12-2014

Issues of Social Data Analytics with a New Method for Sentiment Analysis of Social Media Data

Z. WANG

V. J. C. TONG

David CHAN
Singapore Management University, davidchan@smu.edu.sg

2014 IEEE 6th International Conference on Cloud Computing Technology and Science

Issues of social data analytics with a new method for sentiment analysis of social media data

Zhaoxia WANG, Victor Joo Chuan TONG
Social and Cognitive Computing (SCC) Department
Institute of High Performance Computing (IHPC)
Agency for Science, Technology and Research (A*STAR)
Singapore, 138632
{wangz, tongjc}@ihpc.a-star.edu.sg

David CHAN
Behavioural Sciences Institute
Singapore Management University
Singapore 188065
davidchan@smu.edu.sg

***Abstract*—Social media data consists of feedback, critiques and other comments that are posted online by internet users. Collectively, these comments may reflect sentiments that are sometimes not captured in traditional data collection methods such as administering a survey questionnaire. Thus, social media data offers a rich source of information, which can be adequately analyzed and understood. In this paper, we survey the extant research literature on sentiment analysis and discuss various limitations of the existing analytical methods. A major limitation in the large majority of existing research is the exclusive focus on social media data in the English language. There is a need to plug this research gap by developing effective analytic methods and approaches for sentiment analysis of data in non-English languages. These analyses of non-English language data should be integrated with the analysis of data in English language to better understand sentiments and address people-centric issues, particularly in multilingual societies. In addition, developing a high accuracy method, in which the customization of training datasets is not required, is also a challenge in current sentiment analysis. To address these various limitations and issues in current research, we propose a method that employs a new sentiment analysis scheme. The new scheme enables us to derive dominant valence as well as prominent positive and negative emotions by using an adaptive fuzzy inference method (FIM) with linguistics processors to minimize semantic ambiguity as well as multi-source lexicon integration and development. Our proposed method overcomes the limitations of the existing methods by not only improving the accuracy of the algorithm but also having the capability to perform analysis on non-English languages. Several case studies are included in this paper to illustrate the application and utility of our proposed method.**

Keywords—Social data; social media; Twitter; Weibo; sentiment analysis; fuzzy inference; multi-source lexicon; multilingual sentiment

I. INTRODUCTION

With the advent of the Internet and social media, users can readily offer their feedback and critiques by posting their comments online. They can also post reviews and opinions of services, products and policies through personal blogs, social networks and social media platforms such as Twitter, Facebook and Google+ [1] [2] [3]. Conversely, organizations may also disseminate information online to inform or influence online users. Such social data, which we will refer to collectively as social media data, provide a potentially rich source of information that can generate insights on how people think, feel and behave. Collectively, these comments may reflect sentiments that are sometimes not captured in traditional data collection methods such as administering a survey questionnaire.

Not surprisingly, social media data has played a critical role in advancing the field of social data analytics [4]. Collection, analysis and utilization of social media data have attracted research interest from both academics and industries. Social media data reflects users' emotions and attitudes on almost every topic for which they can find readers and listeners [5]. As a result, sentiment analysis of social media data has emerged as key issues to make sense of the social media data and to be utilized by public organizations and governments, as well as private organizations and citizen groups.

If adequately analyzed and understood, social media data offers a rich source of information that can offer useful insights for addressing both theoretical issues (e.g., testing theories of internet behaviors), practical problems in businesses (e.g., predicting consumer preferences) and public policy (e.g., sensing public sentiments). To translate the large volume and variety of social media data into useful information, we need to apply relevant data analytics to perform a rigorous sentiment analysis of social media data.

In this paper, we provide a survey with extant research literature on sentiment analysis and discuss the main limitations of the current analytical methods. We discuss the need to study social media data in other languages, in addition to the current focus on social media data in the English language. There is also a need for more accurate methods of sentiment analysis on social media. To address these various limitations and issues in current research, we propose a method that employs a new sentiment analysis scheme. We also discuss several case studies to illustrate the application and utility of our proposed method.

This work is supported by the Social Technologies+ Programme, which is granted by the Joint Council Office (JCO) at the Agency for Science Technology and Research (A*STAR).

978-1-4799-4093-6/14 $31.00 © 2014 IEEE
DOI 10.1109/CloudCom.2014.40

II. RELATED WORKS

In the past decade, there has seen a rapid growth in studies on sentiment analysis from text data. Much work has been done to focus on classifying text emotion polarity in terms of positive, negative and neutral [6]. Sentiment analysis has been reported in studies in several business domains such as forecasting daily box office revenue of movies [7] and predicting short-term stock market performance of companies [8]. These methods involve lexicon-based approaches [9] or machine-learning algorithms [10].

Lexicon-based methods are commonly used techniques, but the performances of such systems are limited by semantic ambiguity [11]. For instance, Rao et al. developed an algorithm with three pruning strategies to automatically build a word-level emotional dictionary for social emotion detection [9] [12] .

Although Machine-learning algorithms can outperform simple lexicon-based methods [10] [13], they require large training databases to be effective [10] [14]. For most real-world social media contexts which involve huge datasets, it is difficult to obtain the effective size of a sufficient training dataset because the diversity of the social discussion is often not known *a priori*. Rui and Whinston trained a support vector machine (SVM) with a training dataset and the precision for detecting positive, negative, and neutral tweets are 75%, 65% and 75%, respectively [7].

Feldman noted that many of the commercial sentiment analysis systems continue to use simplistic techniques and their performance leaves a lot to be desired [15]. Although there are reports of hybrid methods, which combined rule-based classification, supervised learning and machine learning [16], they suffer from the same limitations as machine-learning methods, including the issue on insufficient training data to be effective.

Another challenge in current research on sentiment analysis involves requirements of topic domain-specific adaptation. Blitzer et al. investigated domain adaptation for sentiment classifiers for different types of products [17]. Their research indicated that sentiments are expressed differently in different domains, and annotating corpora for every possible domain of interest is impractical. As human express their attitudes and emotions very differently in different linguistic groups and social context as well as topic domains, existing sentiment classification methods face this challenge. For example, the word "fast" is a positive word in a sentence describing a train service, as in "today's train service is pleasant and exceptionally fast". While it will be a negative word in a sentence describing the lifespan of cellular phone batteries as in "My cellular phone runs out of power fast".

Compared with English, it is more difficult to perform sentiment analysis in Mandarin Chinese due to the complexity of the language [18]. In China, microblogs are generally referred to as "Weibo". Sentiment analysis on Weibo is more challenging due to the characters of Chinese phonetics. Both Weibo and Twitter have a 140-character limit [18]. Unlike English tweets, there are no spaces between Chinese words and each word consists of one or more Chinese characters. For example, the words "very clear" have 9 characters when typed in English and there is a space between the two words. However, the same words contain four Chinese characters when typed in Chinese "非常清楚" and there are no spaces between the words. Therefore, the text information in a post or a review in Chinese may be more complex than one in English [19]. Consequently, a sentiment analysis method developed for text in English may not be directly equipped to handle the complexities and uniqueness of the Chinese text. However, it can be extended to handle this issue indirectly.

Examples of previous studies on sentiment analysis of Chinese text include using hybrid association-rule based mining to detect product weaknesses [20], using online sequential extreme learning machine methods and intuitionistic fuzzy sets to predict consumer sentiments [21], and assessing changes in public sentiment expressions during the 10-day period after the "7.23 Wenzhou Train Collision" on 23 July 2011 [22]. For these studies, the main data analytic challenges faced by the researchers included semantic ambiguity [20] and the need for a very large training dataset [23].

To date, most efforts in analyzing social opinions have been focused on the English language, and there are very few studies on cross-lingual sentiment sense mapping [24] [25] [26] [27]. For the few studies that examined text in non-English languages, the methods used were previously published methods (sometimes with minor adaptations) for English sentiment analysis[24].

Balahur and his colleagues argued that sentiment analysis methods should include creating corpora for non-English languages in their literature review [25]. Kaur et al. described the survey on performing sentiment extraction on various Indian languages like Bengali, Hindi, Telugu and Malayalam. They noted the lack of research on sentiment analysis for Indian languages and suggested further research on cross-lingual sentiment sense mapping [26].

Boiy and Moens applied several classification models such as Support Vector Machine (SVM), Multinomial Naïve Bayes (MNB), and Maximum Entropy (ME) to analyze sentiments from comments in blogs, reviews and forums in English, Dutch and French. The experiments had achieved reasonable performance for English (accuracy = 83%), but poor performance for Dutch (accuracy = 70%) and French (accuracy = 68%) [27].

III. CURRENT LIMITATIONS OF THE EXISTING WORKS

Our above literature review has revealed the following limitations of current research on sentiment analysis of social media data:

1. Existing methods may not be well equipped to address the difficulties associated with human language complexity and semantic ambiguity.
2. The majority of existing methods are tested using limited annotated corpus. Such datasets can be a cleaned sample that consists of either positive or negative text, or manually filtered and annotated. This is different from real-

world sentiment classification task (e.g., on tweets), which may also contain spam, advertisement, and bipolar texts.

3. Existing methods have to deal with the requirements of topic domain-specific adaptation and training database. It is important to develop high accuracy domain-specific adaptation methods that can function without the need of customized training datasets.

4. There is a lack of sentiment analytics for non-English languages. This research gap is evident given the rise of social media in non-English languages. Multilingual sentiment analysis is also important for social sentiment analysis in multilingual societies.

IV. Proposed Scheme

We have developed a new sentiment analysis scheme to address the above limitations. By using an adaptive fuzzy inference method (FIM) with linguistics processors to minimize semantic ambiguity as well as multi-source lexicon integration and development, our new scheme enables us to derive dominant valence as well as prominent positive and negative emotions.

The English-based sentiment analysis has shown that this is a viable approach which achieves significant improvements in predictive performance as compared to existing machine-learning techniques without the need of a training dataset. The difficulties of existing machine-learning methods to analyze real-world data are attributable to the limited annotated corpus used in systems training. Such datasets are usually cleaned samples that contain either positive or negative text, or manually filtered and annotated ones. This is different from real-world sentiment classification tasks (e.g., tweets) which may also contain spam, advertisement, and bipolar texts. For this study, real-world datasets will be used instead of limited historical datasets.

To maximize the representativeness of the characteristics of social media expressions in our method, our reference techniques include: linguistic inquiry and word count (LIWC) method [28] [29]; the pleasure, arousal and dominance (PAD) model of emotional states [6]; the affective norms for English words (ANEW) approach for assigning normative emotional ratings to text [30] [31]; and fuzzy logic [32] [33], which permits inexact or many-valued inferences to be obtained from textual sources. The fuzzy rules follow the typical format of: IF <Condition A1> AND <Condition A2> and <Condition A3>, THEN <Conclusion B1> [32].

For each piece of text, we define the *feature vector* $F = \{f_1, f_2, \ldots, f_k, \ldots, f_K\}$, where f_k is the k^{th} feature. Each feature $f_k \in F$ is expressed with a finite set of words, phrases or their abbreviations. We further define the *M intermediate categories* $\{m_1, m_2, \ldots, m_m, \ldots, m_N\}$. Quintuples ($\boldsymbol{O_j}$, $\boldsymbol{F_j}$, $\boldsymbol{V_j}$, $\boldsymbol{h_j}$, t_j) [34] are leveraged by using the objects ($\boldsymbol{O_j}$), opinion holders ($\boldsymbol{h_j}$), representative features ($\boldsymbol{F_j}$), values of the features ($\boldsymbol{V_j}$), and the time (t_j) to describe a piece of target text. The values of the features ($\boldsymbol{V_j}$) are described according to fuzzy theory [32] [33] as well as the theory of core affect and emotion concepts [6] [30] [31]. The information of the quintuples is mined towards the target object. For each quintuple towards the target object, the sentiment orientation is classified into *intermediate categories* (M) for each piece of target text. Fuzzy union operator [32] [33] is performed on the sentiment *intermediate category* classifications towards each target object for each piece of text.

The above procedure of the scheme can be conducted and extended adaptively through analyzing the characteristics of different languages, such as English and Chinese, which will be discussed in the next section.

V. Preliminary Results and Discussion

An English-based sentiment analysis engine has been developed accordingly and applied to study social sentiments on tweets related to public transportation (see Fig. 1).

We combined adaptive fuzzy inference method (FIM) with linguistics processors to minimize semantic ambiguity. We then merged this combined method with multi-source lexicon integration and development. As a result, we were able to derive dominant valence as well as their prominent emotions without the need of a training dataset. The system can classify sentiments related to public transportation with an accuracy of 88.9% on a pre-classified dataset which was identified by the two domain experts and classified by eight annotators into positive and negative items [10]. Machine-learning methods such as support vector machines (SVM), naïve Bayes (NB) and Maximum Entropy (MaxEnt) have also been explored with accuracy of 75.0%, 69.4%, and 75.0% respectively for analyzing the same datasets [10].

Fig. 1 shows some of the results obtained by the proposed sentiment analysis method. Fig. 1 (a) shows the daily number of comments collected. Currently, the proposed method can automatically classify a twitter message (i.e. tweet) into primary sentiment categories (i.e., Positive, Negative, Neutral and Ambivalence). For the Ambivalence outcomes, we conducted further analysis to classify them into Positive or Negative sentiments. If the ambivalent tweet expresses more positive than negative emotions/attitudes, it will be classified into the Positive sentiment category. Vice versa, if the ambivalent tweet expresses more negative than positive emotions/attitudes, it will be further classified into the Negative sentiment category as shown in Fig. 1 (b). The proposed method can also identify its prevailing negative emotion subcategories (e.g., Anger, Sadness, and Anxiety) as shown in Fig. 1 (c). More work needs to be done to generalize the engine and increase the number of emotion subcategories.

Weibo, also known as Chinese Twitter, which was launched in 2009, has accumulated more than 500 million registered users in less than four years. Everyday there will be more than 100 million Chinese tweets posted [35].

For analyzing Chinese tweets, a prototype topic-specific Chinese-based sentiment analysis engine has been developed and applied to study Chinese consumer preferences for smart phone products. It is a knowledge-based system enriched with Chinese lexicons as well as Chinese sentence composing techniques. Initial analysis has shown that the basic system is capable of handling Chinese Mandarin with an average accuracy of 80.0% on a benchmark dataset of 4,995 entries derived from http://www/datatang.com.

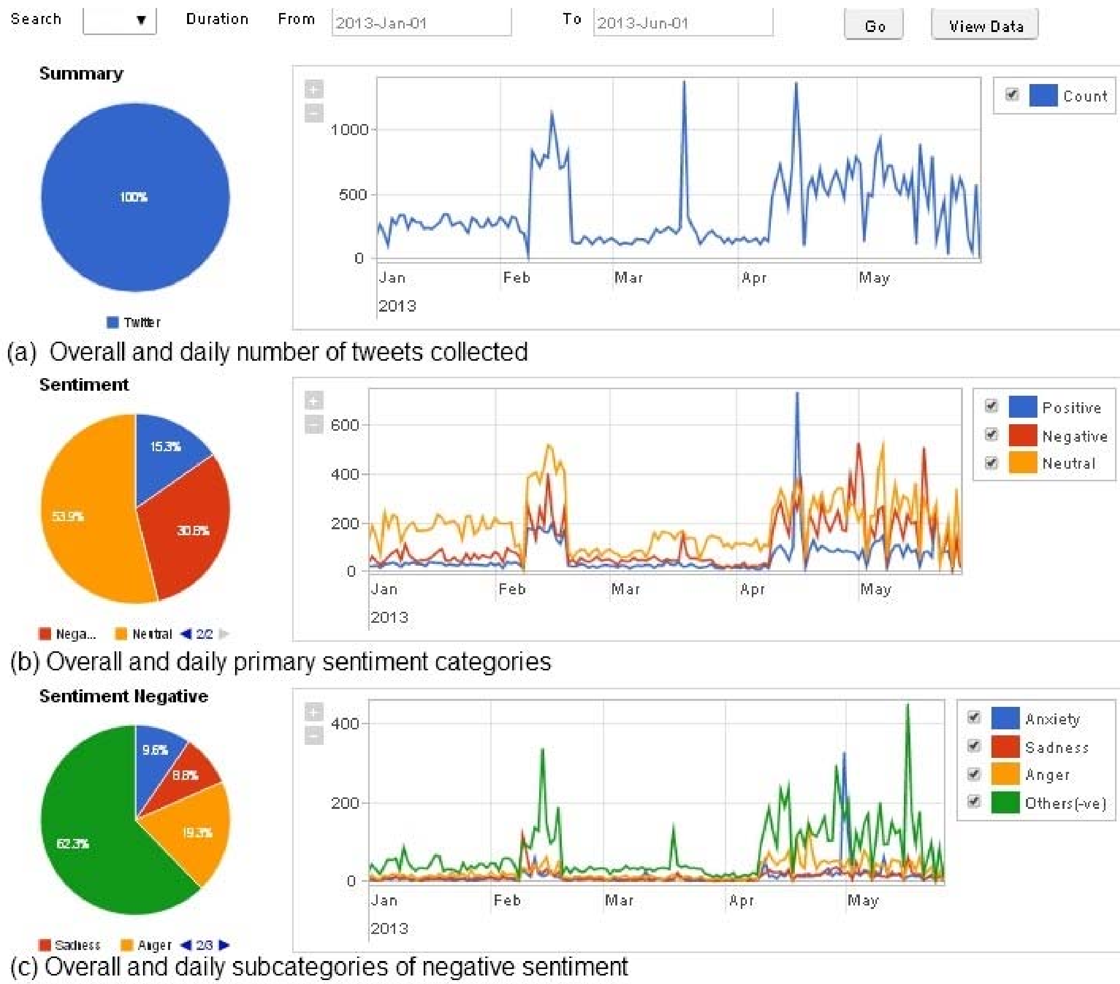
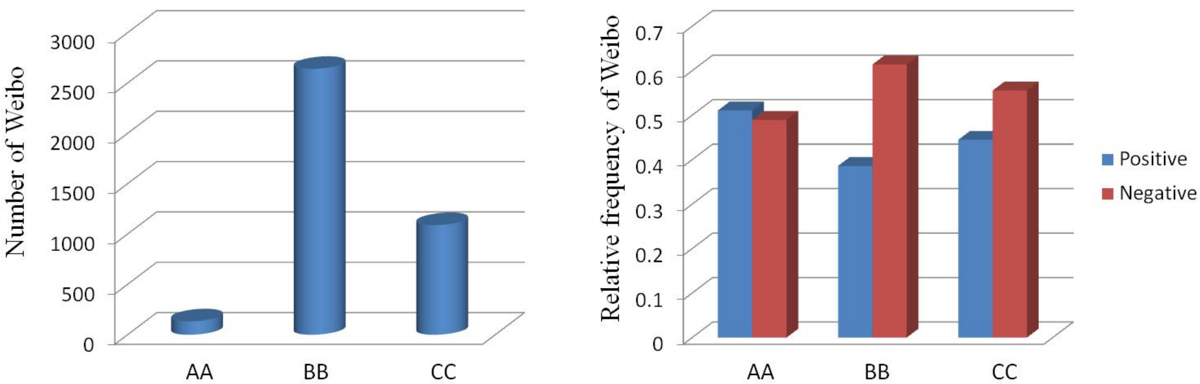


Fig. 1. Sentiment classification by the SentiMo engine. (a) Summary of the tweet collected; (b) Division of sentiments into three primary categories of positive, negative and neutral; (c) Subcategories of negative sentiment results.

Fig. 2 shows the preliminary results of how often each brand of smart phones are mentioned on Weibo during the sample period from 1 June to 31 June 2014 (Here we mask the product, service and companies' name to keep the information confidential to respect their privacy and named them AA, BB, and CC).

A particular AA product is discussed less by Weibo users during the sample period, but there are more positive comments than negative comments as shown in Fig. 2 (a) and (b). Although the targeted BB product is more commonly mentioned on Weibo, the relative frequency of negative comments is higher than the other two brands that we have studied during the sample period. The proposed method can automatically classify a message (i.e. Weibo) into primary sentiment categories, i.e. positive (正面), negative (负面), neutral (中性) and ambivalence (矛盾). However, the accuracy is lower than the English-based method. Improving the performance of the method is our main priority. Detecting the sentiments in Chinese texts to extract the information according to the language logic is a challenge. This is because the syntax and structure of Chinese varies from English. More subcategories of positive (正面) and negative (负面) emotion subcategories should also be included in future investigations, i.e. happiness (快乐), satisfaction (满意), admiration (赞赏), longing (魂牵梦萦), disappointment (大失所望), fear (恐惧), criticism (批评) and disgust (厌恶).

Our proposed scheme has enabled significant achievements in sentiment analysis. We are currently working to improve the effectiveness and utility of the methods. We are also developing a platform, which enables real-time data collection and analysis, to make our above methods available for use by layman to analyze the data on social media websites (Fig.1 is a screenshot of current platform displaying results of sentiment analysis).



(a) (b)

Fig. 2. Chinese-based sentiment analysis on Weibo for Chinese consumer preferences for smart phone products. (a) Number of discussions during the sample period; (b) Categorization of sentiments based on Positive and Negative remarks about the products. Results based on sample period from 1 June to 31 June 2014.

VI. CONCLUDING REMARKS

In this paper, we have outlined several key issues faced by researchers when performing sentiment analysis of social media data. These include problems with commonly used methods (i.e., lexicon based and machine-learning algorithms) and their causes. The main issues include a) The difficulties in handling the complexity and semantic ambiguity of languages; b) The difficulties to handle real-world data sets raised from real-world problems due to fact that existing methods being tested by using limited annotated corpus; c) The requirement of topic domain-specific adaptive methods to obtain high accuracy readings as well as dependence of training dataset; and d) the lack of sentiment analytics to analyze data in other languages besides English.

These key issues need to be addressed urgently and adequately because they could affect the accuracy of sentiment analysis and hence, the appropriateness of applying a specific analytic method and the validity of the inferences from the results. Our research findings provided preliminary evidence that our proposed scheme is a reliable and valid method for sentiment analysis and outperforms existing machine-learning methods. It can achieve higher accuracy for analyzing data in English language and can be extended to handle data in Chinese. We are currently working to improve the performance of the proposed method as well as add more subcategories of positive and negative emotions into future methods. We will also explore the possibility of extending our methods for analyzing social media data in other languages.

ACKNOWLEDGMENT

The authors would like to thank Dr. GOH Siow Mong Rick, Dr. YANG Yinping, Dr. QUEK Boon Kiat, Dr. SAKAMOTO Kayo, Dr. Ilya Benjamin FARBER, Dr. Martin SAERBECK, Dr. Joseph SIMONS and Dr. Sebastian FELLER for their valuable discussions, comments and invaluable help in conducting this study. The authors would like to thank Dr. QUEK Boon Kiat, Dr. YIN Xiao Feng, Mr. XIN Xin and Mr. ONG Boon Som Jimmy for their efforts on the platform development. Many thanks to Mr. LU Sifei for the database support.

REFERENCES

[1] M. Naaman, "Social multimedia: highlighting opportunities for search and mining of multimedia data in social media applications," Multimedia Tools and Applications, vol. 56, no. 1, pp. 9-34, May 2010.

[2] W.-ying S. Chou, Y. M. Hunt, E. B. Beckjord, R. P. Moser, and B. W. Hesse, "Social media use in the United States: implications for health communication.," Journal of medical Internet research, vol. 11, no. 4, p. e48, Jan. 2009.

[3] M. Salathé and S. Khandelwal, "Assessing vaccination sentiments with online social media: implications for infectious disease dynamics and control.," PLoS computational biology, vol. 7, no. 10, p. e1002199, Oct. 2011.

[4] B. Gunter, N. Koteyko, and D. Atanasova, "Sentiment analysis: A market-relevant and reliable measure of public feeling?," International Journal of Market Research, vol. 56, no. 2, p. 231, 2014.

[5] A. Go, R. Bhayani, and L. Huang, "Twitter sentiment classification using distant supervision," CS224N Project Report, Stanford, vol. 1–12, 2009.

[6] Y. Bae and H. Lee, "Sentiment analysis of Twitter audiences: Measuring the positive or negative influence of popular twitterers," Journal of the American Society for Information Science and Technology, vol. 63, no. 12, pp. 2521-2535, 2012.

[7] H. Rui and A. Whinston, "Designing a social-broadcasting-based business intelligence system," ACM Transactions on Management information Systems, vol. 2, no. 4, p. 22:1, 2011.

[8] Y. Yu, W. Duan, and Q. Cao, "The impact of social and conventional media on firm equity value: A sentiment analysis approach," Decision Support Systems, vol. 55, no. 4, pp. 919-926, Nov. 2013.

[9] Y. Rao, J. Lei, L. Wenyin, Q. Li, and M. Chen, "Building emotional dictionary for sentiment analysis of online news," World Wide Web, vol. 17, pp. 723-742, Jun. 2014.

[10] Z. Wang, V. J. C. Tong, and H. C. Chin, “Enhancing machine-learning methods for sentiment classification of Web data,” in 10th Asia Information Retrieval Society Conference, 2014.

[11] A. Balahur, R. Mihalcea, and A. Montoyo, “Computational approaches to subjectivity and sentiment analysis: Present and envisaged methods and applications,” Computer Speech & Language, vol. 28, no. 1, pp. 1-6, Jan. 2014.

[12] V. Loia and S. Senatore, “A fuzzy-oriented sentic analysis to capture the human emotion in Web-based content,” Knowledge-Based Systems, vol. 58, pp. 75-85, Mar. 2014.

[13] T. Wilson, J. Wiebe, and P. Hoffmann, “Recognizing contextual polarity: An exploration of features for phrase-level sentiment analysis,” Association for Computational linguistics, vol. 35, no. 3, 2009.

[14] S. Kim, N. W. Cho, B. Kang, and S.-H. Kang, “Fast outlier detection for very large log data,” Expert Systems with Applications, vol. 38, no. 8, pp. 9587-9596, Aug. 2011.

[15] R. Feldman, “Techniques and applications for sentiment analysis,” Communications of the ACM, vol. 56, no. 4, p. 82, Apr. 2013.

[16] R. Prabowo and M. Thelwall, “Sentiment analysis: A combined approach,” Journal of Informetrics, vol. 3, no. 2, pp. 143-157, Apr. 2009.

[17] J. Blitzer, “Biographies, Bollywood, Boom-boxes and Blenders: Domain adaptation for sentiment classification,” ACL, vol. 7, pp. 440-447, 2007.

[18] W. Wang, H. Xu, and W. Wan, “Implicit feature identification via hybrid association rule mining,” Expert Systems with Applications, vol. 40, no. 9, pp. 3518-3531, Jul. 2013.

[19] B. Yuan, Y. Liu, and H. Li, “Sentiment Classification in Chinese Microblogs : Lexicon-based and Learning-based Approaches,” pp. 1-6, 2013.

[20] W. Zhang, H. Xu, and W. Wan, “Weakness Finder: Find product weakness from Chinese reviews by using aspects based sentiment analysis,” Expert Systems with Applications, vol. 39, no. 11, pp. 10283-10291, Sep. 2012.

[21] H. Wang, G. Qian, and X.-Q. Feng, “Predicting consumer sentiments using online sequential extreme learning machine and intuitionistic fuzzy sets,” Neural Computing and Applications, vol. 22, no. 3–4, pp. 479-489, Feb. 2012.

[22] W. Shi, H. Wang, and S. He, “Sentiment analysis of Chinese microblogging based on sentiment ontology: a case study of ‘7.23 Wenzhou Train Collision’,” Connection Science, vol. 25, no. 4, pp. 161-178, Dec. 2013.

[23] X. Xiong, G. Zhou, Y. Huang, H. Chen, and K. Xu, “Dynamic evolution of collective emotions in social networks: a case study of Sina weibo,” Science China Information Sciences, vol. 56, no. 7, pp. 1-18, Jul. 2013.

[24] A. Ortigosa, J. M. Martín, and R. M. Carro, “Sentiment analysis in Facebook and its application to e-learning,” Computers in Human Behavior, vol. 31, pp. 527-541, Feb. 2014.

[25] A. Balahur and M. Turchi, “Comparative experiments using supervised learning and machine translation for multilingual sentiment analysis,” Computer Speech & Language, vol. 28, no. 1, pp. 56-75, Jan. 2014.

[26] A. Kaur and V. Gupta, “A survey on sentiment analysis and opinion mining techniques,” Journal of Emerging Technologies in Web Intelligence, vol. 5, no. 4, pp. 367-372, Nov. 2013.

[27] E. Boiy and M.-F. Moens, “A machine learning approach to sentiment analysis in multilingual Web texts,” Information Retrieval, vol. 12, no. 5, pp. 526-558, Sep. 2009.

[28] Y. R. Tausczik and J. W. Pennebaker, “The psychological meaning of words: LIWC and computerized text analysis methods,” Journal of Language and Social Psychology, vol. 29, no. 1, pp. 24-54, Dec. 2010.

[29] P. Korenek and M. Šimko, “Sentiment analysis on microblog utilizing appraisal theory,” World Wide Web, vol. 17, no. 4, pp. 847-867, Aug. 2013.

[30] A. Trilla and F. Alías, “Sentence-based sentiment analysis for expressive text-to-speech,” Audio, Speech, and Language Processing, IEEE Transactions on, vol. 21, no. 2, pp. 223-233, 2013.

[31] P. Gonçalves and M. Araújo, “Comparing and combining sentiment analysis methods,” In Proceedings of the first ACM conference on Online social networks. ACM., pp. 27-38, 2013.

[32] C. S. Chang, Z. Wang, F. Yang, and W. W. Tan, “Hierarchical fuzzy logic system for implementing maintenance schedules of offshore power systems,” IEEE Transactions on Smart Grid, vol. 3, no. 1, pp. 3-11, Mar. 2012.

[33] J. M. Mendel and D. Wu, “Challenges for perceptual computer applications and how they were overcome,” IEEE computational intelligence magazine, vol. 7, no. 3, pp. 36 - 47, 2012.

[34] B. Liu, “Sentiment analysis: A multi-faceted problem,” IEEE Intelligent Systems, vol. 25, no. 3, pp. 76-80, 2010.

[35] J. Sullivan, “China’s Weibo: Is faster different?,” New Media & Society, vol. 16, no. 1, pp. 24-37, Feb. 2013.