Chapter III

Using Web Link Analysis to Detect and Analyze Hidden Web Communities

Edna O.F. Reid
Nanyang Technological University, Singapore

ABSTRACT

A great deal of current as well as previous studies on web links has focused mostly on improving the performance of information retrieval systems. The vast but untapped wealth of information from link-related messages generated by online communities has yet to attract the attention of the competitive intelligence researchers and practitioners. The latter groups have depended mainly on traditional intelligence sources while cognizant that much of the information which impinges upon their competitive strengths is shaped by events external to the firm. In view of the foregoing, we present in this chapter an exploratory framework for extracting and exploiting patterns of self-organizing, hyperlinked web communities for...
corporate intelligence purposes. More specifically, this chapter summarizes how the proposed analytical framework has been applied to MicroStrategy Inc.’s website to give us a glimpse of stakeholder communities’ reactions to the enterprise’s activities and identify some early warning signals. The framework can thus be considered as a prototypical approach for exploiting the Web’s structure and content for Web intelligence purposes.

INTRODUCTION

The World Wide Web contains a huge amount of interconnected Web pages that are authored and made available by millions of different individuals. Consequently, it provides new opportunities for using the hyperlink structure as a creative data source for identifying and analyzing snapshots, continuously and historically, of a company’s implicit stakeholder communities who share common interests in a firm. The snapshots provide bits and pieces of a jigsaw puzzle that can be assembled into patterns of relations, activities and early warning signs of developments about a firm’s external environments. More importantly, capturing dispersed intelligence data about a company’s implicit online communities of stakeholders should therefore be possible given the right Web mining and analytical methodologies.

The hyperlinks are reflections of social interactions between stakeholders and a company. Collectively, they represent an underlying social structure of linked communities. These hyperlinked communities are implicit and “natural” in that they are self-organizing (Flake et al., 2002).

To illustrate the notion of implicit hyperlinked communities, consider the community of Web users interested in Porsche Boxster cars (Kumar et al., 1999). There are numerous explicit online communities that are readily available and easy to identify, such as the Porsche newsgroup and Porsche owners’ listserv. Porsche owners subscribe to the listserv and receive their messages via e-mail. On the other hand, there are Web users who create content and provide hypertext links to Porsche Boxster resources (www.porsche.com/english/boxster). These users may not be registered parties in the Porsche newsgroup or listserv but they share in various areas of interests in Porsche Boxster cars. This gave rise to different Web pages, e.g., www.porschefaq.com and www.autopicture.com/porsche, that are hyperlinked to a website (www.porsche-com/english/boxster). According to Kumar et al., these users can be considered as members of implicit hyperlinked communities because the communities are not obvious (hidden), lack definite membership, have spontaneously evolved, and have an implied social structure.
Trends and Research of Wikis' Quality and Governance Based on Bibliometric and Content Analysis
Qinghua Zhu, Linghe Huang, Jia Tina Du and Hua Liu (2014). *Information Quality and Governance for Business Intelligence* (pp. 148-166).

Reliability Analysis for Degrading Systems with 100% Quality Inspection after Burn-In
[www.igi-global.com/article/reliability-analysis-for-degrading-systems-with-100-quality-inspection-after-burn-in/115519?camid=4v1a](http://www.igi-global.com/article/reliability-analysis-for-degrading-systems-with-100-quality-inspection-after-burn-in/115519?camid=4v1a)