

# Linguistic Template Extraction for Recognizing Reader-Emotion

Yung-Chun Chang<sup>\*,+</sup>, Chun-Han Chu<sup>\*</sup> Chien Chin Chen<sup>+</sup>, and

Wen-Lian Hsu<sup>\*</sup>

## Abstract

Previous studies on emotion classification mainly focus on the emotional state of the writer. By contrast, our research emphasizes emotion detection from the readers' perspective. The classification of documents into reader-emotion categories can be applied in several ways, and one of the applications is to retain only the documents that trigger desired emotions to enable users to retrieve documents that contain relevant contents and at the same time instill proper emotions. However, current information retrieval (IR) systems lack the ability to discern emotions within texts, and the detection of reader's emotion has yet to achieve a comparable performance. Moreover, previous machine learning-based approaches generally use statistical models that are not in a human-readable form. Thereby, it is difficult to pinpoint the reason for recognition failures and understand the types of emotions that the articles inspired on their readers. In this paper, we propose a flexible emotion template-based approach (TBA) for reader-emotion detection that simulates such process in a human perceptive manner. TBA is a highly automated process that incorporates various knowledge sources to learn an emotion template from raw text that characterize an emotion and are comprehensible for humans. Generated templates are adopted to predict reader's emotion through an alignment-based matching algorithm that allows an emotion template to be partially matched through a statistical scoring scheme. Experimental results demonstrate that our approach can effectively detect reader's emotions by exploiting the syntactic structures and semantic associations in the context, while

---

\* Institute of Information Science, Academia Sinica, Taipei, Taiwan

E-mail: {changyc, johannchu, hsu}@iis.sinica.edu.tw

<sup>+</sup> Department of Information Management, National Taiwan University, Taipei, Taiwan

E-mail: patonchen@ntu.edu.tw

outperforming currently well-known statistical text classification methods and the state-of-the-art reader-emotion detection method.

**Keywords:** Reader-Emotion Detection, Emotion Template, Template-based Approach, Text Classification, Sentiment Analysis.

## 1. Introduction

With the rapid growth of the Internet, the web has become a powerful medium for disseminating information. People can easily share information of daily experiences and their opinion anytime and anywhere on the social media, such as blogs, Twitter and Facebook. Therefore, sentiment analysis studies have gained increasing interest in recent years with academia and business corporations attempting to analyse and predict public trends by mining opinions that are subjective statements and reflect people's sentiments or perceptions about topics (Pang *et al.*, 2002). Moreover, human feelings can be quickly collected through emotion detection (Quan & Ren, 2009; Das & Bandyopadhyay, 2009). While previous researches on emotions mainly focused on detecting the emotions that the authors of the documents were expressing, it is worthy of noting that the reader emotions, in some aspects, differ from that of the authors and may be even more complex (Lin *et al.*, 2008; Tang & Chen, 2012). Regarding a news article for instance, while a journalist objectively reports up-going oil price and does not express his or her emotion in the text, a reader could yield angry or negative emotions. On the other hand, an infamous politician's blog entry describing his miserable day may not cause the opposing readers to feel the same way. While the author of an article may directly express his/her emotions through sentiment words within the text, a reader's emotion possesses a more complex nature, as even general words can evoke different types of reader's emotions depending on the reader's personal experience and knowledge (Lin *et al.*, 2007).

Instead of detecting writer's emotion, which has been already investigated extensively in previous studies (Zhang & Liu, 2010; Mukherjee & Liu, 2010; Si *et al.*, 2013), this paper aims to uncover the emotions of readers triggered by the document. Such research holds great potential for novel applications. For instance, an enterprise that possesses the business intelligence that is capable of identifying the emotional effect that a document inflicts on its readers can provide services to retain only the documents that evokes the desired emotions, enabling users to retrieve documents with relevant contents and meanwhile being instilled the proper emotions. As a result, users benefit by obtaining opportunities and advantages in the competitive market through a more efficient and quick manner. However, current information retrieval systems lack the ability to discern emotion within texts, and reader-emotion detection has yet to achieve comparable performance (Lin *et al.*, 2007). Machine learning-based approaches are widely used for sentiment analysis and emotion detection related researches. These approaches can usually generate accurate classifiers that assign a category label for

each document with much lower labour cost. Nevertheless, the statistical models used by these classifiers are generally not in a human-readable form. Thus, it is difficult to pinpoint the reason for recognition failures and understand what exact emotion of the reader is triggered.

In light of this rationale, we proposed a flexible template-based approach (TBA) for reader-emotion detection that simulates such process in human perception. TBA is a highly automated process that integrates various types of knowledge to generate discriminative linguistic patterns for document representation. These patterns can be acknowledged as the essential knowledge for humans to understand different kinds of emotions. Furthermore, TBA recognizes reader's emotions of documents using an alignment-based algorithm that allows an emotion template to be partially matched through a statistical scoring scheme. Our experiments demonstrate that TBA can achieve a higher performance than other well-known text categorization methods and the state-of-the-art reader-emotion detection method.

The remainder of this paper is organized as follows. The next section contains a review of related works on reader-emotion detection approaches. We introduce the proposed emotion template-based approach for reader-emotion detection in Section 3, and its evaluation is described in Section 4. Finally, we provide some concluding remarks and potential future avenues of research in Section 5.

## 2. Related Work

Textual articles are one of the most common ways for persons to convey their feelings. Identifying essential factors that affect emotion transition is important for human language understanding. With the rapid growth of computer-mediated communication applications, such as social websites and micro-blogs, the research on emotion classification has been attracting more and more attention recently from enterprises toward business intelligence (Chen *et al.*, 2010; Purver & Battersby, 2012). In general, a single text may possess two types of emotions: writer-emotion and reader-emotion. The research of writer-emotion investigates the emotion expressed by the writer when writing the text. Pang *et al.* (2002) designed an algorithm to classify movie reviews into positive and negative emotions. Mishne (2005), and Yang and Chen (2006) used emoticons as tags to train SVM (Cortes & Vapnik, 1995) classifiers at the document or sentence level, respectively. In their studies, emoticons were taken as mood or emotion tags, and textual keywords were considered as features. Wu *et al.* (2006) proposed a sentence level emotion recognition method using dialogs as their corpus, in which "Happy", "Unhappy", or "Neutral" was assigned to each sentence as its emotion category. Yang *et al.* (2006) adopted Thayer's model (1989) to classify music emotions. Each music segment can be classified into four classes of moods. As for sentiment analysis research, Read (2005) used emoticons in newsgroup articles to extract relevant instances for training polarity classifiers.

Nevertheless, the research of reader-emotion concerns the emotions expressed by a reader after reading the text. As the writer and readers may view the same text from different perspectives, they do not always share the same emotion. Since the recent increase in the popularity of Internet, certain news websites, such as Yahoo! Kimo News, incorporate the Web 2.0 technologies that allow readers to express their emotions toward news articles. Classifying emotions from the readers' point of view is a challenging task, and research on this topic is relatively sparse as compared to those considering the writers' perspective. While writer-emotion classification has been extensively studied, there are only a few studies on reader-emotion classification. Lin *et al.* (2007) first described the task of reader-emotion classification on news articles and classified Yahoo! News articles into 8 emotion classes (e.g. happy, angry, or depressing) from the readers' perspectives. They combined bigrams, words, metadata and word emotion categories to train a classifier for determining the reader-emotions toward news. Yang *et al.* (2009) automatically annotated reader-emotions on a writer-emotion corpus with a reader-emotion classifier, and studied the interactions between writers and readers with the writer-reader-emotion corpus.

Our approach differs from existing reader-emotion detection approaches in a number of aspects. First, we proposed an emotion template-based approach that mimics the perceptual behaviour of humans in understanding. Second, the generated emotion templates are human readable, and can be represented as the domain knowledge required for detecting reader-emotion. Therefore, it is helpful in elucidating how articles trigger certain types of emotions in their readers in a more comprehensive manner. Finally, in addition to syntactic features, TBA further considers the surrounding context and semantic associations to efficiently recognize reader-emotions.

### 3. Learning Reader-emotion Template from Raw Text

In this paper, we present a template-based approach (TBA) for detecting the reader-emotion of documents. We model reader-emotion detection as a classification problem, and define the reader-emotion detection task as the following. Let  $W = \{w_1, w_2, \dots, w_k\}$  be a set of words,  $D = \{d_1, d_2, \dots, d_m\}$  be a set of documents, and  $E = \{e_1, e_2, \dots, e_n\}$  be a set of reader-emotions. Each document  $d$  is a set of words such that  $d \subseteq W$ . The goal of this task is to decide the most appropriate reader-emotion  $e_i$  for a document  $d_j$ , although one or more emotions can be associated with a document. Our proposed method is different in that we take advantage of multiple knowledge sources, and implement an algorithm to automatically generate templates that represent discriminative patterns in documents. Our system, using the proposed method, mainly consists of three components: Crucial Element Labelling (CEL), Emotion Template Generation (ETG), and Emotion Template Matching (ETM), as shown in Figure 1. The CEL first uses prior knowledge to mark the semantic classes of words in the corpus. Then the ETG

collects frequently co-occurring elements, and generates templates for each emotion. These templates are stored in the emotion-dependent knowledge base to provide domain-specific knowledge for our emotion detection. During the detection process, an article is first labelled by the CEL as well. Subsequently, the ETM applies an alignment-based algorithm that utilizes our knowledge base to calculate the similarity between each emotion and the article to determine the main emotion of this article. Details of these components will be disclosed in the following sections.

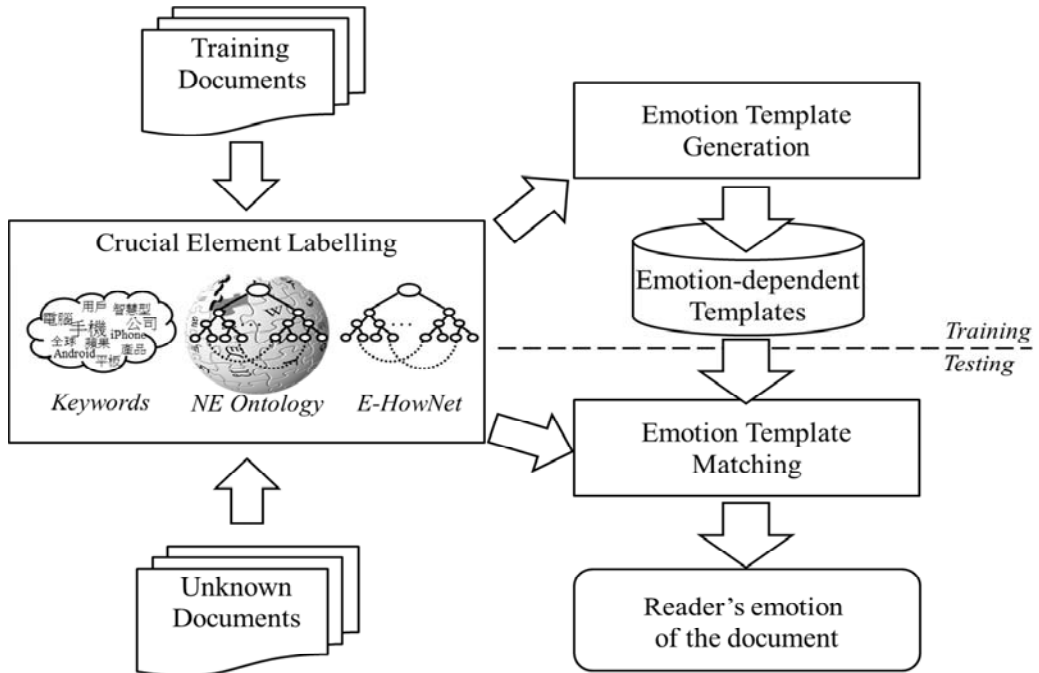


Figure 1. Architecture of our system.

### 3.1 Crucial Element Labelling (CEL)

TBA attempts to simulate the human perception of an emotion through the recognition of crucial elements. In this work, we capture crucial elements within documents by adopting a three-layer labelling approach that utilizes various knowledge sources, such as lexical dictionaries and Wikipedia, to induce template elements. First of all, since keywords within a reader-emotion are often considered as important information, we used the log likelihood ratio (LLR) (Manning & Schütze, 1999), an effective feature selection method, to learn a set of reader-emotion specific keywords. Given a training dataset, LLR employs Equation (1) to calculate the likelihood of the assumption that the occurrence of a word  $w$  in reader-emotion  $E$  is not random. In (1),  $e_i$  denotes the specific reader's emotion in the training dataset;  $N(e_i)$  and  $N(-e_i)$  are the numbers of on-emotion and off-emotion documents, respectively.  $N(w^{\wedge}e_i)$ ,

denoted as  $k$ , is the number of on-emotion documents containing  $w$ ; the number of off-emotion documents containing  $w$ ,  $N(w|\neg e_i)$ , is denoted as  $l$ . Altogether, the formula expresses the ratio of two likelihood functions. To simplify the formula, we also define  $m$  as the number of on-emotion documents with no word  $w$  (that is,  $m = N(e_i) - k$ ), and  $n$  as that of off-emotion documents ( $n = N(\neg e_i) - l$ ). The probabilities  $p(w)$ ,  $p(w|e_i)$ , and  $p(w|\neg e_i)$  are estimated using maximum likelihood estimation. A word with a large LLR value is closely associated with the reader-emotion. We rank the words in the training dataset based on their LLR values and select the top 200 to compile an emotion keyword list.

$$LLR(w, e_i) = 2 \log \left[ \frac{p(w|e_i)^k (1 - p(w|e_i))^m p(w|\neg e_i)^l (1 - p(w|\neg e_i))^n}{p(w)^{k+l} (1 - p(w))^{m+n}} \right] \quad (1)$$

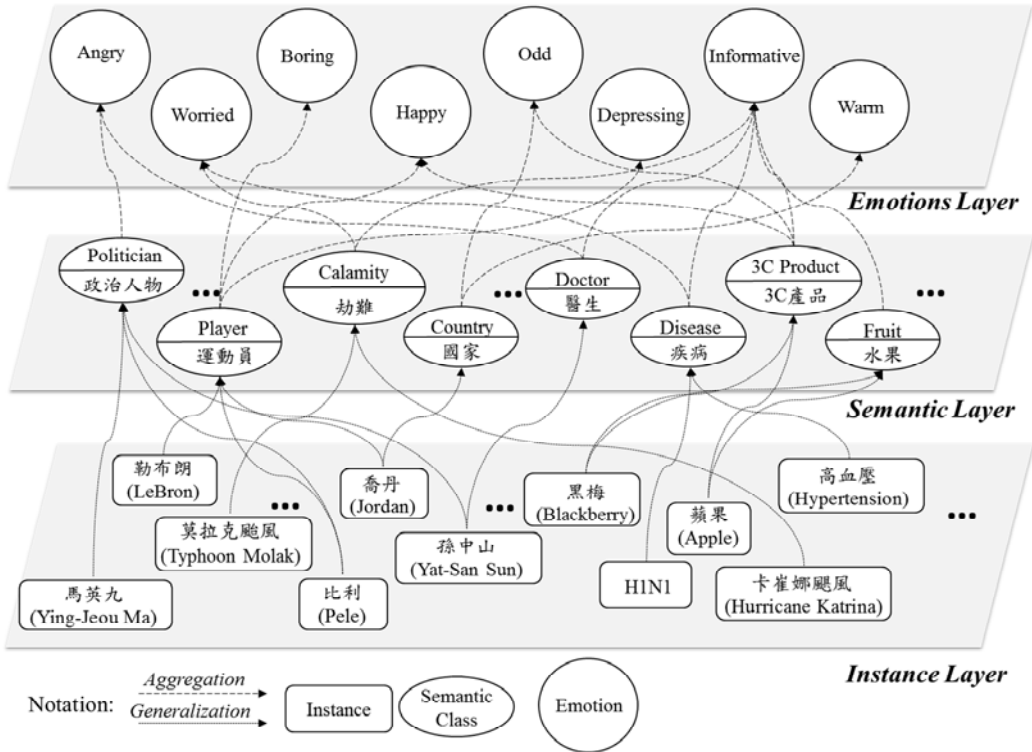


Figure 2. Architecture of named entity ontology.

Next, we aim to recognize named entities (NEs) from text to facilitate document comprehension and improve the performance of identifying topics (Bashaddadh & Mohd, 2011). Our labelling algorithm uses a string matching scheme to single out the keywords (if they exist in the sequences); therefore, segmentation is not required in the preprocess. However, exact-matching NEs may overlook many template elements since it omits semantic

context. To remedy this problem, this paper adopts a novel structure to construct the NE ontology (NEO) for labelling crucial elements based on the levels of organization mentioned in (Lee *et al.* 2005) and (Wang *et al.* 2010). Figure 2 depicts the architecture of the NE ontology, which includes an emotion layer, a semantic layer, and an instance layer. There are eight types of emotions in the emotion layer, namely “Angry”, “Worried”, “Boring”, “Happy”, “Odd”, “Depressing”, “Informative” and “Warm”. Moreover, each semantic class in the semantic layer denotes a general semantic meaning of named entities that can be aggregated from many emotions, including “政治人物 (Politician)”, “疾病 (Disease)” and others. The instance layer represents 6323 named entities extracted from documents across eight emotions by the Stanford NER. In order to minimize the labour cost of instance generalization, we utilize Wikipedia to semi-automatically label NEs with their semantic classes as a way of generalization. Only NE labels for persons, places and organizations are taken into consideration, and Wikipedia’s category tags are used to label NEs recognized by the Stanford NER<sup>1</sup>. We select the category tag to which the most *topic paths* are associated as the main semantic label for NEs in documents. A topic path is the classification hierarchy of a certain category; it can be considered as the traversal from general categories to more specific ones. A category name with more associated topic paths is considered to be more suitable to represent a NE for its appropriate scope of semantic coverage. For example, a query “歐巴馬 (Obama)” to the Wikipedia would return a page titled “巴拉克·歐巴馬 (Barack Obama)”. Within this page, there are a number of category tags such as “民主黨 (Democratic Party)” and “美國總統 (Presidents of the United States)”. These two category tags contain three and seven topic paths, respectively. Suppose “美國總統(Presidents of the United States)” is the one with more topic paths, our system will label “巴拉克·歐巴馬 (Barack Obama)” with the tag “[美國總統 (Presidents of the United States)]”. Domain experts further annotated each named entity by their corresponding semantic classes for the purpose of generalization if the NE term not included in Wikipedia to its category tags. Each instance in the instance layer can connect to multiple semantic classes according to the generalized relations. For example, named entity “喬丹 (Jordan)” can be generalized to “國家 (country)” and “人名 (people)”. In this manner, we can transform plain NEs to a more general class and increase the coverage of each label.

Finally, to incorporate even richer semantic context into our semantic template, we use the Extended HowNet (Chen *et al.*, 2005), which is an extension of the HowNet (Dong *et al.* 2010). Extended HowNet contains a structured representation of knowledge and semantics. It connects approximately 90000 words in the CKIP Chinese Lexical Knowledge Base and HowNet, and includes additional highly frequent words that are specific in Traditional Chinese. It also contains a different formulation of each word to better fit its semantic

---

<sup>1</sup> <http://nlp.stanford.edu/software/CRF-NER.shtml>

representation, as well as the distinct definition of function and content words. A total of four basic semantic classes were applied, namely object, act, attribute, and value. Moreover, in comparison to HowNet, EHowNet possesses a layered definition scheme and complex relationship formulation, and uses simpler concepts to replace schemes as the basic element when defining a more complex concept or relationship. To illustrate the content of the EHowNet, let us take the definition of “血癌 (leukemia)” in EHowNet in Definition 1 as an example. In the most compact sense, leukemia is a type of cancer that originates from disorder of blood (cells); therefore, EHowNet presents the phenomenon and its occurring position for the term in the *Simple Definition*. To further detail its meaning, EHowNet goes an extra mile and explains both “cancer” and “blood” in the *Expanded Definition*. We can see that the EHowNet not only contains semantic representation of a word, but also its relations to other words or entities. This enables us to combine or dissect the meaning of words by using its semantic components. Following the method in (Shihet *al.* 2012), we extracted the main definition of each word as the semantic class label.

### Definition 1:

#### *Simple Definition:*

{癌症|cancer:position={血液|blood}}

#### *Expanded Definition:*

{disease|疾病:position={BodyFluid|體液:telic={transport|送:patient={gas|氣:predication={respire|呼吸:patient={~}}},instrument={~}}},qualification={serious|嚴重}}

We exploit the taxonomies in EHowNet, which include lexical categories, synonyms, and semantic relations between different words or sets of words. With the resources stated, the CEL can transform words in the original documents into their corresponding semantic labels. Our research assigns clause as the unit for semantic labelling. To illustrate the process of CEL, consider the clause  $C_n$  = “希拉蕊順利代表民主黨贏得美國總統選舉 (Hillary Clinton represents the Democratic Party and won the Presidential election in the U.S.)”, as shown in Figure 3. First, “希拉蕊(Hillary Clinton)” is found in the keyword dictionary and tagged. Then NEs like “民主黨 (Democratic Party)” and “總統選舉 (Presidential elections)” are recognized and tagged as “{政黨 (Party)}” and “{總統選舉 (Presidential elections)}”. Subsequently, other terms such as “代表 (represent)” and “贏得 (won)” are labelled with their corresponding E-HowNet senses if they exist. Finally we can obtain the sequence: “[希拉蕊 ]:{代表 }:{政黨 }:{得到 }:{國家 }:{總統選舉 } ([Hillary Clinton]:{represent}:{party}:{got}:{country}:{Presidential elections}”], where domain keywords (represented by square-bracketed slots) are matched verbatim and semantic classes



(represented by curly-bracketed slots) match all words that belong to their specific class. Since different annotation knowledge bases are processed through in a specific-to-general order, the proportion of overlapping annotations are relatively low. In the occasion when there is an overlap between the matched keywords, the longest keyword will be reserved. The labelling process not only effectively prevents errors caused by Chinese word segmentation, but also groups the synonyms altogether by using semantic labels. This enables us to generate distinctive and prominent semantic templates in the next stage.

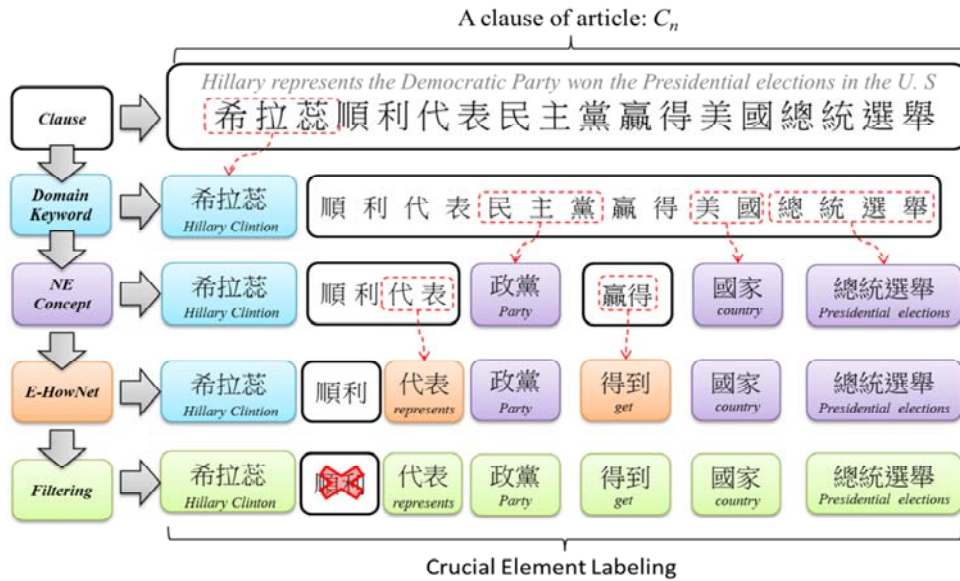


Figure 3. Crucial element labelling process.

### 3.2 Emotion Template Generation

In our framework, a reader's emotion is represented by a set of semantic templates, which are sequences of crucial elements consisting of crucial elements and keywords. The emotion template generation (ETG) process aims at automatically generating  $N$  representative templates from sequences of crucial elements in the documents. These representative (or dominating) templates can be used as background knowledge for each reader's emotion when recognizing documents. More importantly the representative templates can be easily understood by humans. To illustrate, consider the emotion "Happy" and one of the automatically generated semantic templates, "{運動員<sub>player</sub>}: {得到<sub>get</sub>}: [冠軍<sub>championship</sub>]." We can think of various semantically similar sentences that are covered by this semantic template, e.g., "柯里帶領勇士贏得了 NBA 總冠軍賽 (Stephen Curry led the Warriors to win the NBA championship)" or "費德勒擊敗安迪·穆雷獲得溫普敦冠軍 (Roger Federer defeated Andy

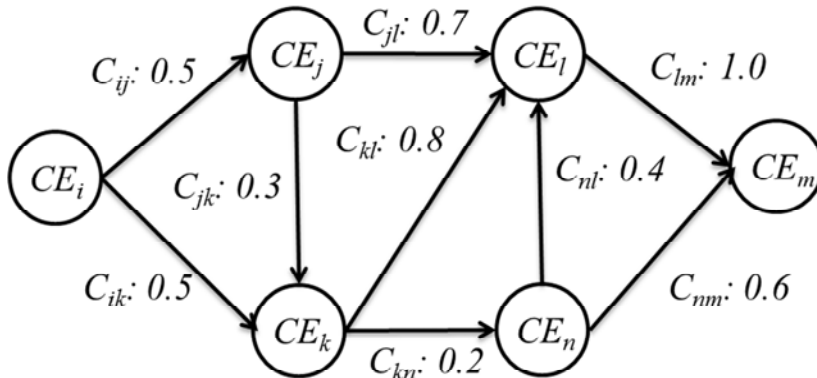
Murray and won the Wimbledon championship)”. Likewise, a similar template “{運動員<sub>player</sub>}: {得到<sub>get</sub>}: [分數<sub>score</sub>]” is capable of representing sentences like “布萊恩在最終賽贏得六十分 (Kobe Bryant won sixty points at his final game.)”. This sort of human-interpretable knowledge cannot be easily obtained by ordinary machine learning models.

The ETG process is described as follows. We formulate reader-emotion template generation as a frequent pattern mining problem. Based on the co-occurrence of crucial elements, we can construct a *crucial element graph* (*CE graph*) to describe the strength of relations between them. Since crucial elements are of an ordered nature, the graph is directed and can be made with association rules. In order to avoid the generation of templates with insufficient length, we empirically set the minimum support of a crucial element as 100 and minimum confidence as 0.5 in our association rules. This is because we observed that the rank-frequency distribution of semantic classes followed Zipf’s law (Manning & Schütze, 1999), so does the normalized frequency of semantic templates. Crucial elements of low frequency usually identify semantic that are irrelevant to the emotion. Hence, for each reader’s emotion, we selected the first frequent crucial elements that accumulated frequencies reached 80% of the total crucial element frequency count in the reader-emotion documents. Thus, an association rule can be represented as (2):

$$\text{confidence}(CE_i \Rightarrow CE_j) = P(CE_j|CE_i) = \frac{\text{support}(CE_i \cup CE_j)}{\text{support}(CE_i)}, \quad (2)$$

where  $\text{support}_{\min}=100$ ,  $\text{confidence}_{\min}=0.5$ .

Figure 4 is the illustration of a CE graph. In this graph, vertices ( $CE_x$ ) represent crucial elements, and edges represent the co-occurrence of two elements,  $CE_i$  and  $CE_j$ , where  $CE_i$  precedes  $CE_j$ . The number on the edge denotes the confidence of two connecting vertices.



Notation:

CE: Crucial Element     $C_{pq}$ : Confidence ( $CE_p \Rightarrow CE_q$ )

**Figure 4.** A CE graph for template generation.

After constructing all CE graphs, we then generate emotion templates by applying the random walk theory (Lovász, 1993) in search of high-frequency and representative elements for each reader's emotion. Let a CE graph  $G$  be defined as  $G=(V,E)$  ( $|V|=p$ ,  $|E|=k$ ), a random walk process consists of a series of random selections on the graph. Every edge  $(CE_n, CE_m)$  has its own weight  $M_{nm}$ , which denotes the probability of a crucial element  $CE_n$ , followed by another element  $CE_m$ . For each element, the sum of weight to all neighboring elements  $N(CE_n)$  is defined as (3), and the whole graph's probability matrix is defined as (4). As a result, a series of a random walk process becomes a Markov Chain. According to (Li *et al.*, 2010), the cover time of a random walk process on a normal graph is  $\forall CE_n, C_{CE_n} \leq 4k^2$ . We can conclude that using random walk to find frequent patterns on CE graphs would help us capture even the low probability combinations and shorten the processing time.

$$\forall CE_n \quad \sum_{m \in N(CE_n)} M_{nm} = 1 \quad (3)$$

$$Pr = \left[ \begin{array}{c} X_{t+1} = CE_m \\ \dots \\ X_0 = CE_i \end{array} \middle| \begin{array}{c} X_t = CE_n \\ X_{t-1} = CE_k, \\ \dots \\ X_0 = CE_i \end{array} \right] = Pr[X_{t+1} = CE_m | X_t = CE_n] = M_{nm} \quad (4)$$

Although the random walk process can help us generate templates from frequent patterns in CE graphs, it can also create some redundancy. Hence, a merging procedure is required to eliminate the redundant results by retaining only the templates with the longest length and highest coverage, and dispose of those that are completely covered by another template. For example, the template [歐巴馬]:{政黨}:{總統選舉} is completely covered by template [歐巴馬]:{代表}:{政黨}:{得到}:{國家}:{總統選舉}. Thus, the former template is removed. Moreover, the reduction of the crucial element space provided by template selection is critical. It allows the execution of more sophisticated text classification algorithms, which lead to improved results. Those algorithms cannot be executed on the original crucial element space because their execution time would be excessively high, making them impractical (Ricardo & Berthier, 2011). Therefore, selecting crucial elements closely associated with an emotion would improve the performance of reader-emotion detection. We use the log likelihood ratio (LLR) again to differentiate crucial elements for each emotion. Given a training dataset comprised of different reader's emotions, the LLR calculates the likelihood of the occurrence of a crucial element in the emotion. A crucial element with a large LLR value is thought to be closely associated with the emotion. Lastly, we rank the crucial elements in the training dataset based on a sum of crucial elements LLR values and retain the top 100 for this reader's emotion.

### 3.3 Emotion Template Matching

We believe the human perception of an emotion is obtained through recognizing important events or semantic contents to rapidly evoke their emotional response. For instance, when an article contains strongly correlated words such as “Japan (country)”, “Earthquake (disaster)” and “Tsunami (disaster)” simultaneously, it is natural to conclude that the article has a much higher chance of eliciting emotions like *depressed* and *worried* rather than *happy* and *warm*. TBA uses an alignment algorithm to measure the similarity of templates, since alignment enables a single template to match multiple semantically similar expressions with appropriate scores. During matching, a document is first labelled with crucial elements. Afterwards, an alignment-based algorithm (Needleman & Wunsch, 1970) is applied to determine to what degree a semantic template fits in a document. For each clause within a given document  $d_j$ , we first label crucial elements  $cs = \{ce_1, \dots, ce_n\}$ , followed by the matching procedure that compares all sequences of crucial element in  $d_j$  to all emotion templates  $ET = \{et_1, \dots, et_j\}$  in each emotion, and calculates the sum of scores for each emotion. The emotion  $e_i$  with the highest sum of scores defined in (5) is considered as the winner.

$$Emotion = \arg \max_{e_i \in E} \sum_{et_n \in ET_{e_i}, cs_m \in CS_{d_j}} \Delta(et_n, cs_m) = \sum_k \sum_l \Delta(et_n \cdot st_k, cs_m \cdot ce_l) \quad (5)$$

$st_k$  and  $ce_l$  represent the  $k^{th}$  slot of  $et_n$  and  $l^{th}$  crucial element of  $cs_m$ , respectively. Scoring of the matched and unmatched components in semantic templates is as follows: If  $et_n \cdot st_k$  and  $cs_m \cdot ce_l$  are identical, we add a matched score ( $MS$ ) obtained from the LLR value of  $ce_l$  if it matches a keyword. Otherwise, the score is determined by multiplying the frequency of the crucial element in emotion  $e_i$  by a normalizing factor  $\lambda = 100$  as in (6). On the contrary, if an element is not matched, the score of *insertion* or *deletion* is calculated. An insertion ( $IS$ ), defined as a label that is present in the article but not in the template, can be accounted for by the inversed entropy of this crucial element (7), which can be thought of as the uniqueness or generality of this label. On the other hand, a deletion ( $DS$ ), representing the label that exists only in the template and not in the article, is computed from the log frequency of this element as (8). Both types give negative scores to the sum. The matching algorithm is then applied to determine the reader-emotion of the article by comparing the sequence of crucial elements  $C = \{c_1, c_2, \dots, c_n\}$  in each clause of an article to every template  $F = \{C_1, C_2, \dots, C_m\}$  in each emotion. An illustration of the matching process of a sequence of semantic classes to an emotion template is shown in Figure 5. A match between the two sequences is given a positive score obtained from the LLR score of the semantic class in an emotion. Finally, the sum of scores of each emotion was computed, and the emotion with the highest score is considered as the winner. Through this method, each individual crucial element label is given a different weight according to its characteristics. Thus, the order of these labels is not the only determining factor in matching. The detailed algorithm is described in Algorithm 1.

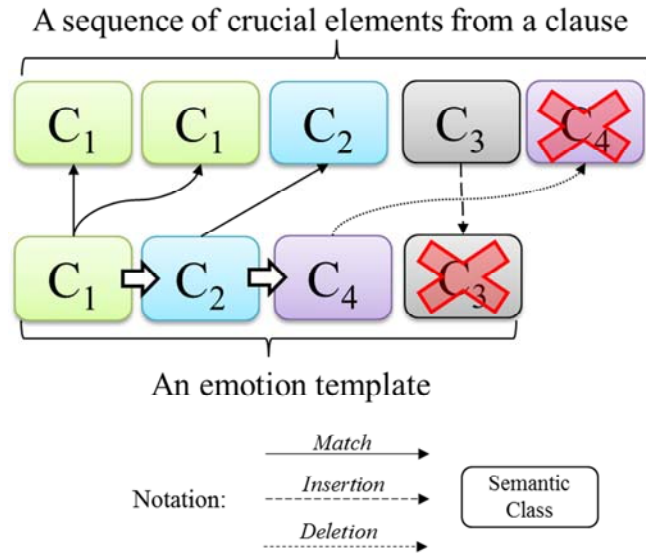


Figure 5. Illustration of an emotion template matching process.

---

### Algorithm 1: Emotion Template Matching

---

**Input:** A emotion template  $et = \{st_1, \dots, st_m\}$ ,  $st$ : slots; A sequence of crucial elements from a clause  $cs = \{ce_1, \dots, ce_n\}$

**Output:** Matching score  $\sigma$  between  $et$  and  $cs$

**BEGIN**

1:  $pos \leftarrow 0$ ;  $\sigma \leftarrow 0$ ;

2: **FOR**  $i = 1$  to  $m$  **DO**

3:  $isMatched \leftarrow false$ ;

4:  $pos \leftarrow$  current matched position in  $CE$ ;

5:   **IF** found  $ce_j$  **EQUAL TO**  $st_i$  after  $pos$  **THEN**

6:      $\sigma \leftarrow \sigma + MatchedScore(ce_j)$ ;

7:      $isMatched \leftarrow true$ ;

8:   **END IF**

9:   **IF**  $isMatched \neq true$  **THEN**

10:      $\sigma \leftarrow \sigma - DeletionScore(st_i)$ ;

11:      $\sigma \leftarrow \sigma - InsertionScore(ce_j)$ ;

12:   **END IF**

13: **END FOR**

14: Output  $\sigma$

**END**

---

$$MS(ce_l) = \begin{cases} LLR_{ce_l}, & \text{if it matches a keyword} \\ \lambda \frac{f_{ce_l}}{\sum_{i=1}^m f_{ce_i}}, & \text{otherwise} \end{cases} \quad (6)$$

$$IS(ce_l) = \frac{1}{-\sum_{i=1}^m P(ce_{l_{e_i}}) \cdot \log_2(P(ce_{l_{e_i}}))} \quad (7)$$

$$DS(ce_l) = \log \frac{f_{ce_l}}{\sum_{i=1}^m f_{ce_i}} \quad (8)$$

## 4. Experiment

### 4.1 Experiment Setup

To the best of our knowledge, there is no publicly available corpus for reader-emotion detection. Therefore, we compiled a data corpus for performance evaluation, as shown in Table 1. The data corpus contains news articles spanning a period from 2012 to 2014 collected from Yahoo News<sup>2</sup>. It is an independent common resource for performance evaluation among reader-emotion research (e.g. Lin *et al.* (2007)), since it has a special feature which allows a reader of a news article to select from eight emotions the one that represent how the reader feels after reading a news article, i.e., “Angry”, “Worried”, “Boring”, “Happy”, “Odd”, “Depressing”, “Warm”, and “Informative”. To ensure the quality of the corpus, only articles with a clear statistical distinction between the highest vote of emotion and others determined by t-test with a 95% confidence level were retained. Finally, a total of 47,285 articles were retained from the original 68,026 articles, and they were divided into the training set consisting of 11,681 articles and the testing set consisting of 35,604 articles, respectively.

**Table 1. The distribution of data corpus.**

	<i>Angry</i>	<i>Worried</i>	<i>Boring</i>	<i>Happy</i>	<i>Odd</i>	<i>Depressing</i>	<i>Warm</i>	<i>Informative</i>
#Training	2,001	261	1,473	2,001	1,536	1,573	835	2,001
#Test	4,326	261	1,473	7,334	1,526	1,573	835	18,266
#Total	6,327	522	2,946	9,345	3,062	3,146	1,670	20,267

<sup>2</sup> <https://tw.news.yahoo.com/>

A comprehensive performance evaluation of the TBA with other methods is provided. The first is an emotion keyword-based model which is trained by SVM to demonstrate the effect of our keyword extraction approach (denoted as *KW-SVM*). Another is a probabilistic graphical model which uses the LDA model as document representation to train an SVM to classify the documents as either emotion relevant or irrelevant (Blei *et al.*, 2003) (denoted as *LDA-SVM*). The last one is a state-of-the-art reader-emotion recognition method which combines various feature sets including bigrams, words, metadata, and emotion category words (Lin *et al.*, 2007) (denoted as *CF-SVM*). To serve as a standard for comparison, we also included the results of Naïve Bayes (McCallum & Nigam, 1998) as a baseline (denoted as *NB*). Details of the implementations of these methods are as follows. We employed CKIP<sup>3</sup> for Chinese word segmentation. The dictionary required by Naïve Bayes and LDA-SVM is constructed by removing stop words according to a Chinese stop word list provided by Zou *et al.* (2006), and retaining tokens that make up 90% of the accumulated frequency. In other words, the dictionary can cover up to 90% of the tokens in the corpus. As for unseen events, we use Laplace smoothing in Naïve Bayes, and an LDA toolkit<sup>4</sup> is used to perform the detection of LDA-SVM. Regarding the CF-SVM, the words outputted by the segmentation tool were used. The information related to news reporter, news category, location of the news event, time (hour of publication) and news agency were used as the metadata features. The extracted emotion keywords were used as the emotion category words, since the emotion categories of Yahoo! Kimo Blog was not provided. To evaluate the effectiveness of these systems, we adopted the accuracy measures used by Lin *et al.* (2007). We used macro-average and micro-average to compute the average performance. These measures are defined based on a contingency table of predictions for a target emotion  $E_k$ . The accuracy  $A(E_k)$ , macro-average  $A^M$ , and micro-average  $A^\mu$  are defined as follows:

$$A(E_k) = \frac{TP(E_k) + TN(E_k)}{TP(E_k) + FP(E_k) + TN(E_k) + FN(E_k)} \quad (9)$$

$$A^M = \frac{1}{m} \sum_{k=1}^m A(E_k) \quad (10)$$

$$A^\mu = \frac{\sum_{k=1}^m TP(E_k) + TN(E_k)}{\sum_{k=1}^m (TP(E_k) + FP(E_k) + TN(E_k) + FN(E_k))} \quad (11)$$

where  $TP(E_k)$  is the set of test documents correctly classified to the emotion  $E_k$ ,  $FP(E_k)$  is the set of test documents incorrectly classified to the emotion,  $FN(E_k)$  is the set of test documents wrongly rejected, and  $TN(E_k)$  is the set of test documents correctly rejected.

---

<sup>3</sup> <http://ckipsvr.iis.sinica.edu.tw/>

<sup>4</sup> <http://nlp.stanford.edu/software/tmt/tmt-0.4/>

## 4.2 Results and Discussion

The performances of emotion detection systems are listed in Table 2. As a baseline, the Naïve Bayes classifier is a keyword statistics-based system which can only accomplish a mediocre performance. Since it only considers surface word weightings, it is difficult to represent inter-word relations. The overall accuracy of the Naïve Bayes classifier is 36.84%, with the emotion “Warm” only achieving 15% accuracy.

On the contrary, the LDA-SVM included both keyword and long-distance relations, and greatly outperforms the Naïve Bayes with an overall accuracy of 76.1%. It even achieved the highest accuracy of 92.83% and 85.40% for the emotion “Worried” and “Odd”, respectively, among all five methods. As we can see, the KW-SVM can bring about substantial proficiency in detecting the emotions with 77.70% overall accuracy. This indicates that reader’s emotion can be recognized effectively by using only the LLR scores of keywords, since the likelihood of a word existing in a certain emotion is not random. Those with a larger LLR value are considered as closely associated with the emotion (Manning & Schütze, 1999). The TBA can further improve the basic keyword-based method with rich context and semantic information, thus achieving the best overall accuracy of 84.72%.

**Table 2. Accuracy of emotion detection systems.**

<i>Topic</i>	<i>Accuracy (%)</i>				
	<b>NB</b>	<b>LDA-SVM</b>	<b>KW-SVM</b>	<b>CF-SVM</b>	<b>TBA</b>
Angry	47.00	74.21	79.21	83.71	<b>83.92</b>
Worried	69.56	<b>92.83</b>	81.96	87.50	80.12
Boring	75.67	76.21	84.34	87.52	<b>87.88</b>
Depressing	73.76	81.43	85.00	87.70	<b>90.13</b>
Happy	37.90	67.59	80.97	86.27	<b>89.50</b>
Warm	15.09	87.09	79.59	85.83	<b>88.56</b>
Odd	73.90	<b>85.40</b>	77.05	84.25	85.32
Informative	20.60	44.02	74.74	<b>83.59</b>	82.10
$A^M$	51.69	76.10	80.36	85.80	<b>85.94</b>
$A^H$	34.52	58.68	77.68	84.61	<b>84.72</b>

It is worth noting that the CF-SVM achieved a satisfactory accuracy around 80% among all emotions. This is because the combined lexicon feature sets (i.e. character bigrams, word dictionary, and emotion keywords) of CF-SVM improved the classification accuracy. In addition, the metadata of the articles are also associated with reader-emotion. For instance, we found that many sports related news articles evoke “Happy” emotion. In particular, 45% of all



“Happy” instances belong to the news category sports. It is also observed that an instance with the news category sports has a 31% chance of having the true class “Happy”. Hence, the high accuracy of “Happy” emotion can be the result of people’s general enthusiasm over sports rather than the result of a particular event. On top of that, the TBA can generate distinct semantic templates to capture alternations of similar combinations to achieve the optimal outcome. For instance, a semantic template generated by our proposed system, “{國家 country}: {發生 occur}: {地震 earthquake}: {劫難 disaster}”, belongs to the emotion “Depressing”. It is perceivable that this template is relaying information about disastrous earthquakes that occurred in a certain country, and such news often induce negative emotions among the readers. This example demonstrates that the automatically generated semantic templates are comprehensible for humans and can be utilized to effectively detect reader’s emotion. Nevertheless, our system could not surpass the LDA-SVM in the emotion “Worried”. It may be attributed to the fact that semantic templates generated in this emotion have inadequate quality. We examined some of the templates within this emotion and found that they mostly contain very general semantic classes, such as “{機構 institution}: {組織 organization}: {政黨 party}: {實現 realize}: {程度 degree}: {念頭 thought}”, thereby reducing its accuracy. Despite the “Worried” emotion, we were able to identify distinctive semantic templates for the other emotions. For instance, the template “[婦女 women]: {救助 help}: [小孩 child]: {當作 treat}: {人 human}: [認為 consider]” was generated for the emotion “Warm”, and it is reasonable that news about a woman helping a child would evoke a warm feeling in the readers’ mind. The ability to generate such emotion-specific templates is considered as the main reason for TBA to outperform other systems.

As a last touch, we dig into the top keywords in each emotion category as an analysis of the news trend, listed in TABLE 3. As stated in the previous section, we observe that keywords related to “Happy” are mostly terms about sports, such as team names (e.g., “熱火 Miami Heat” and “紅襪 Boston Red Sox”) or player names (e.g., “陳偉殷 Wei-Yin Chen”, a pitcher for the Baltimore Orioles). Similar findings had been made before as well (Chen *et al.* 2008). This is certainly a good indication for the performance of a sports team over a specified period of time. On the other hand, “Angry”-related keywords consist largely of political parties or public issues. For instance, the most noticeable word “美牛 United States beef” indicates the heated dispute on the import policy of beef from United States to Taiwan, which has been an issue in Taiwan-U.S. relations and leads to a domestic political unrest. Numerous political terms showed up in the top list too, such as “國民黨 Kuomintang”, “立法院 legislature”, and “立委 legislator”. This highlights the extracted emotion keywords are highly correlated with reader’s emotion, thus tagging them in the emotion template helps our method discriminate reader-emotion.

As for the Depressing category, the keywords mostly relate to social events that involve

severe weather or casualties. The phrase “大炳 Da Bing” refers to a Taiwanese performer who died in 2012 (around the time of our data retrieval) due to drug abuse. On the off chance that sports players suffered from low performance temporarily, they might show up in the category too due to readers’ compassion. Theoretically if we have sufficient patterns containing negation terms, we will be able to generate representative templates that can give rise to the scores in negative emotions. However, in our current data set, the relatively low portion of negative cases affects little more than the score of positive emotion, hence affecting the system performance. We acknowledge that this is an important task and should be studied in our future work. Lastly, not surprisingly, the Warm category contains words associated with mostly social care or volunteer providing charity or assistance to social vulnerable groups, while economic news dominates the Informative emotion.

**Table 3. A top 10 keyword list of each reader-emotion generated from TBA.**

Keywords discovered by the TBA	
Angry	美牛 American beef, 立委 Legislator, 立法院 Legislative Yuan, 國民黨 Kuomintang Party, 政府 Government, 證所稅 Stock exchange income tax, 中油 CPC, 總統 President, 馬英九 Ma Ying-jeou, 黨團 Political party
Worried	颱風 Typhoon, 氣象局 Central Weather Bureau, 病毒 Virus, 感染 Infection, 豪雨 Rain storm, 地震 Earthquake, 土石流 Mudslide, 餘震 Aftershock, 疾病 Disease, 病例 Patient record
Boring	蘇貞昌 Su Jen Chang, 陳水扁 Chen Shui-bian, 名嘴 Critics, 章子怡 Zhang Ziyi, 陳文茜 Chen Wen-chien, 宋正宇 Sung Jen-yu, 節目 TV show, 吳宗憲 Jacky Wu, 藝人 Celebrity, 女星 Female celebrity
Depressing	大炳 Da Bing, 民進黨 Democratic Progressive Party, 遺體 Remain, 送醫 To hospital, 溺水 Drown, 急救 Emergency medical service, 不治 Dead, 失蹤 Missing, 豪雨 Rain storm, 曾雅妮 Tseng Yani
Happy	陳偉殷 Chen Wei-yin, 安打 Strike, 比賽 Game, 熱火 Heat, 紅襪 Red Sox, 宏達電 HTC, 殷仔 Yin, 雷霆 Thunder, 太空人 Astros, 冠軍 Championship
Warm	孩子 Children, 家扶 Fund for family, 媽媽 Mother, 父親 Father, 志工 Volunteer, 關懷 Care, 基金會 Foundation, 行善 Charity, 幫助 Assistance, 弱勢 Social vulnerable
Odd	男子 Man, 警方 Police, 女子 Woman, 台灣 Taiwan, 發現 discover, 竟然 surprisingly, 監視器 Security camera, 網友 Internet user, 立委 Legislator, 美牛 American beef
Informative	工商時報 Commercial Times, 市場 Market, 報導 Report, 營收 Revenue, 成長 Growth, 需求 Necessity, 金融 Financial, 產品 Product, 自由時報 Liberty Times, 投資 Investment

To summarize, the proposed TBA integrates the syntactic, semantic, and context information in text to identify reader-emotions and achieves the best performance among the compared methods. It also demonstrates the capabilities of our approach to integrate statistical and knowledge-based models. Notably, in contrast to models used by previous machine

learning-based methods which are generally not human understandable, the generated templates can be acknowledged as the fundamental knowledge for each emotion and are comprehensible to the human mind.

## 5. Conclusion

With the rapid growth of computer mediated communication applications, the research on emotion classification has been attracting more and more attention recently from enterprises toward business intelligence. Recognizing reader-emotion concerns the emotion expressed by a reader after reading the text, and it holds the potential to be applied in fields that differ from writer-emotion detection applications. For instance, users are able to retrieve documents that contain relevant contents and at the same time produce desired feelings by integrating reader-emotion into information retrieval. In addition, reader-emotion detection can assist writers in foreseeing how their work will influence readers emotionally. In this research, we presented a flexible template-based approach (TBA) for detecting reader-emotion that simulates the process of human perception. By capturing the most prominent and representative pattern within an emotion, TBA allows us to effectively recognize the reader-emotion of text. Results of our experiments demonstrate that the TBA can achieve a higher performance than other well-known methods of reader-emotion detection. In the future, we plan to refine our TBA and employ it to other natural language processing applications. Also, further studies can be done on combining statistical models into different components in our system.

## Acknowledgement

This research was supported by the Ministry of Science and Technology of Taiwan under grant MOST 103-3111-Y-001-027.

## Reference

- Bashaddadh, O.M.A. & Mohd, M. (2011). Topic detection and tracking interface with named entities approach. In *Proceedings of International Conference on Semantic Technology and Information Retrieval*, 215-219.
- Blei, D.M., Ng, A.Y., & Jordan, M.I. (2003). Latent dirichlet allocation. *Journal of Machine Learning Research*, 3, 993-1022.
- Chen, K.J., Huang, S.L., Shih, Y.Y., & Chen, Y.J. (2005). Extended-HowNet: A representational framework for concepts. In *Proceedings of OntoLex – Ontologies and Lexical Resources IJCNLP-05 Workshop*, 1-6.

- Chen, Y., Lee, S., Li, S., & Huang, C. (2010). Emotion cause detection with linguistic constructions. In *Proceedings of the 21th International Conference on Computational Linguistics*, 179-187.
- Cortes, C. & Vapnik, V. (1995). Support vector networks. *Machine Learning*, 20, 1-25.
- Das, D. & Bandyopadhyay, S. (2009). Word to Sentence Level Emotion Tagging for Bengali Blogs. In *Proceedings of the 47th Annual Meeting of the Association for Computational Linguistics*, 149-152.
- Dong, Z.D., Dong, Q., & Hao, C.L. (2010). Hownet and its computation of meaning. In *Proceedings of the 23rd International Conference on Computational Linguistics: Demonstrations*, 53-56.
- Garey, M.R. & Johnson, D.S. (1979). *Computers and intractability: A Guide to the Theory of NP- Completeness*. W. H. Freeman & Co. New York, NY, USA.
- Johnson, D.S. (1974). Approximation algorithms for combinatorial problems. *Journal of Computer and System Sciences*, 9(3), 256-278.
- Read, J. (2005). Using Emotions to Reduce Dependency in Machine Learning Techniques for Sentiment Classification. In *Proceedings of the 43th Annual Meeting of the Association for Computational Linguistics Student Research Workshop*, 43-48.
- Lovász, L. (2005). Random walks on graphs: A survey. *Combinatorics, Paul erdos is eighty*, 2(1), 1-46.
- Lee, C.S., Jian, Z.W., & Huang, L.K. (2005). A fuzzy ontology and its application to news summarization. *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics*, 35(5), 859-880.
- Lin, H.Y., Yang, C.H., & Chen, H.H. (2007). What Emotions Do News Articles Trigger in Their Readers? In *Proceedings of 30th Annual International ACM SIGIR Conference*, 23-27.
- Lin, K.H., Yang, C.H., & Chen, H.H. (2008). Emotion Classification of Online News Articles from the Reader's Perspective. In *Proceedings of International Conference on Web Intelligence*, 220- 226.
- Manning, C.D. & Schütze, H. (1999). *Foundations of Statistical Natural Language Processing*. MIT Press.
- McCallum, A. & Nigam, K. (1998). A comparison of event models for Naïve Bayes text classification. In *Proceedings of AAAI/ICML-98 Workshop on Learning for Text Categorization*, 41-48.
- Mishne, G. (2005). Experiments with mood classification in blog posts. In *Proceedings of the 1st Workshop on Stylistic Analysis of Text for Information Access*.
- Mukherjee, A. & Liu, B. (2010). Improving Gender Classification of Blog Authors. In *Proceedings of Conference on Empirical Methods in Natural Language Processing*, 207-217.

- Pang, B., Lee, L., & Vaithyanathan, S. (2002). Thumbs up?: sentiment classification using machine learning techniques. In *Proceedings of the ACL-02 conference on Empirical methods in natural language processing*, 79-86.
- Patwardhan, S. & Peterson, T. (2003). Using WordNet-based Context Vectors to Estimate the Semantic Relatedness of Concepts.
- Purver, M. & Battersby, S. (2012). Experimenting with distant supervision for emotion classification. In *Proceedings of the 12th Conference of the European Chapter of the Association for Computational Linguistics*, 482-491.
- Quan, C. & Ren, F. (2009). Construction of a Blog Emotion Corpus for Chinese Emotional Expression Analysis. In *Proceedings of Conference on Empirical Methods in Natural Language Processing*, 1446-1454.
- Ricardo, B.Y. & Berthier, R.N. (2011). *Modern Information Retrieval: The Concepts and Technology Behind Search*. New York: Addison Wesley.
- Salton, G. & Buckley, C. (1988). Term-weighting approaches in automatic text retrieval. *Information Processing and Management*, 24(5), 513-523.
- Shih, C.W., Lee, C.W., Tsai, T.H., & Hsu, W.L., (2012). Validating Contradiction in Texts Using Online Co-Mention Pattern Checking. *ACM Transactions on Asian Language Information Processing*, 11(4), 17. doi: 10.1145/2382593.2382599 <http://doi.acm.org/10.1145/2382593.2382599>
- Si, J., Mukherjee, A., Liu, B., Li, Q., Li, H., & Deng, X. (2013). Exploiting Topic based Twitter Sentiment for Stock Prediction. In *Proceedings of The 51st Annual Meeting of the Association for Computational Linguistics*, 24-29.
- Tang, Y.J. & Chen, H.H. (2012). Mining Sentiment Words from Microblogs for Predicting Writer-Reader-emotion Transition. In *Proceedings of 8th International Conference on Language Resources and Evaluation*, 1226-1229.
- Thayer, R.E. (1989). *The Biopsychology of Mood and Arousal*, Oxford University Press.
- Wang, M.H., Lee, C.S, Hsieh, K.L., Hsu, C.Y., Acampora, G., & Chang, C.C. (2010). Ontology-based multi-agents for intelligent healthcare applications. *Journal of Ambient Intelligence and Humanized Computing*, 1(2), 111-131.
- Wu, C.H., Chuang, Z.J., & Lin, Y.C. (2006). Emotion recognition from text using semantic labels and separable mixture models. *ACM Transactions on Asian Language Information Processing*, 5(2), 165-183.
- Yang, C.H., Lin, H.Y., & Chen, H.H. (2009). Writer Meets Reader: Emotion Analysis of Social Media from both the Writer's and Reader's Perspectives. In *Proceedings of International Conference on Web Intelligence*, 287-290.
- Yang, C.H., Lin, K.H.Y., & Chen, H.H. (2007). Building emotion lexicon from Weblog corpora. In *Proceedings of 2007 IEEE/WIC/ACM International Conference on Web Intelligence*, 275-278.

- Yang, Y.H., Liu, C. C., & Chen, H. H. (2006). Music emotion classification: A fuzzy approach. In *Proceedings of the 14th Annual ACM International Conference on Multimedia*, 81-84.
- Zhang, L. & Liu, B. (2010). Extracting and Ranking Product Features in Opinion Documents. In *Proceedings of the 23rd International Conference on Computational Linguistics*, 1462-1470.
- Zou, F., Wang, F.L., Deng, X., Han, S., & Wang, L.S. (2006). Automatic construction of Chinese stop word list. In *Proceedings of the 5th WSEAS International Conference on Applied Computer Science*, 1010-1015.