

数据融合在传感器网络协议中的节能性分析

王娟¹,王汝传^{1,2},孙力娟¹

(1. 南京邮电大学 计算机科学与技术系, 江苏 南京 210003;

2. 南京大学 计算机软件新技术国家重点实验室, 江苏 南京 210093)

摘要:传感器网络受到带宽、能量和吞吐量因素的限制,制约了点对点信息传送的数量。网内处理冗余数据或信息、组合更为高效,符合用户需求的数据融合技术很好地解决了这些问题。阐述数据融合在传感器网络中的作用,针对现有数据融合在网络协议中提出的各种算法思想,分析比较了数据融合的节能性能。由此可知,基于生成树的融合方法更能平衡节点负载,延长网络的生命周期,从而解决无线传感器网络存在的问题。

关键词:无线传感器网络;数据融合;节能

中图分类号:TP393;TP31

文献标识码:A

文章编号:1673-629X(2006)11-0004-03

Energy - Saving Analysis of Data Aggregation in Sensor Network Protocol

WANG Juan¹, WANG Ru-chuan^{1,2}, SUN Li-juan¹

(1. Dept. of Computer, Nanjing University of Posts and Telecommunications, Nanjing 210003, China;

2. State Key Laboratory for Novel Software Technology, Nanjing University, Nanjing 210093, China)

Abstract: Sensor networks suffer bandwidth, energy, and throughput constraints that limit the quantity of information transferred from end-to-end. Data aggregation resolved such problems, dealing with redundant source data to get more efficient and useful information. According to existing data aggregation algorithms in network layer, this paper expatiated on the function of data aggregation technology, analysed and compared energy-saving performance of data fusion in sensor network. Data gathering based on spanning tree balance the load of nodes, prolong the lifetime of network, solve the limitation in wireless sensor network.

Key words: wireless sensor networks; data aggregation; energy saving

0 引言

随着通信技术、嵌入式技术和传感器技术的飞速发展和日益成熟,具有感知能力、计算能力和通信能力的微型传感器开始在世界范围内出现。由这些微型传感器构成的传感器网络引起人们的极大关注。然而,网络受到带宽、能量和吞吐量等因素的限制。如何高效使用能量来最大化网络生命周期、如何设计网络通信机制满足传感器网络的通信需求、如何利用有限的计算和存储资源完成诸多

协同任务成为传感器网络设计的挑战。数据融合技术使得在节点收集数据的过程中,利用节点本地计算和存储能力处理数据,去除冗余的信息,尽量减少传输量,达到节能目的。文中对无线传感器网络定义、节点局限性进行概括,对数据融合技术的概念理解、作用和分类加以总结,针对传感器网络协议中数据融合技术的应用,分析比较了几种协议的节能性能。

1 数据融合

传感器网络是由一组传感器以 Ad Hoc 方式构成的有线或无线网络,其目的是协作地感知、采集和处理网络覆盖的地理区域中感知对象的信息,从而发布给观察者^[1]。数据融合是将采集的多份数据或信息进行处理,组合出更有效、更符合用户需求的数据的过程,它利用计算机技术对按时序获得的若干传感器的观测信息在一定准则下加以自动分析、综合以完成所需的决策和估计任务而进行的信息处理过程。

在传感器网络中,数据融合起着十分重要的作用,主要表现在节省整个网络的能量、增强所收集数据的准确性

收稿日期:2006-02-06

基金项目:国家自然科学基金(60573141, 70271050);江苏省自然科学基金(BK2005146);江苏省高技术研究计划(BG2004004, BG2005038, BG2006001);国家高科技“八六三”项目(2005AA775050);南京市高科技项目(2006软资105);江苏省计算机信息处理技术重点实验室基金(kjs050001, kjs06);江苏省高校自然科学基金研究计划(04KJB520095)

作者简介:王娟(1982-),女,河南开封人,硕士研究生,研究方向为计算机软件、无线传感器网络技术;王汝传,教授,博士生导师,研究方向为计算机软件、计算机网络和网络、信息安全、无线传感器网络、移动代理和虚拟现实技术等。

以及提高收集数据的效率三个方面。

(1) 节省能量。

传感器网络是由大量的传感器节点覆盖到检测区域而组成的,鉴于单个传感器节点的监测范围和可靠性是有限的,在部署网络时,需要使传感器节点到达一定的密度以增强整个网络的鲁棒性和监测信息的准确性,有时甚至需要使多个节点的监测范围互相交叠,这种监测区域的相互重叠导致临近节点报告的信息存在一定程度的冗余。

数据融合针对上述情况对冗余数据进行网内处理,即中间节点在转发传感器数据之前,首先对数据进行综合,去掉冗余信息,在满足应用需求的前提下将需要传输的数据量最小化。网内处理利用的是节点的计算资源和存储资源,其能量消耗与传送数据相比要小得多。

(2) 获得更准确的信息。

仅收集少数几个分散的传感器节点的数据较难确保得到信息的正确性,需要通过对监测同一对象的多个传感器所采集的数据进行综合,来有效地提高所获得信息的精度和可信度。另外,由于临近的传感器节点监测同一区域,其获得的信息之间差异性很小,如果个别节点报告了错误或误差较大的信息,很容易在本地处理中通过简单的比较算法进行排除。

(3) 提高数据收集效率。

在网内进行数据融合,可以在一定程度上提高网络收集数据的整体效率。数据融合减少了需要传输的数据量,可以减轻网络的传输拥塞,降低数据的传输时延;即使有效数据量并未减少,但通过对多个数据分组进行合并减少了数据分组个数,可以减少传输中的冲突碰撞现象,也能提高无线信道的利用率。

传感器网络中的数据融合技术可以从不同的角度进行分类^[2]。根据数据进行融合操作前后的信息含量,可以将数据融合分为无损失融合和有损失融合两类;根据数据融合是否基于应用数据的语义,将数据融合技术分为三类:依赖于应用的数据融合(ADDA)、独立于应用的数据融合(AIDA)^[3],以及结合以上两种技术的数据融合;根据对传感器操作的级别,可将数据融合技术分为数据级融合、特征级融合和决策级融合三类。

虽然对信息融合理论及其应用的研究正在以前所未有的广度和深度进行着,但目前在国际上还没有形成有关信息融合的一套完整的理论框架和融合算法。利用多传感器数据确定目标的位置、速度、属性和辨识是信息融合中最成熟的领域,其中确定目标的位置和速度是一个相对陈旧的问题。Gauss 和 Legendre 提出了确定卫星轨道问题的最小方差方法。利用有限的资源如计算能力、有限的通信带宽和其他限制进行合并或建立模型以得到最优化的性能是非常困难的^[4]。

无线传感器网络有别于多传感器的数据融合,网络的拓扑结构因环境或电能耗尽造成的传感器节点出现故障、新节点的加入等而改变,具有自组织的能力,能够根据自

动进行配置和管理,通过拓扑控制机制和网络协议自动形成转发监测数据的多跳无线网络系统的特点,对网络内部的数据进行网内处理,减少通信的数据量,从而延长了节点的寿命。通过数据融合平衡网络内部各节点的能量消耗,提高网络的生命周期,由于其独特的特点,原来对数据融合的方法不适合于现有的无线传感器网络。针对该情况,提出了很多关于与网络协议层有关的数据融合技术。

2 传感器网络层协议中的数据融合

定向扩散^[5]路由中的数据融合包括路径建立阶段的任务(又称兴趣, interest)融合和数据发送阶段的数据融合,这两种融合都通过缓存机制得以实现。定向扩散中的兴趣融合得益于它基于属性的命名方式,类型相同,监测区域完全覆盖的兴趣在某种情况下就可以融合成为一个兴趣。定向扩散路由的数据融合采用的是“抑制副本”的方法,即对转发过的数据进行缓存,发现重复的数据将不予以转发。

LEACH^[6]与 TEEN^[7]都是基于层次的路由,它们使用分簇的方法使得数据融合的地位突显出来。每个簇头在收到本簇成员的数据后进行融合处理,并将结果发送给 sink 节点。

PEGASIS^[8]及其高阶算法对 LEACH 的数据融合方式进行了改进。它基于两个假设:一是所有节点距离 sink 节点都很远;二是每个节点都能将接受到的数据分组与自己的数据融合成一个大小不变的分组。PEGASIS 算法在收集数据前,首先利用贪心算法将网络中的所有节点链接成一条单链,然后随机选取一个节点作为首领。首领向链的两端发出收集数据的请求,数据从单链的两个端点向首领流动,中间节点在传递数据前要执行融合操作,最终由首领节点将结果数据传送给 sink 节点。

3 无线传感器网络中的节能数据融合

传感器节点最大的局限性在于其能量的有限性,从而限制了传感器网络的生命周期和质量。基于这个原因,运行在传感器网络中的协议为了延长网络寿命,必须采取措施有效利用节点能量。当考虑到通信能量有限时,有效地传送数据,协作各节点工作,延长节点寿命,减少带宽要求,变得至关重要。

传感器网络中节点产生大量的数据需要最终用户处理,有必要研究将数据综合的方法。最简单的操作是合计(求和,求平均,最大值,最小值,计数)来自不同节点的数据。更为准确的解决方案是数据融合。数据融合是指通过增强公共信号,减少非关联的噪声信号,对不可靠数据进行聚合产生更精确数据。这些方法可以通过不同的协议来实现,从而减少了系统中数据传输的数量,改善了传感器网络的性能。

3.1 通信模型

一个传感器节点传送接收数据需要消耗的能量表示

为 $E_{\text{elec}} = 500\text{nJ/bit}$, 经过信号放大电路消耗的能量为 $E_{\text{amp}} = 100\text{pJ/bit/m}^2$, 向距离 d 的节点发送 kbit 信息需要消耗能量为

$$\text{ETX}(k, d) = E_{\text{elec}} * k + E_{\text{amp}} * d^2 * k \quad (1)$$

接收这个信息需要消耗的能量表示:

$$\text{ERX}(k) = E_{\text{elec}} * k \quad (2)$$

同样, 假定通信信道是对称的, 意味着从节点 A 到节点 B 传送信息消耗的能量与节点 B 向 A 相等。

3.2 问题描述

传感器节点随机散布在感兴趣的区域, 假定每个传感器节点位置固定, 且 sink 节点预先知道它们的位置。传感器节点可以和自己通信半径内的节点直接通信, 同样假定它们都可以和 sink 节点直接通信。传感器节点周期性地感知环境并在每轮(round)通信中发送数据, 可以将来自其他节点的数据和自身数据进行融合, 将数据包融合为一个数据包。

问题是寻找合适路径将节点收集的数据传送到 sink 节点, 延长基于上述假定的整个网络的寿命。然而, 寿命的定义由于网络提供服务的不同而存在差异。若每个传感器的协同性强, 即当第一个传感器节点能量耗尽时, 系统的性能显著下降, 此时, 该系统的寿命是指第一个传感器能量耗尽之前系统执行的轮数。另一种情况, 网络中节点部署密集, 邻近节点记录着同样或相关的数据, 只有当网络中大量节点死亡(能量耗尽)才影响到系统的质量。在这种情况下, 网络的寿命定义为半数节点或网络中特殊节点死亡所需要的时间。一般来说, 当网络中最后一个节点耗尽自身能量系统所执行的轮数称为网络的寿命。

相对于对称信道, 发送 kbit 数据对系统而言能量消耗公式为

$$C_{ij}(k) = 2 * E_{\text{elec}} * k + E_{\text{amp}} * d_{ij}^2 * k \quad (3)$$

$$C_i'(k) = E_{\text{elec}} * k + E_{\text{amp}} * d_{ib}^2 * k \quad (4)$$

最简单的方法是将每个传感器节点收集的数据直接传输给基站, 但是如果基站与节点距离远, 传输代价将很大, 节点会很快耗尽能量死亡。

数据通过中间节点进行发送, 中间节点选取发送放大电路能量消耗小的。

(1) 每个信息经过 n 次传送接收到达 sink 节点。系统消耗的总能量相对于直接通信协议小。

(2) 中间节点(多为距离 sink 节点近的节点)将用来传送大量数据而很快死亡, 因此无法对它们周围的环境进行监测。

利用接收和传送电路消耗的能量大于利用放大器电路消耗的能量, 提出一种链的方案, 该链包含所有节点, 相邻节点距离尽可能最小, 每个传感器仅发送接收邻节点的数据。它将与 sink 节点通信的传感器节点减少到一个。每一轮中, 随机选取一个节点将数据进行融合发送给 sink 节点。当 sink 节点距离每个节点都很远时, 每个传感器节点在这种情形下发送数据消耗的能量几乎相同, 该算法

很有效。无论每轮中哪个节点向 sink 节点发送数据, 算法都尽力减少通信节点消耗的能量, 从而延长节点寿命。

在 PEDAP-PA^[9] 协议中考虑到每个传感器节点的剩余能量, 平衡网络中节点能量消耗。当 sink 节点距离每个节点都很远时, 最先耗尽能量的是那些将融合数据传送给 sink 的节点, 剩余能量低的节点不再向 sink 传送数据, 利用它们转发邻节点的数据, 增大系统的寿命。同样, 这些节点不再接收其他节点发来的数据(接收大量数据耗能很大), 尽量只转发数据。对式(3)、(4)进行调整得到:

$$C_{ij}(k) = (2 * E_{\text{elec}} * k + E_{\text{amp}} * d_{ij}^2 * k) / e_i \quad (5)$$

$$C_i'(k) = (E_{\text{elec}} * k + E_{\text{amp}} * d_{ib}^2 * k) / e_i \quad (6)$$

其中, e_i 为节点 i 的剩余能量, 可以看出, 节点间的通信并不对称, 节点发送数据消耗的能量随剩余能量的减少而增大。

3.3 分析与比较

(1) 直接通信协议: 当处于 sink 节点相对于每个节点距离远时, 或者接收数据消耗的能量大时, 该方法适于应用。

(2) MTE 协议: 通过利用中间节点, 经过 n 次传送与接收, 系统总消耗量比直接通信小。

(3) LEACH: 基于层次的路由, 通过簇首节点与 sink 通信, 其余节点选择距离自己近的簇首节点发送自己的数据。为了延长这个系统的寿命, 簇首节点随机方式选择。然而, 因为算法为纯随机性, 不够理想。

(4) PEGASIS: 通过将网络中的全部节点构造成一条链, 若链上某一节点失败则导致数据全部丢失, 由于链长和节点数量有关, 在数据收集时会产生很大的延迟。而且二者都没有考虑到系统中节点的平衡负载问题和减少网络的总体耗能。

(5) PEDAP 协议: 尽可能延长最后一个节点的生命周期, 同样也没有考虑节点平衡负载的问题。在 sink 节点中经过节点剩余能量的计算, 通过最小生成树, 用最新的路由信息配置, 平衡网络中节点的负载, 延长了系统的寿命。基于融合树的数据融合技术更能平衡节点负载, 延长无线传感器网络的生命期, 有效解决网络存在的问题。

4 结束语

数据融合技术与网络层紧密结合, 尤其是路由技术的结合, 起到了对无线传感器网络带来的节能作用, 克服了节点存在能量约束的局限性。

数据融合技术使得节点收集数据的过程中, 利用节点本地计算和存储能力处理数据, 去除冗余的信息, 尽量减少传输量, 达到节能目的。该技术对多份数据进行综合, 提高了信息的准确度。但是, 它也带来了由于寻找易于进行数据融合的路由、进行融合操作、为融合等待数据到来而引起网络的平均时延, 这也是今后研究的一个重要方面。

(下转第 83 页)

3.2 模块间的关系

图 5 显示了论坛系统的整体的模块关系。整个论坛的主要模块根据帖子、板块、用户这 3 个实体联系为 3 条主线,并且,由于板块和帖子具有父子关系,所以浏览帖子模块、发帖回帖模块与板块模块之间也具有关系。

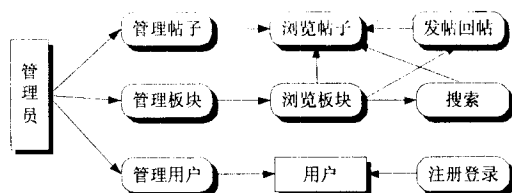


图 5 论坛系统模块关系图

4 数据库设计

数据库是整个系统的重要后台支撑^[6],也是整个系统设计的关键环节,需要仔细设计以避免将来扩充和更新功能时引起大量的重复工作。具体来说,数据库系统要面向 3 种对象,即系统用户、版主、管理员。

表 1 是论坛系统数据库中的主要数据项和数据结构的总结。

表 1 金融论坛系统数据库结构

数据表	数据项
用户信息	用户名、密码、头衔、头像、等级……
版面信息	版面类型、版面分类号、版主、父版面、版面样式……
帖子信息	所属版面、父帖 ID、子帖 ID、用户、主题、内容、时间……
公告信息	所属版面、标题、内容、用户名、时间……
论坛参数表	论坛名称、URL、Email、Logo、IP、禁止语、背景颜色、版本、版权、用户数……

除了表中所列的重要数据表项之外,还可以包括一些其他表项,比如 log 日志、好友表、相关论坛表、短消息表、在线用户记录,等等,这里不再详细列出。

(上接第 6 页)

参考文献:

- [1] Tilaks, Abu - Ghazaleh N B, Heinzelman W. A taxonomy of wireless micro - sensor network models[J]. Mobile Computing and Communications Review, 2002, 1(2): 1 - 8.
- [2] 孙利民, 李建中. 无线传感器网络[M]. 北京: 清华大学出版社, 2005.
- [3] He T, Blum B M, Stankovic J A, et al. AIDA: Adaptive application independent data aggregation in wireless sensor networks[J]. ACM Transactions on Embedded Computing Systems (Special Issue on Dynamically Adaptable Embedded System), 2004, 3(2): 426 - 457.
- [4] 何 友, 王国宏. 多传感器信息融合[M]. 北京: 电子工业出版社, 2000.
- [5] Intanagonwiwat C, Govindan R, Estrin D. Directed diffusion: A scalable and robust communication paradigm for sensor networks[C]//In: Proc 6th Annual ACM/IEEE Int'l Conf on Mobile Computing and Networking (MobiCOM'00). Boston,

5 系统实现

所实现的面向金融领域的论坛系统可以采用 Windows 2000 Server, Windows 2003 Server 等服务期操作系统, 数据库采用 Microsoft SQL Server 2000, Web 服务器采用 Tomcat 5.5, 运行时需要 JRE 5 的运行环境。客户端可以支持所有主流浏览器, 比如 IE6, Firefox 1.5, 等等。

实际上, 如果采用跨平台的数据库, 例如 MySQL, 那么该系统可以方便地迁移到所有主流服务器操作系统上, 比如 Unix, Linux, 等等。

6 结束语

该论坛系统在增进企业与用户的信息交流、扩展交流渠道方面发挥了重要作用, 采用 MVC 模式的设计提高了论坛维护的质量和效率, 取得了较好的效果。

参考文献:

- [1] 董卫军, 周警伟. MVC 在 Web 系统中的模式与应用[J]. 计算机仿真, 2003, 20(12): 111 - 114.
- [2] 周警伟. MVC 在 Web 系统中的模式与应用[EB/OL]. 2004. <http://tech.ccidnet.com>.
- [3] 梁 循, 曾月卿. 网络金融[M]. 北京: 北京大学出版社, 2005.
- [4] 孙 莹, 许俊华, 张 毅, 等. MVC 编程模型在 Web 程序中的应用及 Java 实现[J]. 计算机工程与应用, 2001(17): 160 - 163.
- [5] 于孜清, 冉蜀阳, 李 胜. 基于 MVC 的远程教材管理系统的设计与实现[J]. 计算机技术与发展, 2006, 16(1): 185 - 189.
- [6] 易法令, 王同喜. 基于 SQL Server 的 LAN 办公系统的开发应用[J]. 计算机技术与发展, 2006, 16(1): 167 - 169.

MA: [s. n.], 2000.

- [6] Heinzelman W R, Chandrakasan A, Balakrishnan H. Energy - efficient communication protocol for wireless micro - sensor networks[C]//In: Proc 33rd Hawaii Int'l Conf on System Science. Hawaii: [s. n.], 2000.
- [7] Manjeshwar A, Agrwal D P. TEEN: A routing protocol for enhanced efficiency in wireless sensor networks[C]//In: Proc 15th Int'l Parallel and Distributed Processing Symp (IPDPS '01). San Francisco, CA: [s. n.], 2001.
- [8] Lindsey S, Raghavendra C S. PEGASIS: Power efficient gathering in sensor information systems [C]//In: Proc IEEE Aerospace conference. Washington: IEEE Computer Society, 2002.
- [9] Tan H O, Korpeoglu I. Power efficient data gathering and aggregation in wireless sensor networks[Z]. Ankara, Turkey: Department of Computer Engineering, Bilkent University, 2003.