Page Value Computation and Network Analysis by Link Structure Analysis

Huang LinChun

School of Knowledge Science,
Japan Advanced Institute of Science and Technology
March 2000

Keywords: WWW information space, web robot, object directivity, artificial network, link value.

As the largely growth of Internet infrastructure, many people communicate useful information or knowledge in WWW. However, the distribution mechanism of them is not yet clarified. Also, it is still difficult to select appropriate information from huge lists of many search engines, for the demand of information resources discovery. There are many problems in such keyword-based method, because user's needs are usually not well-defined, and sometime changed depending on various situations. In this research, we will focus the links structure which suggest something of values of pages in a network community. We propose an efficient distributed computing calculation method of the link structure, and analyze the characteristics of the network.

First, We define two types of links called Authority link and Hub link, and set the values of them as the numbers of in-degrees and out-degrees. We assume person's sense and knowledge are contained in the link. Indeed, a interesting page is linked by many pages. If a page links many pages, this page will be regarded as a useful or convenient page. Therefore, it is possible to think that the link value of the page depends on it. To extract the hidden Hub link structure from the visible Authority-link as references in each HTML page, we proposed an algorithm to analyze Hub link structure.

Next, we developed a system using multi-threading technology on Java. It's performance with powerful network function API on multi-platforms. We analyzed 25 artificial networks and 12 actual networks in this experiment, then analyzed about the characteristics of the

network. To explain the combination situation of the link, we made some artificial networks. The connection structures ranging from regular, middle, and to disorder are considered. As the result, we found there is a relation between network structure and user preference depend on each community.

In particular, the network could be mainly classified into the following 4 types, which are evaluated by the items: the randomness of the link, the combination percentage of the link, the degree of the network of the opening, average AUT value and average HUB value.

- Open network, structure with highly Hub-connections.
- Close network, structure with highly Hub-connections.
- Open network, structure with highly Authority-connections.
- Close network, structure with highly Authority-connections.

The proposed system will be contributed for the automatic recommendation by the links value, another application of web robots, and support of creating a new community.