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

ABSTRACT:

Association link network (ALN) can organize massive web information to provide many intelligent services in the era of Big Data. 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. To solve this problem, 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 permerment 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.

INTRODUCTION:

web information contains plentiful, significant knowledge including explicit and implicit knowledge. 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 . Xu et al. have studied on cloud environment for surveillance data management using video structural description, generating temporal semantic context of concepts. Zhu et al. present discovering and learning communities and emerging semantics in semantic link network. With the rapid development of information technology, human kinds are more likely to read and share information by similar intelligent applications. For example, the distributed and collaborative learning, semantic representation of scientific documents for supporting e-learning , discovering and searching of correlation between shared resources , and smart component technologies for human-centric computing ETC