A20

Keyword analysis of the Relationship between Tweets and Shaking intensity after 3.11 Tohoku Earthquake and Tsunami Disaster

○Jietao Wu, Haruno Hayashi

Introduction

Twitter, as one of the most important ICT systems and the most popular social network, not only widely used by people, but also attracts the researcher in the field of information transfer and social informatics deeply. Twitter is a micro-blog service, allow people to broadcast information on their account, which could be seen by the other people who followed this account. Twitter became the most important way for people to do communication or public their information now.

Because of such kinds of characteristic, twitter now is studying by the researcher from the informatics. Tweet, which is the meaning of twitter message, are attracted us most when in the disaster. We could know what the people feel and do when the disaster happened and analyze their behaviors and the real environment after the disaster. That could help us to train people to take refuge and plan for the disaster prevention.

On March 11, 2011, the East Japan Earthquake happened. In this disaster, although the mobile communication (phone) stopped in the affected area, but because the mobile network still worked, twitter became the main method for people to communicate with each other. These tweets became the most important material for us to study this damaged earthquake.

In this research, we study about the tweet of March.11 2011, the first ten hours after the earthquake happened. We cared about the content of these tweets, and we supposed that the tweet would be affected by the time and shaking intensity. In this paper, we could see how the content of the tweet changed during the shaking intensity and time, and try to see the relationship between tweet, time and shaking intensity.

Method and Data Set

In this research, we used the data from the Great East Japan Earthquake Big Data Workshop (Project 311), which contained of more than one billion tweets, from 9:00, March 11 to 23:59, March 18, 2011. But we just used the data with shaking intensity and geography, from 9:00 to 23:59 in March 11, about 16,000 items of tweet.

Also, we used the Trend Reader as the analysis tool. Trend Reader is a keywords detection and analysis tool. By using it, we could get the keywords from the document set easily and have a simple keywords analysis at the same time.

After we got the keyword, we would make the keyword into 7 groups to see how these groups changed in the condition of time and shaking intensity.

Conclusion

In this research, we could see that when earthquake happened, people broadcast more tweets, but this phenomenon doesn’t last long, just about 10 hours. And for different shaking intensity and time, the number of tweet and the behavior of the group of keyword are quite different.

At the same time, by analyzing the three special types of tweets (tweet of website, Repeat tweet and tweet with @), we could also see the difference not only occurs in these three types but also different from the overall situation.