

数字图书馆

基于叙词表的主题爬虫技术研究*

夏崇镛¹;康丽^{1,2}

中国农业大学计算机系¹

收稿日期 2007-2-6 修回日期 2007-3-26 网络版发布日期 2007-5-25 接受日期

摘要 将叙词表同传统的信息检索技术相结合,提出用叙词表的族对爬虫的主题进行描述的方法并用该方法设计实现一主题爬虫。实验结果证明本文提出模型和算法的有效性。

Abstract Combining the thesaurus with the traditional information retrieval technology, a new method is presented that family in thesaurus is used to describe the predefined topic. Also develops a focused-crawler based on this method. On which we compare its efficiency with other well-known Web search engine. The experimental results shows the effectiveness of our models and algorithms.

关键词 [主题爬虫](#) [叙词表](#) [搜索引擎](#)

Key words Focused crawler; Thesaurus; Search engine

分类号 [TP393](#)

DOI:

通讯作者:

夏崇镛 xiachongpu@cau.edu.cn

作者个人主页: 夏崇镛 康丽

扩展功能

本文信息

▶ [Supporting info](#)

▶ [PDF](#) (OKB)

▶ [\[HTML全文\]](#) (OKB)

▶ [参考文献\[PDF\]](#)

▶ [参考文献](#)

服务与反馈

▶ [把本文推荐给朋友](#)

▶ [加入我的书架](#)

▶ [加入引用管理器](#)

▶ [引用本文](#)

▶ [Email Alert](#)

▶ [文章反馈](#)

▶ [浏览反馈信息](#)

相关信息

▶ [本刊中 包含“主题爬虫”的 相关文章](#)

▶ 本文作者相关文章

• [夏崇镛](#)

• [康丽](#)

•