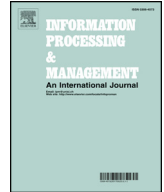




ELSEVIER

Contents lists available at ScienceDirect

# Information Processing and Management

journal homepage: [www.elsevier.com/locate/infoproman](http://www.elsevier.com/locate/infoproman)

## Tracking and recognizing emotions in short text messages from online chatting services

Chih-Hao Chen, Wei-Po Lee\*, Jhieh-Yuan Hwang

Department of Information Management, National Sun Yat-sen University, Kaohsiung, Taiwan

### ARTICLE INFO

#### Keywords:

Online chatting  
Plurk  
Emotion analysis  
Text context  
Valence-arousal space

### ABSTRACT

To automate the process of emotion recognition, in this study, we develop a computational approach for continuously tracking and analyzing users' emotions while chatting online. Our work has several unique features: it provides relative probabilities of possible emotions for a word, constructs a distribution for each chatting message accordingly, performs a clustering procedure for the message distribution, and aggregates the emotions of continuous chatting sentences to draw the conclusion. To evaluate the proposed approach, we conducted experiments in two phases. The first phase was to evaluate the effectiveness of the proposed computational approach in analyzing the chatting sentences. The participants were asked to focus on tagging emotions toward each sentence for a pre-designed dialogue. The second phase involves a real-time chatting between two online users. The participants were asked to choose topics and freely chat with each other. The messages were analyzed, and the results were provided to the users for their evaluations. The results show that our approach is both effective and efficient in tracking the emotions of chatting users. Additional analyses and further discussions were carried out to further evaluate the quantitative experimental results. All the findings confirmed the usefulness and feasibility of the presented approach.

### 1. Introduction

Sentiment analysis refers to determining the evaluative nature of a text. Traditional studies on sentiment analysis usually aim to detect the polarity (positive, negative or neutral) of a given text (Pang & Lee, 2015; Piryani, 2017; Ravi & Ravi, 2015). Other kinds of studies exist that recognize multiple differentiated affective manifestations in texts, such as joy, anger, and fear, and that explore sentences with compound emotions as well as the uses of language in expressing complex feelings (Mohammad, 2015; Sintsova & Pu, 2016). Sentiment analysis systems have been applied to different application domains and on both long and short texts depending on the application. Given the popularity of social networking services, people like to share information and interact with others on different service platforms. Consequently, a large number of social networking texts are produced, and tremendous interest and effort is being invested into applying sentiment analysis to these texts (e.g., (da Silva, Coletta, & Hruschka, 2016; Giachanou & Crestani, 2016; Martínez-Cámara, Martín-Valdivia, Ureñalópez, & Montejoárez, 2012; Meo & Sulis, 2017; Saif, He, Fernández, & Alani, 2016)). Automatically recognizing a user's affective states can enhance the quality of interactions, thereby making an application more usable and effective. Thus, this work presents an approach to tracking and analyzing emotions in a specific text type—online chatting messages, which require instant and continuous processing for chatting users.

Based on their writing styles, online or offline texts can be presented in both formal and informal ways (Mohammad, 2016).

\* Corresponding author

E-mail address: [wplee@mail.nsysu.edu.tw](mailto:wplee@mail.nsysu.edu.tw) (W.-P. Lee).

<https://doi.org/10.1016/j.ipm.2018.05.008>

Received 18 November 2017; Received in revised form 18 May 2018; Accepted 20 May 2018

0306-4573/© 2018 Elsevier Ltd. All rights reserved.

Formal texts, such as documents, lectures or speeches, are written for more serious or official occasions; they are less personal. In contrast, informal texts appear in more relaxed situations and are often written in more familiar or intimate tones, similar to personal letters or conversations. People choose distinct kinds of vocabularies and grammars for different types of texts, and several sentiment analysis methods have been employed to address them, respectively (these methods are reviewed in the next section). Informal texts often appear in online communication in various forms, ranging from blogs, forums, and user feedback to chatting. Over the past decade in particular, there has been substantial growth in the use of microblogs such as Twitter, Plurk and instant messaging; consequently, a prodigious amount of informal text has been produced. Such communications can occur in real-time or instant communication (i.e., where the communication takes place in real time, and the communicating users respond to each other instantly) or in non-instant message postings.

Online chatting refers to any interactive communication on the internet that offers real-time transmission of text messages between a sender and a receiver. In contrast to other text-based online communication forms, chat messages are generally short and informal, that is advantageous for making quick responses. However, short and informal chatting messages pose new challenges to sentiment analysis, and they increase the difficulty of analysis, because they are not well-structured, they are limited in length (usually one or only a few sentences), and they often reflect spoken language with grammatical and spelling mistakes, internet slang, and shortened forms of words. In addition, authors of these messages may use many oblique techniques for covert communication, for example, the use of emoticons to mimic human expressions that complement the text message. Moreover, the semantics of words used within a chat session may be different from their common meaning.

In addition to the above difficulties, it is notable that chatting is a type of continuous activity among the participants; consequently, an analysis should not be concerned solely with the textual and linguistic aspects (Balahur, Mihalcea, & Montoyo, 2014; Garas, Garcia, Skowron, & Schweitzer, 2012; Giachanou & Crestani, 2016). The emotion of a chatting message may relate to its context (e.g., location or previous messages) (Mohammad et al., 2016; Saif et al., 2016). Although various textual affective features and methods have been studied extensively, most previous studies have not created an efficient approach for using contextual emotional clues for emotion recognition in text. Sentiment classification algorithms may be able to effectively classify user text at the page or paragraph level but may struggle at the sentence level (Appel, Chiclana, Carter, & Fujita, 2016; Kiritchenko, Zhu, & Mohammad, 2014). Therefore, the methods often used for sentiment analysis, such as lexical or machine learning methods, cannot be applied directly to the online chatting domain. More comprehensive considerations are required.

In this study, we developed a system for continuously tracking and analyzing users' emotions while chatting online. Our work has several unique features: it deals with short chatting messages instantly, and accounts for multiple emotions while simultaneously considering contextual information. More specifically, our approach provides relative probabilities of possible emotions for a word, constructs a distribution for each chatting message accordingly, performs a clustering procedure for the message distribution, and aggregates the emotions of continuous chatting sentences to draw the conclusion about the emotions conveyed by a chat sentence. To evaluate the proposed approach, we conducted a series of experiments in which a group of participants were invited to engage in the online chatting process and to perform evaluations afterward. The feedback from our participants was very positive. These results show that the proposed computational mechanism is effective and efficient in tracking the emotions of chatting users. Additional analyses and discussions were performed to further evaluate the quantitative experimental results. All the results confirmed the usefulness and feasibility of the proposed approach.

## 2. Background and related works

As mentioned above, various methods have been proposed for sentiment analysis. These can be categorized into different types, depending on the text level (feature, sentence, or document), text length (long or short), and text type (formal or informal) to be investigated (Mohammad, 2016; Thelwall, Buckley, Paltoglou, Cai, & Kappas, 2010). The feasibility and performance of these methods may vary from level to level. Therefore, these methods need to be characterized to ensure the best performance for the target text. This section briefly introduces and analyzes three issues most relevant to our work, including textual emotion recognition, emotion space model, and Chinese text segmentation.

### 2.1. Textual emotion recognition

Although sentiment and emotion have nuanced differences, for simplicity's sake, and without losing generality, the following review does not specifically differentiate between them, but instead regards emotion as a multiple-affect extension of sentiment and focuses on the computational techniques for different types of text processing. In general, the dominant techniques used in sentiment analysis or textual emotion recognition fall into two broad categories: language processing-based methods and machine learning-based methods. The first category, the language processing-based approach, can be divided further into two types: lexical-based (or named keyword-spotting) methods and semantic-based (or named concept-based) methods. Lexical-based methods rely on the presence of obvious sentimental words. In this type of approach, a dictionary is prepared to store the polarity values of lexicons. The polarity of a text is calculated by aggregating the polarity scores of each word in the text, and then classified based on the total score. Because the polarity of the text depends on the score given to each lexicon, it is important to determine the weighting factors of the lexicons. Therefore, much research has been devoted to discovering which lexical information is most efficient, for example, (Deng, Luo, & Yu, 2014; Cambria, Olsher, & Rajagopal, 2018; Khan, Qamar, & Bashir, 2017).

The sentiment measurement in keyword-spotting methods can be extended for emotion measurement. This method detects certain affect words (from a dictionary) such as "happy", "angry", "sad" and "depressed" in statements and assigns an emotion category for

Download English Version:

<https://daneshyari.com/en/article/10998010>

Download Persian Version:

<https://daneshyari.com/article/10998010>

[Daneshyari.com](https://daneshyari.com)