

工作交流

### Web中文文本聚类研究及实现\*

杨学明<sup>1,2</sup>

宁波大学网络中心<sup>1</sup>

收稿日期 2006-9-14 修回日期 2006-9-18 网络版发布日期 2007-1-24 接受日期

**摘要** 采用Web文本自动聚类技术, 提出一种将HAC (Hierarchical Agglomerative Clustering) 聚类算法与K-Means聚类算法相结合的两阶段文本聚类框架, 并通过实验对该框架进行评估。

**Abstract** The text automatic clustering has been put forward and studied in application. This paper proposes a text clustering framework by coalescent the HAC and K-Means clustering algorithm, and evaluates this framework in an experiment.

**关键词** [文本自动聚类](#) [信息检索](#) [HAC](#) [K-Means](#)

**Key words** Text automatic cluster; Information retrieval; HAC; K-Means

**分类号** [TP18](#)

**DOI:**

通讯作者:

杨学明 [yangxueming@nbu.edu.cn](mailto:yangxueming@nbu.edu.cn)

作者个人主页: 杨学明

#### 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“文本自动聚类”的相关文章](#)

▶ 本文作者相关文章

• [杨学明](#)

•