Available online at www.sciencedirect.com

ScienceDirect

Computer Speech & Language 55 (2019) 1–25

www.elsevier.com/locate/csl

Exploiting social and local contexts propagation for inducing Chinese microblog-specific sentiment lexicons☆

Chuanjun Zhao[a], Suge Wang*,[a,b], Deyu Li[a,b]

[a] School of Computer and Information Technology, Shanxi University, Taiyuan 030006, Shanxi, China

[b] Key Laboratory of Computational Intelligence and Chinese Information Processing of Ministry of Education, Taiyuan 030006, Shanxi, China

Received 30 December 2017; received in revised form 3 July 2018; accepted 26 October 2018

Available online 8 November 2018

Abstract

Sentiment lexicons including opinion words, sentiment phrases, and idioms with sentiment polarities play an important role in sentiment analysis tasks. Apart from explicit sentiment features, extracting implicit sentiment features is a challenging research issue. The sentiment expression is very domain-specific, and constructing a general sentiment lexicon that is suitable for all domains is hard or even impossible. In this paper, we propose a novel sentiment unit context propagation framework to extract Chinese microblog-specific explicit and implicit sentiment features. In the process of the selection of seed sentiment units, we select the seed sentiment units that have a large standard degree of centrality with other units, and mark these units with sentiment labels using general sentiment lexicons and manual calibrations. To realize sentiment label propagation from a small amount of labeled sentiment units to unlabeled ones, we exploit local contexts, topic features, and so‘cial relationships among users in microblog social networks. After that, the sentiment scores of units are calculated using unit context sentiment propagation. Experiments on two real-world microblog data sets demonstrate that our method can generate microblog-specific sentiment lexicons effectively. Furthermore, the sentiment classification accuracies significantly outperform state-of-the-art baselines.

Keywords: Implicit sentiment features; Social relationships; Context propagation; Microblog-specific sentiment lexicons; Sentiment classification

1. Introduction

Within the background of social media, user generated data is growing explosively on platforms, such as tweets, Sina Weibo, BBS, and Wechat. Users are likely to express their opinions and reviews about products, evaluations, or services through these social utilities (Chao and Yang, 2018; Yang et al., 2017; Siddique and Fung, 2017). The interest that individual users demonstrate via their online opinions about products and services is something to which vendors of these items are paying attention (Bertero et al., 2016). In short, sentiment analysis (SA) is the process of detecting the contextual polarities of microblogs, reviews, or other perspective texts (Bertero and Fung, 2017; Zhao et al., 2016). To our knowledge, explicit sentiment features such as opinion words, sentiment phrases, and idioms

☆ This paper has been recommended for acceptance by Pascale Fung.

* Corresponding author at: School of Computer and Information Technology, Shanxi University, Taiyuan 030006, Shanxi, China
E-mail address: wsg@sxu.edu.cn (S. Wang).

https://doi.org/10.1016/j.csl.2018.10.004

0885-2308/ 2018 Elsevier Ltd. All rights reserved.

are strong indictors of sentiment polarities. Sentiment lexicons via the labeling of words or phrases with their sentiment polarities are important for sentiment analysis (Kim and Provost, 2016; Zhang et al., 2016b). Lexicon-based approaches involve calculating orientations from the semantic polarities of words or phrases in documents. More recently, some researchers have proposed many methods such as linguistic-rules-based approaches, corpora-based approaches, and dictionary-based approaches to construct sentiment lexicons automatically (Zhang et al., 2018; 2016a). There are many domains existing on the internet. Therefore, high-quality domain-specific sentiment lexicons can improve the effectiveness of fine-grained sentiment analysis (Zhao et al., 2014).

It is worth mentioning that there are three main challenges in the construction of microblog-specific sentiment lexicons: informal language expressions, diverse internet vocabularies, and sentiment domain-dependency. First, the language expressions are very casual and informal. There are a large number of repeated words, modifications, transliterations, and pauses in microblogs (Hu and Li, 2011). For example, "图样图森破" (too young too simple) is a transliterated phrase and expresses a negative sentiment. Second, there are diverse internet vocabularies in microblogs. These new sentiment features also express sentiment polarities (Wang and Liu, 2015). For example, "洪荒之力" (primordial force), "吃瓜群众" (onlooker), "辣眼睛" (unwatchable), and "水军" (spammer) are frequently used online vocabularies by microblog users. Third, the sentiment expression is domain-dependent. For example, "轻薄" (thin and light) expresses positive sentiment in the electronic domain, but it expresses negative sentiment in the kitchen domain. It is hard or even impossible to collect and maintain general sentiment lexicons for all application domains. It is necessary to construct a domain-specific sentiment lexicon for a special domain. Therefore, many more advanced methods add domain knowledge to the construction of microblog-specific sentiment lexicons.

Previous studies focus on explicit sentiment features such as "美好" (beautiful), "厌恶" (disgust), and "喜欢" (like). An obvious characteristic of these features is having significant sentiment indicators (Zhang et al., 2016b). These sentiment words, phrases, and idioms are therefore named explicit sentiment features, which express a direct opinion on an entity or aspect. In fact, many users employ linguistic rhetoric or factual statements to express implicit sentiments indirectly. Implicit sentiment features generally refer to features that express positive or negative sentiments while not having significant sentiment indicator words. These noun features usually demonstrate a fact and express a sentiment indirectly. An implicit opinion is an objective statement that implies a regular or comparative opinion (Duwairi et al., 2015). Four microblogs including four implicit sentiment features "油老虎" (gas guzzler), "洪荒之力" (primordial force), "水军" (spammer), and "五毛特效" (cheap special effects) are shown in Fig. 1. Without any obvious sentiment indicator, the detection of implicit sentiment features has also been a challenging issue. Some researchers leveraged world or commonsense knowledge to detect implicit sentiment features (Balahur et al., 2012). The commonsense knowledge required manual collection and construction, and were labor-sensitive and time-consuming. Zhang and Liu (2011) detected nouns and noun phrases that indicated product features may also imply opinions. They adopted the surrounding local sentiment context and designed candidate identification to identify implied opinion features. In the social media environment, a large amount of social information such as user information can provide rich social contexts for detecting explicit and implicit sentiment features.

In this paper, we propose a novel unit context sentiment propagation-based approach to generate microblog-specific sentiment lexicons (L). We use online microblogs as the source data, and adopt social and local contexts to propagate sentiment using a probability transition matrix (P). First, we extract sentiment units that may express sentiment polarities from microblog data sets. These units not only contain candidate explicit sentiment features but also include implicit noun or noun phrase sentiment features. After that, we construct a graph of general sentiment units and microblog-specific sentiment units using their social relationships, topic features, and local contexts. The seed sentiment units (S) are selected from all sentiment units using a selection provided by a seed units algorithm. These seed units are labeled with sentiment labels (positive, neutral, or negative) using general sentiment lexicons (G) and manual calibrations. Moreover, the sentiment scores of all sentiment units are calculated using unit context sentiment propagation. Finally, we obtain the microblog-specific sentiment lexicons (L) and apply them into two real-world microblog data set sentiment classification tasks. Compared to previous methods, experimental results demonstrate that our method can generate microblog-specific sentiment lexicons and improve sentiment classification accuracies effectively.

Fig. 1. Four microblogs including four implicit sentiment features “油老虎” (gas guzzler), “洪荒之力” (primordial force), “水军” (spammer), and “五毛特效” (cheap special effects) posted by “东方网” (Eastday.com), “羊城晚报” (Yang Cheng Evening News), “新浪网” (Sina.com), and “腾讯网” (Tencent.com). “油老虎” (gas guzzler) expresses a negative sentiment about a gas firing guzzler, ‘洪荒之力” (primordial force) conveys a positive sentiment about great strength, “水军” (spammer) are hired web writers who publish specific information about particular content in a network, “五毛特效” (cheap special effects) describes poor movie special effects.

Our contributions in this paper can be summarized as follows.

- We propose a novel unit context sentiment propagation framework to generate microblog-specific sentiment lexicons for Chinese microblog data sets. The generated lexicons not only contain explicit sentiment features, but also include implicit sentiment features.
- We construct a sentiment propagation graph and its adjacency matrix using social relationships, topic features, and local contexts. The sentiment propagation algorithm (SPA) propagates sentiments from labeled sentiment units to unlabeled ones. Through sentiment label propagation, we can obtain the sentiment scores of explicit and implicit sentiment features.
- We verify the effectiveness of our framework on two real-world microblog data sets. Experimental results demonstrate that our framework can obtain high-quality microblog-specific sentiment lexicons and outperform state-of-the-art sentiment lexicon generating methods in terms of sentiment classification results.

The remainder of this paper is organized as follows. Section 2 introduces existing sentiment lexicon generating methods and implicit sentiment feature detecting strategies. We report the unit context sentiment propagation framework in Section 3. The selection of the seed sentiment unit process is presented in Section 4. The process of unit context sentiment propagation is shown in Section 5. Section 6 describes the experimental setup and evaluation performance of our proposed approach. Section 7 summarizes the entire study and provides directions for the next study.

2. Related work

In this section, we briefly review three kinds of sentiment lexicon generating techniques including linguistic rules-based, corpora-based, and dictionary-based methods. In addition, previous implicit sentiment feature detection approaches are introduced and summarized.

2.1. Sentiment lexicon generating techniques

Sentiment lexicons are widely used in sentiment analysis, especially for word-level and phrase-level sentiment analysis tasks. More recently, researchers have contributed many effective methods to generate domain-specific sentiment lexicons. In general, these methods can be broadly classified into three categories, linguistic rules-based approaches, corpora-based statistical learning approaches, and dictionary-based approaches (Fu et al., 2017; Huang et al., 2014; Zhang and Singh, 2014).

2.1.1. Linguistic rules-based approaches

Linguistic rules-based approaches usually use contextual information and linguistic rules to judge the sentiment orientation of sentiment words and generate specific sentiment lexicons. Linguistic rules refer to dependency relationships, specific language phenomena, conjunction words, and so on. Previous methods make full use of known linguistic rules, while the sentiment expression of free texts is arbitrary. Sometimes, people do not follow grammatical rules to organize microblogs, and this adds to the difficulties in generating domain-specific sentiment lexicons (Cho et al., 2014).

Jijkoun et al. (2010) proposed a bootstrapping method to generate a topic-specific lexicon and formed a general-purpose polarity lexicon. The generation of topic-specific lexicons could be divided into three steps: extracting syntactic contexts, selecting potential targets, and generating topic-specific lexicons. To test the validity of the proposed method, they applied the generated lexicons to opinionated blog post retrieval tasks. A fully automatic and unsupervised algorithm was presented by Kanayama and Nasukawa (2006) to generate a domain-oriented sentiment lexicon for extracting polar clauses. They used the overall density and precision of coherency in the corpus, and acquired the appropriate polar atoms automatically. Similarly, a domain sentiment word extraction approach was proposed by Qiu et al. (2009) based on the propagation of both known sentiment lexicons and extracted product features. They leveraged the dependency relations between features and sentiment words, and used the direct and indirect relations between them to expand the sentiment lexicons. A new method was also proposed to assign polarities to newly discovered sentiment words in a domain. Experimental results demonstrated that the double propagation approach was able to extract a large number of new sentiment words.

2.1.2. Corpora-based approaches

Motivated by different real-world applications, researchers have considered a wide range of problems over a variety of different types of corpora, while the sentiment expression is highly dependent on the domain. Corpora-based approaches make full use of the mutual reinforcement between documents and words or the co-occurrence relationships of words in the corpora. It is worth noting that the effectiveness is highly dependent on the scale and quality of the corpora.

To reduce the sentiment tagging efforts of domain-specific corpus in the construction of domain-specific sentiment lexicons, (Park et al., 2015) extracted the subset of documents that were selected by an active learner initialized by a diverse text analysis. Specifically, active learning techniques were adopted to find the most significant document that determined the sentiment classification boundary. In addition, diverse machine learning algorithms were used to find the best initialization documents. An adapted information bottleneck method was proposed by Du et al. (2010) for the construction of domain-oriented sentiment lexicons. Unlike the existing work that only considered the type of relationships between words, they used iterative reinforcement and an information bottleneck method to construct a domain-oriented sentiment lexicon. Turney and Littman (2003) introduced a method for inferring the semantic orientation of a word from its statistical association with a set of positive and negative paradigm words. The semantic orientation of a given word was calculated from the strength of its association with a set of positive words, minus the strength of its association with a set of negative words. They used two different statistical measures of word association: pointwise mutual information (PMI) and latent semantic analysis (LSA). To retrieve opinionated blog posts, He et al. (2008) proposed a novel and effective dictionary-based statistical approach, which automatically derived evidence for subjectivity from the blog collection itself. The proposed framework did not require any manual effort; Therefore, it was easy to carry out. Tan et al. (2008) proposed a novel scheme for sentiment classification (without labeled examples), which combined the strengths of both the learn-based and lexicon-based approaches. Specifically, they first leveraged an unsupervised technique to label some informative unlabeled examples in a new domain, then trained a new supervised classifier such as centroid classifier over these selected examples.

2.1.3. Dictionary-based approaches

Dictionary-based approaches usually make use of previous sentiment lexicon resources. These methods use semi-supervised technologies to generate domain-specific sentiment lexicons. Most of these approaches depend on the scale and quality of general sentiment lexicons (Liu et al., 2013).

Rao and Ravichandran (2009) treated polarity detection as a semi-supervised label propagation problem in a graph. They used WordNet and OpenOffice resources to achieve labeled examples for unlabeled words. In addition, synonymy and hypernymy were first used to improve label propagation results. The label propagation improved significantly over the baselines and other semi-supervised learning methods. Zhang and Singh (2014) proposed a semi-supervised framework called ReNew to generate a domain-specific sentiment lexicon and infer sentiments at the segment level. In their work, they used segments as the basic sentiment elements and used conditional random fields (CRF) to predict the sentiment label for each segment. To capture the contextual sentiment of words, ReNew adopted dependency relation pairs as basic elements in the generated sentiment lexicon. A graph propagation framework of lexical graphs was constructed by Velikovich et al. (2010). Specifically, they investigated the viability of polarity lexicons that were derived solely from unlabeled web documents. Thus, the method they investigated could be considered a combination of methods for propagating sentiments across lexical graphs and methods to build sentiment lexicons based on the distributional characteristics of phrases in raw data. Weichselbraun et al. (2011) proposed a semi-automatic approach for the creation of sentiment lexicons, which assigned sentiment values to sentiment terms via crowd-sourcing. Furthermore, they introduced a bootstrapping process operating on unlabeled domain documents to extend the created lexicons, and to customize them according to the particular case.

2.2. Implicit sentiment features detecting approaches

Compared to explicit sentiment features, detecting implicit sentiment features is a more challenging task. Within internet environments, individuals become accustomed to expressing their sentiment using nouns or noun phrases. Some researchers find that nouns and noun phrases that indicate product features may also imply opinions in some domains. Implicit sentiment features are objective statements that express desirable or undesirable facts. Sometimes, these features do not occur in the previous sentiment lexicons (Mukherjee and Bhattacharyya, 2012). Detecting these features can help us judge the sentiment polarities and understand author intentions more comprehensively. Identifying such nouns or noun phrases and their polarities is challenging but critical for effective opinion mining in these domains (Pang and Lee, 2008).

Previous studies regarding detecting implicit sentiment features mainly focus on world or commonsense knowledge. Balahur et al. (2011) presented an approach toward automatically detecting emotions from contexts in which no clues of sentiment appeared based on commonsense knowledge. They built a commonsense knowledge base (EmotiNet) representing situations to detect implicit expressions of sentiment in texts. In addition to that, Balahur et al. (2012) presented a comparative analysis between the performance of supervised and lexical knowledge-based well-established methods for emotion detection. They extended the commonsense knowledge stored in the EmotiNet knowledge base. Greene and Resnik (2009) introduced an approach to implicit sentiment motivated by theoretical work in lexical semantics, presenting evidence for the role of semantic properties in human sentiment judgments. They established a strong predictive connection between linguistically well-motivated features and implicit sentiment, and then demonstrated how computational approximations of these features could be used to improve on existing state-of-the-art sentiment classification results. In addition, Van de Kauter et al. (2015) investigated the viability of a new fine-grained sentiment annotation scheme. They mined the factual information in financial news articles at a fine-grained level. The fine-grained approaches was based on the detection of explicit as well as implicit sentiment expressions.

A new feature-based opinion mining model was proposed by Zhang and Liu (2011), they determined feature polarity not only by opinion words that modified these features, but also by the local contexts. They first identified the candidate noun features by determining the surrounding sentiment context of each noun feature. Then they pruned the listed noun features that were modified by only positive or negative opinion words. In this paper, our proposed sentiment unit context propagation framework belongs to corpora-based approaches. We judge the relationship between two sentiment units mainly based on the co-occurrence relationships of local and social contexts. The social context contains topic features, and user relationship information including the follower in a social network. Besides that, we take advantage of the existing sentiment dictionary resources and linguistic semantic rules (LSR) for Chinese microblog domains. In the mining of implicit sentiment features, we are more concerned with the calculation of sentimental score of nominal sentiment features in social media.

3. Basic framework

The construction of sentiment lexicons is a basic and important aspect for sentiment analysis tasks. The sentiment of a word or phrase is dependent on a specific domain. With the goal of achieving the domain dependency of sentiment expressions, the purpose of microblog-specific sentiment lexicons is to collect and detect opinion words, sentiment phrases, and idioms with sentiment polarities. Microblog-specific sentiment lexicons play an important role in overall sentiment polarity and fine-grained sentiment analysis tasks. The mining of implicit sentiment features (mainly referring to noun sentiment features) is an important and challenging task. The semantic distribution hypothesis declares that a word is characterized by the company it keeps. One step further, words or phrases under similar contexts have similar sentiment polarities. This paper aims to construct microblog-specific sentiment lexicons by extracting explicit and implicit sentiment features, and classifies the microblogs according to their sentiment polarities (positive, negative, and neutral).

In this paper, we build a framework of microblog-specific sentiment lexicons based on sentiment label propagation. As shown in Fig. 2, our framework can be divided into four steps: (a) data preprocessing, (b) selection of seed sentiment units, (c) unit context sentiment propagation, and (d) task verifications including sentiment lexicon generation and sentiment classification. To demonstrate our framework visually, we use three microblogs as examples in Fig. 3.

In the preprocessing stage, we first filter URLs, links, stop words, repeated letters, and other noise information in microblogs. For specific characters in microblogs, we extract useful characters such as hashtags, which are labeled by # and #, user relationships, which are labeled by @, emojis, and emoticons. Candidate explicit and implicit sentiment features are defined as words or phrases that may express sentiment polarities. Lexical parsing can analyze the parts of speech (POS) of words tagging information, topic features, and user relationships in social networks. We use the part-of-speech tagging (POS) tagger and dependency parsing (DP) tools in the language technology platform (LTP)[1]. After that, we extract adjectives, adverbs, verbs, and noun words and phrases as candidate sentiment features. At the same time, we filter the words that are below a certain frequency. Through analyzing the dependency relations between language components, dependency parsing reveals their syntactic structures. We extract the negative modification (mNeg) and degree modifier relationships (mDegr) from the dependency parsing results. Then, we use the Chinese degree adverbs list and negative indicators list in the Hownet sentiment lexicon[2] to generate the sentiment units.

In the selection of seed sentiment unit process, we judge the relationship between different sentiment units using local contexts and social contexts. Local contexts mainly refer to contextual information in the corresponding

[1] http://www.ltp-cloud.com/

[2] http://www.keenage.com/html/c_bulletin_2007.htm

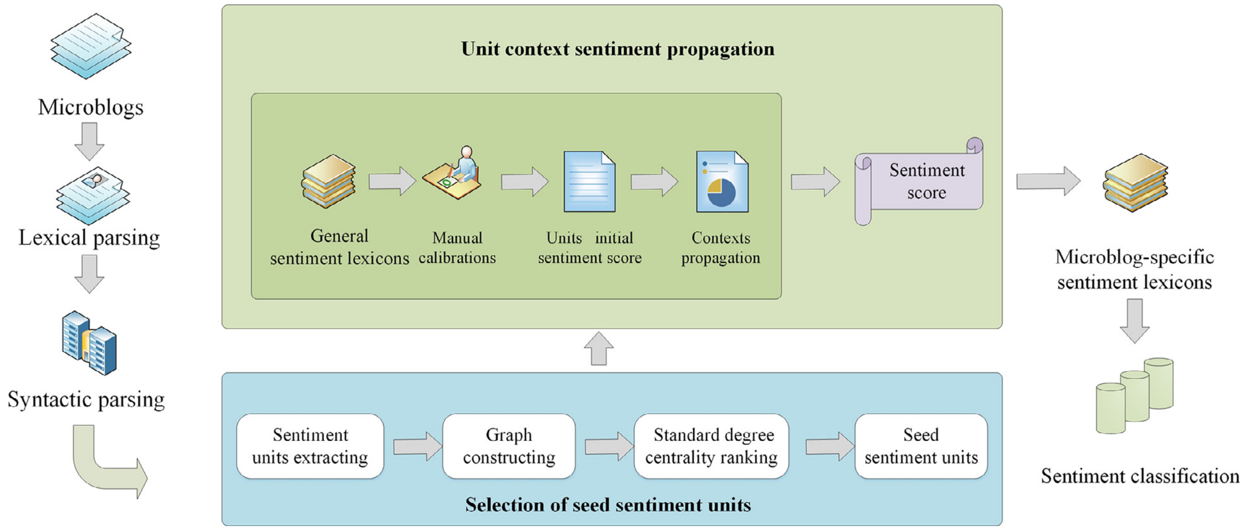


Fig. 2. Overview of the unit context sentiment propagation framework. At the preprocessing stage, lexical parsing and syntactic parsing are adopted to extract the lexical and sentence structure information. Then, we construct the sentiment units connection graph to select seed units according to the standard degree centrality (SDC). For unit context sentiment propagation, we accomplish the sentiment propagation processes from the labeled seed units set to the unlabeled units through an SPA. Finally, we obtain the microblog-specific sentiment lexicons and sentiment classification results.

microblogs. The local contexts between units can be determined by cohesion orders, organization structures, and semantic relationships between the clauses or sentences in microblogs. Given that the relationships between microblog users are directional, followee and follower relationships may exist between two users at the same time. The social relationships information includes the microblog information released by follower users, topic features that are included in microblogs, and microblog information released by microblog users, which is mentioned in the microblogs. Another key point is how to select seed sentiment units. The selected seed units should have a large

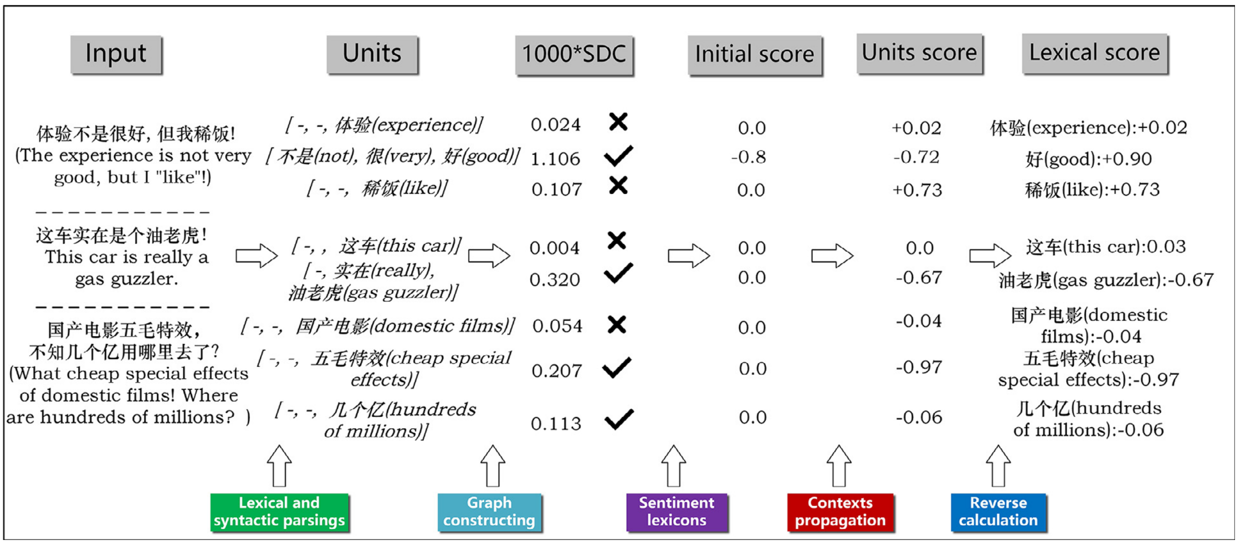


Fig. 3. The processing flow of three microblogs in the unit contexts sentiment propagation framework. The inputs of the framework are the sentences of the microblogs. After lexical and syntactic parsings, we can obtain the candidate sentiment units. Through SDC ranking, we select the seed units whose SDC values are greater than a certain threshold. According to general sentiment lexicons and manual calibrations, we obtain the initial score of the seed units set. Through sentiment label propagation, we obtain the sentiment score of all units. The lexical score is calculated via a reverse calculation.

standard degree centrality (SDC) value of and a large linkage degree with other nodes in the graph. Selecting these features and reinvesting them into the sentiment score can constitute the sentiment resources for propagation in the sentiment propagation graph.

Although implicit sentiment features alone indicate no sentiment, these features can express positive or negative sentiments in particular contexts. These features express strong sentiment polarities but are not included in previous traditional sentiment lexicons. We use the following hypothesis to detect the implicit sentiment features. Non-implicit features (noun or noun phrases without sentiment polarity) usually occur with positive and negative sentiment units, while implicit sentiment features only occur with positive or negative sentiment units alone. In this paper, we judge the sentiment intensities of units according to the semantic distribution hypothesis, which states that two units have similar sentiment polarities if they frequently appear in the same window. Existing sentiment lexicon resources include abundant general explicit opinion words, sentiment phrases, and idioms, which provide us with rich resources.

We leverage three assumptions when propagating sentiment using social and local contexts as described in the following paragraphs.

- Similar microblog users tend to express similar views, interests, acquaintances, and preferences about the same products, evaluations, or services. Users constitute a huge mesh structure graph in social networks. In this graph, an information stream can be transmitted along the edges, which are composed of user relationships. Previous studies have proven that microblog social networks and user communities are basic organizations that disseminate information.
- User networks and topic features are important for sentiment propagation in social networks. Users can focus on other users and have many fans, or even take the initiative in publishing the microblogs inspired by other users. Users usually insert the @ labels when mentioning other users and insert # labels under the same topic.
- Two sentiment units that appear in the same window usually have the same sentiment polarity and similar sentiment intensity. Besides that, linguistic semantic rules (LSR) can judge the semantic relationships directly between sentiment units. For example, conjunctions such as “但是” (but) and “和” (and) have the functions of linking discourse and representing the attitude and sentiment of the speakers.

After obtaining seed sentiment units, we label the seed units set with general sentiment lexicons and manual calibrations. In allusion to the previous statement that sentiment lexicons often do not contain internet vocabulary, we add the internet catchwords such as “山寨” (fake) and “桑心” (sad) with their sentiment polarities into the general sentiment lexicons. Manual calibrations are used to check and correct sentiment labels of seed sentiment units. In the stage of unit context sentiment propagation, an intuitive idea is to propagate sentiment in accordance with the semantic relationships between sentiment units. We use the labeled sentiment units to speculate the unlabeled sentiment units including explicit and implicit sentiment units. Sentiments are propagated through the edges (similarity) between nodes in the graph. We calculate the sentiment score of explicit and implicit sentiment features in the microblog-specific sentiment lexicons. Then, we apply the generated sentiment lexicons into microblog sentiment classification tasks.

4. Selection of seed sentiment units

In this section, we introduce the selection of seed sentiment unit stage. The key points are to construct the relationships between different sentiment units and determine how seed sentiment units are selected. Lexical parsing and semantic parsing are adopted to extract candidate explicit and implicit sentiment features. After that, the connection graph between sentiment units is constructed using local and social relationship contexts.

4.1. Basic symbols and notions

The microblogs data set is denoted by $D = \{d_1, d_2, \ldots, d_l\}$. $W = \{w_1, w_2, \ldots, w_n\}$ is the latent sentiment features set, where w_k is a sentiment word or phrase, $1 \le k \le n$. u is a negative indicator word, such as “不” (not) and “没有” (nothing). v is a degree adverb word, such as “完全地” (absolutely) and “相当地” (fairly).

Definition 1. *Sentiment unit*. A sentiment unit is defined as $T = (N, D, E, P)$, where N is a negative indicator, D is a degree adverb, E is an evaluation word, P is the sentiment polarity. The evaluation word refers to the subjective word or phrase that evaluates or expresses sentiment. It can be an adjective, a verb, or a noun.

Degree adverbs and negative indicators affect sentiment polarities and sentiment intensities. For example, in the sentiment unit (不(not), 很(very), 好(good), negative), "好" (good) is a positive word, while the sentiment unit (不(not), 很(very), 好(good), negative) is a negative unit because of the negative indicator "不" (not), the degree adverb "很" (very) enhances the degree of a negative sentiment.

Definition 2. *Sentiment score*. The sentiment score describes the sentiment intensity and polarity of a word, a phrase, or a unit in great detail. The sentiment score of a unit is denoted as $score(T_i)$, and $-1 \leq score(T_i) \leq 1$.

If the positive sentiment intensity of T_i is stronger, then $score(T_i)$ is closer to 1, and vice versa. If $score(T_i) = 0$, then T_i is a neutral unit. The sentiment intensity of T_i is denoted as $strength(T_i)$, and $0 \leq strength(T_i) \leq 1$.

Definition 3. *Local contexts*. Local contexts of sentiment units include other units in the corresponding microblogs, comments of corresponding microblogs, and other microblogs released by the same author.

Definition 4. *Social contexts*. The social relationship context of sentiment units contains forwards and replies information, topic features, and user relationship information, including the follower in a social network. Forwarding a microblog indicates a node to view in a microblog, and replies mainly refer to the comments of microblogs.

The local and social relationship contexts of a special unit T_i in the microblog d can be seen in Fig. 4. According to three assumptions as listed in the basic framework section, local and social contexts can determine the sentiment of units in which they appear. Two units under common topics and in the processing of forwarding and replying microblogs have similar sentiment polarities and intensities.

4.2. Co-occurrence relations and linguistic semantic rules

Definition 5. *Propagation window*. To determine the neighboring relationships between two units, the propagation window is defined as forward n-units and backward n-units of a specific unit T_i.

The co-occurrence of sentiment expressions is calculated under local contexts and social contexts. When calculating co-occurrence relations in local contexts, co-occurrence window is only extracted in the corresponding

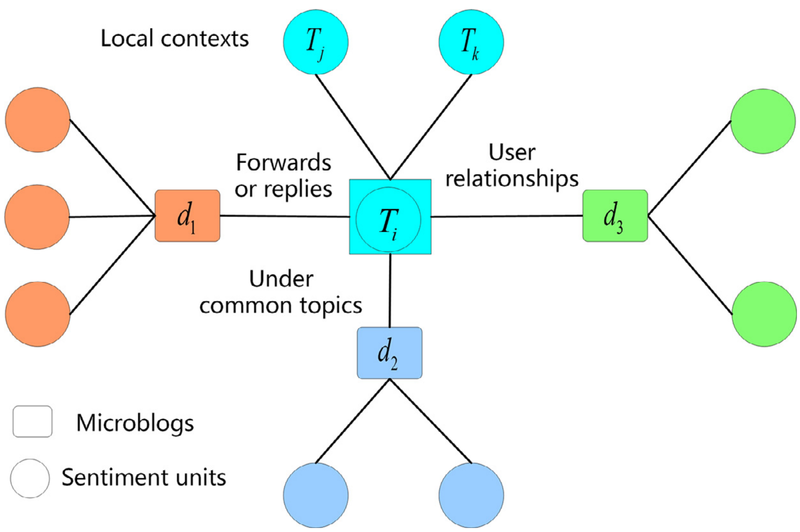


Fig. 4. Schematic diagram of local contexts and social contexts of the target unit T_i in the microblog d. T_j and T_k are the local contexts units of T_i. The social contexts of T_i include d_1, d_2, and d_3, where d_1 are the forwards or replies microblogs, d_2 are the microblogs under common topics with T_i, and d_3 are the microblogs that have user relationships with d.

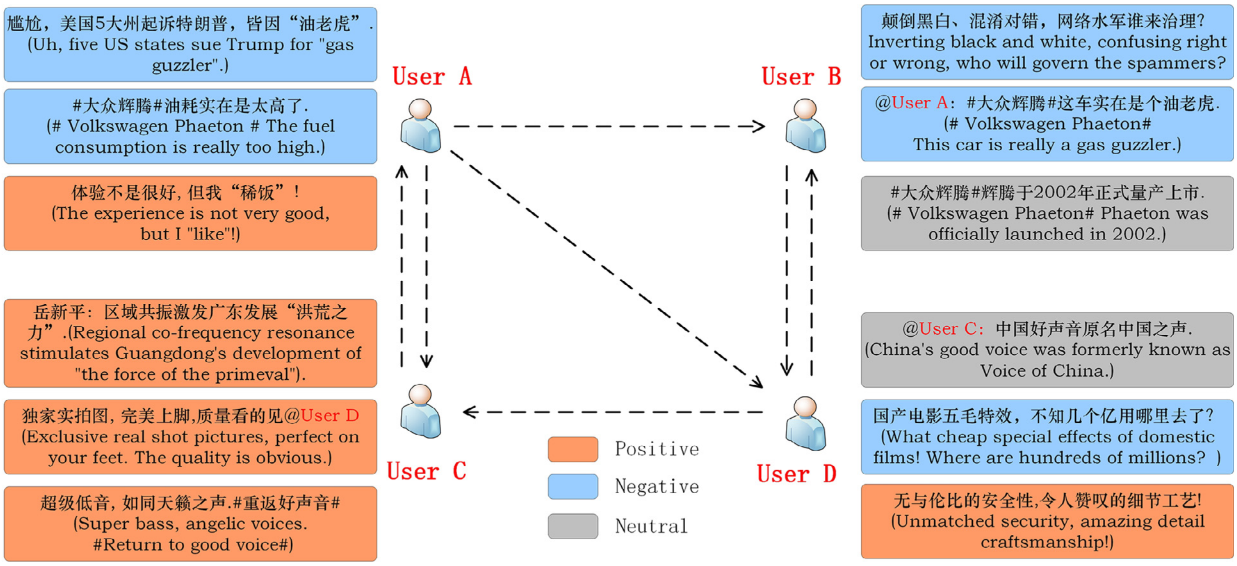


Fig. 5. User relationships and posting microblogs with sentiment polarities from Sina Weibo. The user relationships contain the follower relationships that are brought in the data set and the @ relationships mentioned in the microblogs. Different users form directional follower and followee relationships, for example, *User A* is a follower of *User B*, while *User B* does not pay close attention to *User A*. The label “@*User A*” indicates that the related context is about *User A*. The hashtag labeled by “#” and “#” indicates the topic of microblogs. In the microblog “#大众辉腾#这车实在是个油老虎,” (#Volkswagen Phaeton# This car is really a gas guzzler.), we can easily infer that it is about #*大众辉腾*(Volkswagen Phaeton)#. In addition, different users can discuss the same topic just like #*大众辉腾*(Volkswagen Phaeton)# and #*重返好声音*(Return to good voice)#.

microblogs, comments of corresponding microblogs, and other microblogs released by the same author. The social relationship context of sentiment units contains forwards and replies information, topic features, and user relationship information, including the follower in a social network. When calculating co-occurrence relations in local contexts, co-occurrence window is only extracted in the microblogs that posted by the followers, posted by the user mentioned by @ label, and under the same topics.

Co-occurrence relations mainly refer to word co-occurrences that occur in the same window. Co-occurrence relations and semantic rules can be used to infer sentiment polarities of unlabeled sentiment units. User relationships and posting microblogs with sentiment polarity from Sina Weibo are shown in Fig. 5. In Fig. 5, both “无与伦比” (unparalleled) and “令人赞叹” (amazing) express positive sentiment polarity in the microblog “无与伦比的安全性, 令人赞叹的细节工艺!” (Unparalleled security, amazing detail craftsmanship!) posted by *User D*. Apart from co-occurrence relations, transitional words can help us judge the semantic relationships between units directly. Two sentiment units that are connected by coordinate conjunctions have a larger probability of possessing the same sentiment polarity and similar sentiment intensity. For example, (-,超级(super),低音(bass)) and (-, -,天籁(angelic voices)), which are connected by “如同” (as) are both positive units in the microblog “超级低音,如同天籁之声#重返好声音#” (Super bass, angelic voices. #Return to good voice#) posted by *User C*. In addition, two sentiment units connected by adversative conjunctions have a larger probability of possessing opposite sentiment polarity. For example, two clauses connected by “但” (but) express different sentiment polarities in the microblog “体验不是很好,但我“稀饭”! (The experience is not very good, but I “like”!) posted by *User A*.

In addition to transitional words, we leverage LSR to infer the semantic relationships of units as shown in Table 1. We adjust and extend 13 compositional semantic rules proposed by Xie et al. (2013) into Chinese microblog domains, which include a full presentation of the semantic rules approach. Rules 1–11 show sentiment element composition rules with the parts of speech tagging and combinational rules. Rules 12–17 display six rules of conditional clause sentiment composition. In addition to these semantic rules, we also need to consider the subjunctive mood, interrogative sentences, irony, and sarcasm phenomena. Three types of sentences that require special treatment are listed in rules 18–20. The subjunctive mood is used to indicate what for the speaker is not a fact, but a hypothesis, wishes, doubt, or speculation. Although there may be sentiment indicator words in the question sentences, the sentences do not express any sentiment. Irony or sarcasm leads to the reversal of the sentiment polarity of microblogs.

Table 1
Twenty linguistic semantic rules (LSR) for Chinese microblog domains.

LSR	Description	Examples	Translations
Rule 1	Polarity(Negative indicators + fragment)= -Polarity (fragment)	尚未成功	Not yet successful
Rule 2	Polarity(degree adverbs + fragment)= Degree(adverbs) * Polarity(fragment)	极其讨厌	Extremely disgusting
Rule 3	Polarity(verb + noun) = Compose(verb, noun)	丢掉坏习惯	Throw away bad habits
Rule 4	Polarity(verb1 + verb2) = Compose(verb1, verb2)	拒绝投降	Refuse to surrender
Rule 5	Polarity(adjective + verb) = Compose(adjective, verb)	不可能消失	Never go away
Rule 6	Polarity(noun1 + noun2) = Compose(noun1, noun2)	缺乏信任	Lack of trusts
Rule 7	Polarity(noun + verb) = Compose(noun, verb)	犯罪率下降	Crime has decreased
Rule 8	Polarity(noun + adjective) = Compose(noun, adjective)	将 伤害最小化	Minimizing damage
Rule 9	Polarity(noun + verb) = Compose(noun, verb)	Lack of killing in rural areas	
Rule 10	Polarity(as adjective as noun)=Polarity(ADJ)	像 石头一样 丑陋	As ugly as a rock
Rule 11	Polarity(not as adjective as noun) = -Polarity(ADJ)	没有?来的 好	That was not as good as the original
Rule 12	If the sentence contains “coordination”, disregard all previous sentiment and only take the sentiment of the part after “but”.	屏幕不太好, 但是我还是喜欢这部手机	The screen is not good, but I like this phone.
Rule 13	If the sentence contains “adversative”, only take the sentiment of the part before “despite”.	我喜欢这部电影, 除了讨厌这部电影的导演.	I love this movie, despite the fact that I hate that director.
Rule 14	If the sentence contains “causality” is followed by a negative clause, disregard the “unless” clause.	你会迟到的,除非你现在抓紧时间.	You will be late unless you hurry up.
Rule 15	If the sentence contains “assumption”, neglect the sentiment of assumption clause.	假如价格在低点,就完美了.	If the price is lower, it will be perfect.
Rule 16	If the sentence contains “selection”, take the sentiment of the original clause or selection clause.	这些产品不是太低端, 就是价格 贵.	These products are either low-end or expensive.
Rule 17	If the sentence contains “progressive”, take the sentiment of both original clause and progressive clause.	这件衣服不仅漂亮,而且价格合适.	This dress not only is beautiful, but also its price is a good fit.
Rule 18	Subjunctive mood	我希望他们没有迟到.	I **wish** they were **not so late**.
Rule 19	Interrogatory	快递速度很快吗?	Is the speed of express **fast**?
Rule 20	Irony or sarcasm	这真的是一场 “极好”的电影啊!	It is really a “**great**” movie!

4.3. Sentiment propagation graph

Definition 6. *Sentiment propagation graph*. We construct the sentiment propagation graph as $G = (V, E, W)$, where each node is a sentiment unit, V is the set of sentiment units, E is the edges set, and W is the weight matrix between sentiment units.

If the probability of two sentiment units that appear in the same window is high, we think that the mutual information between two units is big, and vice versa. After obtaining the relationships between candidate sentiment units, the edge between T_i and T_j is defined as $p_{i,j} = PMI(T_i, T_j)$ (Pointwise Mutual Information).

$$PMI(T_i, T_j) = \log_2 \frac{n \times hits(T_i, T_j)}{hits(T_i) \times hits(T_j)}, \tag{1}$$

where n is the total number of sentiment units in the corpus, $hits(T_i)$ is the number of units T_i in the corpus, $hits(T_j)$ is the number of unit T_j in the corpus, $hits(T_i, T_j)$ is the number of T_i and T_j that occur in the same window within local and social contexts. The relationship matrix between sentiment units is denoted as $P_{n \times n}$.

4.4. Standard degree centrality (SDC) ranking

To measure the linkage degree of nodes (sentiment units) in the sentiment propagation graph G, we have the following three definitions.

Definition 7. *Link probability*. Let n_i be the number of nodes that have connections with T_i, the link probability $p'_{i,j}$ of T_i connected to T_j is defined as follows,

$$p'_{i,j} = p_{i,j} / (\sum_{i=1}^{n_i} p_{i,j}). \tag{2}$$

A large value of $p'_{i,j}$ indicates that T_i is more likely to be connected to T_j, and T_i has a large probability of being connected to T_j. The probability transition matrix is denoted as $P'_{n\times n}$.

Definition 8. *Standard degree centrality*. The standard degree centrality (SDC) of T_i is defined as follows,

$$H(T_i) = \frac{1}{n-1}\sum_{j=1}^{n_i} p_{ij}, \tag{3}$$

where n is the total number units, and n_i is the degree of the unit T_i in the graph G. A greater standard degree centrality of a node in the graph signifies a higher importance of the node in the network. We use the SDC to measure the connection degree of a node in the graph G. If $H(T_i)$ is big, the linkage degree of T_i for the graph G is high, and vice versa.

Definition 9. *Seed sentiment units*. The seed sentiment units are defined as units with high SDC values.

The high value of $H(T_i)$ indicates that T_i makes great contributions to sentiment label propagation. Seed sentiment units can offer sentiment sources in the propagation graph G. Then, we arrange all the units according to the values of $H(T_i)$. T^s is the seed sentiment units set, and the scale of T^s is M, that is $|T^s| = M$. On that basis, we select the $top-M$ units as the seed sentiment units subset T^s.

4.5. Algorithm for selection of seed sentiment units

The complete algorithm for the selection of seed sentiment units is described in Algorithm 1.

Algorithm 1. Algorithm for selection of seed sentiment units.

Input: Microblogs set $D = \{d_1, d_2, \cdots, d_l\}$, selected number of seed units M, Chinese transitional words, degree adverbs, and negative indicators words lists.
Output: The seed units set T^s.
Filter noise and unrelated information in the microblogs set D;
Use parts-of-speech tagging (POS) and dependency parsing (DP) tools to extract candidate sentiment features;
Extract candidate explicit and implicit sentiment units and set $T = \{T_1, T_2, \cdots, T_n\}$;
Extract local contexts, topic features, and social relationships of sentiment units;
Construct the relationship matrix $P_{n\times n}$ between sentiment units, and $P_{i,j} = PMI(T_i, T_j)$ is calculated as Eq. 1;
Construct the probability transition matrix $P'_{n\times n}$, and $P'_{i,j}$ is calculated as Eq. 2;
Calculate $H(T_i)$ for each T_i using Eq. 3;
Rank all the units according to the value of $H(T_i)$;
Select the $top-M$ sentiment units as the sentiment seed units;
Return the seed units set T^s.

5. Unit context sentiment propagation and applications

In this section, we first describe the sentiment propagation process and microblog-specific sentiment lexicon. Then we discuss the convergence of the SPA. At the sentiment classification stage, the sentiment polarities of microblogs are judged in accordance with the final sentiment scores. The flow chart of the unit context sentiment propagation and verification processes can be observed in Fig. 6.

5.1. Sentiment propagation process

We use three the hypotheses, which are defined in Section 3, to propagate a sentiment in the sentiment propagation graph G. Then, the SPA accomplishes the sentiment propagation process from the labeled seed units set T^s to the unlabeled units set.

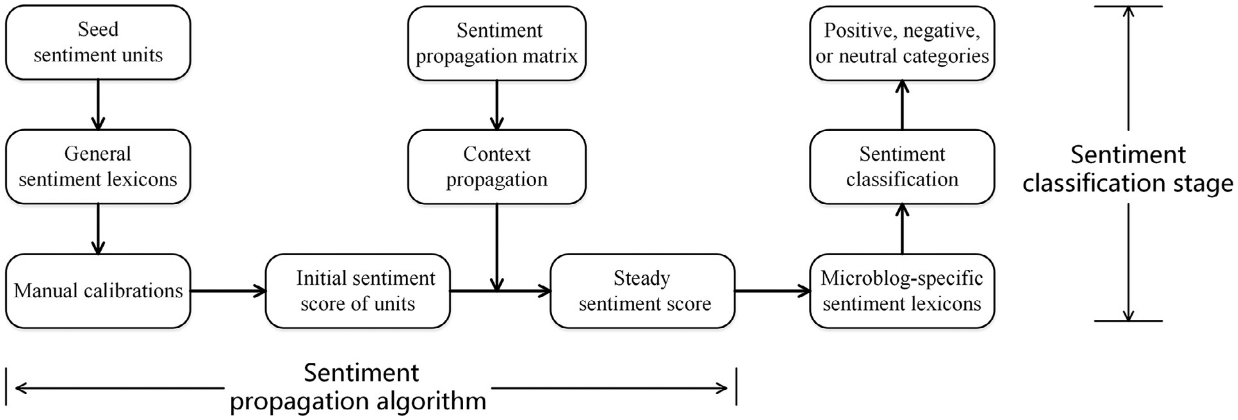


Fig. 6. Flow chart of unit context sentiment propagation framework. The input of the framework is the seed units set. After general sentiment lexicons and manual calibrations, we obtain the initial sentiment score of the seed units set. We propagate the sentiment from labeled seed units to unlabeled ones through context propagation. Then, we obtain the steady sentiment score of all units and construct the microblog-specific sentiment lexicons. Finally, we obtain the sentiment classification results of microblogs.

After using standard degree centrality (SDC) to choose the candidate sentiment units, we use three general sentiment dictionaries (Hownet, NTUSD, and SXU Lexicon) to label the sentiment polarities (positive, negative, and neutral). In the process of automatic labeling, there may be inaccuracies in the polarity of some words, or some words do not appear in existing sentiment lexicons. For different data sets, separate manual annotations are required. This step can provide the correct source of sentiment transmission for the next sentiment propagation.

The sentiment unit set is denoted by $T = \{T_1, T_2, \ldots, T_n\}$, where T_i is a sentiment unit, $1 \le i \le n$. The initial sentiment score vector of sentiment units is denoted as $score(T)$,

$$score(T) = [score(T_1), score(T_2), \ldots, score(T_n)]. \tag{4}$$

Normalize the $score(T)$ using Eqs. (5) and (6),

$$\sum_{T_i \in T_{pos}} score(T_i) = 1, \tag{5}$$

$$\sum_{T_i \in T_{neg}} score(T_i) = -1, \tag{6}$$

where T_{pos} is the set of all positive units, and T_{neg} is the set of all negative units.

Definition 10. *Graph pruning*. To reduce unnecessary operations and improve algorithm efficiency, we prune the graph G to remove the small connectivity edges. Specifically, we maintain k major correlation units for each unit as its sentiment neighbors. That is to keep k major values in each row of the matrix P', while the other value is assigned to 0.

Definition 11. *Probability transition matrix*. To maintain the convergence of SPA, we define the sentiment propagation probability transition matrix as follows:

$$\widehat{P} = \beta(P' + A) + (1-\beta)J, \tag{7}$$

where $\beta \in [0, 1]$ is the adjustable parameter. A is a matrix where part of the columns are $1/n$, and J is a matrix of all elements are $1/n$. Adding A is to ensure that the matrix P' has zeroes in all columns.

Definition 12. *Sentiment label propagation*. The sentiment label propagation process is shown as follows:

$$score(T_i^{t+1}) = \alpha \times score(T^{t+1}) \cdot \widehat{P}_i + (1-\alpha) \times score(T_i^t), \tag{8}$$

where $score(T_i^{t+1})$ is the sentiment score of T_i in the $t+1$ iteration, $\alpha \in (0, 1]$ is the weight parameter, $\widehat{P}_i$ is the i column of $\widehat{P}$, and $score(T^t)$ is the score vector of T in the t iteration. In each iteration, we calculate $score(T_i^{t+1})$ following the order $i = 1 : n$, and update $score(T_i^{t+1})$ when obtaining a new $score(T_{i-1}^{t+1})$.

When the iteration stops, we normalize $score(T)$ according to Eq. (9),

$$score(T_i) = \begin{cases} score(T_i) / \max\left(score(T_i)\right) & score(T_i) \geq 0 \\ -score(T_i) / \min\left(score(T_i)\right) & score(T_i) < 0 \end{cases} \tag{9}$$

5.2. Microblog-specific sentiment lexicons

We obtain the sentiment score of sentiment features in accordance with the sentiment score of units as follows,

$$score(w_i) = \left(\sum_{i=1}^{n(w_i)} \frac{score(T_i)}{N \times D} \right) / n(w_i), \tag{10}$$

where $n(w_i)$ is the number of w_i occurrences in the corpus, N is the negative indicator word in the sentiment unit T_i, and D is the degree adverb word in T_i. Finally, we obtain the microblog-specific lexicon L^{spe} including explicit and implicit sentiment features. In the experiment, we found that we can get good sentiment lexicon and sentiment classification results when the sentiment intensity of score is set to be greater than 0.15. So the word w_i is considered as an "evaluation word" and can be added to the sentiment dictionary.

5.3. Sentiment propagation algorithm

After the above descriptions including the sentiment propagation graph and sentiment label propagation, we obtain the microblog-specific sentiment lexicons. Our approach is capable of propagating sentiment from labeled sentiment units to unlabeled ones. The complete algorithm is described in Algorithm 2.

Algorithm 2. Sentiment propagation algorithm (SPA).

Input: Candidate sentiment units set $T = \{T_1, T_2, \cdots, T_n\}$, seed units set T^s, general sentiment lexicons L^{gen}.
Output: Microblog-specific sentiment lexicon L^{spe}.
Initialize initial sentiment score of unit vector $score(T)$;
Initialize parameters β, α;
Mark seed feature set T^s using general sentiment lexicons L^{gen} and manual calibrations;
Normalize $score(T)$ using Eq. 5 and Eq. 6;
Keep k major values of each row in $P'_{n\times n}$, or else assign to 0;
Normalize each row in P', and construct the sentiment propagation matrix $\hat{P}$ using Eq. 7;
repeat
Calculate $score(T_i^{t+1})$ $(1 \leq i \leq n)$ using Eq. 8;
Update and normalize $score(T)$;
until convergence /* the convergence condition is that two consecutive change values of each $score(T_i)$ are less than 0.0001; */
Normalize each $score(T_i)$ using Eq. 9;
for $1 \leq i \leq n$ **do**
    Obtain $score(w_i)$ using Eq. 10;
    if $|score(w_i)| \geq 0.150$ **then**
        Add w_i into L^{spe};
    end
end
Return the microblog-specific sentiment lexicon L^{spe}.

5.4. Algorithm convergence

In the SPA, the changes of $score(T)$ form a Markov chain. $score(T^{t+1})$ is only associated with $score(T^t)$ and unrelated to the previous sentiment vector $score(T^{t-1}), \ldots, score(T^0)$. When the algorithm converges, $score(T^{t+1}) = score(T^t)$. Substituting it into Eq. (8), we can obtain $score(T_i^t) = \alpha \times score(T^{t+1}) \cdot \widehat{P}_i + (1-\alpha) score(T_i^t)$, that is $\widehat{P}_i \cdot score(T^t) = score(T_i^t)$. This proves that the convergence of $score(T)$ is directly related with $\widehat{P}$, and the final steady sentiment score $score(T^t)$ is related to the initial sentiment score vector $score(T)$, transition matrix $\widehat{P}$, and parameters α, β.

To ensure the convergence of SPA, we construct a sentiment transition matrix $\widehat{P}$ as mentioned in Eq. (7). Adding matrix A ensures all nodes in G have neighbors. Adding matrix J captures all nodes that connect randomly. $\beta \in [0, 1]$ is the weight adjustment parameter. When β is closer to 1, $score(T^{t+1})$ is more likely to be influenced by its sentiment neighbors. When β is closer to 0, $score(T^{t+1})$ is more likely to be spread randomly. According to the above settings, there must be a path constituted by other units to make T_i connected to T_j for any given T_i and T_j in the sentiment propagation graph G. This indicates that graph G is strongly connected. According to Brin and Page (1998) and Austin (2006), $score(T)$ must be able to converge to a stable value; Therefore, our proposed SPA is convergent.

5.5. Sentiment classification applications

After obtaining the generated microblog-specific sentiment lexicons L^{spe}, we apply them into sentiment classification tasks. $score(d_i)$ is calculated by summing all the sentiment units in the microblog d_i and considering the explicit and implicit sentiment features, and the twenty LSR, which are listed in Table 1 for Chinese microblog domains. Then, we judge the sentiment polarities of the microblogs according to their sentiment scores. Finally, we compare the prediction label with the true sentiment label of the microblogs.

6. Experiments and evaluations

In this section, we first introduce the experimental setup, optimized model parameters, and parameter sensitivity. Next, we demonstrate the generated microblog-specific sentiment lexicons and implicit sentiment features mining results. Then, we present the experimental results of the sentiment units context propagation framework and compare the results with the baselines. Finally, we discuss the experimental results, explicit and implicit sentiment features detection, and three limitations that require improvement.

6.1. Experimental setup

6.1.1. Microblog data sets

UCI data set [3]: this data set was collected from UC Irvine Machine Learning Repository. It contained 781 users, 311636 follower and followee relationships. It contained 40445 microblogs including 5333 negative samples, 18542 neutral samples, and 16570 positive samples.

Weibo data set [4]: this data set contained 335950 microblogs, including 183930 negative samples and 152020 positive samples. Each microblog contained user IDs, microblog texts, sentiment labels, and retweeted microblog user IDs.

We adopted a 5-fold cross-validation method to divide the sample set into five parts, of which four parts were the training data set and the other part was the verification data set. Then we repeated this process five times and used the average as the final result.

6.1.2. Sentiment resources

General sentiment lexicons:

(1) *Hownet*. The Hownet sentiment lexicon contains 3116 negative evaluative terms and 1254 negative sentiment words, 3730 positive evaluative terms and 836 positive sentiment words.

[3] http://archive.ics.uci.edu/ml/datasets/microblogpcu
[4] http://goo.gl/iXzoXm

(2) *NTUSD*. NTUSD sentiment lexicon[5] contains 2810 positive terms and 8276 negative terms, which are collected by the National Taiwan University.

(3) *SXU lexicon*. Shanxi university sentiment lexicon contains 3284 strong positive words, 13012 positive words, 10829 negative words, and 2863 strong negative words. The network catchwords lexicon contains 250 positive words, 320 neutral words, and 576 negative words.

Transitional words, degree adverbs, and negative indicators words: we use the transitional words, degree adverbs, and negative indicator words lists in the Hownet lexicon. We use the seven Chinese transitional types including coordination words, adversative words, causality words, assumption words, selection words, comparison words, and progressive words, which are listed in Table 1. We use Chinese seven transitional types including coordination words such as "和" (and) and "同" (together); adversative words such as "却" (while) and "然而" (however), causality words such as "因为" (because) and "因此" (therefore), assumption words such as "若" (if) and "倘若" (in case), selection words such as "或" (or) and "还是" (still), comparison words such as "像" (like) and "如同" (as), progressive words such as "不但" (not only) and "而且" (as well as). We use 219 degree adverbs lists in the Hownet sentiment lexicon. The degree adverbs can be divided into five degrees (0.2, 0.4, 0.6, 0.8, and 1.0) including most degree (1.0) such as "最" (most) and "极其" (extremely), very degree (0.8) such as "很" (very) and "相当" (fairly), more degree (0.6) such as "更" (even more) and "更加" (still more), slightly degree (0.4) such as "略微" (appreciably) and "一点" (a little), insufficient degree (0.2) such as "不怎么" (not very) and "轻度" (slight). The negative indicators words can be divided into absolute degree words (-1.0) such as "不" (no) and "没有" (none), extreme degree words (-0.5) such as "尚未" (not yet) and "不必" (not necessary).

6.1.3. Comparing methods

we design the comparison experiments as follows to verify the effectiveness of our proposed framework.

Lexicon: we calculate the sentiment score of microblogs using sentiment features and general sentiment lexicons directly. In this process, the lexicon method considers the impacts of degree adverbs and negative indicators words. Then, we judge the sentiment polarities according to the sentiment scores of the microblogs.

ReNew: ReNew is a semi-supervised framework for generating a domain-specific sentiment lexicon and inferring sentiments at the segment level, which is proposed by Zhang and Singh (2014). We follow the original parameters setting of confidence thresholds of 0.9 for positive labels, 0.7 for negative and neutral labels, and 4 for the minimum frequency.

SO-LSA: We have compared a corpus-based approach SO-LSA, semantic orientation from latent semantic analysis (Turney and Littman, 2003). SO-LSA first uses singular value decomposition (SVD) to analyze the statistical relationships among words in a corpus. SO-LSA applies LSA to calculate the strength of the semantic association between words. We follow the original experimental parameter settings and use SXU Lexicon as the list of benchmark words.

Local: Draw on the method which was proposed by Zhang and Liu (2011) to identify noun product features that imply opinions, the local contexts method propagates sentiment using only local contexts including the corresponding microblogs, comments of microblogs, and other microblogs released by the same author while considering explicit and implicit features.

Social: The social contexts method propagates sentiment only using social information, including forwards and replies microblogs from friends regarding common topics, and user information by considering explicit and implicit features.

Explicit: The explicit features method propagates sentiment only considering the explicit sentiment features, which contain adjectives, adverbs, and verbs using local and social relationship contexts.

The selected parameters of *Local*, *Social*, and *Explicit* methods can be seen in Table 2.

SUCPF: Our proposed sentiment units context propagation framework considering explicit and implicit sentiment features using local and social contexts. The selected parameters can be seen in Table 2.

6.1.4. Evaluation metrics

We evaluate our proposed method by classifying the microblogs into three categories (positive negative, and neutral) in the UCI data set, two categories (positive and negative) in the Weibo data set. This paper uses precision,

[5] http://goo.gl/forms/1jBAxw7aFW

Table 2
Parameter settings on UCI and Weibo data sets with Hownet, NTUSD, and SXU lexicons in our proposed framework.

Parameters	UCI			Weibo		
	Hownet	NTUSD	SXU	Hownet	NTUSD	SXU
H	7	6	7	5	6	6
M	20%	20%	20%	20%	20%	20%
k	20	30	35	25	30	25
T	160	200	180	180	200	200
α	0.20	0.20	0.20	0.20	0.20	0.20
β	0.85	0.85	0.85	0.85	0.85	0.85

recall, and the F1-score as the basic evaluation metrics. For an overall evaluation, we use accuracy as the evaluation metric (Wang et al., 2011).

6.2. Model parameters

The parameter sets for the sentiment units context propagation framework are denoted as $\Phi = \{H, M, k, T, \alpha, \beta\}$. H is the window size in the selection of the seed units algorithm, and it determines the surrounding context of a specific unit. M is the number of seed sentiment units. k is number of sentiment neighbors in the sentiment propagation graph, and it affects the number of iterations and the stability of the calculation. T indicates the iteration times. In Section 5.4, we discuss the convergence analysis of the SPA, and the convergence of $score(T)$ is directly related to the sentiment transition matrix $\widehat{P}$. α is the updating rate, and β is the weight parameter. Generally speaking, when α is closer to 1, $score(w^{t+1})$ is more easily affected by its sentiment neighbors, and SPA is easier to converge. When α is closer to 0, $score(w^{t+1})$ is more vulnerable to the impact of $score(w^t)$, and SPA is more difficult to converge. According to Brin and Page (1998) and Austin (2006), we set $\alpha = 0.20$ and $\beta = 0.85$. In this paper, we have tested combinations of each of the parameter values to find the best parameters for our proposed framework. The best model parameters of our framework are shown in Table 2.

6.3. Parameter sensitivity

6.3.1. Window size influence analysis

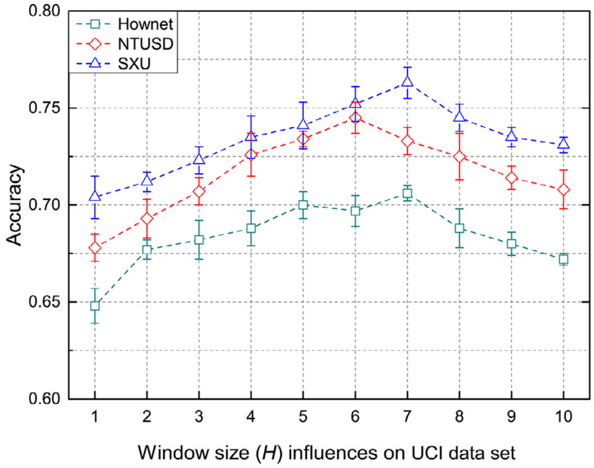
The first experiment is to test the influence of different window sizes H which change from 1 to 10 as shown in Fig. 7(a) and (b). As we can see from Fig. 7(a) and (b), the overall accuracies firstly increase and then decrease with the increment of H. When the overall accuracies reach their peak values, the quality threshold starts to limit the number of reference samples. The accuracies get maximum values when H= 7, 8, 7 with Hownet, NTUSD, and SXU lexicons on UCI data set, when H= 5, 6, 6 with Hownet, NTUSD, and SXU lexicons on Weibo data set.

The window size H determines the surrounding contexts of specific units. So the parameter H is an essential factor to the classification performances. At the beginning, the accuracies increase as the selected units obtained more accurate semantic information. Then the performances decrease for too much irrelevant neighbors units. Taken together, these results suggest that there is an association between the accuracies and window sizes H.

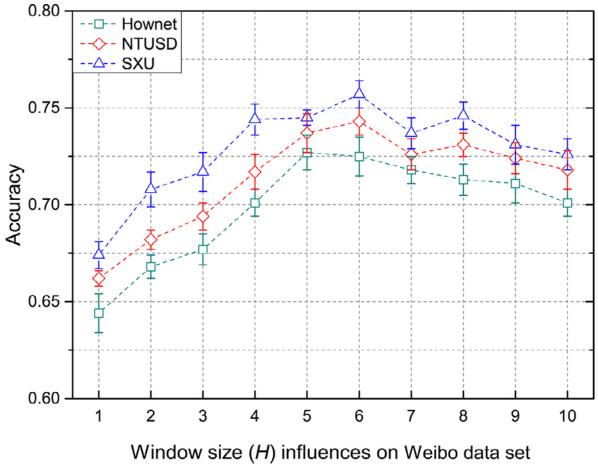
6.3.2. Size of seed sentiment units influence analysis

We test the relationship between the accuracies and the size of seed sentiment units M which change from 2% to 20% shown in Fig. 7(c) and (d). As Fig. 7(c) and (d) shown, the overall accuracies increase with the size of seed sentiment units. The proper values of size of seed sentiment units with Hownet, NTUSD, and SXU lexicons are 20%, 20%, and 20% on UCI data set, 20%, 20%, and 20% on Weibo data set.

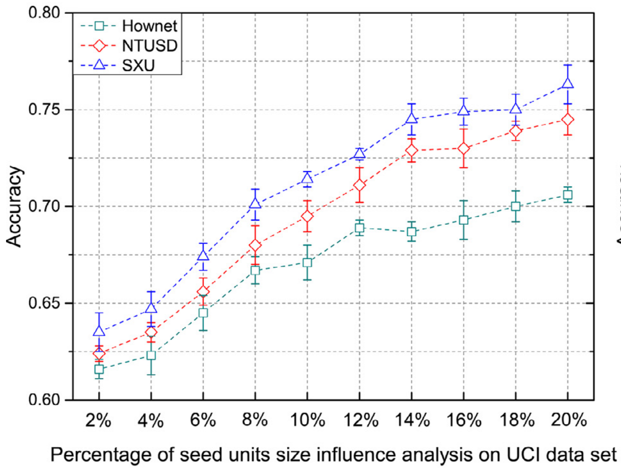
In fact, more seed sentiment units can get more sentiment information from existing sentiment dictionary resources. Unfortunately, existing sentiment dictionaries are generic rather than domain-specific, and some of labeled sentiment features are not accurate in specific domains. And some of domain-specific sentiment features do not appear



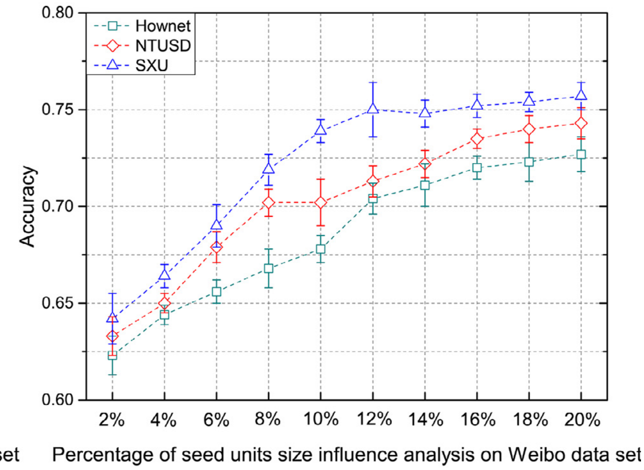
(a) Window size (H) influence analysis on UCI data set



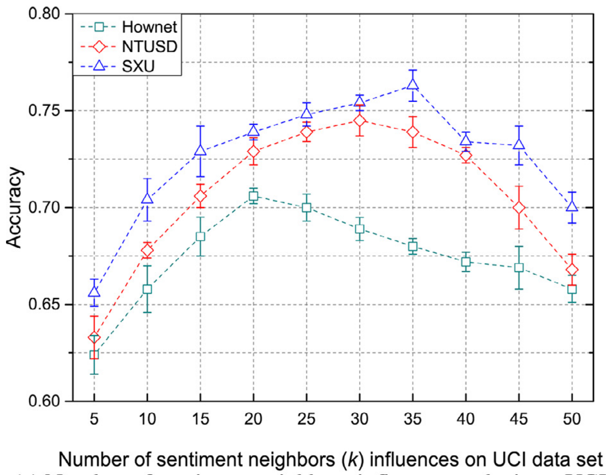
(b) Window size (H) influence analysis on Weibo data set



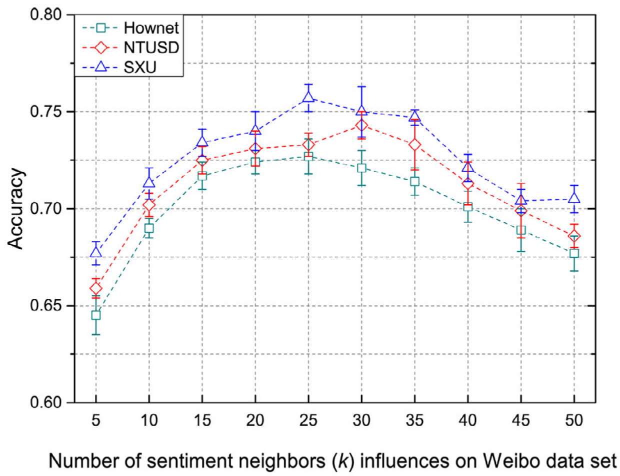
(c) Percentage of seed units size influence analysis on UCI data set



(d) Percentage of seed units size influence analysis on Weibo data set



(e) Number of sentiment neighbors influence analysis on UCI data set



(f) Number of sentiment neighbors influence analysis on Weibo data set

Fig. 7. Parameters influence analysis including size of seed sentiment units *H*, percentage of seed units size, number of sentiment neighbors *k*.

in the general dictionaries. Although these shortcomings can be made up for manual calibrations, this requires a lot of manual work.

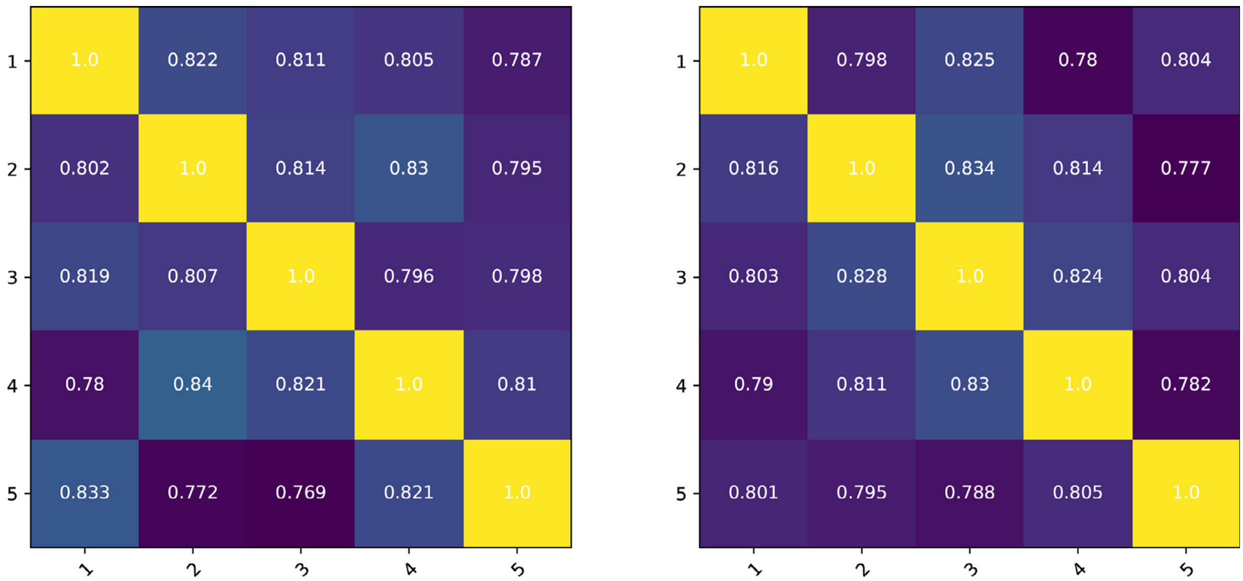
6.3.3. Number of sentiment neighbors influence analysis

To test the influence of sentiment neighbors, we fix other parameters unchanged and keep k changing from 5 to 50. It can be seen from Fig. 7(e) and (f) that the accuracies of SUCPF are increasing when the number of sentiment neighbors k increases and finally leveling off at a certain value. In addition, UCI and Weibo data sets have different stable points. The accuracies get maximum value when $k = 20, 30, 35$ with Hownet, NTUSD, and SXU lexicons on UCI data set, $k = 25, 30, 25$ with Hownet, NTUSD, and SXU lexicons on Weibo data set.

The number of sentiment neighbors affects the number of iterations, the stability of the calculation, and the final accuracies of the model greatly. Limited sentiment neighbors are difficult to capture semantic information effectively. If the number of sentiment neighbors is too big, the target units may be mixed with unrelated semantic information.

6.4. Generated sentiment lexicons

Through sentiment label propagation, we obtain the sentiment polarities and intensities of explicit and implicit sentiment features. The consistency ratio of sentiment dictionaries obtained by 5-fold cross-validation can be seen in seen in Fig. 8. Consistency ratio refers to the percentage of sentiment features and polarities contained in two sentiment dictionaries. The size of the sentiment dictionary obtained by different data sets is inconsistent. Therefore, the consistency ratio matrix is asymmetrical. We synthesize the sentiment dictionary of 5-fold cross-validation as the final sentiment dictionary. Specifically, the voting method is used to determine the sentiment polarity of sentiment features. The details of the generated sentiment lexicons on the UCI and Weibo data sets with the SXU general lexicon can be seen in Table 3. As we can see from Table 3, the UCI sentiment lexicon contains 11428 sentiment features, including 8665 explicit sentiment features and 2763 implicit features, and the Weibo sentiment lexicon contains 24415 sentiment features, including 20189 explicit sentiment features and 4226 implicit features. The majority of general sentiment lexicons are verbs and adjectives, while noun sentiment features account for 2763 (24.2%) and 4226 (17.5%) in the generated microblog-specific lexicons, respectively, on the UCI and Weibo data sets. This shows that the noun implicit sentiment features are important sentiment indicators in microblog sentiment



(a) Consensus ratio (%) of sentiment dictionaries of the UCI data set (b) Consensus ratio (%) of sentiment dictionaries of the Weibo data set

Fig. 8. Consensus ratio (%) of sentiment dictionaries obtained by 5-fold cross-validation.

Table 3
Generated sentiment lexicons including four parts of speeches (POS) and percentages.

POS	UCI			Weibo		
	Overall	Positive	Negative	Overall	Positive	Negative
Verbs	5358(46.9%)	2420(21.2%)	2938(25.7%)	12521(51.3%)	6270(25.7%)	6251(25.6%)
Adjectives	3122(27.3%)	1269(11.1%)	1853(16.2%)	7320(30.0%)	3211(13.2%)	4109(16.8%)
Adverbs	185(1.6%)	65(0.6%)	120(1.0%)	308(1.3%)	149(0.6%)	159(0.7%)
Nouns	2763(24.2%)	1355(11.9%)	1408(12.3%)	4266(17.5%)	1850(7.6%)	2416(9.9%)
Sum	11428(100%)	5106(44.8%)	6322(55.2%)	24415(100%)	11480(47.0%)	12935(53.0%)

expressions. A small number of adverbs such as “胡乱地” (casually), “幸好” (luckily), and “节俭地” (sparely) also express sentiment. Adverb sentiment features only occupy 185 (1.6%) on UCI data set and 308 (1.3%) on Weibo data set. Negative sentiment features are more common than positive sentiment features, for example, negative and positive sentiment features account for 6322 (55.2%) and 5106 (44.8%) respectively on UCI data set, and 12935 (53.0%) and 11480 (47.0%) respectively, on the Weibo data set.

As we can see from the extracted sentiment features, the method in this article demonstrates an excellent adaptability and can be adapted to different microblogs corpora. Our generated sentiment lexicons contain 11428 and 24415 sentiment features respectively, on the UCI and Weibo data sets, while the general SXU lexicon contains 29988 sentiment words. The scale of the generated sentiment lexicons is smaller than the general sentiment lexicons. This is because the microblog-specific sentiment lexicons are limited to distinctive sentiment features. In addition, the larger the corpus size, the greater the scale of the particular sentiment lexicons obtained. This proves that large-scale microblog corpora have abundant and diverse sentiment features, particularly, network vocabularies such as “土豪” (rich man), “拼装机” (assembled machine), and “贱行者” (slut).

The second focus of experimental research is to study implicit sentiment features mining tasks. We use a word cloud of fifty noun sentiment features on two microblog data sets as shown in Fig. 9. From Fig. 9, we can observe that these words and expressions strongly imply positive or negative sentiments and opinions. Different microblogs data sets have common implicit sentiment feature sets, such as “坏蛋” (badass), “流弊” (abuses), and “香菇” (mushroom). In comparing the results with baselines as shown in Tables 4 and 5, the accuracy of **SUCPF** is better than the **Explicit features** method. Our method can also find new sentiment features including explicit sentiment features such as “坏蛋” (badass), “神经病” (neuropathy), and “弱者” (the weak), implicit sentiment features such as “救火队员” (fireman), “潮男” (fashionable man), and “行业病” (occupational disease). Individual users are accustomed to expressing their sentiments about products, evaluations, or services indirectly using implicit sentiment features, especially in social media, such as microblogs. Therefore, the proposed method has good domain adaptation and recognition stability of explicit and implicit sentiment features.

6.5. *Sentiment classification results*

The sentiment classification results of the experimental system and the baseline systems on UCI and Weibo data sets are depicted in Tables 4 and 5, and the overall accuracy comparisons are shown in Fig. 10. From Tables 4, 5, and Fig. 10, we can observe that our method **SUCPF** outperforms the other five baseline methods, in terms of

(a) Twenty-five noun sentiment features of the UCI data set (b) Twenty-five noun sentiment features of the Weibo data set

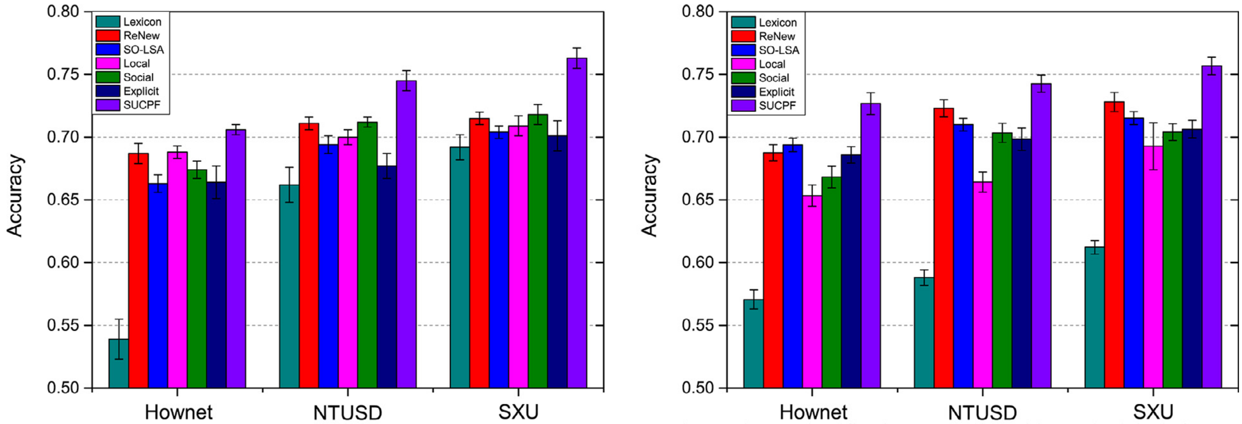
Fig. 9. Word cloud of fifty noun sentiment features on UCI and Weibo data sets.

Table 4
Sentiment classification results with standard deviation (%) on UCI data set under three different general sentiment lexicons.

Category	Evaluation	Lexicon	Renew	SO-LSA	Local	Social	Explicit	SUCPF
Hownet	Accuracy	53.9 ± 1.6	68.7 ± 0.8	66.3 ± 0.7	68.8 ± 0.5	67.4 ± 0.7	66.4 ± 1.3	**70.6 ± 0.4**
	Precision	63.1 ± 2.1	75.8 ± 1.6	72.9 ± 0.8	**77.7 ± 0.6**	74.4 ± 1.0	76.1 ± 1.4	76.4 ± 0.9
Positive	Recall	55.2 ± 1.8	71.0 ± 0.9	70.4 ± 1.2	71.2 ± 0.7	71.5 ± 1.3	68.9 ± 1.5	**77.2 ± 0.6**
	F1-score	58.9 ± 1.8	73.3 ± 1.1	71.6 ± 0.9	74.3 ± 0.5	72.9 ± 0.8	72.3 ± 1.4	**76.8 ± 0.6**
	Precision	47.8 ± 2.1	67.1 ± 1.0	65.4 ± 0.6	66.8 ± 0.9	66.1 ± 1.0	62.9 ± 1.2	**68.0 ± 0.7**
Neutral	Recall	61.1 ± 1.0	74.2 ± 0.3	70.6 ± 0.7	72.3 ± 0.3	73.4 ± 0.3	73.4 ± 0.5	**74.4 ± 0.2**
	F1-score	53.6 ± 1.7	70.5 ± 0.7	67.9 ± 0.6	69.5 ± 0.6	69.6 ± 0.7	67.7 ± 0.8	**71.1 ± 0.5**
	Precision	46.6 ± 2.7	51.8 ± 2.3	49.0 ± 3.1	47.7 ± 2.0	50.3 ± 3.1	48.6 ± 2.1	**61.6 ± 1.6**
Negative	Recall	35.5 ± 1.4	46.4 ± 1.6	42.8 ± 2.4	**48.6 ± 1.7**	41.2 ± 1.4	41.7 ± 2.9	46.3 ± 0.8
	F1-score	40.3 ± 1.5	48.9 ± 1.2	45.7 ± 2.7	48.1 ± 1.8	45.3 ± 2.0	44.9 ± 2.5	**52.9 ± 0.8**
NTUSD	Accuracy	66.2 ± 1.4	71.1 ± 0.5	69.4 ± 0.7	70.0 ± 0.6	71.2 ± 0.4	67.7 ± 1.0	**74.5 ± 0.8**
	Precision	72.1 ± 3.2	77.5 ± 0.8	76.8 ± 0.8	77.1 ± 0.9	78.6 ± 0.5	76.6 ± 1.1	**79.6 ± 0.7**
Positive	Recall	69.5 ± 2.8	74.3 ± 0.5	72.6 ± 0.8	73.5 ± 0.8	74.9 ± 0.7	71.6 ± 1.3	**81.6 ± 1.1**
	F1-score	70.7 ± 1.9	75.9 ± 0.6	74.6 ± 0.8	75.2 ± 0.8	76.7 ± 0.4	74.0 ± 1.2	**80.6 ± 0.8**
	Precision	65.8 ± 3.1	69.5 ± 0.4	67.3 ± 0.6	69.2 ± 0.6	68.0 ± 0.5	62.1 ± 0.9	**71.6 ± 0.7**
Neutral	Recall	74.0 ± 0.6	74.8 ± 0.1	72.5 ± 0.2	73.0 ± 0.2	75.7 ± 0.1	72.6 ± 0.3	**76.6 ± 0.4**
	F1-score	69.7 ± 2.0	72.1 ± 0.3	69.8 ± 0.4	71.0 ± 0.4	71.6 ± 0.3	66.9 ± 0.6	**74.0 ± 0.5**
	Precision	49.0 ± 2.1	56.5 ± 2.2	53.5 ± 2.3	50.8 ± 1.6	59.3 ± 1.7	59.9 ± 2.8	**68.9 ± 2.1**
Negative	Recall	39.0 ± 2.9	50.8 ± 1.8	50.2 ± 1.9	49.2 ± 2.1	49.4 ± 0.8	46.5 ± 2.3	**52.9 ± 1.3**
	F1-score	43.4 ± 1.9	53.5 ± 1.9	51.8 ± 2.0	50.0 ± 1.8	53.9 ± 1.1	52.3 ± 2.6	**59.8 ± 1.5**
SXU	Accuracy	69.2 ± 1.0	71.5 ± 0.5	70.4 ± 0.5	70.9 ± 0.8	71.8 ± 0.8	70.1 ± 1.2	**76.3 ± 0.8**
	Precision	77.0 ± 1.1	78.2 ± 0.4	79.0 ± 0.3	77.2 ± 0.8	79.3 ± 0.7	78.4 ± 1.0	**81.0 ± 0.6**
Positive	Recall	71.4 ± 1.2	74.7 ± 0.7	73.3 ± 0.9	75.9 ± 1.1	75.7 ± 1.3	73.1 ± 1.5	**82.0 ± 1.2**
	F1-score	74.1 ± 1.2	76.4 ± 0.5	76.1 ± 0.5	76.6 ± 0.9	77.5 ± 0.7	75.6 ± 1.3	**81.5 ± 0.9**
	Precision	66.9 ± 0.7	70.1 ± 0.6	68.0 ± 0.9	68.8 ± 1.0	68.3 ± 1.2	63.6 ± 0.8	**73.9 ± 0.9**
Neutral	Recall	74.1 ± 0.2	75.0 ± 0.2	72.7 ± 0.3	74.2 ± 0.3	75.8 ± 0.3	74.5 ± 0.2	**79.7 ± 0.4**
	F1-score	70.3 ± 0.5	72.4 ± 0.4	70.3 ± 0.6	71.4 ± 0.7	71.9 ± 0.8	68.6 ± 0.6	**76.7 ± 0.5**
	Precision	53.0 ± 2.0	55.3 ± 1.1	51.8 ± 1.6	58.5 ± 2.3	60.3 ± 3.0	67.0 ± 3.1	**70.0 ± 2.7**
Negative	Recall	48.4 ± 2.4	51.2 ± 0.8	52.8 ± 0.9	48.8 ± 1.2	50.4 ± 1.4	52.1 ± 2.3	**54.2 ± 3.1**
	F1-score	50.6 ± 2.2	53.2 ± 0.9	52.3 ± 1.2	53.2 ± 1.5	54.9 ± 1.9	58.6 ± 2.7	**61.1 ± 3.0**

Table 5
Sentiment classification accuracies with standard deviation (%) on Weibo data set under three different general sentiment lexicons.

Methods	Positive			Negative			Total
	Precision	Recall	F1-score	Precision	Recall	F1-score	Accuracy
Hownet							
Lexicon	52.6 ± 0.8	52.5 ± 0.3	52.5 ± 0.6	60.8 ± 0.6	60.8 ± 1.1	60.8 ± 0.8	57.1 ± 0.8
Renew	68.4 ± 0.6	57.6 ± 0.8	62.5 ± 0.8	69.0 ± 0.5	78.0 ± 0.5	73.2 ± 0.6	68.8 ± 0.6
SO-LSA	69.1 ± 0.9	58.5 ± 0.4	63.4 ± 0.5	69.6 ± 0.3	78.4 ± 0.8	73.7 ± 0.5	69.4 ± 0.5
Local	62.1 ± 1.3	60.2 ± 0.8	61.1 ± 0.5	68.0 ± 0.3	69.5 ± 2.0	68.7 ± 1.2	65.3 ± 0.8
Social	64.2 ± 1.0	60.4 ± 1.2	62.2 ± 1.0	68.8 ± 0.8	72.2 ± 0.9	70.4 ± 0.8	66.8 ± 0.9
Explicit	67.5 ± 1.0	58.9 ± 0.5	62.9 ± 0.6	69.3 ± 0.5	76.6 ± 1.0	72.8 ± 0.7	68.6 ± 0.7
SUCPF	**73.0 ± 1.2**	**63.0 ± 0.8**	**67.6 ± 1.0**	**72.5 ± 0.6**	**80.7 ± 1.0**	**76.4 ± 0.8**	**72.7 ± 0.9**
NTUSD							
Lexicon	54.7 ± 0.8	52.2 ± 0.6	53.4 ± 0.6	61.9 ± 0.5	64.2 ± 0.9	63.1 ± 0.6	58.8 ± 0.6
Renew	74.2 ± 1.1	59.4 ± 0.5	66.0 ± 0.7	71.2 ± 0.5	82.9 ± 0.8	76.6 ± 0.6	72.3 ± 0.7
SO-LSA	71.7 ± 0.7	59.4 ± 0.4	64.9 ± 0.5	70.6 ± 0.3	80.6 ± 0.6	75.3 ± 0.5	71.0 ± 0.5
Local	63.0 ± 1.0	**62.6 ± 0.5**	62.8 ± 0.7	69.2 ± 0.5	69.6 ± 1.1	69.4 ± 0.8	66.4 ± 0.8
Social	70.2 ± 1.2	60.0 ± 0.5	64.7 ± 0.7	70.4 ± 0.4	78.9 ± 1.2	74.4 ± 0.8	70.3 ± 0.8
Explicit	69.1 ± 1.3	60.3 ± 0.7	64.4 ± 0.9	70.3 ± 0.7	77.7 ± 1.1	73.8 ± 0.8	69.8 ± 0.9
SUCPF	**77.1 ± 1.2**	61.3 ± 0.6	**68.3 ± 0.7**	**72.6 ± 0.4**	**85.0 ± 1.0**	**78.4 ± 0.7**	**74.3 ± 0.7**
SXU							
Lexicon	57.6 ± 0.6	53.9 ± 0.7	55.7 ± 0.6	63.8 ± 0.4	67.2 ± 0.7	65.5 ± 0.5	61.2 ± 0.5
Renew	74.9 ± 1.4	60.0 ± 0.2	66.7 ± 0.7	71.6 ± 0.4	83.4 ± 1.2	77.1 ± 0.7	72.8 ± 0.8
SO-LSA	72.8 ± 0.9	59.2 ± 0.2	65.3 ± 0.5	70.8 ± 0.3	81.7 ± 0.7	75.9 ± 0.5	71.5 ± 0.5
Local	67.0 ± 2.2	**64.5 ± 0.8**	67.0 ± 2.6	72.0 ± 1.1	74.1 ± 2.0	72.3 ± 2.1	69.3 ± 1.9
Social	69.8 ± 1.1	60.9 ± 0.9	65.0 ± 0.7	70.8 ± 0.5	78.2 ± 1.3	74.3 ± 0.7	70.4 ± 0.7
Explicit	69.8 ± 1.0	62.0 ± 0.5	65.7 ± 0.7	71.2 ± 0.5	77.8 ± 1.0	74.4 ± 0.7	70.6 ± 0.7
SUCPF	**79.0 ± 1.0**	62.9 ± 1.0	**70.0 ± 1.0**	**73.7 ± 0.6**	**86.2 ± 0.6**	**79.5 ± 0.6**	**75.7 ± 0.7**



(a) Sentiment classification results with standard deviation on UCI data set

(b) Sentiment classification accuracies with standard deviation on Weibo data set

Fig. 10. Sentiment classification accuracies with standard deviation on UCI and Weibo data sets under three different general sentiment lexicons.

accuracy. For example, **SUCPF** improves 7.1%, 4.8%, 5.9%, 5.4%, 4.5%, and 6.2% respectively, when compared to the baselines on UCI data set with the SXU lexicon, 14.5%, 2.9%, 4.2%, 6.4%, 5.3%, and 5.1% respectively, when compared to baselines on the Weibo data set with the SXU lexicon. This proves that **SUCPF** can obtain a better domain-specific sentiment lexicon and sentiment classification results. This is mainly because our proposed **SUCPF** method is a semi-supervised framework utilizing existing sentiment lexicons and manual calibrations effectively. **SUCPF** can extract both implicit and explicit sentiment features using local contexts, social relationships, and topic features of sentiment units.

As shown in Fig. 10, general sentiment lexicon-based methods (**Lexicon**) obviously perform worse than other domain-specific lexicon generating methods. This proves that the sentiment expression is highly dependent on domains. In detail, the same word may express different sentiments in different domains, and different domains usually use different sentiment words. Therefore, we need to consider domain knowledge when generating domain-specific sentiment lexicons. From Tables 4 and 5, **SUCPF** improves 1.8%, 4.5%, and 5.4% with the Hownet, NTUSD, and SXU lexicons on the UCI data set, 7.4%, 7.9%, and 6.4% on the Weibo data set compared to the **Local** method. **SUCPF** improves 3.2%, 3.3%, and 4.5% with the Hownet, NTUSD, and SXU lexicons on the UCI data set, 5.9%, 4.0%, and 5.3% on the Weibo data sets when compared to the **Social** method. This proves that sentiment propagation requires both local and social context information.

SUCPF improves 4.2%, 6.8%, and 6.2% more than the **Explicit** features method on UCI data sets, 4.1%, 4.5%, and 5.1% than the **Explicit** method on Weibo data sets with Hownet, NTUSD, and SXU lexicons. This proves that implicit sentiment features are important sentiment indicators in microblogs.

Compared to **ReNew, SUCPF** improves 1.9%, 3.4%, and 4.8% on UCI data sets, 3.9%, 2.0%, and 2.9% on Weibo data sets. Unlike **ReNew**, which uses segments, our method uses the sentiment units as the basic sentiment elements. At the same time, we use the semantic relationships between different sentiment units rather than only using the relationships between different segments. On the other hand, we use an SPA to propagate sentiment from general sentiment features to microblog-specific sentiment features instead of using a CRF model.

Comparing with **SO-LSA, SUCPF** improves 4.3%, 5.1%, and 5.9% on UCI data sets, 3.3%, 3.3%, and 4.2% on Weibo data sets with Hownet, NTUSD, and SXU lexicons. The **SO-LSA** method judges the emotional polarity of candidate words based on the sentiment polarity of the seed words. The final effect often depends on the quality of the seed word. There is no iterative update process for this method. Our **SUCPF** method is to standardize the seed sentiment units according to the general sentiment dictionary, and then conduct manual calibration to provide the correct sources for sentiment propagation. And we get the accurate sentiment score of the sentiment units in the process of iterating.

6.6. Discussions

This paper uses a hybrid of rules and statistical strategies to construct microblog-specific sentiment lexicons including implicit and explicit sentiment features. Twenty linguistic semantic rules listed in Table 1 can determine the semantic relationships between sentiment units using grammar and syntactic knowledge. Statistical regularities summarize the probability statistic of sentiment units that occur in the same window. Local contexts, topic features, and user relationships in social networks are used to determine the relationships of candidate sentiment units. Our approach can propagate sentiment through edges in the sentiment propagation graph effectively. The experimental results, as shown in Tables 4 and 5, indicate that our methods outperform baseline methods significantly. This is an effective attempt to extract implicit sentiment features such as “油老虎” (gas-guzzler) and “拼装机” (erector machine) for microblog sentiment analysis. Through experimental verification, implicit sentiment features such as nouns or noun phrases can improve the effectiveness of sentiment classification. In general, this paper describes the basic work of generating domain-specific sentiment lexicons, which can improve microblog sentiment classification results effectively.

To verify the effectiveness of our framework, we carry out experimental verifications and detail analyses. Our framework is validated on two real microblog data sets, UCI and Weibo, respectively. We select the best parameters according to existing experiences and the experimental results as shown in Table 2. Table 3 shows the generated sentiment lexicons including the four parts of speech (POS) and percentages of the UCI and Weibo data sets. In the comparison of the results as shown in Fig. 10, our method outperforms previous microblog-specific sentiment lexicon methods. Experimental results demonstrate that our framework can construct domain-specific sentiment lexicons and extract implicit sentiment features effectively. We propose a novel sentiment unit context propagation framework (SUCPF) to generate microblog-specific sentiment lexicons for Chinese microblog data sets. While ReNew and SO-LSA methods adopt segments and word as the basic sentiment elements, respectively. Our generated microblog-specific lexicons not only contain explicit sentiment features, but also include implicit noun or noun phrase sentiment features. The two methods only consider explicit verbs or adverbs sentiment features. We construct a sentiment propagation graph and its adjacency matrix using social relationships, topic features, and local contexts. At the same time, we use the semantic relationships between different sentiment units, rather than only using the relationships between different segments in ReNew method and applying LSA to calculate the strength of the semantic association between words in SO-LSA method.

The sentiment expression of texts in social media are very complex, and most of the sentiment sentences do not have explicit sentiment features. On the contrary, users often use rhetorical expressions or factual descriptions. In traditional sentiment analysis of microblogs, researchers mainly focused on explicit sentiment features rather than implicit sentiment features. Therefore, it cannot effectively deal with the implicit sentiment expression in social media. It is important to note that implicit sentiment features are important for sentiment expression. As we can see from Tables 4 and 5, we can improve the recall rate of sentiment classification tasks when adding implicit sentiment features. This proves that implicit sentiment features are also important sentiment indicators in the sentiment expression of microblogs. People are accustomed to expressing sentiment with nouns or noun phrases. Some implicit sentiment features may indicate no sentiment initially, but imply strong sentiment in particular domains, such as the microblog domain. The sentiment expressions are context-dependent, and the features can be inferred as positive (negative) sentiments, which occur with positive (negative) sentiments frequently. Local and social contexts can capture the contextual information of sentiment features. Through context propagation, we can obtain the sentiment polarities and intensities of implicit sentiment features.

In this paper, there are three limitations for which improvement is required, data capacity restriction, sentiment lexicon resources constraint, and sensitive parameters selection, respectively. Co-occurrence relationships between sentiment units are relatively weak. Therefore, these relationships may be mixed with certain noises. If our corpora are large enough, the relationships between units have a certain statistical significance. It is difficult to achieve satisfactory results without enough data. Therefore, our method requires a large amount of microblogs data. The scale of data sets should have statistical significance of semantic and context relationships between sentiment units. Large-scale data sets contain rich sentiment features that can generate domain-specific sentiment dictionaries, and can contain enough co-occurrence semantic information. Larger data sets can be richer in explicit and implicit sentiment features, leading to larger and more comprehensive sentiment lexicons. In selecting different propagation resources, high-quality general sentiment lexicons can cause more accurate sentiment classification results. Our method makes full use of previous general sentiment lexicon resources and semantic relationships between sentiment units. Therefore, our proposed framework depends on linguistic data and general sentiment lexicons. If the quality of the general sentiment lexicon is not good or the corpus scale is limited, it would be difficult to obtain satisfactory microblog-specific sentiment lexicons. In the process of parameter experimental verifications, our framework is sensitive to parameter selection. For example, the number of sentiment neighbors affects the accuracy of our framework greatly.

7. Conclusions and future work

In this paper, we propose a sentiment unit context propagation framework to generate microblog-specific sentiment lexicons. This approach can overcome the five listed difficulties in the generation of microblog-specific sentiment lexicons, which are listed in the basic framework section. First, the selection of a seed sentiment units stage selects the seed sentiment units. After this, the seed sentiment units are labeled with general sentiment lexicons and manual calibrations. Then, the SPA is constructed to propagate the sentiment from the labeled sentiment units to the unlabeled ones. The sentiment score of words is calculated through their corresponding sentiment units. Finally, we obtain a microblog-specific sentiment lexicon for particular microblog domains. We apply the generated sentiment lexicons on two real-world microblog sentiment classification data sets, UCI and Weibo respectively. Experimental results demonstrate that our methods can extract implicit and explicit sentiment features and improve sentiment classification results effectively.

We intend to study other implicit sentiment tasks, such as implicit sentiment sentences extraction, fine-grained sentiment analysis, and emotion classification of implicit sentiment sentences. Moreover, the applications of the automatically generated microblog-specific sentiment lexicons to other context-sensitive sentiment analysis tasks will be explored, such as sentiment summarization, aspect-level sentiment analysis, and so on. Due to the scarcity of social network information and tagged microblogs data sets in Chinese, we have validated our framework under only two data sets. Next we will collect and annotate such data sets and validate our approach under more data sets.

Acknowledgments

This study was supported by the National Natural Science Foundation of China (61632011, 61573231, 61672331, 61432011, 61603229); the Shanxi Province Graduate Student Education Innovation Project (2016BY004, 2017BY004).

References

Austin, D., 2006. How google finds your needle in the web's haystack. Am. Math. Soc. Feature Column 10, 1–13.

Balahur, A., Hermida, J.M., Montoyo, A., 2011. Detecting implicit expressions of sentiment in text based on commonsense knowledge. In: Proceedings of the 2nd Workshop on Computational Approaches to Subjectivity and Sentiment Analysis. ACL, pp. 53–60.

Balahur, A., Hermida, J.M., Montoyo, A., 2012. Detecting implicit expressions of emotion in text: a comparative analysis. Decis. Support Syst. 53 (4), 742–753.

Bertero, D., Fung, P., 2017. A first look into a convolutional neural network for speech emotion detection. In: Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing. IEEE, pp. 5115–5119.

Bertero, D., Siddique, F.B., Fung, P., 2016. Towards a corpus of speech emotion for interactive dialog systems. In: Proceedings of Conference of The Oriental Chapter of International Committee for Coordination and Standardization of Speech Databases and Assessment Techniques (O-COCOSDA). IEEE, pp. 241–246.

Brin, S., Page, L., 1998. The anatomy of a large-scale hypertextual web search engine. Comput. Netw. Syst. 30 (1), 107–117.

Chao, A.F., Yang, H.-L., 2018. Using chinese radical parts for sentiment analysis and domain-dependent seed set extraction. Comput. Speech Lang. 47, 194–213.

Cho, H., Kim, S., Lee, J., Lee, J.-S., 2014. Data-driven integration of multiple sentiment dictionaries for lexicon-based sentiment classification of product reviews. Knowl. Based Syst. 71, 61–71.

Du, W., Tan, S., Cheng, X., Yun, X., 2010. Adapting information bottleneck method for automatic construction of domain-oriented sentiment lexicon. In: Proceedings of the 3rd ACM International Conference on Web search and Data Mining. ACM, pp. 111–120.

Duwairi, R., Ahmed, N.A., Al-Rifai, S.Y., 2015. Detecting sentiment embedded in arabic social media−a lexicon-based approach. J. Intel. Fuzzy Syst. 29 (1), 107–117.

Fu, X., Liu, W., Xu, Y., Cui, L., 2017. Combine hownet lexicon to train phrase recursive autoencoder for sentence-level sentiment analysis. Neurocomputing 241, 18–27.

Greene, S., Resnik, P., 2009. More than words: Syntactic packaging and implicit sentiment. In: Proceedings of the Annual Conference of the North American Chapter of the Association for Computational Linguistics. ACL, pp. 503–511.

He, B., Macdonald, C., He, J., Ounis, I., 2008. An effective statistical approach to blog post opinion retrieval. In: Proceedings of the 17th ACM Conference on Information and Knowledge Management. ACM, pp. 1063–1072.

Hu, Y., Li, W., 2011. Document sentiment classification by exploring description model of topical terms. Comput. Speech Lang. 25 (2), 386–403.

Huang, S., Niu, Z., Shi, C., 2014. Automatic construction of domain-specific sentiment lexicon based on constrained label propagation. Knowl. Based Syst. 56, 191–200.
Jijkoun, V., de Rijke, M., Weerkamp, W., 2010. Generating focused topic-specific sentiment lexicons. In: Proceedings of the 48th Annual Meeting of the Association for Computational Linguistics. ACL, pp. 585–594.
Kanayama, H., Nasukawa, T., 2006. Fully automatic lexicon expansion for domain-oriented sentiment analysis. In: Proceedings of the 2006 Conference on Empirical Methods in Natural Language Processing. ACL, pp. 355–363.
Van de Kauter, M., Breesch, D., Hoste, V., 2015. Fine-grained analysis of explicit and implicit sentiment in financial news articles. Expert Syst. Appl. 42, 4999–5010.
Kim, Y., Provost, E.M., 2016. Emotion spotting: Discovering regions of evidence in audio-visual emotion expressions. In: Proceedings of the 18th ACM International Conference on Multimodal Interaction. ACM, pp. 92–99.
Liu, L., Lei, M., Wang, H., 2013. Combining domain-specific sentiment lexicon with hownet for chinese sentiment analysis. J. Comput. 8 (4), 878–883.
Mukherjee, S., Bhattacharyya, P., 2012. Feature specific sentiment analysis for product reviews. In: Proceedings of the 13th International Conference on Computational Linguistics and Intelligent Text Processing-Volume Part I. Springer, pp. 475–487.
Pang, B., Lee, L., 2008. Opinion mining and sentiment analysis. Found. Trends Inform. Retrieval 2 (1−2), 1–135.
Park, S., Lee, W., Moon, I.-C., 2015. Efficient extraction of domain specific sentiment lexicon with active learning. Pattern Recog. Lett. 56, 38–44.
Qiu, G., Liu, B., Bu, J., Chen, C., 2009. Expanding domain sentiment lexicon through double propagation. In: Proceedings of the 21st International Joint Conference on Artificial Intelligence, pp. 1199–1204.
Rao, D., Ravichandran, D., 2009. Semi-supervised polarity lexicon induction. In: Proceedings of the 12th conference of the European Chapter of the Association for Computational Linguistics. ACL, pp. 675–682.
Siddique, F.B., Fung, P., 2017. Bilingual word embeddings for cross-lingual personality eecognition using convolutional neural nets. In: Proceedings of the INTERSPEECH, pp. 3271–3275.
Tan, S., Wang, Y., Cheng, X., 2008. Combining learn-based and lexicon-based techniques for sentiment detection without using labeled examples. In: Proceedings of the 31st Annual International ACM SIGIR Conference on Research and Development in Information Retrieval. ACM, pp. 743–744.
Turney, P.D., Littman, M.L., 2003. Measuring praise and criticism: Inference of semantic orientation from association. ACM Trans. Inf. Syst. 21 (4), 315–346.
Velikovich, L., Blair-Goldensohn, S., Hannan, K., McDonald, R., 2010. The viability of web-derived polarity lexicons. In: Proceedings of the Annual Conference of the North American Chapter of the Association for Lomputational Linguistics. ACL, pp. 777–785.
Wang, D., Liu, Y., 2015. Opinion summarization on spontaneous conversations. Comput. Speech Lang. 34 (1), 61–82.
Wang, S., Li, D., Song, X., Wei, Y., Li, H., 2011. A feature selection method based on improved fisher's discriminant ratio for text sentiment classification. Expert Syst. Appl. 38 (7), 8696–8702.
Weichselbraun, A., Gindl, S., Scharl, A., 2011. Using games with a purpose and bootstrapping to create domain-specific sentiment lexicons. In: Proceedings of the 20th ACM International Conference on Information and Knowledge Management. ACM, pp. 1053–1060.
Xie, Y., Chen, Z., Zhang, K., Cheng, Y., Honbo, D.K., Agrawal, A., Choudhary, A.N., 2013. Muses: multilingual sentiment elicitation system for social media data. IEEE Intel. Syst. 29 (4), 34–42.
Yang, Y., Ma, X., Fung, P., 2017. Perceived emotional intelligence in virtual agents. In: Proceedings of the CHI Conference Extended Abstracts on Human Factors in Computing Systems. ACM, pp. 2255–2262.
Zhang, B., Essl, G., Mower Provost, E., 2016. Automatic recognition of self-reported and perceived emotion: does joint modeling help? In: Proceedings of the 18th ACM International Conference on Multimodal Interaction. ACM, pp. 217–224.
Zhang, B., Provost, E.M., Essi, G., 2016. Cross-corpus acoustic emotion recognition from singing and speaking: A multi-task learning approach. In: Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). IEEE, pp. 5805–5809.
Zhang, C., Li, D., Liang, J., 2018. Hesitant fuzzy linguistic rough set over two universes model and its applications. Int. J. Mach. Learn. Cybern. 9 (4), 577–588.
Zhang, L., Liu, B., 2011. Identifying noun product features that imply opinions. In: Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies. ACL, pp. 575–580.
Zhang, Z., Singh, M.P., 2014. Renew: A semi-supervised framework for generating domain-specific lexicons and sentiment analysis. In: Proceedings of the 52nd Annual Meeting of Association for Computational Linguistics. ACL, pp. 542–551.
Zhao, C., Wang, S., Li, D., 2014. Fuzzy sentiment membership determining for sentiment classification. In: Proceedings of the IEEE International Conference on Data Mining Workshop. IEEE, pp. 1191–1198.
Zhao, C., Wang, S., Li, D., 2016. Determining fuzzy membership for sentiment classification: a three-layer sentiment propagation model. PLoS One 11 (11), e0165560.