Advanced Science and Technology Letters
Vol.51 (CES-CUBE 2014), pp.224-228
http://dx.doi.org/10.14257/astl.2014.51.51

Public Opinion Mining on Social Media: A Case Study of Twitter Opinion on Nuclear Power[1]

DongSung Kim[2] and Jong Woo Kim[2,3]

[2] 222 Wangsimni-ro, Seongdong-gu Seoul 133-791, Korea
{paulus82, kjw}@hanyang.ac.kr
[3] Corresponding Author

Abstract. Social media, including micro-blogs such as Twitter, have become main channels to communicate and share public social opinion among people. The opinion on nuclear power is an important social issue because nuclear plant construction needs national consensus. This paper proposes an opinion mining approach to monitor public sentiments on nuclear related issues using tweets on Twitter. The proposed process consists on (1) crawling related tweets, (2) text preparation, (3) sentiment dictionary construction, and (4) sentimental scoring. Based on experiment using nuclear related tweets in Korean between 2009 and 2013, we verify the usefulness on the proposed approach and confirm the changes on national opinion on nuclear generation depending on critical events such as Fukushima Daiichi nuclear disaster.

Keywords: public opinion mining, sentiment analysis, nuclear power

1 Introduction

The popularity of social media such as Facebook and Twitter turned them into main channels to communicate and share opinions on political, economic, social, and cultural issues. Even though social media contribute to changing consumers to prosumers (producers plus consumers), there are also some drawbacks on public opinion conversation and convergence such as fraudulent and biased messages, witch hunting, and extrusion of personal information. Reflecting the increase of interest on opinion on social media, there has been trials and experiments to monitor and analyze public opinion on specific issues on social media [4, 8].

Nuclear power is a significant national issue because it is a double-edged sword. Economic efficiency of power generation is the most important benefit of nuclear power. However, the potential risk of radiation leakage is the biggest difficulty of nuclear energy. So, nuclear power plant construction requires national consensus and agreement of residents in construction area. Traditional survey approach has been

[1] This work was supported by the Nuclear Power Core Technology Development Program of the Korea Institute of Energy Technology Evaluation and Planning (KETEP) granted financial resource from the Ministry of Trade, Industry & Energy, Republic of Korea (No. 201300000003242).

ISSN: 2287-1233 ASTL

used to monitor and investigate public opinion on nuclear power; however, it takes excessive cost and time. Opinion mining approach can provide an alternative way to monitor public opinion on nuclear power. This paper aims to suggest and verify opinion mining approach on nuclear power. The structure of the paper is as follows. In section 2, the researches of opinion mining approaches to monitor customer opinion and public opinion are described briefly. In section 3, the proposed opinion mining approach will be presented. In section 4, the experimental results using Twitter data between 2009 and 2013 are provided to demonstrate the usefulness of the proposed approach. Section 5 includes conclusion remarks.

2 Related Works

Recently, there have been many researches to monitor public opinion and social trends [1, 6]. They include election prediction using Twitter data [3, 11], monitoring of customer sentiment on a certain brand [7], movie performance prediction using Twitter [2, 9], disease and disaster tracking using Internet information [10], and unemployment benefit prediction using Internet search information [5]. Especially, public opinion monitoring is useful in sensing public opinion trends and reduction of potential social risks and conflicts [8]. In this study, we mainly focus on public opinion on nuclear power.

3 Proposed Approach and Experimental Design

The proposed procedure for public opinion mining consists of four phases: (1) crawling social media data, (2) cleansing and preprocessing texts, (3) construction of a sentiment dictionary, and (4) tweets sentiment prediction. The detail of such experimental procedure is shown in Fig. 1.

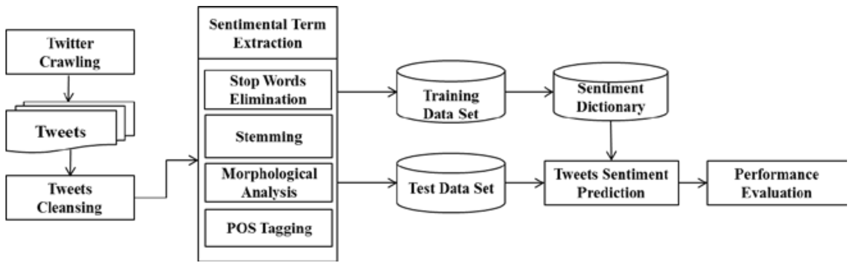

Fig. 1. Experimental Procedure

In the crawling phase, tweets including terms “Nuclear” or “Nuclear power” in Korean are crawled by a crawling tool, LocoySpider[2]. The tweets from 1st 2009 to

[2] www.locoy.com

December 31th 2013 are within the scope of crawling. After excluding irrelevant tweets on nuclear power issue, potential sentimental terms are extracted though stop words elimination, stemming, morphological analysis, and POS (Part of Speech) tagging using Korean morphological analyzer, KoNLP (Korea Natural Language Processing)[3]. Finally, nouns are extracted as potential sentimental terms.

The tweets from 2009 to 2011 are used to construct sentimental dictionary. The extracted nouns are reviewed by human evaluators and classified to positive terms and negative terms. The number of positive terms is 1,012 and that of negative terms is 3,291, which reflects that negative tweets are dominant in nuclear power tweets.

The tweets between 2012 and 2013 are used to evaluate the performance of sentimental classification. Sentimental classification is based on sentimental scores of tweets. The sentimental score of a tweet is calculated based on the number of positive terms and the number of negative terms in the tweet. The range of sentimental score is between -1 and 1. Tweets are classified as positive tweets when sentimental scores are greater than 0, and as negative tweets when those are less than 0. The tweets with 0 sentimental scores which mean that there are no positive terms and negative terms, or the number of positive terms and negative terms are the equal.

$$Sentimental\_Score(t) = \frac{N(Positive\_terms(t)) - N(Negative\_terms(t))}{N(Positive\_terms(t)) + N(Negative\_terms(t))} \quad (1)$$

4 Experimental Results

Table 1 exhibits the sentimental prediction results. To evaluate the accuracy, before applying proposed approach, human evaluators are classified tweets between 2012 and 2013 into three categories, Positive, Negative, and Neutral. The result shows that the proposed approach provides more than 50% prediction accuracy on positive and negative tweets.

Table 1. Sentimental prediction accuracy

Sentiment	2012 Accuracy Rate	2012 No of Tweets	2013 Accuracy Rate	2013 No. of Tweets
Positive	51.58%	948	50.55%	991
Negative	61.19%	2067	64.08%	2289
Neutral	38.96%	2066	21.67%	1375

To trace the changes on public opinion on nuclear power, we propose a measure, monthly Nuclear Opinion Index (NOI). As shown in formula (2), a monthly NOI is

[3] Heewon Jeon (2013). KoNLP: Korean NLP Package. R Package Version 0.76.9. http://CRAN.R-project.org/package=KoNLP

defined based on the number of positive tweets, negative tweets, and total tweets of the month. Fig. 2 shows monthly NOIs between 2009 and 2013. In Fig. 2, we can see the dramatic changes of nuclear power opinion in March 2010 due to Fukushima nuclear disaster.

$$Nuclear\ Opinion\ Index(m) = \frac{N(Positive\_tweets(m)) - N(Negative\_tweets(m))}{N(Total\_tweets(m))} \times 100 + 100 \tag{2}$$

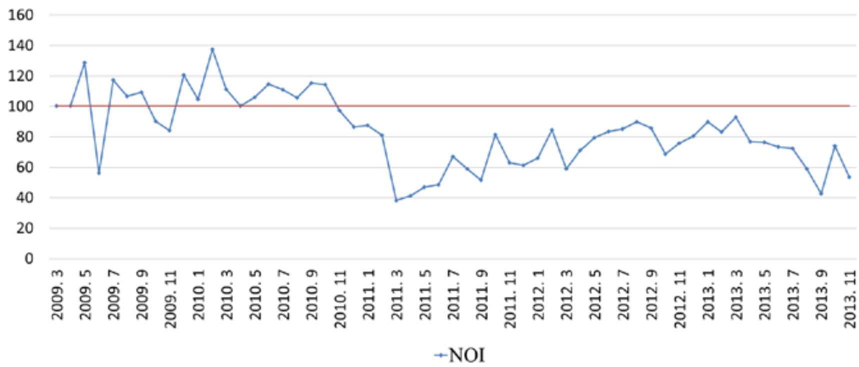

Fig. 2. Changes on monthly nuclear opinion index

5 Conclusions

In this paper, we propose public opinion mining approach to monitor nuclear power opinion on Twitter. The procedure consists of four phases: (1) crawling related tweets, (2) extracting potential sentimental terms, (3) building sentiment dictionary, and (4) tweets sentiment scoring and prediction. The experiments using tweets between 2009 and 2013 showed that the proposed approach provided acceptable performance on sentimental prediction. Also, NOI (Nuclear Opinion Index) is proposed to visualize the sentimental changes on nuclear power opinion.

References

1. Akcora, C.G., Bayir, M.A., Demirbas, M., Ferhatosmanoglu, H.: Identifying Breakpoints in Public Opinion. In: 1st Workshop on Social Media Analysis, pp. 62--66. Washington, DC (2010)
2. Baek, H.M., Ahn, J.H., Oh, S.W.: Impact of Tweets on Movie Sales: Focusing on the Time when Tweets are Written. J. ETRI. (2014)
3. Boutet, A., Kim, H., Yoneki, E.: What's in Your Tweets? I Know Who You Supported in the UK 2010 General Election. In: The International AAAI Conference on Weblogs and Social Media (2012)

4. Choi, H., Varian, H.: Predicting the Present with Google Trends. Technical Report, Google (2009)
5. D'Amuri, F., Marcucci, J.: "Google it!" Forecasting the US Unemployment Rate with a Google Job Search Index. In: Conference on Urban and Regional Economics (2009)
6. Diakopoulos, N., Shamma, D.A.: Characterizing Debate Performance via Aggregated Twitter Sentiment. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 1195--1198. ACM (2010)
7. Liu, Y., Huang, X., An, A., Yu, X.: ARSA: a Sentiment-aware Model for Predicting Sales Performance Using Blogs. In: Proceedings of the 30th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, pp. 607--614. ACM(2007)
8. Lee, C.H., Hur, J., Oh, H.J., Kim, H.J., Ryu, P.M., Kim, H.K.: Technology Trends of Issue Detection and Predictive Analysis on Social Big Data. J. Electronics and Telecommunications Trends. 28, 62--71 (2013)
9. Rui, H., Liu, Y., Whinston, A.: Whose and What Chatter Matters? The Effect of Tweets on Movie Sales. Decision Support Systems. 55, 863--870 (2013)
10. 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, pp. 851--860. ACM (2010)
11. Tumasjan, A., Sprenger, T.O., Sandner, P.G., Welpe, I. M.: Election Forecasts with Twitter How 140 Characters Reflect the Political Landscape. J. Social Science Computer Review. 29, 402--418 (2011)

