

无线传感器网络的数据存储和索引技术

毕学军, 杨朝红

(装甲兵工程学院信息工程系, 北京 100072)

摘要: 研究影响无线传感器网络数据存储和查询效率的应用指标, 如查询比率和采样频率。对无线传感器网络数据存储方法与各种应用指标的关系进行定量分析和比较, 给出各种方法的典型应用场景, 为无线传感器网络的数据存储和索引技术设计提供了参考。

关键词: 无线通信; 无线传感器网络; 数据存储

Data Storage and Index Technique for Wireless Sensor Network

BI Xue-jun, YANG Zhao-hong

(Department of Information Engineering, Academy of Armored Forces Engineering, Beijing 100072)

【Abstract】 This paper proposes the parameters that can influence the data storage and index of wireless sensor networks. It applies these parameters to analyze current storage methods of wireless sensor network, and gives theoretical quantitative analysis results. According to the analysis, the typical usage scenarios of each method are introduced, which can help the design of the data management of wireless sensor networks.

【Key words】 wireless communication; wireless sensor network; data storage

1 概述

无线传感器网络是目前备受关注、涉及多学科高度交叉、知识高度集成的前沿热点研究领域, 其数据管理具有如下特点: (1)传感器节点能源有限; (2)数据不可靠; (3)每个传感器节点产生连续的无限数据流; (4)查询技术具有新特点, 如时效性、地理性等。上述特点给无线传感器网络的数据存储和索引技术设计带来新挑战。文献[1]介绍传感器网络数据管理的概念和特点。康奈尔大学在感知数据查询处理技术方面进行了较多的研究工作, 研制了一个测试感知数据查询技术性能的 COUGAR 系统^[2]。加州大学伯克利分校研究传感器网络的数据查询技术, 并研制了感知数据库系统 TinyDB^[3]。现有无线传感器网络数据存储方法主要包括 3 种, 即外部存储方法、本地存储方法和以数据为中心的存储方法。在现实的传感器网络中采用哪种方法取决于应用特点, 目前没有相关的定量理论分析。

2 影响数据存储和查询效率的应用指标分析

先定义用户发起的查询频率 q , 每个节点感知数据的采样频率 w , 感知数据总的传输次数 C 。假设如图 1 所示, 节点均匀分布在某个区域, $sink$ 节点在区域中心, 为方便计算, 假设节点数目为 $n = (2k)^2, k \in N$, 节点 P_i 与 $sink$ 节点的距离为 $D(P_i, sink)$, 节点 P_i 的能量消耗为 $E(P_i)$, 由于传感器节点通信能量消耗与节点处理能量消耗通常不是一个数量级, 因此只考虑节点通信能量消耗。

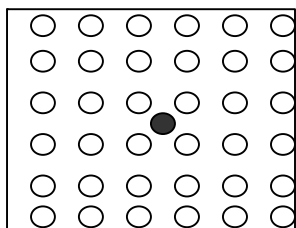


图 1 无线传感器节点分布

以通信次数为基本度量单位, C 是所有节点的能量消耗总和。为方便分析, 以一跳为基本距离单位, 不考虑分组输出错情况和控制信息传输消耗。

2.1 内部存储

对于内部存储情况, 如果每次用户查询对所有节点进行一次遍历, 即每个节点将数据传输给 $sink$ 节点, 则每次遍历的数据传输次数为

$$C/q = \sum_{m=1}^k ((2m)^2 - (2(m-1))^2)m$$

$$C = q \times 4(k^3 - \sum_{m=1}^{k-1} m^2)$$

可见, 感知数据总的传输次数 C 取决于查询次数 q , 与感知数据的采样频率 w 无关, 传感器节点的能量消耗分布如下:

$$E(P_i) = \frac{n - (2D)^2}{(2D)^2 - (2(D-1))^2} + 1 = \frac{n - (2D)^2}{8D - 4} + 1$$

其中, D 为节点 P_i 与 $sink$ 节点的距离。

当 $n=64$ 时, 考虑一次遍历情况, 内部存储方式的传感器节点能量消耗如图 2 所示。

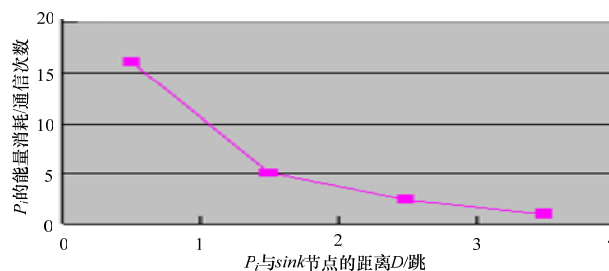


图 2 内部存储方式的传感器节点能量消耗

作者简介: 毕学军(1975—), 男, 讲师、硕士, 主研方向: 软件工程, 无线传感器网络; 杨朝红, 讲师、博士

收稿日期: 2008-09-28 **E-mail:** bxj168@sohu.com

由图 2 可以看出, 能量消耗随节点距离 sink 节点的距离减少而急剧增加, 因此, 很容易使 sink 节点旁的传感器节点能量快速消耗, 影响整个网络的可工作时间。

2.2 外部存储

对于外部存储情况, 每个传感器节点采样一次就将最新的采样结果发送到 sink 节点, 对整个网络的查询在 sink 节点上进行, 则有

$$C = w \times 4(k^3 - \sum_{m=1}^{k-1} m^2)$$

可见, 感知数据总的传输次数 C 取决于采样频率 w , 与查询的次数 q 无关, 外部存储的能量消耗情况与内部存储情况类似。

2.3 以数据为中心的存储

以地理散列方法^[4]为例进行研究。地理散列函数方法根据数据项的某个特定事件类域进行二维散列操作, 根据其结果由相应的地理路由协议将其传输并存储到对应地理位置的节点中。

同一事件类型的值域被散列到相同地理位置上, 因此, 对某个事件感兴趣的查询操作可以迅速利用地理路由协议定位到相应节点上。

如图 3 所示, 假设传感器网络对该地域的温度进行采集, 按采集温度范围在整个地域设置 4 个复制点, 则每次传感器节点采样后, 按相应温度范围将采样值发送给最近的复制点。假设每个传感器节点的温度采样都是采集范围内的一个随机分布, 由于复制点没有位于地域中心, 因此以数据为中心的存储相对于外部存储而言, 其 C 没有减少, 反而增大了。但由于 4 个复制点分担了 sink 节点部分负荷, 因此其能量消耗分布将得到较大改善。

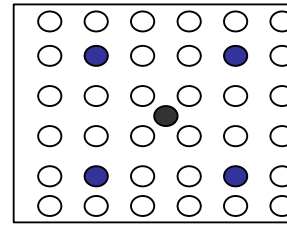


图 3 地理散列存储的节点分布

3 结束语

由上述分析可以看出, 如果感知数据的访问频率远高于产生这些数据的频率, 则外部存储方法是可用的。本地存储适用于感知数据产生频率远高于访问频率的情况。以数据为中心的存储方法使用数据名字存储和查询数据, 没有降低总体能量消耗, 但能较好地改善节点的能量消耗分布情况。

参考文献

- [1] 李建中, 李金宝, 石胜飞. 传感器网络及其数据管理的概念、问题与进展[J]. 软件学报, 2003, 14(10): 1717-1727.
- [2] Yao Yong, Gehrke J. The Cougar Approach to In-network Query Processing in Sensor Networks[J]. SIGMOD Record, 2002, 31(3): 9-18.
- [3] Madden S R, Franklin M J. Fjording the Stream, An Architecture for Queries over Streaming Sensor Data[C]//Proceedings of the ICDE Conference. Los Alamitos, USA: [s. n.], 2002.
- [4] Ratnasamy S, Karp B, Yin Li, et al. GHT: A Geographic Hash Table for Data-centric Storage in Sensornets[C]//Proc. of the 1st ACM International Workshop on Wireless Sensor Networks and Applications. Atlanta, Georgia, USA: ACM Press, 2002.

编辑 陈 晖

(上接第 80 页)

表 1 感知算法性能比较

信噪比/dB	多数准则	$\bar{D}(\Delta k = 3)$	$\bar{D}(\Delta k = 7)$
0	0.001 6	6e-4	3e-4
5	6.5e-4	3e-4	2e-4

从表 1 可以看出, 与基于多数判决准则的感知算法相比, 信噪比为 0 dB, Δk 为 3 时误判率改进了 2.67 倍, 相当于 4.26 dB, 而 Δk 为 7 时误判率改进了 5.33 倍, 相当于 7.27 dB; 信噪比为 5 dB, Δk 为 3 时误判率改进了 2.16 倍, 相当于 3.35 dB, 而 Δk 为 7 时误判率改进了 3.25 倍, 相当于 5.11 dB。表 1 还表明, 信噪比越低, 改善效果越明显。原因是, 信噪比越高, 误码越低, 误判的可能性也越小。

4 结束语

深空探测网络的异构特性和探测过程的长时间性, 要求深空探测网络具有感知通信协议和自重配置通信协议的能力, 以满足深空探测多任务目标的需求。

以 HDLC 协议和 IEEE802.3 协议为例, 本文研究了一种基于感知状态信息平均的链路层协议感知算法。这种算法考虑了深空通信信道的时变特性, 协议感知判决性能较基于多数判决准则的感知算法有了明显改进。

参考文献

- [1] Clare L, Clement B, Hutcgerson J, et al. Space-based Networking Technology Development in the Interplanetary Network Directorate Information Technology Program[C]//Proc. of the 2nd IEEE Conference on Space Mission Challenges for Information

Technology. Pasadena, CA, USA: [s. n.], 2006

- [2] Dobson S, Massacci F, Denazis S, et al. A Survey of Autonomic Communication[J]. ACM Trans. on Autonomous and Adaptive Systems, 2006, 1(2): 223-259.
- [3] Patouni E, Merentitis A, Polychronopoulos C, et al. Exploiting Protocol Reconfiguration for Autonomic Communication Environments[C]//Proc. of IST Mobile Summit'06. Mykonos, Greece: [s. n.], 2006.
- [4] Bhasin K, Hayden J. Developing Architectures and Technologies for an Evolvable NASA Space Communication Infrastructure[C]//Proc. of the 22nd AIAA International Communications Satellite Systems Conference and Exhibit. Monterey, CA, USA: [s. n.], 2004.
- [5] Okino C, Lee C, Gray A, et al. Protocol Sensing Across Multiple Space Missions[C]//Proc. of IEEE Aerospace Conference. Montana, USA: [s. n.], 2005.
- [6] Zhao Na, Meng Xin, Hu Shengbo. A Modification to GLRT-type SSME Data Rate Estimation Algorithm for Deep Space Communication[C]//Proc. of the 2nd International Conference on Space Information Technology. Wuhan, China: [s. n.], 2007.
- [7] Malakooti B, Kim H, Bhasin K. Human & Robotics Technology Space Exploration Communication Scenarios: Characteristics, Challenges & Scenarios for Developing Intelligent Internet Protocol[C]//Proc. of the 2nd IEEE Conference on Space Mission Challenges for Information Technology. Pasadena, CA, USA: [s. n.], 2006.

编辑 顾逸斐

