

网络资源与建设

基于XML的Web日志挖掘研究*

潘有能^{1,2};

浙江大学信息资源管理系¹

收稿日期 2006-2-16 修回日期 网络版发布日期 2006-6-13 接受日期

摘要 设计一个基于XML的Web日志挖掘体系结构, 简要介绍XGMML和LOGML, 并在此基础上讨论LOGML文档的生成方法及利用Apriori算法对日志文档进行频繁集、频繁序列和频繁子图挖掘。

Abstract This paper builds a framework of Web usage mining based on XML technology, introduces XGMML and LOGML briefly. Then, the author discusses the method to generate LOGML documents. At last, the Apriori algorithm was used to mine frequent sets, frequent sequences and frequent sub-graphic in Web usage documents.

关键词 [数据挖掘](#) [Web日志](#) [XML](#)

Key words Data mining; Web usage data; XML

分类号 [TP311](#)

DOI:

通讯作者:

潘有能 ynpan@zju.edu.cn

作者个人主页: 潘有能

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“数据挖掘”的 相关文章](#)
- ▶ 本文作者相关文章
 - [潘有能](#)
 -
 -