

# CT-SPA: Text Sentiment Polarity Prediction Model Using Semi-automatically Expanded Sentiment Lexicon

**Tao-Hsing Chang**

Department of Computer Science  
and Information Engineering  
National Kaohsiung University of  
Applied Sciences  
changth@gm.kuas.edu.tw

**Ming-Jhih Lin**

Department of Computer Science  
and Information Engineering  
National Kaohsiung University of  
Applied Sciences  
1101108139@kuas.edu.tw

**Chun-Hsien Chen**

Department of Computer Science  
and Information Engineering  
National Kaohsiung University of  
Applied Sciences  
1101108103@kuas.edu.tw

**Shao-Yu Wang**

Department of Computer Science  
and Information Engineering  
National Kaohsiung University of  
Applied Sciences  
1101108143@kuas.edu.tw

## Abstract

In this study, an automatic classification method based on the sentiment polarity of text is proposed. This method uses two sentiment dictionaries from different sources: the Chinese sentiment dictionary CSWN that integrates Chinese WordNet with SentiWordNet, and the sentiment dictionary obtained from a training corpus labeled with sentiment polarities. In this study, the sentiment polarity of text is analyzed using these two dictionaries, a mixed-rule approach, and a statistics-based prediction model. The proposed method is used to analyze a test corpus provided by the Topic-Based Chinese Message Polarity Classification task of SIGHAN-8, and the F1-measure value is tested at 0.62.

## 1 Introduction

The automatic text sentiment analysis method is an essential part in many big data analytics applications. For example, in opinion mining applications, the reviews for a certain movie in an online movie community are classified into positive or negative opinions (Kennedy & Inkpen, 2006). In addition, there are commercial organizations that analyze real-time social media content. When a large number of positive or negative posts on the user experience of a client's product appear suddenly in social media, these organizations automatically create an analysis report and send it to their client, thus allowing the client to gain more time for crisis response (Feldman, 2013).

There have been numerous studies on how to analyze sentiment tendency expressed in text. Most such algorithms rely on lexicon-based methods (Taboada, Brooke, Tofiloski, Voll, & Stede, 2011) that normally comprise assigning positive or negative sentiment values to words in the documents according to a sentiment dictionary, and then evaluating the sentiment orientation of the text according to different classification methods, such as weighting method and k-means. These methods can obtain very good results in certain standard tests, such as Epinions' positive and negative product review corpus. A sentiment dictionary is constructed in such a way that every word in the dictionary is assigned to a sentiment category (also called polarity), either positive or negative. Sentiment polarity labeling for these dictionaries is performed manually, semi-automatically, or automatically. Manually labeled sentiment dictionaries have been developed for many years, but most dictionaries are only labeled with polarities without polarity strength. SO-CAL, proposed by Taboada et al., is an English dictionary labeled with both polarity and strength.

In view of the restrictions associated with manually built dictionaries, several researchers have adopted semi-automatic or fully automatic methods to build sentiment dictionaries based on existing resources or large amounts of linguistic data. For example, SentiWordNet (Baccianella, Esuli, & Sebastiani, 2010) is a WordNet-based sentiment dictionary where the polarity strength of every sentiment word is labeled after analysis of the documents labeled with sentiment polarities. The Chinese dictionary NTUSD (Ku & Chen,

2007) relies on an analysis of reader positive and negative opinions on the linguistic data of news to add the polarity of every word. In a previous study, we attempted to create a Chinese SentiWordNet based on the associations among ChineseWordNet, WordNet 1.6, WordNet 3.0, and SentiWordNet, but there are some restrictions on the use of SentiWordNet. For example, a word in SentiWordNet could have both a positive value of 0.3 and a negative value of 0.1 because the word is used in text with different sentiment orientations. Although this information is correct in general, it causes a problem in how to determine the value to be used in text sentiment orientation analysis. Several methods use both values, and several methods only consider the orientation with a relatively high value. These methods cause considerable estimation errors, and thus they cannot properly achieve the intended results in practical applications.

The purpose of this study is to propose a lexicon-based text sentiment analysis method called the Chinese Text Sentiment Polarity Analyzer (CT-SPA). This method uses two sentiment dictionaries from different sources: a Chinese sentiment dictionary that integrates Chinese WordNet with SentiWordNet, and a sentiment dictionary obtained by training a text corpus labeled with sentiment polarities. In this study, text sentiment polarity is analyzed using these two dictionaries, a mixed-rule approach, and a statistics-based prediction model. The content of the remaining sections is as follows. Section 2 presents a review of related studies. Section 3 describes the method for creating two sentiment dictionaries. Section 4 proposes the algorithm for predicting text sentiment using the aforementioned sentiment dictionaries. In Section 5, the proposed method is confirmed using the test corpus provided by the Topic-Based Chinese Message Polarity Classification task of ACL-SIGHAN 2015. Section 6 includes the conclusions of this study and a description of possible future study topics.

## 2 Related Works

Automatic text sentiment classification has been studied extensively in the last five years. The classification methods are divided into supervised learning that uses text labeled with sentiment values as training data (Moraes, Valiati, & Gavião Neto, 2013), unsupervised learning that requires only unlabeled data (Paltoglou & Thelwall, 2012; Turney, 2002), and semi-supervised learning that

combines a small amount of labeled data with a large pool of unlabeled data (Liu, Chang, & Li, 2013). The supervised learning approach delivers the best performance in classification accuracy, but collecting a large amount of labeled data in every domain is not feasible. Unsupervised learning is readily applicable to every domain, but delivers low classification accuracy. However, it is worth noting that many unsupervised learning methods rely on the characteristics of big data (for example, Paltoglou & Thelwall (2012) used a huge number of posts on social web) to improve clustering accuracy.

With regard to the content for classification, the most frequently analyzed data are posts on social networking sites, such as Twitter, Facebook (Thompson, Poulin, & Bryan, 2014; Thelwall & Buckley, 2013; Martínez-Cámara, Martínez-Valdivia, Urena-Lopez, & Montejo-Ráez, 2014), followed by long reviews, especially movie reviews (Martínez-Cámara, Martínez-Valdivia, Urena-Lopez, & Montejo-Ráez, 2006; Liu et al., 2013). It can be seen that different methods are used for different content, but most methods employ sentiment dictionaries to a certain extent.

Sentiment dictionaries can be divided into three categories according to their sources: those selected from regular dictionaries, where their sentiment polarities and strengths are defined by experts; those where the dictionaries are automatically generated through machine learning; and those where the dictionaries are semi-automatically created using manually built dictionaries as seeds or their extended definitions. There have been few studies on manually built dictionaries because creating such dictionaries is time consuming and usually results in the problem where one word can have different sentiments. However, almost all studies on automatically generated dictionaries require a comparison with manually built dictionaries, and semi-automatically generated dictionaries are considerably dependent on manually built dictionaries. Examples of well-known sentiment dictionaries include SO-CAL (Taboada, Brooke, Tofiloski, Voll, & Stede, 2011), General Inquirer (Stone, Dunphy, & Smith, 1966), and ANEW (Bradley & Lang, 1999).

Most automatic dictionary generation methods use a semantic relationship algorithm that explores the semantic relationship between two words in a large amount of text and analyzes the sentiment polarities of the words based on this relationship. For example, Turney (2002) created a dictionary that consists of adjectives and adverbs using the PMI-IR algorithm and text from the

search engine AltaVista. Kilgarriff (2007) built a Google-PMI dictionary with a similar method.

Semi-automatic extension methods have been used in building most sentiment dictionaries, such as SentiWordNet (Baccianella, Esuli, & Sebastiani, 2010), SenticNet (Cambria, Speer, Havasi, & Hussain, 2010), WordNet-Affect (Strapparava, & Valitutti, 2004), and Chinese NTUSD (Ku & Chen, 2007). The most typical process is the procedure proposed by Whitelaw, Garg, & Argamon (2005). First, seed words with polarities or a small dictionary are used. Second, synonym resources (such as WordNet, HowNet, Chinese Thesaurus, and others) and extension algorithms are used to extend a small amount of sentiment data with labeled words to other words. Third, correct words are obtained through manual detection or screening. In several methods, the third step is performed using rule-based screening (for example, only retaining a few word classes), rather than manual screening.

### 3 Chinese SentiWordNet

We built a Chinese sentiment dictionary (CSWN) based on the relationship among four dictionaries, Academia Sinica Bilingual Ontological WordNet (BOW), WordNet1.6, WordNet 3.0, and SentiWordNet. BOW is a bilingual dictionary that corresponds to WordNet Version 1.6. SentiWordNet is an extended sentiment dictionary built on the WordNet Version 3.0 lexical database. Because there are considerable differences among different versions of WordNet and the same word does not correspond to another in different versions, we established a method for associating different versions of the same word in different WordNet versions based on several rules. For a Chinese word in BOW, its corresponding English word in SentiWordNet can be found through this association. For every Chinese word, its sentiment value can be obtained according to the given sentiment polarity and strength (hereinafter referred to as “sentiment value”) of its corresponding English word in SentiWordNet, where the sentiment value of every word consists of a pair of numbers that represent positive and negative sentiment strength. It is especially worth noting that several words may have both positive and negative values because they may have different sentiment polarities in different context.

Although this method can be used to establish the sentiment values for a considerable number of Chinese words, BOW does not contain a large number of such words, and the sentiment values

still have not been established for numerous Chinese words. To increase the number of Chinese words with sentiment values, the sentiment labels for the English words in E-HowNet are used, and the sentiment value of every English word in E-HowNet is assigned to its corresponding Chinese word. Meanwhile, the sentiment value of every Chinese word with sentiment value is assigned to other Chinese words without sentiment value in the synonym set through the synonym labels in E-HowNet.

The data set of Chinese words with sentiment values obtained by the aforementioned method is called Chinese SentiWordNet (CSWN). Because sometimes there might be errors in the sentiment values obtained by the aforementioned method, NTUSD is used to correct all possible errors. NTUSD is a sentiment dictionary with high labeling accuracy, but all Chinese words in the dictionary are only labeled with sentiment polarity without sentiment strength. Therefore, we use the following rules to correct the sentiment values of the Chinese words obtained previously.

Assuming a word has a positive polarity in NTUSD:

1) If both the positive and negative strengths in CSWN are greater than zero, but the positive strength is greater than the negative strength, the negative strength is adjusted to zero;

2) If both the positive and negative strengths are greater than zero, and the positive strength is equal to the negative strength, the positive strength is set to 0.125 and the negative strength is set to zero.

3) If both the positive and negative strengths are equal to zero, the positive strength is set to 0.25, and the negative strength is set to zero.

4) If both the positive and negative strengths are greater than zero, but the negative strength is greater than the positive strength, the negative strength is adjusted to zero.

5) If the negative strength is greater than zero and the positive strength is equal to zero, the positive strength is set to the average positive strength of all words with unadjusted sentiment values, and the negative strength is set to zero.

If the word has a negative polarity in NTUSD, its polarity is corrected by using rules contrary to those mentioned above.

#### 4 Data-driven sentiment dictionary

A common method for building a sentiment dictionary is to use documents with sentimental labels. A basic prerequisite for such method is as follows: if a word appears more frequently in positive documents than negative or neutral documents, this word is prone to convey a positive sentiment, and vice versa. Therefore, we define three parameters for a corpus, All-Pos, All-Neu, and All-Neg, which represent the numbers of positive, neutral, and negative documents in a corpus, respectively. In addition, we define three parameters for a word in a corpus, Pos, Neu, and Neg, which represent the numbers of positive, neutral, and negative documents that contain the word, respectively.

Based on these six parameters, the frequency of each word occurring in different labeled documents can be calculated. Another three parameters, PosSS, NeuSS, and NegSS can be given by formula (1)-(3)

$$\begin{aligned} \text{PosSS} &= \text{Pos}/\text{All-Pos} & (1) \\ \text{NeuSS} &= \text{Neu}/\text{All-Neu} & (2) \\ \text{NegSS} &= \text{Neg}/\text{All-Neg} & (3) \end{aligned}$$

The sentiment value of a word can be determined according to the aforementioned parameters. Because the words that appear in a corpus are not necessarily contained in CSWN, the following rules are used to establish their sentiment values:

1) If a word only appears in positive and neutral documents, the positive sentiment strength is set to the  $y$  value given by formula (4), and the negative strength is set to zero.

$$y = \log(\text{PosSS}/\text{NeuSS}) * \text{Pos}/(\text{Pos} + \text{Neu}) * \alpha \quad (4)$$

where  $\alpha$  is the strength adjustment parameter. For example, for the corpus used in the experiments for this study, if a word appears in NTUSD,  $\alpha$  is set to 0.3343; otherwise,  $\alpha$  is set to 0.2343.

2) If a word only appears in negative and neutral documents, the negative sentiment strength is set to the  $y$  value given by formula (5), and the positive strength is set to zero.

$$y = \log(\text{Neg}/\text{NeuSS}) * \text{Neg}/(\text{Neg} + \text{Neu}) * \alpha \quad (5)$$

where  $\alpha$  is set similarly to the previous rule.

3) If a word only appears in positive and negative documents, the sentiment value is given by formula (6).

$$y = \log(\text{Pos}/\text{Neg}) \quad (6)$$

If  $y$  is greater than zero, the positive strength of the word is set to  $y$  and its negative strength is set to zero; if  $y$  is less than zero, the negative strength of the word is set to  $y$  and the positive strength is set to zero; and if  $y$  is equal to zero, both the positive and negative strengths are set to zero.

4) If a word appears in documents with various labels, and its PosSS value is greater than its NegSS value, the positive sentiment strength is set to the  $y$  value given by formula (7), and the negative strength is set to zero. If its PosSS value is less than the NegSS value, the negative sentiment strength is set to the  $y$  value given by formula (8), where  $\alpha$  is the strength adjustment parameter, and the positive strength is set to zero. For example, for the corpus used in the experiments for this study, if a word appears in NTUSD,  $\alpha$  is set to 0.7; otherwise,  $\alpha$  is set to 0.2343.

$$\begin{aligned} y &= \log(\text{PosSS}/\text{NegSS}) * \text{Pos}/(\text{Pos} + \text{Neg} + \text{Neu}) * \alpha & (7) \\ y &= -\log(\text{PosSS}/\text{NegSS}) * \text{Neg}/(\text{Pos} + \text{Neg} + \text{Neu}) * \alpha & (8) \end{aligned}$$

For the words contained in CSWN, the following corrective rules are used to suit the characteristics of the linguistic data. Assume that the sentimental score of a word in CSWN is  $G$ :

1) If a word only appears in neutral documents in a corpus, and its  $G$  value is greater than zero, the  $y$  value given by formula (9) is calculated. If the  $y$  value is greater than zero, the positive strength of the word is adjusted to  $y$ ; otherwise, the positive strength is adjusted to zero. The negative strength is set to zero regardless of the  $y$  value. On the other hand, if the  $G$  value of the word is less than zero, the  $y$  value given by formula (9) is calculated. If the  $y$  value is greater than zero, the negative strength of the word is adjusted to  $y$ ; otherwise, the negative strength is adjusted to zero. The positive strength is set to zero regardless of the  $y$  value.

$$y = (1 - \log(\text{Neu} * \omega)) * |G| \quad (9)$$

2) If a word only appears in positive documents in a corpus, and the number of positive documents is greater than the value of parameter  $\delta$ , the positive strength is set to the  $y$  value given by formula (10), and its negative strength is set to zero.

$$y = |(1 - \log(\text{Pos} * \beta)) * G| \quad (10)$$

3) If a word only appears in negative documents in a corpus, and the number of negative documents is greater than the value of parameter  $\delta$ , but the  $G$  value of the word in CSWN is greater than zero, the negative strength of the word is set to the  $y$  value given by formula (10), and the positive strength is set to zero.

$$y = |(1 - \log(\text{Neg} * \beta)) * G| \quad (11)$$

The values of the aforementioned parameters  $\delta$ ,  $\omega$ , and  $\beta$  are related to the number of documents in a corpus, and increase with the increasing number of documents. For example, for the corpus used in the experiments for this study,  $\delta$ ,  $\beta$ , and  $\omega$  are set to 3, 3.3, and 1, respectively. In addition, if the  $y$  value given by any of the formulas (4) to (11) is greater than one, all these parameters are set to one.

## 5 Text sentiment classification method

In the method proposed herein, the difference between the positive and negative strengths for every word in the text is defined as the sentiment score of the word. A positive sentiment score means positive polarity and vice versa. The sum of the sentiment scores for all words is the sentiment score for the text. If the sentiment score for the text is a positive value greater than a threshold value, the sentiment orientation of the text is positive, and vice versa. If the sentiment score does not exceed the threshold value, the text is considered neutral. However, because sentiment words express different sentiment strengths or even opposite sentiments in different word classes and syntactic structures, the following five correction rules are used to calculate the sentiment value for the text.

First, the sentiment value shall be adjusted according to the weight of the word class. A Chinese word may appear in the text as different word classes. Several Chinese word classes impose slight or even no effect on the sentiment value of the text. Therefore, if a word appears as such word classes, its sentiment score shall be adjusted by multiplying it by a weight to obtain the new sentiment value. For example, for words of classes  $N_f$  and  $Neu$ , the weight is set to zero. The weight value can be obtained from the corpus training experiments.

Second, a weighting calculation shall be performed for words that collocate with degree adverbs. Degree adverbs may strengthen or weaken the sentiments of words. For example, “very

happy” expresses stronger sentiment strength than “happy.” Therefore, we select words from the word class  $D_{fa}$  in E-HowNet and manually screen and define the weights of degree adverbs. If a degree adverb precedes a sentiment word in a sentence, the sentiment score of the word shall be multiplied by the weight of the degree adverb to obtain the new sentiment score for this word.

Third, the sentiment scores of words in interrogative sentences and rhetorical questions shall be corrected. The sentiment of an interrogative sentence or rhetorical question is normally contrary to the sentiment score obtained. For example, in the sentence “everyone has tried their best for this. How can you still accuse who shall be blamed for his or her fault?” The “accuse” and “fault” in this sentence express negative sentiment, but their use in this interrogative sentence reverses the entire sentence to positive sentiment. Therefore, the sentiment score of interrogative sentences and rhetorical questions shall be multiplied by -1.

Fourth, the sentiment value for any word that collocates with a negative word shall be reversed to its opposite. When a negative word precedes a word, its overall sentiment polarity is normally contrary to the word. For example, the polarity of “not happy” is contrary to the polarity of “happy.” Therefore, the sentiment score for a word that occurs after negative words shall be multiplied by -1.

Fifth, the sentiment value for any transition sentence shall be corrected. When a transition sentence pair with “but,” “nevertheless,” or “although” appears in the text, the real sentiment of the sentence pair is expressed in the sentence after, rather than before, the transition. For example, in the sentence “this way of doing things is undesirable, but the result is surprisingly good,” obviously the sentiment of the entire sentence pair is identical to that of the latter sentence, but contrary to that of the former sentence. Therefore, when calculating the sentiment score of the text, the sentiment score generated by the sentence before the transition of a transition sentence pair is not considered.

## 6 Experiments

The Topic-Based Chinese Message Polarity Classification task of SIGHAN-8 (hereinafter referred to as SIGHAN-8) provided a corpus labeled with sentiment polarities for training. This training corpus consists of short messages classified into five different topics collected from various social networking sites. SIGHAN-8 also provided a test

corpus from the same source as that for the training corpus, but includes 20 different topics. The numbers of positive, neutral, and negative documents in the two corpora are listed in Table 1. In this study, this training corpus is used to train the proposed prediction model according to the method mentioned in Section 4, and the sentiment polarities of the text in the test corpus are tested with this model.

	Positive	Negative	Neutral
Training corpus	394	538	3,973
Test corpus	1,152	3,639	14,678

Table 1 Number of documents with different polarities in the corpus provided by SIGHAN-8

Table 2 shows the predicted numbers of positive, neutral, and negative documents obtained by the proposed method. According to the SIGHAN evaluation, the prediction results of the proposed method for the test corpus are expressed by three performance indicators, recall, precision, and F1-measure, and all the three values are 0.62.

	Positive	Negative	Neutral
Predicted number of documents	993	5,054	13,422

Table 2 Number of documents with different polarities in the test corpus predicted by the proposed method

## 7 Discussions and future works

The experiment results show that the proposed method can predict the sentiment polarity of certain text, but results in incorrect predictions for other text. We analyzed the causes for prediction errors and made three conclusions.

First, all the test data used are short messages, with each document containing only a limited number of words. This means that whether the judgment about the sentiment value for every word is right or wrong affects the final result. The CSWN dictionary established in this study contains a large number of Chinese words, but numerous words still have not been included, such as specialized terms and unknown words. The sentiment values of these words are inputted manually, and thus developing an automatic labeling method for such words is a very important task.

Second, several prediction errors are caused by the fact that the sentiment values of the words are

highly correlated to the domain of the text. Several words have strong sentiment connotations in some domains, but are neutral in other domains. Several words even exhibit a different or opposite sentiment value in the same domain under different context. Therefore, the predictive ability of a model might be improved by developing methods for solving the problem of the ambiguous sentiment value of words.

Third, there are considerable numbers of English corpora labeled with sentiment values, but very few Chinese corpora are available. Because of insufficient training corpus, combined with the short length of the document, the proposed method barely predicted the correct sentiment values for words not included in the sentiment dictionary, and many documents could not be predicted correctly. How to rapidly develop a corpus with sentiment labels through semi-automatic methods is one of the focused areas for future studies.

## Acknowledgements

This work is supported in part by the Ministry of Science and Technology, Taiwan, R.O.C. under the Grants MOST 103-2511-S-151-001. It is also partially supported by the “Aim for the Top University Project” and “Center of Learning Technology for Chinese” of National Taiwan Normal University (NTNU), sponsored by the Ministry of Education, Taiwan, R.O.C. and the “International Research-Intensive Center of Excellence Program” of NTNU and Ministry of Science and Technology, Taiwan, R.O.C. under Grant MOST 104-2911-I-003-301.

## References

- Baccianella, S., Esuli, A., & Sebastiani, F. 2010. SentiWordNet 3.0: An Enhanced Lexical Resource for Sentiment Analysis and Opinion Mining. *Proceedings of LREC*, 10:2200-2204.
- Bradley, M. M., & Lang, P. J. 1999. Affective norms for English words (ANEW): Instruction manual and affective ratings. *Technical Report C-1*, 1-45. The Center for Research in Psychophysiology, University of Florida.
- Cambria, E., Speer, R., Havasi, C., & Hussain, A. 2010. SenticNet: A Publicly Available Semantic Resource for Opinion Mining. *Proceedings of AAAI Fall Symposium: Commonsense Knowledge*, 10:14-18.
- Feldman, R. 2013. Techniques and applications for sentiment analysis. *Communications of the ACM*, 56(4):82-89

- Kennedy, A., & Inkpen, D. 2006. Sentiment classification of movie reviews using contextual valence shifters. *Computational Intelligence*, 22(2):110-125.
- Ku, L. W., & Chen, H. H. 2007. Mining opinions from the Web: Beyond relevance retrieval. *Journal of the American Society for Information Science and Technology*, 58(12):1838-1850.
- Liu, C. L., Chang, T. H., & Li, H. H. 2013. Clustering Documents with Labeled and Unlabeled Documents using Fuzzy Semi-Kmeans. *Fuzzy Sets and Systems*, 221(16): 48–64.
- Martínez-Cámara, E., Martínez-Valdivia, M. T., Urena-Lopez, L. A., & Montejo-Ráez, A. R. 2014. Sentiment analysis in twitter. *Natural Language Engineering*, 20(1):1-28.
- Moraes, R., Valiati, J. F., & Gavião Neto, W. P. 2013. Document-level sentiment classification: An empirical comparison between SVM and ANN. *Expert Systems with Applications*, 40(2), 621-633.
- Paltoglou, G., & Thelwall, M. 2012. Twitter, MySpace, Digg: Unsupervised sentiment analysis in social media. *ACM Transactions on Intelligent Systems and Technology*, 3(4):66.
- Stone, P. J., Dunphy, D. C., & Smith, M. S. 1966. The General Inquirer: A Computer Approach to Content Analysis.
- Strapparava, C., & Valitutti, A. 2004. WordNet Affect: an Affective Extension of WordNet. *Proceedings of LREC*, 4:1083-1086.
- Taboada, M., Brooke, J., Tofiloski, M., Voll, K., & Stede, M. 2011. Lexicon-based methods for sentiment analysis. *Computational linguistics*, 37(2):267-307.
- Thelwall, M., & Buckley, K. 2013. Topic-based sentiment analysis for the social web: The role of mood and issue-related words. *Journal of the American Society for Information Science and Technology*, 64(8):1608-1617.
- Thompson, P., Poulin, C., & Bryan, C. J. 2014. Predicting military and veteran suicide risk: Cultural aspects. *Proceedings of ACL 2014*, 1-6.
- Whitelaw, C., Garg, N., & Argamon, S. 2005. Using appraisal groups for sentiment analysis. *Proceedings of the 14th ACM international conference on Information and knowledge management*, 625-631.