2015

Semantic Frame-Based Approach for Reader-Emotion Detection

Yung-Chun Chang
National Taiwan University, changyc@iis.sinica.edu.tw

Cen-Chieh Chen
National Chengchi University, can@iis.sinica.edu.tw

Wen-Lian Hsu
Academia Sinica, hsu@iis.sinica.edu.tw

SEMANTIC FRAME-BASED APPROACH FOR READER-EMOTION DETECTION

Yung-Chun Chang, Department of Information Management, National Taiwan University, and Institute of Information Science, Academia Sinica, Taipei, Taiwan, R.O.C., changyc@iis.sinica.edu.tw

Cen-Chieh Chen, Department of Computer Science, National Chengchi University, Taipei, Taiwan, R.O.C., can@iis.sinica.edu.tw

Wen-Lian Hsu, Institute of Information Science, Academia Sinica, Taipei, Taiwan, R.O.C., hsu@iis.sinica.edu.tw

Abstract

Previous studies on emotion classification mainly focus on the writer's emotional state. By contrast, this 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 IR systems lack the ability to discern emotions within texts, and the detection of reader-emotion has yet to achieve a comparable performance. Moreover, previous machine learning-based approaches are generally not human understandable. Thereby, it is difficult to pinpoint the reason for recognition failures and understand the types of emotions articles inspire in their readers. In this paper, we propose a flexible semantic frame-based approach (FBA) for reader-emotion detection that simulates such process in a human perceptive manner. FBA is a highly automated process that incorporates various knowledge sources to learn semantic frames from raw text that characterize an emotion and are comprehensible for humans. Generated frames are adopted to predict reader-emotion through an alignment-based matching algorithm that allows a semantic frame to be partially matched through a statistical scoring scheme. Experimental results demonstrate that our approach can effectively detect reader-emotions by exploiting the syntactic structures and semantic associations in the context, while outperforming currently well-known statistical text classification method sand the stat-of-the-art reader-emotion detection method.

Keywords: Reader-Emotion Detection, Semantic Frame, Frame-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 Blog, Twitter and Facebook. Therefore, sentiment analysis studies have gained increasing interest in recent years with an attempt to obtain trends in public by mining opinions that are subjective statements that reflect people's sentiments or perceptions about topics (Pang et al., 2002). Moreover, human feelings can be quickly collected through emotion detection, as these emotions reflect an individual's feelings and experiences toward some subject matters (Quanand Ren, 2009; Das and 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 note that the reader-emotions, in some aspects, differ from that of the authors and may be even more complex (Lin et al., 2008; Tang and Chen, 2012). Regarding a news article, while its writer just objectively reports the news and thus does not express his/her emotion in the text, a reader could yield angry or boring emotions. For instance, 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, reader-emotion possesses a more complex nature, as even general words can evoke different types of reader-emotions depending on the reader's personal experience and knowledge (Lin et al., 2007).

Instead of writer-emotion detection, which has been widely investigated in previous studies, this paper aims to uncover the emotions documents trigger in their readers. 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 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 instill proper emotions. As a result, users benefit by obtaining opportunities and advantages in the competitive market though 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 labor cost. Nevertheless, the models used by these classifiers are generally not human understandable. Thus, it is difficult to pinpoint the reason for recognition failures and understand what emotions the articles trigger in their readers.

In light of this rationale, we proposed a flexible semantic frame-based approach (FBA) for reader-emotion detection that simulates such process in human perception. FBA 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, FBA recognizes reader-emotions of documents using an alignment-based algorithm that allows a semantic frame to be partially matched through a statistical scoring scheme. Our experiments demonstrate that FBA 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 semantic frame-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

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 and 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 and 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 a semantic frame-based approach that mimics the perceptual behavior of humans in understanding. Second, the generated semantic frames 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, FBA further considers the surrounding context and semantic associations to efficiently recognize reader-emotions.

3 READER-EMOTION DETECTION USING SEMANTIC FRAMES

In this paper, we present a frame-based approach (FBA) 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 $W= \{w_1, w_2, \ldots, w_k\}$ be a set of words, $D= \{d_1, d_2, \ldots, d_m\}$ be a set of documents, and $E= \{e_1, e_2, \ldots, e_n\}$ 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 automatic generation algorithm to generate frames that represent discriminative patterns in documents. FBA mainly consists of three components: Crucial Element Labeling (CEL), Semantic Frame Generation (SFG), and Semantic Frame Matching (SFM), as shown in Figure 1. The CEL first uses prior knowledge to mark the semantic classes of words in the corpus. Then the SFG collects frequently co-occurring elements, and generates frames for each emotion. These frames 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 labeled by the CEL as well. Subsequently, the SFM 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.

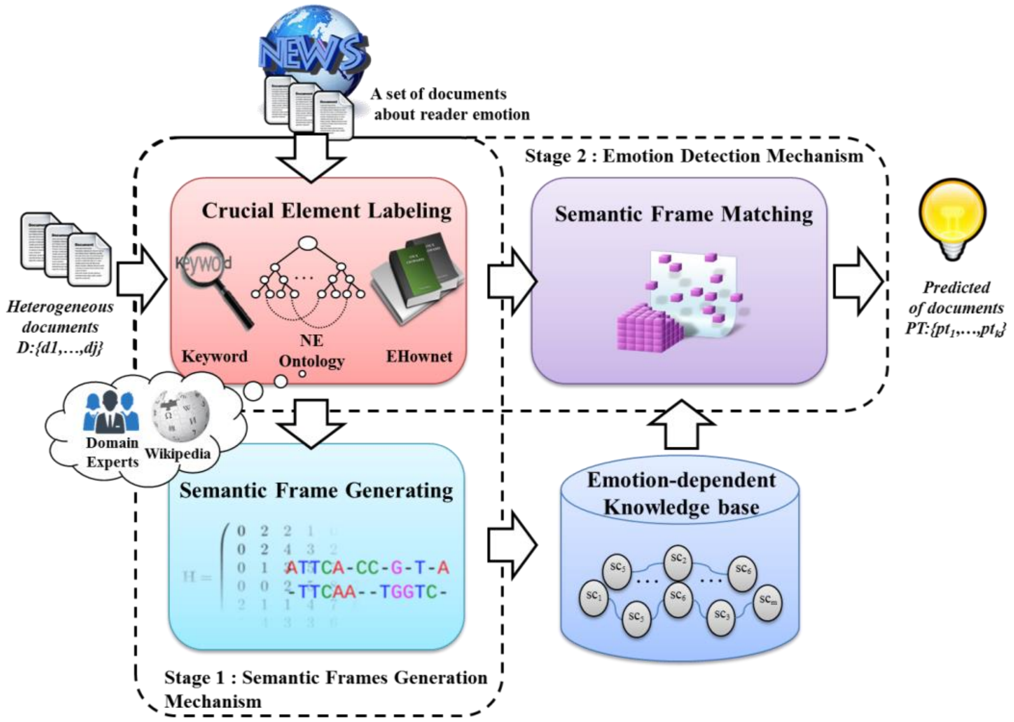

Figure 1. Architecture of our semantic frame-based emotion detection system.

3.1 Crucial Element Labeling (CEL)

FBA 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 labeling approach that utilizes various knowledge sources, such as dictionaries and the Wikipedia, to induce frame elements. First of all, since keywords within a reader-emotion are often considered as important information, we used the log likelihood ratio (LLR) (Manning and 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 denotes the set of documents of the reader-emotion in the training dataset; $N(E)$ and $N(\neg E)$ are the numbers of on-emotion and off-emotion documents, respectively; and N(w^E) is the number of document on-emotion having w. The probabilities $p(w)$, $p(w|E)$, and $p(w|$^$E)$ 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.

$$-2\log\left[\frac{p(w)^{N(w\wedge E)}(1-p(w))^{N(E)-N(w\wedge E)}\,p(w)^{N(w\wedge\neg E)}(1-p(w))^{N(\neg E)-N(w\wedge\neg E)}}{p(w|E)^{N(w\wedge E)}(1-p(w|E))^{N(E)-N(w\wedge E)}\,p(w|\neg E)^{N(w\wedge\neg E)}(1-p(w|\neg E))^{N(\neg E)-N(w\wedge\neg E)}}\right] \quad (1)$$

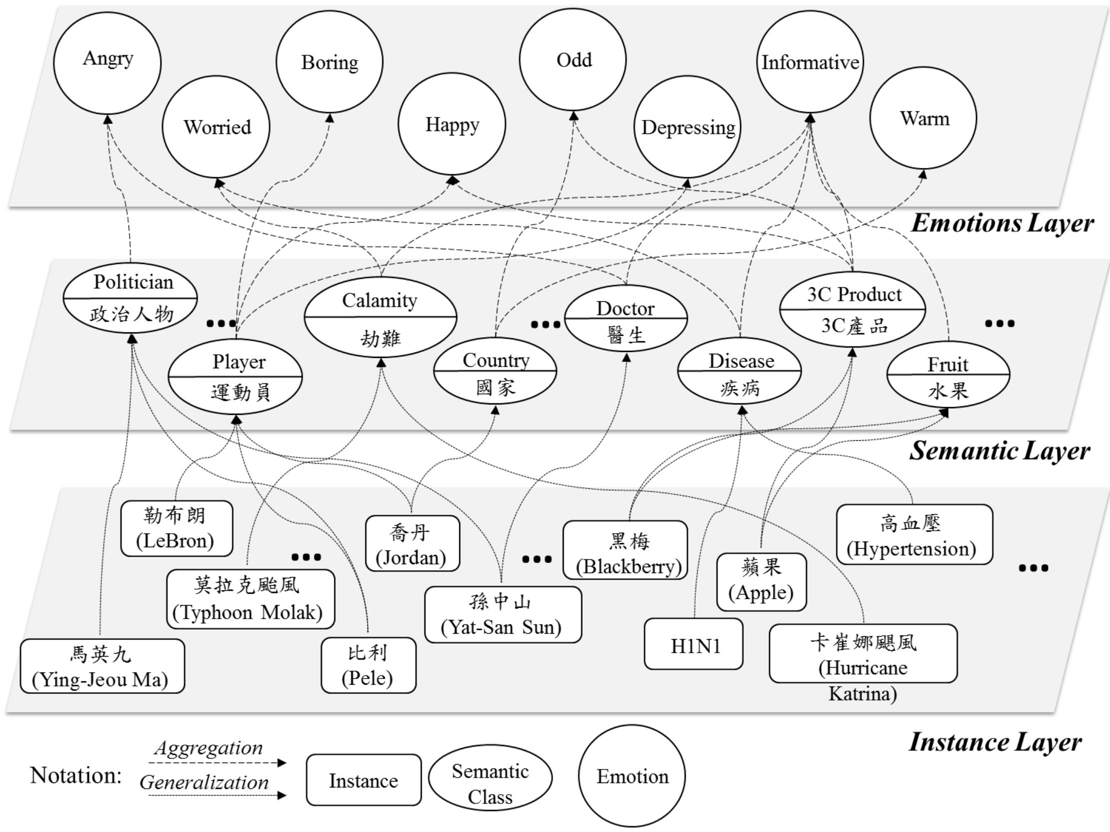

Figure 2. Architecture of named entity ontology.

Next, we aim to recognize named entities (NEs) from text to facilitate document comprehension and improve the performance of identifying topics (Bashaddadh and Mohd, 2011). However, exact match NEs may overlook many frame elements since it omits semantic context. To remedy this problem, this paper adopts a novel structure to construct the NE ontology (NEO) for labeling 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 labor 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[1]. We select the category tag to which the most *topic paths* are associated, and use them to represent the main semantic label of NEs in documents. Topic paths can be considered as the traversal from general categories to more specific ones. Thus, more topic paths may indicate that this category is more general. 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)]". We also annotate NE

[1]http://nlp.stanford.edu/software/CRF-NER.shtml

terms 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 frame, we exploit taxonomies that include lexical categories, synonyms, and semantic relations between different words or sets of words. In this paper, we use the Extended HowNet (Chen et al., 2005), which is an extension of the HowNet (Dong et al. 2010) for a structured representation of knowledge and semantics. It connected approximately 90 thousand words of the CKIP Chinese Lexical Knowledge Base and HowNet, and included additional highly frequent words that are specific in Traditional Chinese. It also contains a different formulation of each word to better fit its semantic 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’s take “血癌 (leukemia)” defined in Definition 1 as an example. 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}}

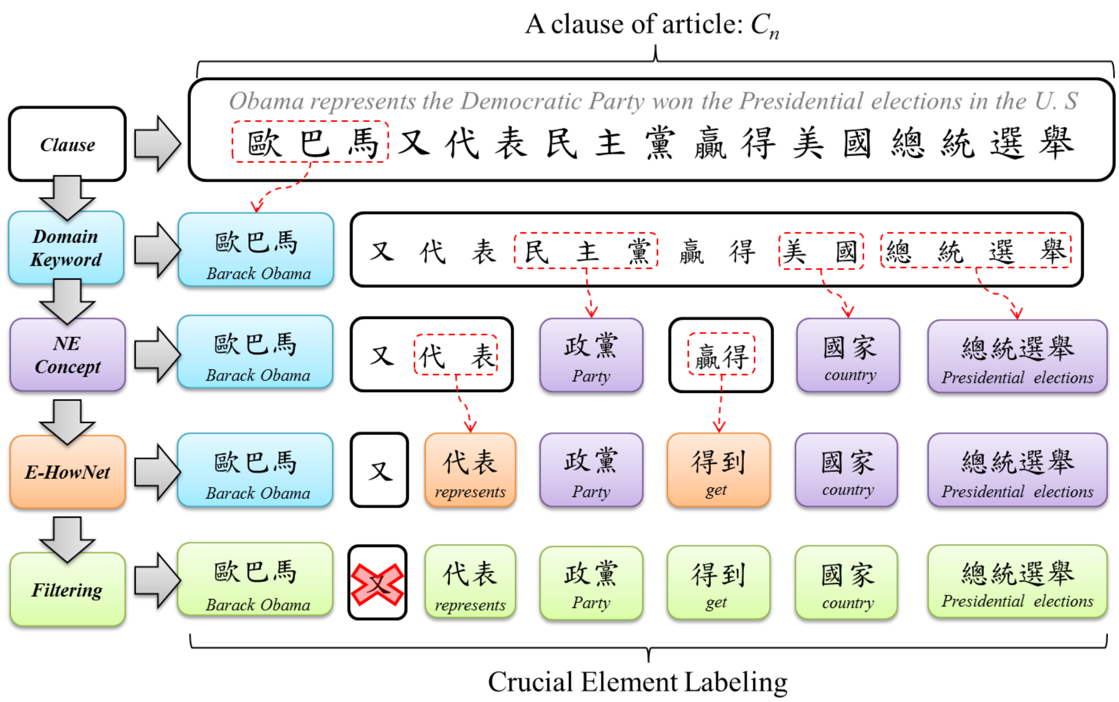

Figure 3. *Crucial element labeling process.*

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 labeling. To illustrate the process of CEL, consider the clause C_n = "歐巴馬又代表民主黨贏得美國總統選舉 (Obama represents the Democratic Party won the Presidential elections in the U. S)", as shown in Figure 3. First, "歐巴馬 (Obama)" 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 (represents)" and "贏得 (won)" are labeled with their corresponding E-HowNet senses if they exist. Finally, we can obtain a sequence like "[歐巴馬]:{代表}:{政黨}:{得到}:{國家}:{總統選舉} ([Obama]:{represents}:{party}:{got}:{country}:{ Presidential elections })." This labeling process can not only prevent errors caused by Chinese word segmentation, but also group the synonyms together using semantic labels. It enables us to generate distinctive and prominent semantic templates in the next stage.

3.2 Semantic Frame Generation (SFG)

In our framework, a reader-emotion is represented by a set of semantic frames that consist of crucial elements and keywords. The semantic frame generation (SFG) process aims at automatically generating *N* representative frames from sequences of crucial elements in the documents. These representative (or dominating) frames can be used as background knowledge for each reader-emotion when recognizing documents. More importantly the representative frames can be easily understood by humans. To illustrate, consider the emotion "Happy" and one of the automatically generated semantic frames "{運動員$_{\text{player}}$}:{得到$_{\text{get}}$}: [冠軍$_{\text{championship}}$]." We can think of various semantically similar sentences that are covered by this semantic frame, e.g., "科比帶領湖人贏得了NBA總冠軍賽 (Kobe Bryant led the Lakers 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)". This sort of interpretable knowledge cannot be easily obtained by ordinary machine learning models.

The SFG process is described as follows. We observed that the rank-frequency distribution of the elements in templates follow the Zipf's law (Manning and Schütze, 1999). Thus, we adopted the dominating set algorithm to use only the frames in the Top 20% after ranking by frequency to cover the rest of the frames. Since it is an NP-hard problem (Garey and Johnson, 1979), we implemented an approximation using a greedy algorithm based on Johnson (1974). First, we constructed a directed graph $G= \{V, E\}$, in which vertices V contains all crucial element sequences $\{CES_1, \ldots, CES_m\}$ in each reader-emotion, and edges E represent the dominating relations between sequences. If CES_x dominates CES_y, there is an edge $CES_x \rightarrow CES_y$. The definition of a dominating relation is as follows. 1) Highly frequent crucial element sequences were selected as the dominators. 2) Longer sequences dominate shorter ones if their head and tail elements were identical. The intermediate elements could be skipped, as they can be identified as insertions and given scores based on their distribution in this category during the matching process. Use of the dominating set can help us capture the most prominent and representative sequences within a category. Afterwards, the dominating sequences further undergo a selection process that is similar to our keyword selection method mentioned previously. Lastly, we retain the frames based on its dominating rate and preserved the top 100 from approximately 55,000 crucial element sequences. As a result, we were able to reduce the number of frames while keeping the most prominent and distinctive ones, which assists the execution of our matching algorithm.

3.3 Semantic Frame Matching (SFM)

We believe the human perception of an emotion is obtained through recognizing important events or semantic contents to rapidly evoke their emotion. 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. FBA uses an alignment algorithm to measure the similarity of frames, since alignment enables a single frame to match multiple semantically similar expressions with appropriate scores. During matching, a document is first labeled with crucial elements. Afterwards, an alignment-based algorithm (Needleman and Wunsch, 1970) is applied to determine to what degree a semantic frame fits in a document. For each clause within a given document d_j, we first label crucial elements $cs = \{ce_1, \ldots, ce_n\}$, followed by the matching procedure that compares all sequences of crucial element in d_j to all semantic frames $SF = \{sf_1, \ldots, sf_j\}$ in each emotion, and calculates the sum of scores for each emotion. The emotion e_i with the highest sum of scores defined in (2) is considered as the winner.

$$Emotion = \arg\max_{e_i \in E} \sum_{sf_n \in SF_{c_i}, cs_m \in CS_{d_j}} \Delta(sf_n, cs_m) = \sum_k \sum_l \Delta(sf_n \cdot fe_k, cs_m \cdot ce_l) \tag{2}$$

where fe_k and ce_l represent the k^{th} frame element of sf_n and l^{th} crucial element of cs_m, respectively. Scoring of the matched and unmatched components in semantic frames is as follows. If $sf_n \cdot fe_k$ 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 (3). On the contrary, if an element is not matched, the score of insertion or deletion is calculated. An insertion (*IS*), defined as (4), can be accounted for by the inversed entropy of this crucial element, which represents the uniqueness or generality of this element among emotions. And a deletion (*DS*), defined as (5), is computed from the log frequency of this crucial element in this emotion. The detailed algorithm is described in Algorithm 1.

Algorithm 1:Semantic Frame Matching

Input: A semantic frame $sf = \{fe_1, \ldots, fe_m\}$, fe: frame elements; A sequence of crucial elements from a clause $cs = \{ce_1, \ldots, ce_n\}$
Output: Matching score σ between *sf* and *cs*

```
BEGIN
pos← 0; σ← 0;
FOR i = 1 to m DO
isMatched ← false;
pos← current matched position in CE;
    IF found ce_j EQUAL TO fe_i after pos THEN
        σ ← σ + MatchedScore(ce_j);
        isMatched ← true;
    END IF
    IF isMatched != trueTHEN
        σ ← σ- DeletionScore (fe_i);
        σ ← σ- InsertionScore (ce_j);
    END IF
END FOR
Output σ
END
```

$$MS(ce_l) = \begin{cases} LLR_{ce_l}, \text{if it matches a keyword} \\ \lambda \dfrac{f_{ce_l}}{\sum_{i=1}^{m} f_{ce_i}}, \text{otherwise} \end{cases} \tag{3}$$

$$IS(ce_l) = \frac{1}{-\sum_{i=1}^{m} P(ce_{l_{e_i}}) \cdot \log_2(P(ce_{l_{e_i}}))} \tag{4}$$

$$DS(ce_l) = \log \frac{f_{ce_l}}{\sum_{i=1}^{m} f_{ce_i}} \tag{5}$$

4 EXPERIMENT

4.1 Experiment Setup

4.1.1 Datasets

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[2]. It is an independent common resource for performance evaluation among reader-emotion researches (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.

	Angry	*Worried*	*Boring*	*Happy*	*Odd*	*Depressing*	*Warm*	*Informative*
#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

Table 1. The distribution of data corpus.

4.1.2 Comparison Settings and Evaluation Metrics

A comprehensive performance evaluation of the FBA 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 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 Naive Bayes (McCallum and Nigam, 1998) as a baseline (denoted as *NB*). Details of the implementations of these methods are as follows. We employed CKIP[3] for Chinese word segmentation. The dictionary required by Naïve Bayes and LDA-SVM is constructed

[2]https://tw.news.yahoo.com/
[3] http://ckipsvr.iis.sinica.edu.tw/

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[4] 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^μ are defined as follows:

$$A(E_k) = \frac{TP(E_k) + TN(E_k)}{TP(E_k) + FP(E_k) + TN(E_k) + FN(E_k)} \tag{6}$$

$$A^M = \frac{1}{m}\sum_{k=1}^{m} A(E_k) \tag{7}$$

$$A^\mu = \frac{\sum_{k=1}^{m} TP(E_k) + TN(E_k)}{\sum_{k=1}^{m} (TP(E_k) + FP(E_k) + TN(E_k) + FN(E_k))} \tag{8}$$

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.

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-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 and Sch¨utze, 1999). The FBA can further improve the basic keyword-based method with rich context and semantic information, thus achieving the best overall accuracy of 84.65%.

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.

[4]http://nlp.stanford.edu/software/tmt/tmt-0.4/

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 FBA can generate distinct semantic frames to capture alternations of similar combinations to achieve the optimal outcome. For instance, a semantic frame generated by our system, "{國家$_{\text{country}}$}:[發生$_{\text{occur}}$]:[地震$_{\text{earthquake}}$]:{劫難$_{\text{disaster}}$}", belongs to the emotion "Depressing". It is perceivable that this frame is relaying information about disastrous earthquakes that occurred in a certain country, and such news often makes readers depressed. This example demonstrates that the automatically generated semantic frames are comprehensible for humans and can be utilized to effectively detect reader-emotion. Nevertheless, our system could not surpass the LDA-SVM in the emotion "Worried". It may be attributed to the fact that semantic frames generated in this emotion have inadequate quality. We examined some of the frames within this emotion and found that they mostly contain very general semantic classes, such as "{機構$_{\text{institution}}$}:{組織$_{\text{organization}}$}:{政黨$_{\text{party}}$}:{實現$_{\text{realize}}$}:{程度$_{\text{degree}}$}:{念頭$_{\text{thought}}$}", thereby reducing its accuracy. Despite the "Worried" emotion, we were able to identify distinctive semantic frames for the other emotions. For instance, the frame "[婦女$_{\text{women}}$]:{救助$_{\text{help}}$}:[小孩$_{\text{child}}$]:{當作$_{\text{treat}}$}:{人$_{\text{human}}$}:[認為$_{\text{consider}}$]" was generated for the emotion "Warm", and it is understandable that news about a woman helping a child would evoke a warm feeling in the readers' heart. The ability to generate such emotion-specific frames is considered as the main reason for FBA to outperform other systems.

Topic	***Accuracy (%)***				
	NB	**LDA-SVM**	**KW-SVM**	**CF-SVM**	**FBA**
Angry	47.00	74.21	79.21	83.71	**87.83**
Worried	69.56	**92.83**	81.96	87.50	75.80
Boring	75.67	76.21	84.34	87.52	**90.52**
Happy	37.90	67.59	80.97	86.27	**88.94**
Odd	73.90	**85.40**	77.05	84.25	83.34
Depressing	73.76	81.43	85.00	87.70	**92.15**
Warm	15.09	87.09	79.59	85.83	**91.91**
Informative	20.60	44.02	74.74	**83.59**	80.92
A^M	51.69	76.10	80.36	85.80	**86.43**
A^μ	34.52	58.68	77.68	84.61	**84.63**

Table 2. Accuracy of emotion detection systems.

To summarize, the proposed FBA 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 frames 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 frame-based approach (FBA) for detecting reader-emotion that simulates the process of human perception. By capturing the most prominent and representative pattern within an emotion, FBA allows us to effectively recognize the reader-emotion of text. Results of our experiments demonstrate that the FBA can achieve a higher performance than other well-known methods of reader-emotion detection. In the future, we plan to refine our FBA and employ it to other NLP 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.

References

Blei, D.M., Ng A.Y., and Jordan, M.I. (2003). Latent dirichlet allocation. Journal of Machine Learning Research, 3, 993-1022.

Chen, K.J., Huang, S.L., Shih, Y.Y., and 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., and Huang, C. (2010). Emotion cause detection with linguistic constructions. In Proceedings of the 21th International Conference on Computational Linguistics, 179-187.

Cortes, C. and Vapnik, V. (1995). Support vector networks. Machine Learning, 20, 1-25.

Das D. and 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. and 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. and 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.

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. and 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., and 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. and Sch¨utze, H. (1999) Foundations of Statistical Natural Language Processing. MIT Press.

McCallum, A. and 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.

Bashaddadh, O.M.A and 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.

Pang, B., Lee, L., and 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.

Purver M. and S. Battersby. 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. and 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, pp.1446-1454.

Salton, G. and 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), 17:1-12:21.

Tang, Y.J. and 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., and 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., and 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. and 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.

Zou, F., Wang, F.L., Deng, X., Han, S., and 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.