

Sentiment-based spatial-temporal event detection in social media data



by LIN CHE

The emergence of social media provides a new source of sensing society [1]. Traditional event detection methods can only detect specified types of events based on relevant corpus.

In this thesis, a novel event detection method for detecting a wide variety of real-world social events based on population sentiment orientation (PSO) from social media check-in data is proposed.

The workflow of this method is composed of the following three main parts: sentiment analysis, spatial-temporal analysis based event detection and event interpretation (Fig. 1).

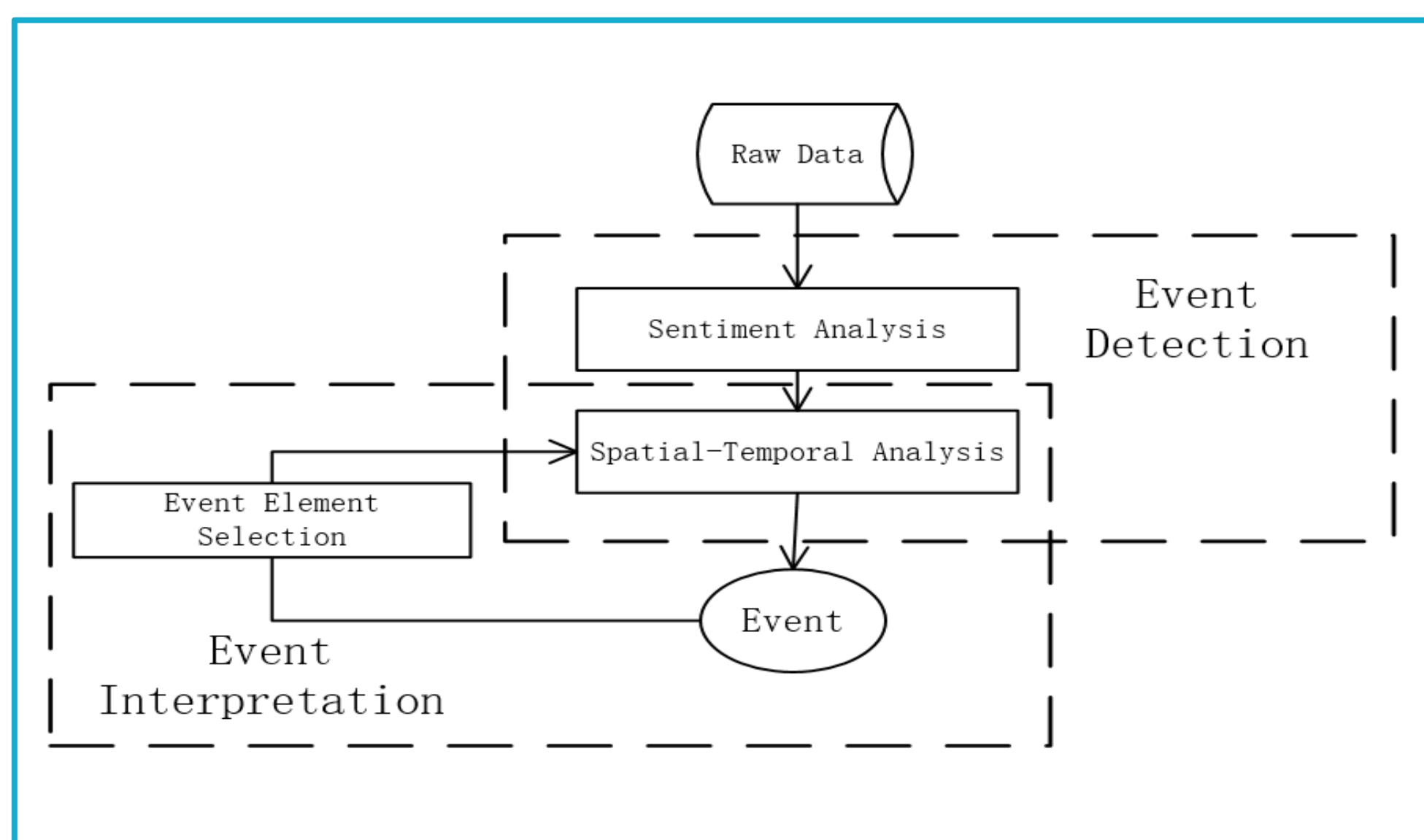


Fig.1: Thesis workflow chart

THESIS CONDUCTED AT

Chair of Cartography
Department of Civil, Geo and
Environmental Engineering
Technische Universität München



THESIS ASSESSMENT BOARD

Univ.Prof. Dr. Liqiu Meng, TUM

M.Sc. Juliane Cron, TUM

M.Sc. Ruoxin Zhu, TUM

Dr. Eva Hauthal, TUD

YEAR

2019

KEYWORDS

Spatial-temporal analysis, Event detection, Social media, Sentiment analysis, Machine Learning, Data mining

SENTIMENT ANALYSIS

The ratio of the number of positive and negative records is chosen to indicate the PSO within a specified period and geographical area. In order to calculate PSO, the dataset needs to be classified by the sentiment polarity. Lexicon-based and Machine learning-based classification methods [2] are compared and tested in this thesis.

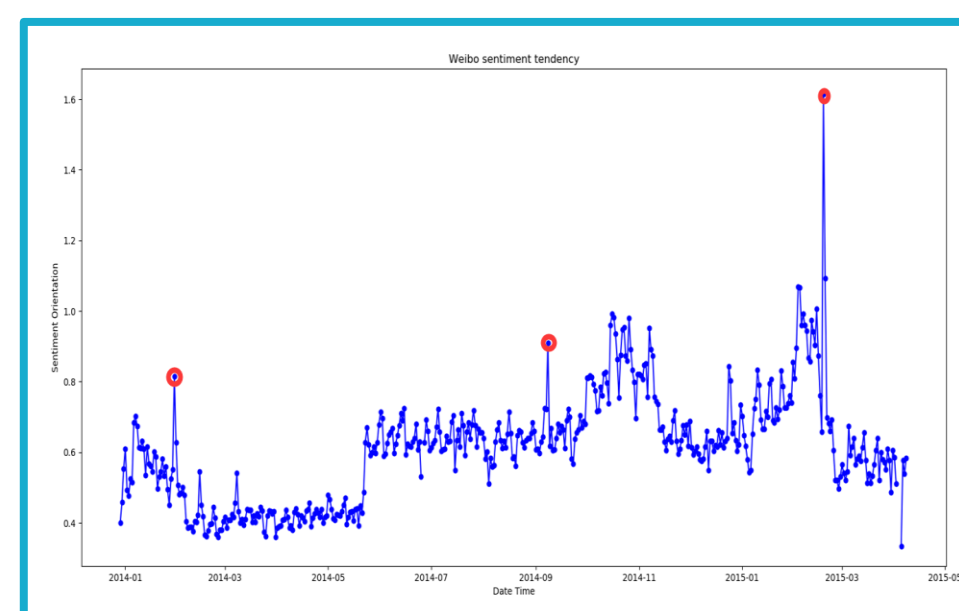


Fig. 2: Time series plot of PSO in Shanghai, 2014



Fig. 3: The Word Cloud of 18.02.2015 ("Happy New Year", "Spring Festival Gala", "New Year's Eve")

SPATIAL-TEMPORAL ANALYSIS-BASED EVENT DETECTION

A case study on Shanghai is conducted for testing the methodology. The data is from Sina Weibo, a Chinese microblogging website. Fig. 2 shows the time series plot and the extrema indicating the events. Fig. 3 shows the corresponding Word Cloud generated from the first local maximum point. We can easily visually learn that that day is the Chinese New Year's Eve.

EVENT INTERPRETATION

For sensing negative events on the Chinese New Year's Eve of 2014, a Word Cloud is generated as well. Negative words found are related to traffic like "Working" or "Going home. Therefore, a spatial clustering analysis [3] is conducted for exploring the traffic related negative microblogs spatial patterns. Fig. 5 shows the clustering results as sample multiples. Fig. 6 shows the clusters are distributed around Airports and railway stations.

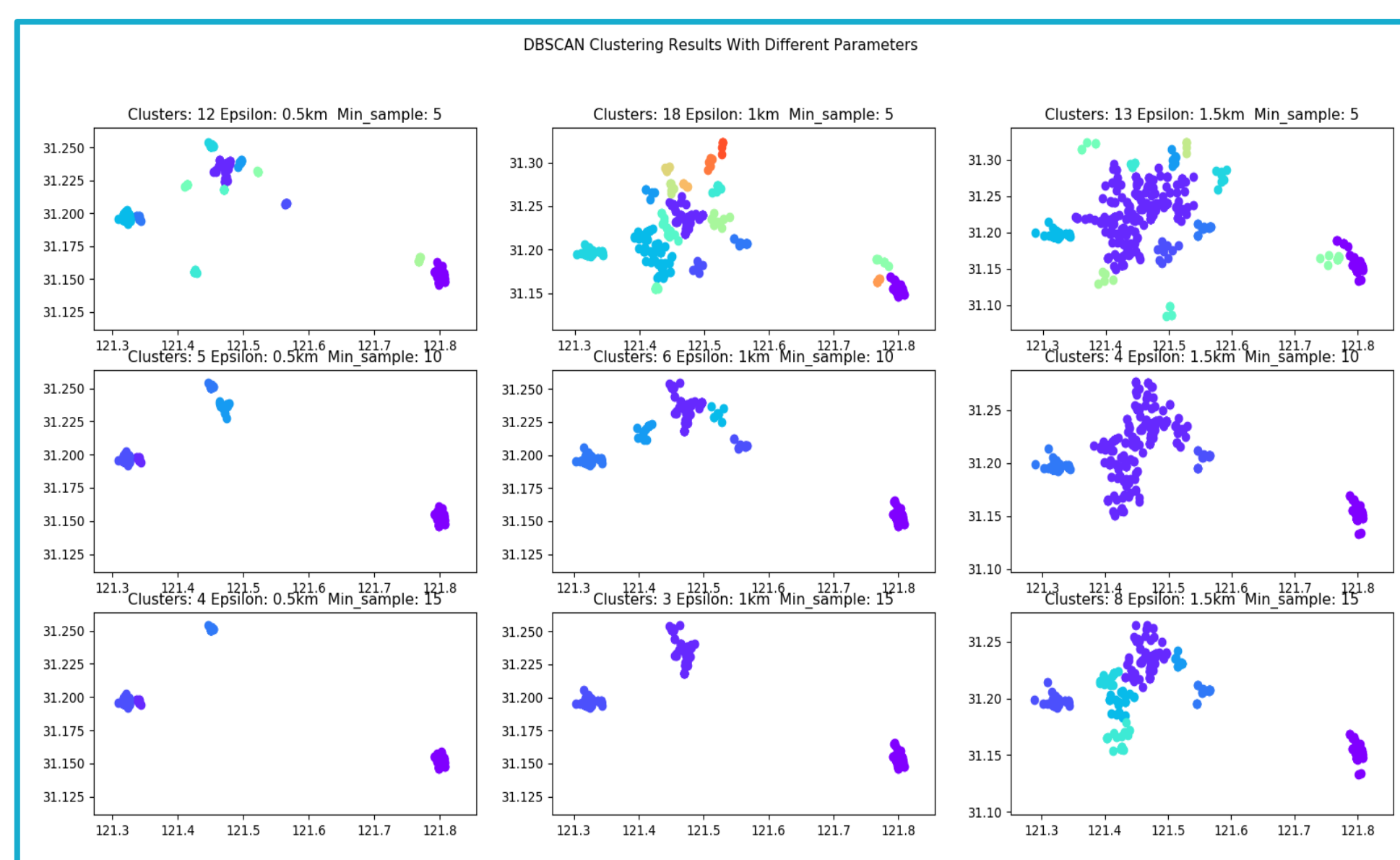


Fig. 5: Sample multiple of clustering results

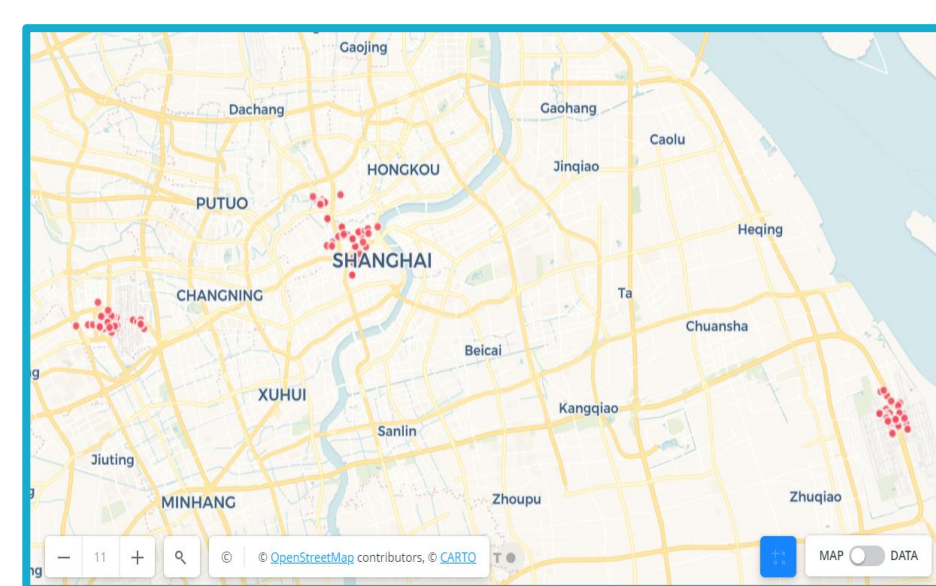


Fig. 6: The cluster map

CONCLUSION

In this thesis, it is first proven that PSO can be used to detect events from social media data. Besides, the ratio of positive and negative records can serve as an indicator for PSO. Second, this event detection method can successfully detect multi-scale events. Finally, this method can also be used to interpret different events.

REFERENCES

- [1] Liu, Y. et al. (2015). Social sensing: A new approach to understanding our socioeconomic environments. *Annals of the Association of American Geographers*, 105(3), 512-530.
- [2] Medhat, W., Hassan, A., and Korashy, H. (2014). Sentiment analysis algorithms and applications: A survey. *Ain Shams engineering journal*, 5(4), 1093-1113.
- [3] Ester, M. et al. (1996). A density-based algorithm for discovering clusters in large spatial databases with noise. *Kdd*, 96(34), 226-231.