The Electronic Library

Research on feature-based opinion mining using topic maps

Lixin Xia Zhongyi Wang Chen Chen Shanshan Zhai

Research on feature-based opinion mining using topic maps

Lixin Xia, Zhongyi Wang, Chen Chen and Shanshan Zhai
School of Information Management, Central China Normal University, Wuhan City, People's Republic of China

Received 13 November 2014
Revised 26 February 2015
27 April 2015
Accepted 10 June 2015

Abstract

Purpose – Opinion mining (OM), also known as "sentiment classification", which aims to discover common patterns of user opinions from their textual statements automatically or semi-automatically, is not only useful for customers, but also for manufacturers. However, because of the complexity of natural language, there are still some problems, such as domain dependence of sentiment words, extraction of implicit features and others. The purpose of this paper is to propose an OM method based on topic maps to solve these problems.

Design/methodology/approach – Domain-specific knowledge is key to solve problems in feature-based OM. On the one hand, topic maps, as an ontology framework, are composed of topics, associations, occurrences and scopes, and can represent a class of knowledge representation schemes. On the other hand, compared with ontology, topic maps have many advantages. Thus, it is better to integrate domain-specific knowledge into OM based on topic maps. This method can make full use of the semantic relationships among feature words and sentiment words.

Findings – In feature-level OM, most of the existing research associate product features and opinions by their explicit co-occurrence, or use syntax parsing to judge the modification relationship between opinion words and product features within a review unit. They are mostly based on the structure of language units without considering domain knowledge. Only few methods based on ontology incorporate domain knowledge into feature-based OM, but they only use the "is-a" relation between concepts. Therefore, this paper proposes feature-based OM using topic maps. The experimental results revealed that this method can improve the accuracy of the OM. The findings of this study not only advance the state of OM research but also shed light on future research directions.

Research limitations/implications – To demonstrate the "feature-based OM using topic maps" applications, this work implements a prototype that helps users to find their new washing machines.

Originality/value – This paper presents a new method of feature-based OM using topic maps, which can integrate domain-specific knowledge into feature-based OM effectively. This method can improve the accuracy of the OM greatly. The proposed method can be applied across various application domains, such as e-commerce and e-government.

Keywords Feature extraction, Sentiment classification, Topic map, Feature-based opinion mining

Paper type Research paper

Introduction

With the development of Web 2.0, which emphasizes the participation of users, the number of online opinion sources is growing rapidly. More and more websites, such as

This study is supported by National Social Science Foundation of China: "Research on Multi-granularity Integration Knowledge Services of Digital Library Based on Linked Data" (14CTQ003) and is the major project of National Social Science Foundation of China: "Research on Knowledge Discovery of Internet Resource base on Multi-dimensional Aggregation" (No. 13&ZD183).

The Electronic Library
Vol. 34 No. 3, 2016
pp. 435-456

0264-0473
DOI 10.1108/EL-11-2014-0197

Amazon (www.amazon.com) and Epinions (www.epinions.com), encourage users to write opinions for the products they are interested in. These online opinions are useful for customers and manufacturers. The customers refer to these opinions to help decide about their purchases and manufacturers can gather feedback from their customers to improve the quality of their products. However, these online opinions are text information generated by a large number of users, which results in unstructured and unmanaged content. The opinions can number in the hundreds or even in the thousands, so it may be difficult for people to use them. Therefore, automatic extraction and summarization of these opinions has become an urgent need. To meet this requirement, the technology of opinion mining (OM – also known as "sentiment classification"), which aims to discover common patterns of user opinions from their textual statements automatically or semi-automatically, initially proposed by Hatzivassiloglou and McKeown (1997), has now become a significant area of research in the field of data mining. The method involves techniques from different disciplines, including information retrieval, natural language processing and data mining (Vijaya and Sudha, 2013). Meanwhile, a considerable number of varied research results (Choi *et al.*, 2005; Dave *et al.*, 2003; Ghose *et al.*, 2007; Hu and Liu, 2004a, 2004b; Liu, 2010; Pang and Lee, 2008; Turney, 2002) have been achieved, which can be mainly divided into three directions: document-level OM, sentence-level OM and feature-level OM.

For document-level OM (Pang *et al.*, 2002), the entire document is classified as positive, negative or neutral. However, this method is too broad. In most cases, both positive and negative opinions can appear in the same document, so the sentiment orientation at the document level is not sufficient. At the sentence level (Wiebe and Riloff, 2005; Wiebe *et al.*, 2004; Wilson *et al.*, 2005), sentiment classification is applied to individual sentences in a document. Although OM at the sentence level is useful in many cases, it still leaves much to be desired. A positive evaluative sentence on a particular entity does not mean that the author has positive opinions on every aspect of the entity. Researchers, such as Hu and Liu (2004a, 2004b) and Liu *et al.* (2005), have worked on finer-grained OM which predicts the sentiment orientation related to different review features. The task is known as feature-level OM. In feature-level OM, structured opinions on individual features of a whole object are extracted from subjective texts, so it is important to determine the reviewers' opinions towards different product features instead of the overall opinion in those reviews. For this paper, the authors focus on feature-based OM. The authors propose an OM approach and develop an OM system based on the proposed approach to conduct sentiment analysis on the reviews of washing machines. The authors also provide an interface for showing summaries of OM results and report the results. Conclusions and future work are presented.

Literature review

Feature-based OM has been studied by many researchers in recent years. The first definition of the problem can be found in Hu and Liu (2004a):

> […] given a set of customer reviews of a particular product, the task involves three subtasks: identifying features of the product; for each feature, identifying review sentences; and producing a summary using the discovered information.

Many different kinds of approaches for feature-based OM have been proposed, such as Abbasi (2003), Baccianella *et al.* (2010), Ding and Liu (2007) and Zhu *et al.* (2011).

Depending on the implementation method, feature-based OM can be divided into different types of OM, based on keyword spotting, syntax parsing, statistical approaches, taxonomy, concept-level techniques and other methods. However, different approaches have disadvantages requiring optimization.

Opinion mining based on keyword spotting

Some approaches (Esuli and Sebastiani, 2005; Kanayama and Nasukawa, 2006; Turney and Littman, 2003) emphasized opinion words and did not consider features. A few works (Kim and Hovy, 2004; Meena and Prabhakar, 2007; Nasukawa and Yi, 2003) performed sentence-level sentiment analysis using words but did not extract representative features.

Opinion mining based on syntax parsing

Pang and Lee (2008) proposed a method using association rule mining to extract the most frequent features; however, association rule mining has some drawbacks and challenges. Association rule mining generates many features which do not actually represent features of the product, but are just frequently occurring noun phrases (Scaffidi *et al.*, 2007). Zhuang *et al.* (2006) classified and summarized movie reviews by extracting high-frequency feature keywords and high-frequency opinion keywords. Feature-opinion pairs were identified by using a dependency grammar graph. However, this work used a fixed list of keywords to recognize high-frequency feature words and, thus, the system capability is limited. Ding *et al.* (2008) further improved the system by adding some rules to handle different kinds of sentence structures; however, the capability of recognizing phrase features is limited by the accuracy of recognizing noun-group boundaries. Their approach also lacked an effective way to address infrequent features.

Opinion mining based on statistical approaches

Popescu and Etzioni (2005) proposed a relaxation labelling approach to find the semantic orientation of words. However, their approach only extracted feature words with frequency greater than an experimentally set threshold value and ignored low frequency feature words. The main limitation of the approach is that there are many extracted features and there is a lack of organization. Thus, similar features are not grouped together and possible relationships between features of an object are not recognized.

Opinion mining based on taxonomy

To improve the quality of extracted features, Blair-Goldensohn *et al.* (2008) developed a system which extracts information about services, aggregates the sentiments expressed on every aspect and produces a summary. The automatic feature extraction combines a dynamic method, where the different aspects of services are the most common nouns, and a static method, where a taxonomy grouping of the concepts considered to be the most relevant by the user is used to manually annotate sentences. The main limitation is that the taxonomy is restricted to a hierarchical relationship between concepts and, thus, cannot describe other types of paradigmatic relations, such as synonymy or more complex relationships, including compositional or spatial relationships.

Opinion mining based on concept-level techniques

To solve the above problems, Cruz *et al.* (2010) described a domain-specific, resource-based opinion extraction system and the results suggest that domain-specific knowledge is a valuable resource to build precise opinion extraction systems. Zhao and Li (2009) also solved these problems by using an ontology. In their method, the feature extraction phase is guided by domain ontology. They extracted features corresponding exclusively to terms contained in the ontology, which was built manually. However, this method exploited the ontology as a taxonomy using only the "is-a" relation between concepts. It did not use any of the other semantic relationships available in a true ontology, such as the lexical components and other types of relationships.

Motivation for the approach

From the above discussion, it can be seen that, in feature-level OM, most of the existing research associates product features and opinions by their explicit co-occurrence taxonomy or use syntax parsing to judge the modification relationship between opinion words and product features within a review unit. These are mostly based on the structure of language units without considering domain knowledge. Only a few methods based on ontology incorporate domain knowledge into feature-based OM, but they only use the "is-a" relation between concepts. In addition, some online reviews posted by customers fail to comply with syntax specification, which also causes difficulties when OM. Therefore, in feature-based OM, there are still some problems that need to be addressed, including:

- *Domain dependence of sentiment words*: The same sentiment word can mean two very different things in different domains; for example, *big* trucks are good, but *big* phones are bad. This problem has a major effect on the correctness of OM. However, up to the present, there has been not an effective way to solve this problem.
- *Extraction of implicit features*: Implicit features are words or phrases referring to properties which are sometimes connotative (e.g. "tiny phone" refers to the phone size). They are often implied by the opinion words in its context. Neither explicit adjacency nor syntactic analysis is the correct approach to this kind of problem.
- *Identification synonyms of features*: It is common for people to use different words or phrases to describe the same feature. These words and phrases are often synonyms. How to solve this problem effectively is also very important to improve the accuracy of OM.
- *Determining the sentiment strengths of features*: Features are used to represent both components and attributes of an object. These features together form a hierarchy. In this hierarchy, the sentiment polarity of a child feature has an effect on its parent-feature's sentiment polarity. However, when determining the sentiment strengths of features, no method has taken this fact into account, which has reduced the correctness of OM.

From the above discussion, it can be seen that domain-specific knowledge is a key for feature-based OM to solve the problems noted above.

Ontology as a form of knowledge representation about the world has been used by some researchers to integrate knowledge into opinion extraction systems (Cruz *et al.*, 2010; Zhao and Li, 2009). Ontology is an explicit specification of a conceptualization (Gruber, 1995). It is composed of two parts: semantic network and the catalogue. The semantic network is made of abstract concepts and their relations. The catalogue is composed of concrete information items. Ontology can be used to share a common understanding of the structure of information among people (Gruber, 1993; Musen, 1992), to analyse domain knowledge (McGuinness *et al.*, 2000), to make domain assumptions explicit (Rothenfluh *et al.*, 1996) and so on. That is why ontology is used by researchers to implement OM systems. However, the creation process for an ontology is very complex, time-consuming and requires a lot of human and financial resources. That is because during ontology creation, it is necessary to specify the entire semantic network and all the information items that are going to populate it (Ramalho *et al.*, 2006), such as defining classes, arranging the classes in a taxonomic hierarchy, defining slots, describing allowed values for these slots and filling in the values for slots.

Topic maps (Members of the TopicMaps.Org Authoring Group, 2001), as an ontology framework, is not a fixed ontology. It is composed of topics, associations, occurrences and scopes. Topics represent the things the topic map is about. Associations are the relationships between the topics. Occurrences are information resources relevant to a topic. One note is that topics, associations and occurrences may also have types. Scopes allow you to qualify a statement and can be attached to any topic, association or occurrence in a topic map (Pepper and Moore, 2010). Thus, a topic map can describe concepts and relations, not just types, properties and relationship types, which are described by ontology (Park and Cheyer, 2006). The applications of topic maps fall into four broad categories: information integration (II), knowledge management (KM), e-learning and Web publishing. From the II perspective, topic maps offer a meta-model for integrating information. From a KM perspective, topic maps provide the ability to capture and manage some degree of human knowledge, enabling it to be shared and reused across systems. In the domain of e-learning, topic maps bridge the gap between information and knowledge. In the domain of Web publishing, topic maps offer a readymade information architecture.

Topic maps can represent a class of knowledge representation schemes, among the simplest possible architectures, which facilitates representation of subject identity combined with representation of relationships among subjects; thus, it can be used to integrate domain-specific knowledge into OM to solve the problems mentioned above. On the other hand, compared with ontologies, topic maps have many advantages. First, they are more flexible and are able to represent any kind of ontological framework (Vatant, 2004). Second, topic maps are also easy for users to grasp, presumably because they derive from artefacts, such as indexes, that humans have used for centuries. Last but not least, topic maps can describe any concept, not just types, properties and relationships, which are described by ontology. Therefore, topic maps are more suitable than ontologies for feature-based OM, especially when describing features, relationships between features, sentiments and the relationships between features and sentiments. Therefore, this paper proposes a novel OM method based on the International Organization for Standardization (ISO) topic map standard. As the proposed method can make full use of domain knowledge in feature-based OM, it will improve the accuracy of OM to a certain extent.

Aims

To discover common patterns of user opinions from their textual statements accurately, this paper, based on analysing the problems existing in OM, concentrates on improving the accuracy of the feature-based OM method in a practical way. Specifically, the aims of this paper are as follows:

- To solve problems existing in feature-based OM, such as domain dependence of sentiment words, extraction of implicit features and similar, this paper proposes a novel OM method based on topic maps, which use a simplified ontology to implement OM easily and cheaply.
- Develop a topic map which enumerates domain topics and relationships among these topics.
- Integrate topic maps into feature-based OM. Under the direction of a topic map, one can extract implicit features, determine sentiment polarity of a sentiment word in a different context and group synonyms of features.
- Calculate the final sentiment strengths of features based on the features hierarchy provided by the topic map and propose a formula to integrate the child-features' sentiment polarities into their parent-feature, which can improve the veracity of the calculation on the features' sentiment strengths.

Methodology

In this study, the authors focus on feature-based OM and propose a novel OM method based on topic map. The main steps of the algorithm include:

- constructing a topic map on a subject which describes the features, sentiments, the relationships among features and modifying relations between features and sentiments;
- under the direction of the topic map, extracting feature words and sentiment words and determining the initial sentiment strengths of features according to the sentiment words;
- revising sentiment strengths of features if the sentiment words are modified by negative words or adverbs of degree;
- calculating the final sentiment strengths of features to integrate the child-features' sentiment polarities into their parents-feature's sentiment polarity; and
- Summarizing the results of the opinion mining and showing them to the end-users.

The proposed system architecture is shown in Figure 1.

Topic maps, as a new ISO standard for describing knowledge structures, aim to provide knowledge about specific domains that are understandable by both developers and computers. The full use of topic maps has several advantages in the domain of feature-based OM to:

- *Determine sentiment polarity of a sentiment word*: Determine sentiment polarity of a sentiment word in different contexts as the same sentiment word can mean two very different things in different domains; for example, big trucks are good, but big phones are bad. In topic maps, each topic is considered to be valid only within a certain scope. Thus, it is possible to define a scope of sentiment words to

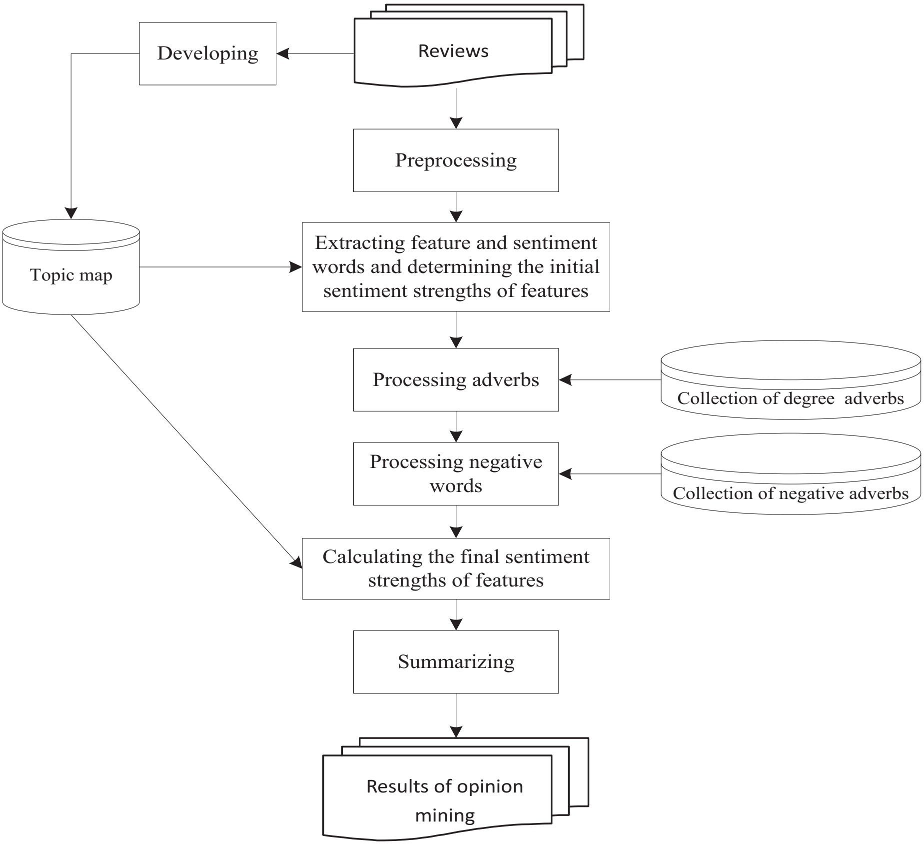

Figure 1.
Method of opinion mining based on topic maps

specify the context in which the sentiment words have the only sentiment polarity. In this way, it is easy to correctly determine the sentiment polarity of a sentiment word.

- *Extract features*: Topic maps can define different relationships between concepts. These relationships are a valuable resource for extracting explicit and implicit features. For example, if the topic "student" is linked to the topic "school" by the relation "to study at", a positive opinion towards the school can be extracted from the review: we study well.
- *Group synonyms of features*: It is common for people to use different words or phrases to describe the same feature. These words and phrases are often synonyms. In a topic map, a topic can be given any number of names and all of these names are synonyms. This means that every topic is really a synonym ring. Therefore, the ability to be able to specify more than one name to the same feature can be used to group synonyms of features.
- *Improve the veracity of the calculation on the features' sentiment strengths*: A topic map provides a structure for features through the concepts hierarchy and the

ability to define relationships linking these concepts. When calculating the sentiment strengths of features, using the structure provided by topic map, the child-features' sentiment polarities can be integrated into their parent-features, which can improve the veracity of the calculation on the features' sentiment strengths.

- *Produce a summary of the review*: The topic map has described the relationships among topics which will provide a summary of the review which presents all of the opinion expressions associated with the main topic and its features.

The above analysis indicates that a topic map allows people to interpret a text review at a finer granularity with shared meanings. Topic maps are especially promising for OM of end-user reviews. In the following discussion, this paper will detail the main steps of this method.

Developing the topic map on a subject

A topic map enumerates domain topics and relationships among the topics, and provides a sound semantic ground of machine-understandable description of digital content. As the functionality and feasibility of an automatically generated topic map still remains unclear, the topic map was constructed by semi-automatic means. The process is defined below.

First, a topics extracting tool (CaseTool), developed by the researchers to extract feature words and sentiment words from end-user reviews, was applied. CaseTool was developed using the JAVA development language. The algorithm of CaseTool is an iterative process. First, sentences of reviews are pre-processed by means of a word processing tool called ICTCLAS (http://ictclas.org) which can realize word segment and part of speech tagging. Second, sentiment words are extracted based on their speech and frequency. Then feature words are extracted not only according to their speech and frequency but also based on their associated sentiment words. The interdependent relationships between feature words and sentiment words are identified by the dependency parsing method which is an important natural language processing technology. Finally, the process is repeated until all the extracted feature words and sentiment words are identified.

The second step is to use the lexicon HowNet, WordNet and Chinese Internet Language Dictionary to expand the synonyms of the feature words and sentiment words. HowNet is an online common-sense knowledgebase unveiling inter-conceptual relationships and inter-attribute relationships of concepts as connoting in lexicons of Chinese and their English equivalents. WordNet is a lexical database. It groups words into sets of synonyms, called *synsets*, provides short definitions and usage examples, and records relationships among these synonym sets or their members. However, Internet language also includes a variety of slang languages, such as abbreviations ("CU" for "see you"), keyboard symbols and others, which are not contained in mainstream dictionaries, such as Hownet and WordNet. To address this problem, the Chinese Internet Language Dictionary is adopted as an additional synonym extension of feature words and sentiment words. The Chinese Internet Language Dictionary is a Chinese dictionary focused on Internet terms. The collection of entries includes general terms and commonly used special words used on or related to the Internet. Additionally, for the purpose of practicality, this dictionary contains commonly used English words and abbreviations related to the Internet.

Next, experts are expected to accept or refuse each candidate, and define the relationships among all of these accepted feature words and sentiment words. The object of this step is to construct a scheme which defines the rules for what can go into the instances.

Finally, a topic map is implemented under Ontopoly which is a topic map visualization development tool published by Ontopia (Figure 2). Ontopoly's primary purpose is to enable the manual creation and maintenance of topic maps. The procedure of topic map construction based on Ontopoly includes two steps and involves two configuration pages. Initially, Ontopoly's Type Configuration Pages are used by the schema designer to define the topic, name, association, association role and occurrence types that will be available to creators of the instance topics. Second, once a scheme is created, it can be used to guide authoring the topic map instances by means of the Instance Editing Pages.

Extracting feature words and sentiment words

Once the topic map has been constructed, the next step is to extract feature words and sentiment words under its guidance. Extracting feature words and sentiment words based on a topic map can effectively address the problems still existing in feature-based OM, including domain dependence of sentiment words, extraction of implicit features and identifying synonyms of features. The process of extracting feature words and sentiment words based on a topic map is shown in Figure 3:

- To implement this process, the first step is to perform sentence splitting. It generally considers ".", "!", "?" and related characters as sentence delimiters for the splitting process.
- After sentence splitting, the next step is to segment and tag words for each sentence by using the ICTCLAS.
- Search each word in the sentiment words collection of the topic map checking whether the sentence has sentiment words or not. If this step fails to find the sentiment word in the topic map, it means that this sentence is not a sentiment sentence, so the sentiment analysis work is unnecessary for this sentence.
- If successful at finding the sentiment word in the topic map, then the next step is to conduct dependency parsing using the Stanford Parser to check whether the sentiment word has a dependency with a certain feature word or not. If there is a

Figure 2.
Interface of Ontopoly for topic map construction

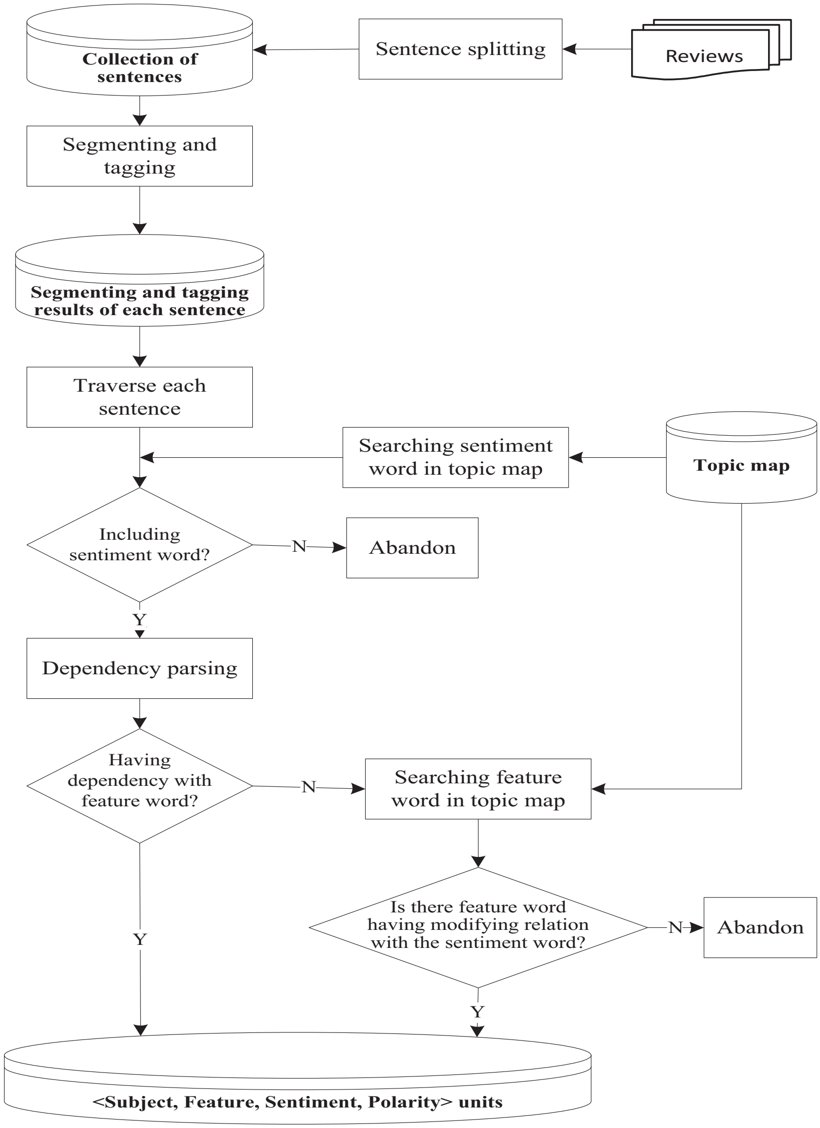

Figure 3.
Extracting feature words and sentiment words

feature word having a dependency with the sentiment word, the subject, feature and sentiment word with the feature's initial sentiment polarity will be considered as a unit and stored in the database.

- If there is no feature word having dependency with the sentiment word, it means that this feature is an implicit feature. Then, the final step is to search the feature words collection of the topic map for the purpose of checking whether some feature word has a modifying relationship with the sentiment word or not. If successfully finding the feature word, the subject, feature and sentiment word with the feature's initial sentiment polarity will be considered as a unit and stored in the database. If the search fails, the sentence will be abandoned.

Revising sentiment strengths of features

Processing adverbs of degree. Adverbs of degree tell us the strength or intensity of something that happens. Basically, they answer the sort of question that asks "How much […]?" or "How little […]?" Adverbs of degree include: very, almost, highly, strongly, totally and similar words. These words play an important role in people's sentiment expression. They can intensify or weaken the strength of the sentiment polarity. Therefore, dealing with adverbs of degree is an essential part of OM. The process of the processing degree adverbs is shown in Figure 4.

The first step is to match the adverbs in the sentence with the collection of degree adverbs. If the match fails, it means that the adverb is not a degree adverb; thus, the strength of the feature's sentiment polarity remains unchanged. If the match works, proceed to the next step. The second step is to perform dependency parsing to check whether the degree adverb has dependency with sentiment word or not. If there is no sentiment word having dependency with this degree adverb, the strength of the

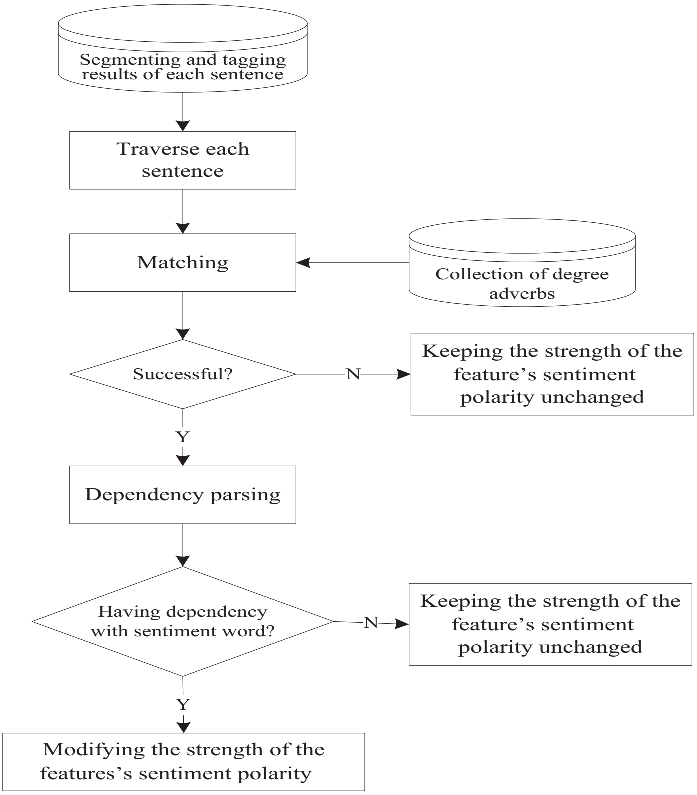

Figure 4.
Processing adverbs

feature's sentiment polarity remains unchanged. Otherwise, the strength of the feature's sentiment polarity should be modified.

Processing negative words. A so-called negative word can reverse a feature's sentiment polarity. Negative words include no, not, never, hardly, rarely, few, little, seldom and similar. Processing these words is a crucial step during OM. The process of dealing with negative words is shown in Figure 5.

As shown in Figure 5, the first step is to match the negative word in the sentence with the negative words collection. If the match fails, the strength of the feature's sentiment polarity remains unchanged. If the match succeeds, continue to the second step. The second step is to analyse the dependency between the negative adverb and the sentiment word. If it has no dependency with a sentiment word, the feature's sentiment polarity should be unchanged; otherwise, if the negative word also has dependency with another negative word, the strength of the feature's sentiment polarity remains unchanged. If the negative word has no dependency with another negative word, reverse the feature's sentiment polarity. The third step is to analyse dependency between the negative adverb and degree adverb. If the negative word has no dependency with a degree adverb, keep the strength of the feature's sentiment unchanged. If the position of the negative adverb is in front of the degree adverb, the strength of the feature's sentiment should be modified. If the negative adverb is after the degree adverb, the feature's sentiment polarity should be reversed.

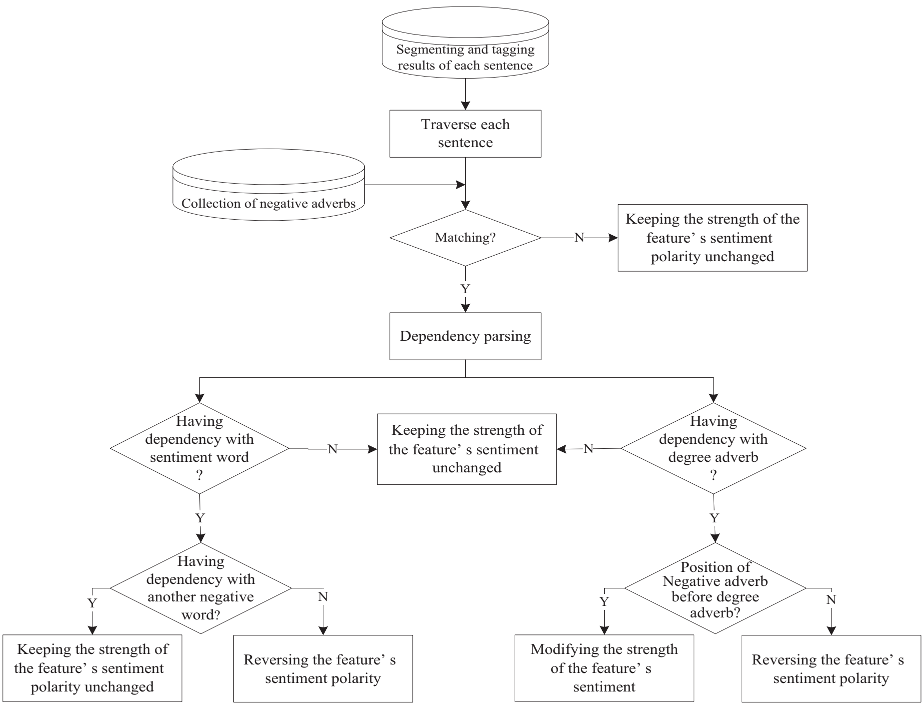

Figure 5.
Processing negative words

Calculating the final sentiment strengths of features

Through the above steps, the initial sentiment strengths of features are obtained and the next step is to calculate the final sentiment strength of each feature. Features in the topic map are hierarchically organized (Figure 6). The subject class (for example, a product) itself is the root node of the hierarchy, with a set of features hanging on it. Each feature can be recursively decomposed in a set of sub-features. The feature hierarchy is a useful resource for both aggregating opinions to produce summaries and calculating the final sentiment strengths of features. This section will focus on how to use this hierarchy to calculate the final sentiment strength of features.

In this hierarchy, the sentiment polarity of a child-feature has an effect on its parent-feature's sentiment polarity. However, when determining the sentiment strengths of features, no existing method has taken account of this fact which will reduce the accuracy of OM. To solve this problem, this next section will introduce how to integrate the child-features' sentiment polarities into their parent-feature based on the topic map. The calculation of the final sentiment strength of each feature follows a bottom-up approach, as shown in equation (1):

$$P_{\mathrm{t}}(\text{parent\_feature}) = P(\text{parent\_feature}) + \sum_{\mathrm{i}=1}^{\mathrm{n}} P_{\mathrm{t}}(\text{child\_feature i}) \times W(\text{child\_feature i}) \tag{1}$$

The P_t(feature) denotes the final sentiment strength of this feature; P(feature) denotes the initial sentiment strength of this feature, while W(feature) denotes the weight of this feature's sentiment strength to its parent-feature.

In equation (2), f_i denotes the value of child-feature i, which is set by the end-user. The sum of f_i equals 100. From equation (2), it can be seen that the system grants the end-user the right to change these weights according to need, which thus can effectively meet personalized demands of end-users. This is because the same feature may have a different importance for different end-users. For example, some people may think that the "speed" of a car is the most important feature, while other people may believe that "fuel consumption" is more important when deciding to buy a car:

$$W(\mathit{child}\text{\_feature i}) = f_i/100 \tag{2}$$

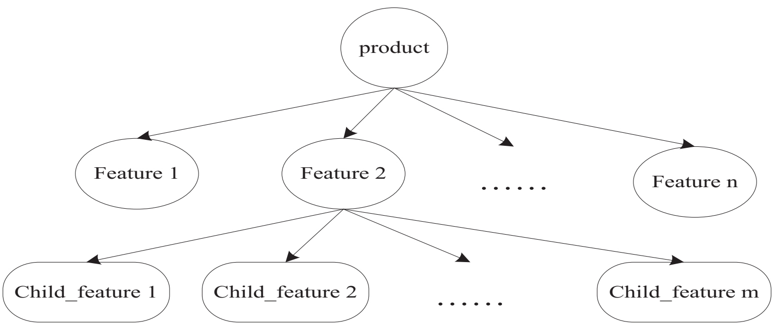

Figure 6. The feature hierarchy

Experiments and results

Data selection

The authors conducted this experiment using customer reviews of washing machines collected from the website JingDong Mall (www.360buy.com), which is one of the more famous e-commerce websites in China.

Washing machines were selected as the test domain mainly for two reasons. First, as washing machines have become a necessity in daily life, the demand for washing machine has grown quickly. However, there are so many diverse brands of washing machines that customers often encounter difficulties. They often do not have enough information about each brand of washing machine to help them make an informed purchase decision. Secondly, in online communities and Q&A platforms, there are a large number of product reviews shared by users. Digging out information from these reviews has great value for customers.

JingDong Mall, as one of the most popular e-commerce websites in China, allows customers to post reviews on the quality, price, service and so forth of products. Thus, a great deal of subjective information has been accumulated on the site, creating an effective data source. To make the data more representative, only 20 different styles of washing machine with more than 100 reviews were selected. The reviews from the JingDong Mall website were extracted. This resulted in 4,025 reviews.

Construction of the topic map

According to the topic map development method proposed above, a topic map on the washing machines (Figure 7) has been developed. This topic map consists of features on the washing machines, the relations among these features, sentiment words and the relationships between features and sentiment words. Because of size limitations, Figure 7 only shows part of the topic map on the LG WD-N10230D washing machine.

Implementation of the system

According to the method proposed in this paper and based on the topic map on the product of washing machine, the researchers implemented an OM system on the Windows XP platform by means of JavaServer Pages. The entire system can be divided into two parts: knowledge resources and the toolkit used in feature-based OM.

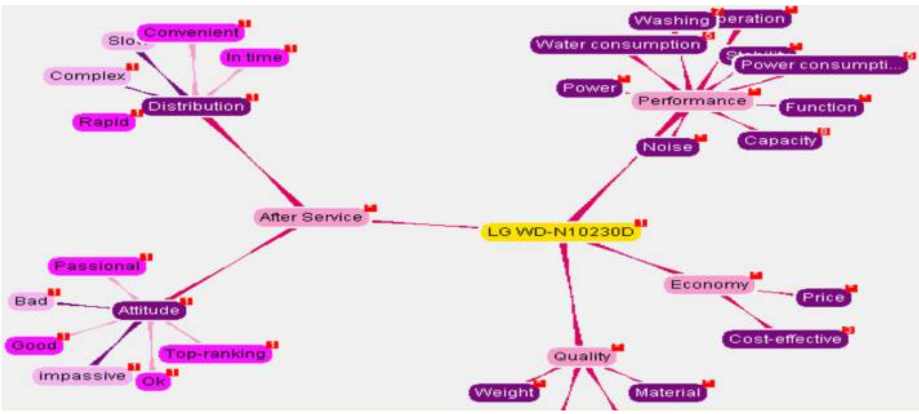

Figure 7.
Partial topic map for the LG WD-N10230D washing machine

Knowledge resources are the basis of feature-based OM. They include the topic map, a collection of negative words and a collection of degree adverbs which provide data supporting the feature-based OM. The toolkit provides tools for OM, including ICTCLAS, an annotation tool, the feature and sentiment words extracting tool (Casetool), processors of negative words and adverbs and the Stanford Parser. The tools in the toolkit linked by interfaces are independent of each other, which improves the scalability of the system and facilitates future improvements and upgrades.

Interface for showing summaries of opinions

To meet the personalized demands of end-users, the system provides a customizable interface (Figure 8) which is used to show the summaries of opinions to the end-users. End-users can choose their own preferred brands of washing machines or their features of concern and the system will then display the summaries of opinions related to their choices.

The summaries of opinions can be shown by two methods: comparison lists and ranking lists.

Comparison lists

After end-users select their preferred brands of washing machines, the system shows the summaries of opinions about them for comparison. Figure 9 shows comparison results about the features of appearance, performance, economy and after service for washing machines SANYO XQB50-S805Z, TCL XQB50-32S and HAIER XQB45-10B.

Ranking lists

The system provides two kinds of ranking lists: ranking list for washing machines and ranking list for each feature of the washing machines. Each kind of ranking list is divided into three types: positive ranking list, negative ranking list and hot ranking list. The *positive ranking list* is a list ranked according to the sentiment strength of positive opinions. The *negative ranking list* is a list which is ranked according to an estimation of the sentiment strength of negative opinions. The *hot ranking list* is ranked based on the number of opinions, with both negative and positive opinions on a feature. Figure 10

Figure 8.
Opinion mining system interface

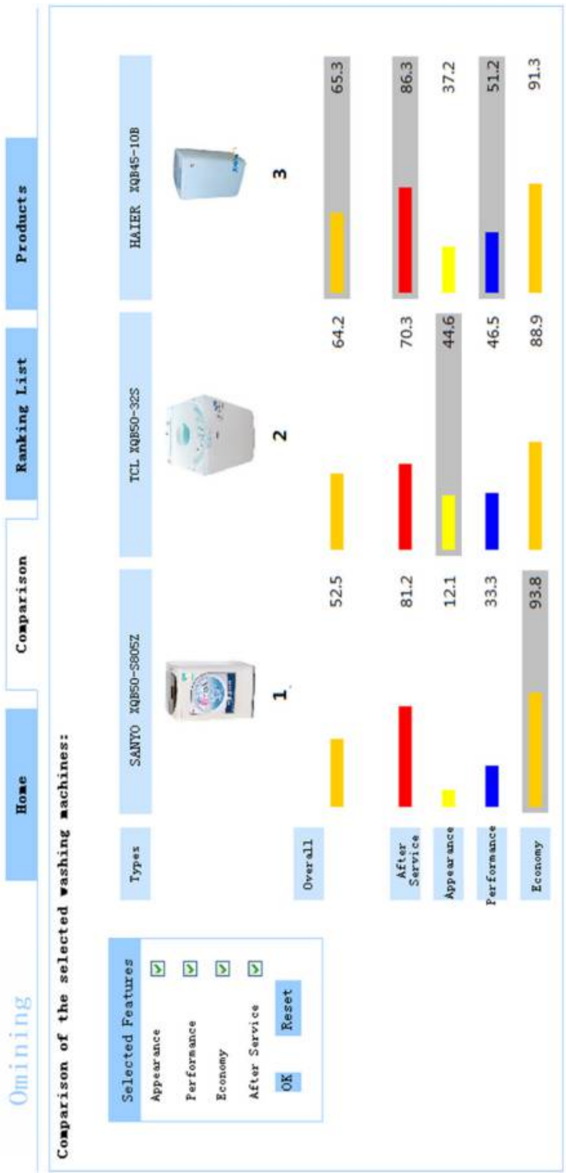

Figure 9.
Comparison interface

shows the positive, negative and hot ranking lists for the feature "configuration", while the ranking lists for the different brands of washing machines are shown in Figure 11. Using AUCMA XPB20-2208 as an example (Figure 10), AUCMA XPB20-2208 is on the top of the hot ranking list for configuration, which means that AUCMA XPB20-2208 has the highest number of comments on configuration among all the washing machines. However, it is the fourth place in the positive ranking list of configuration and ranks first on the positive ranking list for configuration, which means that although it gets the most reviews, the sentiment strength of positive opinions is bigger than the sentiment strength of negative opinions.

Experimental results

The authors conducted several experiments to empirically evaluate the performance of the system. In these experiments, 500 positive reviews and 500 negative reviews were randomly selected from the 4,025 collected reviews. For each selected review, the researchers manually labelled the subjects, features, sentiments and polarities. These labelled <subject, feature, sentiment, polarity> units are used as a gold standard in these experiments. For simplicity, "SFSP" was substituted for the <subject, feature, sentiment, polarity> unit.

Figure 10.
Ranking lists on an individual feature of each washing machine

Figure 11.
Ranking lists on washing machines

The results were evaluated by recall R, as given in equation (3), and precision P, a sgiven in equation (4), defined as follows:

$$R = \frac{\text{correctly extracted SFSP units}}{\text{Total number of SFSP units}} \tag{3}$$

$$P = \frac{\text{correctly extracted SFSP units}}{\text{Total number of SFSP units extred by the system}} \tag{4}$$

The F-measure, a sgiven in equation (5) which is the harmonic mean of precision and recall was also used:

$$\mathrm{F} = \frac{2 \times \mathrm{R} \times \mathrm{P}}{\mathrm{R} + \mathrm{P}} \tag{5}$$

To evaluate the added value of using a topic map with feature-based OM, the system was compared with the well-known approach of Hu and Liu (2004a) that did not use any knowledge representation and the approach of Zhao and Li (2009) which uses an ontology to integrate domain-specific knowledge into feature-based OM. The results are shown in Table I.

From Table I, it is clear that the precision of the proposed system is stronger than the approach of Hu and Liu (2004a) that extracts too many irrelevant features, and also provides more effective recall than the study by Hu and Liu (2004a), as the proposed system can extract implicit features because of the use of a topic map.

When compared to Zhao and Li (2009), the proposed approach has a higher recall and a similar precision. Both the proposed system and the approach of Zhao and Li (2009) combine domain-specific knowledge into OM, thus, the precision is similar and stronger than the approach of Hu and Liu (2004a). However, as the approach of Zhao and Li (2009) just exploited the ontology as a taxonomy using only the "is-a" relation between concepts and did not use any of the other semantic relationships available in a true ontology, such as the lexical components and other types of relations, the experimental system provides a more effective recall compared to the approach of Hu and Liu (2004a) on the recall.

The experimental results indicate that the proposed OM method based on a topic map had a positive impact on the performance of polarity mining. There are both theoretical contributions and practical implications to OM. Regarding the theoretical contributions, this method can address problems existing in feature-based OM effectively, such as domain dependence of sentiment words, extraction of implicit features and similar. The approach uses a topic map to provide semi-automatic implementation of OM, which not only improves the accuracy of OM but also makes it easy and cheap to implement OM.

Table I.
Results of experiments

Method	Precision (P)	Recall (R)	F-measure (F)
Hu and Liu (2004a)	0.6682	0.7665	0.7140
Zhao and Li (2009)	0.7832	0.7245	0.7527
Proposed system	0.8128	0.7946	0.8036

As for practical implications, the system can be widely applied in many fields. First, it can help marketers evaluate the success of an ad campaign or new product launch, determine which versions of a product or service are popular and identify which demographics like or dislike particular product features. Second, the experimental system can be used to help users digest the vast availability of opinions in an easy manner. Third, the system can be adopted by governments to capture the societal impact of public sector regulations in an attempt to decipher the public's stance towards governmental decisions. Finally, the system can also be used in other fields where sentiment analysis is involved, such as in trends analysis in research, population behaviour analysis in information consumption and so on.

Conclusions and future work

Feature-based OM has proven to be a promising technique to mine subjective information. Many researchers have been doing studies on it in recent years and valuable research results have been gained. However, because of the complexity of natural language, there are still some problems that need to be addressed in feature-based OM, such as the problems discussed above. To solve these problems, this research proposes a novel OM method based on a topic map. The contributions of the work include: developing a topic map on washing machines which enumerated domain topics and relationships among the topics and integrating it into feature-based OM. Under the direction of this topic map, the computer can extract implicit features; determine sentiment polarity of a sentiment word in a different context and group synonyms of features. Additionally, based on the features hierarchy provided by the topic map, the researchers propose a formula to integrate the child-features' sentiment polarities into their parent-feature which can improve the veracity of the calculation on the features' sentiment strengths.

OM is a difficult task. Although the novel proposed OM method based on a topic map can address some of the problems which still exist in feature-based OM and can improve the performance, there is still much work needed. One of the problems urgently needed to be addressed in the proposed method is the problem of automatic development of the topic map. In future, the researchers will concentrate on addressing this problem to further improve this method and evaluating the proposal on other languages.

References

Abbasi, A. (2003), "Intelligent feature selection for opinion classification", *Technology*, Vol. 54 No. 14, pp. 1269-1277.

Baccianella, S., Esuli, A. and Sebastiani, F. (2010), "SentiWordNet 3.0: an enhanced lexical resource for sentiment analysis and opinion mining", *LREC, Malta*, available at www.lrec-conf.org/proceedings/lrec2010/slides/769.pdf (accessed 31 January 2016).

Blair-Goldensohn, S., Hannan, K., McDonald, R., Neylon, T., Reis, G.A. and Reynar, J. (2008), "Building a sentiment summarizer for local service reviews", *WWW2008 Workshop: Natural Language Processing Challenges in the Information Explosion Era (NLPIX 2008)*, ACM, New York, NY, p. 14.

Choi, Y., Cardie, C., Riloff, E. and Patwardhan, S. (2005), "Identifying sources of opinions with conditional random fields and extraction patterns", *Proceedings of the Human Language Technology Conference and the Conference on Empirical Methods in Natural Language Processing (HLT/EMNLP)*, Association for Computational Linguistics, Stroudsburg, PA, pp. 355-362.

Cruz, F.L., Troyano, J.A., Enríquez, F., Ortega, F.J. and Vallejo, C.G. (2010), "A knowledge-rich approach to feature-based opinion extraction from product reviews", *Proceedings of the 2nd International Workshop on Search and Mining User-Generated Contents*, ACM, New York, NY, pp. 13-20.

Dave, K., Lawrence, S. and Pennock, D.M. (2003), "Mining the peanut gallery: opinion extraction and semantic classification of product reviews", *Proceedings of WWW*, ACM, New York, NY, pp. 519-528.

Ding, X. and Liu, B. (2007), "The utility of linguistic rules in opinion mining", *Proceedings of the 30th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, ACM, New York, NY, pp. 811-812.

Ding, X., Liu, B. and Yu, P.S. (2008), "A holistic lexicon-based approach to opinion mining", *Proceeding of the International Conference on Web Search and Web Data Mining (WSDM '08)*, ACM, New York, NY, pp. 231-239.

Esuli, A. and Sebastiani, F. (2005), "Determining the semantic orientation of terms through gloss classification", *Proceedings of the 14th ACM International Conference on Information and Knowledge Management*, ACM, New York, NY, pp. 617-624.

Ghose, A., Ipeirotis, P.G. and Sundararajan, A. (2007), "Opinion mining using econometrics: a case study on reputation systems", *Proceedings of the 45th Annual Meeting of the Association of Computational Linguistics, Prague*, ACM, New York, NY, pp. 416-423.

Gruber, T.R. (1993), "A translation approach to portable ontology specifications", *Knowledge Acquisition*, Vol. 5 No. 2, pp. 199-220.

Gruber, T.R. (1995), "Toward principles for the design of ontologies used for knowledge sharing?", *International Journal of Human-Computer Studies*, Vol. 43 No. 5, pp. 907-928.

Hatzivassiloglou, V. and McKeown, K.R. (1997), "Predicting the semantic orientation of adjectives", *Proceedings of the 35th Annual Meeting of ACL*, Association for Computational Linguistics, Stroudsburg, PA, pp. 174-181.

Hu, M. and Liu, B. (2004a), "Mining and summarizing customer reviews", *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD '04)*, ACM, New York, NY, pp. 168-177.

Hu, M. and Liu, B. (2004b), "Mining opinion features in customer reviews", *Proceedings of AAAI*, AAAI Press, Palo Alto, CA, pp. 755-760.

Kanayama, H. and Nasukawa, T. (2006), "Fully automatic lexicon expansion for domain-oriented sentiment analysis", *Proceedings of the 2006 Conference on Empirical Methods in Natural Language Processing*, Association for Computational Linguistics, Stroudsburg, PA, pp. 355-363.

Kim, S.M. and Hovy, E. (2004), "Determining the sentiment of opinions", *Proceedings of the 20th International Conference on Computational Linguistics*, Association for Computational Linguistics, Stroudsburg, PA, p. 1367.

Liu, B. (2010), "Sentiment analysis and subjectivity", available at: www.cs.uic.edu/~liub/FBS/NLP-handbook-sentiment-analysis.pdf (accessed 31 January 2016).

Liu, B., Hu, M. and Cheng, J. (2005), "Opinion observer: analyzing and comparing opinions on the web", *Proceedings of the 14th International Conference on World Wide Web (WWW '05)*, ACM, New York, NY, pp. 342-351.

McGuinness, D.L., Fikes, R., Rice, J. and Wilder, S. (2000), "An environment for merging and testing large ontologies", *Proceedings of the Seventh International Conference on Principles of Knowledge Representation and Reasoning (KR '00), Breckenridge, CO, 12-15 April 2000*, pp. 483-493.

Meena, A. and Prabhakar, T.V. (2007), "Sentence level sentiment analysis in the presence of conjuncts using linguistic analysis", in Amati, G., Carpineto, C. and Romano, G. (Eds), *Advances in Information Retrieval*, Springer, Berlin, pp. 573-580.

Members of the TopicMaps.Org Authoring Group (2001), "XML Topic Maps (XTM) 1.0", available at: www.topicmaps.org/xtm/ (accessed 12 October 2014).

Musen, M.A. (1992), "Dimensions of knowledge sharing and reuse", *Computers and Biomedical Research*, Vol. 25 No. 5, pp. 435-467.

Nasukawa, T. and Yi, J. (2003), "Sentiment analysis: capturing favorability using natural language processing", *Proceedings of the 2nd International Conference on Knowledge Capture*, ACM, New York, NY, pp. 70-77.

Pang, B. and Lee, L. (2008), "Opinion mining and sentiment analysis", *Foundations and Trends in Information Retrieval*, Vol. 2 No. 1, pp. 1-135.

Pang, B., Lee, L. and Vaithyanathan, S. (2002), "Thumbs up?: sentiment classification using machine learning techniques", *Proceedings of the ACL-02 Conference on Empirical Methods in Natural Language Processing*, Association for Computational Linguistics, Stroudsburg, PA, pp. 79-86.

Park, J. and Cheyer, A. (2006), "Just for me: topic maps and ontologies", in Maicher, L. and Park, J. (Eds), *Charting the Topic Maps Research and Applications Landscape*, Springer, Berlin, pp. 145-159.

Pepper, S. and Moore, G. (2010), "Topic maps", *Encyclopedia of Library and Information Sciences*, CRC Press, Boca Raton, FL, pp. 5247-5259.

Popescu, A. and Etzioni, O. (2005), "Extracting product features and opinions from reviews", *Proceedings of Human Language Technology Conference and Conference on Empirical Methods in Natural Language Processing (HLT/EMNLP)*, *Vancouver*, *October*, pp. 339-346.

Ramalho, J.C., Librelotto, G.R. and Henriques, P.R. (2006), "Metamorphosis – a topic maps based environment to handle heterogeneous information resources", in Maicher, L. and Park, J. (Eds), *Charting the Topic Maps Research and Applications Landscape*, Springer, Berlin, pp. 14-25.

Rothenfluh, T.E., Gennari, J.H., Eriksson, H., Puerta, A.R., Tu, S.W. and Musen, M.A. (1996), "Reusable ontologies, knowledge-acquisition tools, and performance systems: PROTÉGÉ-II solutions to Sisyphus-2", *International Journal of Human-Computer Studies*, Vol. 44 No. 3, pp. 303-332.

Scaffidi, C., Bierhoff, K., Chang, E., Felker, M., Ng, H. and Jin, C. (2007), "Red Opal: product-feature scoring from reviews", *Proceedings of the 8th ACM Conference on Electronic Commerce*, ACM, New York, NY, pp. 182-191.

Turney, P. (2002), "Thumbs up or thumbs down? Semantic orientation applied to unsupervised classification of reviews", *Proceedings of the Association for Computational Linguistics (ACL)*, ACM, Philadelphia, PA, pp. 417-424.

Turney, P.D. and Littman, M.L. (2003), "Measuring praise and criticism: inference of semantic orientation from association", *ACM Transactions on Information Systems (TOIS*), Vol. 21 No. 4, pp. 315-346.

Vatant, B. (2004), "Ontology-driven topic maps", *Proceedings of XML Europe*, *Europe*, pp. 18-21.

Vijaya, D.M.S. and Sudha, V.P. (2013), "Research directions in social network mining with empirical study on opinion mining", *CSI Communication*, Vol. 37 No. 9, pp. 23-26.

Wiebe, J. and Riloff, E. (2005), "Creating subjective and objective sentence classifiers from unannotated texts", in Gelbukh, A. (Ed.), *Computational Linguistics and Intelligent Text Processing*, Springer, Berlin, pp. 486-497.

Wiebe, J., Wilson, T., Bruce, R., Bell, M. and Martin, M. (2004), "Learning subjective language", *Computational Linguistics*, Vol. 30 No. 3, pp. 277-308.

Wilson, T., Wiebe, J. and Hoffmann, P. (2005), "Recognizing contextual polarity in phrase-level sentiment analysis", *Proceedings of the Conference on Human Language Technology and Empirical Methods in Natural Language Processing*, Association for Computational Linguistics, Vancouver, pp. 347-354.

Zhao, L. and Li, C. (2009), "Ontology based opinion mining for movie reviews", in Karagiannis, D. and Jin, Z. (Eds), *Knowledge Science, Engineering and Management*, Springer, Berlin, pp. 204-214.

Zhu, J., Wang, H., Zhu, M., Tsou, B.K. and Ma, M. (2011), "Aspect-based opinion polling from customer reviews", *IEEE Transactions on Affective Computing*, Vol. 2 No. 1, pp. 37-49.

Zhuang, L., Jing, F. and Zhu, X. (2006), "Movie review mining and summarization", *Proceedings of the International Conference on Information and Knowledge Management (CIKM '06), Arlington, VA*, pp. 43-50.

Corresponding author
Zhongyi Wang can be contacted at: wzywzy13579@163.com