

Construction of a Blog Emotion Corpus for Chinese Emotional Expression Analysis

Changqin Quan

Faculty of Engineering
University of Tokushima
2-1 Minamijosanjima Tokushima Japan
quan-c@is.tokushima-u.ac.jp

Fuji Ren

Faculty of Engineering
University of Tokushima
2-1 Minamijosanjima Tokushima Japan
ren@is.tokushima-u.ac.jp

Abstract

There is plenty of evidence that emotion analysis has many valuable applications. In this study a blog emotion corpus is constructed for Chinese emotional expression analysis. This corpus contains manual annotation of eight emotional categories (expect, joy, love, surprise, anxiety, sorrow, angry and hate), emotion intensity, emotion holder/target, emotional word/phrase, degree word, negative word, conjunction, rhetoric, punctuation and other linguistic expressions that indicate emotion. Annotation agreement analyses for emotion classes and emotional words and phrases are described. Then, using this corpus, we explore emotion expressions in Chinese and present the analyses on them.

1 Introduction

Textual emotion analysis is becoming increasingly important due to augmented communication via computer mediated communication (CMC) internet sources such as weblogs, email, websites, forums, and chat rooms. Especially, blogspace consists of millions of users who maintain an online diary, containing frequently-updated views and personal remarks about a range of issues.

Despite the increased focus on analysis of web content, there has been limited emotion analysis of web contents, with the majority of studies focusing on sentiment analysis or opinion mining. Classifying the mood of a single text is a hard task; state-of-the-art methods in text classification achieve only modest performance in this domain (Mishne, 2005). In this area, some of the hardest problems involve acquiring basic resources. Corpora are fundamental both for developing sound conceptual analyses and for training these emotion-oriented systems at different levels: to recognize emotions, to express appropriate

emotions, to anticipate emotions, and other emotion processing applications.

In this study we propose a relatively fine-grained annotation scheme, annotating emotion in text at three levels: document, paragraph, and sentence. We select eight emotion classes (expect, joy, love, surprise, anxiety, sorrow, angry and hate) for this annotation, and explore various linguistic expressions that indicate emotion in Chinese. The annotation scheme has been employed in the manual annotation of a corpus containing 1,487 documents, with 11,255 paragraphs, 35,096 sentences, and 878,164 Chinese words. Then, using this corpus, we explore and present data analyses on emotions, involving emotion states, accompanying emotions, transfer emotions, independent emotions in texts.

The remainder of this paper is organized as follows. Section 2 describes the emotion corpus annotation scheme. Section 3 presents the inter-annotator agreement study. Section 4 describes the analysis of emotion expressions. Section 5 presents a review of current emotion corpora for textual emotion analysis. Section 6 concludes this study with closing remarks and future directions.

2 Blog Emotion Corpus Annotation Scheme

Weblogs are an increasingly popular mode of communication in the ever changing online world. Writing suits the recording of facts and the communication of ideas, and their textual basis makes them equally suitable for recording emotions and opinions. So, we select blogs as object and data source for this emotion corpus annotation.

2.1 Emotional Expression in Text

An important starting point in constructing this corpus is to represent emotion in text. One of the biggest questions in affect recognition is, "What

2.3 Sentence Level Annotation

Sentences are basic units for emotional expression. The central aim of sentence level annotation is to explore as much linguistic expressions for reflecting emotion in Chinese as possible.

a) Emotion holder/target

In the task of opinion analysis, the problem of opinion holder identification has also been studied, (Bethard, Steven et al., 2004; Choi, Cardie, et al., 2005; Kim and Hovy, 2005). As for emotion holder/target identification, little research has been conducted, but we believe it is important for exploring emotional expression and emotion analysis. Emotion holder is the one who holds the emotions, and an emotion target is the object of an emotion holder. For instance,

(1) 我喜欢这个老师。(English: I like this teacher.) In sentence (1),我“(English: I)” is the emotion holder, and 这个老师“(English: this teacher.)” is the emotion target.

In this corpus, not every sentence is annotated with emotion holder or emotion target, and emotion holder or emotion target may not appear in pairs in one sentence. If one sentence has more than one emotion holders or emotion targets, they are all annotated.

b) Emotional words and phrases

Lexicon-based methods have received a lot of attention in opinion analysis task. There are many lexical resources for these tasks. For emotion analysis tasks, the function of words is equally fundamental. In most sentimental lexicons, the words usually bear direct emotions or opinions, such as happy or sad, good or bad. However, there are a lot of sentences can evoke emotions without direct emotional words, for example,

(2) 春天在孩子们的眼里、在孩子们的心里。(English: Spring is in children’s eyes, and in their hearts.)

In sentence (2), we may feel joy, love or expect delivered by the writer. Indeed, as (Ortony, Andrew, et al., 1987) indicates, besides words directly referring to emotional states and for which an appropriate lexicon would help, there are words that act only as an indirect reference to emotions depending on the context.

In this annotation scheme, direct emotional words and indirect emotional words in a sentence are all annotated. In sentence (2), 春天“(English: spring)”, 孩子们“(English: children)” are labeled. An emotional word or phrase

is represented as a vector to record its intensities of the eight basic emotional classes. For instance, the vector for the word 喜欢“(English: like)” $\vec{w} = (0.0, 0.3, 0.9, 0.0, 0.0, 0.0, 0.0, 0.0)$ indicates the emotions of weak joy and strong love. For indirect emotional words, we annotate their emotion vectors according to their contexts, for example, the possible emotion vector for the word “春天(English: spring)” $\vec{w} = (0.1, 0.3, 0.3, 0.0, 0.0, 0.0, 0.0, 0.0)$ indicates the emotions of weak expect, joy and love. (The emotions and intensity values may be different because of different annotators).

Emotional phrases are combination of words, such as Chinese proverbs, like “世上无难事, 只要肯攀登(English: Where there is a will, there is a way)”. For an emotional phrase, the positions of its first and character in a sentence are labeled, and also for emotional words if there are Chinese word segmentation mistakes.

The statistics show that 84.9% of all emotional words have one emotion, and 14.7% have two emotions, only 0.4% have three or four emotions, but they are indispensable for expressing complex feelings in use of language.

Table 2 shows the numbers of emotional words with different POS (part-of-speech) tags. The set of POS includes 35 classes; Table 2 lists the top five classes.

POS	Num. of words (have repeat)
Verb	37,572
Noun	21,308
Adj.	20,265
Adv.	4,223
Gerund	2,789

Table 2: Emotional words with different POS

As shown in Table 2, verbs, nouns, adjectives and adverbs are strong markers of emotion in Chinese.

c) Degree words, negative words, conjunctions

Degree words are associated with the intensities of emotions. In Chinese, degree words appear with high frequency. In this corpus, there are 1,039 different degree words annotated, the total occurring number of them is 16,713, in which, 8,294 degree words modify emotional words or phrases directly. Degree words and the modifying

contents are all labeled.

Negative words can be placed almost everywhere in a sentence to change the meaning, also to change the emotions. Negative words are frequently used in Chinese. The statistical data shows that there are 645 different negative words annotated in this corpus, the total occurring number of them is 13,750, in which, 3,668 negative words modify emotional words or phrases directly.

Besides, conjunctions may change the emotion of a sentence. for example,

(3) 尽管我们喜欢这个老师, 但她已经离开了我们。(Jin guan wo men xi huan zhe ge lao shi, dan ta yi jing li kai le wo men; English: Although we like this teacher, she has leaved.)

Sentence (3) uses the conjunctions “尽管...但...(jin guan...dan..., English: although)” express emotions of love and sorrow. There are 297 different conjunctions annotated in this corpus. Conjunctions and the modifying contents are all labeled. If conjunctions appear in pairs in a sentence, the position of pairing words for each conjunction are also labeled. For the above sentence (3), conjunctions are annotated as follows (Figure. 2).

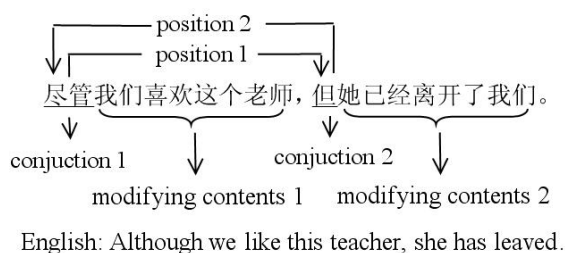


Figure 2: An example of conjunctions annotation

Figure 3 shows the growth curve of word number with document number from 300 to 1487. As can be seen from Figure 3, the increase numbers of emotional words/phrases slow down with the increase in the number of documents, and the numbers of negative words, degree words and conjunctions basically remained stable. We can look forward to containing most of common emotional expressions in weblogs articles.

d) Rhetorics, punctuations

Chinese rhetoric has been well studied from the view of linguistics and literature. We select nine common rhetoric categories to annotate: 比喻(English: metaphor), 夸张(exaggeration), 拟人(English: personification), 对偶(English: antithesis or parallel), 排

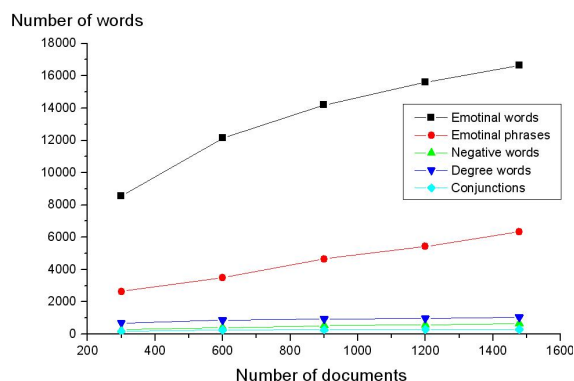


Figure 3: Growth curve of word number

比(English: parallelism sentence), 设问(English: rhetorical question with answer), 反问(English: rhetorical question), 重复(English: repeat), 讽刺(English: irony). Especially, 讽刺(English: irony) is a way as to imply the contrary of what one says, if a sentence is annotated with irony, its emotions maybe totally different from the emotions of words that it contains. We annotate rhetoric category and the corresponding emotion category.

Punctuation is the use of standard marks and signs in writing to separate words into sentences, clauses, and phrases in order to clarify meaning. Some punctuation marks can express emotions, for example, an exclamation mark (!) or a question mark (?) is used at the end of a sentence to show strong emotion. Balog, Mishne, et al. (2006) suggests that people relied on four strategies including punctuation to express happiness versus sadness. Punctuation effect is also shown in (Leshed and Kaye, 2006) to extend to emoticon placement in website text messages. We annotate punctuation with emotion and the corresponding emotion category.

e) Emotion objective/subjective, emotion polarity

Distinguishing a sentence between factual and subjective information could support for many natural language processing applications. Objective and subjective in our annotation scheme is to distinguish a sentence between writer’s emotion and non-writer’s emotion.

There is a positive side or a negative side on emotion. We call this an emotional polarity. Emotion polarity of a sentence is determined by integrating its emotions. A sentence without emotion is annotated with neutral.

An annotation tool is developed for this corpus

annotation. Input files are text files with Chinese segmentation and part-of-speech tags, the annotated output files are XML files.

3 Annotation Agreement analysis

Emotion annotation is a hard task because the nature of emotion is inherently ambiguous. In the process of annotation, annotators were encouraged to follow their “first intuition”. To measure agreement on various aspects of the annotation scheme, three annotators independently annotated 26 documents with a total of 270 paragraphs, 701 sentences.

3.1 Agreement for Emotion Classes

The kappa coefficient of agreement is a statistic adopted by the Computational Linguistics community as a standard measure for this purpose (Carletta, 1996). We measured two agreements for emotion classes’ annotation:

Agreement (a): the agreement on classification of containing or not containing some emotions. In this case, we distinguish two classes: emotion intensity $e_i \in \{0.0\}$ or $e_i \in \{0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1.0\}$;

Agreement (b): the agreement on classification of emotion intensity. In this case, we distinguish four classes: $e_i \in \{0.0\}$ or $e_i \in \{0.1, 0.2, 0.3, 0.4\}$, or $e_i \in \{0.5, 0.6, 0.7\}$, or $e_i \in \{0.8, 0.9, 1.0\}$.

Table 3 shows Agreement (a) and (b) measure on documents, paragraphs and sentences.

	Agreement(a)	Agreement(b)
documents	0.831	0.695
paragraphs	0.705	0.616
sentences	0.756	0.648
Average	0.764	0.653

Table 3: Agreement on emotion classes

As shown in Table 3, it is easier for annotators to agree at the coarser levels of granularity, and it is more difficult to agree on the level of emotion intensity.

3.2 Agreement for Emotional Words and Phrases

Measuring agreement for emotional words and phrases is to verify that annotators agree on which expressions should be marked. To illustrate this agreement problem, consider the emotional words

and phrases identified by annotators **a** and **b**. This sentence was preprocessed by Chinese segmentation and tagged with part-of-speech.

(4) 今晨/t, /w 当/d 我/v 沐浴/n 着/u 阳光/n 前往/v 会场/n 时/Ng, /w 脑中/v 突然/ad 浮现/v 出/v 多年/m 不用/d 的/u 优美/a 词语/n, /w 那/r 就/d 是/v : /w 秋高气爽/n、/w 金光/n 璀璨/z。(English: This morning, when I walked to the meeting with sunshine, some wonderful words that have not been used for many years crossed my mind, which are “the autumn sky is clear, the air is crisp” and “shinning with gold color”)

a: 阳光, 优美, 秋高气爽, 金光, 璀璨;

b: 沐浴, 阳光, 优美, 秋高气爽, 璀璨;

In sentence (4), the two annotators agree that “阳光”, “优美”, “秋高气爽” and “璀璨” can express emotion. In addition, annotator **a** marked the word “金光”, and annotator **b** marked the word “沐浴”.

In this task, there is no guarantee that the annotators will identify the same set of expressions. Thus, to measure agreement we want to consider how much intersection there is between the sets of expressions identified by the annotators. We use the following voting-agreement metric to measure agreement in identifying emotional words and phrases.

Metric voting-agreement is defined as follows. Let A, B and C be the sets of expressions annotated by annotators a, b and c respectively. The expert coder is the set of expressions that agreed by at least two annotators, see Equation 2.

$$voting_agreement = Avg\left(\frac{count(t_i = e_j)}{count((t_i))}\right) \quad (2)$$

In which, $t_i \in T, e_j \in E, T = A \cup B \cup C, E = (A \cap B) \cup (A \cap C) \cup (B \cap C)$.

The agreement for emotional words and phrases is 0.785.

4 Emotional Expressions Analysis

4.1 Emotion State

“Emotion state in text” is the state of combined emotions in a text unit. An emotion state is represented by 8 binary digits, each digit corresponding to a basic emotion class respectively. As an example, a document emotion state “01100000” is the state of combined emotions by joy and love.

The statistics show that, in this corpus, there are 149 different emotion states in all of the 1,487

documents, 165 different emotion states in all of the 11,255 paragraphs, and 143 different emotion states in all of the 35,096 sentences respectively. That indicates the set of emotion state in texts is relatively small. We also found some basic emotions tend to combine together, such as {expect, joy, love}, {anxiety, sorrow}, {angry, hate}. However, some emotions have small or scarce possibility appear together, such as joy and hate, surprise and angry.

4.2 Accompanying Emotions

In an emotion state, some basic emotions are mixed together. When an emotion e_j arise, emotion $e_i (i \neq j)$ arise with accompany, then, e_i is an accompanying emotion of e_j . To compute the probability of the accompanying emotion given an emotion e_j , we count the cooccurrence of e_i and e_j in a text unit (a document, a paragraph, or a sentence).

$$P(e_i|e_j) = \frac{\text{count}(e_i \text{ with } e_j)}{\text{count}(e_j)} \quad (3)$$

Table 4 shows the accompanying emotions with the highest probabilities for the eight basic emotions in documents, paragraphs and sentences.

Emotions	Docs	Paras	Sens
Expect	Love	Love	Love
Joy	Love	Love	Love
Love	Joy	Joy	Joy
Surprise	Anxiety	Love	Love
Anxiety	Sorrow	Sorrow	Sorrow
Sorrow	Anxiety	Anxiety	Anxiety
Angry	Anxiety	Hate	Hate
Hate	Anxiety	Sorrow	Angry

Table 4: Accompanying emotions

In Table 4, the accompanying emotions has shown a high uniformity in the 3 units of text.

4.3 Transfer Emotions

When emotion change from one emotion class to another one, we call this emotion transfer. Using the context relation of paragraphs and sentences, we compute the probability $P(e_i \rightarrow e_j)$.

$$P(e_i \rightarrow e_j) = \frac{\text{count}(e_t = e_i, e_{t+1} = e_j)}{\text{count}(e_t = e_i)} \quad (4)$$

In which, e_t is an emotion class in paragraph t (or sentence t), and e_{t+1} is another emotion class

in paragraph $t + 1$ (or sentence $t + 1$). Table 4 shows the transfer emotions with the highest probabilities for the eight basic emotions in paragraphs and sentences.

Emotions	Paras	Sens
Expect	Love	Expect
Joy	Love	Love
Love	Love	Love
Surprise	Love	Love
Anxiety	Anxiety	Anxiety
Sorrow	Sorrow	Sorrow
Angry	Anxiety	Angry
Hate	Hate	Hate

Table 5: Transfer emotions

Similar to this, we can compute the probability of emotion state transfer $P(e_state_i \rightarrow e_state_j)$. This may help a lot for emotion prediction, for example, if we know the current emotion state is “00000110” (sorrow an angry), we can estimate the probability of this emotion state to another emotion state “00000001” (hate).

4.4 Independent Emotion

When a text unit (a document, a paragraph, or a sentence) only contains one emotion class, this emotion class is an independent emotion. The statistics show that emotion of love has high independence, however, joy, surprise and angry has relative low independence. The intuition is love can be the only topic emotion in a text unit, but emotions of joy, surprise and anxiety more incline to combine with other emotions.

5 Related work

Previous approaches to textual emotion analysis have employed some different corpora. Mishne (2005) experimented mood classification in blog posts on a corpus of 815,494 blog posts from Livejournal (<http://www.livejournal.com>), a free weblog service with a large community. Livejournal also used as data source for finding happiness (Mihalcea and Liu, 2006), capturing global mood levels (Mishne and De Rijke, 2006), classifying mood (Jung, Park, et al., 2006; Jung, Choi, et al., 2007), discovering mood irregularities (Balog, Mishne, et al., 2006), recognizing affect (Leshed and Kaye, 2006). A similar emotion corpus in Chinese is Yahoo!’s Chinese news (<http://tw.news.yahoo.com>), which is used

for Chinese emotion classification of news readers (Lin, Yang, et al., 2007) and emotion lexicon building (Yang, Lin, et al., 2007). Tokuhima (2008) also use web as data resources to obtain a huge collection of emotion-provoking event instances for Japanese emotion classification. More and more weblogs have added mood column to record blog users' moods when they read or write a blog.

Two merits let them well accepted as emotion corpora: a large number of weblogs contained and moods annotated by blog users. However, there is a great inconsistency on emotion categories given by different websites. Livejournal gives a pre-defined list of 132 common moods, while Yahoo!'s Chinese news provides readers 8 emotion categories. Too many mood classes may confuse users, and Mishne (2005) also pointed out one obvious drawback of the mood "annotation" in this corpora is that they are not provided in a consistent manner; the blog writers differ greatly from each other, and their definitions of moods differ accordingly. In addition, some words are not fitted to be taken as emotion classes, such as "useful" in Yahoo!'s emotion categories. These corpora may be helpful for analyzing the global moods on a full text, but the inconsistent emotion categories is a problem, and no more labeled information can be exploited from them.

The emotion analysis on sentence level may also be important for more detailed emotion analysis systems. Alm, Roth, et al. (2005) explore the text-based emotion prediction problem; they annotated a corpus of 22 Grimms' tales on sentence level with eight emotion categories (angry, disgusted, fearful, happy, sad, positively surprised, negatively surprised), contain 1580 sentences. Neviarouskaya, Prendinger et al. (2007) address the tasks of recognition and interpretation of affect communicated through text messaging. They collected 160 sentences labeled with one of nine emotions categories (anger, disgust, fear, guilt, interest, joy, sadness, shame, and surprise) from a corpus of online diary-like blog posts and a corresponding intensity value. Aman and Szpakowicz (2007) classify emotional and non-emotional sentences based on a knowledge-based approach. They used a corpus with tags of emotion category, emotion intensity and the words/phrases that indicate emotion in text. An emotion corpus for Japanese was built for rec-

ognizing emotions and emotion estimation (Ren, 2009; Matsumoto, 2006). However, the sizes of these corpora seem not enough for large scale textual emotion analysis, a lot of linguistic features are not reflected from them. A more fine-grained opinion and emotion corpus is the MPQA Corpus (Wiebe, Wilson, et al., 2005), which contains 535 news articles (10,000-sentence) from a wide variety of news sources, manually annotated at the sentential and subsentential level for opinions and other private states. But emotion categories are not included in it.

To the best of our knowledge, at present, there's no relatively large corpora annotated with detailed linguistic expressions for emotion in Chinese, and we believe that such corpora would support the development and evaluation of emotion analysis systems.

6 Conclusions and Future Work

In this study we proposed an emotional expression space model. Emotion of a document, a paragraph, a sentence, or even a word is represented by an emotional vector. Based on this model, we described a relatively fine-grained annotation scheme and annotated emotion in text. We also gave the inter-annotator agreement study on annotation. Then, we explore the emotional expressions in texts.

This annotated dataset can be obtained for free with license ¹. Eleven annotators made efforts on it spanning a period of ten months (They are Ph.D and M.S. candidates specialize in Natural Language Processing and Emotion Analysis). To ensure the quality of this dataset, each document was performed a three pass annotation, in which the first pass is annotated by one annotator and then the second and the third verification pass were performed by other two annotators. The process of this corpus annotation is easy to make mistakes because of a lot of information should be annotated. The verification pass is to check the annotation mistakes (such as the start and end positions of emotional phrases in sentences), but not to change the choices of emotion classes or emotional words which had been annotated by other annotators.

Using this corpus, we will make a more extensive study of textual emotion analysis in Chinese,

¹<http://a1-www.is.tokushima-u.ac.jp/member/ren/Ren-CECps1.0/Ren-CECps1.0.html>

for example, the influence of degree words, negative words, or other elements on emotional expression; the difference between subjective emotion and objective emotion; emotion transfer tracking. More applications also will be explored, such as emotional summarization, emotional question answering; emotional topic discovering. At the same time, new research problems will arise, for examples, how to acquiring more emotional words and to generate their emotional vectors automatically; how to generate emotional vectors for sentences, paragraphs and documents with known emotional elements in them? There is need to immerge further into these problems.

Acknowledgments

We are grateful to our annotators: Huana Li, Ye Wu, Lei Chen, Yu Zhang, Ji Li, Ziliang Du, Yuanlu Fu, Rong Mu, Yan Sun, Cheng Wang, Yunong Wu, and other participants and supporters. We are also grateful to Dr. Suzuki and Dr. Matsumoto for the helpful advice. This research has been partially supported by Ministry of Education, Science, Sports and Culture, Grant-in-Aid for Challenging Exploratory Research, 21650030.

References

- Alena Neviarouskaya, Helmut Prendinger, Mitsuru Ishizuka. 2007. Textual Affect Sensing for Social and Expressive Online Communication. *Proceedings of the 2nd international conference on Affective Computing and Intelligent Interaction*, pp. 218-229.
- Bethard, Steven, Hong Yu, Ashley Thornton, Vasileios Hatzivassiloglou, and Dan Jurafsky. 2004. Automatic Extraction of Opinion Propositions and their Holders. *AAAI Spring Symposium on Exploring Attitude and Affect in Text: Theories and Applications*, pp. 133 - 136.
- Changhua Yang, Kevin Hsin-Yih Lin, Hsin-Hsi Chen. 2007. Building Emotion Lexicon from Weblog Corpora. *Proceedings of the ACL 2007 Demo and Poster Sessions*, pp. 133 - 136.
- Cecilia Ovesdotter Alm, Dan Roth, Richard Sproat. 2005. Emotions from text: Machine learning for text-based emotion prediction. *Proceedings of Human Language Technology Conference and Conference on Empirical Methods in Natural Language Processing*, pp. 579-586, Vancouver, British Columbia, Canada.
- Fuji Ren. 2009. Affective Information Processing and Recognizing Human Emotion. *Electronic Notes in Theoretical Computer Science*, 225: 39-50.
- Gilad Mishne. 2005. Emotions from text: Machine learning for text-based emotion prediction. *Proceedings of Style2005 in SIGIR'05*, pp. 15-19.
- Gilad Mishne and Maarten de Rijke. 2006. Capturing global mood levels using blog posts. *AAAI 2006 Spring Symposium on Computational Approaches to Analysing Weblogs*, pp.145-152.
- Gilly Leshed and Joseph Kaye. 2006. Understanding how bloggers feel: recognizing affect in blog posts. *Conference on Human Factors in Computing Systems CHI '06 extended abstracts on Human factors in computing systems*, pp. 1019- 1024.
- Janyce Wiebe, Theresa Wilson, Claire Cardie. 2005. Annotating expressions of opinions and emotions in language. *Language Resources and Evaluation*. 39: 164 - 210.
- Jean Carletta. 1996. Assessing Agreement on Classification Tasks: The Kappa Statistic. *Computational Linguistics*. 22(2):249-254.
- Kazuyuki Matsumoto, Fuji Ren, Shingo Kuroiwa. 2006. Emotion Estimation System based on Emotion Occurrence Sentence Pattern. *Computational Intelligence, Lecture Notes in Computer Sciences*, pp.902-911.
- Kevin Hsin-Yih Lin, Changhua Yang, Hsin-Hsi Chen. 2007. What emotions do news articles trigger in their readers? *Annual ACM Conference on Research and Development in Information Retrieval*, pp. 733- 734.
- Krisztian Balog, Gilad Mishne, Maarten de Rijke. 2006. Why are they excited? identifying and explaining spikes in blog mood levels. *Proceedings 11th Meeting of the European Chapter of the Association for Computational Linguistics*, pp. 207-210.
- Ortony, Andrew, Gerald L. Clore, and Mark A. Foss. 1987. The referential structure of the affective lexicon. *Cognitive Science*, 11: 341-364.
- Rada Mihalcea and Hugo Liu. 2006. A corpus-based approach to finding happiness. *Proceedings of the AAAI Spring Symposium on Computational*, pp. 139-144.
- Rosalind Picard. 1997. *Affective Computing*. The MIT Press, MA, USA.
- Ryoko Tokuhsa, Kentaro. Inui, and Yuji. Matsumoto. 2008. Emotion Classification Using Massive Examples Extracted from the Web. *Proceedings of COLING 2008*, pp. 881-888.
- Saima Aman and Stan Szpakowicz. 2007. Identifying Expressions of Emotion in Text. *Lecture Notes in Computer Science*. 4629: 196-205.
- Soo-Min Kim and Eduard Hovy. 2005. Identifying Opinion Holders for Question Answering in Opinion Texts. *Proceedings of AAAI-05 Workshop on Question Answering in Restricted Domains*, pp. 1367-1373.

- Yejin Choi, Claire Cardie, Ellen Riloff, and Siddharth Patwardhan. 2005. Identifying Sources of Opinions with Conditional Random Fields and Extraction Patterns. *Proceedings of HLT/EMNLP-05*, pp. 355-362.
- Yuchul Jung, Hogun Park, Sung Hyon Myaeng. 2006. A Hybrid Mood Classification Approach for Blog Text. *Lecture Notes in Computer Science*, pp.1099-1103.
- Yuchul Jung, Yoonjung Choi, Sung-Hyon Myaeng. 2007. Determining Mood for a Blog by Combining Multiple Sources of Evidence. *IEEE/WIC/ACM International Conference on Web Intelligence*, pp. 271-274.