Fine-Grained Opinion Mining by Integrating Multiple Review Sources

Qingliang Miao and Qiudan Li
Institute of Automation, Chinese Academy of Sciences, Beijing, China. E-mail: qingliang.miao@gmail.com; qiudan.li@ia.ac.cn

Daniel Zeng
MIS Department, University of Arizona, Tucson, AZ. E-mail: zeng@email.arizona.edu

With the rapid development of Web 2.0, online reviews have become extremely valuable sources for mining customers' opinions. Fine-grained opinion mining has attracted more and more attention of both applied and theoretical research. In this article, the authors study how to automatically mine product features and opinions from multiple review sources. Specifically, they propose an integration strategy to solve the issue. Within the integration strategy, the authors mine domain knowledge from semistructured reviews and then exploit the domain knowledge to assist product feature extraction and sentiment orientation identification from unstructured reviews. Finally, feature-opinion tuples are generated. Experimental results on real-world datasets show that the proposed approach is effective.

Introduction

With the dramatic adoption of Web 2.0, user-generated content (UGC) such as product reviews, blogs, and Web forums have become extremely valuable sources for mining people's opinions on products and services. In business and trade, these UGC could greatly affect the decision of consumption. According to a survey, 81% of Internet users have read online reviews when they want to purchase a product, and between 73% and 87% users report that online product reviews had a significant influence on their purchase decisions (Pang & Lee, 2008). Unfortunately, the large number of reviews makes it difficult for online users to read to make an informed decision on whether to purchase the product or not. On the other hand, the large amount of reviews also increases the difficulty for product manufacturers to keep track of customers' opinions of their products (Miao, Li, & Dai, 2009). One way to solve the above problem is to mine customers' opinions automatically (Hu & Liu, 2004a, 2004b). Because customers may express their opinions on every attribute or property of a product, a customer's negative opinion on one property of product or service does not mean that the customer dislikes everything about the product, and vice-versa. Traditional sentiment analysis approaches such as document-level classification and sentence-level analysis are too coarse to provide deep analysis of product reviews. Consequently, fine-grained opinion mining is urgently demanded (Hu & Liu, 2004b; Miao, Li, & Dai, 2008a; Popescu & Etzioni, 2005).

Generally speaking, for a given product or service, there are often two main kinds of reviews on the Web. The first kind of reviews is written in a semistructured style. Epinions.com and Cnet.com use this format. The second kind of review is written in a free format and there is no separation of Pros and Cons, the reviewer can write freely. Amazon.com uses this format (Liu, Hu, & Cheng, 2005). For convenience of discussion, we refer to the first one as semistructured reviews and the second as unstructured reviews. Figures 1 and 2 show a sample semistructured and unstructured review, respectively.

From Figures 1 and 2, we can see that unstructured reviews tend to provide more abundant and detailed opinion information than semistructured reviews. However, due to the lack of effective natural language processing techniques, analyzing unstructured reviews pose significant challenges. In contrast, semistructured reviews are relatively easy for mining product features and opinions because semistructured reviews can be split into small opinion segments, and product features and opinions are explicit in each segment. In addition, because opinions are written separately in Pros and Cons, identifying the sentiment orientation is relatively easier. Unfortunately, semistructured reviews may not provide enough evidence to support customers' opinions. For example, we cannot find reasons why customers hold positive or negative opinions from most semistructured reviews. Therefore, integrating the advantages of multiple review sources may provide efficient solutions to fine-grained opinion mining.

Received April 9, 2010; revised June 7, 2010; accepted June 8, 2010

 • Published online 19 July 2010 in Wiley Online Library (wileyonlinelibrary.com). DOI: 10.1002/asi.21400

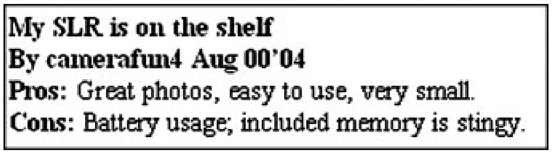
My SLR is on the shelf
By camerafun4 Aug 00'04
Pros: Great photos, easy to use, very small.
Cons: Battery usage; included memory is stingy.

FIG. 1. A sample semistructured review.

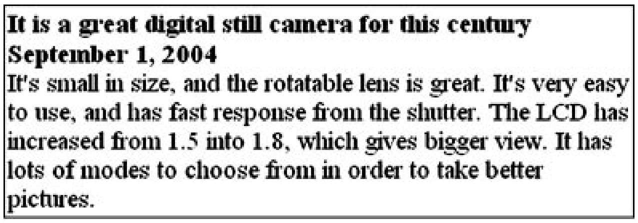
It is a great digital still camera for this century
September 1, 2004
It's small in size, and the rotatable lens is great. It's very easy to use, and has fast response from the shutter. The LCD has increased from 1.5 into 1.8, which gives bigger view. It has lots of modes to choose from in order to take better pictures.

FIG. 2. A sample unstructured review.

In the field of fine-grained opinion mining (e.g., Hu & Liu, 2004a; Popescu & Etzioni, 2005), most existing work has mainly focused on mining product features and opinions from a single review source, and has missed the point that the issue can be boosted by integrating multiple review sources. Consequently, it is important to investigate how to integrate multiple review sources for fine-grained opinion mining. Recently, researchers have started to investigate whether domain knowledge could help in fine-grained opinion mining. For instance, Carenini, Ng, and Zwart (2005) exploited user-defined taxonomy of product features as domain knowledge to extract product features and demonstrated that domain knowledge can improve the accuracy and reduce semantic redundancy of crude features. Shi and Chang (2006) also adopted a hierarchical product feature concept model to extracted product features. Lu and Zhai (2008) studied how to automatically integrate opinions expressed in a well-written expert review with various opinions scattering sources such as blog spaces and forums. However, how to extract and integrate product features and opinions from multiple review sources in a systematic manner remains to be investigated. Our first attempt in using integration strategy to extract product features and opinions shows promising results (Miao, Li, & Dai, 2008b). Following this research stream, we hypothesize that integrating semistructured reviews could help improve the performance of fine-grained opinion mining. More specifically, we will aim to answer the following research questions:

1. Can semistructured reviews be incorporated to improve the performance of fine-grained opinion mining from unstructured reviews?
2. How do we systematically exploit semistructured reviews to improve the performance of the fine-grained opinion mining?

Our investigation is based on an empirical study using two real-world datasets under different experimental conditions. To answer the first research question, we have built and compared two kinds of mining algorithms, one with utilizing semistructured reviews and the other without. To answer the second question, we propose an integration strategy to mine product features and opinions from multiple review sources. Within the integration strategy, we have developed a systematic approach to obtain and transform semistructured reviews into domain knowledge, and then exploit the knowledge to assist product feature extraction and opinion polarity identification from unstructured reviews.

The remainder of the article is organized as follows. In the next section we review the existing literature on sentiment analysis. In the third section, we introduce the basic terminologies of fine-grained opinion mining. In the fourth section, we introduce the architecture of the mining system. In the fifth section, we introduce the proposed approach and give an example. We have conducted comparative experiments and present the results in the sixth section. Last, we conclude the article with a summary of our work and future directions.

Literature Review

Speaking from a broad sense, "opinion mining" and "sentiment analysis" denote the same field of study (Pang & Lee, 2008). Due to the different objectives of applications, previous work tackled the sentiment analysis issue at different levels of granularity, from document-level sentiment classification, sentence-level sentiment analysis to fine-grained opinion mining. In this section, we will review the related work from three granularities.

Document-Level Sentiment Classification

In the field of document-level sentiment classification, one assumption is that opinionated document expresses positive or negative opinions on a single object and the task of document-level sentiment classification is to determine whether opinionated document expresses a positive or negative opinion based on the overall sentiments (Liu, 2009). Many approaches for document-level sentiment classification have been proposed, and they are usually categorized into supervised and unsupervised learning methods.

For supervised learning methods, Pang, Lee, and Vaithyanathan (2002) adopted machine learning techniques to classify movie reviews into positive and negative classes. Based on experiments, they found that Naïve Bayesian and Support Vector Machines (SVM) performed well in document-level sentiment classification when using unigrams as learning features. Mullen and Collier (2004) proposed a hybrid model that combined unigram with syntactic features. They used the SVM classifier and found that superior performance may be obtained by incorporating topic information into unigram models. Ng, Dasgupta, and Niaz Arifin (2006) incorporated linguistic knowledge into their supervised polarity classification system and their results suggested that bigrams and trigrams as well as manually tagged polarity information were effective learning features for

DOI: 10.1002/asi

document-level sentiment classification. Gamon (2004) also demonstrated that linguistic features consistently contributed to the accuracy in sentiment classification. Supervised methods are highly dependent upon the size and quality of training data, which is very costly and time-consuming to obtain. In addition, supervised methods are subjective to overtraining (Zhou & Chaovalit, 2008).

Turney (2002) proposed an unsupervised learning approach, which used PMI-IR (Turney, 2001) to calculate the sentiment orientation. The algorithm included three steps: (a) extracting phrases that contain adjectives or adverbs; (b) estimating the sentiment orientation of each phrase by point-wise mutual information (PMI)—the sentiment orientation of a phrase is computed based on its association with a positive or negative reference word such as *excellent* or *poor*; and (c) classifying reviews into positive or negative classes based on the average sentiment orientation of the phrases. Zhang, Zeng, Li, Wang, & Zuo (2009) proposed a rule-based approach on sentiment analysis of Chinese reviews. Unsupervised methods are efficient and require little training data, but their performance is generally lower than supervised methods.

Sentence-Level Sentiment Classification

Sentence-level sentiment analysis contains two subtasks. One is subjectivity classification: classify a sentence as a subjective or an objective sentence. The other subtask is a sentence-level sentiment classification, which determines the sentiment orientation of a sentence, positive or negative.

Most existing techniques for subjectivity classification are based on supervised learning. Hatzivassiloglou and Wiebe (2000) studied different adjective features for predicting subjectivity and found that lexical features such as sentiment orientation were a good indicator of subjective sentences. Riloff and Wiebe (2003) proposed a bootstrapping approach to learn extraction patterns for subjective expressions. Yu and Hatzivassiloglou (2003) adopted Bayesian classifier to determine the sentiment orientation of sentences.

For the subtask of sentence-level sentiment classification, Yu and Hatzivassiloglou (2003) determined the sentiment of a sentence by calculating the average sentiment orientation of opinion words in a sentence. The sentiment orientation of opinion words was calculated based on log-likelihood ratio. Kim and Hovy (2004) developed a system that used sentiment words to determine whether a sentence expresses a positive or negative opinion.

Fine-Grained Opinion Mining

Document-level sentiment classification and sentence-level sentiment analysis can provide qualitative analysis of sentiment information, whereas neither document-level nor sentence-level sentiment analysis can provide deep analysis of customers' opinions. Fine-grained opinion mining includes two subtasks: one is to identify product features that have been commented on by opinion holders and the other one is to determine the sentiment orientation of the opinions on features (Liu, 2009).

In the field of fine-grained opinion mining, many approaches for extracting product features and opinions have been proposed. Hu and Liu (2004b) proposed a two-staged approach: (a) extracting product features using association rule mining, and then extracting adjectives near product features as opinions; and (b) identifying the polarity of opinion words based on a lexicon. Popescu and Etzioni (2005) used a "Know It All" system to compute a confidential score for each product feature candidate and filtered some frequent noun phrases that may not be features and obtained better precision. Su, Xiang, Wang, Sun, and Yu (2006) studied the problem of extracting implicit features from customer reviews and developed a feature-based point-wise mutual information algorithm. Utilizing a polarity lexicon, they mapped each adjective in the lexicon to a set of predefined product features. Ghani, Probst, Liu, Krema, and Fano (2006) viewed the product features extraction problem as a classification problem. They adopted a semisupervised learning algorithm, which reduced the need for labeled data. Wang and Wang (2007) considered the fact that product properties and opinion words usually co-occur with high frequency in product reviews and proposed to bootstrap both of them using cross training. Feldman, Fresko, Netzer, and Ungar (2007) presented a case study in extracting information about comparisons. All of the above studies are based on the structure of language units or co-occurrence of features and opinions without considering domain knowledge, which had been proved to be essential to improve the performance of opinion mining systems (Bao, Li, Yu, & Cao, 2008; Raymond, Chapmann, Jian, & Yuefeng, 2009). Carenini et al. (2005) and Shi and Chang (2006) used user-defined taxonomy of features as domain knowledge to extract features from unstructured reviews. However, these user-defined taxonomies are difficult, time-consuming, and costly to produce, making such approaches difficult to scale and generalize to deal with rapidly increasing review sources and new products and associated new features. In this article, we take a different approach to automatically construct domain knowledge through exploiting semistructured reviews.

Lexicon-based and corpus-based approaches are two dominant methods for sentiment orientation identification (Liu, 2009). Hu and Liu (2004a) manually selected a set of seed adjectives, whose orientations are known, and then enlarged the seed set by searching in the WordNet. Zhou and Chaovalit (2008) developed an ontology-based approach to identify the sentiment orientation of opinions. Lexicon-based approaches often highly rely on expert-defined dictionaries of subjective words. However, subjective words are often context-dependent. For instance, *unpredictable* has a positive orientation in the context of *movie plot*, while the same sentiment has a negative orientation in the context of *automotive*. As such, using a general sentiment lexicon alone may not provide an effective solution for context-sensitive opinion mining. Corpus-based approaches rely on syntactic or co-occurrence patterns as well as an initial seed list of

DOI: 10.1002/asi

subjective words to identify other subjective words and their orientations in a large corpus. Hatzivassiloglou and McKeown (1997) adopted a log-linear regression model to infer the sentiment orientation of adjectives based on conjunctions. Wilson, Wiebe, and Hwa (2004) exploited subjectivity and syntax clues to predict the strength of sentiment orientation under a classification-based formulation. Ding, Liu, and Yu (2008) proposed a holistic lexicon-based approach for sentiment orientation identification and found that external evidences such as context dependent and linguistic conventions and language constructs are good clues for sentiment orientation identification. Such corpus-based approaches are able to find domain-specific opinion words and their orientations in limited contexts. Note, however, it is impossible to develop a comprehensive corpus to cover all possible subjective words in a general sense.

In this article, we study the problem of fine-grained opinion mining from multiple review sources. Specifically, we investigate how to obtain domain knowledge from semistructured reviews and exploit the domain knowledge to assist fine-grained opinion mining from unstructured reviews.

Terminology Definitions

In this section, we define the basic terminology of fine-grained opinion mining.

Definition 1: Semistructured Reviews

In semistructured reviews (referring to Figure 1), Pros and Cons are described separately, and the contents of Pros and Cons are often short phrases. Short phrases are separated by commas, periods, semicolons, hyphens, etc. We can view each semistructured review as a sequence of segments, where a segment could be a product feature or an opinion word or feature opinion pairs.

Formally, we denote a semistructured review by $sR = \{sg_1, sg_2, \ldots sg_m\}$, where $sg_i (1 \leq i \leq m)$ is a segment. Because we can always treat a phrase as a segment, this definition is quite reasonable.

Definition 2: Unstructured Reviews

Unstructured reviews (referring to Figure 2) are written in a free format and composed of several sentences. Each sentence contains product features and opinions and there is no separation of Pros and Cons.

Similarly, we formally define the unstructured review as $uR = \{s_1, s_2, \ldots s_n\}$, where $s_i (1 \leq i \leq n)$ is a sentence. This assumption is reasonable because a sentence may contain one or more than one product feature.

We give the definition of *feature*, *explicit feature*, *implicit feature*, and *opinion* the same as Ding et al. (2008).

Definition 3: Product Feature

A product contains a set of components and attributes, and each component may also have its subcomponents. For example, a digital camera has several components, such as a battery, a memory card, and a viewfinder. The viewfinder also has its attributes such as size and resolution. The components and attributes of products are generally called *features*.

Definition 4: Explicit Feature

If a feature f appears in the segment of review sR or uR, the feature is called an explicit feature of a product. For example, in the segment "the picture is wonderful," *picture* is an explicit feature.

Definition 5: Implicit Feature

If a feature f does not appear in the segment of review sR or uR, but is implied, the feature is called an implicit feature of a product. For example, in the segment "it is very expensive," *price* is an implicit feature, and *expensive* is a feature indicator.

Definition 6: Opinion Word

An opinion word on a feature f is a subjective word or phrase that expresses a positive or negative opinion. For example, the following segment expresses a positive opinion, "the picture quality is excellent," and in this segment the opinion word is *excellent*.

Definition 7: Sentiment Orientation

Sentiment orientation is the polarity of opinion words, namely positive or negative. For example, the sentiment orientation of the opinion word *excellent* in the above paragraph is positive, so we assign "+" to it. Note that sentiment orientation of opinion word is context sensitive, which means that opinion word may have totally different sentiment polarity in different circumstances. For example, the opinion word *small* is positive in segment "the size of the camera is small"; however, it is negative in segment "the memory is small."

Definition 8: Feature-Opinion Tuple

Given a collection of semistructured and unstructured reviews about a product or service, our goal is to extract and integrate product features and opinions from multiple reviews sources and identify the sentiment orientation of opinions. The mining and integrating results are represented as tuples:

$$\{(f_1, n_{f1} \ldots f_q, n_{fq}), (o_1^+ \ldots o_p^-)\},\ f_1 \ldots f_q \in F,$$
$$o_1^+ \ldots o_p^- \in O$$

where F is the set of product features and O is the set of opinion words; f_i is the i_{th} feature of a product and n_{fi} is the frequency of feature f_i in the collection of reviews; o_i is the opinion word of feature f_i and the superscript of o_i is the sentiment orientation. We use "+" and "−" to represent the positive and negative polarity, respectively.

DOI: 10.1002/asi

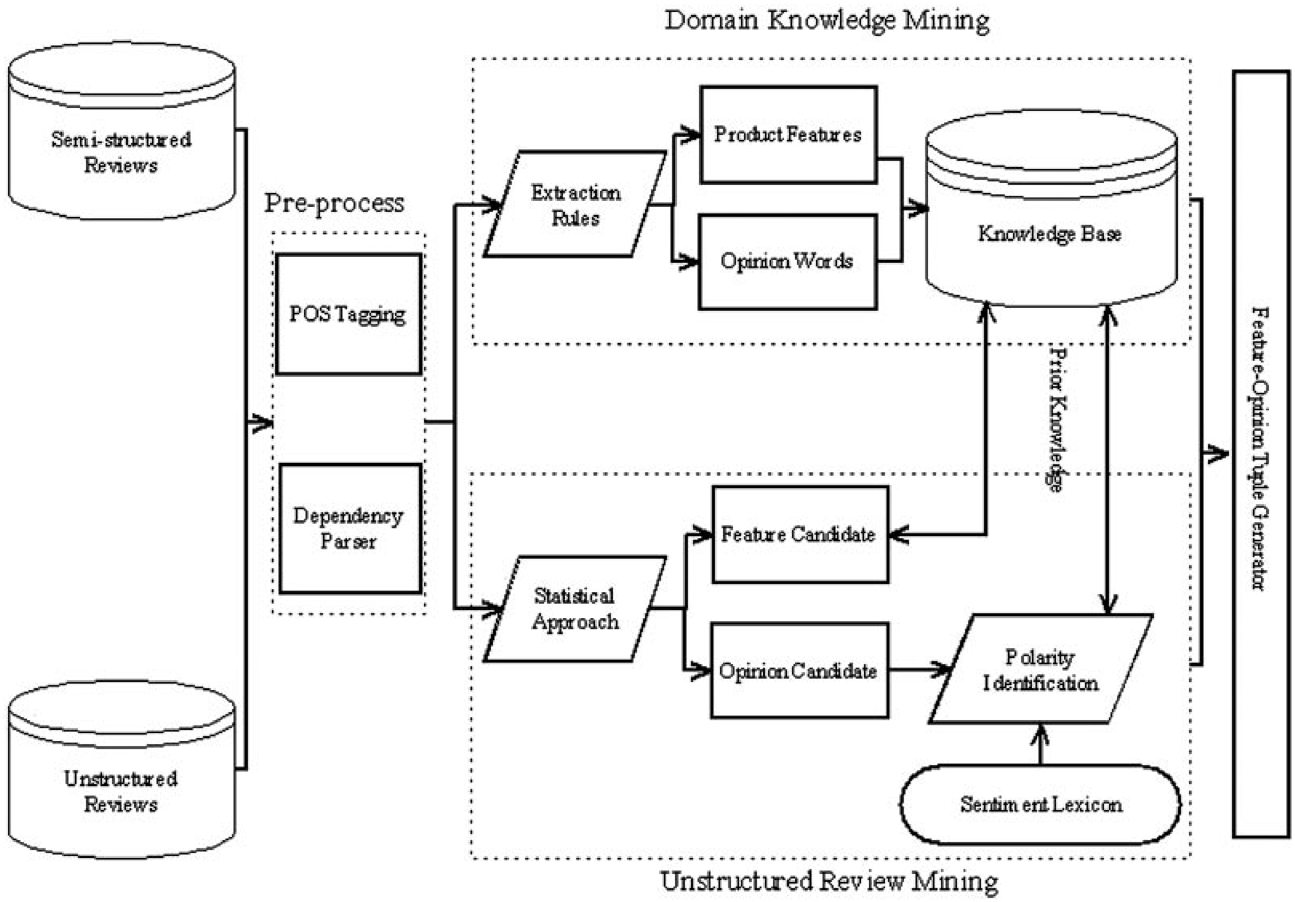

FIG. 3. The design framework of the proposed approach.

System Architecture

In this section, we will introduce the architecture of the system. The inputs of the system are semistructured and unstructured reviews, the outputs are product feature-opinion tuples. In particular the system consists of four parts: (a) review data server (RDS), (b) domain knowledge mining engine (DKME), (c) unstructured review mining engine (URME), and (d) feature-opinion tuple generator (FOG).

RDS collects review pages by periodically crawling the review Web sites such as Epinion.com and Amazon.com. These review pages are then cleaned to remove HTML tags, after that, reduplicate reviews are removed and the rest of the reviews are stored in the semistructured and unstructured review data server.

DKME extracts product features and opinion words from semistructured reviews. First, DKME extracts product features and opinion words by predefined extraction rules. Then synonyms are identified and grouped and sentiment orientations of opinion words are identified. Finally, feature-opinion tuples are generated and stored in a domain knowledge base.

URME has two inputs, unstructured reviews from RDS and domain knowledge from the knowledge base. URME exploits the domain knowledge to assist product feature extraction and opinion polarity identification.

FOG merges the newly discovered product features and opinion words from unstructured reviews into a knowledge base and generates feature-opinion tuples.

An Integration Approach for Fine-Grained Opinion Mining From Multiple Review Sources

In this section, we present the integration approach for mining product features and opinions from multiple review sources. Figure 3 shows the design framework of the proposed approach including two steps: First, we mine product features and opinions from semistructured reviews and build a domain knowledge base. Second, we use the domain knowledge to assist product features extraction and opinion polarity identification.

Domain knowledge mining

In this section, we introduce how to mine domain knowledge from semistructured reviews. Through observation, we find that semistructured reviews have the following two properties: First, semistructured reviews consist of subjective segments and each subjective segment includes explicit product features and opinion words. Second, because opinions are written separately in Pros and Cons, it is easier to identify

DOI: 10.1002/asi

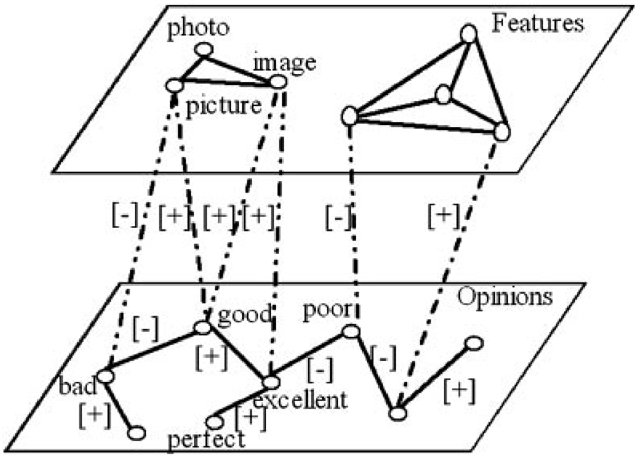

FIG. 4. The structure of domain knowledge base.

the polarity of opinions. With these properties in mind, the domain knowledge base can be built in three steps.

1. Extracting product features and opinions
2. Propagating product features and opinions
3. Associating product features and opinions

There are many semistructured reviews on the Web such as Epinion.com and CNet.com; they provide enough material for constructing our domain knowledge base.

Figure 4 illustrates the structure of our domain knowledge base. From Figure 4, we can see that the knowledge base includes a feature and opinion layer. The associations between product features and opinion words are represented as interlayer links. The feature layer contains product features that are distributed in cliques based on their semantic similarity. Each clique on a feature layer is a complete graph, and the weight of each link on the feature layer is equal. For example, *image*, *picture*, and *photo* are synonyms in the digital camera domain; therefore, they are in the same clique. The opinion layer contains opinion words and their sentiment associations. We use [+] and [−] to represent the positive and negative associations, respectively. The link on the opinion layer is generated based on feature cliques and feature-opinion pairs. In particular, given two features in the same clique, if the associations of the two feature-opinion pairs are the same, then a positive link is assigned between the two opinion words. Otherwise a negative link was assigned. For example, in Figure 4, given a feature clique *photo*, *image*, and *picture*, and two feature-opinion pairs *good image* and *bad picture*, the associations of the two pairs are different (feature-opinion pair *good image* has positive association, while *bad picture* has negative association); therefore, the link between *good* and *bad* is negative.

As product features and opinions are explicit in semistructured reviews, we manually develop some extraction rules to extract them. Some key extraction rules and corresponding instances are shown in Table 1. We first use NLProcessor linguistic parser (http://www.infogistics.com/textanalysis.html) and Stanford parser (http://nlp.stanford.edu/software/lex-parser.shtml) to generate part of speech and dependency tree for each subjective segment in semi-structured reviews. These two NLP processing tools can generate part of speech tags and dependency trees automatically. And then, according to the part of speech tags and syntactic chunks, we extract high-frequency nouns, noun phrases, adjectives, adverbs, and verbs that match the extraction rules. We manually examine the results and remove those that are not a product feature and opinion based on their definitions given in the Terminology Definition section. Finally, product features that have semantic similarity are grouped in cliques by domain experts. For example, in the digital camera domain, *image*, *photo*, and *picture* are grouped together. We also map implicit features to its actual feature, for example, for implicit features like *expensive*, we associate the actual feature *price* to it. The sentiment orientations of opinion words are identified based on Pros or Cons. Note that, if negative words such as NOT, or NO exist in the segment, we inverse the sentiment orientation of opinion words. We also add some subjective verbs such as *love*, *recommend*, or *hate* into the opinion layer. Consequently, we obtain product features, opinion words, and their associations.

TABLE 1. Key extraction rules and instances.

Rules	POS Tags	Instances
N → F	NN, NNS	Picture
NN → F	NN NN	Battery life
JN → O F	JJ NN(NNS)	Poor quality
JNN → O F	JJ NN NN	Small memory card
JTB → O to F	JJ TO VB	Difficult to use

Note. F = Product feature; O = opinion word; NN = noun; NNS = noun plural; JJ = adjective; VB = verb.

To enlarge the coverage of our knowledge base, we need to propagate product features and opinion words to their synonyms and antonyms. By propagating each product feature and opinion to their semantically related words, new associations can be formed. Figure 4 illustrates an example. In Figure 4, from the given product feature and opinion pair *good picture*, we can propagate *picture* to its synonym *image*. Similarly, we can obtain the synonyms and antonyms of the opinion word *good*. Consequently, we can acquire a new product feature and opinion pairs such as *excellent image*, *good image*, etc.

Obviously, the strengths of association between different product features and opinion words are different. Moreover, some new feature–opinion pairs generated from the propagation might contain noise. To reduce the impact of noise, interlayer links should be assigned different weights to reflect different confidence. In this research, we use the co-occurrence frequency of feature–opinion pairs in the entire review collection as confident weights and assign them to interlayer links.

Finally, we build the knowledge base including product features, opinion words, and their relations. This knowledge

DOI: 10.1002/asi

base is used to assist product feature extraction and opinion polarity identification.

Unstructured Reviews Mining

In this section, we will make use of the domain knowledge derived from semistructured reviews to assist product features extraction. Previous work points out that product features in unstructured reviews have the following three characteristics: (a) product features are usually nouns and noun phrases such as *battery* and *picture quality* (e.g., Liu, 2009; Popescu & Etzioni, 2005); (b) product features occur more frequently than background words in product reviews (e.g., Hu & Liu, 2004a; Liu et al., 2005); (c) product features are similar words or have a semantic similarity, e.g., *image*, *photo*, and *picture* refer to the same product feature in the digital camera domain (e.g., Carenini et al., 2005; Miao et al., 2008b). With these properties in mind, we developed a hybrid approach to extract product features from unstructured reviews. Specifically, we extract high-frequency words and phrases that match extraction rules and then filter out the feature candidates, which are not true product features based on our domain knowledge base. In particular, each product feature candidate is given a confident score that shows how confident the candidate is a product feature. The confident score is computed by linguistic similarities defined below. The unstructured review mining algorithm is given in Figure 5.

Considering the properties of product features, we adopt two kinds of similarity metrics, one is semantic similarity and the other is string similarity.

Semantic similarity. We use WordNet to compute the semantic similarity between two words. In WordNet, the lexical items are organized according to the part of speech and sense (Budanitsky & Hirst, 2001; Miller, 1995). Pirro and Seco (2008) proposed a novel semantic similarity metric. They exploited the notion of intrinsic information content which quantifies information content values by scrutinizing how concepts are arranged in an ontological structure. Their semantic similarity is defined as follows:

$$semanticSim(c_i, c_j) = 3IC(msca(c_i, c_j)) - IC(c_i) - IC(c_j)$$

$$IC(c) = 1 - \frac{\log(hypo(c)) - 1}{\log(\max_{wn})}$$

where max_{wn} is a constant that indicates the total number of concepts in the noun taxonomy of WordNet. Function *hypo* returns the number of hyponyms of a given concept c and *msca* is most specific common abstraction that subsumes both concepts.

Phrase similarity. As discussed above, we can see that some product features are phrases; therefore, phrase similarity measure is required. Gal, Modica, Jamil, and Eyal (2005) proposed a symmetric measure of phrase similarity, in their measure, the similarity of two phrases is defined as the ratio between the number of common words in phrases p_i and p_j (γ_{com}) and the total number of unique words in phrases p_i and p_j (γ_{uni}). The measure is defined as follows:

$$phraseSim(p_i, p_j) = \frac{\gamma_{com}}{\gamma_{uni}}$$

Unstructured review mining algorithm

```
INPUT: Unstructured Reviews (uR) and Domain Knowledge (DK)
OUTPUT: Feature-Opinion Tuples
FOR each sentence s ∈ uR
do pos tagging and dependency parser
extract words and phrases match extraction rules
compute the frequency of words f(w) and phrases f(p)
IF f(w)>threshold, put w into feature candidate set (FCS)
IF f(p)>threshold, put p into feature candidate set (FCS)
FOR each fc ∈ FCS
  IF (fc is a word)
     FOR each feature f ∈ DK
       IF (fc matches f)
        merge fc to f
        DK=DK+fc
        FCS=FCS-fc
       ELSE
       compute the semanticSim (fc ,f)
       merge fc to f =max (semanticSim (fc , f))
       DK=DK+fc
       FCS=FCS-fc
 ELSE
    IF (fc is a phrase)
    FOR each feature f ∈ DK
     IF (fc matches f)
        merge fc to f
        DK=DK+fc
        FCS=FCS-fc
      ELSE
      compute the phrase Sim (fc , f)
      merge fc to f =max (phrase Sim (fc , f))
      DK=DK+fc
      FCS=FCS-fc
```

FIG. 5. Unstructured review mining algorithm.

Identifying sentiment orientation

In this section, we will exploit the domain knowledge derived from semistructured reviews to assist opinion polarity identification.

There are several public sentiment thesauruses from the community of linguistics, such as the General Inquirer (Stone, Dunphy, Smith, & Ogilvie, 1966). Unfortunately, these public sentiment thesauruses cannot be applied directly because the sentiment orientation of the opinion word is context sensitive. For example, the opinion word *high* in the sentence "the product is of high quality" expresses a positive sentiment, whereas in the sentence "the price is very high" indicates negative opinions. As discussed above, the domain knowledge base not only contains opinion words and their polarities, but also has product features and feature opinion associations, which is essential for context sensitive opinion words. We divide the problem into four cases in accordance with whether product features and opinions can be found in the domain knowledge base. In Cases 1 and 2,

DOI: 10.1002/asi

Sentiment orientation identification algorithm

```
INPUT: Product Feature (f), Opinion Word (o), Domain Knowledge (DK)
       Sentiment Lexicon (SL) and WordNet (WN)
OUTPUT: Sentiment Orientation of Opinion Word o
1:  CASE 1: (f ∈ DK and o ∈ DK)
2:  determine the sentiment orientation of o
3:  CASE 2: (f ∉ DK and o ∈ DK)
4:  IF (o has associated opinion words)
5:    find opinion word set oS associated with o
6:    determine the sentiment orientation of o by the polarity of oS
7:  ELSE
8:  GOTO STEP 11
9:  CASE 3: (f ∈ DK and o ∉ DK)
10: CASE 4: (f ∉ DK and o ∉ DK)
11:   IF (o ∈ SL)
12:   determine the sentiment orientation of o
13:   ELSE
14:     search synonym of o in WN and form synonym set (sS)
15:     search antonym of o in WN and form antonym set (sA)
16:     FOR each word w ∈ sS
17:     IF (w ∈ DK or w ∈ SL)
18:     the sentiment orientation of o is the same as w
19:     add o to DK
20:     ELSE
21:     FOR each word w ∈ sA
22:       IF (w ∈ DK or w ∈ SL)
23:       the sentiment orientation of o is the opposition of w
24:       add o to DK
25:       ELSE
26:         compute score = PMI-IR(o, POSITIVE)
27:                         - PMI-IR(o, NEGATIVE)
28:         IF score>0
29:         o=POSITIVE
30:         ELSE
31:         o=NEGATIVE
32:         add oW to DK
```

FIG. 6. Sentiment orientation identification algorithm.

opinion words can be found in the domain knowledge base, therefore, the polarity can be identified directly. In Cases 3 and 4 because opinion words are not in the domain knowledge base, the polarity of opinion words can be identified by using external sentiment lexicons and WordNet. Figure 6 shows the sentiment orientation identification algorithm.

An Example

In this section, we will illustrate the proposed approach by using digital camera reviews shown in Figures 1 and 2.

Figure 7 shows the part of speech analysis results of the semistructured review shown in Figure 1. Product features of semistructured review are *photo*, *use*, *battery*, and *memory*, and opinion words are *great*, *easy*, *small*, and *stingy*. We can obtain three feature-opinion pairs: {photo, great}, {use, easy}, {memory, stingy}, but feature *battery* and opinion word *small* cannot form feature–opinion pairs because there are no corresponding features or opinion words. For implicit feature *small*, we associate actual feature *size* to it. For the product feature *battery* usage, there is no opinion word because *battery usage* is in Cons, so we assign *battery usage* with opinion word *bad*. The results are shown in Table 2.

Treating the mining results of semistructured reviews as domain knowledge, we can exploit the domain knowledge to assist feature extraction from unstructured reviews. First,

```
<P><S><W chunk='NGstart' C='JJ' T='w'
S='Y'>Great</W> <W chunk='NGend'
C='NNS'>photos</W><W C=','>,</W> <W
chunk='VGstart' C='RB'>easy</W> <W chunk='VGin'
C='TO'>to</W> <W chunk='VGend' C='VB'>use</W><W
C=','>,</W> <W C='RB'>very</W> <W
C='JJ'>small</W><W C='.' T='.'>.</W></S>
<S><W chunk='NGstart' C='NN' T='w'
S='Y'>Battery</W> <W chunk='NGend'
C='NN'>usage</W><W C=':'>;</W> <W chunk='NGstart'
C='JJ'>included</W> <W chunk='NGend'
C='NN'>memory</W> <W chunk='VGstart_end'
C='VBZ'>is</W> <W C='JJ'>stingy</W><W C='.'
T='.'>.</W></S>
</P>
```

FIG. 7. Part of speech analysis results of the semi-structured review shown in Figure 1.

TABLE 2. The mining results of semistructured reviews.

Product features	Opinion words	Orientations
Photo	Great	Positive
Use	Easy	Positive
Size	Small	Positive
Memory	Stingy	Negative
Battery usage	Bad	Negative

we extract 11 candidate product features and six adjective words as candidate opinion words based on POS tags and syntax structure. In the candidate feature set, *size* and *use* can be easily identified as product features based on domain knowledge. Candidate feature *picture* is identified as product feature by computing the semantic similarity. Using the algorithm in Figure 5, we obtain nine feature opinion pairs. Based on the domain knowledge and sentiment thesauruses, we can identify the sentiment orientations of opinion words by algorithm in Figure 6. If there is no opinion word associates with product feature in a sentence, we assign sentiment orientation according to the product feature's context. Specifically, we do it automatically by considering the conjunctions such as *and*, *but*, *either or,* and *neither nor*. For example, in the sentence "It's very easy to use, and has fast response from the shutter," there is no opinion word associate with product feature *shutter*; however, the polarity of opinion words *easy* and *fast* are positive in the sentence, therefore, we can assign opinion word *good* to product feature *shutter*. Table 3 shows the mining results of unstructured reviews.

Table 4 shows the results of feature-opinion tuples. From Table 4, we can see that product features *photo* and *picture* denote the same product feature, and they are merged into one cluster. The frequency of product features is also presented in Table 4. From the frequency of product features, we can intuitively know which product feature is most concerned by customers.

Experiment

In this section, we report an experimental study aimed at answering the research question: Can semistructured

DOI: 10.1002/asi

reviews be incorporated as domain knowledge to improve the performance of fine-grained opinion mining?

We first describe two datasets used in our study and then we give our experiment results to demonstrate the effectiveness of our integrating strategy for fine-grained opinion mining from multiple review sources.

TABLE 3. The mining results of unstructured reviews.

Product features	Opinion words	Orientations
Size	Small	Positive
Lens	Great	Positive
Use	Easy	Positive
Response	Fast	Positive
Shutter	Good	Positive
LCD	Good	Positive
View	Bigger	Positive
Mode	Good	Positive
Picture	Better	Positive

TABLE 4. The results of feature–opinion tuples.

Product features	Opinion words	Orientations
Photo [1], picture [1]	Great, good	Positive
Size [2]	Small	Positive
Use [2]	Easy	Positive
Memory [1]	Stingy	Negative
Len [1]	Great	Positive
Response [1]	Fast	Positive
Shutter [1]	Good	Positive
View [1]	Bigger	Positive
Mode [1]	Good	Positive
LCD [1]	Good	Positive
Battery usage [1]	Bad	Negative

Data Preparation and Evaluation Metrics

We conduct our experiments using two datasets. The first dataset concerns electronic products including digital cameras, cell phones, mp3 players, and DVD players. The second dataset is on movies and music. Semistructured reviews are crawled from Epinion.com. The Amazon Web Services APIs are used to collect unstructured reviews from Amazon.com. The Amazon standard identification numbers (ASINs), associated with the products under study, are fed through the Amazon Web services API, which returns unstructured reviews. Some key statistics about the two datasets are shown in Table 5. Note that there are some duplicate reviews and low quality reviews in our data set; we built a simple filter to filter the duplicate reviews and low quality reviews.

Three graduate students with a background in information systems were recruited to annotate the reviews contained in our datasets. All the students were given detailed instructions about what are product features and opinion words. The statistics of the annotated results are shown in Table 6.

To verify the degree of agreement among the three annotators, we adopt Fleiss' kappa (Sim & Wright, 2005) to evaluate the consistency of annotated results. Table 7 summarizes the kappa values.

From Table 7, we note that for both datasets, the average kappa scores are 0.85 and 0.76, respectively, which indicate strong agreements. To construct the final gold standard, we adopted the following procedure. For reviews that

TABLE 5. Descriptive statistics of the semistructured and unstructured review dataset.

	Semistructured reviews		Unstructured reviews	
Product	No. of reviews	Average length (words)	No. of sentences	Average length (words)
Camera	464	8	234	20
Cell phone	480	7	226	17
MP3 player	446	7	275	17
DVD player	410	7	209	16
Movie	480	6	283	17
Music	482	6	244	15
Average	464	8	234	20

TABLE 6. Descriptive statistics of the annotated results of semiunstructured and unstructured review dataset.

	Semistructured reviews		Unstructured reviews	
Product	No. of positive reviews	No. of negative reviews	No. of positive sentences	No. of negative sentences
Camera	271	193	184	50
Cell phone	258	222	162	64
MP3 player	296	150	161	114
DVD player	294	116	107	102
Movie	235	245	210	73
Music	242	240	196	48
Average	266	194	170	75

TABLE 7. The values of Fleiss' kappa of semistructured and unstructured review dataset.

	Fleiss' kappa	
Product name	Semistructured review dataset	Unstructured review dataset
Camera	0.85	0.78
Cell phone	0.86	0.76
Mp3 player	0.88	0.77
DVD player	0.85	0.76
Movie	0.82	0.74
Music	0.83	0.75
Average	0.85	0.76

TABLE 8. The mining results of semistructured and unstructured reviews.

	F-measure		
Product	Semistructured review mining	Unstructured review mining with domain knowledge	Unstructured review mining without domain knowledge
Camera	0.892	0.861	0.826
Cell phone	0.906	0.887	0.841
MP3 player	0.922	0.888	0.814
DVD player	0.936	0.881	0.858
Movie	0.952	0.865	0.766
Music	0.939	0.863	0.815
Average	0.925	0.874	0.820

TABLE 9. p-Values for pairwise t-test on F-measure of entertainment and electronic product domain.

Product name	p-Value
Camera	<0.01
Cell phone	<0.01
MP3 player	<0.01
DVD player	<0.01
Movie	<0.01
Music	<0.01

Note. p-Values significant at corrected threshold 0.01.

have received the same labels or annotations from all three annotators, we assigned this agreed-upon label. For a small number of reviews that have received differing assessment from these three annotators, we had all three annotators go through these reviews and discuss their assessment with each other in a face-to-face meeting. We then used their consensual assessment as the final label. In all cases, these three annotators were able to converge on the label with which everyone feels comfortable.

Based on the above manually constructed gold standard, the F-measure is used in our experiments to evaluate the effects of integrating semi-structured reviews. The F-measure is defined as follows:

$$F - measure = \frac{2precision \times recall}{precision + recall}$$

$$precision = \frac{Ra}{A} \quad recall = \frac{Ra}{R}$$

where Ra is the relevant product features which has been retrieved by the system, R is all relevant product features about the product in a review, and A is the product features the system retrieved from a review.

Can semistructured reviews be incorporated as domain knowledge to improve the performance of product feature extraction?

To answer this research question, we have built and compared two kinds of mining algorithms, one with utilizing semistructured reviews and the other without.

Table 8 shows the mining results of semistructured and unstructured reviews, respectively. From Table 8 we can see that the F-measure in semistructured reviews is relatively high in both datasets, which indicates that we can obtain domain knowledge from semistructured reviews effectively. As semistructured reviews are too short to provide enough opinion information for many interesting applications such as comparative opinion mining and causal relation mining, mining and integrating more relevant opinion information from unstructured reviews is crucial.

In Table 8, columns 3 and 4 show the mining results of unstructured reviews without and with domain knowledge. We can see that when incorporating domain knowledge, the average F-measure can be improved by 7.4% and 4.5% in entertainment and electronic product domain, respectively. The improvement of the F-measure is likely attributable to the domain knowledge, which is derived from semistructured reviews.

To validate whether the improvement is significant, we adopt pair-wise t-tests on the F-measure. Table 9 shows the p-values of the pair-wise t tests on the $F-$measure. For all t-tests, only p-values less than 0.01 are considered significant. From Table 9, we can see that p-values of two datasets are all less than 0.01, therefore the improvement is significant. We confirm that the improvement of the $F-$measure is due to incorporating semistructured reviews and we believe that it will achieve better results if we incorporate enough high-quality semistructured reviews to the system.

Can semistructured reviews be incorporated as domain knowledge to improve the performance of sentiment orientation identification?

To evaluate the effectiveness of the knowledge base on opinion polarity identification, we adopt two lexicon-based approaches as baselines. For the first baseline, we use a public sentiment lexicon General Inquirer (Stone et al., 1966), which contains 1,915 positive words and 2291 negative words. For the second baseline, we first select some subjective adjectives with explicit orientations as a seed set and then enlarge the seed set by searching the synonyms and antonym from

TABLE 10. Results of sentiment classification of entertainment and electronic product domain.

Product	*F*-measure		
	General Inquirer based approach	WordNet-based approach	The proposed approach
Camera	0.818	0.888	0.897
Cell phone	0.869	0.849	0.888
MP3 player	0.720	0.766	0.866
DVD player	0.686	0.705	0.874
Movie	0.764	0.852	0.90
Music	0.791	0.884	0.915
Average	0.775	0.824	0.890

WordNet. Finally, we use the enlarged adjective set to identify the sentimental orientation of opinions.

Table 10 shows the results of sentiment identification of entertainment and electronic product domain respectively. From Table 10 we can see that our knowledge-based approach outperforms the baselines in all products. Through analyzing the results, we find that the baselines fail to recognize some domain sensitive opinion words, whereas our approach can identify the polarity of these opinions because product feature information is considered in our approach.

From the above analysis, it is evident that semistructured reviews can be incorporated as domain knowledge to improve the performance of sentiment orientation identification.

The preliminary experimental results show that our proposed approach is effective in fine-grained opinion mining; however, there are still some mistakes in product feature extraction and sentiment orientation identification. Through careful analysis, we find that errors in product feature and opinion extraction are due to complex and unusual sentence structure and ambiguous pronouns. For example, our models fail to recognize "it" in the sentence "it's so worth it!." Therefore, resolutions on pronoun are needed. In addition, the part of speech tagger and dependency parser failed to label some weird expressions. Errors in sentiment orientation identification are due to the limited coverage of the domain knowledge base. We fail to recognize some cases when idiomatic or vague expressions were used to express opinions. For example, in the sentence "It cost me an arm and a leg," our sentiment lexicon cannot identify the sentiment orientations correctly. We need to enlarge our domain knowledge base with some idioms. Moreover, we find that some subjective verbs, such as *recommend* and *love* are important indicators to sentiment identification.

Conclusion and Future Directions

In this study, we investigate how to mine product features and opinions from multiple review sources and propose an integration strategy. The intuition of this strategy is to combine the advantages of both semistructured and unstructured reviews. Within this strategy, we first extract product features and opinions from semistructured reviews and build domain knowledge base, and then we exploit the domain knowledge to boost up the mining process of unstructured reviews. Finally, feature-opinion tuples are generated. Experimental results on two real-world datasets show promising results and demonstrate semistructured reviews can be incorporated as domain knowledge to improve the performance of fine-grained opinion mining from unstructured reviews. We believe that this study is just the first step in integrating multiple review sources and much more work is needed to further explore the issue. In future research, we plan to use more sophisticated natural language processing techniques such as anaphora resolution to the mining process. We also plan to address the problem of integrating customer reviews with other free-form evaluative texts such as blog articles and communities.

Acknowledgment

This research is supported by the 863 project (No. 2006AA010106), 973 project (No. 2007CB311007), NNSFC projects (70890084, 60875049 and 60703085), and the CAS grant (2F07C01).

References

Bao, S., Li, R., Yu, Y., & Cao, Y. (2008). Competitor Mining with the Web. IEEE Transactions on Knowledge and Data Engineering, 20(10), 1297–1310.

Budanitsky, A., & Hirst, G. (2001, June). Semantic distance in Wordnet: An experimental, application-oriented evaluation of five measures. Paper presented at the Workshop on WordNet and Other Lexical Resources, Second Meeting of the North American Chapter of the Association for Computational Linguistics, Pittsburgh, PA.

Carenini, G., Ng, R.T., & Zwart, E. (2005). Extracting knowledge from evaluative text. In Proceedings of the Third International Conference on Knowledge Capture (pp. 11–18). New York: ACM Press.

Ding, X., Liu, B., & Yu, P.S. (2008). A holistic lexicon-based approach to opinion mining. In Proceedings of the First ACM International Conference on Web Search and Data Mining (pp. 231–240). New York: ACM Press.

Feldman, R., Fresko, M., Netzer, O., & Ungar, L. (2007). Extracting product comparisons from discussion boards. In Proceedings of the Seventh IEEE International Conference on Data Mining (pp. 469–474). Piscataway, NJ: IEEE.

Gal, A., Modica, G., Jamil, H., & Eyal, A. (2005). Automatic ontology matching using application semantics. AI Magazine, 26(1), 21–31.

Gamon, M. (2004). Sentiment classification on customer feedback data: Noisy data, large feature vectors, and the role of linguistic analysis. In

Proceedings of the 20th International Conference on Computational Linguistics (pp. 841–847). College Park, MD: Association for Computational Linguistics.
Ghani, R., Probst, K., Liu, Y., Krema, M., & Fano, A. (2006). Text mining for product attribute extraction. Proceedings of ACM SIGKDD Explorations Newsletter, 8(1), 41–48.
Hatzivassiloglou, V., & McKeown, K.R. (1997). Predicting the semantic orientation of adjectives. In Proceedings of the Eighth Conference on European Chapter of the Association for Computational Linguistics (pp. 174–181). College Park, MD: Association for Computational Linguistics.
Hatzivassiloglou, V., & Wiebe, J. (2000). Effects of adjective orientation and gradability on sentence subjectivity. In Proceedings of the International Conference on Computational Linguistics (pp. 299– 305). College Park, MD: Association for Computational Linguistics.
Hu, M., & Liu, B. (2004a). Mining and summarizing customer reviews. In Proceedings of the Tenth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (pp. 168–177). New York: ACM Press.
Hu, M., & Liu, B. (2004b). Mining opinion features in customer reviews. In Proceedings of the 19th National Conference on Artificial Intelligence (pp. 755–760). Cambridge, MA: MIT Press.
Kim, S.M., & Hovy, E. (2004). Determining the sentiment of opinions. In Proceedings of the International Conference on Computational Linguistics (Article No. 1367). College Park, MD: Association for Computational Linguistics.
Lau, R.Y.K., Chapmann, C.L.L., Jian, M., & Yuefeng, L. (2009, December). Automatic domain ontology extraction for context-sensitive opinion mining. Paper presented at the International Conference on Information Systems, Phoenix, AZ.
Liu, B. (2009). Sentiment analysis and subjectivity. In N. Indurkhya, & F.J. Damerau (Eds.), the Handbook of Natural Language Processing (2nd ed., pp. 627–666). New York: Chapman & Hall.
Liu, B., Hu, M., & Cheng, J. (2005). Opinion observer: Analyzing and comparing opinions on the web. In Ellis, A., & Hagino, T. (Eds.), Proceedings of the 14th International World Wide Web Conference (pp. 342–351). New York: ACM Press.
Lu, Y., & Zhai, Ch. (2008). Opinion integration through semi-supervised topic modeling. In Proceedings of the 17th International World Wide Web Conference (pp. 121–130). New York: ACM Press.
Miao, Q.L., Li, Q.D., & Dai, R.W. (2008a). A unified framework for opinion retrieval. In Proceedings of the 2008 IEEE/WIC/ACM Conferences on Web Intelligence (pp. 739–742). Piscataway, NJ: IEEE.
Miao, Q.L., Li, Q.D., & Dai, R.W. (2008b). An integration strategy for mining product features and opinions. In Shanahan et al. (Eds.), Proceedings of the International Conference on Information and Knowledge Management (pp. 1369–1370). New York: ACM Press.
Miao, Q.L., Li, Q.D., & Dai, R.W. (2009). AMAZING: A sentiment mining and retrieval system. Expert Systems with Applications, 36(3), 7192–7198.
Miller, G.A. (1995). WordNet: A lexical database for English. Communications of the ACM, 38(11), 39–41.
Mullen, T., & Collier, N. (2004). Sentiment analysis using support vector machines with diverse information sources. In M. Lapata & H.-T. Ng (Eds.), Proceedings of the Conference on Empirical Methods in Natural Language Processing (pp. 412–418). College Park, MD: Association for Computational Linguistics.
Ng, V., Dasgupta, S., & Niaz Arifin, S.M. (2006). Examining the role of linguistic knowledge sources in the automatic identification and classification of reviews. In Proceedings of the 44th Annual Meeting of the Association for Computational Linguistics (pp. 611–618). College Park, MD: Association for Computational Linguistics.
Pang, B., & Lee, L. (2008). Opinion mining and sentiment analysis. Foundations and Trends in Information Retrieval, 2(1–2), 1–135.
Pang, B., Lee, L., & Vaithyanathan, S. (2002). Thumbs up? Sentiment classification using machine learning techniques. In M. Lapata & H.-T. Ng (Eds.), Proceedings of the Conference on Empirical Methods in Natural Language Processing (pp. 79–86). College Park, MD: Association for Computational Linguistics.
Pirro, G., & Seco, N. (2008). Design, implementation and evaluation of a new semantic similarity metric combining features and intrinsic information content. In Proceedings of the OTM 2008 Confederated International Conferences, CoopIS, DOA, GADA, IS, and ODBASE 2008. Part II On the Move to Meaningful Internet Systems. Lecture Notes in Computer Science, Vol. 5332, 1271–1288.
Popescu, A.M., & Etzioni, O. (2005). Extracting product features and opinions from reviews. In R.J. Mooney (Ed.), Proceedings of the Conference on Human Language Technology and Empirical Methods in Natural Language Processing (pp. 339–346). College Park, MD: Association for Computational Linguistics.
Riloff, E., & Wiebe, J. (2003). Learning extraction patterns for subjective expressions. In M. Lapata & H.-T. Ng (Eds.), Proceedings of the Conference on Empirical Methods in Natural Language Processing (pp. 105–112). College Park, MD: Association for Computational Linguistics.
Shi, B., & Chang, K. (2006). Mining Chinese reviews. In Proceedings of the Sixth IEEE International Conference on Data Mining (pp. 585–589). Piscataway, NJ: IEEE.
Sim, J., & Wright, C.C. (2005).The kappa statistic in reliability studies: Use, interpretation, and sample size requirements. Physical Therapy, 85(3), 257–268.
Stone, P.J., Dunphy, D.C., Smith, M.S., & Ogilvie, D.M. (1966). The general inquirer: A computer approach to content analysis. Cambridge, MA: MIT Press.
Su, Q., Xiang, K., Wang, H., Sun, B., & Yu, Sh. (2006). Using pointwise mutual information to identify implicit features in customer reviews. In Y. Matsumoto (Ed.), Computer Processing of Oriental Languages: Beyond the Orient: The Research Challenges Ahead: 21st International Conference (ICCPOL 2006). Lecture Notes in Computer Science, 4285, 22–30.
Turney, P.D. (2001). Mining the web for synonyms: PMI-IR versus LSA on TOEFL. In L. De Raedt & P.A. Flach (Eds.), Proceedings of the 12th European Conference on Machine Learning (pp. 491–502). Berlin, Germany: Springer-Verlag.
Turney, P.D. (2002). Thumbs up or Thumbs down? Semantic orientation applied to unsupervised classification of reviews. In Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics (pp. 417–424). College Park, MD: Association for Computational Linguistics.
Wang, B., & Wang, H. (2007). Bootstrapping both product properties and opinion words from Chinese reviews with cross-training. In Proceedings of 2007 IEEE/WIC/ACM International Conference on Web Intelligence (pp. 259–262). Piscataway, NJ: IEEE.
Wilson, T., Wiebe, J., & Hwa, R. (2004). Just how mad are you? Finding strong and weak opinion clauses. In Proceedings of the 19th National Conference on Artificial Intelligence (pp. 761–769). Cambridge, MA: MIT Press.
Yu, H., & Hatzivassiloglou, V. (2003). Towards answering opinion questions: Separating facts from opinions and identifying the polarity of opinion sentences. In M. Lapata & H.-T. Ng (Eds.), Proceedings of the Conference on Empirical Methods in Natural Language Processing (pp. 129–136). College Park, MD: Association for Computational Linguistics.
Zhang, Ch., Zeng, D., Li, J., Wang, F.-Y., & Zuo, W. (2009). Sentiment analysis of Chinese documents: From sentence to document level. Journal of the American Society for Information Science and Technology, 60(2), 2474–2487.
Zhou, L., & Chaovalit, P. (2008). Ontology-supported polarity mining. Journal of the American Society for Information Science and Technology, 59(1), 1–13.

DOI: 10.1002/asi