

数据库、信号与信息处理

改进的遗传算法在Web使用挖掘中的应用

雷亮¹, 李善君², 彭军¹

1.重庆科技学院 电子信息工程学院, 重庆 401331

2.重庆市科协技术信息中心, 重庆 401147

收稿日期 2008-11-3 修回日期 2009-1-14 网络版发布日期 2009-3-10 接受日期

摘要 Web使用挖掘是近年来Web数据挖掘中的研究热点。针对传统遗传算法在提取关联规则问题时常采用固定染色体交叉概率和染色体变异概率, 容易出现早熟、收敛速度较慢的问题, 提出了改进的遗传算法, 并在关联规则的提取中增加了用户页面兴趣度这一阈值, 成功地运用到某商业网站服务器日志挖掘。实验证明, 这种改进的遗传算法能够有效避免早熟收敛现象, 是一种有效的方法。

关键词 [Web数据挖掘](#) [Web使用挖掘](#) [遗传算法](#) [兴趣度](#)

分类号

Application of Web usage mining based on improved genetic algorithm

LEI Liang¹, LI Shan-jun², PENG Jun¹

1.School of Electronic Information Engineering, Chongqing University of Science and Technology, Chongqing 400050, China

2.Information Center of Chongqing Association for Science and Technology, Chongqing 401174, China

Abstract

Web usage mining is a hot research direction of Web data mining. In this paper, an improved genetic algorithm is proposed to overcome the shortage of early convergence and stagnation in the traditional genetic algorithm, which is based on unconvertible rate of crossover operator and mutation operator. Moreover, the user page interest measure threshold is introduced into the association rules mining. Lastly, the improved genetic algorithm is successfully applied to a commerce Web server log mining and the experiment results indicate that the proposed algorithm is an effective method to avoid early convergence.

Key words [Web data mining](#) [Web usage mining](#) [genetic algorithm](#) [interest measure](#)

DOI: 10.3778/j.issn.1002-8331.2009.08.041

通讯作者 雷亮 cqlei.l@163.com

扩展功能

本文信息

▶ [Supporting info](#)

▶ [PDF\(692KB\)](#)

▶ [\[HTML全文\]\(0KB\)](#)

▶ [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 [“Web数据挖掘”](#) 的 相关文章

▶ 本文作者相关文章

· [雷亮](#)

· [李善君](#)

· [彭军](#)