

eMood: Modeling Emotion for Social Media Analytics on Ebola Disease Outbreak

Research-in-Progress

Wingyan Chung

Institute for Simulation and Training
University of Central Florida
Orlando, FL, U.S.A.
wchung@ucf.edu

Saike He

State Key Laboratory of Management and
Control for Complex Systems,
Chinese Academy of Sciences
Beijing, China
saike.he@ia.ac.cn

Daniel Zeng

Eller College of Management
The University of Arizona
Tucson, AZ, U.S.A.
and
Chinese Academy of Sciences
Beijing, China
zeng@email.arizona.edu

Abstract

Social media enable users to express their emotion promptly, helping health policy makers to gauge public sentiment of disease outbreak. In this research, we developed an approach to social-media-based public health informatics and built a proof-of-concept system named eMood that helps to collect, analyze, and visualize Ebola outbreak discussions on Twitter. Our approach uses a comprehensive lexicon to identify emotion categories and present analysis findings of users' network relationship and influence patterns. We compared two methods of identifying user influence, user centrality and emotion entrainment, by using 255,118 tweets posted by 210,900 users in January 2015. Experimental results show that both methods identified highly influential users. Regression analysis of user influence rank and emotion scores demonstrates significant relationship between user influence and each emotion category. These results should provide strong implication for understanding social actions and for collecting social intelligence for public health informatics.

Keywords

Social media analytics, sentiment analysis, social network analysis, emotion extraction, influence measure, social intelligence, public health informatics, Ebola disease outbreak.

Introduction

Public emotion often influences business decisions and policy development. Business leaders, politicians, and healthcare organizations create their agendas and offerings based partly on public emotion. Social media, a growing online platform for sharing opinions and ideas, provides many opportunities for decision makers to understand public emotion. For example, healthcare professionals and public-health policy makers may identify from social media the spread of an infectious disease in advance of an epidemic, helping them to react in a timely manner.

Since the Ebola outbreak in West Africa was first reported in March 2014, more than 11,284 people have died from the disease (BBC 2015). The outbreak is the largest and most complex in the history of the disease, which was first discovered in 1976 (CDC 2015; WHO 2015). There have been more cases and deaths in this outbreak than all previous outbreaks combined. The disease was transmitted first from

animals to humans, and then from humans to humans across countries. As of August 2015, the World Health Organization reports that no licensed Ebola vaccines are available (WHO 2015). Fear about the disease spreads widely on social media, raising public panic and concerns about this outbreak.

Despite potential advantages of using social media to track the Ebola outbreak, the rapidly-growing volume of social media (SM) makes it difficult for decision makers to understand the sentiment and relationships of social media users in the discussion about the outbreak. How to gauge the emotion of the public on the Ebola outbreak and related health policy issues becomes a concern for decision makers and healthcare professionals. Unfortunately, there is little work on social-media-based public health informatics that could potentially benefit related policy decision making.

By analyzing the content and linkage information in social media, decision makers and healthcare professionals could possibly obtain useful answers to such questions as: Who are the most influential leaders in the Ebola SM community? How are different types of emotion changing over time as new events emerge? How does emotion spread and entrain in the SM community? What are the network relationships among participants in Ebola outbreak SM discussion?

In this research, we developed an approach and a new system to collecting and analyzing emotion expressed in infectious disease social media. We compared two methods of emotion analysis to identify influential users and to trace their contagion effects on public emotion, and report empirical results of analyzing the emotion of 210,900 users who posted 255,118 tweets on Ebola outbreak during January 2015. The results provide strong implication for understanding social actions and for collecting social intelligence for public health informatics.

Literature Review

Social media analytics (SMA) is a rich set of tools and technologies most recently applied to public health informatics (Liang et al. 2015; Zeng et al. 2011; Zeng et al. 2010). Traditional domains that SMA is used include business intelligence, online product review analysis, security informatics, and advertising (Chen et al. 2012; Chung 2014; Chung et al. 2012; Hansen et al. 2007). The analysis of emotion in social media relates to literature in sentiment analysis, social network analysis, and emotion entrainment. We review below these domains with a view of informing the analysis of emotion in social-media-based public health informatics.

Emotion Analysis

Emotion extraction and sentiment analysis are important tools for understanding large amounts of social media data (Abbasi et al. 2008). Researchers have developed a specialized lexicon to support emotion extraction (Mohammad et al. 2013). The lexicon contains over 13,000 entries (each entry being a combination of term-category). A term may belong to one or more of eight emotion categories (anger, anticipation, disgust, fear, joy, sadness, surprise, trust) identified by psychological theories (Plutchik 1980). Social network analysis is an SMA application in which users within a large community are analyzed to reveal their positions and levels of influence. Traditionally, social network analysis examines network dynamics using small datasets (typically with fewer than 20 actors) (Wasserman et al. 1994). Advances in information technologies enable new capabilities of network analysis, such as identifying opinion leaders in discussion forums (Song et al. 2007), finding expertise in networks (Zhang et al. 2007), analyzing business stakeholder networks (Chung et al. 2009), and visualizing networked relationships of infectious disease transmission (Zeng et al. 2011). Although social media are increasingly important information sources for understanding public sentiment, little work is found in analyzing emotion to support intelligence and public health policy decision making.

Measuring Influence

Previous studies have explored various aspects of influence across social networks. A study by the Hewlett-Packard Company uses a count of the number of re-tweets on Twitter to determine influence, finding that popularity does not imply influence (Romero et al. 2011). Others explored the traits of tweets that are commonly retweeted in an effort to discover the basis behind the spread of an individual message (Kiciman 2012). Along the same path of content-based information, a study has also found that diverse

messages and focused messengers have the greatest impact in a social network (Weng et al. 2014). Each of these approaches has some amount of variance in yielding valid results, as was made clear by critical analyses (Probst et al. 2013). Due to the inherent variability present in opinion- and influence-based research, steps should be taken to quantify and evaluate the results.

Traditional social network approaches quantify user influence by using various structural measures, such as in/out-degree, betweenness, and PageRank (Jackson 2008). The idea is based on social linkages developed in human interaction that facilitate spread of human emotion. These linkages play a critical role in large social media communities by facilitating exchange of information and confirmation of authoritative and influential positions. However, they have not been examined widely in social-media-based public health informatics.

Recent works also attempt to model the spread of human behavior in social networks (Bond et al. 2012; Kane et al. 2014; Liang et al. 2015; Schreck et al. 2013). Emotion entrainment, which accounts for the synchronous convergence of human emotions, has been examined in neuroscience (Bastiaansen et al. 2011) and in human social networks (Kramer et al. 2014). Whether human emotion entrains in large-scale networks of importance to international public health policy decision making is relatively unknown. While previous methods provide quantitative analysis of social interactions, they often require explicit causal knowledge that is scarce in many scenarios (Anderson et al. 1992; Bailey 1975; Bakshy et al. 2011; Romero et al. 2011).

Social-Media-Based Public Health Informatics

In this research, we developed an approach to social-media-based public health informatics and built a proof-of-concept system named eMood that helps to collect, summarize, and visualize Ebola discussions on Twitter. Our approach consists of emotion extraction and entrainment of a large-scale social media network of significance to international public health policy decision making. Through applying our approach, we examined the process of identifying influential opinion leaders in social media networks, studied changes in emotion over time, and revealed network relationships among participants. We describe below our dataset, analyses, and preliminary findings of the two approaches to emotion analysis.

Our Approach and the eMood System

Our approach performs relevant data collection, cleaning, filtering, and analysis of social media postings. Input sources include social media data, a specialized lexicon for emotion extraction and sentiment analysis, an algorithm to compute emotion indices, and program modules for performing network analysis and visualization. While previous works did not address public health informatics (e.g., (Bond et al. 2012; Lampe et al. 2011)) or focused primarily on certain aspects (e.g., visualization (Schreck et al. 2013) or network analysis (Kane et al. 2014)), our framework supports a comprehensive analysis of public-health social media content and networks and provides a dynamic, user-friendly visualization of the user sentiment, reputable opinion leaders, and their social activities.

Based on our approach, we developed the eMood system (see Figure 1), an online web portal that extracts real-time sentiment and emotion from tweets related to Ebola and discover network relationship among these tweet authors. An information system artifact (Hevner et al. 2004), eMood implements the steps of our approach in the context of Ebola discussion on Twitter.com, a major social media website where users post short messages called “tweets” on issues that they are interested in. Each tweet consists of no more than 140 alphanumeric characters. We developed the eMood system and its two influence-identification methods (described below) by following design science research guidelines (Hevner et al. 2004), which include producing a design as an artifact, solving important and relevant business / social problems, demonstrating utility and quality through evaluation, contributing to social-media-based public health informatics, and presenting the research to both technology-oriented and management-oriented audiences. We treated design as a search process to discover an effective solution to our problem and tested alternatives against our requirements and constraints in the domain of health informatics.

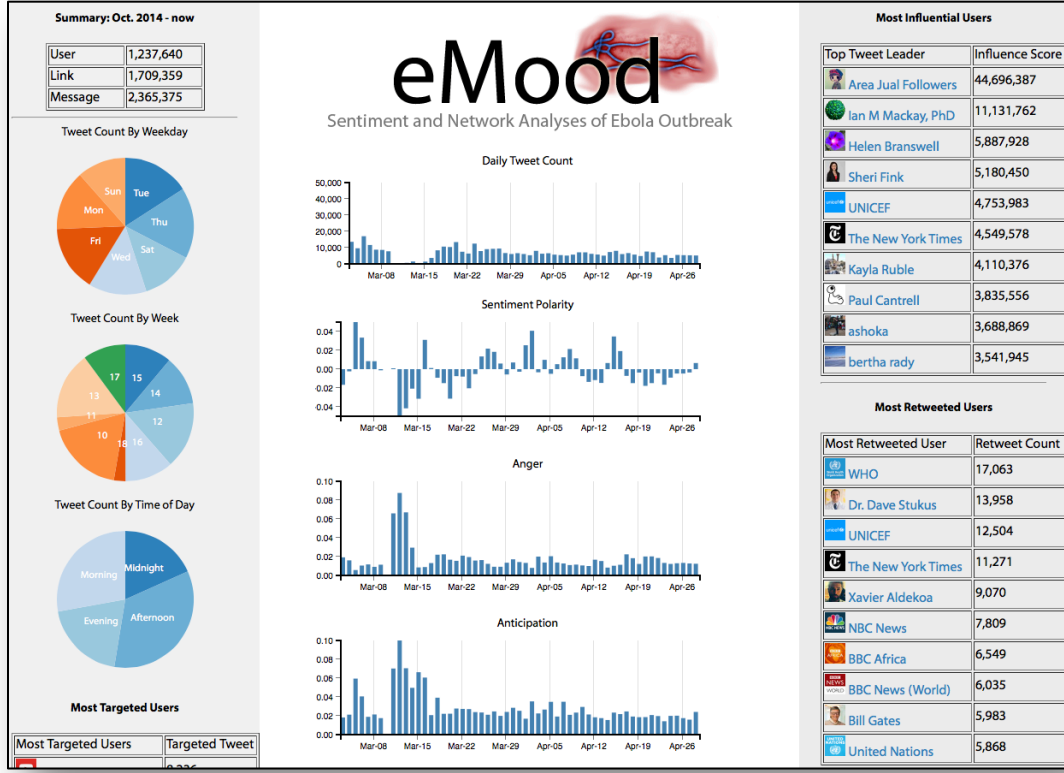


Figure 1. The eMood System

Dataset

To collect the tweets related to Ebola discussion, we constructed carefully a list of queries by reviewing recent literature published on the topic (BBC 2015; CDC 2015; WHO 2015). Filtering and collection testing helped to reduce the original 12 queries to 6 domain-specific queries: “ebola,” “ebola patient,” “ebola outbreak,” “ebola virus,” “ebola vaccine,” and “ebola epidemic.” Our eMood system collects tweet postings automatically and periodically since October 2014.

As of September 8, 2015, our dataset consists of 2,980,918 tweets posted by 1,460,389 unique users. In this study, we focus on the recent spike of Ebola cases in January 2015 (BBC 2015). We selected a subset of the data consisting of 255,118 tweets posted by 210,900 users between January 1st and January 31st of 2015. That month also recorded the highest number of tweets and the most user activities in 2015. From the users whose tweets were collected in January 2015, we selected a random sample of 5,000 users for further examination of the emotion extracted from the tweets and of emotion entrainment.

Extracting Emotion Scores

To extract emotion from each tweet, we computed values of indices in eight emotional categories defined in (Plutchik 1980) (i.e., anger, disgust, fear, sadness, surprise, anticipation, joy, and trust). Each

$$M_{ij} = \frac{\sum_{k=1}^{n_i} W_k}{n_i} \quad \forall i \in \{1 \dots N\}, \forall j \in \{1 \dots k\} \quad \dots (1)$$

where

M_{ij} = Index of emotion j of Tweet i

n_i = Number of words in Tweet i

k = Number of different emotion categories

N = Total number of tweets being considered

$$W_k = \begin{cases} 1 & \text{if word}(k) \text{ is in category } j, \text{ where} \\ & k \in \{1 \dots n_i\} \\ 0 & \text{otherwise} \end{cases}$$

index represents the emotional intensity of a tweet. We used the emotion lexicon developed in (Mohammad et al. 2013) to identify emotional words that appear in the tweets. The lexicon contains 13,901 word-emotion entries tagged by over 2,216 human users who completed a total of 38,726 tagging assignments. Each word was assigned a sentiment polarity (positive (+1) or negative (-1)) and is categorized into one (or more) of the eight emotional categories. Equation (1) shows the computation of the index of emotion j (ranging from 1 to 8) of tweet i , which contains n_i words.

Identifying User Influence

We developed two methods for identifying user influence. Below we describe the two methods, user centrality and emotion entrainment, that we developed and their design rationale.

User Centrality

Twitter users interact with each other by sending tweets or replying to their tweets. These interactions enable researchers and policy makers to analyze user activities and roles in the community. To support the analysis, we constructed a user interaction network by using the interactions and roles of the users. A node in the network represents a user; a link represents an interaction between two users. The eMood system identifies a link between User A and User B when User A sends a tweet targeted to User B, or retweets another tweet written by User B, or modifies and then sends out a tweet written by User B.

The network of user interaction changes over time. We define an interaction window to be a specific time frame during which user activities are considered in building the network. This time frame is a sliding window of 36 days that end on the day on which new data are collected. While other lengths are possible, we chose the 36-day length to reflect a typical time span of user activity spike based on empirical observation. The eMood system automatically constructs a new user interaction network for the most recent interaction window, and does it two times per day.

With the network of relationships constructed from the data, we measured user influence by using the *Betweenness Centrality* score. While other measures of centrality are available (Freeman 1977; Jackson 2008), *Betweenness Centrality* (BC) identifies the extent to which a user serves as a bridge in Ebola SM discussion. A user with high BC score is able to bridge diverse opinions and to find common grounds among highly polarized views.

Emotion Entrainment

Emotion entrainment is the socialization of feelings among people who interact with each other. To examine how this process takes place in a large-scale social media network and how emotion evolves dynamically, we developed a model-free approach to detect causal relationships and to identify influential users based on the theory of emotion entrainment (Clayton et al. 2005). This theory suggests that human emotion rhythmically converges through social interactions. In this process, we consider users that are most entrained by others as influential. We represent the emotion time series of two users x and y with two Markov processes $X = x_t$ and $Y = y_t$. Then the strength of User y who entrains User x is given by Equation (2).

$$En(X \rightarrow Y) = H(x_{t+1} | \mathbf{x}_t^m) - H(x_{t+1} | \mathbf{x}_t^m, \mathbf{y}_t^n) \quad (2)$$

where, $\mathbf{x}_t^m = (x_t, \dots, x_{t-m+1})$, $\mathbf{y}_t^n = (y_t, \dots, y_{t-n+1})$, while m and n are the orders of each of the Markov processes; $H(*)$ calculates the entropy of the probability distribution enclosed. In our experiments, we set the Markov orders $m=n=3$. According to our empirical analysis, this order is high enough while higher order increases computational cost without present much quantitative differences. Equation (2) can be estimated based on Simpson's rule (Kaw et al. 2008) with a computational complexity of $O(N \log(N))$. This cost is acceptable for entrainment analysis on large-scale datasets.

$$I(x) = \sum_{x \neq y} En(y \rightarrow x) \quad (3)$$

Using Equation (3), we calculated the entrainment score of a user x to measure the user's influence based on how other users (y) are entrained by him. It follows that the higher the value of $I(x)$ is, the better User x can influence others because of the ability to entrain other users.

Empirical Results

In this section, we present empirical results of using the emotion analysis methods to examine influence of Twitter users who posted about Ebola outbreak in January 2015, and to provide descriptive and inferential statistics of the emotion scores.

Top Influential Users

Table 1 lists the top 10 most influential users identified by each method. The first method identified several authorities in infectious diseases (shown on left of the table). World Health Organization (ranked #1) and Center for Infectious Disease Research and Policy at University of Minnesota (ranked #4) are two major organizations that conduct various activities to help Ebola patients. Médecins Sans Frontières (ranked #7) is an independent organization providing medical care for Ebola patients. Experts and authorities in Ebola are identified, including Ilona Kickbusch (ranked #2) and Ian M. Mackay (ranked #3), who are respectively renowned public health expert and virologist. The remaining users on the list are activists, medical news correspondents, and major news sources related to Ebola such as BBC Africa.

Table 1. Top Influential Users

| Influential Users based on User Centrality | | | Influential Users based on Emotion Entrainment | | |
|--|---|-------------|--|---|---------------|
| User Name | URL | BC Score | User Name | URL | Entrain Score |
| World Health Organization | http://twitter.com/WHO | 55148797.6 | Joel Selanikio | http://twitter.com/jselanikio | 31.89448276 |
| Ilona Kickbusch | http://twitter.com/IlonaKickbusch | 51890152.59 | CNN Mexico | http://twitter.com/CNNMex | 30.86975574 |
| Ian M Mackay, Ph.D. | http://twitter.com/MackayIM | 34222688.91 | Debra A. Barrath, MN | http://twitter.com/dbarrath | 29.88568909 |
| Ctr. for Infectious Disease Rsch. and Policy, U. Minn. | http://twitter.com/CIDRAP | 15711893.61 | SE-KY | http://twitter.com/KG_Wabbit | 29.58278792 |
| Laurie Garrett | http://twitter.com/Laurie_Garrett | 15473363.91 | Channel NewsAsia | http://twitter.com/ChannelNewsAsia | 29.50283438 |
| Alex Thomson | http://twitter.com/alextomo | 14881153.38 | Worldwide News | http://twitter.com/Koran_Inggis | 29.02798354 |
| Médecins Sans Frontières, UK | http://twitter.com/MSF_uk | 14720134.42 | ★ οπ*(~φασηκ~)*π | http://twitter.com/iam_frankovic | 29.00249686 |
| Richard Besser | http://twitter.com/DrRichardBesser | 14454347.22 | Patriots Billboard | http://twitter.com/PatriotsOrg | 28.97650483 |
| D. Marie | http://twitter.com/itsmepanda1 | 14393567.04 | Vivek Virendra Singh | http://twitter.com/VivekSViewS | 28.72227653 |
| BBC Africa | http://twitter.com/BBCAfrica | 11543295.31 | Sigueme y te Sigo | http://twitter.com/SiguesTeSigo | 28.63721573 |

In contrast with the results found by the user centrality method, the entrainment method identified a diverse set of influential users. The top 3 users are Joel Selanikio (a pediatrician and Ebola responder), CNN Mexico, and Debra A. Barrath (a healthcare consultant). Each of them has posted at least 20 messages during the outbreak and successfully involved emotional engagement in large-scale. Other top-ranked users include journalists or news organizations (Channel News Asia #5, WorldWide News #6), public health workers (Vivek Singh #9), and celebrities (SiguesTeSigo #10). These users have large numbers of followers or following, but may not be all focused on Ebola outbreak and relief effort. Compared with the User Centrality method, the Entrainment method identified users who appear to have broader media exposure but less authority in the network.

Distribution of Emotion Ratio

To understand the relative strength of the emotion, we computed the ratio for the j th emotion e_j in top k users (sorted by descending order of influence score calculated by (3)) as:

$$R(e_j)_k = \frac{\sum_{i=1}^k e_{ij}}{\sum_{i=1}^k \sum_{j=1}^m e_{ij}}, \quad j = 1, \dots, m; k = 1, \dots, n \quad (4)$$

where, e_{ij} is the sum of all the j th emotion scores (among m ($= 8$) categories defined in (Plutchik 1980)) of the tweets written by the i th user in the ranked list of n ($= 5,000$) users. Intuitively, $R(e_j)_k$ characterizes the emotion distribution over the top k users. If $R(e_j)_k$ varies significantly as k increases (as more less-influential users are involved in the calculation of ratio), the j th emotion should be considered as a critical factor for distinguishing influential users from normal users.

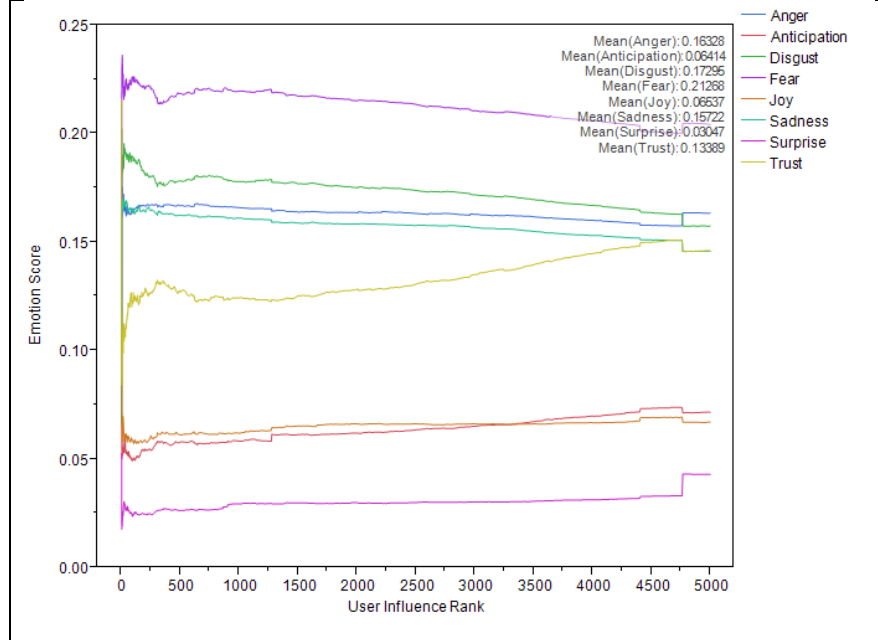


Figure 2. Relative Strengths of Eight Emotion Categories

Figure 2 illustrates the change in ratio of emotions in top n ($= 5000$) most influential users. There are three aspects worth mentioning. *First*, as users' influence decreases (i.e., lower in ranks), the ratios of fear, disgust, anger, and sadness decrease while the ratios of trust, joy, anticipation, and surprise increases (the changes were found to be statistically significant as described in the following section). These results indicate that influential users are more likely to be those high in emotions related to fear, disgust, anger, and sadness. *Second*, fear was the most dominant emotion (with the highest ratio) in the posted tweets while surprise was the least dominant emotion (with the lowest ratio). *Third*, the ratio for trust increases dramatically as users' ranks change from 1 to n ($= 5000$). This means the relative amount of trust expressed in social media increasingly distinguishes highly-influential users from less-influential users. This implies that expressing less trust in social media increases user influence in the community.

Regression Analysis of Emotion Scores and User Ranks

To find out whether users' influence ranks can be used to predict different types of emotion they expressed, we conducted a regression analysis to examine the relationship. The dependent variables (DVs) are the scores of the eight emotion types (y_i , where $i = 1, 2, \dots, 8$). The independent variable (IV) is the users' influence rank. These results demonstrate the usability of the eMood system and the emotion analysis methods for social-media-based public health informatics. Health professionals, public health policy makers, and health organizations administrators should benefit from the identification of influential opinion leaders, who can inform situational awareness during a disease outbreak and can provide necessary social support to other SM users. Findings from regression analysis support a deeper understanding of the way user influence rank predicts emotion scores.

Table 2 shows the eight linear regression equations, p -value of testing the significance of the regression slope coefficient, and R^2 . The results show that all the regression slopes are statistically significant, meaning that user influence rank is a significant predictor of the score of each emotion type. The R^2 values are higher than 84% for the emotion of anticipation, disgust, fear, sadness, and trust, showing that

variations in these emotion scores were explained mostly by user influence rank. Reversing the positions of IV and DVs, though possible, would not produce meaningful regression results because user influence rank is a general attribute that should be used to predict specific emotion scores.

These results demonstrate the usability of the eMood system and the emotion analysis methods for social-media-based public health informatics. Health professionals, public health policy makers, and health organizations administrators should benefit from the identification of influential opinion leaders, who can inform situational awareness during a disease outbreak and can provide necessary social support to other SM users. Findings from regression analysis support a deeper understanding of the way user influence rank predicts emotion scores.

Table 2. Linear Regression Statistics

| Emotion (y_i) | Linear Regression Equation ($x = \text{UserRank}$) | p -value of testing $H_0 : \beta_1 = 0$ | R^2 |
|------------------------|---|--|--------|
| Anger (y_1) | $y_1 = 0.1673 - 1.61\text{e-}6 * x$ | $< .0001^*$ | 75.88% |
| Anticipation (y_2) | $y_2 = 0.055 + 3.78\text{e-}6 * x$ | $< .0001^*$ | 88.78% |
| Disgust (y_3) | $y_3 = 0.184 - 4.33\text{e-}6 * x$ | $< .0001^*$ | 85.13% |
| Fear (y_4) | $y_4 = 0.223 - 4.18\text{e-}6 * x$ | $< .0001^*$ | 85.44% |
| Joy (y_5) | $y_5 = 0.062 + 1.50\text{e-}6 * x$ | $< .0001^*$ | 65.94% |
| Sadness (y_6) | $y_6 = 0.165 - 0.000003 * x$ | $< .0001^*$ | 88.80% |
| Surprise (y_7) | $y_7 = 0.026 + 1.83\text{e-}6 * x$ | $< .0001^*$ | 58.65% |
| Trust (y_8) | $y_8 = 0.119 + 6.00\text{e-}6 * x$ | $< .0001^*$ | 84.10% |

Conclusion

Social media enable users to express their emotion publicly and promptly, thereby helping health policy makers and healthcare workers to gauge public sentiment of infectious disease outbreak. In this research, we developed an approach to social-media-based public health informatics and built a proof-of-concept system named eMood that helps to collect, summarize, and visualize Ebola discussions on Twitter. Our approach uses a comprehensive lexicon to identify emotion categories, includes algorithms to compute emotion scores and user influence scores (based on both social ties and entrainment), and present analysis findings of users' network relationship and influence patterns. We also developed two methods for identifying user influence: user centrality and emotion entrainment. To compare the two methods, we used a dataset consisting of 255,118 tweets posted by 210,900 users in January 2015.

Empirical results show that both methods identified highly influential users from the Twitter network of Ebola outbreak discussion. We also found, among the top influential users, the relative strengths of fear, disgust, anger, and sadness decrease with users' influence while the relative strengths of trust, joy, anticipation, and surprise increases with the influence. Regression analysis of user influence rank and emotion scores demonstrate significant relationship between influence and each emotion score. These results should provide strong implication for understanding social actions and for collecting social intelligence for public health informatics. Our ongoing works include conducting a thorough evaluation of the emotion methods, extending the analysis to a larger dataset covering a longer time span, developing metrics to understand performance, and developing new visualization of the data and findings. This research should contribute to research in sentiment and social media network analyses, and specifically to the domain of public health informatics.

Acknowledgments

This paper is based upon work supported in part by funding from the UCF Institute for Simulation and Training and from the National Natural Science Foundation of China under Grant No. 71402177. Any opinions, findings, and conclusions or recommendations expressed in this paper are those of the authors and do not necessarily reflect the views of the funding agencies. We thank the conference chairs, reviewers, and coordinators for their assistance and valuable suggestions.

References

- Abbasi, A., Chen, H., Thoms, S., and Fu, T. 2008. "Affect analysis of Web forums and blogs using correlation ensembles," *IEEE Transactions on Knowledge and Data Engineering* (20:9), pp. 1168-1180.
- Anderson, R. M., May, R., and Anderson, B. 1992. "Infectious diseases of humans: dynamics and control," *Australian Journal of Public Health* (16), pp. 208-212.
- Bailey, N. T. 1975. *The mathematical theory of infectious diseases and its applications*, Charles Griffin & Company Ltd, 5a Crendon Street, High Wycombe, Bucks HP13 6LE.
- Bakshy, E., Hofman, J. M., Mason, W. A., and Watts, D. J. 2011. "Everyone's an Influencer: Quantifying Influence of Twitter," in *Fourth ACM International Conference on Web Search and Data Mining*, WSDM '11: Hong Kong.
- Bastiaansen, J., Thioux, M., and Keysers, C. 2011. "Evidence for mirror systems in emotions," *Philosophical Transactions of The Royal Society B* (364), pp. 2391-2404.
- BBC 2015. "Ebola: Mapping the outbreak," British Broadcasting Company:
<http://www.bbc.com/news/world-africa-28755033>.
- Bond, R. M., Fariss, C. J., Jones, J. J., Kramer, A. D. I., Marlow, C., and Settle, J. E. 2012. "A 61-million-person experiment in social influence and political mobilization," *Nature* (489:7415), pp. 295-298.
- CDC 2015. "Ebola: Current Outbreaks," Centers for Disease Control and Prevention,
<http://www.cdc.gov/vhf/ebola/outbreaks/index.html>.
- Chen, H., Chiang, R., and Storey, V. 2012. "Business Intelligence and Analytics," *MIS Quarterly* (36:4), pp. 1165-1188.
- Chung, W. 2014. "BizPro: Extracting and Categorizing Business Intelligence Factors from Textual News Articles," *International Journal of Information Management* (34:2), pp. 272-284.
- Chung, W., Chen, H., and Reid, E. 2009. "Business stakeholder analyzer: An experiment of classifying stakeholders on the Web," *Journal of the American Society for Information Science and Technology* (60:1), pp. 59-74.
- Chung, W., and Tseng, T.-L. 2012. "Discovering business intelligence from online product reviews: A rule-induction framework," *Expert Systems with Applications* (39:15), pp. 11870-11879.
- Clayton, M., Sager, R., and Will, U. 2005. "In time with the music: the concept of entrainment and its significance for ethnomusicology," *European Meetings in Ethnomusicology* (11), pp. 1-82.
- Freeman, L. C. 1977. "A set of measures of centrality based on betweenness," *Sociometry* (40:1), pp. 35-41.
- Hansen, F., and Christensen, S. R. 2007. *Emotions, Advertising and Consumer Choice*, Copenhagen Business School Press.
- Hevner, A. R., March, S. T., Park, J., and Ram, S. 2004. "Design science in information systems research," *Management Information Systems Quarterly* (28:1), pp. 75-105.
- Jackson, M. O. 2008. *Social and Economic Networks*, Princeton University Press: Princeton, NJ.
- Kane, G. C., Alavi, M., Labianca, G. J., and Borgatti, S. P. 2014. "What's different about social media networks? A framework and research agenda," *MIS Quarterly* (38:1), pp. 275-304.
- Kaw, A., and Kalu, E. 2008. *Numerical Methods with Applications*, (1st. ed.) <http://www.autarkaw.com>.
- Kiciman, E. 2012. "OMG, I Have to Tweet That! A Study of Factors that Influence Tweet Rates," in *The Sixth International AAAI Conference on Weblogs and Social Media*, Association for the Advancement of Artificial Intelligence.
- Kramer, A. D., Guillory, J. E., and Hancock, J. T. 2014. "Experimental evidence of massive-scale emotional contagion through social networks," *Proceedings of the National Academy of Sciences* (111:29), pp. 8788-8790.
- Lampe, C., LaRose, R., Steinfield, C., and DeMaagd, K. 2011. "Inherent barriers to the use of social media for public policy informatics," *The Innovation Journal: The Public Sector Innovation Journal* (16:1), p. Article 6.
- Liang, Y., Zheng, X., Zeng, D. D., Zhou, X., Leischow, S. J., and Chung, W. 2015. "Exploring How the Tobacco Industry Presents and Promotes Itself in Social Media," *Journal of Medical Internet Research* (17:1).
- Mohammad, S. M., and Turney, P. D. 2013. "Crowdsourcing a Word-Emotion Association Lexicon," *Computational Intelligence* (29:3), pp. 436-465.

- Plutchik, R. 1980. "A general psychoevolutionary theory of emotion," *Emotion: Theory, research, and experience* (1:3), pp. 3-33.
- Probst, F., Grosswiele, L., and Pfleger, R. 2013. "Who will lead and who will follow: Identifying Influential Users in Online Social Networks - A Critical Review and Future Research Directions," *Business & Information Systems Engineering* (5:3), pp. 179-193.
- Romero, D. M., Galuba, W., Asur, S., and Huberman, B. A. 2011. "Influence and Passivity in Social Media," in *Proceedings of the 20th International Conference Companion on World Wide Web*, ACM Press, New York, NY.
- Schreck, T., and Keim, D. 2013. "Visual analysis of social media data," *IEEE Computer* (46:5), pp. 68-75.
- Song, X., Chi, Y., Hino, K., and Tseng, B. 2007. "Identifying opinion leaders in the blogosphere," in *Proceedings of the sixteenth ACM conference on Conference on information and knowledge management*, ACM pp. 971-974.
- Wasserman, S., and Faust, K. 1994. *Social Network Analysis: methods and applications*, Cambridge University Press.
- Weng, L., and Menczer, F. 2014. "Topicality and Social Impact: Diverse Messages but Focused Messengers," in *Cornell University Library*, Cornell University Library: Web.
- WHO 2015. "Ebola virus disease," World Health Organization, <http://www.who.int/mediacentre/factsheets/fs103/en/>.
- Zeng, D., Chen, H., Castillo-Chavez, C., Lober, W. B., and Thurmond, M. 2011. *Infectious Disease Informatics and Biosurveillance*, Springer.
- Zeng, D., Chen, H., Lusch, R., and Li, S.-H. 2010. "Social Media Analytics and Intelligence," *IEEE Intelligent Systems* (25:6), pp. 13-16.
- Zhang, J., Ackerman, M. S., and Adamic, L. 2007. "Expertise networks in online communities: structure and algorithms," in *Proceedings of the 16th international conference on World Wide Web*, ACM New York, NY, USA, Banff, Alberta, Canada, pp. 221 - 230.