

移动多 Sink 传感器网络数据查询和收集技术研究现状

石高涛 赵增华

(天津大学计算机科学与技术学院 天津 300072)

摘要 在 MMS(Mobile Multi-Sink)型传感器网络中,多个移动的 Sink 节点会随时随地查询和收集数据,因此如何有效解决数据的查询和收集是 MMS 型传感器网络最基本和具有挑战性的工作。首先根据无线传感器网络的结构将其类型进行划分,之后给出数据查询和收集协议的评价原则,在此基础上综述了当前的具有典型特点的数据查询和收集协议,并按照已给的原则进行评价。最后给出若干需要解决的问题,为 MMS 型传感器网络的协议设计提供参考。

关键词 传感器网络,出版者/订阅者模型,数据查询,数据收集

Overview of Data Dissemination Protocols in Wireless Sensor Networks

SHI Gao-tao ZHAO Zeng-hua

(School of Computer Science and Technology, Tianjin University, Tianjin 300072, China)

Abstract For MMS (Mobile Multi-Sink) sensor networks, there exists many mobile sinks roaming in the area that will be likely to query the network at any time and at any place. Data Query and Collection from sources to Base station is the main functions in sensor networks. Many protocols have been proposed. The evaluation criterions of the performance for data query and collection protocols were presented. And then the existing protocols were reviewed and evaluated. Finally, the issues and the future trends in this field were pointed out.

Keywords Sensor networks, Publish/subscribe model, Data query, Data collection

最近,具有感知能力、通信能力和计算能力的无线传感器网络(wireless sensor networks, WSNs)开始在全世界范围内引起了人们的特别关注^[1],它是由具有传感、数据处理和短距离无线通信功能的传感器节点组成,能够协作地实时监测、感知和采集网络分布区域内的各种环境或监测对象的信息,可以使人们在任何时间、地点和任何环境条件下获取大量详实而可靠的信息,从而为用户与外界环境的交流提供了一个有效的桥梁。

无线传感器网络的一个最重要的任务和功能就是监测和采集环境信息,并将感知到的信息发布给用户。比如传感器网络可以监测野生动物栖息地,研究生物作息规律^[2];采集战场信息并向战场士兵或后方基地报告^[3,4],协助士兵制定更加可靠的作战方案。根据网络体系结构的不同,可以将传感器网络分成以下几种:(1)SSS(Static Single Sink)网络;(2)SMS(Static Multi-Sink)网络;(3)MSS(Mobile Single-Sink)网络;(4)MMS(Mobile Multi-Sink)网络。传感器网络的类型决定了数据收集方式。比如在 SSS 型网络中,传感器节点会主动将采集到的数据上传到 Sink;在 MSS 型网络中,移动的 Sink 采用 data mules 的方式充当数据收集单元,在移动中主动收集数据;在 MMS 型网络中,传感器节点一般将收集的数据暂时缓存下来,移动的 Sink 会随时随地发出查询,查找感兴趣的数据。可见,不同的应用类型和网络类型决定了数据收集的有效性。当基站静止或者个数很少时,当前的研究成

果相对比较成熟。但对 MMS 型的网络,问题更加复杂,当前尚未有一个理想的解决方案,在消息负载、可扩展性、负载平衡和实时性方面还存在诸多问题。为此,本文对 MMS 型网络中的数据收集技术进行评述,并指出存在的问题和可能的发展方向。

1 MMS 型网络数据收集及评价准则

MMS 型网络中的数据收集符合传统分布式系统中的 Publish&Subscribe 模型^[5],即 Sink 可以发出查询进行订阅(Pull),探测到数据的节点也可以根据订阅信息发布(Push)数据。由于 MMS 型网络中存在多个随机移动的 Sink 节点,因此传统的为 Sink 预先创建数据收集路径的方法会带来较大的负载;此外,这些 Sink 节点也会随时随地发出查询收集感兴趣的数据,传统的通过借助洪泛查询信息的策略也会带来较大的负载,特别是当 Sink 的数目较多或者查询频繁时,网络性能将进一步恶化。综合考虑各方面的性能指标,比如延迟、负载等,MMS 型网络一般采用 Pull 和 Push 相结合的数据收集策略。一般来说,一个 MMS 型网络数据收集协议的设计应当考虑以下因素:

(1) 能源消耗:能源是传感器网络中最宝贵的资源,主要体现在 3 个方面,即数据缓存代价 C_{cache} 、数据查询代价 C_{query} 和数据传输代价 C_{trans} 。其中数据缓存代价是数据 Push 所需代价;数据查询代价是指协议为转发查询所需要的代价。数

到稿日期:2008-09-23 本文受国家自然科学基金(No. 90604013 和 No. 60702038)资助。

石高涛(1978—),男,博士,讲师,主要研究领域为无线传感器网络、无线 Ad hoc 网络,E-mail: shgt@tju.edu.cn;赵增华(1974—),女,博士,副教授,主要研究领域为无线传感器网络、多媒体传输、无线 Mesh 网络。

据传输代价是指数据从缓存地点(包括探测地点)传输到基站所需代价。令 f 代表某一操作的频率,那么在一个较长的时间里,一个数据收集协议总的代价 C_{total} 可以表示如下:

$$C_{total} = f_{event} C_{cache} + f_{query} (C_{query} + C_{trans}) \quad (1)$$

(2) 查询成功率:在基于查询驱动的信息发布协议中,查询成功率是很重要的一个准则。查询成功率是指当被查询的事件信息已经存在的情况下,每次查询命中的概率。由式(1)可以看出,查询影响整个信息发布协议的代价,因此理想情况下,协议一方面要降低查询代价,另一方面要提高查询的成功率。

(3) 负载平衡:在分布式系统中应当尽量避免热点(Hotspot)问题,传感器网络同样需要避免某一区域负载过高而其它区域没有负载或负载偏低的现象。由于传感器网络的特点,要求所有节点都能够几乎在同一时刻死亡,从而能延长网络寿命。负载不平衡会导致某些节点过早死亡,从而降低网络的可用性。因此,信息发布协议应当考虑负载平衡问题。

(4) 时间延迟:时间延迟是衡量数据发布协议的另一个重要准则,它是提供给用户高服务质量中的重要因素。当查询发出或者事件发生之后,用户不应当等待过长时间。一方面传感器网络中的事件具有一定的时效性,另一面及时将数据传输给用户有利于网络资源的再利用。一般时间延迟需要考虑查询延迟 d_{query} 和数据传输延迟 d_{trans} 两部分。

(5) 可扩展能力:传感器网络数据发布协议性能应当随着网络规模的增大而具有一定的可扩展性。没有可扩展性的协议,当网络规模增大时性能会严重下降,这样的协议应用范围受限。未来的传感器网络势必是大规模的,因此具有可扩展能力的协议是必需的。

(6) 动态适应能力:传感器网络本质上是一个动态的网络,随着节点的死亡、新节点的加入以及节点位置上的改变,整个网络处于动态变化之中。特别是一些节能算法^[6,7]要求节点周期性轮流睡眠。因此,信息发布协议应当考虑网络的动态变化情况,使数据发布协议能够适应这种动态性。

这些性能指标不仅是数据发布协议评价标准,而且是其追求的目标。同时,这些指标之间也相互关联,一个性能指标的提高有时会带来另一个性能指标的提高,有时也会使另一个指标降低。因此,协议的设计要根据具体应用的要求进行权衡。

2 典型数据查询和收集协议

在无线传感器网络中,根据查询的种类可以分成基于 Pull、基于 Push 和混合型 3 种查询策略,其中基于 Pull 的策略是 Sink 节点发出查询,基于 Push 的策略是感知到数据的节点主动向 Sink 上报数据,将两者结合在一起就是混合式的查询。不同的策略适用于不同的应用类型中。在 MMS 型传感器网络中, Sink 随机游走,因此基于 Push 的方法无法有效工作。此外,由于 Sink 无法预先了解到数据的存储位置,基于 Pull 的查询策略是最直接的方法,但是考虑到 Sink 的数目和查询频率,这种策略会带来较大的负载,因此最常用的策略是混合式的查询。

本节介绍几种最具代表的数据查询和收集方法,并讨论其优缺点,为进一步的研究提供参考。思想接近的协议归并到同一类别中阐述。

2.1 基于 Pull 的数据查询和收集

这一类型的查询可以分成 3 种,分别是基于 flood 的查询、基于受限 flood 和随机转发。基于 flood 的查询方法将查询 flood 到整个网络中,满足要求的数据被收集上来,是最典型的数据收集方式。在 MMS 型网络中,由于数据和查询者都无法预先知道对方的位置,因此这种方法是最简单的。然而,当 Sink 数量较多或者查询量大时,网络内会承受较多查询负载,这类协议的性能就会下降。文献[8]对这一问题进行了详细的分析。

基于受限 flood 的方法每次发出的查询设定最大的转发次数,节点接收到查询后只有转发次数没有超过最大限制才继续转发。这种方法可能需要多次发送查询,每次逐渐增大最大转发次数,会带来一定的延迟。文献[18,19]对这类类型的查询进行了研究。

在随机转发查询,每个节点按照一定的概率来转发查询。这种方法类似于随机广播算法。研究结果表明,如果转发概率值 P 大于某一个常数 P_c ,那么每个节点都能以 1 概率收到查询,其中 P_c 依赖于网络大小等参数。

2.2 结构化的数据查询和收集

结构化的数据收集方式是借助于 Hash 表或者索引等机制来对数据进行映射,查询过程适用同样的机制。GHT^[9]是一个最典型的基于 Hash 表的数据查询和收集方法,它将数据按照关键字和地理位置一一对应。当 Sink 查询信息时,根据相同的映射规则得到特定类型事件的存储位置,然后使用基于地理位置的路由协议将查询转发到存储点上,从而获取事件信息。

ART^[10]将事件分为若干类,每类事件都有一个关键字,通过 Hash 函数将每类事件映射到网络中的一个或多个位置,所有相同类型的事件在该位置(称为索引中心)周围建立一个索引环,记录着每个事件的存储节点位置(发生事件的位置)。这样,对每个查询首先通过相同的 Hash 函数找到该类事件的索引中心位置,通过顺时针在索引环中转发查询,直到找到所需的事件对应的索引节点,然后通过该索引节点转发查询至存储节点上,进而将数据传输给查询者。其主要困难在于如何创建和维护索引以及如何转发查询。

2.3 基于副本存储的数据收集

基于副本存储的数据发布一般都是 Push-Pull 相结合的发布策略,它将感知到的数据以副本的形式存储在某些节点上或者某一固定区域中,查询时采用贪心转发的方式直接到存储区。

Artery^[11]将所有节点按照地理位置关系划分成网格,如图 1(a)所示,每个网格内选择一个 Leader(灰色节点),同时在网络内开辟一个连续的封闭区域作为缓存区,缓存区内的 Leader(黑色实心节点)缓存着网络内的事件信息。Artery 在初始化时为所有节点建立到达该缓存区的路径,查询消息直接转发到该区域内,并在缓存区中沿逆时针转发,直至查询到数据。

SEAD^[12]和 DP^[13]根据基站的分布建立一个树形结构的副本缓存结构,该树形结构使每个数据源的副本放置节点到基站的传输能耗最小化。SEAD 支持基站的移动,而 DP 协议不支持基站移动。这两个协议的基本思想都是根据事件的更新频率和基站请求数据的频率创建以数据源为根的 d-树,

如图 1(b) 所示, 并且通过选择副本节点来减少数据传输能耗。基站第一次发出请求时加入该 d-树, 每一个基站有一个入口节点(access node)来代表和跟踪基站的位置, 此后不断更新的数据沿着该 d-树发布给基站。

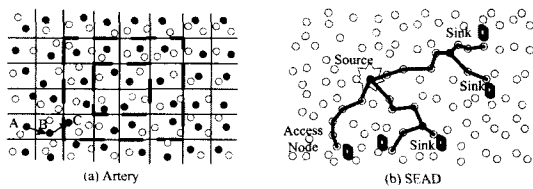


图 1 Artery 和 SEAD 协议示意图

2.4 基于路线控制的数据收集

基于路线控制的数据收集在转发查询和缓存数据时都严格按照特定的路线进行, 有效解决了随机转发和 flood 转发的缺点。这类数据收集方式中借用最多的思想是基于 quorum 交叉属性, 从而可以查找到感兴趣的信息。基于 quorum 的发布模式同样是结合 Pull-Push 策略的。

TTDD^[14] 的基本思想是每个数据源主动创建一个网络结构, 如图 2 所示, 整个网络被划分成若干 Cell, 数据源通过本身位于的 Cell 的 4 个交叉点将数据元信息发布到网格中的每个交叉点上, 这些节点在转发消息时记录下通向数据源的路径。当基站查询信息时, 直接在基站所处的 Cell 中 flood 查询消息, 然后由最近的一个交叉点将查询消息按照事先建立好的路由转发到数据源, 最终数据沿着相反的路径传输给基站。由于 TTDD 每个数据源都要创建和维护一个网格, 因此有些不会被查询的数据产生了不必要的浪费, 而且在事件发生频率比较高的网络中, 该协议性能值得怀疑, 并且网格中每个 Cell 的大小将影响数据发布的性能。

Liu^[15] 提出了一个梳子-针(Comb-needle)模型, 每个事件信息缓存成一个“针”, 而查找信息按照梳子形状进行转发, 这样梳子齿就有可能从草堆中梳理出事件信息。然而这种方法查询成功率低。

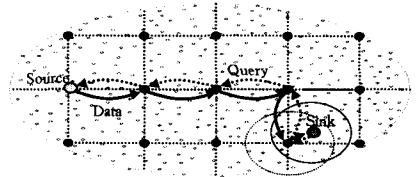


图 2 TTDD 协议示意图

GLIDER^[21] 协议基于 Landmark 的思想将网络区域划分成很多具有以 Landmark 为中心的 Voronoi 分片, 感知到的数据采用 GHT 映射到某一个 Landmark, 并在转发过程中缓存到经过的 Voronoi 边上。查询消息使用同样的 GHT 映射到 Landmark, 并在转发查询时察看经过的 Voronoi 边是否有满足条件的数据存在。这一方法需要具有全局网络信息。

3 协议比较和存在的问题

3.1 协议比较

为了更加清楚地察看每个协议性能, 我们在表 1 中参照第 1 节中的评价准则给出了不同协议的比较。根据评价准则将相关协议的性能分成了若干等级, 分别对应差、中、良和最优。

表 1 WSN 数据查询和收集协议比较

Protocol	Energy Cost	Success Rate	Load Balance	Scaleability	Time Delay	Dynamic Adaptability	Delivery Rate
TTDD ^[14]	Depends	Moderate	Moderate	Moderate	Low	Bad	Moderate
Artery ^[11]	High	High	Bad	Bad	Low	Bad	Bad
SEAD ^[12]	Moderate	Moderate	Bad	Moderate	Moderate	Bad	Good
Liu ^[15]	Low	Low	Good	Moderate	Low	Bad	Moderate
ARI ^[10]	Moderate	Moderate	Moderate	Moderate	Moderate	Bad	Moderate
DP ^[13]	High	Moderate	Bad	Bad	Low	Bad	Good
TP ^[16]	Low	High	Good	Moderate	Moderate	Moderate	Moderate
DIFS ^[17]	Low	High	Moderate	Moderate	Low	Bad	Moderate
Double Rulings ^[20]	Low	High	Good	Best	Low	Bad	Good
GLIDER ^[21]	Low	High	Good	Moderate	Low	Bad	Moderate

从表 1 中可以看出, 所有协议在动态适应能力方面都比较欠缺, 但在可扩展性方面大部分都有所考虑, 而且大部分协议的能耗相对都比较低。通过以上对各种模式的介绍可以发现, 几乎所有协议都依赖于节点的地理位置信息, 这是当前协议设计的一个最大缺点。众所周知, 节点的位置信息一方面难以获得, 另一方面受定位协议或设备的闲限制置, 所得位置信息不可能是精确的。因此, 设计一个无位置依赖的数据查询和收集协议更有意义。

3.2 亟待解决的问题

尽管目前已有众多数据收集协议出现, 但是以下问题仍然是亟待解决的, 它们也将成为未来的主要研究方向。

(1) 移动性支持: 越来越多的研究开始转向移动传感器网络上, 这不仅仅包括基站的移动, 而且涉及到网络内的节点移动。目前已经有移动的传感器节点出现。在传感器网络中引入移动节点能够提高网络覆盖和网络连通性, 但也给数据查询和收集带来了一定的困难。如何利用节点的移动性, 将移动性的劣势转化为优势来协助数据的收集和发布, 是一个有挑战性并且有意义的工作。

(2) 动态性影响: 动态性包括移动性, 但并不限于移动性。动态性是指节点状态的变化, 包括节点位置移动、节点睡眠/唤醒状态转换等。由于节能算法的影响, 节点在工作 and 休眠两种状态的转换更为普遍。目前的数据查询和收集模式较少考虑节点状态变化时对协议的影响, 这使得一些数据发布模式不适用于在这种情况下工作。比如, 以数据为中心的存储和基于索引的存储, 这些方式需要节点状态稳定。如何在节点状态不断变化下有效地完成数据发布, 是另一个需要解决的问题。

(3) 异构网络支持: 传感器网络是人类和自然界交流的桥梁, 未来的传感器网络中将越来越多地出现各种无线或有线设备, 比如手机、手提电脑等。这样, 整个系统的异构性将大大增强。这种现象既给数据的查询和收集带来挑战, 也为其提供了机会。

(4) 数据发布的安全性: 安全性在传感器网络中是一个很重要的问题。尽管与安全有关的工作已经开展并取得一定的结果, 但是在数据发布中考虑安全目前还没有出现。事实上, 数据查询和收集更容易引入攻击, 从而使网络受到破坏。比如敌人不停地发送查询, 使节点疲于应答, 这样可以很快消耗掉这些节点的能源。安全应当在设计数据查询和收集协议的开始就考虑进去, 而不是之后再单独解决。

结束语 无线传感器网络提供了一个全新的分布式信息处理平台和计算模式, 将自然世界和人有机的地结合起来。本文总结了 MMS 型网络中现有的数据查询和收集协议, 给出了其性能评价指标, 并对现有的协议进行了分类, 详细介绍其中典型的模式及其协议。通过对现有的协议和算法的综述可以看到, 传统分布式系统中的思想仍然影响着 MMS 型传感器网络查询和收集的设计。比如传统分布式系统中已有数据的 Pull 和 Push 相结合的策略; 以数据为中心的存储则借鉴了 P2P 系统的文件管理; 基于索引的发布模式则能在文件管理系统和数据库系统中找到它的影子。这为将来设计新的发布模式提供了参考。

参 考 文 献

- [1] Akyildiz I F, Su W, Sankarasubramaniam Y, et al. A survey on sensor networks[J]. IEEE Communications Magazine, 2002, 40 (8): 102-114
- [2] Mainwaring A, Polastre J, Szewczyk R, et al. Wireless sensor networks for habitat monitoring[C]//Proceedings of the ACM International Workshop on Wireless Sensor Networks and Applications, 2002
- [3] Gui C, Mohapatra P. Power conservation and quality of surveillance in target tracking sensornetworks[C]//International Conference on Mobile Computing and Networking (ACM MobiCom). Philadelphia, PA, 2004: 129-143
- [4] He Tian, Krishnamurthy S, Stankovic J A, et al. Energy-efficient Surveillance System Using Wireless Sensor Networks[C]//2nd Int. Conf. on Mobile Systems, Applications, and Services (MobiSys04). Boston, MA, June 2004
- [5] Caporuscio M, Carzaniga A, Wolf A L. Design and evaluation of a support service for mobile, wireless publish/subscribe applications[J]. IEEE Trans. on Software Engineering, 2003, 29 (12): 1059-1071
- [6] Ye F, Zhong G, Lu S, et al. PEAS: A Robust Energy Conserving

- Protocol for Long-lived Sensor Networks[C]//23rd International Conference on Distributed Computing Systems, 2003: 28-37
- [7] Kumar S, Lai T H, Balogh J. On k-coverage in a mostly sleeping sensor network[C]//Proc. of the 10th Annual International Conference on Mobile Computing and Networking, 2004: 144-158
- [8] Heidemann J, Silva F, Estrin D. Matching Data Dissemination Algorithms to Application Requirements[R]. ISI-TR-571. USC/Information Sciences Institute, April 2003
- [9] Shenker S, Ratnasamy S, Karp B, et al. Data-Centric Storage in Sensornets[C]//the First ACM SIGCOMM Workshop on Hot Topics in Networks (HotNets 2002). Princeton, NJ, October 2002
- [10] Zhang W, Cao G, La Porta T. Data Dissemination with Ring-based Index for Sensor Networks[C]//IEEE International Conference on Network Protocol (ICNP). November 2003
- [11] Lin Lan, Wu Hailin. Artery: A Data-Centric Architecture for Wireless Sensor Networks[C]//CCCT '03. Orlando, USA, 2003
- [12] Kim Hyung Seok, Abdelzaher T F, Kwon Wook Hyun. Minimum-energy asynchronous dissemination to mobile sinks in wireless sensor networks[C]//Proceedings of the 1st International Conference on Embedded Networked Sensor Systems. Los Angeles, California, USA, 2003: 193-204
- [13] Bhattacharya S, Kim H, Prabh S, et al. Energy-conserving Data Placement and Asynchronous Multicast in Wireless Sensor Networks[C]//Proceedings of The First International Conference on Mobile Systems, Applications, and Services (MobiSys). San Francisco, CA, May 2003
- [14] Ye Fan, Luo Fan, Cheng J, et al. A Two-tier Data Dissemination Model for Large-scale Wireless Sensor Networks[C]//Proceedings of the Eighth International ACM Conference on Mobile Computing and Networking (MOBICOM). 2002
- [15] Liu X, Huang Q, Zhang Y. Combs, Needles, and Haystacks: Balancing, Push and Pull in Large Scale Sensor Networks[C]//the Second ACM Conference on Embedded Networked Sensor Systems (SenSys). 2004
- [16] 石高涛, 廖明宏, 一种大规模传感器网络节能数据发布协议[J]. Journal of Software, 2006, 17(8): 1785-1795
- [17] Greenstein B, Estrin D, Govindan R, et al. DIFS: A Distributed Index for Features in Sensor Networks[C]//First IEEE International Workshop on Sensor Network Protocols and Applications, May 2003
- [18] Chang N B, Liu M. Revisiting the ttl-based controlled flooding search: Optimality and randomization[C]//Proceedings of the Tenth Annual International Conference on Mobile Computing and Networks (ACM MobiCom). 2004
- [19] Cheng Z, Heinzelman W. Flooding Strategy for Target Discovery in Wireless Networks. ACM/Baltzer Wireless Networks, 2005
- [20] Sarkar R, Zhu Xianjin, Gao Jie. Double Rulings for Information Brokerage in Sensor Networks[C]//The 12th Annual International Conference on Mobile Computing and Networking (MobiCom'06). September 2006: 286-297
- [21] Fang Qing, Gao Jie, Guibas L J. Landmark-based Information Storage and Retrieval in Sensor Networks[C]//The 25th Conference of the IEEE Communication Society (INFOCOM'06). April 2006