

无线传感器网络时空数据聚集查询算法

徐 艳, 孙 燕

(南京师范大学 数学与计算机科学学院, 江苏 南京 210097)

[摘要] 提出了基于属性簇的时空数据聚集查询算法. 数据聚集以簇为 WSNB-TREE 存取目标, 将无线传感器网络 (WSN) 所覆盖的区域划分为多个属性簇, 同质传感器的数据便于实时统计处理和分时段查询; 对于事先确定监测的区域, 以区域划分数据聚集簇, 将异质传感器的聚集数据进行汇集, 从而快速过滤获取需查询的“事件”, 进而实现对全网的时空数据聚集查询.

[关键词] 无线传感器网络, 时空, 数据聚集, WSNB-TREE

[中图分类号] TP 393 [文献标识码] A [文章编号] 1672-1292(2008)04-0039-04

Spatio-Temporal Data Aggregation Query Algorithm in Wireless Sensor Networks

Xu Yan, Sun Yan

(School of Mathematics and Computer Science, Nanjing Normal University, Nanjing 210097, China)

Abstract In this paper, based on attribute-cluster, a new spatio-temporal data aggregation query algorithm is proposed. Data aggregation uses clusters as WSNB-TREE's storage object. Areas which are covered by wireless sensor network are divided into attribute-clusters. The data of the homogeneity sensors are easy to do real-time statistics and segmental query. Areas which are confirmed in advance are divided into a few data aggregation clusters, then the data of the heterogeneity sensors are aggregated. Hence, less time is taken to filter out the events which are needed to be queried, and then spatio-temporal data aggregation queries are conducted in the whole network.

Key words wireless sensor network, spatio-temporal data aggregation, WSNB-TREE

时空数据聚集查询主要研究无线传感器网络在某区域某时间段上各类数据存储和访问机制. 核心是确定节点产生的数据在网络中的存放位置和数据组织形式, 包括如何计算和保存过去数据的汇总信息, 如何最有效利用多个数据粒度分类属性簇, 如何应用分治策略将数据分块合并, 如何提供符合“事件”的查询方式等. 除此之外, 无线传感器网络还要求在通信量、存储资源占用和存取延时等方面有所兼顾. 无线传感器数据分析和处理研究已经取得了初步成果. 以汇聚节点为中心的数据聚集, 文献 [1] 提出以数据为中心的存取方法, 重点研究数据传输路径的能量损耗最少; 以存取速度平衡下的数据聚集策略, 文献 [2] 提出将数据存取的位置放置到查询访问需要的地方, 自适应更改数据存取位置; 以多节点数据聚集的树结构, 文献 [3] 提出 SB-TREE 的树结构的插入、删除和查询等聚集数据算法. 文献 [4] 提出查询谓词相似性优化策略及文献 [5] 提出的在线相关性分析等. 这些方法很难直接满足实际的无线传感器网络时空数据查询需要. 本文提出基于属性簇的时空数据聚集查询算法, 数据聚集以簇为 WSNB-TREE 存取目标, 将无线传感器网络所覆盖的区域划分为多个属性簇, 同质传感器的数据便于实时统计处理和分时段查询; 而对于事先确定监测的区域, 再以区域划分数据聚集簇, 将异质传感器的聚集数据进行汇集, 可以快速过滤获取需查询的“事件”, 进而实现对全网的时空数据聚集查询.

1 问题描述

1.1 概念定义

定义 1 簇头节点 C_i : 负责接收簇内节点传送的数据帧, 并在此作数据聚集处理.

收稿日期: 2008-06-18

通讯联系人: 孙 燕, 副教授, 研究方向: 分布式网络 and 智能主体. E-mail: sunyan@njjnu.edu.cn

定义 2 传感器节点 N_{in} : 每个传感器节点具有不同的感知属性 A_i , 由于每种属性变化敏感度不同, 其传感器的数据采集速度 V_{Ai} 不同, 当传感器节点以滑动窗口大小 W_{Ai} 和匀速度 V_{Ai} 采集数据时, 节点就将一个窗口快照为数据帧发送到簇头节点.

1.2 属性数据聚集模型

如图 1 所示. 将无线传感器网络划分为多属性层的聚集平面, 以时间序列建立数据聚集 WSNB-TREE, 簇头再将聚集数据流传输到上层汇集. 它与复式数据流^[6]相比, 具有数据类型单一, 聚集容易, 处理简单, 快速响应等特点; 与单式数据流相比, 通过数据聚集, 消除了多点的冗余数据, 减少了存储资源, 降低了网络节点的发送能耗.

定义 3 属性聚集树 T_{A_i} : 用 WSNB-TREE 树结构, 以属性 A_i 的滑动窗口快照周期为时间区域 (T_{S-A_i}, T_{E-A_i}) , 插入每个节点的时段感知数值所得到的聚集树和子簇树.

1.3 区域时间序列聚集模型

如图 2 所示. 将无线传感器网络划分为一个个观测区域, 对观测区域设置捕捉“事件”的条件. 每个观测区域可分时处理或设陷阱处理, 将空间上问题转化为区域内时间序列异质数据聚集 WSNB-TREE 分区域监测对分布式网络管理“事件”具有较高的灵活性和扩展性, 这种方法与不等速率多维数据流相关性分析^[5]相比, 大大降低了计算的复杂性, 与分布式数据交换汇总后再计算“事件”相比, 减少了数据传输, 从而降低能耗.

定义 4 区域聚集树 T_{S_k} : 用 WSNB-TREE 合并算法, 在区域 S_k 中, 以多属性的滑动窗口快照周期的公倍数为时间区域 (T_{S-S_k}, T_{E-S_k}) , 插入每个节点的时段感知数值所得到的多属性区域聚集树.

时空数据聚集模型是上述两种模型的组合, 当观察者对某个时段的某个属性值 A_i 感兴趣时, 查询条件为 $\{N_{Tree-Root}, t, A_i\}$, 查询属性聚集树 T_{A_i} ; 当观察者要监测某个区域 S_k 的“事件”时, 查询条件为 $\{N_{Tree-Root}, t, \{V_{A_1}, V_{A_2}, \dots, V_{A_i} \dots\}\}$, 查询区域聚集树 T_{S_k} .

定义 5 WSNB-TREE

(1) WSNB-TREE 属性聚集树 T_{A_i} 结构: 它保存着在快照窗口 W_{A_i} 的倍数时间区域中同质属性簇的聚集值. 每个属性簇树 T_{A_i} 由根结点、内结点和叶结点 3 种结点组成, 根结点包含 k 个时间间隔连续的时间段 $N, I_b, N, I_2, \dots, N, I_b$. 每个时间段对应一个属性聚集值 N, V_{A_i} 和一个相应的子结点 N, C_i , 且 k 个时间段之和为时间区域 (T_{S-A_i}, T_{E-A_i}) 的倍数, 时间区域的大小由传感器的属性决定. 内结点类似于根结点, 叶结点是一种没有子结点的内结点.

(2) WSNB-TREE 区域聚集树 T_{S_k} 结构: 它保存着在区域 S_k 中多属性时间区域中各种属性聚集值. 与属性聚集树的不同之处在于其时间区域是各种属性快照窗口 W_{A_i} 的公倍数 (T_{S-comb}, T_{E-comb}) , 聚集值包括传感器节点的所有属性聚集值 $N, V_{A_1}, N, V_{A_2}, \dots, N, V_{A_i} \dots$.

2 算法描述

2.1 算法思想

(1) 将无线传感器网络根据传感器的属性划分为 j 个属性簇; (2) 基于 k-means 的聚类算法, 将空间区域划分为 n 个子区域, 分别为 $S_1, S_2 \dots S_n$; (3) 每个区域里构建 j 棵属性聚集树 $T_{A_i} (i = 1 \dots j)$; (4) 每个区域里构建一棵区域聚集树 $T_{S_i} (i = 1 \dots n)$; (5) 区域事件查询.

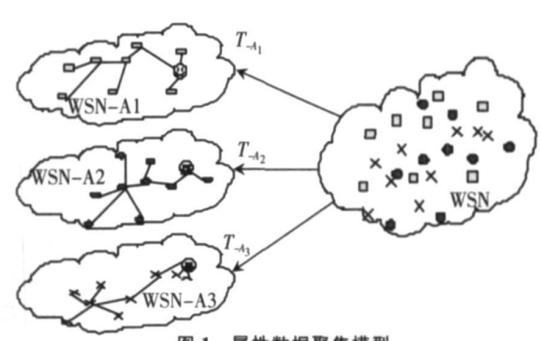


图 1 属性数据聚集模型
Fig.1 Attribute data aggregation model

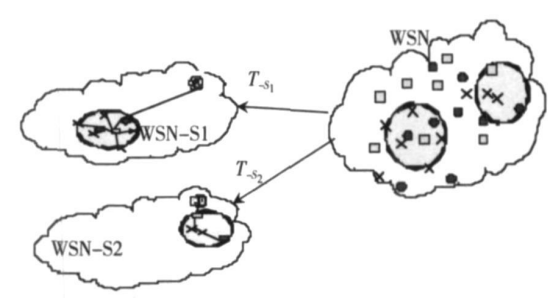


图 2 区域时间序列聚集模型
Fig.2 Area time sequence aggregation model

2.2 WSNB-TREE的生成和查询算法

WSNB-TREE的生成和查询是无线传感器网络时空数据聚集存取的两个方面, 下面对部分算法作一介绍.

算法 1 聚集树算法.

输入: Tree-Type聚集树类型; 插树节点集 $N; A_i$ 属性的时间区域 $[T_{S-A_i}, T_{E-A_i}]$.

输出: 生成符合条件的 WSNB-TREE聚集树.

```

1  if (Tree-Type = "属性  $A_i$  聚集树")
2       $T_S = T_{S-A_i}$ ;
3       $T_E = T_{E-A_i}$ ;
4      Insert ( $N, V_{A_i}, T_S, T_E$ );
5  else
6       $T_S = \text{最小公倍数} \{T_{S-A_1}, T_{S-A_2}, \dots, T_{S-A_j}\}$ ;
7       $T_E = \text{最小公倍数} \{T_{E-A_1}, T_{E-A_2}, \dots, T_{E-A_j}\}$ ;
8      Insert-comb ( $N, \{V_{A_1}, V_{A_2}, \dots, V_{A_j}\}, T_S, T_E$ );
9  end if

```

函数 Insert算法 Node Insert (N, V_{A_i}, T_S, T_E) 类似于 SB-TREE^[3] 中的 Insert算法.

当聚集数据以预先定义的方式保存在相关簇根节点中, 可以提供相应时段的查询或“事件”访问算法, 例如: 在 WSNB-TREE 的根节点中查询 t 时刻所有属性聚集值.

算法 2 时空事件的聚集查询 Query ($N_{\text{Tree-Root}}, t$).

输入: $N_{\text{Tree-Root}}$ 表示根结点, t 表示查询时刻.

输出: 时刻 t 的各种属性聚集值.

```

1  for ( $i = 1; i \leq \text{当前结点时间区域段的个数}; i++$ )
2      if ( $t \in N_{\text{Tree-Root}}.I_i$ )
3          if (当前结点不是叶结点)
4              result = Query ( $N_{\text{Tree-Root}}, C_i, t$ );
5          end if
6          for ( $m = 1; m \leq \text{当前结点的属性数}; m++$ )
7              result [ $m$ ] = accumulate (result [ $m$ ],  $N_{\text{Tree-Root}}.V_i[m]$ );
8          end for
9          break
10     end if
11 end for
12 return result

```

3 实验与分析

3.1 实验场景

本文采用 MATLAB 与 C 语言实现算法, 在 100×100 的 WSN 模拟场景下随机生成 100 个节点, 设有 4 种类型传感器属性, 每种属性的节点各 25 个, 数据采集的最小时间段 V_{A_i} 相同为 1 个元组 /ms, W_{A_i} 分别为 1s, 2s, 3s, 6s, 采集数据帧长都是 1000 元组.

3.2 实验分析

实验结果表明, 基于属性聚集与基于时空聚集模型在可观测相同数据量时, 基于属性聚集所花费的时间较少, 对捕捉单属性“事件”具有较高的实时性, 如图 3 所示; 而基于时空聚集模型由于聚集数效率较高, 在同样的时间内, 获得全部属性值的情况下, 时空聚集传输量比属性聚集的传输数据量少, 全网所消耗的能量少, 如图 4 所示.

4 结语

本文提出了基于属性簇的时空数据聚集查询算法, 利用 WSNB-TREE 的结构分而治之的聚集策略, 并

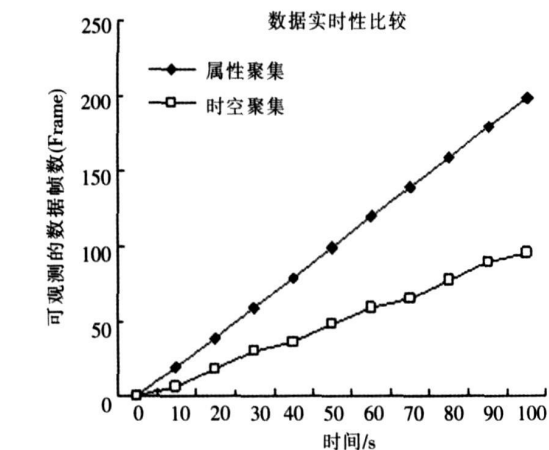


图 3 聚集时间-聚集量关系
Fig.3 Aggregation time-aggregation quantity

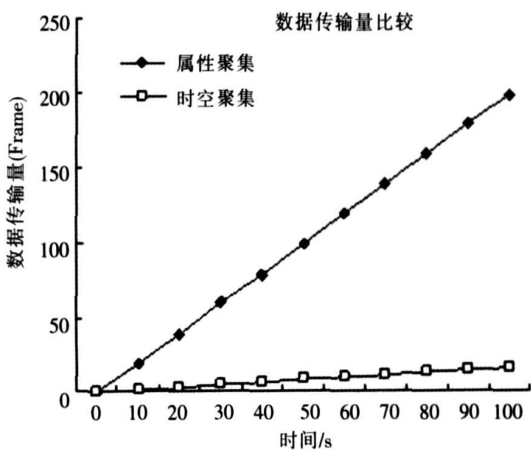


图 4 聚集时间-传输量关系
Fig.4 Aggregation time-transmission quantity

提出了属性数据聚集模型以及区域时间序列聚集模型, 通过这两个模型结合得到的时空数据聚集模型可以方便地对无线传感器网络实现时空数据聚集查询.

[参考文献] (References)

[1] 蔚赵春, 周水庚, 肖斌. 无线传感器网络中自适应数据存取 [J]. 软件学报, 2008, 19(1): 103-115.
Yu Zhaochun, Zhou Shuigeng, Xiao Bin. Adaptive information brokerage in wireless sensor networks [J]. Journal of Software, 2008, 19(1): 103-115. (in Chinese)

[2] 郭龙江, 李建中, 李贵林. 无线传感器网络环境下时-空查询处理方法 [J]. 软件学报, 2006, 17(4): 794-805.
Guo Longjiang, Li Jianzhong, Li Guilin. Spatio-temporal query processing method in wireless sensor networks [J]. Journal of Software, 2006, 17(4): 794-805. (in Chinese)

[3] Yang Jun, Widom Jennifer. Incremental computation and maintenance of temporal aggregates [C] // The 17th Int'l Conf Data Engineering. New York: Springer-Verlag, 2001.

[4] 魏雪云, 廖惜春. 基于卡尔曼的无线传感器网络时空融合研究 [J]. 传感器与微系统, 2007, 26(9): 72-75.
Wei Xueyun, Liao Xichun. Research on spatial-temporal fusion for WSNs based on Kalman filter [J]. Transducer and Microsystem Technologies, 2007, 26(9): 72-75. (in Chinese)

[5] 杨雪梅, 董逸生, 徐宏炳, 等. 高维数据流的在线相关性分析 [J]. 计算机研究与发展, 2006, 43(10): 1 744-1 750.
Yang Xuemei, Dong Yisheng, Xu Hongbing, et al. Online correlation analysis for multiple dimensions data streams [J]. Journal of Computer Research and Development, 2006, 43(10): 1 744-1 750. (in Chinese)

[6] 张冬冬, 李建中, 王伟平, 等. 分布式复合数据流处理 [J]. 计算机研究与发展, 2004, 41(10): 1 780-1 785.
Zhang Dongdong, Li Jianzhong, Wang Weiping, et al. Distributed compound-data streams processing [J]. Journal of Computer Research and Development, 2004, 41(10): 1 780-1 785. (in Chinese)

[责任编辑: 严海琳]