

An Effective Graph-based Framework for Name Disambiguation

Jianyong Wang*

Department of Computer Science and Technology
Tsinghua University, Beijing, China
jianyong@tsinghua.edu.cn

Abstract:

Name ambiguity is very common in real life and may cause much confusion in many cases if we use names to identify specified persons. In general, the problems caused by name ambiguity can be classified into two categories, namely, reference disambiguation and name disambiguation. In this talk, I will introduce our recent work on name disambiguation. Based on the observation that graph structure can be used to naturally model complex relationships among entities (and their attributes too), we proposed a very effective graph-based framework for name disambiguation in both digital library and Web search settings. The framework involves graphic modeling, similarity computation between any two nodes, and graph clustering. Under this unified framework, I will introduce two algorithms, GHOST and GRAPE, proposed by our group for name disambiguation in digital library and Web search settings, respectively. If time permits, I would also like to demonstrate the GRAPE system for disambiguating people appearances in Web search.

Keywords: Name Disambiguation, Graph Clustering, GHOST, GRAPE

References:

- Xiaoming Fan, Jianyong Wang, Xu Pu, Lizhu Zhou, Bing Lv. On Graph-based Name Disambiguation. *ACM Journal of Data and Information Quality*, Vol. 2, No. 2, Article 10. February 2011. (**ACM JDIQ**)
- Lili Jiang, Wei Shen, Jianyong Wang, Ning An. GRAPE: A System for Disambiguating and Tagging People Names in Web Search. *Proc. 19th International World Wide Web Conference, Raleigh, NC, USA, April 26-30, 2010.* (**WWW'10**, demo paper)
- Lili Jiang, Jianyong Wang, Ning An, Shengyuan Wang, Jian Zhan, Lian Li. GRAPE: A Graph-Based Framework for Disambiguating People Appearances in Web Search. *Proc. the 2009 IEEE International Conference on Data Mining, Miami, FL, USA, Dec. 6-9, 2009.* PP: 199- 208. (**IEEE ICDM'09**)

* corresponding author. Tel: