



Forecasting the daily outbreak of topic-level political risk from social media using hidden Markov model-based techniques



Jong Hwan Suh *

Research Center for Future Strategy, Graduate School of Future Strategy, KAIST, 291 Daehak-ro, Yuseong-gu, Daejeon 305-701, Republic of Korea

ARTICLE INFO

Article history:

Received 17 July 2013

Received in revised form 20 August 2014

Accepted 22 August 2014

Available online 13 September 2014

Keywords:

Political risk

Social media

Topic extraction

Sentiment

Social network

Hidden Markov model

Markov switching model

ABSTRACT

Nowadays, as an arena of politics, social media ignites political protests, so analyzing topics discussed negatively in the social media has increased in importance for detecting a nation's political risk. In this context, this paper designs and examines an automatic approach to forecast the daily outbreak of political risk from social media at a topic level. It evaluates the forecasting performances of topic features, investigated among the previous works that analyze social media data for politics, hidden Markov model (HMM)-based techniques, widely used for the anomaly detection with time-series data, and detection models, into which the topic features and the detection techniques are combined. When applied to South Korea's Web forum, Daum Agora, statistical comparisons with the constraints of false positive rate of <0.1 and timeliness of <0 show that, for accuracy, social network-based feature and, for sensitivity, energy-based feature give the best results but there is no single best detection technique for accuracy and sensitivity. Besides, they demonstrate that the detection model using Markov switching model with jumps (MSJ) with social-network based feature is the best combination for accuracy; there is no single best detection model for sensitivity. This paper helps make a move to prevent the national political risk, and eventually the predictive governance benefits the people.

© 2014 Elsevier Inc. All rights reserved.

1. Introduction

Social media is an electronic communication platform that conveys content, generated and exchanged by the networks of users. It ranges in form and purpose; the ever-changing universe of social media includes blogs, e.g., Blogspot, microblogs, e.g., Twitter, discussion forums, e.g., Epinions, media sharing sites, e.g. YouTube, and social networks, e.g., Facebook (Auer, 2011; Kaplan and Haenlein, 2010). People daily use social media to share information, express their opinions, and construct their social networks; beyond that usage, social media is now a vital open platform for political and social innovations in a nation (Abbott, 2012). Hence, politicians and governments around the world are trying to take advantage of social media to hear the voice of people concerns, improve life for the people, and build national stability (Baxter and Marcella, 2012; Suh et al., 2010).

However, using social media is a double-edge sword because social media diffuses information and mobilizes the people, so topics, initiated by certain political events and proliferating on negative sentiments, in social media occasionally propel people into aggressive actions and violent demonstrations, e.g., South Korea candlelight protest of 2008 (Lee et al., 2010) and the United States Occupy Wall Street of 2011 (Penney and Dadas, 2013). Furthermore, political risk, aroused through social media, sometimes results in disruptive political and social changes, e.g., Jasmine's Revolution (Khondker, 2011). Hence, not only monitoring topics, pertaining to political risk, in social media but also forecasting the daily outbreak of political risk related to the topics is essential to avoid the national instability and disruptive political/social changes by well-timed responses.

Nevertheless, no study has been made for the task of forecasting the daily outbreak of political risk relating to the topics in social media. Besides, building the automatic approach for the forecasting task has problems and challenging issues, summarized into three matters as follows. First, although three

* Tel.: +82 42 350 4025; fax: +82 42 350 4020.
E-mail address: suhjonghwan@kaist.ac.kr.

topic features, i.e., energy-based, sentiment-based, and social network-based features, are investigated among the previous works that analyze social media data for politics, their usefulness has been uninvestigated so it is unclear which topic feature works better than the others for the forecasting task. Second, according to the literature review, hidden Markov model (HMM)-based techniques are expected to fit well to the forecasting task, but it is unknown which HMM-based technique outperforms the others for the forecasting task. Third, neither guidance nor reference is available for selecting a topic feature and a HMM-based technique to combine for the forecasting task.

Therefore, this paper suggests a research framework to find a better way in forecasting the daily outbreak of political risk from social media at a topic level. To explain, its research framework extracts the political risk-related topics (hereinafter, PRRTs) from the test bed social media. Then, it measures the three topic features of the extracted PRRTs. Using the topic feature value, measured up to time t , it estimates the daily outbreak of political risk at time t on the extracted PRRT with four kinds of HMM-based techniques. Next, it evaluates the forecasting performances of twelve detection models, resulted by combining the three topic features and the four HMM-based techniques, with respect to four metrics, mostly used for evaluating the anomaly detection techniques: false positive rate, timeliness, accuracy, and sensitivity. In addition, it statistically compares the forecasting performances with the other detection models.

To sum up the contributions of this paper, it is the first work to deal with forecasting the daily outbreak of political risk from social media at a topic level. As the first reference, it is helpful for researchers, governments, politicians, etc., concerned with the social media analytics and intelligence for monitoring and forecasting the nationwide political risk. In other words, applying its research framework to the real world test bed gives the practical guidelines by resolving problems and challenging issues, mentioned above. Eventually, it benefits the people by urging governments and politicians to make a move to prevent the national political risk, i.e., predictive governance.

The rest of the paper is organized as follows: [Section 2](#) briefly introduces and reviews the relevant literature. [Section 3](#) outlines the research framework, and explains it in detail. Subsequently, [Section 4](#) demonstrates the results of applying the suggested research framework to the Web forum of South Korea, Daum Agora, chosen as a test bed, and the forecasting performances of topic features, detection techniques, and detection models combining them are evaluated with statistical comparisons. [Section 5](#) discusses the application results. Finally, [Section 6](#) concludes the paper with a reflection on limitations and further works.

2. Literature review

2.1. National stability and political risk

On portraying a nation's stability, two common viewpoints are the measurements of the risk, i.e., country risk, and the measurement of capability to manage the risk, i.e., good governance. First, country risk is measured by practitioners through a daunting task in their selection of variables and evaluation systems to represent and interpret the various

economic, social, and political factors. In a business climate, assessing country risk is a crucial area because country risk affects all firms in a given country and encourages foreign investments (Oetzel et al., 2001). Previous researches flourish on identifying the important factors that affect country risk, and political risk has been recognized as the most influential factor among various economic, social, and political factors that can help in predicting country risk (Vij, 2005). Second, good governance manifests itself in political risk, rule of law, control of corruption, government effectiveness, regulation quality, and voice and accountability of those in power. Good governance can improve the performance of firms whereas poor governance is a major obstacle to the economic development. Like country risk, previous researches show that political risk is more important to good governance than other indicators (Ngobo and Fouda, 2012). Thus, political risk is the most important factor in measuring both country risk and good governance.

Political risk is explained from the aforementioned two common viewpoints on representing a nation's stability. First, from the perspective of country risk, political risk is defined as the impact of politics on markets for foreign investors (Bremmer, 2005). In other words, political risk is the risk or probability of occurrence of some political events that will change the prospects for the profitability of a given investment (Vadlamannati, 2012). Second, in a viewpoint of good governance, political risk is political instability and the presence of violence, and it refers to the perceptions of the likelihood that the government will be destabilized or overthrown by unconstitutional or violent means, including political violence and terrorism (Ngobo and Fouda, 2012). Putting together these two viewpoints on political risk, broadly political risk is defined as the political instability and presence of violence in a nation, namely political protest (Bremmer, 2005; Click, 2005).

2.2. Social media analytics and intelligence for politics

[Table 1](#) explains previous studies on social media analytics and intelligence for politics, grouped by their research topics into four categories: forecasting polling and election, investigating leaning and polarization of users, modeling user behavior and information diffusion, and mining opinion and mood. Most of them are limited to checking the potential of exploiting social media in aiding decisions or making predictions about politics. Hardly any previous work systematically helps to forecast the daily outbreak of political risk from social media. However, the informational use of social media for news has a significant and positive association with both online and offline political participations (de Zuniga, 2012; Lim and Golan, 2011). In addition, information and opinion, shared and shaped in social media, influence views, and incite political actions and protests (Auer, 2011; Bennett, 2012; Kushin and Yamamoto, 2010; Payton and Kvasny, 2012; Se Jung et al., 2010; Tufekci and Wilson, 2012). To date, several media have also reported that social media influences the people to coordinate action and participate in political protests (Gross, 2011; Preston, 2011). For these reasons, social media analytics and intelligence, which develop and evaluates tool and frameworks to gain practical information from social media data and help in decision-making (Zeng et al., 2010), are required for this paper's task of forecasting the political risk of a nation from social media at a topic level.

Download English Version:

<https://daneshyari.com/en/article/7256639>

Download Persian Version:

<https://daneshyari.com/article/7256639>

[Daneshyari.com](https://daneshyari.com)