

## Social Media and Influenza Emergency: Content Analysis of Tweets during the 2015 MERS Outbreak in Korea

**Byoungkwan Lee**<sup>\*</sup>

(Professor, Department of Advertising & PR, Hanyang University)

**Hyun Jung Oh**<sup>\*\*</sup>

(Assistant Professor, Department of Health and  
Strategic Communication, Cha University)

**Hyunmi Baek**<sup>\*\*\*</sup>

(Assistant Professor, School of Media & Communication,  
Korea University)

**Sangrock Lee**<sup>\*\*\*\*</sup>

(Adjunct Professor, Department of Advertising & PR, Hanyang University)

The risk perception is a key issue for effective health risk communication. This study gathers and analyzes MERS-related tweets in order to monitor and understand the public perception on Twitter during the MERS outbreak of 2015 in South Korea. The main purpose of study is to identify and

---

\* gogreen@hanyang.ac.kr

\*\* janeoh82@gmail.com

\*\*\* lotus1225@korea.ac.kr, corresponding author

\*\*\*\* prstat@naver.com

describe the changes in the number of MERS-related tweets. Using Naïve Bayesian classifiers, this study investigates the public sentiments on Twitter and examines how the sentiments affect the spreading volume, duration, and speed of MERS-related tweets. The results show dramatic changes in the amount of MERS-related tweets during these sequences of time periods. The results of sentiment analysis also showed that the emotional response of the public on Twitter had changed in a variety of ways depending on the types of triggering events. Moreover, an emotional (positive or negative) message in social media seems to receive more attention and feedback than a non-emotional message and induce cognitive and arousal-related effects that trigger sharing behavior in social media communication.

*Key words : Social Media, Health Risk Communication, MERS in South Korea, Public Sentiments, Information Diffusion, Tweets*

## 1. Introduction

“What each man does”, Walter Lippmann (2017) argued, “is based not on objective knowledge, but on subjective pictures made himself or given to him.” In the context of public health emergency, individuals' actions against risk are vulnerable to their subjective and (or) intuitive perception rather than objective and (or) scientific facts. Indeed, the public perception of risk is in discord with its scientific assessment by experts as revealed by previous risk communication studies (Sjöberg,

2000; Slovic, 1987; Slovic, Peters, Finucane, & MacGregor, 2005). Risk, as Slovic (1987) mentioned, means different things to different people.

With the advancement of communication technologies, the public derives its perception of health from a variety of sources. With the recent advancement of social media, the public can gain grounds to judge the situation. Particularly in relation to health issues, the public exchanges opinions on social media, obtains information, and determines health trends. As more and more people tweet about health topics, Twitter becomes good sources of public opinion and sentiment (Ghiassi, Skinner, & Zimbra, 2013). Such information can be a great help to researchers with an enormous amount of dataset that can be easily accessed (Hornikx, & Hendriks, 2015). Therefore, it is important to understand public opinion in social media and establish an effective risk communication strategy.

This study gathered and analyzed MERS-related tweets in order to monitor and understand the public perception on Twitter during the MERS outbreak of 2015 in South Korea. MERS is one of the newly discovered respiratory diseases caused by a new kind of coronavirus, originally identified in Saudi Arabia in 2012. Most MERS patients developed severe acute respiratory illness with symptoms of fever, cough and shortness of breath, and about 3 to 4 out of every 10 patients diagnosed with MERS have died (CDC, 2016). In 2015, the MERS outbreak put South Korea in an immense social and economic turmoil. It resulted in 186 confirmed cases of infections with 38 deaths during the 218-day period from May 20th, 2015, when a person passing

through a region in the Middle East became the first confirmed patient to be diagnosed with the MERS, to the official date of termination on the 24th of December, 2015. This was the second largest MERS outbreak in the world, and the largest outbreak recorded outside Saudi Arabia (WHO, 2015).

Therefore, the aim of the present study is two-fold. Firstly, this study investigated the number of MERS-related tweets during the MERS outbreak in South Korea. Specifically, this study longitudinally tracks the changes in the number of MERS-related tweets during the MERS outbreak and identifies the reasons for such changes. Secondly, this study explored the public's responses on Twitter during the outbreak. Using Naïve Bayesian classifiers, the study investigates the public sentiments on Twitter during the MERS outbreak and, in addition, examines how the sentiments affect the spreading volume, duration, and speed of MERS-related tweets. Understanding the relationship between the characteristics of tweets and their spread pattern is very crucial to practitioners during public health crisis. Although many studies have been designed to reveal the relationship between the characteristics of messages and that of information spread (Hoang, & Lim, 2011; Son, Kim, & Lee, 2012; Yang, & Counts, 2010), few studies have examined how emotion, especially in tweets, affects the characteristics of RTs (Stieglitz, & Dang-Xuan, 2013). This study provides additional information to the body of knowledge on risk perception and risk communication by showing how public sentiment affects the pattern of information spread during the outbreak of a pandemic disease. Given

that controlling the spread of news on epidemic diseases on online social networking platforms during a public health emergency is one of the most important and challenging tasks for public health practitioners, the findings of the current study will provide implications for the practitioners.

## 2. Literature Review

The risk perception is a key issue for effective health risk communication. As many previous empirical studies and real-world incidents such as SARS (Severe Acute Respiratory Syndrome: hereafter SARS), H1N1, and MERS (Middle East Respiratory Syndrome Coronavirus: hereafter MERS) have shown (Brug et al., 2004; Gidengil, Parker, & Zikmund-Fisher, 2012; Seale et al., 2010), a better understanding of and response to the public perception of the risk is a crucial prerequisite for effective risk management during a public health emergency. The public's risk perception, which is driven from various sources, is a social and cultural construct beyond the individual (Weinstein, 1980). While traditional mass media has been one of the most powerful communication channels that shape public perception, the recent progress of communication technologies have made social media (i.e., Twitter, Facebook, FourSquare, Instagram etc.) as another important source that risk communicators should consider during the public health emergency. For example, during the H1N1 outbreak of 2010, more than

2 million tweets containing the keywords “H1N1”, “swine flu” or “swineflu” were posted, and health authorities used Twitter content as a complementary source of information to monitor public opinion and respond to public concerns in real time (Signorini, 2014).

Social media, however, is a double-edged sword for health risk communicators. Social media is a useful tool to disseminate and share health information in an efficient way during a public health emergency. Social media can be employed to “disseminate time-sensitive health information, promote information sharing to encourage behavioral changes (including corrective changes during potential health crises), be a platform for conversation between agencies and constituents (rather than just as an information provider) and allow the public to provide useful information and feedback” (Dosemagen, & Aase, 2016). At the same time, social media challenges the very tasks that health risk communicators perform. It may become difficult for practitioners to manage health crises if social media create biased public responses. During the 2010 earthquake in Chile, it was found that rumors posted and re-posted on Twitter contributed to an increase the sense of chaos and insecurity among the local population (Castillo, Mendoza, & Poblete, 2011).

Consequently, monitoring public opinion on social media and responding to it in a timely manner has become necessary for effective risk communication during a public health emergency. But, technically, traditional survey or content analysis as a monitoring tool for social media has its limits. As new data scientific approaches, namely,

“infoveillance” that mine, aggregate, and analyze the social media data in real-time are becoming available, researchers and practitioners are provided with an instantaneous comprehension of the public’s opinions and behavioral responses on social media (Chew, & Eysenbach, 2010).

As a result, sentiment analysis has attracted much attention from researchers. Sentiment analysis is a natural language processing technique that extracts subjective information from texts created using algorithms and can identify the degree of positive or negative emotion of the message (Beaunoyer, Arsenault, Lomanowska, & Guitton, 2017; Gohil, Vuik, & Darzi, 2018). Sentiment analysis methods are often used to predict various social phenomena. These techniques have been used to analyze and predict people’s behavior on elections and the launch of new products or services (Borras-Morell, 2015; Hirschberg, & Manning, 2015). It has also been used in the health field to identify the degree of sentiment (D’Alfonso, Santesteban-Echarri, Rice, Wadley, Lederman, R., & Alvarez-Jimenez, 2017; Greaves, Ramirez-Cano, Millett, Darzi, & Donaldson, 2013; Mazzocut, Truccolo, Antonini, & Tasso, 2016).

In healthcare, sentiment analysis can help you better understand how people talk and feel about a particular health topic or health condition. The use of sentiment analysis is particularly relevant to the social media context where social media has become a natural environment in which people can share and obtain health information (Gabarron, Lau, & Wynn, 2015; Li, Wang, Lin, & Hajli, 2018; Nath, Huh, Adupa, & Jonnalagadda, 2016). Social media content influences individual decision-making, as it not only finds information but also can grasp the

degree of positive or negative emotions in the message (Oyeyemi, Gabarron, & Wynn, 2014; Wynn, Oyeyemi, Johnsen, & Gabarron, 2017). Few publications currently study the use of emotional analysis in relation to MERS. Understanding the emotions expressed by social media users about MERS can provide information on the knowledge framework of risk awareness and risk communication.

### 3. Methods

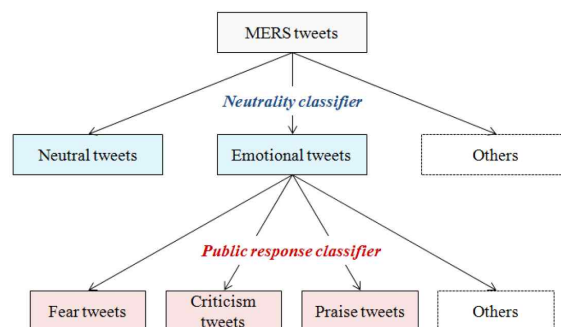
#### 1) Data Collection and Analysis

##### (1) Tweet Data

Twitter is a social networking and microblogging service that allows users to post real time messages, called tweets. Tweets are short messages, restricted to 140 characters in length. Due to the nature of this microblogging service (quick and short messages), researchers use Twitter as indicator of emotional expression. We collected 3,785,125 tweets, which included the keyword “MERS (middle east respiratory syndrome)” or “Korean term of MERS,” from May 5, 2015 to August 7, 2015. We collected daily tweets and user information on the Twitter API (application programming interface) by using the programming language Python. The metadata include tweet texts, times of tweets, retweet counts, user ID’s, etc. Tweets were searched continuously by including target keywords at the appointed time every day.

## (2) Classification of Tweets for Sentiment Analysis

We developed neutrality classifier and public response classifier to classify MERS-related tweets automatically by using machine learning approach (see <Figure 1>). There are two main approaches for automatic text classification: machine learning approach and lexicon-based approach. The lexicon-based approach uses a dictionary of words, whereas the machine learning approach requires a training stage based on manually coded training data set. Neutrality classifier was developed to classify tweets into neutral tweets and emotional charged tweets.(Why? Provide sources of information and explain/justify) During the MERS outbreak, the public responded emotionally, expressing criticism, fear, and praise etc. The contents of MERS-related emotional charged tweets were mainly classified into tweets to 1) criticize the government's incompetence in managing the risk, 2) express the fear of MERS infection, and 3) encourage medical staffs and emergency workers. Therefore, we tried to classify emotion charged tweets into three



<Figure 1> Tweet Classifiers: neutrality tweet classifier and public response tweet classifier

categories of major public reactions - fear, criticism, and praise - by using public response classifier (Citation needed; provide an explanation about this classifier). When a tweet includes the other emotion except for “criticism”, “fear”, and “praise”, it is classified as “others”.

There are several kinds of machine learning approach for text classification, such as decision tree, Naïve Bayes, Support Vector Machines, etc. (Zhang et al. 2011). By developing two Naïve Bayes classifiers (neutrality classifier and public response classifier), we classify tweets into neutral, fear, criticism, praise tweets. For the classification, two human coders, independently from one another, classified a total of 2,121 tweets, which have more than 200 retweets, into emotional /neutral tweets to develop neutrality classifier and criticism/fear/praise to develop public response classifier. The two coders produced highly congruent classifications with 95%. 1,999 tweets (1,500 tweets are used as the training set and 499 as the test set) for which identical classification occurs were utilized for manual labeling data. Second, the neutrality tweet classifier and the public response classifier were developed by applying strong independence assumptions to Bayes’ theorem (Bhargavi, & Jyothi, 2009) on the 1,500 manually coded tweets (training set). Morpheme (excluding postpositional particle) tokenize method and 2,000 feature sets with the highest accuracy among various combinations are used for the Naive Bayesian classifiers. Naïve Bayes is commonly used for text classification because it is simple and efficient (Oscar et al., 2017). To evaluate the accuracy of developed two classifiers, we used the 499 manually coded tweets (test set). The neutrality classifier showed

71.1% accuracy and the public response classifier exhibited 76.8% accuracy, which is consistent with other previous studies (Lee et al., 2011; Read, 2005; Rui, Liu, & Whinston, 2013). Finally, we applied the neutrality classifier and the public response classifier to a full set of collected tweets to classify the remaining tweets automatically.

### (3) Retweet Analysis and Characteristics of Information Spread

Acknowledging that retweeting can be a powerful mechanism to measure the public sentiments on Twitter and the spread of information (Salathé, & Khandelwal, 2011), the present study analyzed the diffusion of retweets to reveal how the sentiments of MERS-related tweets are related to the characteristics of information spread such as the amount of spread, speed of spread, and duration of spread. To analyze the retweets, this study identified the tweets that were shared over 50 times among the entire MERS tweets. The tweets were composed of 2,189,942 tweets holding 8,748 tweet contents, which was 57% of all MERS tweets (3,785,125). 1,653 of the tweet contents that did not fit the categories of sentiment analysis and drawing the characteristics of spread of RTs were excluded and the remaining 7,095 tweet contents were analyzed among 8,748 tweet contents retweeted over 50 times. By obtaining enough tweet data diffused profusely during MERS outbreak, the relationship between sentiment of message and characteristics of information spread could be analyzed.

Based on the previous studies (Hoang, & Lim, 2011; Son, Kim, & Lee, 2012; Stieglitz, & Dang-Xuan, 2013; Yang, & Counts, 2010),

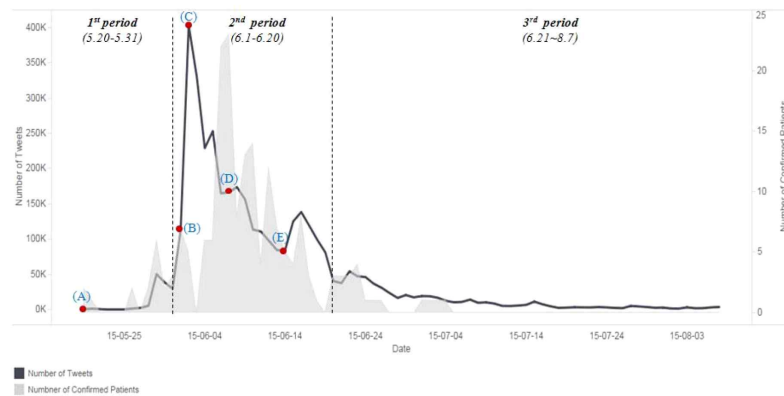
characteristics of information spread - diffusion volume, diffusion speed, and diffusion duration - were defined and measured as follows. The diffusion volume was operationalized as the total number of RTs spread during MERS outbreak; the diffusion speed was operationalized as the number of RT spread during one hour of peak time when the speed of spread was the fastest. This study defined the highest speed of RT per hour (number of RTs spread during one hour of peak time when the spread speed was the fastest) within the target period --when many RTs were made to form public opinion-- as diffusion speed. Diffusion duration was defined as the number of days from the first day when the number of RTs per hour was over 10 to the last day when the number of RTs per hour is over 10. The reason that the days with more than 10 RTs per hour was considered is to operationalize the duration with significant diffusion speed as the diffusion period. Additionally, the diffusion duration was calculated by considering the number of days with more than 3 or 5 RTs per hour, and similar results were produced as compared to the result, calculated with the number of days with more than 10 RTs.

#### **4. Results**

##### **1) MERS-related tweets during the MERS outbreak**

First of all, this study attempts to identify and describe the changes

in the number of MERS-related tweets during the MERS outbreak in South Korea. In order to achieve such purpose, the study longitudinally tracked the changes in the number of MERS-related tweets during the MERS outbreak and identified important events that triggered the dramatic increase in the number of tweets. For more in-depth analysis, the total period of the MERS outbreak was divided into three phases based on the severity of the issue itself. The first phase starts on May 20th, when the first MERS patient was reported. The second phase



<Figure 2> Changes in the Number of MERS Tweets During the Outbreak

- (A) The first MERS patient occurs (15-05-20: 494 tweets, 2 patients)
- (B) The first death occurs (15-06-01: 113970 tweets, 7 patients)
- (C) The first tertiary case occurs, Schools started to be temporarily closed due to the MERS outbreak (15-06-02: 403274 tweets, 5 patients)
- (D) The names of MERS exposed medical institutions were disclosed (15-06-07: 166401 tweets, 23 patients)
- (E) The first quaternary case occurred (15-06-14: 84648 tweets, 5 patients)

starts on June 1st, when the first death occurred. The third phase starts on June 21st. This is the first day when no confirmed case of MERS was announced since its outbreak. The third phase continues until the monitoring officially ended on August 7th. <Figure 2> shows the trends of MERS-related tweets in each of the three phases.

In order to identify the triggering events that contributed to the dramatic change in the number of tweets over time, the 100 most-retweeted tweets, spread most frequently in each phase, were extracted and classified by the topics. During the first phase, 100 most-retweeted tweets were shared 58,418 times in total, accounting for 45% of the entire tweets during the period. In the second phase, top 100 tweets were retweeted 353,840 times, accounting for 11.5% of the total tweets in the second phase. In the last phase, top 100 tweets were shared 114,551 times, equivalent to 20.1% of the total tweets during the period.

During the first phase, 18 patients were reported to have been diagnosed with MERS. The most retweeted topics were about (1) the state of MERS spread, (2) prevention, (3) a MERS patient who went on a business trip to China, (4) workplace issues, and (5) government's management. Tweets on the state of MERS spread were consistently updated and were shared the most. The most shared tweet was "A MERS patient is found at downtown Seoul. The fear is spreading" (710 retweets). Tweets on MERS prevention were mostly about complaining the lack of information on preventive measures. The most shared tweet also indicated the lack of information in general regarding the outbreak

(721 retweets). Tweets about a patient who traveled to China were also widely shared on Twitter; the press interview of the patient's wife was retweeted the most (624 retweets). Workplace issues were also discussed on Twitter, and the tweet "government should fine managers who do not report employee's infection" was shared the most (1,355 retweets). Finally, government's disease management was the most prominent agenda on Twitter. The most frequently retweeted message was about government's policy that would impose fines on people who refuse self-quarantine (1,066 retweets).

After the first death occurred, people became more aggressive in expressing their fear and anger on Twitter. At the second phase, the number of reported case had dramatically increased; in total, 154 people were reported to having been infected. Accordingly, many agendas and incidences were discussed on Twitter, including 'the prevention rule involving camels', 'Seoul Mayor's tweets', 'disclosing the hospitals where the MERS patient visited', 'government's response', 'temporary closure of schools', and 'deaths of medical staffs'. In particular, tweets on governmental preventive measures to enforce avoidance of contact with camels for MERS prevention (4,823 retweets), tweets of the Mayor of Seoul that urged the government to transparently disclose MERS-related information (8,950 retweets), and tweets criticizing the government for not revealing the name of the hospital where the MERS infection was found (7,934 retweets) were mostly retweeted.

In the final phase, the number of infected people had significantly decreased; in total, 10 people were infected with MERS. As the situation

became alleviated, tweets in the final phase shifted their focus towards evaluating government's MERS management and encouraging health workers. The most retweeted message on government was about government's disciplinary action on television programs that criticized government's incompetence in managing the risk (2,495 retweets). Also, more tweets were posted and shared to praise medical staffs and emergency workers. The most shared tweet was also about encouraging other users to cheer for and praise medical staffs (2,168 retweets).

## 2) Public sentiments and MERS-related tweets

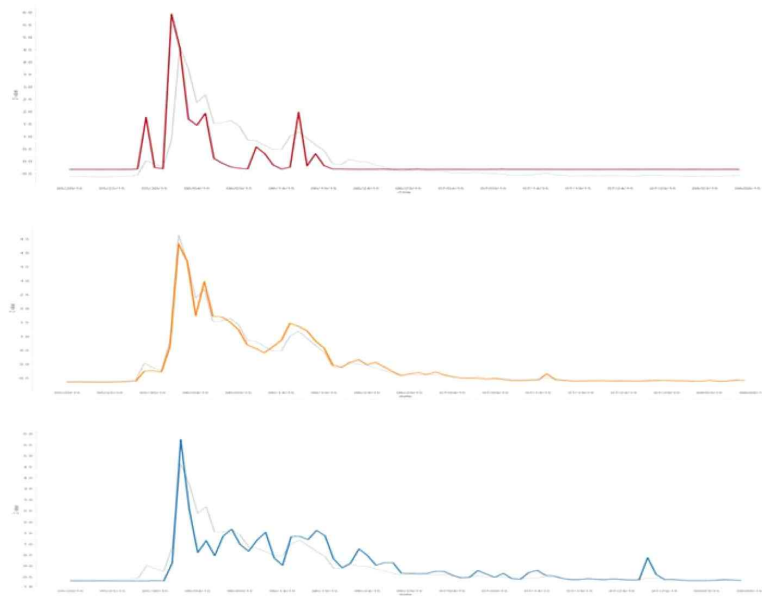
During public health crises such as the MERS outbreak, public's emotional responses spread rapidly on social media and influence public opinion. To take strategic measures to respond to the public opinion, it is essential to attentively observe the public's emotional responses on Twitter during the MERS outbreak. The second interest of the current study was to investigate the public sentiments on Twitter during the MERS outbreak and to analyze how it influence the spreading speed of MERS-related tweets. To this end, a sentiment analysis was conducted by using machine learning approach. <Table 1> shows the results of the analysis.

As shown in <Table 1>, 42.7% of the total MERS tweets were emotional responses, while 31.3% were neutral responses. Compared to that of the neutral responses, the proportion of emotional responses was higher in the second and the third phase. Among the emotional

<Table 1> Number of Tweets and Public Sentiments

	1st Phase (5.20-5.31)	2nd Phase (6.1-6.20)	3rd Phase (6.21-8.7)	Total (5.20-8.7 )
Number of Days	12 days (15%)	20 days (25%)	48 days (60%)	80 days (100%)
Number of Confirmed Patients	18 (9.68%)	151 (81.18%)	17 (9.14%)	186 (100%)
Number of Tweets	130,645 (3.5%)	3,085,711 (81.5%)	568,769 (15.0%)	3,785,125 (100%)
Neutral	48,169 (36.9%)	979,477 (31.7%)	157,881 (27.8%)	1,185,527 (31.3%)
Emotionally Charged	40,670 (31.1%)	1,298,582 (42.1%)	278,831 (49.0%)	1,618,083 (42.7%)
Fear	1,588 (1.2%)	16,593 (0.5%)	63 (0.3%)	18,244 (0.5%)
Criticism	30,126 (23.1%)	994,461 (32.2%)	213,193 (37.5%)	1,237,780 (32.7%)
Praise	36 (0%)	72,802 (2.4%)	26,660 (4.7%)	99,498 (2.6%)
Total Number of Tweets	130,645 (100%)	3,085,711 (100%)	568,769 (100%)	3,785,125 (100%)

responses, the highest rate of emotion embedded was 'criticism', which occupies 32.7% of the entire MERS tweets. Especially, the volume of critical tweets increased with the increasing number of super-spreaders that tweeted about the disease right after the first death was announced in the second phase. The percentage of 'fear' scored the highest in the first phase and gradually decreased, whereas 'praise' increased over time as it went onto the final phase. According to <Figure 3>, the sentiment of fear peaked just before the first deaths were reported, whereas praise showed a high percentage after June 11th. Criticism was noticeably consistent with the changing trends of the entire MERS tweets.



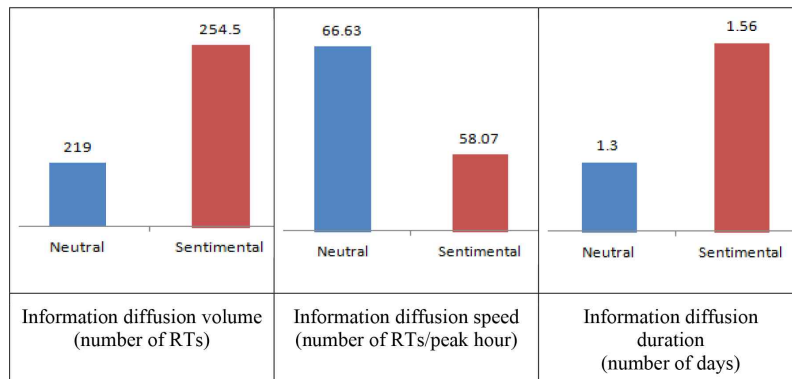
(a) Fear, (b) Criticism, (c) Praise

〈Figure 3〉 Trends of Three Sentiments and overall Tweets (Z-value)

### 3) Influences of Public sentiments on Volume, Duration, and Speed of Retweet Spread

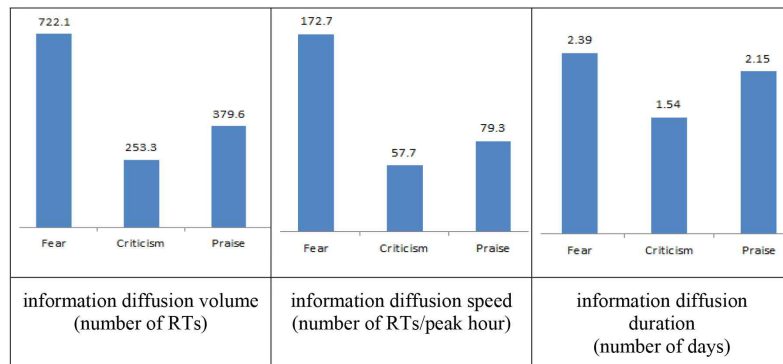
The final interest of this study was to examine the relationship between sentiment of contained within messages and characteristics of information spread. As a result of analyzing the tweet contents, it was found that 3,112 tweet contents were neutral and 3,983 tweet contents were emotional (including both positive and negative). When analyzing the characteristics of retweet, it appeared that emotional tweets had more

diffusion volume than did neutral tweets. As shown in <Figure 4>, emotional tweets were retweeted 254.5 times on average (SD = 533.2), which was more than average retweet frequency of neutral tweets (M = 219.0; SD = 441.0). It is found, however, that neutral tweets spread more quickly than emotional tweets. To look at the number of RTs spread at peak hour, the fastest diffusion speed of neutral tweets was 66.6 times per hour on average (SD = 112.1), whereas that of emotional tweets was 58 times per hour (SD = 118.5) on average. Regarding the diffusion duration, emotional tweets spread much longer than neutral tweets. Emotional tweets lasted 1.56 days (SD = 2.62), and neutral tweets lasted 1.3 days on average (SD = 1.82).



<Figure 4> Public sentiments and Retweets Spread  
 (number of neutral tweet contents = 3112, number of emotionally charged tweet contents = 3983)

Among the three emotional tweets, it was found that average retweet frequency of fear (M = 722.1; SD = 795.7) was significantly higher than average retweet frequencies of both criticism (M = 253.3; SD = 528.9) and praise (M = 379.6; SD = 723.2). In addition, the retweet speed of fear tweets was 172.7 times per peak hour (SD = 213.6), which was significantly faster than those of criticism tweets (M = 57.7; SD = 117.4) and praise tweets (M = 79.3; SD = 160.6). Looking at the persistency of the spread, fear tweets and praise tweets lasted for 2.39 days (SD = 3.06) and 2.15 days (SD = 5.20), respectively, while criticism tweets recorded the lowest period of spread, lasting for 1.54 days (SD = 2.37).



〈Figure 5〉 (number of fear tweets = 54, number of criticism tweets = 3223, number of praise tweets=275)

## 5. Discussion and Conclusions

The main purpose of the present study was to describe the changes in the number of MERS-related tweets and to explore the reasons for such changes during the MERS outbreak in South Korea. In addition, this study investigated the public sentiments on Twitter and examined how the sentiments affect the spreading volume, duration, and speed of MERS-related tweets.

First of all, we found dramatic changes in the amount of MERS-related tweets during these sequences of time periods. These changes were affected by various triggering events such as releases of new cases and deaths announcements, updates on government's policy and response strategies, and stakeholder actions. The results of sentiment analysis also showed that the emotional response of the public on Twitter had changed in a variety of ways depending on the types of triggering events. The announcement of the government's inappropriate response lead to immediate criticism, and the government policy for not revealing the lists of MERS infected hospitals had aggravated the sense of public confusion and fear. These results are consistent with the findings of previous researches on how public response on social media fluctuates during the infectious disease-related crisis (Chew, & Eysenbach, 2010; Salathé, & Khandelwal, 2011; Szomszor, Kostkova, & Louis, 2011).

Secondly, this study examined how the sentiments influence the spreading volume, duration, and speed of MERS-related tweets.

Understanding how the characteristics of tweets are related to the spread is vital for practitioners during public health crisis. Although many studies have been designed to reveal the relationship between the characteristics of message and that of information spread (Hoang, & Lim, 2011; Son, Kim, & Lee, 2012; Yang, & Counts, 2010), few have examined how emotion, especially in tweets, affects the characteristics of RTs (Stieglitz, & Dang-Xuan, 2013). The findings of the current study are consistent with previous findings that emotion-charged or simply emotional tweets tend to be more frequent and to last longer than neutral tweets (Stieglitz, & Dang-Xuan, 2013). These results can be interpreted as follows. An emotional (positive or negative) message in social media might receive more attention and feedback than a non-emotional message (Huffaker, 2010) or might induce cognitive and arousal-related effects that trigger sharing behavior in social media communication (Stieglitz, & Dang-Xuan, 2013). On the other hand, this study found that neutral tweets spread faster than emotional tweets, which is in contrast with previous findings that emotional tweets spread faster (Chen, & Sakamoto, 2013; Stieglitz, & Dang-Xuan, 2013). A possible explanation might be that when faced with a risk situation, people decide to share information depending on the utility of the information rather than the emotional impact of the message.

Given that controlling the commotion due to user-driven spread of news on epidemic diseases during a public health emergency is one of the most important and challenging tasks for public health practitioners, the findings of this study will provide several implications for the

practitioners in the future. First, the public response on Twitter is closely related to the government's response or policy. During the MERS crisis, most of the Korean government's responses were passive, such as reporting the status of patients and confirmed patients and explaining the news articles. It seems that the government's efforts to communicate the MERS-related information to the public, to share countermeasures with the public, and support procedures were relatively insufficient. Indeed, information uncertainty plays a major role in risk perception. Since information is processed on the basis heuristic (Tversky, & Kahneman, 1974), if the information communicated to the people is inadequate, it can inevitably lead to cognitive biases and, in turn, affect their decision-making. In addition, if people feel that they are not getting adequate information from the official source of government, they will find other sources of information outside the government. In this process, the government will lose public trust and credibility, the losses of which will make it harder for the government to manage such risks.

Second, it is also important to manage the public's emotions about the risk as well as to control the risk itself when it situations of crisis occur. Emotions are important explanatory factors in risk perception, attitudes, and actions, and without proper management of the emotions during health emergency, we cannot expect to achieve effective risk communication (Sjöberg, 2007; Xie et al., 2011). Covello and his colleagues (2001) argues that when the communication environment becomes emotionally charged, familiar and traditional approaches do not work or can make the situation worse. Indeed, it was found within the

course of this study that people mainly expressed negative emotional responses such as criticism and fear on Twitter during the MERS outbreak, and emotionally charged tweets tended to be more frequent and lasted longer than neutral tweets. In this situation, it is important for health practitioners to respond instantly and actively to public anxiety through social media. Social media is a useful channel for identifying and responding to public reactions, and the practitioners can minimize the spread of unnecessary rumors by providing immediate information about public anxieties.

Finally, the public wants to establish an emotional connection with the government agency as well as to be provided with accurate information during the public health emergency. Social media technically has characteristics of mass media, but with regards to its role as a conveyor of emotion, it has strong characteristics as an interpersonal communication channel. Therefore, health practitioners need to be aware that social media such as Twitter can be employed as powerful channels that minimize the negative emotion of the public by expressing sympathy and anxiety in a timely manner and communicating with the public in real time during public health emergency.

Several limitations of this study should be noted. The neutrality classifier and the public response classifier were developed using the Naive Bayesian classification method, but the accuracy was 71.1% and 76.8%, respectively. Compared with the accuracy of contents analysis through current machine learning approach, the accuracy is not low, but it is necessary to improve the accuracy to derive research findings that

are more accurate and, thus, more credible. The purpose of this study is to examine the trend of tweets related to MERS, and to examine the diffusion characteristics of public opinion through Twitter. However, Twitter alone cannot represent various types of social media. To generalize the research findings, it is necessary to look at various types of social media platforms such as Facebook and blogs.

## References

- Beaunoyer, E., Arsenault, M., Lomanowska, A. M., & Guitton, M. J. (2017). Understanding online health information: Evaluation, tools, and strategies. *Patient education and counseling*, 100(2), 183-189.
- Bhargavi, P., & Jyothi, S. (2009). Applying naive bayes data mining technique for classification of agricultural land soils. *International journal of computer science and network security*, 9(8), 117-122.
- Borras-Morell, J. E. (2015). Data mining for pulsing the emotion on the web. In *Data Mining in Clinical Medicine* (pp. 123-130). Humana Press, New York, NY.
- Brug, J., Aro, A. R., Oenema, A., de Zwart, O., Richardus, J. H., & Bishop, G. D. (2004). SARS risk perception, knowledge, precautions, and information sources, the Netherlands. *Emerging infectious diseases*, 10(8), 1486.
- Castillo, C., Mendoza, M., & Poblete, B. (2011). Information credibility on twitter. In *Proceedings of the 20th international conference on World wide web* (pp. 675-684). ACM.
- Centers for Disease Control and Prevention (CDC). (2016). *About MERS*. Retrieved from <https://www.cdc.gov/coronavirus/mers/about/index.html>.
- Chen, R., & Sakamoto, Y. (2013). Perspective matters: Sharing of crisis information in social media. In *System Sciences (HICSS), 2013 46th Hawaii International Conference on* (pp. 2033-2041). IEEE.
- Chew, C., & Eysenbach, G. (2010). Pandemics in the age of Twitter: content analysis of Tweets during the 2009 H1N1 outbreak. *PLoS one*, 5(11), e14118.

- Covello, V. T., Peters, R. G., Wojtecki, J. G., & Hyde, R. C. (2001). Risk communication, the West Nile virus epidemic, and bioterrorism: responding to the communication challenges posed by the intentional or unintentional release of a pathogen in an urban setting. *Journal of Urban Health*, 78(2), 382-391.
- D'Alfonso, S., Santesteban-Echarri, O., Rice, S., Wadley, G., Lederman, R., Miles, C., ... & Alvarez-Jimenez, M. (2017). Artificial intelligence-assisted online social therapy for youth mental health. *Frontiers in psychology*, 8, 796.
- Dosemagen, S., & Aase, L. (2016). How Social Media Is Shaking Up Public Health and Healthcare. *Published on The Huffington Post website*, 27.
- Gabarron, E., Lau, A. Y., & Wynn, R. (2015). Is there a weekly pattern for health searches on Wikipedia and is the pattern unique to health topics?. *Journal of medical Internet research*, 17(12), e286.
- Ghiassi, M., Skinner, J., & Zimbra, D. (2013). Twitter brand sentiment analysis: A hybrid system using n-gram analysis and dynamic artificial neural network. *Expert Systems with applications*, 40(16), 6266-6282.
- Gidengil, C. A., Parker, A. M., & Zikmund-Fisher, B. J. (2012). Trends in risk perceptions and vaccination intentions: a longitudinal study of the first year of the H1N1 pandemic. *American Journal of Public Health*, 102(4), 672-679.
- Gohil, S., Vuik, S., & Darzi, A. (2018). Sentiment analysis of health care tweets: review of the methods used. *JMIR public health and surveillance*, 4(2), e43.
- Greaves, F., Ramirez-Cano, D., Millett, C., Darzi, A., & Donaldson, L. (2013). Use of sentiment analysis for capturing patient experience from free-text comments posted online. *Journal of medical Internet research*,

15(11), e239.

- Hirschberg, J., & Manning, C. D. (2015). Advances in natural language processing. *Science*, 349(6245), 261-266.
- Hoang, T. A., & Lim, E. P. (2012). Virality and Susceptibility in Information Diffusions. *Proceedings of the Sixth International AAAI Conference on Weblogs and Social Media*, Palo Alto, CA, 2011, 146-153.
- Hornikx, J., & Hendriks, B. (2015). Consumer tweets about brands: A content analysis of sentiment tweets about goods and services. *Journal of Creative Communications*, 10(2), 176-185.
- Huffaker, D. (2010). Dimensions of leadership and social influence in online communities. *Human Communication Research*, 36(4), 593-617.
- Lippmann, W. (2017). *Public opinion*. Routledge.
- Li, Y., Wang, X., Lin, X., & Hajli, M. (2018). Seeking and sharing health information on social media: A net valence model and cross-cultural comparison. *Technological Forecasting and Social Change*, 126, 28-40.
- Lee, K., Palsetia, D., Narayanan, R., Patwary, M. M. A., Agrawal, A., & Choudhary, A. (2011). Twitter trending topic classification. In *Data Mining Workshops (ICDMW), 2011 IEEE 11th International Conference on* (pp. 251-258). IEEE.
- Mazzocut, M., Truccolo, I., Antonini, M., Rinaldi, F., Omero, P., Ferrarin, E., ... & Tasso, C. (2016). Web conversations about complementary and alternative medicines and cancer: content and sentiment analysis. *Journal of medical Internet research*, 18(6), e120.
- Nath, C., Huh, J., Adupa, A. K., & Jonnalagadda, S. R. (2016). Website sharing in online health communities: a descriptive analysis. *Journal of medical Internet research*, 18(1), e11.
- Oscar, N., Fox, P. A., Croucher, R., Wernick, R., Keune, J., & Hooker, K.

- (2017). Machine learning, sentiment analysis, and tweets: An examination of Alzheimer's disease stigma on Twitter. *Journals of Gerontology: Psychological Sciences*, 72, 742-75
- Oyeyemi, S. O., Gabarron, E., & Wynn, R. (2014). Ebola, Twitter, and misinformation: a dangerous combination?. *Bmj*, 349, g6178.
- Read, J. (2005). Using emoticons to reduce dependency in machine learning techniques for sentiment classification. In *Proceedings of the ACL student research workshop* (pp. 43-48). Association for Computational Linguistics.
- Rui, H., Liu, Y., & Whinston, A. (2013). Whose and what chatter matters? The effect of tweets on movie sales. *Decision Support Systems*, 55(4), 863-870.
- Salathé, M., & Khandelwal, S. (2011). Assessing vaccination sentiments with online social media: implications for infectious disease dynamics and control. *PLoS computational biology*, 7(10), e1002199.
- Seale, H., Heywood, A. E., McLaws, M. L., Ward, K. F., Lowbridge, C. P., Van, D., & MacIntyre, C. R. (2010). Why do I need it? I am not at risk! Public perceptions towards the pandemic (H1N1) 2009 vaccine. *BMC infectious diseases*, 10(1), 99.
- Signorini, A. (2014). Use of social media to monitor and predict outbreaks and public opinion on health topics.
- Sjöberg, L. (2000). Factors in risk perception. *Risk analysis*, 20(1), 1-12.
- Sjöberg, L. (2007). Emotions and risk perception. *Risk management*, 9(4), 223-237.
- Slovic, P. (1987). Perception of risk. *Science*, 236(4799), 280-285.
- Slovic, P., Peters, E., Finucane, M. L., & MacGregor, D. G. (2005). Affect, risk, and decision making. *Health psychology*, 24(4S), S35.
- Som, I., Kim, E., & Lee, D. (2012). Information Diffusion in Online Social

Media: The Role of Message and User Characteristics in Retweeting on Twitter. *Paper presented at: Korea Intelligent Information System Society*, Seoul, Korea.

- Stieglitz, S., & Dang-Xuan, L. (2013). Emotions and information diffusion in social media-sentiment of microblogs and sharing behavior. *Journal of management information systems*, 29(4), 217-248.
- Szomszor, M., Kostkova, P., & Louis, C. S. (2011). Twitter informatics: tracking and understanding public reaction during the 2009 swine flu pandemic. In *Proceedings of the 2011 IEEE/WIC/ACM International Conferences on Web Intelligence and Intelligent Agent Technology-Volume 01* (pp. 320-323). IEEE Computer Society.
- Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *science*, 185(4157), 1124-1131.
- Weinstein, N. D. (1980). Unrealistic optimism about future life events. *Journal of personality and social psychology*, 39(5), 806.
- World Health Organization. (2015). Summary and risk assessment of current situation in Republic of Korea and China. Retrieved from [http://www.who.int/csr/disease/coronavirus\\_infections/risk-assessment-19june-2015/en/](http://www.who.int/csr/disease/coronavirus_infections/risk-assessment-19june-2015/en/)
- Wynn, R., Oyeyemi, S. O., Johnsen, J. A., & Gabarron, E. (2017). Tweets are not always supportive of patients with mental disorders. *International J Integrated Care*, 17(3), A149-A149.
- Xie, X. F., Wang, M., Zhang, R. G., Li, J., & Yu, Q. Y. (2011). The role of emotions in risk communication. *Risk Analysis: An International Journal*, 31(3), 450-465.
- Yang, J., & Counts, S. (2010). Predicting the Speed, Scale, and Range of Information Diffusion in Twitter. *Proceedings of the Fourth International*

*AAAI Conference on Weblogs and Social Media*, Palo Alto, CA, 10(2010), 355-358.

Zhang, H., Gan, W., & Jiang B. (2014). Machine Learning and Lexicon based Methods for Sentiment Classification: A Survey. *Proceedings of the 2014 11th Web Information System and Application Conference*, Tianjin, China, (pp. 262-265). IEEE.

투고일: 2019년 05월 15일  
논문 수정일: 2019년 07월 14일  
게재 확정일: 2019년 07월 22일

## 소셜 미디어와 인플루엔자 비상사태: 한국의 2015년 MERS 사태 트위터 내용 분석

**이병관**

(한양대학교, 광고홍보학과, 교수)

**오현정**

(차 의과대학교, 의료홍보미디어학과, 조교수)

**백현미**

(고려대학교, 미디어학부, 조교수)

**이상록**

(한양대학교, 광고홍보학과, 겸임교수)

위험 인식은 효과적인 헬스 리스크 커뮤니케이션의 주요한 쟁점이다. 본 연구는 2015년 MERS사태에 트위터에서 발생하는 대중의 인식을 확인하고 이해하기 위하여 MERS와 관련된 트윗을 수집 및 분석하였다. 연구의 주요 목적은 MERS와 관련된 트윗의 수의 변화를 확인하고 설명하는 것이다. Naïve Bayesian 분류법을 사용하여 트위터 상에서 대중의 정서를 분석하고 이러한 정서가 확산 정도, 지속성, 속도에 어떻게 영향을 미치는지를 확인하였다. 연구 결과 대중의 정서적 반응이 정부에 대응과 사건의 전개에 따라 다양한 방식으로 변화했음을 확인하였다. 또한 감정적인 트윗이 중립적인 트윗보다 더 오래 지속되었다. 하지만 중립적인 트윗은 감정적인 트윗보다 빠르게 퍼져 나가는 것으로 확인되었는데 위험 상황에 직면하였을 때 사람들은 메시지의 정서적인 영향보다는 정보의 유용성에 따라 정보를 공유하는 경향이 있었다. 본 연구의 실무적 함의로는 위험 상황에서 정부의 대응이 중요하다는 점과 대중의 정서를 관리하는 것이 효과적인 헬스 리스크 커뮤니케이션의 중요한 요인이라는 것을 인식해야 한다.

주제어 : 소셜 미디어, 헬스 리스크 커뮤니케이션, MERS, 정서 분석, 정보 확산, 트위터