

Reading the public mind to emigrate, as expressed over tweets

Hyekyung Woo¹, Eunyoung Shim¹, Youngtae Cho^{1,2*}

¹School of Public Health, Seoul National University, Seoul 151-742, Korea

²Institute of Health and Environment, Seoul National University, Seoul 151-742, Korea

*Corresponding author. E-mail: youngtae@snu.ac.kr (Y Cho)

INTRODUCTION

It is well known from early research that motivation to migrate play an important role in migration intention and behavior. A conceptual model of migration decision-making has provided further theories that migration intentions and behaviors are significantly driven by motivations to migrate.(1) Most theories on migration are based on the assumption that people behave in ways that get benefits in the place to which they are going, without some strong causes of unhappiness where they are.(2) In the current age of globalization, motivations of international migration have increased and diversified. International migration is not only an individual phenomenon, but it is also influenced by social factors and issues. Thus, understanding the motivations is important because these have implications of the social situations.

Most of what is known about motivation of the international migrations has been derived by analyzing survey data, and only limited insights have been gained due to the nature of survey. Recently, unstandardized data collected from the Internet or social media have been used to gain insight into public mind, as expressed in large scale collections of daily posts. Social media, such as twitter, blogs, and Facebook can be channels for personal emotion, thinking, and desire expressions. Thus, it offers a novel opportunity to gain understanding of public mind to emigrate by analyzing weblog texts or documents on the social media platform. Here we explore how Koreans' public mind to migrate has changed

over time by tracking the usage of key words among users of social media services, especially *Twitter*. (We will offer lessons from further findings hereafter in this study.)

DATA and METHODS

The data are collected from posts that were recorded from 01 January 2014 to 30 June 2014 using the social media analyzing tool, SOCIALmetrics™ offered by Daumsoft (<http://www.daumsoft.com/eng/>). This system provides deep level keyword analysis and opinion mining for weblog posts in social media. The SOCIALmetrics™ platform processes texts collected from various social network services using the state-of-the-art natural language processing and text mining technologies. The NLP module splits an input text into sentences and segments wordforms contained in each sentence into a string of morphemes. The segmented morphemes are grouped into syntactic units via syntactic analysis. Once syntactic units are constructed, expressions denoting named entities such as people, location, and organization are recognized. Then association analysis is performed to identify tuples of *<topic keyword, associated keyword>*. The results of the whole analysis in time series fashion are served using API engine to be ready for various querying from users.

International migration weblog count was defined as the daily document frequency mentioning the Korean word '*imin*' (English word international migration) at least once. Trends of these words were analyzed in frequency for every 100,000 post of the daily Twitter posts mentioning the Korean word '*imin*'. And association analysis was performed to identify a variety of motivations of the public mind to migrate.

RESULTS and IMPLCATIONS

Figure 1 shows the trends of emigration-related minds in general population level from 01 January 2014 to 30 June 2014. Surprisingly, emigration-related keyword '*imin*' presented above showed great fluctuations over a month following the South Korean ferry disaster. And, the level was notably higher than its baseline level over a month. The finding suggests that the disaster can lead to immediate reactions in the public mind to emigrate. Our finding may have larger implication for demographic research related to out-migration motivation. In fact,

there had been a rise in the number of Korean emigrants over a three-year period happened a series of the largest manmade disaster, such as the Ferry Boat Sinks in 1993, Songsu Bridge Collapse in 1994 and Sampoong Department Store Collapse in 1995, for three years, from 14,477 to 15,917 Korean emigrants. We may see from words expressing cognitive activity over daily Twitter posts what the people are really thinking. Example would be the amount of words associated with cognitive activity (“emigration”) in the hope of a safer life.

This approach is more immediate and efficient in terms of cost and time than conventional approach through survey. Moreover, the information from continuous data rather than cross-sectional or sequential data is more useful both to understand public mind generation over time and to identify major determinants in public mind transition related to emigration. In looking toward the future, there is no doubt that social media data can play a groundbreaking role as information sources for demographic research. We also suggest that social media data containing emotions, thoughts and desires of peoples has perspectives and possibilities to monitor public mind and to understand determinants of public mind dynamic related to emigration. However, further challenges are empirical justifications of knowledge derived from social media and the development of sophisticated methodology for social media data analysis.

REFERENCE

1. J. R. Weeks, *Population : an introduction to concepts and issues* 9th ed. (Thomson Wadsworth Learning, Belmont, CA, 2005), vol. 9th ed.
2. H. P. van Dalen, K. Henkens, Explaining emigration intentions and behaviour in the Netherlands, 2005-10. *Population studies* **67**, 225 (Jul, 2013).

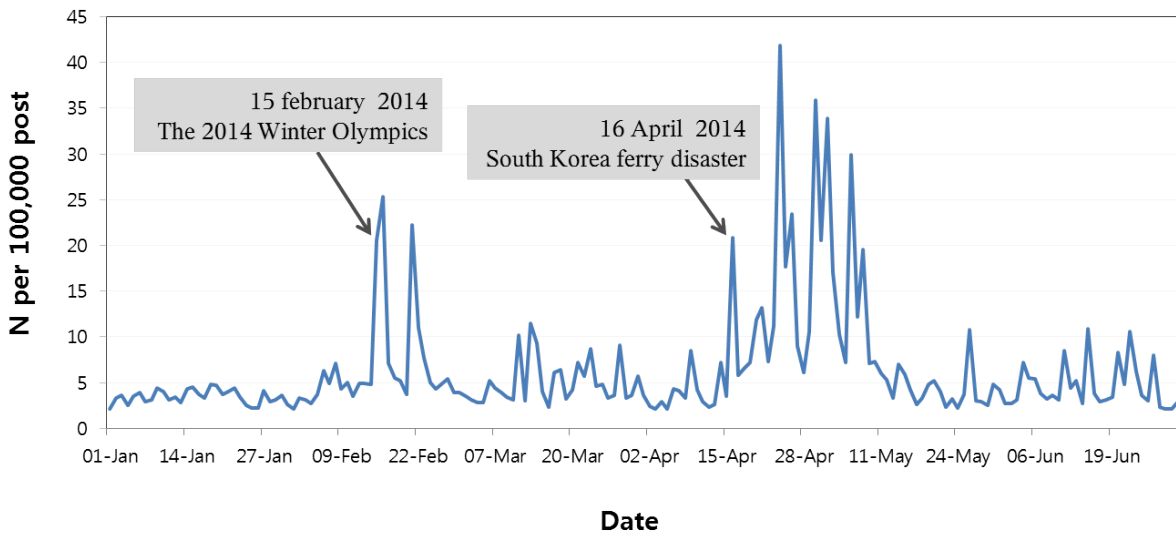


Figure 1. Trend of Public mind to emigrate detected from daily tweets. **Public minds trend detected from daily** tweets show public responses to South Korea ferry disaster. An emigration-related keyword ‘*imin*’ presented above showed great fluctuations over a month following the South Korean ferry disaster. And, the level was notably higher than its baseline level over a month. Trend of the keyword was analyzed in frequency for every 100,000 post of the daily Twitter posts mentioning the Korean words ‘*imin*’ (English word international migration). The data are collected from posts that were recorded between January 2014 and Jun 2014 using the social media analyzing tool, SOCIALmetrics™ offered by Daumsoft (<http://www.daumsoft.com/eng/>).

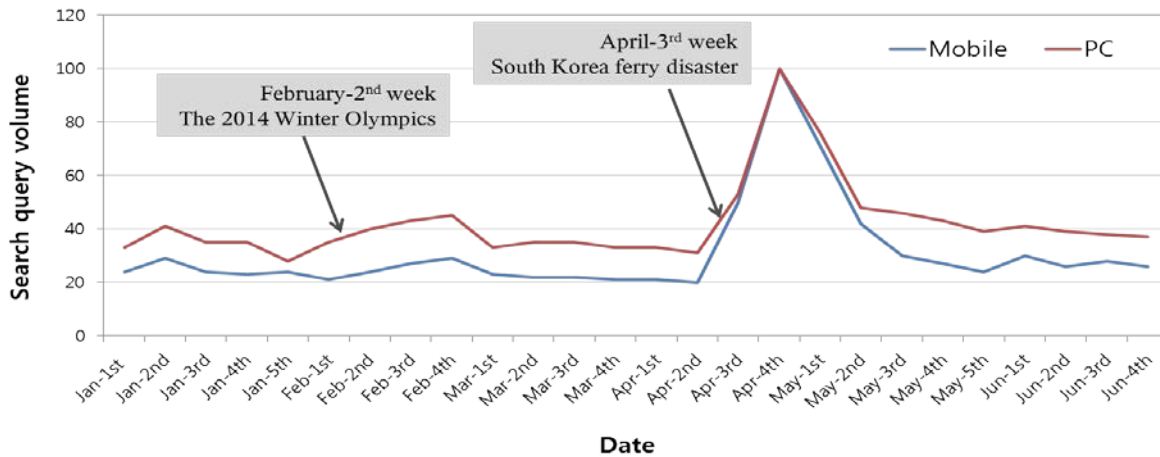


Figure S.1. Search Query Volume of ‘*imin*’ (English word international migration) between January 2014 and Jun 2014. It shows the trends of ‘*imin*’ query volume in the public. The South Korean ferry disaster led to immediate reactions. The emigration-related keyword presented had great fluctuations over a month following the disaster. The data to identify trends are collected from NAVER trend, offered by the biggest portal site in South Korea.