

基于时空关系和关联规则挖掘的上下文信息缺失插补研究

王玉祥 乔秀全 李晓峰 孟洛明*

北京邮电大学网络与交换技术国家重点实验室 北京 100876

An Imputation Technique for Missing Context Data Based on Spatial-temporal and Association Rule Mining

Wang Yu-xiang Qiao Xiu-quan Li Xiao-feng Meng Luo-ming*

State Key Laboratory of Networking and Switching Technology, Beijing University of Posts and Telecommunications, Beijing 100876, China

摘要

参考文献

相关文章

Download: PDF (256KB) [HTML](#) 1KB Export: BibTeX or EndNote (RIS) Supporting Info

摘要 上下文信息的缺失是上下文信息处理中不可避免的问题, 缺失数据插补方法也是数据挖掘中的研究热点。但是, 现有的缺失数据的插补方法不太适合上下文信息这一流数据形式, 没有充分利用各传感器采集数据之间的关联性, 而且在插补的过程中没有考虑传感器数据的时空关系。为了解决现存的缺失数据插补方法的缺陷和不足, 该文提出了基于时空关系和关联规则挖掘的上下文信息缺失插补方法(STARM), 对传感数据进行空间化和时间序列化, 并生成强关联规则对缺失数据进行插补。最后, 通过温度传感器采集数据验证了这一算法合理性和高效性。实验证明, 该算法在上下文信息缺失估计的准确性要高于简单线性回归算法(SLR)和EM算法等, 而且具有较小的时空开销, 能够保证实时应用的服务质量(QoS)。

关键词: 信息处理 时空关系 关联规则挖掘 上下文信息缺失插补 均方根误差

Abstract: The context data missing is an inevitable problem of context information processing mechanism, the imputation technique of missing data also is a research hotspot in data mining. However, the existing imputation technique of missing data is not suitable for the flow data form of context information that does not make full use of data relevance between every collecting sensor. Moreover, that does not take spatial-temporal relationship into account. In order to conquer the shortcomings and deficiencies of the existing imputation technique of missing data, this paper proposes an imputation technique for context data missing based on Spatial-Temporal and Association Rule Mining (STARM) to perform spatiality and time series analysis on sensor data, and generate strong association rules to interpolate missing data. Finally, the simulation experiment verifies the rationality and efficiency of STARM through temperature sensor data acquisition. Experiments show that the algorithm is of high accuracy for the imputation of context data missing, such as Simple Linear Regression (SLR) algorithm and the EM algorithm. In addition, which is smaller time and space overhead and can guarantee Quality of Service (QoS) of real-time applications.

Keywords: Information processing Spatial-temporal relation Association rule mining Imputation for missing context data Root Mean Square Error (RMSE)

Received 2010-01-12;

本文基金:

国家自然科学基金(60802034, 60672122), 高等学校博士学科点专项科研基金(20070013026)和北京市科技新星计划(2008B50)资助课题

通讯作者: 王玉祥 Email: quanlin@163169.net

引用本文:

王玉祥, 乔秀全, 李晓峰, 孟洛明. 基于时空关系和关联规则挖掘的上下文信息缺失插补研究[J] 电子与信息学报, 2010, V32(12): 2913-2918

Wang Yu-Xiang, Qiao Xiu-Quan, Li Xiao-Feng, Meng Luo-Ming. An Imputation Technique for Missing Context Data Based on Spatial-temporal and Association Rule Mining[J], 2010, V32(12): 2913-2918

链接本文:

<http://jeit.ie.ac.cn/CN/10.3724/SP.J.1146.2010.00035> 或 <http://jeit.ie.ac.cn/CN/Y2010/V32/I12/2913>

Service

- ▶ 把本文推荐给朋友
- ▶ 加入我的书架
- ▶ 加入引用管理器
- ▶ Email Alert
- ▶ RSS

作者相关文章

- ▶ 王玉祥
- ▶ 乔秀全
- ▶ 李晓峰
- ▶ 孟洛明