Mining Sentiment Words from Microblogs for Predicting Writer-Reader Emotion Transition

Yi-jie Tang, **Hsin-Hsi Chen**
Department of Computer Science and Information Engineering
National Taiwan University, Taipei, Taiwan
E-mail: tangyj@nlg.csie.ntu.edu.tw; hhchen@ntu.edu.tw

Abstract

The conversations between posters and repliers in microblogs form a valuable writer-reader emotion corpus. In a microblog conversation, the writer of the initial post and the reader who replies to the initial post can both express their emotions. The process of changing from writer emotion to reader emotion is called a writer-reader emotion transition in this paper. Log relative frequency ratio is adopted to investigate the linguistic features that affect emotion transitions, and the results are used to predict writers' and readers' emotions. A 4-class emotion transition predictor, a 2-class writer emotion predictor, and a 2-class reader emotion predictor are proposed and compared.

Keywords: emotion mining, microblogging, microtext classification, sentiment analysis

1. Introduction

People often express their feelings when writing and reading articles. Writers and readers do not always share the same emotions for the same text. The process of changing from writer emotion to reader emotion is called writer-reader emotion transition in this paper. To know which factors affect the emotion transition is important for human language understanding and has many potential applications.

Most of the researches on emotion analysis first focus on the writer's perspective. Pang et al. (2002) classified movie reviews into positive and negative emotions. Wiebe (2000) investigated the subjectivity of adjectives. Aman and Szpakowicz (2007) labeled phrases with emotional categories. Beyond binary classification, Mishne (2005) classified blog posts into 37 emotion classes.

Then, some work begins investigating reader-emotion analysis. Lin, Yang and Chen (2008) classified Yahoo! News articles into 8 emotion classes from readers' perspectives. Lin and Chen (2008) extended their work from reader emotion classification to emotion ranking. Yang, Lin and Chen (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.

This paper collects messages posted in microblogs and annotated with both writers' and readers' emotions by posters and repliers collaboratively. We mine linguistic features from the writer-reader emotion corpus and predict emotion transitions between writers and readers.

2. Plurk Dataset

In microblogging, social interaction is represented by a sequence of messages posting and replying. Figure 1 shows a typical conversation in Plurk[1], a web-based social network that allows users to post short messages limited to 140 characters. A message is written by a poster, i.e., the person who posts the message, and read by multiple repliers, i.e., those who read the initial message and give replies. We call the former a writer and the latter a reader. Both a poster and repliers may express their emotions in writing and reading. That forms a valuable dataset for investigating the emotion transition from a writer to a reader. We will study what factors keep the emotion not change during writing and reading, and what factors change the emotions.

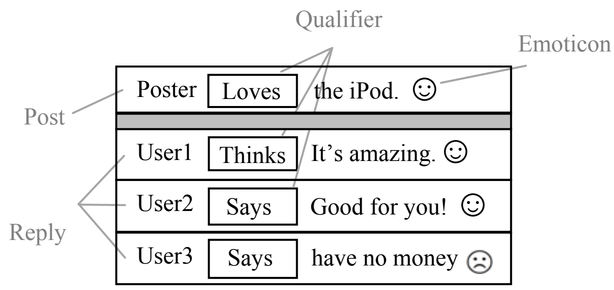

Figure 1: A conversation on Plurk

We use an emotion pair (*writer_emotion*, *reader_emotion*) to formulate the emotion transition, where *writer_emotion* means the emotion expressed by a writer, i.e., a poster, and *reader_emotion* means the emotion expressed by a reader, i.e., a replier. The emotion can be *positive* (*pos*) or *negative* (*neg*), so that there are four kinds of possible emotion transitions including (*pos*, *pos*), (*pos*, *neg*), (*neg*, *neg*) and (*neg*, *pos*).

Plurk provides 78 basic graphic emoticons, which are commonly used in users' messages. We choose 35 of the emoticons and categorize them into the positive and negative groups according to their names and

[1] http://www.plurk.com

common usage. The other 43 are either neutral or cannot be clearly categorized, so we exclude them to minimize uncertainty. Figure 2 lists the Plurk emoticons used in this study.

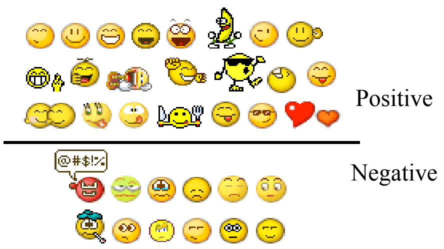

Figure 2: Emoticons as positive and negative labels

In microblogging, a microblogger may play the role of both a writer and a reader. When posting the first message, s/he is a writer. After reading the repliers' messages, s/he may give some feedback. In that case, s/he will serve as a reader to the replies. To avoid the confusion, we collect the first messages of microbloggers and the first reply to the messages to form an emotion corpus. Then we divide this corpus into four datasets based on the emotion transition types, i.e., (*pos*, *pos*), (*pos*, *neg*), (*neg*, *neg*) and (*neg*, *pos*). For clarity, the four datasets are named as *PP*, *PN*, *NN*, and *NP* datasets, respectively. We select 79,042 conversations to form our experimental corpus. The number of instances in each dataset *PP*, *NN*, and *NP* is 20,000. The number of instances in the dataset *PN* is 19,042 because fewer examples of (*pos*, *neg*) can be found. These datasets are also described in Resource Map along with this paper.

3. Sentiment Word Mining

We perform word segmentation and part of speech (POS) tagging on the four datasets with the Yahoo! Segmentation (斷章取義) system. In this way, a dataset is composed of Chinese words along with their POS. We will study the pairs of datasets to see if their word distribution is different, and what make them different.

3.1 Similarity among Emotion Datasets

	PP	*PN*	*NN*	*NP*
PP	1	0.899	0.816	0.871
PN	0.899	1	0.922	0.940
NN	0.816	0.922	1	0.953
NP	0.871	0.940	0.953	1

Table 1: Similarity among Emotion Datasets

We remove those function words with POS *articles*, *prepositions*, and *conjunctions* from the datasets. Each dataset is represented by a word vector (w_1, w_2, …, w_n), where w_i is a normalized weight of the word w^i. The weight w_i in a dataset, i.e., *PP*, *PN*, *NN*, or *NP*, is computed as: total occurrences of w^i divided by total number of words in the dataset. We employ cosine function to measure the similarity among each pair of datasets. Table 1 shows the results.

The dataset *PP* contains no negative emotions and *NN* contains no positive emotions, so it is natural that they differ a lot from each other (i.e., the lowest cosine similarity). The pair with the highest cosine similarity is *NN* and *NP*. It means that negative writer messages are similar, regardless of reader emotion. That is, the same writer message with negative emotion can cause either positive or negative reader emotion. This may make reader emotion more difficult to predict when the writer message is negative. In contrast, *PP* and *PN* have lower cosine similarity. It means that reader emotion is relatively easier to predict if the writer message is positive.

Generally speaking, all the datasets with any negative emotions, including *PN*, *NN* and *NP*, have higher cosine similarity (> 0.9) when compared with each other. *PP*, which contains positive emotions only, has lower cosine similarity (< 0.9) with all the other datasets.

3.2 Log Relative Frequency Ratio

The log relative frequency ratio *lr* of words in two datasets *A* and *B* defined as follows is used to select the critical features that capture the emotion transition.

For each $w^i \in A \cup B$, compute

$$lr_{AB}(w^i) = \log \frac{\dfrac{f_A(w^i)}{|A|}}{\dfrac{f_B(w^i)}{|B|}}$$

where $lr_{AB}(w^i)$ is a log ratio of relative frequencies of word w^i in *A* and *B*, $f_A(w^i)$ and $f_B(w^i)$ are frequencies of w^i in *A* and in *B*, respectively, and $|A|$ and $|B|$ are total words in *A* and in *B*, respectively. The log relative frequency ratios are used to estimate the distribution of the words in datasets *A* and *B*.

The interpretations of $lr_{AB}(w^i)$ are shown as follows.

(1) If w^i has higher relative frequency in *A* than in *B*, then $lr_{AB}(w^i)>0$. Those words of positive ratio form a set *A-B*.

(2) If w^i has higher relative frequency in *B* than in *A*, then $lr_{AB}(w^i)<0$. Those words of negative ratio form a set *B-A*.

(3) If w^i has similar relative frequency in both sets, then $lr_{AB}(w^i)\cong 0$.

A and *B* may have the following combinations.

(1) *A=PP*, *B=PN*
It captures emotion transitions *pos*→*pos* and *pos* → *neg*. Those words in *PN-PP* may be probable to affect the emotion transitions from positive to negative. Those words in *PP-PN* may be probable to keep the emotion unchanged, i.e., in positive state.

(2) *A=NP*, *B=NN*
It captures emotion transitions *neg*→*pos* and *neg*→*neg*. Those words in *NP-NN* may have

some effects on the emotion transition from negative to positive. Those words in *NN-NP* may keep the emotion unchanged, i.e., in negative state.

(3) $A=PP\cup PN$, $B=NN\cup NP$
It captures positive and negative emotion representations of *writers*. A writer emotion dictionary ED_W can be constructed based on the log relative frequency ratio of words in positive dataset $PP\cup PN$ and negative dataset $NN\cup NP$.

(4) $A=PP\cup NP$, $B=NN\cup PN$
It captures positive and negative emotion representations of *readers*. Thus, a reader emotion dictionary ED_R can be constructed from positive dataset $PP\cup NP$ and negative dataset $NN\cup PN$.

3.3 Analysis of the Mined Words

We examine the top 200 words with higher log relative frequency ratios in *PN-PP*, *PP-PN*, *NN-NP*, and *NP-NN*, respectively, and identify their semantic categories in the Chinese thesaurus *Tongyicicilin* (同義詞詞林. Mei et al, 1982), which is abbreviated as *Cilin*. There are 12 categories labeled by letters:

A. Human
B. Object
C. Time and Space
D. Abstract concept
E. Characteristics
F. Movement
G. Mental activity
H. Activity
I. Status
J. Relation
K. Particle words
L. Greetings

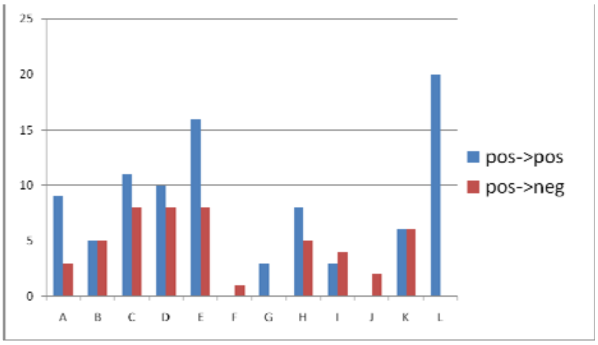

Figure 3: Category Distribution of Emotion Words

We calculate the word counts in each category. Only the words that can be found in *Cilin* are analyzed. Figure 3 shows the distribution of words in the *pos*→*pos* transition and the *pos*→*neg* transition. Words used in positive writer contents are more likely to get positive response, except for the categories B, F, I, J, and K. The most noticeable feature is greeting words (category L) such as 掰掰 (goodbye), 早安 (good morning), 晚安 (good night), which never cause the *pos*→*neg* transition. The words causing the *pos*→*neg* transition include some words in the category K like 難道 (dubiously), 幸好 (fortunately, which usually follows a negative expression), and 到底 (exactly). These words themselves do not contain negative emotion, but are usually used in expressions related to negative emotions.

The distribution of the words used for the *neg*→*pos* transition and the *neg*→*neg* transition is interpreted similarly. If a writer expresses negative emotion but somehow uses a greeting word in his/her message, s/he can still get positive response from the reader most of the time. That is, the words in the group L can cause the *neg*→*pos* transition.

The words used in the *neg*→*pos* transition include personal status like 好累 (tired) and 睏 (sleepy), which belongs to category I and can receive encouragement or other positive responses. As expected, the words in the transition *neg*→*neg*, including 生氣 (angry) (category G), 恐怖 (terrible) (category G), and 可惡 (hateful) (category E), are mostly used to express negative status or characteristics.

Some words in the datasets have no emotional meanings but are related to specific topics that can cause emotions. We identify them manually with reference to *Cilin*, and show some of the high frequency topical words as follows:

1. Political terms: 馬英九 (Ma Ying-jeou), 陳水扁 (Chen Shui-bian), 國民黨 (Kuomintang), 民進黨 (Democratic Progressive Party), etc. The former two are current and past Taiwan presidents, and the latter two are major parties in Taiwan. According to our observations, people tend to have negative emotions to the political entities.
2. News topics: 油價 (oil price), 牛肉 (beef), 健保 (health insurance), etc. The topics are related to government policy. People tend to have negative emotions to the increase of oil price.
3. Special days: 情人節 (Valentine's Day), 中秋節 (Mid-Autumn Festival), etc. These terms are usually associated with activities that affect emotions.

4. Writer-Reader Emotion Prediction

Given a message, we would like to predict which kind of emotion pair (*writer_emotion*, *reader_emotion*) is presented.

4.1 SVM Classifiers

A training set is composed of m messages $t_1, t_2, \ldots, t_m$ sampled from Plurk. Each message t_i is annotated with an emotion pair (e_{i1}, e_{i2}), where e_{i1} and e_{i2} denote poster's and replier's emotions, respectively. Assume there are n unique words w^1, w^2, …, w^n in the training set. At first, we employ the log relative frequency ratio to determine a writer emotion dictionary ED_W, and a reader emotion dictionary ED_R in the way specified in Section 3.2. Then, each message t_i is transformed into a 2n-dimension vector (w_{11}, w_{12}, …, w_{1n}, w_{21}, w_{22}, …, w_{2n}), where w_{1j} and w_{2j} are the weights of word w^j in the writer and reader emotion dictionaries, respectively.

The m 2n-dimension vectors along with their writer-reader emotion labels are used to learn a 4-class SVM classifier. During testing, each message is represented as a 2n-dimension vector, which is the input to the 4-class SVM classifier. The classifier outputs the prediction, which is (*pos*, *pos*), (*pos*, *neg*), (*neg*, *neg*) or (*neg*, *pos*).

An alternative solution is to compose two binary SVM classifiers for writer and reader emotion prediction, respectively. Each training instance corresponds to two n-dimension vectors, (w_{11}, w_{12}, …, w_{1n}) and (w_{21}, w_{22}, …, w_{2n}) with writer emotion label e_1 and reader emotion label e_2, respectively. The weights are determined in the same way as above. For each testing message, the writer emotion prediction is made first and then the reader emotion prediction.

4.2 Experimental Results

We use 10-fold cross-validation for evaluation. Table 2 shows the accuracy of the 4-class SVM classifier and two binary classifiers. In the baselines, the accuracies are 50%, 50% and 25%, since the emotion proportions of each dataset are near equal. All the proposed emotion models are higher than their corresponding baselines significantly. It shows that sentiment word mining in Section 3.2 is useful. The 4-class model outperforms the 2-class model in reader emotion prediction and writer-reader emotion pair prediction. The 2-class model outperforms the 4-class model for the writer emotion prediction task. Paired t-tests show that the performance differences between these two models for the 3 prediction tasks are all significant with p-values 0.017, 0.023, and 0.009, respectively.

Prediction Task → Prediction Model↓	Writer emotion	Reader emotion	Emotion pair
Baseline	50.00%	50.00%	25.00%
4-class classifier	62.04%	**63.85%**	**40.86%**
2-class classifiers	**64.23%**	62.18%	38.31%

Table 2: Accuracy of emotion prediction

5. Conclusion and Future Work

In this paper, we employ a log relative frequency ratio to mine sentiment words, and then use these words to predict emotion transition by building 4-class and 2-class SVM classifiers. The results show the sentiment word mining method is useful. The models considering both features from writer and reader emotion dictionaries are more effective to predict emotion transition, and the models considering features from writer emotion dictionary only is more proper to predict writer emotion.

As shown in our emotion transition analysis, we can further understand how writer emotions transition to reader emotions by examining the mined sentiment words. Some mined words do not carry emotion features, and thus are time-dependent topical words. They can affect the performance of emotion prediction. In the future work, how to distinguish them automatically from common emotion words will be investigated. Datasets that cover a longer period of time will also be collected.

6. Acknowledgements

Research of this paper was partially supported by National Science Council (Taiwan) under the contracts 98-2221-E-002-175-MY3 and 99-2221-E-002-167-MY3, and the Excellent Research Projects of National Taiwan University.

7. References

S. Aman, and S. Szpakowicz. (2007). Identifying Expressions of Emotion in Text. In *Proceedings of 10th International Conference on Text, Speech and Dialogue*. *Lecture Notes in Computer Science* 4629, pp. 196-205.

H.Y. Lin, and H.H. Chen. (2008). Ranking Reader Emotions Using Pairwise Loss Minimization and Emotional Distribution Regression. In *Proceedings of 2008 Conference on Empirical Methods in Natural Language Processing*. pp. 136-144.

H.Y. Lin, C.H. Yang, and H.H. Chen. (2008). Emotion Classification of Online News Articles from the Reader's Perspective. In *Proceedings of International Conference on Web Intelligence*. pp. 220-226.

J. Mei, Y. Zhu, Y. Gao, and H. Yin. (1982). *tóngyìcícílín*. Shanghai Dictionary Press.

G. Mishne. (2005). Experiments with Mood Classification in Blog Posts. In *Proceedings of 1st Workshop on Stylistic Analysis of Text for Information Access.*

B. Pang, L. Lee, and S. Vaithyanathan. (2002). Thumbs up? Sentiment Classification Using Machine Learning Techniques. In *Proceedings of 2002 Conference on Empirical Methods in Natural Language Processing*. pp. 79-86.

J. Wieb. (2000). Learning Subjective Adjectives from Corpora. In *Proceedings of 17th Conference of the American Association for Artificial Intelligence*. pp. 735-740.

C.H. Yang, H.Y. Lin, and H.H. Chen. (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*. pp. 287-290.