

能量高效的传感器网络数据查询路由*

崔艳荣, 曹加恒, 何宁, 朱帆

(武汉大学 计算机学院, 武汉 430072)

摘要: 提出一种能量高效的传感器网络数据查询路由 EEDQ (energy-efficient data query), EEDQ 以 sink 节点为根节点, 构造最小路由生成树, 由 sink 节点发出查询任务, 查询结果由叶子节点向 sink 节点传输, 传输过程中进行数据汇聚。实验表明, EEDQ 相比 direct transmission, 大大提高了传感器网络的生命周期。

关键词: 数据查询; 路由协议; 数据汇聚; 定向扩散

中图分类号: TP393 **文献标志码:** A **文章编号:** 1001-3695(2008)02-0562-03

Energy-efficient data query protocol for wireless sensor network

CUI Yan-rong, CAO Jia-heng, HE Ning, ZHU Fan

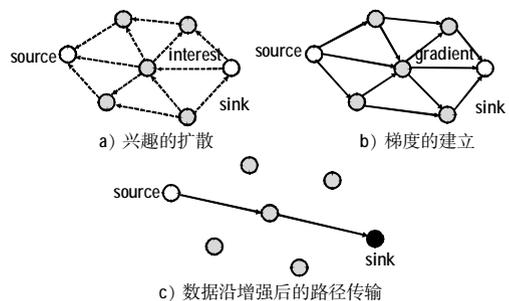
(School of Computer, Wuhan University, Wuhan 430072, China)

Abstract: This paper proposed an energy-efficient data query (EEDQ) protocol for wireless sensor network, in which designed a minimum spanning tree with the sink node as root, query task was flooding from the sink along the tree, and transmitted query results from the leaves nodes toward the sink node, in the same time, aggregated data in the parents nodes. Experimental results show that the EEDQ protocol algorithm can lead to the increase of sensor network lifetime compared with direct transmission.

Key words: data query; routing protocol; data aggregation; direct diffusion (DD)

传感器技术、微机电系统、现代网络和无线通信技术的进步, 推动了具有现代意义的无线传感器网络的产生和发展。无线传感器网络由具有感知、计算存储和通信能力的微型传感器组成, 能实现实时监测、感知和采集网络分布区域内的各种监测对象信息, 并对这些信息进行处理, 传送给需要这些信息的用户^[1,2]。传感器网络节点的资源十分有限, 主要体现在电池能量、处理能力、存储容量以及通信带宽等几个方面, 所以传感器网络中的数据查询处理不同于传统数据库, 传感器网络中的查询路由协议应该能够有效地利用节点有限的能量延长网络的生命周期^[2]。目前已有大量的研究工作从不同角度来力求延长传感器网络的生命^[3~5]。本文则从传感器网络查询路由协议的角度出发, 提出一种能量高效的传感器网络数据查询路由。

议采用休眠机制时可能造成兴趣建立的不一致。图 1 表示了 DD 协议的路由建立过程。



1 DD 协议的路由建立过程

1 相关工作

定向扩散是一种典型的基于查询的路由协议^[6], 该协议用属性/值对命名数据。汇聚节点 (sink) 通过兴趣消息发出查询任务, 采用洪泛方式传播兴趣消息到整个区域或部分区域内的所有传感器。兴趣消息用来表示查询的任务, 表达网络用户对监测区域内感兴趣的信息, 如监测区域内的温度、湿度和光照等环境信息。在兴趣消息的传播过程中, 协议逐跳地在每个传感器节点上建立反向的从数据源到汇聚节点的数据传输梯度 (gradient), 传感器节点将采集到的数据沿着梯度方向传送到汇聚节点。使用查询驱动机制按需建立路由, 避免了保存全网信息, 但是定向扩散路由在路由建立时需要一个兴趣扩散的洪泛传播, 能量和时间开销都比较大。尤其是当底层 MAC 协

有些传感器网络的应用中, 数据传输量较少或者已知事件区域, 如果采用定向扩散路由, 需要经过查询消息的洪泛传播和路由增强机制才能确定一条优化的数据传输路径。因此在这类应用中, 定向扩散路由并不是高效的路由机制。Boulis 等人提出了谣传路由 (rumor routing)^[7], 适用于数据传输量较小的传感器网络。

谣传路由的基本思想是: 事件区域中的传感器节点产生代理 (agent) 消息, 代理消息沿随机路径向外扩散传播, 同时汇聚点发送的查询消息也沿随机路径在网络中传播; 当代理消息和查询消息的传输路径交叉在一起时, 就会形成一条汇聚节点到事件区域的完整路径。图 2 所示为谣传路由原理图。

在 sink 节点多、查询请求数目很大、网络事件很少的情况下, rumor 协议较为有效。但如果事件非常多, 维护事件表和收发 agent 带来的开销会很大。

收稿日期: 2006-10-24; 修回日期: 2007-01-19 基金项目: 国家“863”计划资助项目 (2002AA4Z3450)

作者简介: 崔艳荣 (1968-), 女, 湖北仙桃人, 副教授, 博士研究生, 主要研究方向为高级数据库、传感器网络路由协议、传感器网络数据查询 (cyanr@tom.com); 曹加恒 (1946-), 男, 湖北武汉人, 教授, 博导, 主要研究方向为高级数据库、多媒体技术; 何宁 (1960-), 女, 湖北武汉人, 副教授, 博士研究生, 主要研究方向为高级数据库技术; 朱帆 (1976-), 男, 湖北武汉人, 博士研究生, 主要研究方向为高级数据库技术。

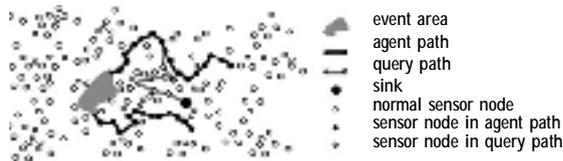


图 2 谣传路由原理图

2 系统模型

1) 网络模型

本文采用文献[8]中的网络模型: N 个传感器节点随机均匀分布在一个正方形区域 A 内, 传感器节点部署后不再移动; 惟一的基站部署在区域 A 以外的一个固定位置, 并且部署后网络不需要人为维护; 所有节点具有相似的能力(处理/通信), 并且地位平等; 节点没有装备 GPS, 也不能通过测量的方法知道其具体位置; 线发射功率可控, 即节点可以根据距离来调整发射功率的大小。

2) 无线通信模型

在无线传输中, 发射功率的衰减随着传输距离的增大而呈指数衰减。文献[9]中提出了两种信道模型: 自由空间 (free space) 和多路径衰减 (multipath fading)。当发送和接收节点之间的距离 d 小于某个值 d_0 时, 采用自由空间模型, 发射功率呈 d^2 衰减; 否则采用多路径衰减模型, 发射功率呈 d^4 衰减。

3) 无线能量模型

本文采用与文献[10]相同的无线能量模型。式(1)为发射 k bit 数据耗损的能量, 由发射电路耗损和功率放大耗损两部分构成。功率放大耗损则根据发送者和接收者之间的距离分别采用自由空间模型和多路径衰减模型。 E_{elec} 为发射电路的耗损能量; ϵ_{fs} 、 ϵ_{amp} 分别为两种信道模型下功率放大所需能量。式(2)为接收 k bit 数据的能量耗损, 仅由电路耗损引起。

$$E_{Tx} = \begin{cases} k \times E_{elec} + k \times \epsilon_{fs} a^2 & d < d_0 \\ k \times E_{elec} + k \times \epsilon_{amp} a^4 & d \geq d_0 \end{cases} \quad (1)$$

$$E_{Rx} = k \times E_{elec} \quad (2)$$

本文假设无线信道是对称的, 即从节点 u 传送消息到 v 消耗的能量等于从 v 传送同样的消息到 u 所消耗的能量。

3 高效节能的查询路由

本文提出一种高效节能的查询路由协议。该协议的主要思想是以 sink 为树根, 传感区域的其他所有节点为 sink 节点的子孙创建一棵最小代价树。查询信息由 sink 发出, 沿树的层次洪泛到各节点, 查询结果由树中的叶子节点沿反方向传送到 sink 节点。其间, 每个节点在各自的父节点处进行数据聚集。

3.1 最小代价生成树

本文所采用的网络模型是 N 个节点随机均匀地分布在正方形区域 A 中。这样, 传感器网络内的所有节点和能够实现直接通信的节点对就构成了一个连通图。因为又假设了无线通信模型是对称的, 所以该连通图是无向图。设此无向连通图 $G = \langle V, \{E\} \rangle$ 。其中: V 为传感器网络中节点的集合; $\{E\}$ 为连接两个节点的边的集合, 给每条边赋上权值; U 是顶点集合 V 的一个非空子集。若 (u, v) 是一条具有最小权值的边。其中: $u \in U, v \in V - U$, 则存在一棵包含边 (u, v) 的最小生成树。用 Prim 算法可以构造最小生成树。假设 $G = (V, \{E\})$ 是连通网, TE 是 G 上最小生成树中边的集合。算法从 $U = \{u_0\}$ (这里 u_0

为 sink 节点), $TE = \{\}$ 开始, 重复执行下述操作: 在所有 $u \in U, v \in V - U$ 的边 $(u, v) \in E$ 中找一条代价最小的边 (u_0, v_0) 并入集合 TE ; 同时 v_0 并入 U , 直到 $U = V$ 为止。此时 TE 中必有 $n - 1$ 条边, 则 $T = (V, \{TE\})$ 为 G 的最小生成树。

3.2 构造查询路由

Sink 节点发出查询任务, 采用洪泛的方式传播查询消息 M 到整个区域的所有传感器节点。任务 M 可以用如下查询语句给出:

```
Select room_no, average(light)
from sensors
group by room_no
having average(light) > 1
epoch duration 5 min
```

传感器网络中的节点收到消息 M 后, 在消息 M 中加上 sink 节点到该节点的路由信息, 然后继续广播, 即用洪泛的方式直到所有节点得到其通向 sink 节点的路由信息。如果在节点收到的路由消息中, 路由跳数大于该节点收到的其他路由消息中的路由跳数, 则丢弃该广播消息; 当广播过程结束后, 每个节点保留了所有 sink 到本节点的最短路径消息, 然后从中选择最近的上游节点作为父亲(根据权值), 并向其发送子孙消息; 最后将生成以 sink 节点为根节点的最小生成树。图 3 为查询路由示意图。叶子节点将它们的数据传送到父节点; 父节点首先利用聚集函数 f 聚集自己与子节点的数据; 然后沿着路由树向上发送部分聚集结果以及需要更新聚集的额外数据, sink 节点对收集到的所有信息进行汇聚。

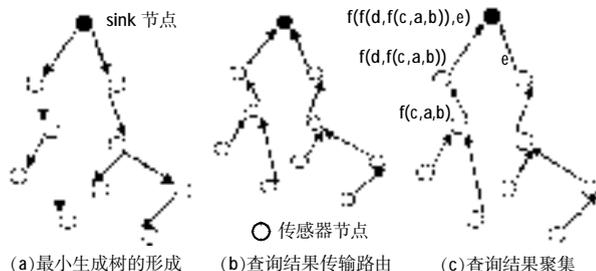


图 3 查询路由示意图

4 模拟与分析

传感器网络模型的主要参数如下: 100 个节点平均分布在 $100 \text{ m} \times 100 \text{ m}$ 的地域范围 A 内, sink 节点在传感区域 A 之外。为了测试 EEDQ 的性能, 将其与 direct transmission 进行比较, 用 C++ 完成模拟。图 4 为初始能量相同时不同协议的网络生命周期。从图中可以看到, EEDQ 比 direct transmission 将传感器网络的生命周期提高了约 11 倍(以最后一个节点死亡为标准)。图 5~7 分别表示传感器网络中初始能量为 0.25、0.5、1 J 时第一个节点死亡(FND)、一半节点死亡(HND)、最后一个节点死亡(LND)的情况。从图 5~7 中可以看出, EEDQ 比 direct transmission 分别将传感器网络的 FND、HND、LND 提高了约 3、20 和 11 倍。

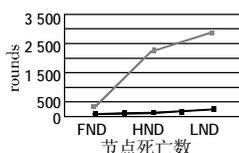


图 4 不同协议下的网络生命周期

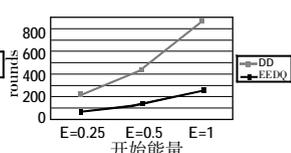


图 5 不同开始能量下的第一个节点死亡的情况

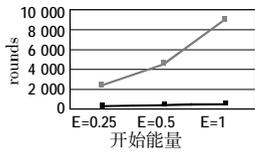


图 6 不同开始能量下的一半

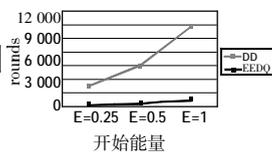


图 7 不同开始能量下的最后

5 结束语

本文提出了一种能量高效的传感器网络数据查询路由 EEDQ,由查询驱动建立以 sink 节点为根节点的最小路由生成树,任务消息由 sink 通过 flooding 给传感区域的每个节点并构造最小生成树,查询结果由叶子节点向 sink 节点传输,传输过程中每个节点在父节点处进行汇聚,sink 节点对每一轮查询结果进行汇聚。每个节点只与它的父节点通信,大大地节约了传感器节点的能量,降低了能量消耗,延长了整个传感器网络的生命周期。

参考文献:

[1] AKYILDIZ I F, SU W, SANKARASUBRAMANIAM Y, et al. Wireless sensor networks: a survey[J]. Computer Networks, 2002, 38(4):393-422.

[2] ESTRIN D, GOVINDAN R, HEIDEMANN J, et al. Next century challenges: scalable coordination in sensor networks[C]//KODESH H. Proc of the 5th Annual ACM/IEEE International Conference on Mobile Computing and Networking. New York: ACM Press, 1999:263-270.

[3] YE Wei, HEIDENMANN J, ESTRIN D. An energy efficient MAC protocol for wireless sensor networks [C]//Proc of the IEEE INFOCOM. San Francisco: IEEE Computer Society, 2002.

[4] KULIK J, HEINZELMAN W R, BALAKRISHNAN H. Negotiation-based protocols for dissemination information in wireless sensor networks[J]. ACM Wireless Networks, 2002, 8(2):169-185.

[5] HEINZELMAN W R, KULIK J, BALAKRISHNAN H. Adaptive protocols for information dissemination in wireless sensor networks[C]//Proc of the 5th Annual ACM/IEEE International Conference on Mobile Computing and Networking. New York: ACM Press:174-185.

[6] INTANAGONWIWAT C, GOVINDAN R, ESTRIN D. Directed diffusion: a scalable and robust communication paradigm for sensor networks[C]//Proc of the 6th Annual ACM/IEEE International Conference on Mobile Computing and Networking. New York: ACM Press, 2000:56-67.

[7] BRAGINSKY D, ESTRIN D. Rumor routing algorithm for sensor networks[C]//Proc of the 1st Workshop on Sensor Networks and Applications. Atlanta: ACM Press, 2002:22-31.

[8] YONIS O, FAHMY S. HEED: a hybrid, energy efficient distributed clustering approach for Ad hoc sensor networks[J]. IEEE Trans on Mobile Computing, 2004, 3(4):366-379.

[9] RAPPAPORT T. Wireless communications: principles and practice [M]. New Jersey: Prentice Hall Inc, 1996.

[10] HEINZELMAN W R, CHANDRAKASAN A, BALAKRISHNAN H. An application specific protocol architecture for wireless microsensor networks[J]. IEEE Trans on Wireless Communications, 2002, 1(4):660-670.

(上接第 561 页)

3) 聚合实验结果

对被测网络 3 h 左右所产生的共计 292 307 条原始流进行聚合。各级聚合结果如表 2 所示。

表 2 聚合实验结果

聚合流结构	一级 (5 min)	二级 (30 min)	三级 (3 h)
基于应用层协议端口的聚合流	5 106 条	872 条	153 条
基于三层协议(UDP/TCP)的聚合流	1 442 条	245 条	43 条

2.3.4 聚合流量存储设计

根据被测网络的实际流量情况,本文设计了三级时间粒度聚合表的单表时间跨度,如表 3 所示。经实验证明,这种时间跨度的设计不会发生 MySQL 单表记录数过大从而导致查询效率低下的现象,同时也不会由于时间跨度过小而出现过多的跨表情况。

3 流量分析展现

在本系统多级聚合与存储策略的支持下,系统数据库为前端分析提供了充足且多样化的数据准备,前端程序只需通过简单的查询语句即可得到所需的数据集,极大简化了查询的工作量。图 7 即为被测网络某路由器接口 HTTP 协议的流出流量趋势实验效果。由图清晰可见每日流量呈周期性变化。

表 3 三级聚合表单表存储

聚合级别	时间粒度	时间跨度
一级	5 min	1 month
二级	30 min	1 year
三级	3 h	3 year

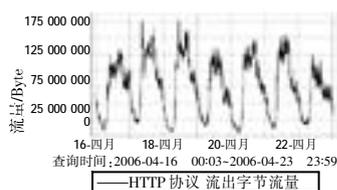


图 7 动态流量趋势分析实例

4 结束语

本文基于 NetFlow 特点,提出了一套适用于高速大流量网络的流量采集与多级聚合存储方案。流量采集通过多线程和双链表机制实现,有效提高了流量采集的可靠性。采集的原始流经内存级聚合和 DB 级聚合,并通过合理的分级存储策略进行存储组织,为前端的数据分析提供了全面支持。本系统在实际应用中取得了良好效果。下一步还将对采集和多级聚合存储方案进行改进,以丰富系统对于子网、主机等信息的静态统计分析功能,并力争为异常流量分析提供较为完善的数据支持。

参考文献:

[1] RFC 2722, Traffic flow measurement: architecture [S]. 1999.

[2] YANG Bo, LI Yi, CHEN Yue-hei, et al. A flow-based network monitoring system used for CSCW in design [C]//Proc of the 9th International Conference on Computer Supported Cooperative Work in Design. 2005:503-507.

[3] 王华. 软件实现 NetFlow 流量处理的关键技术和算法[J]. 计算机工程, 2004, 30(增):232-233.

[4] 袁梅宇. 高效率多线程网络流量采集算法研究及实践[J]. 昆明理工大学学报:理工版, 2006, 31(1):34-36.

[5] 周韶泽, 邵力耕. 高速网络环境下基于 NetFlow 的网络监测系统设计[J]. 大连铁道学院学报, 2005, 26(2):86-89.

[6] 蔡伟祥, 张凌, 宁国宁. 流量采集在网络性能监测与分析系统中的改进和实现[J]. 计算机工程, 2003, 29(15):58-59.

[7] 徐川, 唐红, 赵国锋. 降低高速网络流量测量器报文丢失率技术的研究[J]. 计算机工程, 2006, 32(16):115-117.