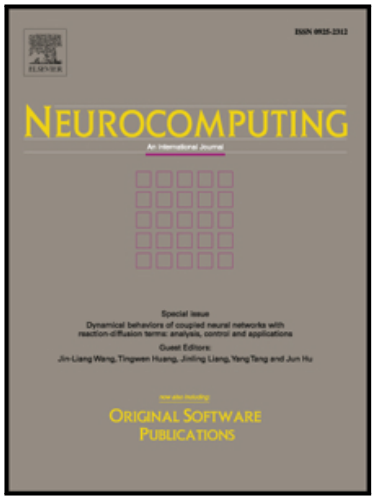
Communicated by Dr. Y. Chang

Accepted Manuscript



Detecting Global and Local Topics via Mining Twitter Data

Huan Liu, Yong Ge, Qinghua Zheng, Rongcheng Lin, Huayu Li

PII: S0925-2312(17)31376-0
DOI: 10.1016/j.neucom.2017.07.056
Reference: NEUCOM 18766

To appear in: *Neurocomputing*

Received date: 24 December 2016
Revised date: 13 May 2017
Accepted date: 31 July 2017

Please cite this article as: Huan Liu, Yong Ge, Qinghua Zheng, Rongcheng Lin, Huayu Li, Detecting Global and Local Topics via Mining Twitter Data, *Neurocomputing* (2017), doi: 10.1016/j.neucom.2017.07.056

Detecting Global and Local Topics via Mining Twitter Data

Huan Liu[a,*], Yong Ge[b], Qinghua Zheng[a], Rongcheng Lin[c], Huayu Li[c]

[a]*MOEKLINNS Lab, Department of Computer Science and Technology, Xi'an Jiaotong University, Xi'an, China*
[b]*Nanjing University of Finance and Economics, Nanjing, China*
[c]*Department of Computer Science, University of North Carolina at Charlotte,Charlotte, USA*

Abstract

Detecting topics from Twitter has been widely studied for understanding social events. There are two types of topics, i.e., global topics attracting widespread tweets with larger volume and local topics drawing attention of limited tweets of somewhere. However, most of existent works neglect the difference between them and suffer from the *Long Tail Effect*, resulting in the inability to detect the local one. In this paper, we distinguish global and local topics by associating each tweet with both of them simultaneously. We propose a probabilistic graphical model to extract global and local topics related to social events in a unified framework at the same time. Our model learns global topics using tweets scattered around all locations, while studies local topics merely utilizing tweets within the corresponding location. We collect two tweet datasets on Twitter from several cities in USA and evaluate our model over them. The experimental results show significant improvement of our model compared to baseline methods.

*Corresponding authors
Email addresses: hliuxjtu@gmail.com (Huan Liu), ygestrive@gmail.com (Yong Ge), qhzheng@mail.xjtu.edu.cn (Qinghua Zheng), rlin4@uncc.edu (Rongcheng Lin), hli38@uncc.edu (Huayu Li)