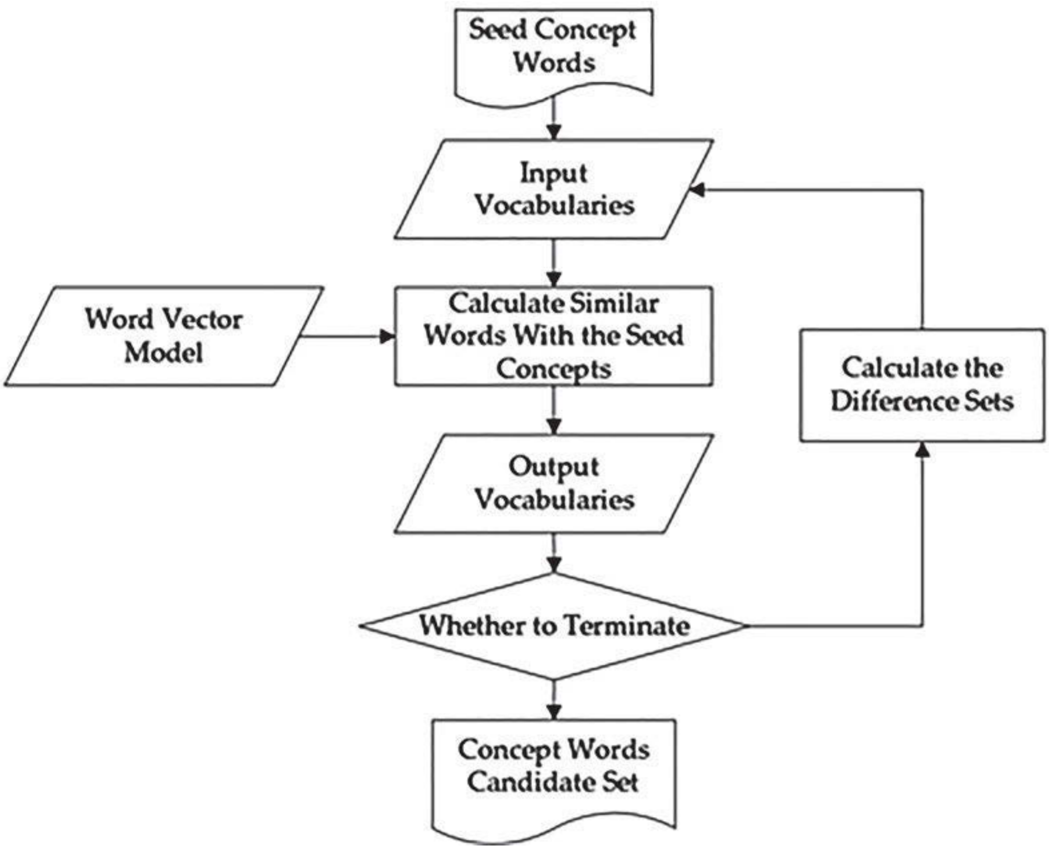

Figure 2. The process of ontology concept words set extension.

3.2 Feature-level sentiment analysis based on rules parsing and fine-grained domain ontology

The previous section introduced the construction process of fine-grained phone product ontology, and this section conducts feature-level sentiment analysis based on rules dependency syntax and this domain ontology. The research model is presented in the lower half of Figure 1. The feature-level sentiment analysis in this study includes an explicit feature extraction module based on rules parsing, an implicit feature extraction module based on fine-grained domain ontology, sentiment dictionary construction module, and the sentiment classification module.

3.2.1 Explicit feature extraction

This module performs dependency syntax analysis on the preprocessed review sentences to obtain semantic dependency relationships between words. All the combinations of feature word and sentiment word in a sentence are regarded as feature-sentiment combinations. By setting a series of extraction rules, the feature-sentiment combinations, including the evaluation object and the sentiment word, are extracted from the dependency relationship. However, the results of the dependency syntax analysis indicate that not all the dependency relationships can extract valid feature-sentiment combinations. Most of the evaluation units exist in only a small number of dependencies. Through the comparative analysis of a large number of Chinese dependency relationships and original sentences, relationships containing valid feature-sentiment combinations are identified that exist only in the dependency relationships, such as the subject-verb relationship (SBV), attribute relationship (ATT), adverbial relationship (ADV), left adjunct relationship (LAD), and right adjunct relationship (RAD).

Therefore, we develop the following series of rule algorithms for extracting Chinese explicit feature-sentiment combinations. The inputs of the following algorithms are all dependency parser documents after segmentation and syntactic analysis of the review corpus. Each line contains a dependency relationship of a word. The total number of lines is set to n, where the first line is t_1 and the nth line is t_n. Each line is treated as a list of three elements: the first element t[0] is the original word, the second element t[1] expresses the dependency relationship, and the third element t[2] represents the word indicated by the dependency relationship. The extraction rules are as follows.

I. ASA extraction rule

The ASA extraction rule is a combination of AS and SA. AS is used mainly for the reviews containing noun phrases, that is, ATT+SBV dependencies; SA is used mainly for the reviews containing sentiment adverbs, that is, SBV+ADV dependencies.

The first part is AS. In Chinese product review sentences, we often see reviews such as "fast delivery speed" and "not enough screen sensitivity," where the feature words are noun phrases composed of two nouns, i.e., "delivery speed" and "screen sensitivity." Word segmentation models are not yet able to automatically recognize these phrases. After these corpora are segmented, noun phrases are often identified separately as two separate words, and thus noun phrases that should be distributed as a whole are assigned to different dependencies.

Then, the SA part, which can be viewed as an optimization of the AS part, is implemented. According to the analysis of a large number of review corpora, many user reviews contain sentiment adverbs such as "very" in "very big screen" and "unsatisfactory" in "unsatisfactory power failure." These sentiment adverbs often play a role in strengthening, weakening, or transforming the sentiment orientation in sentences. For example, "very" can strengthen sentiment, "a little" can slightly weaken sentiment, and "no" directly changes the sentiment. In Chinese syntactic parser, sentiment adverbs and sentiment words are usually divided into two words and thus often do not appear in the same dependency relationship. As a result, the extracted sentiment tendencies might differ from those of the original texts.

Therefore, negative adverbs in the ADV relationship must be extracted to ensure that the original sentiment tendencies are unchanged. Based on a large number of analyses and verifications, we propose the ASA (ATT+SBV+ADV) extraction rule.

ASA extraction rule:
a Set the number of document lines to t_i, i initialized to 1;
b If ATT∈t_i[1], SBV∈t_{i+1}[1], SBV∈t_{i+2}[1] // Determine whether t_i, t_{i+1}, and t_{i+2} meet the extraction rules
Then, extract t_i, t_{i+1} and t_{i+2}; perform step c;
Otherwise, i=i+1; return to step b;
c If t_i [2]=t_{i+1}[0], t_{i+1}[2]=t_{i+2}[2], t_{i+2}[0]= sentiment adverbs // Determine whether the elements in the lists of t_i, t_{i+1}, and t_{i+2} match the rules
Then, output feature-sentiment combinations, where t_i[0]+ t_{i+1}[0] is the feature word, t_{i+2}[0]+t_{i+2}[2] is the sentiment word; i = i + 1, loop step b;
Otherwise, i=i+1; return to step b;
d until i=n-2

II. AAS extraction rule

Because product reviews are unwritten language published by web users, the language style is casual, so some linguistic phenomena, such as omission of punctuation and irregular punctuation, exist. For example, in a Chinese sentence "！手机功能强大质量优越" (that means the mobile phone is powerful and superior in quality, but punctuation is missing from the sentence). In Chinese, "function" and "powerful" form a pair of feature-sentiment combinations, and "quality" and "superior" constitute a pair of feature-sentiment combinations. Due to the nonstandard user review sentences, no punctuation is placed between the two pairs of feature-sentiment combinations, making it impossible for the dependency parser to accurately identify the pairs. The feature words and the sentiment words within the same pair of feature-sentiment combinations are assigned to two dependency relationships. The unit "quality-superior" in the latter half of the sentence is identified as a SBV relationship, while the unit "function-powerful" in the first half of the sentence has not been extracted. Therefore, we need to set a rule to extract the word pairs from the first half of the sentence. After analyzing a large number of these Chinese structural dependencies, we propose the following AAS (ATT+ATT+SBV) extraction rules.

AAS extraction rule:

a. Set the number of document lines to t_i, with i initialized to 1

b. If ATT $\in t_i$ [1], ATT$\in t_{i+1}$[1], SBV$\in t_{i+2}$[1] // Determine whether t_i, t_{i+1} and t_{i+2} meet the extraction rules.
Then, extract t_i, t_{i+1} and t_{i+2}; perform step c;
Otherwise, i = i + 1; return to step b;

c. If $t_i[2] = t_{i+1}[2] = t_{i+2}[0]$ // Determine whether the elements in the lists of t_i, t_{i+1}, and t_{i+2} conform to the rules
Then, output feature-sentiment combinations, where $t_i[0]$ is the feature word, $t_{i+1}[0]$ is the sentiment word; i = i + 1; loop step b;
Otherwise, i=i+1; return to step b;

d. until i = n.

III. AS extraction rule

The mode of AS is included in both the ASA and AAS extraction rules, that is, a combination of AAT and SBV. The ASA rule considers the case of negative adverbs, and the AAS rule includes the case where multiple feature-sentiment combinations are parallel and undivided on the basis of AS. However, a more normal dependency pattern exists. These dependencies have noun phrases and no negative adverbs, and punctuation is more standardized. The dependent syntax analysis results are more accurate and thus can be directly extracted. However, because the combination modes set in the previous ASA and AAS rules already contain a part of the AS mode, direct extraction would result in duplication. Therefore, we add the judgment to exclude the first two rule modes in the AS rules. The specific extraction rules are as follows.

AS extraction rule:

a. Set the number of document lines to t_i, i initialized to 1;

b. If ATT $\in t_i$[1], SBV$\in t_{i+1}$[1],ADV$\notin t_{i+2}$ // When i=1, determine whether t_i, t_{i+1}, and t_{i+2} meet the extraction rules
Then, extract t_i,t_{i+1}; perform step d;
Otherwise, i=i+1; perform step c;

c. If ATT $\notin t_{i-1}$[1], ATT $\in t_i$[1], SBV $\in t_{i+1}$[1], ADV$\notin t_{i+2}$ // When i$\geq$ 2, determine whether t_{i-1}, t_i, t_{i+1},
and t_{i+2} meet the extraction rules
Then, extract t_i, t_{i+1}; perform step d;
Otherwise, i=i+1; perform step c;

d. If $t_i[2] = t_{i+1}[0]$ // Determine whether the elements in the lists of t_i and t_{i+1} match the rules
Then, output feature-sentiment combinations, where $t_i[0]$+ $t_i[1]$ is the feature word, $t_{i+1}[2]$ is the sentiment word; i=i+1; loop step c.
Otherwise, i=i+1; return to step c

e. until i=n-1

IV. SA extraction rule

The SA extraction rule algorithm allows the combined extraction of SBV and ADV relationships, which belong to the generalization based on ASA. In the ASA rules, we consider the situation that punctuation is not standardized. However, some reviews contain more standardized punctuation. The evaluation object and sentiment word are distributed in the same dependency relationship. In this case, only the combination of SBV and ADV can be extracted.

SA extraction rule:

a. Set the number of document lines to t_i, i initialized to 1;

b. If SBV$\in t_i$[1], ADV$\in t_{i+1}$[1] // When i =1, determine whether t_i, t_{i+1}, meet the extraction rules
Then, extract t_i, t_{i+1}; perform step d;
Otherwise, i=i+1; perform step c;

c. If ATT $\notin t_{i-1}$[1], SBV $\in t_i$[1],ADV $\in t_{i+1}$[1] // When i $\geq$ 2, determine whether t_{i-1}, t_i, t_{i+1} meet the extraction rules

Then, extract t_i, t_{i+1}; perform step d;
Otherwise, i=i+1; perform step c;
d. If $t_i[2]=t_{i+1}[2]$, $t_{i+1}[0]$= sentimental adverb // Determine whether the elements in the lists of t_i and t_{i+1} match the rules
Then, output feature-sentiment combinations, where $t_i[0]$ is the feature word, $t_{i+1}[0]+t_{i+1}[2]$ is the sentiment word; i=i+1; loop step c;
Otherwise, i=i+1; return to step c;
e. until i=n-1

V. SL+ extraction rule

LAD is the left additional dependency relationship. Usually, semantic relations such as juxtaposition, comparison, and selection can be expressed according to the keywords such as "and" and "or." Chinese product reviews frequently compare two products, for example, "iOS is more fluid than android" and "color and appearance both are very good." In this sentence pattern, one sentiment word often corresponds to multiple feature words, or multiple feature words may appear in the sentence with only one sentiment word. In the previous sentence, the sentiment word described by "fluid" is iOS; in the latter sentence, the feature words corresponding to the sentiment word "good" are the two words "color" and "appearance." Since the existing dependency parser is not able to accurately identify this Chinese phenomenon, some manual rules must be set to extract these dependencies.

After analyzing a large number of dependency parsing results, we find that the combination patterns of effective semantic components mainly follow the patterns SBV + LAD + SBV, SBV + LAD + ATT, and SBV + LAD + COO. The number of these combinations is small in the dependency parsing results; thus, the extraction rules are designed to extract these combinations. The proposed SL+ extraction rules are as follows.

SL+ extraction rule:
a. Set the number of document lines to t_i, i initialized to 1;
b. If SBV∈ $t_i[1]$, LAD∈$t_{i+1}[1]$, SBV∈$t_{i+2}[1]$ or ATT∈$t_{i+2}[1]$ or COO∈$t_{i+2}[1]$; // when i=1, determine whether t_i, t_{i+1}, and t_{i+2} meet the extraction rules
Then, extract t_i, t_{i+1} and t_{i+2}; i=i+1; loop step c;
Otherwise, i=i+1; return to step c;
c. Until i=n-2

In summary, the ASA extraction rule concentrates on the combination occurrence of noun phrases and negative adverbs. The AAS extraction rule focuses mainly on the feature-sentiment combinations that are assigned to different dependencies due to nonstandard punctuation. The AS extraction rule is used to identify the noun phrases combinations that do not contain negative adverbs. The SA extraction rule is used to extract negative adverb combinations without noun phrases. The SL+ extraction rule applies to the case in which one sentiment word corresponds to multiple feature words.

3.2.2 Implicit feature extraction

Implicit feature extraction includes identification of implicit sentiment words, sentiment weight assignment based on the ontology concept, and matching of implicit features. Some sentiment words in product reviews do not point to obvious product features. For example, in a common Chinese review, "It's too expensive for me," although the feature of "price" is described; from the perspective of Chinese sentences and words, the sentiment word "expensive" does not match the obvious product feature. We call such sentiment words implicit sentiment words. If the sentiment words in review sentences are not explicit sentiment words, they must be implicit sentiment words. Implicit sentiment words usually have two characteristics: sentiment words with clear expression of sentiment tendencies and sentiment words that do not match explicit evaluation objects. Therefore, we propose a method to identify implicit sentiment words based on explicit feature-sentiment words. First, the sentences containing the explicit feature-sentiment combinations are filtered out from the original reviews. Second, by performing segmentation and part-of-speech tagging on the filtered texts, verbs and adjectives are collected for each review to form the implicit sentiment words document. In Chinese, most sentiment words are adjectives or a combination of adverbs and adjectives. In the extraction of implicit sentiment words, we recognize adjectives by default.

To assign implicit sentiment words to their corresponding evaluation objects, we must first construct a corresponding relation library between the sentiment words and product features. According to the feature words, sentiment words and sentiment weights in the library, the most suitable feature words will be matched with the implicit sentiment words.

In Section 3.1, a semiautomatic method is used to construct a fine-grained phone product ontology that includes the concept of mobile phone and subordinate affiliation in phone products. Combining the collocation relationship between these feature words and the sentiment words, we propose a method for assigning sentiment values to the domain ontology concepts based on the weight of the explicit feature sentiment. The process is as follows.

1) Identify sentiment words from explicit feature-sentiment combinations and cluster according to different sentiment words.
2) In each cluster, the feature weight value is given according to the occurrence frequency of each feature word. The assignment formula of the feature weight value is the number of occurrences of feature words in the cluster divided by the total number of word pairs in the cluster. For example, in the cluster of the sentiment word "ugly," there are sixteen pairs of feature-sentiment combinations, and "color-ugly" appears four times. Thus, under the cluster of the sentiment word "ugly," the feature weight value of "color" is 0.25; "notch screen-ugly" appears twelve times, and thus, the feature weight value of "notch screen" is 0.75.
3) Add corresponding sentiment words and weights to each concept in the domain ontology based on the calculated feature weights. Notably, when a sentiment word is matched with multiple feature words, and these words are synonymous under the same concept in the FDO, the weight values should be added.
4) The feature word with the largest weight value in the ontology is assigned to the implicit sentiment word.

3.2.3 Sentiment dictionary

The sentiment dictionary consists of three parts: the General Sentiment Dictionary, the Domain Sentiment Dictionary, and the Context Sentiment Dictionary.

The General Sentiment Dictionary is a universal affective dictionary used in various fields. We use HowNet sentiment dictionary for sentiment analysis. The dictionary contains sentiment words and polarity indications represented by numbers. These general and nonspecific terms have shown their flaws in sentiment analysis for specific fields. For instance, words such as "flashback" and "broken screen" show up frequently in mobile phone reviews, but these sentiment words usually do not appear in the General Sentiment Dictionary. Thus, the Domain Sentiment Dictionary is indispensable when performing sentiment analysis for a specific field.

The construction of the Domain Sentiment Dictionary includes two key steps: the acquisition of domain sentiment words and the judgment of sentiment categories. We select product reviews from e-commerce websites as the source of the Domain Sentiment Dictionary. Analysis of a large number of corpora indicates that most sentiment words are adjectives, and a small number are verbs. Therefore, we set the recognition range of sentiment words as adjectives and verbs in the reviews corpus. First, the reviews corpus is segmented, and parts of speech are tagged; second, the adjectives and verbs are sorted according to their frequency of occurrence, and words whose frequency is greater than a certain threshold are extracted as sentiment benchmark words. Then, the word vector model trained by Word2Vec is used to find high-similarity words to expand the sentiment benchmark words. Finally, sentimental categories are assigned to each sentiment word. The structure of the Domain Sentiment Dictionary is shown in Table 1.

Sentiment words	Sentiment polarity
死机 (crash)	-1
闪退 (flashback)	-1
发烫 (run hot)	-1
黑屏 (black screen)	-1
卡顿 (stuck)	-1
网络延迟 (network delay)	-1
流畅 (smooth)	1
发黄 (yellowing)	-1

Table 1. Part of the Domain Sentiment Dictionary.

Moreover, some sentiment words belong to general sentiment words but in different contexts may show different sentiment polarities. For example, "fast" in "logistics is fast" is a positive sentiment, while in "power out too fast," it is a negative sentiment. Such sentiment words are usually included in the General Sentiment Dictionary and are often given a fixed sentiment tendency. If only one sentiment tendency is inclined to define such sentiment words in different contexts, the accuracy of the sentiment classification will inevitably be reduced. Therefore, a Context Sentiment Dictionary is needed to express the different sentiment polarities for such sentiment words when they are matched with different evaluation objects. For this sentiment dictionary, we manually sort the contextual sentiment words and their evaluation objects from mobile phone reviews and then mark the sentiment polarities for each match. The structure of the Context Sentiment Dictionary is shown in Table 2.

Feature-sentiment classification is the last step in sentiment analysis. The previously extracted explicit and implicit feature-sentiment combinations are summarized. One word pair per line represents a record. Each record is matched to the dictionaries in the following order: Contextual Sentiment Dictionary, Domain Sentiment Dictionary, and General Sentiment Dictionary. If a sentiment word is matched, the sentiment polarity is recorded. If the record matches a certain sentiment dictionary successfully, it is no longer matched against the next sentiment dictionary. All the combinations are clustered according to the feature, the sentiment values of all the records in the feature cluster are added, and the average value is obtained as the final sentiment score for the feature.

Sentiment words	Feature words	Sentiment polarity
大 (big)	屏幕可用率 (screen availability)	1
大 (big)	手机厚度 (phone thickness)	-1
高 (high)	价格 (price)	-1
高 (high)	性价比 (cost performance)	1
简单 (simple)	包装 (package)	-1
简单 (simple)	操作 (operating)	1
快 (fast)	掉电 (power out)	-1
快 (fast)	发货 (delivery)	1
慢 (slow)	功耗 (power consumption)	1
慢 (slow)	物流 (logistics)	-1

Table 2. Part of the Contextual Sentiment Dictionary.

Function	Camera, Photograph, Camera Type, Pixel Camera, Wide Angle, Telephoto, Video, Audio, GPS, Payment, APP, MP3, Entertainment, Game, Sound Effect
System and Hardware	RAM, ROM, Memory Capacity, OS, iOS, Android, WP, Symbian, Battery, mAh, Battery Capacity, Charger, Headset, Data Wire, USB
Appearance	Color, Size, Bar Phone, Clamshell, Slide Phone, Keyboard, Thickness, Weight, Material, Operation Type, Glass Body, Virtual Button Bar
Screen	Screen Size, Screen Style, Screen Color, Resolution, Screen Availability, Main Touch Screen, Touch Panel, MultiTouch, OLED, HD
Operator and Network	Bluetooth, Signal, Network Mode, CPU, Dual SIM, Operator, GSM, GPRS, CDMA, 3G, 4G, 5G, WCDMA, SIM Card, Volte, WiFi, TD-LTE, FDD-LTE
Services	Customer Service, After Sale, Warranty, Logistics, Delivery, Package, Three Guarantees Certificate, Return Policy, Invoice Nationwide Warranty

Table 3. Some important terms of the phone product ontology.

4.0 Experiment and results analysis

4.1 Data description

For the ontology concepts set, we obtained product parameters and manual data from professional portals, such as Mobile China and Pacific Internet. For the word vector model training set, we wrote Python programs to crawl nearly 300,000 mobile phone review data from Chinese websites, such as Jingdong and Taobao, and mobile phone forums.

Due to the different format of review text from various websites, some noise data, such as emoticons, advertisements, and links, were intermixed with the original data. After the preprocessing steps of cleaning and noise reduction, 250,760 pieces of review data remained as the training corpus for Word2Vec word vector model.

The experimental data of the feature extraction and sentiment analysis were obtained from Chinese reviews for the latest mobile phone product, iPhone X, in Jingdong Mall. After the preprocessing steps of cleaning and noise reduction, 10,000 phone review sentences with at least one sentiment word were selected for sentiment analysis.

4.2 Fine-grained domain ontology construction

The mobile phone product parameters and manual data crawled from the official electronic portal are summarized to obtain the seed concept words set of the phone product. Referring to HowNet's Chinese information structure library, the seed concepts are defined by the upper and lower positions and the relationships that form the conceptual relationship framework of the fine-grained phone product ontology. Table 3 shows a selection of important terms involved in the phone product ontology.Then, the concept of each node in the ontology is synonymously extended. We use the Word2Vec toolkit provided by the Python gensim module to train the corpus. Based on the similarity calculation for a large number of words, the similarity threshold is set to 0.634. Words with similarity greater than this threshold are used as synonymous extensions of the seed concept words. Finally, these words are arranged in descending order of Domain Membership Degree. According to the scale of the experimental ontology, the number of synonymous extended words of each ontology concept is set to no more than eight.

The phone product ontology "class" is based on the properties of the phone product. To design the class and class hierarchical structure, the general concepts are defined first, and then the defined concepts are specialized. The hierarchical structure of this ontology mainly includes several major classes, including "brand," "functions," "systems and hardware," "appearance," "screen," "operator and network," and "services." Each major class contains the corresponding subclass. Important attributes involved in the phone product ontology include "is_part_of," "is_attribute_of," and "has_appearance_of," among other object attributes and datatype attributes corresponding to various product parameters.

We adopt the knowledge engineering method (Natalya and Deborah 2001) to construct the phone product ontology. This method assumes that there is no absolutely correct way to model a certain domain, all solutions must be adapted for practical application, and the process of ontology development is one of continued iteration. The generated phone product domain ontology is visualized by Protégé5.2.

Protégé is an open java-based tool that integrates ontology editing and supports knowledge representation of class, class multiple inheritances, class properties, and class individuals. Part of the phone product ontology is shown in Figure 3. The fine-grained synonymous concepts are stored in the ontology by adding individuals. Figure 3 shows the synonymous concepts of "camera," including "photo," "picture," "lens," "shot," and "photography."

4.3 Feature-level sentiment analysis

Dependency syntax analysis is performed on the review texts. We adopt Pyltp, the implementation version of the Harbin Institute of Technology Language Platform (LTP) in Python, which provides rich and efficient natural language processing techniques, such as Chinese word segmentation, part-of-speech tagging, named entity recognition and dependency syntax analysis. Then, according to the five extraction rules proposed in Section 3.2.1, namely, ASA, AAS, AS, SA, and SL+, explicit feature-sentiment combinations are obtained, and the format is shown in Table 4. We can observe that explicit feature-sentiment combinations with negatives, noun phrases, and degree adverbs are correctly extracted.

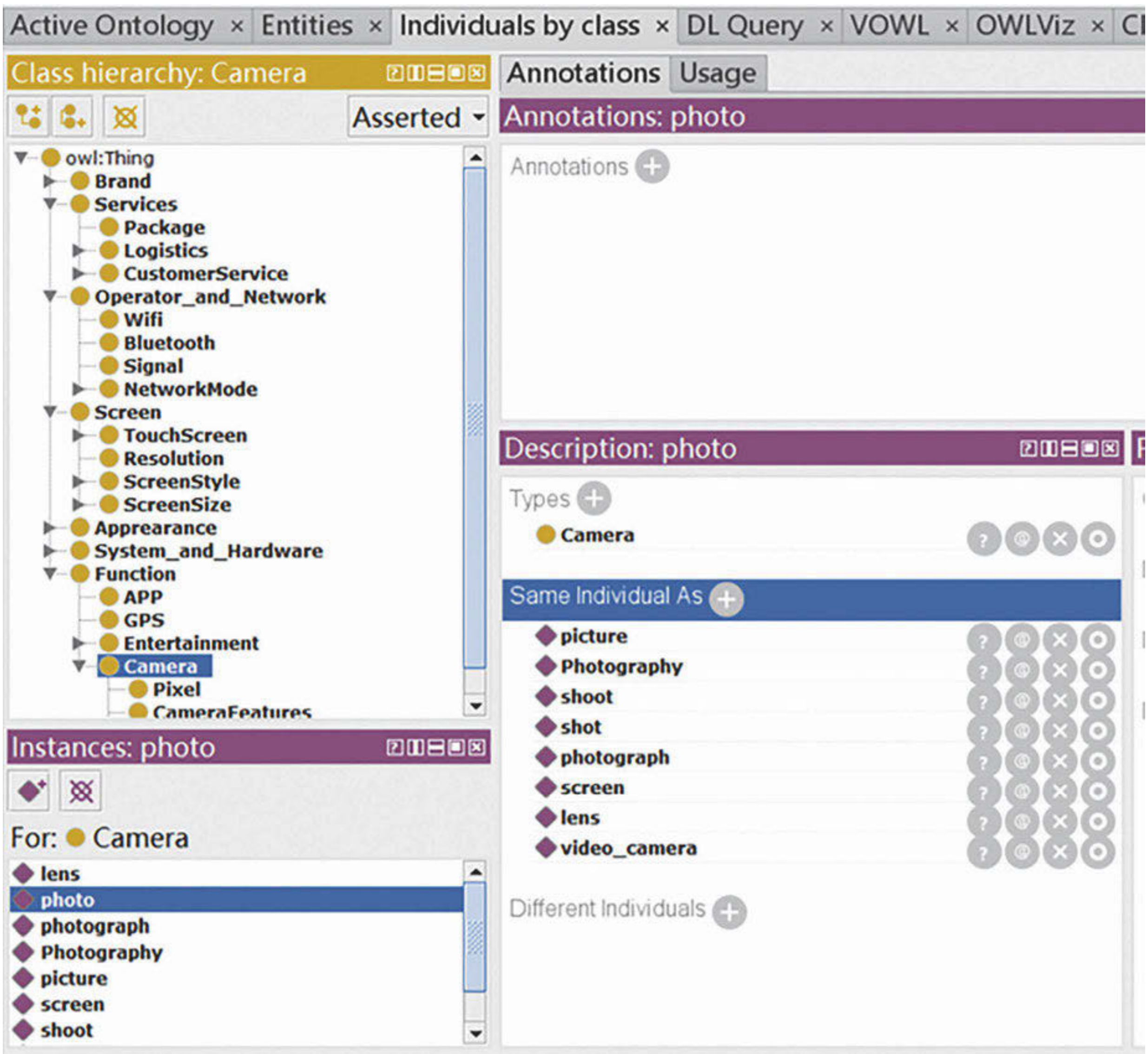

Figure 3. Part of the phone ontology and ontology hierarchy.

Explicit feature words	**Sentiment words**
性价比 (cost performance)	不高 (low)
发货速度 (delivery speed)	很快 (very fast)
反应 (response/ reaction)	灵敏 (responsive /quickly)
运行速度 (running speed)	很快 (very fast)
快递员 (delivery man)	不错 (good)
手感 (touch)	无与伦比 (unbeatable/nice)
OLED屏 (OLED screen)	不错 (good)

Table 4. Part of the explicit feature-sentiment combinations.

To evaluate the performance of our method in extracting explicit feature- sentiment combinations, a total of 30% of records are randomly selected from the experimental corpus. The five proposed rules are used to analyze these 3,000 records by rules dependency syntax and to extract explicit feature-sentiment combinations. At the same time, these 3,000 records are manually identified to extract explicit feature-sentiment combinations. Three students are instructed to perform the identification task, and each feature-sentiment combination is recognized by all the three students.

We use precision (P), recall (R), and F-measure (F) to evaluate the extraction effect. Some parameters are defined as follows: TP is the number of valid feature-sentiment combinations extracted by the rules, TN is the number of invalid feature-sentiment combinations extracted by the rules, and N is the number of feature-sentiment combinations recognized manually.

$$\textit{Precision (P)} = \frac{TP}{TP+TN} \qquad (1)$$

$$\textit{Recall (R)} = \frac{TP}{N} \qquad (2)$$

$$\textit{F-measure (F)} = \frac{2*P*R}{(P+R)} \qquad (3)$$

The total number of explicit feature-sentiment combinations extracted by the rule algorithms is 2,418, and the number of valid explicit feature-sentiment combinations is 2,082. The number of explicit feature-sentiment combinations recognized manually is 2,455. Table 5 shows the evaluation results of our rule algorithms.

Precision	***Recall***	***F***
0.861	0.848	0.854

Table 5. Performance evaluation of the extraction rules.

Although we consider the fact that some punctuation is not standardized in the extraction rules, nonstandard punctuation, unreasonable grammatical structure and typos are abundant. Since the current dependency parser can identify only standardized sentence elements, further rules need to be developed for the identification of these irregular words.

Next, we use the FDO with sentiment weight to match implicit sentiment words and feature words. The steps in 3.2.2 are followed to identify implicit sentiment words and to assign sentiment weights to ontology concepts. The visual results of some ontology concepts after sentiment weight assignment are shown in Figure 4.

According to the sentiment words and the weight of each concept in the ontology, the concept with the highest weight for an implicit sentiment word is used as the feature word. For example, for the Chinese review sentence "too expensive, my heart hurts," the sentiment word "expensive" is matched with the feature "price."

To evaluate the performance of our method in extracting implicit feature-sentiment combinations, explicit feature-sentiment combinations recognized manually in the experimental corpus are filtered out, and 402 pieces of review corpus with adjectives or combination of adverbs and adjectives obtained from the remaining reviews corpus are identified as implicit sentiment words. The FDO with sentiment weights is used to match the implicit sentiment words and feature words. At the same time, the 402 pieces of review sentences and extracted implicit feature-sentiment combinations are judged manually, of which 327 extracted implicit feature-sentiment combinations are identified as valid. These findings indicate that the knowledge-enhanced method with semantic information can improve the effectiveness of feature detection and sentiment analysis.

The final step is to calculate the sentiment value of each product feature. After summarizing the explicit feature-sentiment combinations and the implicit feature-sentiment combinations extracted above, the Contextual Sentiment Dictionary, Domain Sentiment Dictionary, and General Sentiment Dictionary are sequentially matched in order, and the sentiment polarities are marked for the combinations. Part of the sentiment classification results of feature-sentiment combinations are shown in Table 6. We can observe that sentiment polarities of feature-sentiment combinations with negatives, noun phrases, and degree adverbs are correctly judged.

The performance is evaluated on the basis of precision, recall, and F-measure. The parameters are defined as follows: precision is the number of feature-sentiment combinations correctly judged as (+/-) divided by the number of feature-sentiment combinations judged as (+/-) by the rules; and recall is the number of feature-sentiment combinations correctly judged as (+/-) divided by the number of feature-sentiment combinations that actually belong to (+/-).

Comparative experiments are conducted between our method and the method using only the General Sentiment Dictionary. Table 7 shows the sentiment classification results of experiments. We can observe that all the evaluation indicators detected the classification results by our method

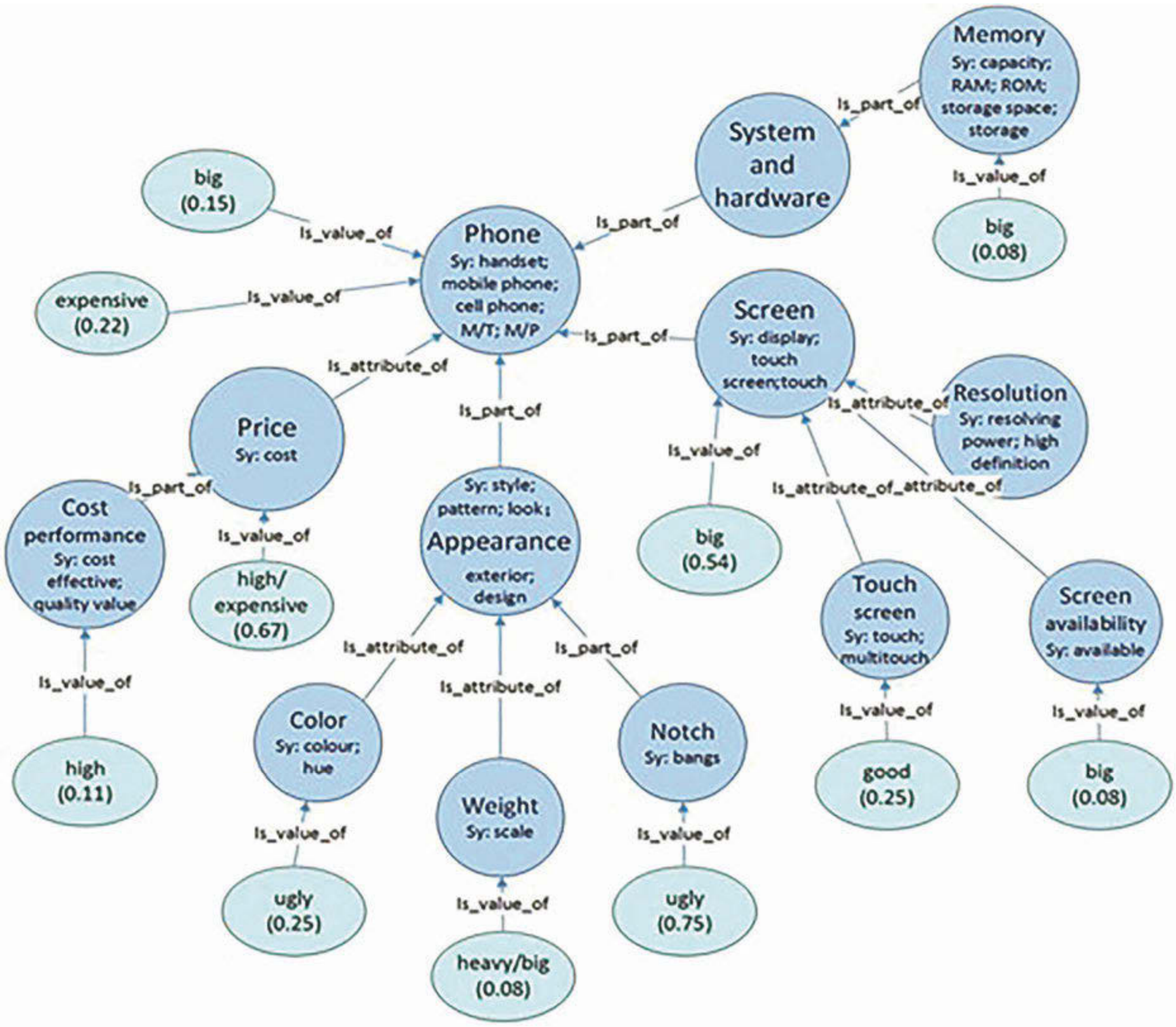

Figure 4. Part of the sentiment weight assignment of the domain ontology concept

Feature-sentiment combinations	Sentiment polarity
外观--惊艳 (appearance--stunning)	1
屏幕可用率--很大 (screen availability--very high)	1
屏幕--够大 (screen--big enough)	1
包装--完美 (package-- perfect)	1
发货--快 (delivery--fast)	1
手感--不错 (touch--nice)	1
性价比--不高 (cost performance--low)	-1

Table 6. Part of the sentiment classification results of feature-sentiment combinations.

Method	Polarity	*Precision*	*Recall*	*F-measure*
Our method	Positive	0.92	0.93	0.92
	Negative	0.90	0.81	0.85
Comparative method	Positive	0.75	0.77	0.76
	Negative	0.67	0.55	0.61

Table 7. Performance evaluation of sentiment classification.

are better than those detected by the comparative method. The precision achieved by the proposed method is nearly 90%, which is an improvement over setting more appropriate rules and more contextualized knowledge mapping for the Chinese corpus. Our approach can provide more detailed sentiment analysis results. Furthermore, the positive sentiment classification results consistently outperform the negative sentiment classification results in both methods. This occurs because review users are more inclined to give a positive evaluation than a negative evaluation, and thus the higher cardinality of positive reviews than of negative reviews causes an imbalance in the experimental results.

Finally, all the feature-sentiment combinations are clustered according to the feature words, and the sentiment values of all the records in each feature word cluster are added and averaged to obtain the final sentiment score of the feature words. Figure 5 presents the calculation results of some feature sentiment values for iPhone X. The direction and values in the histogram discriminate the user's sentiment orientation towards a certain feature. The results show that users positively evaluate the features of "OLED screen," "delivery," "camera," and "appearance" but negatively evaluate the features of "price," "cost performance," and "notch screen." For the features of "customer service" and "screen size," the overall opinions are relatively neutral.

5.0 Conclusions and future work

Pursuant to its aim to obtain sentiment polarity and sentiment scores for the product feature level, this study extracts the explicit features of products based on rules parsing, extracts implicit features of products based on domain ontology, and establishes a series of sentiment lexicons to analyze the sentiment value of the product features, which can provide new ideas for feature-level sentiment analysis. Furthermore, we propose the concept of FDO for review mining, which is used to describe synonyms of the same entity or attribute in reviews. A semiautomatic method is adopted to construct FDO for review mining, which uses machine learning to determine synonymous internet words and improve the efficiency of ontology construction. The method proposed in our study improves the accuracy of feature extraction and the effect of sentiment analysis to some extent.

Our study is also subject to limitations and deficiencies. In the process of domain ontology construction, the acquisition of seed concept words and the upper and lower relationships are still manually constructed. Algorithms instead of manual operations could be used to improve the automation level of domain ontology construction. In addition, a large number of language irregularities exist in actual Chinese review texts, such as omitted punctuation, abbreviated words, and typos. Therefore, future work will focus on proposing rule algorithms for different grammatical irregularities to further improve the extraction performance.

References

Afzaal, Muhammad, Usman Muhammad and Fong Alvis. 2019. "Predictive Aspect-based Sentiment Classification of Online Tourist Reviews." *Journal of Information Science* 45: 341-63.

Arruda, Cafe, Ligia Maria Souza and Renato Rocha. 2017. "Sentiment Analysis and Knowledge Organization: An Overview of the International Literature." *Knowledge Organization* 44: 199-214.

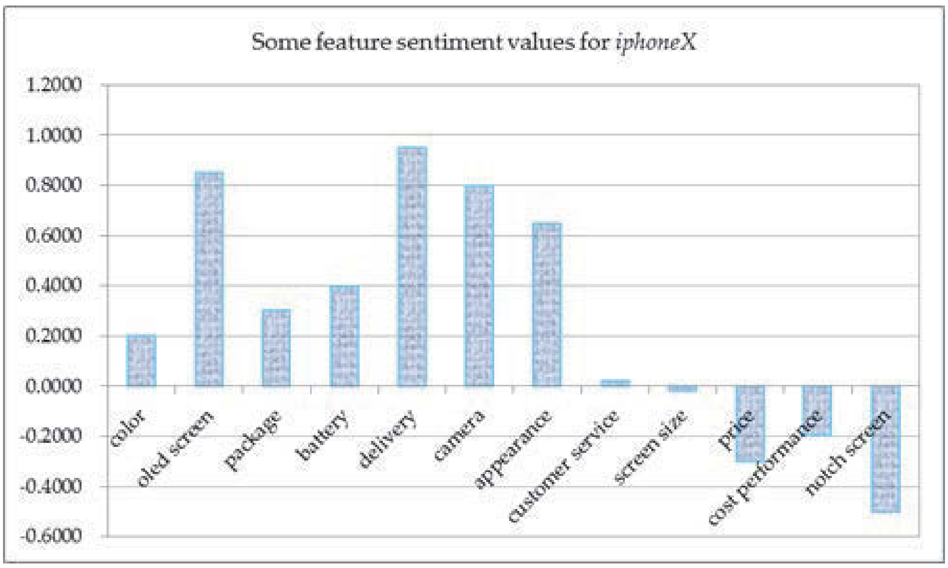

Figure 5. Part of the feature sentiment values.

Chen, Fang and Huang Yongfeng. 2019. "Knowledge-enhanced Neural Networks for Sentiment Analysis of Chinese Reviews." *Neurocomputing* 368: 51-58.

Chen, Runyu, Zheng Yitong, Xu Wei, Liu Minghao and Wang Jiayue. 2018. "Secondhand Seller Reputation in Online Markets: A text Analytics Framework." *Decision Support System* 108: 96-106.

Cui, Xuelian, Narisa and Liu Xiaojun. 2018. "Sentiment Analysis of Online Review based on Topic Similarity." *Journal of Systems and Management* 5: 821-7.

Debashis, Naskar and Subhashis Das. 2019. "HNS Ontology Using Faceted Approach." *Knowledge Organization* 46: 187-98.

Farman, Ali, Daehan Kwak, Pervez Khan, et al. 2019. "Transportation Sentiment Analysis using Word Embedding and Ontology-based Topic Modeling." *Knowledge-Based Systems* 174: 27-42.

Hajar, EI Hannach and Mohammed Benkhalifa. 2016. "Hybrid Approach to Extract Adjectives for Implicit Aspect Identification in Opinion Mining." In *Proceedings of the 11th International Conference on Intelligent Systems: Theories and Applications Mohammedia, Morocco, October 9-20 2016,* New York: IEEE, 289-97. doi:10.1109/SITA.2016.7772284

Hu, Ming and Liu Bing. 2004. "Mining and Summarizing Customer Reviews." In *Proceedings of the 10th ACM SIGKDD International Conference on Knowledge Discovery and Data mining, Seattle, USA*, *August 05-10 2004*. New York: ACM, 168-77. doi:10.1145/1014052.1014073

Huang, Wenming and Sun Yanqi. 2017. "Chinese Short Text Sentiment Analysis based on Maximum Entropy." *Computer Engineering and Design* 1: 138-43.

Kumar, H. I., B. S. Keerthi Harish and I. I. K. Darshan. 2019. "Sentiment Analysis on IMDb Movie Reviews Using Hybrid Feature Extraction Method." *International Journal of Interactive Multimedia and Artificial Intelligence* 5, no. 5: 109-14.

Lazhar, F. and T.G. Yamina. 2017. "Mining Explicit and Implicit Opinions from Reviews." *International Journal of Data Mining, Modeling and Management* 8, no. 1: 75-92.

Li, Shi, Ye Qiang and Li Yijun. 2010. "Mining Features of Products from Chinese Customer Online Reviews." *Journal of Management Sciences in China* 2: 142-52.

Liao, Jian, Wang Suge and Li Deyu. 2019. "Identification of Fact-implied Implicit Sentiment based on Multi-level Semantic Fused Representation." *Knowledge-based Systems* 165: 197-207.

Mankar, S. A. and M. Ingle. 2015. "Implicit Sentiment Identification Using Aspect based Opinion Mining." *International Journal on Recent and Innovation Trends in Computing and Communication* 3: 2184-8.

Medhat, W., A. Hassan and H. Korashy. 2014. "Sentiment Analysis Algorithms and Applications: A Survey." *AIN Shams Engineer Journal* 5: 1093-113.

Mohammad, Tubishata, Norisma Idrisa and Mohammad A.M. Abushariah. 2018. "Implicit Aspect Extraction in Sentiment Analysis: Review, Taxonomy, Opportunities and Open Challenges." *Information Processing and Management* 54: 545-63.

Mukhtar, N., M.A. Khan and Chiragh Nadia. 2018. "Lexicon-based Approach Outperforms Supervised Machine Learning Approach for Urdu Sentiment Analysis in Multiple Domains." *Telematics and Information* 35: 2173-83.

Natalya, F. Noy and Deborah L. McGuinness. 2001. "Ontology Development 101: A Guide to Creating Your First Ontology." Stanford Knowledge Systems Laboratory Technical Report KSL-01-05 and Stanford Medical Informatics Technical Report SMI-2001-0880, March 2001. Palo Alto, CA: Stanford University. http://103.95.217.70/www.ksl.stanford.edu/people/dlm/papers/ontology-tutorial-noy-mcguinness.pdf

Turney, Peter D. 2002. "Thumbs Up or Thumbs Down? Semantic Orientation Applied to Unsupervised Classification of Reviews." In *National Research Council of Canada: Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics, Philadelphia, USA, July 8-10 2002*, ed. Pierre Isabelle, Eugene Charniak and Dekang Lin. https://www.aclweb.org/anthology/P02-1053.pdf

Prasojo, R. E., M. Kacimi and W. Nutt. 2015. "Entity and Aspect Extraction for Organizing News Comments." In *Proceedings of the 24th ACM International on Conference on Information and Knowledge Management, Melbourne, Australia, October 9-12 2015*. New York: ACM, 233-42. doi:10.1145/2806416.2806576

Rana, T. A., Y. N. Cheah and S. Letchmunan. 2018. "Topic Modeling in Sentiment Analysis: A Systematic Review." *Journal of ICT Research and Applications* 10: 76-93.

Santosh, D.T., K.S. Babu, S.D.V. Prasad and A. Vivekananda. 2016. "Opinion Mining of Online Product Reviews from Traditional LDA Topic Clusters using Feature Ontology tree and Sentiwordnet." *International Journal of Education and Management Engineering* 6: 634.

Schouten, Kim and Frasincar Flavius. 2014. "Implicit Feature Detection for Sentiment Analysis." In *Proceedings of the 23rd International Conference on World Wide Web*, *Seoul, Korea, April 2014,* New York: ACM, 367-8. doi:10.1145/2567948.2577378

Schouten, Kim and Frasincar Flavius. 2016. "Survey on Aspect-Level Sentiment Analysis." *IEEE Transactions on Knowledge and Data Engineering* 28: 813-30.

Shepherd, Michael and Tara Sampalli. 2012. "Ontology as Boundary Object." In *Categories, Contexts, and Relations in Knowledge Organization: Proceedings of the 12th International ISKO Conference, 6-9 August 2012, Mysore, India*, ed. A. Neelameghan and K. S. Raghavan. Advances in Knowledge Organization 13. Würzburg: Ergon, 131-7.

Smiraglia, Richard P. 2017. "Practical Ontologies for Information Professionals." *Knowledge Organization* 44: 689-1.

de Kok, Sophie, Linda Punt, Rosita van den Puttelaar, et al. 2018. "Review-Aggregated Aspect-based Sentiment Analysis with Ontology Features." *Progress in Artificial Intelligence* 7: 295-306.

Tang, Feilong, Fu Luoyi and Yao Bin. 2019. "Aspect Based Fine-grained Sentiment analysis for Online Reviews." *Information Science* 488: 190-204.

Tang, Li and Liu Chen. 2019. "Extraction of Feature and Sentiment Word Pair Based on Conditional Random Fields and HITS Algorithm." *Computer Technology and Development* 7: 71-5.

Tang, Xiaobo and Lan Yuting. 2016. "Sentiment Analysis of Microblog Product Reviews Based on Feature Ontology." *Library and Information Service* 16: 121-7.

Toqir, A. Rana and Yu N. Cheah. 2016. "Aspect Extraction in Sentiment Analysis: Comparative Analysis and Survey." *Artificial Intelligence Review* 46: 459-83.

Wang, Hong, Zhou Hao and Sun Jinchuan. 2017. "Research and Application on Domain Ontology Learning Method based on LDA." *Journal of Software* 12: 265-73.

Wan, Y. Nie, H. T. Lan and Z. Wang. 2018. "Fine-grained Sentiment Analysis of Online Reviews." In *Proceedings of the 12th International Conference on Fuzzy Systems and Knowledge Discovery, Zhangjiajie, China, August 15-17 2015,* Washington DC: IEEE, 78-92. doi: 10.1109/ FSKD.2015.7382150

Wang, W., H. Xu and W. Wan. 2013. "Implicit Feature Identification via Hybrid Association Rule Mining." *Expert Systems with Applications* 40: 3518-31.

Wang, Wei, Wang Hongwei and Sheng Xiaobao. 2017. "Product Features and Opinion Recognition in Chinese Online Reviews: A Cross-field Comparative Study." *Journal of Industrial Engineering and Engineering Management* 4: 52-62.

Wei, Jin and Hung Hay Ho. 2009. "A Novel Lexicalized HMM-based Learning Framework for Web Opinion Mining." In *Proceedings of the 26th International Conference on Machine Learning, Montreal, Canada, June 14-18 2009,* New York: ACM, 465-72. https://www.docin.com/p-1656779642.html

Xu, H., F. Zhang and W. Wang. 2015. "Implicit Feature Identification in Chinese Reviews Using Explicit Topic Mining Model." *Knowledge-Based Systems* 76: 166-75.

Yadollahi, A., A.G. Shahraki and O.R. Zaiane. 2017. "Current State of Text Sentiment Analysis from Opinion to Emotion Mining." *ACM Computing Surveys (CSUR)* 50, no. 2: 25.

Yin, Pei, Wang Hongwei and Guo Kaiqiang. 2013. "Feature-Opinion Pair Identification in Chinese Online Reviews Based on Domain Ontology Modeling Method." *Systems Engineering* 1: 68-77.

Yu, Juan and Dang Yanzhong. 2009. "Domain Feature and Its Extracting Approach." *Journal of China Society for Scientific and Technical Information* 28: 368-73.

Zainuddin, N., A. Selamat and R. Ibrahim. 2016. "Improving Twitter Aspect-based Sentiment Analysis Using Hybrid Approach." In *Proceedings of the 2016 Asian conference on intelligent information and database systems March 14-16 2016, Da Nang, Vietnam,* ed. T. N. Ngoc. Berlin: Springer, 151-60.

Zhou, Lifeng. 2016. "Research of Product Feature Extraction and Sentiment Analysis base on Chinese Online Reviews." PhD diss., Dongnan University.