

信息检索技术

基于多语境的相关词自动提取系统的设计与实现

章成志¹; 苏兰芳¹; 苏新宁^{1,2}

南京大学信息管理系¹

收稿日期 2006-6-5 修回日期 2006-6-9 网络版发布日期 2006-9-15 接受日期

摘要 利用语料库、释义词典、用户检索日志作为识别相关词的语境,设计并实现相关词自动提取系统。实验结果表明,虽然面向相同的基本词汇集合,但是基于不同语境提取的相关词之间的重复率很低,各个结果间的互补性很强,说明结果整合非常有必要。在本系统中,通过直接整合途径构建最后的相关词词表。

Abstract This paper chooses corpus, definitions dictionaries and users' query logs as contexts to extract the relevance terms. The experiment results show that the overlap ratio of results in different contexts is very low. So, it is necessary to integrate the different results. All of the relevance terms are integrated to a relevance table through direct integration.

关键词 [相关词](#) [多语境](#) [语料](#) [释义词典](#) [用户日志](#)

Key words Relevance term; Multi-context; Corpus; Definitions dictionary; Query log

分类号 [TP391G252](#)

DOI:

通讯作者:

章成志 zcz51@citiz.net

作者个人主页: 章成志 苏兰芳 苏新宁

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“相关词”的 相关文章](#)

▶ 本文作者相关文章

• [章成志](#)

• [苏兰芳](#)

• [苏新宁](#)

•