

博士论文

一种基于语言概念空间聚类的信息检索方法

吴 晨<sup>1,2</sup>, 张 全<sup>2</sup>

(1. 中国科学院研究生院, 北京 100039; 2. 中国科学院声学研究所, 北京 100080)

收稿日期 修回日期 网络版发布日期 2007-4-6 接受日期

**摘要** 提出了一种以语言概念空间中的概念为聚类对象的信息检索方法以及适合于该方法的聚类算法。该聚类算法通过曲线拟合技术来实现文本的自动阈值确定和聚类划分, 并最终通过聚类间的迭代和结果修正来完成整个聚类过程。概念的引入为解决词语的同义、多义问题提供了有力保障。实验表明, 采用该方法的信息检索系统, 与 Jelinek-Mercer、k-means模型相比有较高的准确率和召回率, 效果理想。

**关键词** [信息检索](#) [语言概念空间](#) [聚类](#) [自动阈值下的聚类划分](#)

分类号

**DOI:**

通讯作者:

作者个人主页: 吴 晨<sup>1,2</sup>;张 全<sup>2</sup>

扩展功能

本文信息

▶ [Supporting info](#)

▶ [PDF](#)(115KB)

▶ [\[HTML全文\]](#)(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“信息检索”的 相关文章](#)

▶ 本文作者相关文章

· [吴 晨<sup>1,2</sup>, 张 全<sup>2</sup>](#)