Design of a Visualized Information Retrieval Assistance System Based On Graphic Clustering

許銘維、陳鴻文

E-mail: 9708161@mail.dyu.edu.tw

ABSTRACT

Each kind of fast distributed information through computer networks mutually interacts and forms new knowledge nowadays. However, all full-text information retrieval systems such as Google have no suitable mechanisms to synthesize correlated glossaries and to illustrate their global information structures. To simulate human knowledge concepts from the connections of a vocabulary graph, it was attempted to construct reasoning patterns of information by applying a graph-based clustering algorithm to classify and induce more information, and by using colored indication to auxiliary mark up important vocabularies. A visualized information retrieval assistance system was designed and implemented based on the mentioned concepts above. The proposed system can assist users to retrieve information more intuitively and following the knowledge constitution way. Finally, a questionnaire survey was used to evaluate the satisfaction degrees of system users. Some achievements are reached as following: 1.A visualized information retrieval assistance system was implemented to visually present information and to moderately guide users in retrieving information. 2.Correlated glossaries are acquired and organized through Internet to assist users to understand the related knowledge of searching keywords. 3.Users are guided to retrieve information more precisely by clustering related glossaries to provide the information of important vocabularies and previous retrieval paths.

Keywords : graph-based clustering ; data visualization ; information retrieval

Table of Contents

REFERENCES

一、中文部份 中文詞彙網路(2008)，中文詞彙網路[線上資料]，來源: http://cwn.ling.sinica.edu.tw/.

[2008, April 4]。 二、英文部份 Beate, D. (2006). A Graph Model for Words and their Meanings. PhD thesis, Institute of Computational Linguistics, University of Stuttgart. Duncan, J. W., & Steven, H. S. (1998). Collective dynamics of ’small-world’ networks. Nature, 393, 440 – 442. Google Labs. (2008). Google Labs [Online]. Available: http://labs. google.com.

[2008, May 2]. Google Sets. (2008). Google Sets [Online]. Available: http://labs. google.com/sets.

[2008, May 10]. Grokker. (2008). Gokker [Online]. Available: http://www.grokker. com.

[2008, May 5]. Stijn, V. D. (2000). Graph Clustering by Flow Simulation. PhD thesis, University of Utrecht. TouchGraph. (2008). TouchGraph

[Online]. Available: http://source forge.net/projects/touchgraph/.
[2007, Oct 28]. TouchGraph Google Browser. (2008). TouchGraph Google Browser [Online]. Available: http://www.touchgraph.com/TGGoogle Browser.html.
[2008, May 2]. WordNet. (2008). WordNet [Online]. Available: http://wordnet. prin-ceton.edu.tw/.
[2007, Nov 2]. Wikipedia. (2008). Wikipedia [Online]. Available: http://www. wiki-pedia.org/.
[2008, May 10].