

数据库、信息处理

基于语义相似度的自动文摘评价方法

张培颖

中国石油大学(华东) 计算机与通信工程学院, 山东 东营 257061

收稿日期 2008-5-14 修回日期 2008-7-28 网络版发布日期 2009-9-8 接受日期

摘要 针对现有自动文摘内部评价方法存在主观性强, 无法体现深层语义的缺陷, 提出利用概率潜在语义分析计算原文和待评价摘要在潜在语义上的贡献, 通过相似度的比较, 得出符合原文语义的文摘。整个评价方法无需人为干预, 实验表明, 基于语义相似度的评价方法实现简单、效果稳定, 是一种更接近自然模型的评价方法。

关键词 [自动文摘](#) [评价方法](#) [概率潜在语义分析](#) [相似度](#)

分类号 [TP391.1](#)

Evaluation method of automatic summarization based on semantic similarity

ZHANG Pei-ying

College of Computer & Communication Engineering, University of Petroleum (East China), Dongying, Shandong 257061, China

Abstract

Against to the shortcoming of strong subjectivity and the lack of deep-layer semantic information in existing internal evaluation of automatic summarization, by using probability latent semantic analysis to calculate the contribution on semantics including the original document and summaries, and then the similarity is compared. The summary which matches to semantic information of original document is got. Without human intervention, the results show that the evaluation method based on the semantic similarity has the feature of simplicity and stability and is closer to the natural model.

Key words [automatic summarization](#) [evaluation method](#) [probabilistic latent semantic analysis](#) [similarity](#)

DOI: 10.3778/j.issn.1002-8331.2009.25.044

通讯作者 张培颖 zhangpeiying@hpu.edu.cn

扩展功能

本文信息

- ▶ [Supporting info](#)
- ▶ [PDF\(622KB\)](#)
- ▶ [\[HTML全文\]\(0KB\)](#)
- ▶ [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](#)
- ▶ [复制索引](#)
- ▶ [Email Alert](#)
- ▶ [文章反馈](#)
- ▶ [浏览反馈信息](#)

相关信息

- ▶ 本刊中 [包含“自动文摘”的相关文章](#)
- ▶ 本文作者相关文章
- [张培颖](#)