IJCLA VOL. 4, NO. 1, JAN-JUN 2013, PP. 159–173
RECEIVED 29/12/12 ACCEPTED 11/01/13 FINAL 15/05/13

Twitter Emotion Analysis in Earthquake Situations

BAO-KHANH H. VO AND NIGEL COLLIER

National Institute of Informatics, Japan

ABSTRACT

Emotion keyword spotting approach can detect emotion well for explicit emotional contents while it obviously cannot compare to supervised learning approaches for detecting emotional contents of particular events. In this paper, we target earthquake situations in Japan as the particular events for emotion analysis because the affected people often show their states and emotions towards the situations via social networking sites. Additionally, tracking crowd emotions in the Internet during the earthquakes can help authorities to quickly decide appropriate assistance policies without paying the cost as the traditional public surveys. Our three main contributions in this paper are: a) the appropriate choice of emotions; b) the novel proposal of two classification methods for determining the earthquake related tweets and automatically identifying the emotions in Twitter; c) tracking crowd emotions during different earthquake situations, a completely new application of emotion analysis research. Our main analysis results show that Twitter users show their Fear and Anxiety right after the earthquakes occurred while Calm and Unpleasantness are not showed clearly during the small earthquakes but in the large tremor.

KEYWORDS: *Twitter, Social media, Emotion recognition, Sentiment analysis, Earthquake, Japan*

1 Introduction

Sentiment analysis and emotion analysis have been increasingly studied in recent years thanks to the population of big text data. Although both emotion analysis and sentiment analysis apply psychology and cognitive science to computer science applications, emotion analysis mainly targets the fine-grained emotions while sentiment analysis detects simple attitudes such as positive and negative [1]. According to Scherer's typology of affective states [2], emotion is a relatively brief episode of synchronized response to the evaluation of an external or internal event as being of major significance; attitude is relatively enduring, affectively coloured beliefs, preferences, and predispositions towards objects or persons. Hence, emotion analysis is more appropriate for particular events rather than polarity sentiment analysis. Accordingly, our work addresses emotion analysis in different earthquake events for tracking and comparing the emotion variations during these events.

Tracking crowd responses, especially opinions and emotions towards an earthquake could provide valuable situational awareness for not only authorities to manage bad situations but also for psychological scientists to understand human behaviors in such situations. When a natural disaster like an earthquake occurs, the public agencies need the up-to-the-minute the affected people's responses to tailor emergency warnings, to aid the victims, and to calm down the public anxiety. With the strong development of social media, tracking emotions becomes easier, faster and more reliable than using the traditional public surveys or polls [3]. The social media service in our research is Twitter that allows users to send and read instant text-based messages or "tweets". This work analyzes Japanese tweets in Tokyo for emotion tracking during earthquakes because of the following reasons:

- Japan is the third in the world for total Twitter usage in 2012, and Tokyo is one of the top three cities in terms of tweets.[1]
- Japan often encounters earthquakes. The Great East Japan Earthquake in 2011 affected millions of people in Japan including Tokyo.
- As Japanese does not contain white space between words like English, Twitter users can convey more information within 140 characters of a tweet in Japanese than English.

[1] See http://techcrunch.com/2012/07/30/analyst-twitter-passed-500m-users-in-june-2012-140m-of-them-in-us-jakarta-biggest-tweeting-city/

We track emotions of the four earthquake dates by the Coordinated Universal Time (UTC): March 11th, April 7th, April 11th and July 10th, 2011 for the sake of comparison between earthquakes. To analyze emotions during earthquake dates, the first important task is to identify tweets related to the earthquake. These earthquake concerned tweets can be recognized in the Twitter data by our proposed supervised classification method. For emotion analysis task, we annotated the training data to Calm, Unpleasantness, Sadness, Anxiety, Fear, and Relief emotion. We use another supervised method for emotion recognition and then plot the categorized emotions into the time interval of each earthquake. Various features and machine learning models are used for both classification methods for selecting the best features and models.

To the best of our knowledge, our work is the first research upon emotion recognition and tracking for Twitter data in earthquake situations whereas the earthquake related Twitter analysis works do not consider the emotion aspects of users. Nevertheless, the choice of emotions and the classification methods with appropriate features for Japanese social media are the first contributions for Twitter emotion analysis applications.

2 RELATED WORK

Sentiment analysis and emotion analysis are the tasks of identifying the attitude and emotion classes of the investigated document [1]. The word "document" we use here has a general meaning as it can refer to a linguistic unit including a single sentence, a paragraph, and a document of many paragraphs. There are three main approaches used for emotion and attitude identification:

- Textual keyword spotting approach: Using a set of emotion words mostly adjectives and adverbs defined by specific lexical resources like Google Profiles of Mood States [4], Linguistic Inquiry and Word Count dictionaries [5] to select documents such as tweets and Facebook statuses containing such keywords in distinct emotions. This method is used widely for social media due to the large scale and the noise of social network data because it can remove a fair amount of irrelevant documents. While this approach is directly applicable for English with no modification in adjectives and adverbs, it is difficult to be applied for Japanese, an agglutinative language.
- Rule-based linguistic approach: Each sentence is processed in stages, including symbolic cue, abbreviations, sentence parsing, and word /

phrase / sentence-level analysis [6]. This rule-based approach often has limitations due to the diversity of natural language, especially the language in social media.

– Feature-based classification approach: This empirical approach has been used from the first applications of sentiment analysis on movie reviews to current sentiment and emotion analysis applications on social media [7]. Alm et al. [8] consider determining emotion of a linguistic unit is a multi-class classification problem. This supervised learning approach generates a function that maps linguistic units to the desired emotion by looking at the features derived from linguistic unit-emotion examples of the function. The features can be n-grams, bag of words, or Twitter features such as re-tweets, hashtags, replies, punctuations, and emoticons [9].

There are a few works on sentiment analysis in crisis contexts similar to earthquake events which are hurricanes [3], and gas explosion [10]. These works also use the feature-based classification method for English sentiment analysis. Mandel et al. [3] experiment with features: two tokenizer alternatives, stop word removal, frequency pruning, worry lexicon, humor lexicon, and emoticon; and classifiers: Maximum Entropy, Decision Tree, and Naive Bayes to choose the best features and model for classifying tweets to Irene Hurricane concerned or unconcerned. Nagy and Stamberger [10] combine the available English sentiment data comprising SentiWordNet, emoticons, AFNN for classification by Bayesian Network.

Most of Japanese emotion analysis applications are for blogs [11] and Japanese emotion-provoking sentences collected in the Web [12], not Twitter data that is shorter and less information. These applications restrict the data to explicit emotive sentences that contain emotive words, not inferred emotive sentences. In another effort to analyze emotions automatically for Japanese, a Japanese WordNet Affect is being developed [13], but it is not completed yet.

For Twitter analysis in earthquake situations, the available works only concentrate on earthquake and rumor detection [14], not sentiment analysis in earthquake situations.

Due to the listed disadvantages and advantages of the related works as well as our wish to handle inferred emotive tweets, the feature-based classification method is feasible for our purpose of emotion analysis in earthquake situations.

3 METHODOLOGY

3.1 *Emotion selection for earthquake situations*

Plenty of English emotion analysis researches classify documents [1, 8] to Ekman's 6 basic emotions [15]: Surprise, Happiness, Anger, Fear, Disgust, and Sadness. Meanwhile, some Japanese works [11] base on Nakamura's Japanese emotion dictionary [16] with 10 emotion types: Excitement, Shame, Joy, Fondness, Dislike, Sorrow, Anger, Surprise, Fear, and Relief; and Tokuhisa et al. [12] use 10 emotion classes: Happiness, Pleasantness, Disappointment, Unpleasantness, Loneliness, Sadness, Anger, Anxiety, Fear, and Relief from Teramura dictionary [17]. It is clear that the emotions used by Tokuhisa et al. are more separate than Ekman's emotions. Besides, Nakamura's emotions have Excitement, Joy, Fondness are quite similar and along with Shame, they are not appropriate in earthquake contexts. Therefore, we construct our emotions from 10 emotions of Tokuhisa et al.

For the reason of emotion analysis in earthquake situations, we need to adjust the 10 emotion classes to match such situations. We remove Happiness and Pleasantness because we think Happiness and Pleasantness are too positive to fit in negative situations like earthquakes. In order to show positive emotion in such situations, Calm emotion may be the most appropriate one. Anger, Disappointment, and Unpleasantness can be grouped into Unpleasantness. Sadness can include Loneliness. To sum up, we use 6 emotion classes for data annotation and classification: Calm, Unpleasantness, Sadness, Anxiety, Fear, and Relief. This choice of emotions is also verified in the data.

3.2 *Earthquake related tweet identification and emotion analysis methods*

Before classifying emotions in the crisis situation, we need to perform the task of selecting only the tweets related to the earthquake. We call these tweets as Concerned tweets, the others are Unconcerned, and the task of filtering Concerned tweets as earthquake related tweet identification. In order to filter the Concerned tweets out of the Unconcerned ones, instead of using the simple keyword spotting approach, we apply the feature-based classification method because the word spotting approach can not cover all the tweets related to the earthquake as the empirical method does. For example, it is difficult to redefine the word list for tweets related to the food shortage, family's safety, etc. For the purpose of emotion comparison after earthquakes, we select the tweets right after the earthquakes on the earthquake dates. Similarly, we apply the feature-based classification method for the emotion analysis task for the reasons mentioned in Section 2.

The two tasks can share the same linguistic features: n-grams, bag-of-words, stop-word removal, and emoticons. The purpose for using various features is comparing applying them in classifiers to select the best feature and classifier for the two classification tasks.

FEATURES

Emoticons Japanese emoticons or better known as kaomoji are much more complex than English emoticons, thus it is hard to fully detect emoticons in Japanese text [11]. For serving our main purpose of classification, we detect emoticons in tweets by using complex regular expressions instead of the techniques mentioned in [11]. The detected emoticons are used for two kinds of testing features: a) they are removed from tweets to become no-emoticon feature or b) they are grouped to 10 emotions of Nakamura [16] thank to the available of CAO preliminary emoticon lists mentioned in [11]; we call this feature as grouped-emoticon feature.

Bag of words Japanese words are not separated by space. Therefore, we need to use a Japanese morphological processing tool to segment words. However, because Twitter language is informal with many new words and slangs, morphological processing tools can not correctly analyze morphemes of tweets. To solve a part of this problem, we need to make a normalization dictionary for convert the wrongly segmented words, out-of-dictionary words to the correct words for adding unique words to the bag. The bag-of-word feature has two options: removing or not removing stop words beside the options of including emotions of emoticons or removing all emoticons.

N-grams Due to the bad performance probability of the Japanese morphological analysis tool for Twitter data, we think we should use n-gram features. We intend to try the uni-gram, bi-gram, and tri-gram features separately as well as the combinations of uni-gram and bi-gram; uni-gram, bi-gram and tri-gram. All of these n-gram features have the options of including emotions of emoticons or removing emoticons.

Stop-words We use the stop-word removal feature accompany with the bag-of-word feature. For the reason that there is no official stop-word list in Japanese, we need to find the most appropriate list for our purpose.

EARTHQUAKE CONCERNED/UNCONCERNED TWEET CLASSIFICATION We use different features in Section 5.1 with Support Vector Machine (SVM), Naive Bayes, Multinomial Naive Bayes (MNB), Decision Tree (J48), and Maximum Entropy (MaxEnt) models for the purpose of comparison to select the best features and models.

EMOTION CLASSIFICATION For this multi-class classification task, we use the Sequential Minimal Optimization (SMO) and Multinomial Naive Bayes (MNB) models with features in Section 5.1 for classification.

4 DATA

4.1 *Data collecting and pre-processing*

We collected Twitter data for five months, starting from March 10th 2011 to July 31st 2011 using Twitter API[2] with the geolocation feature set to track messages originating within Tokyo because Tokyo, the capital city with biggest population and Twitter users of Japan, is near Tohoku area where the 2011 Great East Japan Earthquake occurred which also affected Tokyo residences. In order to select only useful informations for emotion analysis, we need to pre-processed the tweets. From the original tweets, we parsed them and changed their encoding to UTF-8. From this parsed tweets, we selected only the tweets on the days of the big earthquakes or aftershocks according to `http://en.wikipedia.org/wiki/ List_of_foreshocks_ and_aftershocks_of_the_2011_Tohoku_earthquake`. We chose the 6 earthquakes and aftershocks from this site which shown in Table 1. All of the tweets were daily selected by UTC time.

As the result, the corpus consists of 4 files for 4 days of 6 earthquakes. This corpus has totally 110,715 tweets. For the purpose of processing only Japanese tweets, we selected only Japanese tweets by a language detection program. We then removed the spam tweets from these Japanese tweets. The spam tweets in our research context are advertising

[2] http://dev.twitter.com/

Table 1. Significant earthquakes in Tohoku area from March to July, 2011

Japan Time	Magnitude	Intensity (shindo)
2011-03-11 14:46	Mw 9.0	7
2011-03-11 15:15	Mw 7.9	upper 6
2011-03-11 15:25	Mw 7.7	4
2011-04-07 23:32	Mw 7.1	upper 6
2011-04-11 17:16	Mw 6.6	lower 6
2011-07-10 09:57	Mw 7.0	4

tweets, automatic tweets generated by applications, and location check-in information because they do not show human's emotions. Totally, the number of the spam-filtered Japanese tweets is 70,725. Table 2 shows the concrete spam-filtered Japanese tweet numbers of 4 days.

Table 2. Spam-filtered Japanese tweet numbers in 4 days

UTC Date	Tweet number
2011-03-11	19,420
2011-04-07	15,893
2011-04-11	16,700
2011-07-10	18,712
Total tweets	70,725

4.2 *Data annotation*

A part of the spam-filtered-Japanese tweets was annotated with 6 emotions and earthquake not-related tweets following our emotion definitions and annotation guideline. The tweets annotated with 6 emotions are considered as Concerned tweets while earthquake not-related tweets are Unconcerned tweets. The Concerned data includes messages directly convey the emotions with obvious emotion words and messages with inferred emotions. The tweets were annotated to 6 emotions by two annotators. The annotator chose only one emotion class or Unconcerned class for a tweet. The inter-annotator agreement was calculated using Fleiss' Kappa statistics [18]. The measured Kappa coefficients for Concerned and 6 emotions are 0.96 and 0.684, respectively. Only the tweets annotated with the same class were examined as the training tweets.

Table 3 shows the distribution of the annotated data in which the sum of 6 emotion tweets equals the number of Concerned tweets and Unconcerned tweets for the sake of balancing the training set.

Table 3. Distribution of annotated data

Unconcerned	Concerned	Calm	Unpleasantness	Sadness	Anxiety	Fear	Relief
1905	1905	155	310	4	580	635	221

From observing the annotated data, we found that Sadness has only 4 tweets over 1905 tweets. There are two reasons of this low appearance of Sadness tweets in the annotated data: a) The unbalanced selection of data; b) Not many people in Tokyo feel sad because they were not effected severely by the earthquake and tsunami like the people of the Tohoku area. Therefore, we remove Sadness from our training data set for emotion classification. The training data now has 5 emotion classes: Calm, Unpleasantness, Anxiety, Fear, and Relief. Accordingly, these 5 emotions are our targeted emotions for analysis.

5 Experiments and evaluation

5.1 *Experiment settings*

We implemented the proposed methods with following specifications:

Features

Emoticons Although we do not use the emoticon detection methods mentioned in [11], we could identify emoticons effectively with our regular expression for our data. As the Japanese emoticons often start with nonword characters and different kinds of brackets, the beginning of regular expression are the representation of non alphabetical and Japanese word character with a set of brackets. The central part of emoticons are three or more than three word characters. The ending parts of these emoticon regular expressions are similar to their starting parts.

After detecting emoticons in tweets, we remove them for testing nonemoticon feature or assign the emotions for these emoticons using the two emoticon lists of CAO: the list of full characters of each emoticon and the list of only three main characters of each emoticon. Firstly, we used the full character emoticon list for identifying emoticons that appear in this list. If the emoticons did not appear in this list, we then used the list of three main characters to replace the emoticons with special emotion names.

Bag of words Mecab[3], a Japanese dependency structure analyzer was used for word segmentation. We did not change the segmented words to the dictionary form. We only changed the wrongly segmented words to the correct ones by using our manual normalization dictionary. We selected the features with frequencies equal or greater than 5 for the experimental purpose and reducing the numbers of features.

N-grams We applied frequency pruning for bi-gram, tri-gram, the combination of uni-gram and bi-gram, and the combination of unigram, bi-gram and tri-gram. We selected the features with frequencies equal or greater than 5.

CLASSIFIER MODELS We use Weka [19], the collection of machine learning algorithms, for classification tasks. Options of all algorithms were set as default values for 10-fold cross-validation classification.

5.2 *Earthquake Concerned/Unconcerned tweet classification*

Table 4 shows a part of the classification results in Precision - Recall - F-measure order. List of models are in the left column while some features are in the first row. The first listed feature is bi-gram with removing emoticons out of tweets. The second feature is bi-gram with emoticons grouped into 10 emotions. The third feature is the combination of unigram, bi-gram and tri-gram with the emoticons grouped into 10 emoticons. All of these features were selected based on their appearance frequencies in the feature sets.

The best result (F-measure = 87.8) comes to the combination of unigram, bi-gram, and tri-gram with Multinomial Naive Bayes model.

5.3 *Emotion classification*

The combination of uni-gram, bi-gram, and tri-gram with MNB model again bring the best results. We classified 5 emotions with this feature

[3] http://code.google.com/p/mecab

Table 4. A part of earthquake Concerned/Unconcerned tweet classification results (Precision - Recall - F-measure)

Models	No emoticon - Bi	Grouped emoticons - Bi	Grouped emoticons - Uni-bi-tri
Naive Bayes	79.9 - 79.9 - 79.9	80.2 - 80.1 - 80.1	82.2 - 82.1 - 82.1
SVM	25.6 - 50.6 - 34	25.6 - 50.6 - 34	55.1 - 61.7 - 55.3
MaxEnt	69.6 - 69.4 - 69.3	67 - 66.7 - 66.7	55.1 - 61.7 - 55.3
J48	81.2 - 81.2 - 81.2	81.2 - 81.1 - 81.2	84.6 - 82.1 - 83.1
MNB	87.2 - 87.1 - 87	87.2 - 87.2 - 87.2	**87.9 - 87.8 - 87.8**

and obtained **64.7** and 62.2 F-measure for **MNB** and SMO, respectively. However, Calm emotion is classified with 22.4 F-measure by MNB and 22.3 by SMO. This springs from the fuzzy characteristics of this emotion class. This Calm class includes the tweets about the earthquake without distinctive emotions towards earthquake problems.

5.4 *Emotion analysis in earthquake dates*

We used the MNB model of uni-gram, bi-gram, and tri-gram feature for both classifying earthquake Concerned/Unconcerned tweet and emotions of tweets in earthquake dates. Because of the purpose of tracking emotions during earthquake situations, we only consider the emotions from the time when each earthquake occurred until the end of that day by UTC time. More clearly, we plot the time in Japan time from the period of each earthquake until 9 AM of the next day because Japan time is GMT+9. Emotions are tracked in 30 minute time unit. Figure. 1, figure. 2, and figure. 3 show the emotions plotted in March 11th, April 07th, April 11th, and July 10th, 2012.

Table 5. Pearson correlation coefficient of Fear and Anxiety

Date	Pearson coefficient
2011-03-11	0.85
2011-04-07	0.96
2011-04-11	0.96
2011-07-10	0.89

A clear notice of these earthquakes is that Fear emotion is always dominant when earthquakes started and rapidly decreases to nearly cor-

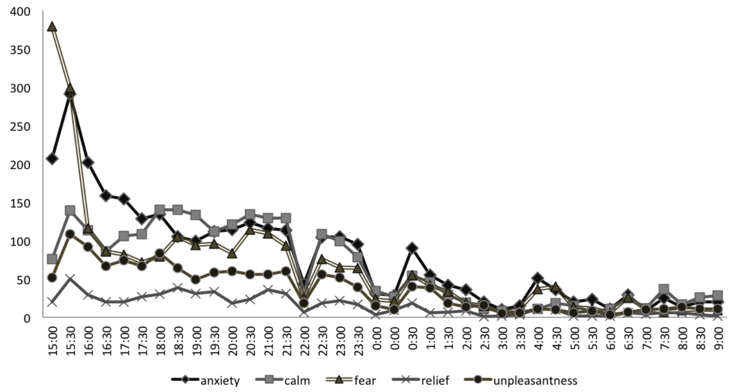

Fig. 1. Twitter emotions in March 11th, 2012

related with Anxiety emotion. Table 5 shows the Pearson correlation coefficient of Fear and Anxiety of 4 earthquake dates. The peaks of Fear and Anxiety mostly associate with the real earthquakes. For example, earthquakes in March 11th, 2011 that showed in Table 1 and a small earthquake around 8 AM of April 12th, Japan time. The earthquakes with higher intensity (March 11th, April 07th, and April 11th) result more tweets than the earthquake with lower intensity (July 10th). Because of the small amount of tweets in July 10th, the small peaks of Fear emotions (around 13:00, 16:00, and 21:00) are not really correlated with the real earthquakes. Therefore, we need to improve the Concerned tweet classification with other features such as the tweet amount in a specific time scale.

Except for March 11th when all the emotions significantly variate, other earthquakes show the low level of Calm, Relief, and Unpleasantness emotion because March 11th earthquake was the biggest earthquake brought various issues including the unpleasantness of transportations, phone communications, and the relief of surviving from this big earthquake. March 11th earthquake also has the Anxiety emotion has higher volumes than Fear emotion because of the intensity of this earthquake and other worries for the Fukushima Power Plant and tsunami. Calm emotion changes from lower than Anxiety and Fear from the earthquakes happened at nearly 15:00 to higher than these emotions at 18:00 that shows Twitter users became calmer after the first three hours of Anxiety and Fear.

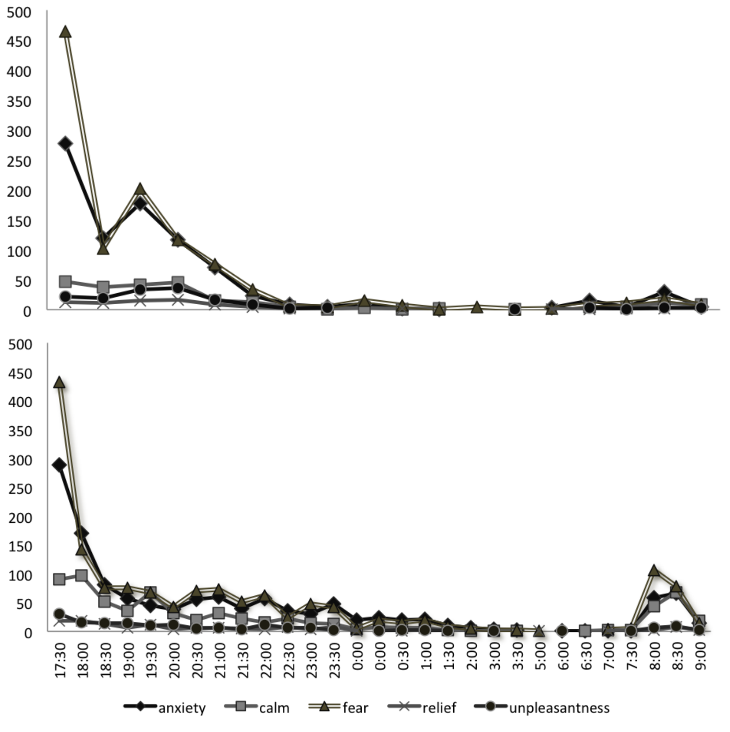

Fig. 2. Twitter emotions in April 07th (left) and April 11th (right), 2012

6 Conclusion

In this paper, we presented a novel application of Twitter sentiment analysis: tracking emotions in earthquakes for better managing the situations. To accomplish this purpose, we are the pioneer to propose the appropriate emotions for earthquake situations. We also propose the earthquake Concerned/Unconcerned tweet classification and emotion tweet classification which are completely different from the available works related to earthquake. Simple n-gram features are the best choice for classifying the agglutinative Japanese language and noisy Twitter language. Emotions in the time interval of earthquake dates reveal the insights of Twitter users during such hard time. Fear and Anxiety emotion always correlate with the occurrences of big earthquakes. Calm emotion will dominant after

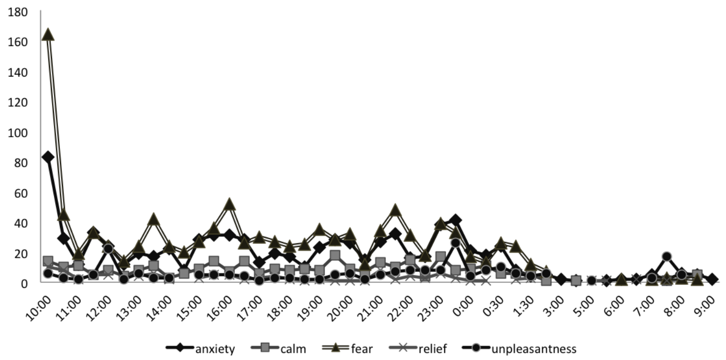

Fig. 3. Twitter emotions in July 10th, 2012

the first hours of Fear and Anxiety because of the earthquakes. Although Relief and Unpleasantness do not present the significant tweet amounts in earthquakes, they are important for the management purposes.

REFERENCES

1. Strapparava, C., Mihalcea, R.: Learning to identify emotions in text. In: 2008 ACM symposium on Applied computing, ACM (March 2008) 1556–1560
2. Scherer, K.R.: Emotion as a multicomponent process: A model and some cross-cultural data. Review of Personality and Social Psych **5** (1984) 37–63
3. Mandel, B., Culotta, A., Boulahanis, J., Stark, D., Lewis, B.: A demographic analysis of online sentiment during hurricane Irene. In: NAACL-HLT Workshop on Language in Social Media. (2012)
4. Bollen, J., Mao, H., Zeng, X.J.: Twitter mood predicts the stock market. Journal of Computational Science **2**(1) (2010) 1–8
5. Kramer, A.D.: An unobtrusive behavioral model of “gross national happiness”. In: 28th International Conference on Human Factors in Computing Systems, April 10-15. (2010)
6. Neviarouskaya, A., Prendinger, H., Ishizuka, M.: Affect analysis model: Novel rule-based approach to affect sensing from text. International Journal of Natural Language Engineering **17**(1) (2011) 95–135
7. Pang, B., Lee, L.: Opinion mining and sentiment analysis. Foundations and Trends in Information Retrieval **2**(1–2) (January 2008) 1–135
8. Alm, C.O., Roth, D., Sproat, R.: Emotions from text: machine learning for text-based emotion prediction. In: Human Language Technology Conference and Conference on Empirical Methods in Natural Language Processing. (2005) 579–586

9. Barbosa, L., Feng, J.: Robust sentiment detection on Twitter from biased and noisy data. In: 23rd International Conference on Computational Linguistics, COLING. (2010) 36–44
10. Nagy, A., Stamberger, J.: Crowd sentiment detection during disasters and crises. In: 9th International Conference on Information Systems for Crisis Response and Management, ISCRAM. (2012)
11. Ptaszynski, M., Dybala, P., Rzepka, R., Araki, K., Momouchi, Y.: YACIS: A five-billion-word corpus of Japanese blogs fully annotated with syntactic and affective information. In: 2nd Symposium on Linguistic and Cognitive Approaches To Dialog Agents, LaCATODA. (2012)
12. Tokuhisa, R., Inui, K., Matsumoto, Y.: Emotion classification using massive examples extracted from the Web. In: 22nd International Conference on Computational Linguistics, COLING. Volume 1. (2008)
13. Torii, Y., Das, D., Bandyopadhyay, S., Okumura, M.: Developing Japanese WordNet Affect for analyzing emotions. In: Workshop on Computational Approaches to Subjectivity and Sentiment Analysis, WASSA 2011, 49th Annual Meeting of the Association for Computational Linguistics (ACL). (2011) 80–86
14. Sakaki, T., Okazaki, M., Matsuo, Y.: Earthquake shakes Twitter users: Real-time event detection by social sensors. In: 19th International Conference on World Wide Web, WWW 2010, April 26-30. (2010)
15. Ekman, P.: Facial expression and emotion. American Psychologist **48**(4) (1993) 384–392
16. Nakamura, A.: Kanjo hyogen jiten (in Japanese). Tokyodo Publishing (1993)
17. Teramura, H.: Japanese Syntax and Meaning (in Japanese). Kurosio Publishers (1982)
18. Fleiss, J.L.: Measuring nominal scale agreement among many raters. Psychological Bulletin **76**(5) (1971) 378–382
19. Hall, M., Frank, E., Holmes, G., Pfahringer, B., Reutemann, P., Witten, H.I.: The weka data mining software: An update. SIGKDD Explorations **11**(1) (2009)

BAO-KHANH H. VO
NATIONAL INSTITUTE OF INFORMATICS,
2-1-2 HITOTSUBASHI, CHIYODA-KU,
TOKYO 101-8430, JAPAN
E-MAIL: <KHANHVO@NII.AC.JP>

NIGEL COLLIER
NATIONAL INSTITUTE OF INFORMATICS,
2-1-2 HITOTSUBASHI, CHIYODA-KU,
TOKYO 101-8430, JAPAN
E-MAIL: <COLLIER@NII.AC.JP>