

A Study on the Spatiotemporal Informatization of Social Media Data for Obtaining Disaster Status

Changkyu Lee^a, Jinmu Choi^{b,*}

^a Ph.D. Student of Department of Geography, Kyung Hee University, Seoul, Republic of Korea, ehlfkrsk@khu.ac.kr

^b Department of Geography, Kyung Hee University, Seoul, Republic of Korea – cjm89@khu.ac.kr

* Corresponding author

Keywords: Disaster, Social Media, Spatiotemporal Informatization, Geocoder

Abstract:

In disaster situation, social media data can be used as a material for real-time crisis maps of natural disasters (Middleton et al., 2013). Social media can collect disaster information from a close distance and serve as a channel of communication that receives detailed on-site information (Doo et al., 2017). For this reason, existing studies performed the spatial analysis of social media data to identify disaster status (De Albuquerque et al., 2015). However, most of them dealt with only geotagged data, so most of the other data was inevitably missing (Meier, 2011). Since only a few of social media data is geotagged, spatiotemporal informatization through text analysis is required to utilize social media data as spatiotemporal information (Samuel and Sharma, 2021).

This study proposed a spatiotemporal method to utilize social media data for identifying disaster damages. Most of the social media data are not geotagged, and some geotags are just toponym entered by the author and do not contain any coordinates. Considering these characteristics of social media data, the proposed method (geocoder) proceeds in three stages (Fig. 1). As shown in Fig. 1(a), in a map search stage, API and crawler are used to give a coordinate to posts that have geotag but have not been coordinated. In this stage, geotagged data that contain coordinates will be ignored. In a dictionary comparison (geoparsing) stage, a gazetteer consisting of district names and Points of Interest (POI) of the South Korea is used. For example, in Fig. 1(b), the hashtags are compared with place names in the dictionary, and the coordinate of the most matched place name is used. In a similarity analysis stage, posts are compared with each other by a sentence similarity analysis method. As shown in Fig. 1(c), the coordinate of the post with the highest sentence similarity is used. This process is carried out for spatial informatization of non-geotagged data.



Figure 1. Procedure of proposed Geocoder method

As an example of applying the geocoder, among the typhoons that affected Korea in 2021, typhoons 'LUPIT', 'OMAS', and 'CHANTHU' which have more than 30 related Instagram posts were selected. As a result of applying the geocoder to posts collected for by each typhoon keyword, it was possible to convert more than 80% of data into spatial information, excluding data that already had location coordinates (Table 1). In the case of LUPIT, 46 out of 57 data were converted

except geotagged 39 out of 96 data. In the case of OMAIS, CHANTHU, most of data were converted, such as 462 out of 536 and 1,351 out of 1,513, except geotagged data.

In addition, As a result of performing spatiotemporal analysis for each typhoon using the obtained location and posting time, it was confirmed that the regional and temporal characteristics of posts changed according to the progress of the typhoon (Fig. 2).

Stage	Typhoon (Total posts)		
	LUPIT (96)	OM AIS (737)	CHAN THU (2,327)
Before application	39	201	814
Map search	3	17	216
Dictionary comparison	38	370	959
Similarity analysis	5	75	176
Geocoding Results	46	462	1,351

Table 1. Geospatial informatization result according to stage of the proposed Geocoder

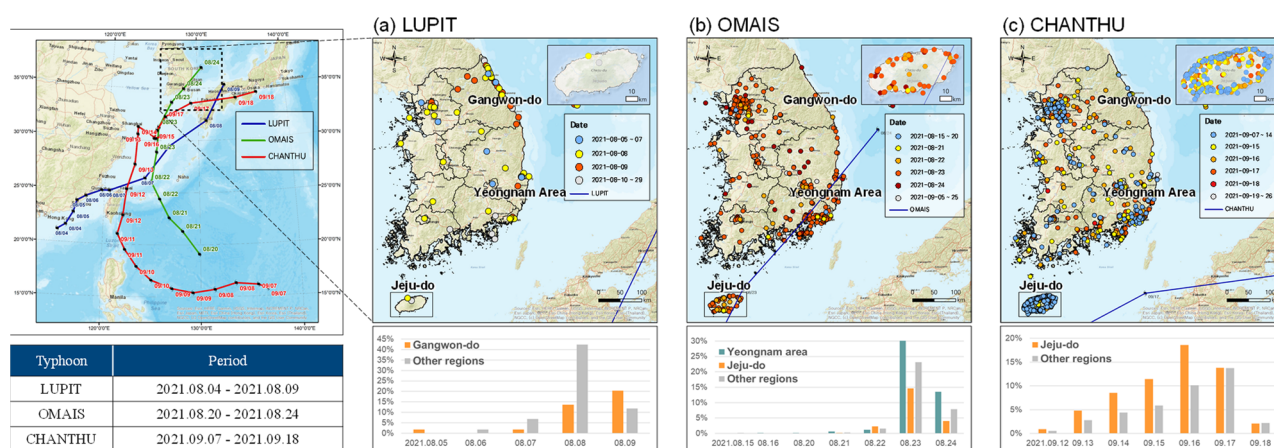


Figure 2. Spatiotemporal information of posts

In the future, it is expected that a more accurate and useful disaster monitoring tool can be produced by upgrading the geocoder and identifying the damage status using the collected images.

Acknowledgements

This research was supported by a grant (20009742) of Ministry-Cooperation R&D program of Disaster-Safety, funded by Ministry of Interior and Safety (MOIS, Korea)

References

- De Albuquerque, J. P., Herfort, B., Brenning, A. and Zipf, A., 2015. A geographic approach for combining social media and authoritative data towards identifying useful information for disaster management. *International journal of geographical information science*, Vol. 29, pp. 667-689.
- Doo, H. C., Park, J. H., Kim, H. Y., Oh, H. J. and Kim, Y., 2017. A Study on the Improvement and Analysis of SNS Operation Status on Disaster Information in Domestic and Foreign Public Institution. *Journal of the Korean BIBLIA Society for library and Information Science*, Vol. 28, pp. 57-78.
- Meier, P., 2011. New information technologies and their impact on the humanitarian sector. *International review of the Red Cross*, Vol. 93, pp. 1239-1263.
- Middleton, S. E., Middleton, L. and Modafferi, S., 2013. Real-time crisis mapping of natural disasters using social media. *IEEE Intelligent Systems*, Vol. 29, pp. 9-17.
- Samuel, A. and Sharma, D. K., 2021. Location estimation of non-geo-tagged tweets. *Evolutionary Intelligence*, Vol. 14, pp. 205-216.