

**doi:**10.1631/FITEE.1800010

**题目:** 基于同类用户的个性化查询词自动推荐方法

**概要:** 查询词自动推荐 (query auto-completion, QAC) 通过预测查询词前缀对应的完整补全查询词帮助用户构造查询词。大多互联网搜索引擎利用用户的行为信息为用户提供个性化的查询词自动推荐列表。为提高推荐成功率,个性化的QAC方法需获取大量关于用户搜索兴趣和搜索意图的上下文信息。因此,这些方法通常受制于用户数据的稀疏性问题。本文提出利用同类用户的搜索记录解决用户数据的稀疏性问题,并提升个性化QAC方法的推荐性能。首先,通过主题模型得到用户的主题兴趣,建立每个用户的兴趣肖像,然后将兴趣肖像相似的用户聚集起来建立同类用户群。由于传统主题模型不能自动识别同类用户,提出两个同类用户主题模型,将主题建模与同类用户识别包含在同一个模型框架内。根据不同的同类用户识别方法,提供4个不同的基于同类用户的个性化QAC方法。所提个性化QAC方法通过同类用户的上下文信息和查询词的频率对补全的查询词排序。在公开的AOL查询词数据集上进行大量实验,并与不采用同类用户上下文信息的方法进行排序性能对比。实验结果显示,本文提出的基于同类用户的个性化QAC方法能有效解决用户数据稀疏性问题,并且相对于基准方法能大幅提升排序结果准确性。

**关键词:** 查询词自动推荐; 基于同类用户的信息检索; 主题模型