Chapter 10
News Recommendation for China Sina Weibo Microblog Service Based on User Social Behaviors

Zuo Yuchu
Sun Yat-sen University, China

You Fang
Sun Yat-sen University, China

Wang Jianmin
Sun Yat-sen University, China

Zhou Zhengle
Sun Yat-sen University, China

ABSTRACT

Sina weibo microblog is an increasingly popular social network service in China. In this work, the authors conducted a study of detecting news in Sina weibo microblog. They found the traditional definition for news can be generalized here. They first expanded the definition of news by conducting user surveys and quantitative analysis. The authors built a news recommendation system by modeling the users, classifying them into four different groups, and applying several heuristic rules, which derived from the generalized definition of news. By applying the new recommendation system, people got newsworthy information, while the funny and interesting tweets, which are popular in Sina weibo microblog, were put in the last ranking list. This study helps us achieve better understanding of heuristic rules about news. Some official organizations can also benefit from the work by supervising the most popular news around civilians.

1. RESEARCH BACKGROUND

As a new type of social media, Sina weibo microblog has become significantly popular in China nowadays. So far, considerable amount of work has been done in order to classify the tweets (or status) automatically. Current classification researches on Chinese microblog are mostly based on supervise learning (Alec Go, 2009; Dong Zhendong, 2011; Z. Liu, Yu, Chen, Wang, & Wu, 2010; Zitao, Wenchao, Wei, Shuran, & Fengyi, 2010) and semantic network (Okazaki & Matsuo,

2010), thus the tweets are classified into two categories—the positive and the negative(Qin, Xin, & Niu, 2010). Nevertheless, both supervise learning and semantic network are not so efficacious because of some inevitable shortcomings. For instance, supervise learning demands quite a large dataset while the constantly changing content in microblog makes it difficult to build a robust system. On the other hand, considering it has been noted that microblog possesses the attributes of news media (Kwak, Lee, Park, & Moon, 2010), it is necessary to classify it as news-worthy type and non-news-worthy type. As it was mentioned in (Kwak, et al., 2010), microblog is more of news media than social network, which can be justified by our basic motives of using microblog. (Java, Song, Finin, & Tseng, 2007) said a single user may have multiple intentions or may even play different roles in different communities. They may be an information source, a social relationship builder, an information seeker, or the one combining some intentions above. In Sina weibo microblog, users' intentions are quite the same with those of the Twitter's users since both have the same mechanism. Firstly, it's quite popular for us to raise and discuss issues or topics on microblog when news happens in the real world. Secondly, we pay attention to or "follow" celebrities who are "verified" as well as our friends in real life because we are eager to know what happen to them recently. Thirdly, we concern a special topic where our own interest lies and communicate with others in the common-interest communities. As a matter of fact, however, we are frequently harassed by piles of spam information, and unfortunately miss much news-worthy one. Hence, it's really imperative and significant for us to come up with one new microblog classification criteria and rule to optimize the microblog management and provide better user experience.

In this paper, we found that some user groups are more possible to post news-worthy content. They are verified-organizational user and unver-

ified-personal user. What's more, we are going to build a system to find these tweets out of piles of non-news-worthy content.

2. RELATED WORKS

In previous content-based classification researches which mainly center on the Chinese sentence structure and syntax analysis, we encountered such a fatal bottleneck: constantly newly-emerged words and usages. In fact, such a problem results from the complexity and flexibility of Chinese which is so different from English. Even some researchers solve this problem by building mathematical model (Abdullah & Wu, 2011), considering users' relations in social network(Magnani, Montesi, & Rossi, 2010), or combining social network with RSS feeds (Phelan, McCarthy, Bennett, & Smyth, 2011a, 2011b; Phelan, McCarthy, & Smyth, 2009) (though they only recommend RSS news by tweets, we can easily reverse the idea by recommending tweets by RSS feeds), a subtle problem which is omitted is that people use microblog not only for the news which would appear on every portal site, but also for some social purpose (Chen, Nairn, & Chi, 2011; Zhao & Rosson, 2009).

Apparently microblog is not a typical medium, and the traditional definition for "news" can be generalized here. As far as we know, this is the first study on finding the characters of the news-worthy content in microblog. We interviewed 37 Sina weibo microblog users, who have used Sina weibo microblog service over a year, and asked them the generalized definition for "news", which drove us to redefine "news" in the microblog-sphere. After that, we found evidences that can be used to decide whether a tweet is news-worthy. Back to previous research of (Guangxia, Hoi, Kuiyu, & Jain, 2010; Jansen, Zhang, Sobel, & Chowdury, 2009; Okazaki & Matsuo, 2010; Qin, et al., 2010), the existing classification method

22 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage: www.igi-global.com/chapter/news-recommendation-for-china-sina-weibo-microblog-service-based-on-user-social-behaviors/88817

Related Content

Kinetic User Interfaces: Physical Embodied Interaction with Mobile Ubiquitous Computing Systems
Vincenzo Pallotta, Pascal Brueggerand Béat Hirsbrunner (2008). *Advances in Ubiquitous Computing: Future Paradigms and Directions (pp. 201-228).*
www.irma-international.org/chapter/kinetic-user-interfaces/4923

Falling Behind: A Case Study in Uncritical Assessment
Jonathan G.M. Pratt (2009). *Risk Assessment and Management in Pervasive Computing: Operational, Legal, Ethical, and Financial Perspectives (pp. 102-133).*
www.irma-international.org/chapter/falling-behind-case-study-uncritical/28452

Pervasive and Ubiquitous Computing Databases: Critical Issues and Challenges
Michael Zoumboulakisand George Roussos (2010). *Ubiquitous and Pervasive Computing: Concepts, Methodologies, Tools, and Applications (pp. 1517-1526).*
www.irma-international.org/chapter/pervasive-ubiquitous-computing-databases/37865

State of the Art Recommendation Approaches: Their Issues and Future Research Direction in E-Learning A Survey
Bhupesh Rawatand Sanjay K. Dwivedi (2018). *International Journal of Advanced Pervasive and Ubiquitous Computing (pp. 51-76).*
www.irma-international.org/article/state-of-the-art-recommendation-approaches/199776

RFID as the Critical Factor for Superior Healthcare Delivery
A. Dwivediand T. Butcher (2010). *Ubiquitous and Pervasive Computing: Concepts, Methodologies, Tools, and Applications (pp. 823-832).*
www.irma-international.org/chapter/rfid-critical-factor-superior-healthcare/37821