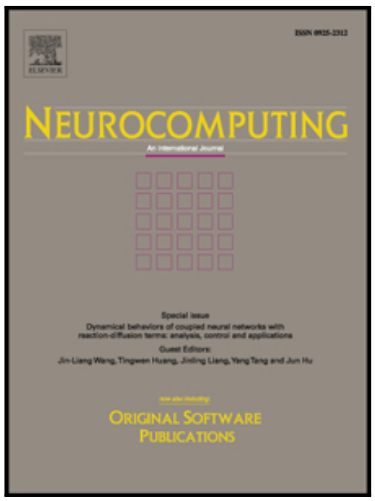

Real-time personalized twitter search based on semantic expansion and quality model

Xiang Zhu, Jiuming Huang, Bin Zhou, Aiping Li, Yan Jia

PII: S0925-2312(17)30403-4
DOI: 10.1016/j.neucom.2016.10.082
Reference: NEUCOM 18153

To appear in: *Neurocomputing*

Received date: 27 January 2016
Revised date: 17 October 2016
Accepted date: 25 October 2016

Please cite this article as: Xiang Zhu, Jiuming Huang, Bin Zhou, Aiping Li, Yan Jia, Real-time personalized twitter search based on semantic expansion and quality model, *Neurocomputing* (2017), doi: 10.1016/j.neucom.2016.10.082

Real-time personalized twitter search based on semantic expansion and quality model

Xiang Zhu[a,*], Jiuming Huang[a], Bin Zhou[a], Aiping Li[a], Yan Jia[a]

[a] *College of Computer, National University of Defense Technology*

Abstract

The vast amount of information in social networks makes it difficult for users to find what they want, users may get drowned in the information flood. It is a challenging problem to retrieve the high quality and relevant information according to a user's searching query. Traditional methods for personalized search become insufficient in social networks due to the high velocity, topic variety, data sparseness and high sociability. To overcome those difficulties, we propose a novel framework for real-time personalized twitter search for twitter stream in this paper. Firstly, we develop a boolean logic keyword filter to enhance the accuracy. Then a tweet quality model is built to distinguish high quality tweets, it could improve the ranking performance. After that, we utilize an external search engine to implement query expansion, which could understand user preferences and interests properly. Our framework integrates the semantic features and social attributes which are utilized to make a comprehensive rank for a tweet. In addition, we adopt a dynamic strategy to push high quality and relevant tweets to a user automatically to avoid information overload. A thorough evaluation is conducted using real twitter stream data in TREC 2015, demonstrating a superior performance against competitive baselines in a variety of metrics.

Keywords: social network, personalized search, semantic computing, quality model

*Corresponding author
Email address: zhuxiang@nudt.edu.cn (Xiang Zhu)