

人工智能及识别技术

基于概率潜在语义分析的中文信息检索

罗 景, 涂新辉

(武汉科技大学计算机学院, 武汉 430065)

收稿日期 修回日期 网络版发布日期 2008-1-10 接受日期

**摘要** 传统的信息检索模型把词看作孤立的单元, 没有考虑自然语言中存在大量的同义词、多义词现象, 对召回率和准确率有不利的影响。概率潜在语义模型使用统计的方法建立“文档-潜在语义-词”之间概率分布关系并利用这种关系进行检索。该文将概率潜在语义模型用于中文信息检索, 实验结果表明, 概率潜在语义模型相对于传统的向量空间模型能够显著地提高检索的平均精度。

**关键词** [概率潜在语义分析](#) [中文信息检索](#) [索引技术](#) [关键词抽取](#)

**分类号** [TP18](#)

**DOI:**

通讯作者:

作者个人主页: [罗 景;涂新辉](#)

扩展功能

本文信息

- ▶ [Supporting info](#)
- ▶ [PDF\(214KB\)](#)
- ▶ [\[HTML全文\]\(0KB\)](#)
- ▶ [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“概率潜在语义分析”的 相关文章](#)
- ▶ [本文作者相关文章](#)
- ▶ [罗 景, 涂新辉](#)