Hierarchy-Cutting Model based Association Semantic for Analyzing Domain Topic on the Web

ABSTRACT

Web information contains plentiful, significant knowledge including explicit and implicit knowledge [1]. How to organize the Web information for facilitating knowledge discovery has been deeply investigated by some researchers. Association link network (ALN) is a kind of semantic link network built by mining the association relations among Web resources for effectively supporting Web intelligent application such as Web semantic association search, Web knowledge discovery, and recommendation [2, 3]. Xu et al. have studied on cloud environment for surveillance data management using video structural description [4], generating temporal semantic context of concepts [5]. Zhu et al. present discovering and learning communities and emerging semantics in semantic link network [6]. With the rapid development of information technology, human kinds are more likely to read and share information by similar intelligent applications.

EXISTING SYSTEM

Effective semantic layered technology not only can provide theoretical support for knowledge discovery in Web resources, but also can improve the searching efficiency of the related information system such as Web information system and industrial information system. How to realize the layer division of association semantic by the hierarcy analysis of ALN is an important research topic.

DRAWBACKS

- The existing methods fall short of multiple-layer model of large-scale Web resources for supporting the discovery of domain topic.
- The existing network analysis mainly focuses on the characteristic analysis; there is no method or algorithm of division of association semantic to satisfy different users' information requirement.

PROPOSED SYSTEM

This paper proposes a hierarchy-cutting model of association semantic. First, some experiments of four types of keywords with different linking roles are conducted to discover the possible distribution law. Experimental results show that these keywords with association role reveal previous power-law distribution. Then, based on the discovered power-law distribution, up-cutting and down-cutting points are presented to divide the association semantic into three layers. At the same time, the theories of hierarcy-cutting model are presented. Finally, the examples of

the current core topic and permernent topics belonging to a domain are given. The experiments show that hierarcy-cutting points have high accuracy. The multilayer theory of association semantic can provide a theoretical support for knowledge recommendation with different particle sizes on ALNs.

ADVANTAGES

- Keywords with association role reveal previous power-law distribution.
- The discovered power-law distribution, up-cutting and down-cutting points are presented to divide the association semantic into three layers.

SYSTEM REQUIREMENTS

- **H/W System Configuration:-**

- Processor - Pentium –IV
- RAM - 4 GB (min)
- Hard Disk - 20 GB
- Key Board - Standard Windows Keyboard
- Mouse - Two or Three Button Mouse
- Monitor - SVGA

- **S/W System Configuration:-**

- Operating System : Windows 7 or 8 32 bit
- Application Server : Tomcat5.0/6.X
- Backend coding : Java
- Tool : Virtual Box
- Environment : Ubuntu
- Technology : Hadoop