Li WEIGANG, Edans F. O. SANDES, Jianya ZHENG, Alba C. M. A. de MELO, Lorna UDEN, 2014. Querying dynamic communities in online social networks. *Journal of Zhejiang University-Science C (Computers & Electronics)*, 15(2):81-90. [10.1631/jzus.C1300281]

Querying dynamic communities in online social networks

在线社交网络内动态群组查询

Key words: Follow Model, Hadoop, MapReduce, Querying, Twitter

关键词：粉丝模型，**Hadoop**，映射和化简，信息查询，**Twitter**微博

Major discussion in this paper

- Querying in a dynamic group from Twitter is a challenge. The Follow Model is proposed to present the basic relationship between users in Twitter, and it combines the MapReduce solution to develop new algorithms with parallel paradigms for querying
- The new solution with the implementation in Hadoop significantly improves the ability to find useful information from Twitter

Due to the dynamic real-time characteristics, querying an incident is a challenge in Twitter

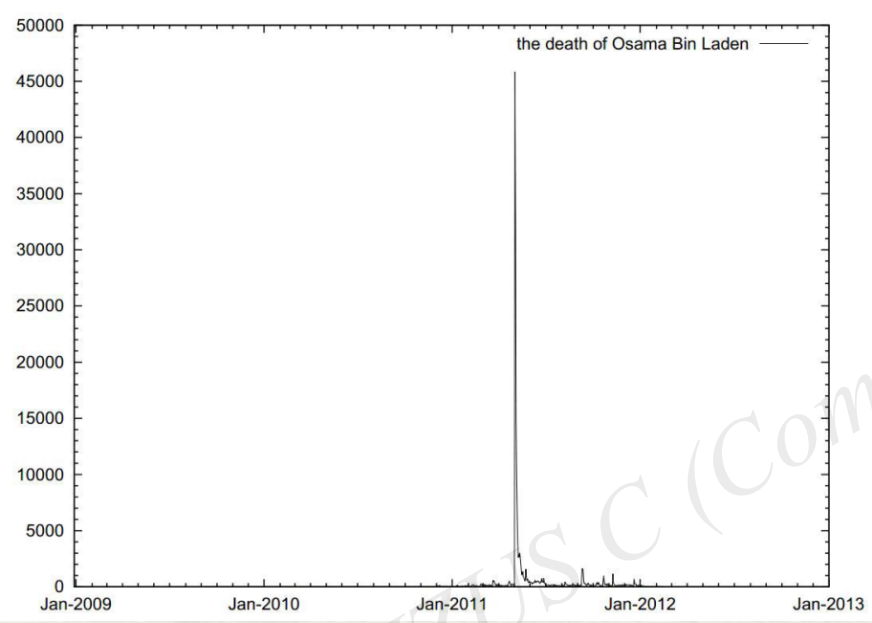


Fig. 1 The distribution of messages about the death of Osama bin Ladin. On Sina Weibo, 54,000 messages were issued by users that day. The distribution of the number of messages showed a strong pulse

Table 4 Top 10 in the group with event #beatcancer in Twitter (June/2009-Dec/2009)

User Name	@Twitter	Followers	Event Followers
Pete Cashmore	@mashable	3,227,280	2,494
Stephanie Pratt	@stephaniepratt	638,059	644
Charlie Brooker	@charltonbrooker	715,148	535
Mikey Way	@mikeyway	284,845	485
Jeanette Joy Fisher	@JeanetteJoy	117,040	409
Richard Bacon	@richardpbacon	1,487,418	405
Calvin Harris	@CalvinHarris	2,149,058	356
Shannon Seek	@shannonseek	59,312	343
Don Lemon	@DonlemonCNN	177,646	331
Calvin Lee	@mayhemstudios	81,211	331

Table 4 The group to discuss the topic #beatcancer consists of 66,846 users. There were 26,930 users' relations available from the dataset. There were a total of 8,668,545 relations related to these users in the entire relation network of Twitter

Conclusions

- The paper described the mechanisms of Twitter and the relationships/actions of the users using the Follow Model, and proposed a new querying solution with the MapReduce concept
- The developed algorithms with parallel computation property was implemented in the Hadoop system
- The performance of computation showed positive results of the algorithms in parallel paradigms for information querying in Twitter