

# 支持无线传感器网络的实时查询技术研究

李芳芳 于 戈 杨晓春 韩 冰 程 研

(东北大学信息科学与工程学院 沈阳 110004)

**摘 要** 无线传感器网络集感知、通信、计算功能于一体,在诸多领域有重要应用,因而已经成为计算机领域研究的热点。目前的研究大都主要以节约能源消耗为目标,提出不同的查询处理策略。本文从提高查询速度,满足用户对查询响应时间的约束的角度,设计了一个支持实时查询的无线传感器网络系统;给出了系统中的关键技术,包括数据模型、查询语言模型、路由策略、查询处理及优化策略,以及双向递交机制。

**关键词** 无线传感器网络,实时查询,路由策略

## Study on Real-time Query Techniques for Wireless Sensor Networks

LI Fang-Fang YU Ge YANG Xiao-Chun HAN Bing CHENG Yan

(School of Information and Engineering, Northeastern University, Shenyang 110004)

**Abstract** Wireless sensor networks, with their primary functions as sensing, communication and computing, are applied in many kinds of fields. They have actually become a real hot topic in the computer science research. At present, most researchers are emphasizing on how to reduce the power consumption, and up to now many query strategies have been brought out. In this paper, we try to solve this problem from the angles of both query acceleration and response time limit presented by the user. Firstly we analyzed the whole problem and the requirements, and then proposed a wireless sensor network system supporting real-time query. There are also some new key techniques for the system here, such as data model, query language model, routing policy, query processing, optimization strategy, and bidirectional delivery mechanism.

**Keywords** Wireless sensor network, Real-time query, Routing strategy

## 1 引言

无线传感器节点是具有一定的感知、计算和通信能力的电子元件,外形轻巧方便,适合的工作环境广泛,由其构成的无线传感器网络在科研、军事、生产和生活等方面都具有重要的应用价值。

在无线传感器网络中,无线传感器节点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,利用无线介质和通信协议进行通信和数据传输,并采用合适的路由策略将信息传递给用户。由于无线传感器可安装在条件比较恶劣、人员较少到达的区域,因而在气候监测、生态环境考察等方面都有较多的应用。在这些应用中,要对观测和研究对象进行长期跟踪和数据搜集,因而目前无线传感器网络的研究焦点主要集中在如何节约能耗来延长系统的生命周期。

但是对于有些实时应用需求,比如抢险、救灾、作战等,系统的反应速度往往是最重要的。当发生空难的时候,救援者最关心的是每时每刻遇难者的位置和状态或事故现场的环境;当进行海底作业的时候,工作人员最关心的是海底作业人员所处位置的水压、温度、是否有凶猛的海底生物等环境信息;在军事上,能够实时知道作战士兵的情况,甚至是敌方的情况是尤为重要的,比如知道敌方潜艇的航行轨迹,能够正确决定如何攻击。因此,在这些情况下,查询的响应时间成为影

响结果价值的最重要的因素,因而查询响应时间也应成为研究的重要对象。

本文主要从提高查询效率,减少响应时间的角度出发,提出一个面向实时应用的无线传感器网络的查询系统。提出系统的逻辑层次模型、数据模型,引入时间因子的查询语言模型,路由策略,相应的查询处理及优化系统,给出了一种双向递交的查询机制,并在满足时间要求的前提下尽可能节约能耗。

## 2 相关工作

目前,国外对无线传感器网络的研究主要集中在如何节约能量。Madden 等人主要是通过调整传感器的采样率和数据包的传输率尽可能使能量的消耗最小化<sup>[1]</sup>,为此他们使用称为 Fjord 的结构来实现多查询管理。在 TinyDB<sup>[2]</sup>中,主要是对长期查询进行数据获得方法的优化。TAG<sup>[3]</sup>用于在低能源的无线环境下的数据聚集处理,它通过对流过传感器节点的数据进行计算,丢弃那些无关的数据而尽可能将相关的读数合并成一个更紧凑的记录,以此减少能量的消耗。上述系统对于查询响应的时间并未做出严格要求,也没有提出满足实时应用的查询策略。

为实现用户的应用程序,大部分无线传感器网络系统<sup>[3,9,10]</sup>都使用了类 SQL 的查询语句,如图 1 所示。其中 agg (expr)表示在所查询的属性上做聚集计算,attrs 表示其它需

李芳芳 博士研究生,主要方向为无线传感器网络及实时数据库;于 戈 教授、博士生导师,主要研究方向为数据库技术、数据挖掘、无线传感器网络等;杨晓春 副教授,主要研究方向为信息安全技术。

要查询的属性, FROM 子句指出搜寻范围。按 WHERE {selpreds}子句进行条件过滤, 利用 GROUP BY 按照 attr 的属性进行分组, 同时找出符合 HAVING 子句中条件 {havingpreds}的元组。在 EPOCH DURATION I 子句中, I 表示时间段, 通常用在长期查询的系统当中, 指出上述查询操作的执行周期。但其中并未引入与实时查询请求相关的时间因素, 无法准确地表示出查询对响应时间、数据时间的要求。

```
SELECT {agg(expr), attr}
FROM sensors
WHERE {selpreds}
GROUP BY {attr}
HAVING {havingpreds}
EPOCH DURATION I
```

图1 无线传感器网络中的查询语句

目前大部分文献中提到的路由策略大都从节约能量的角度考虑问题, 采用查询驱动数据模式或是采用层次结构进行网内聚集, 以减少通信代价, 比如 Directed diffusion<sup>[6]</sup>, Rumor<sup>[11]</sup>, Wavescheduling<sup>[12]</sup>等方法都是用时间换取能量, 在查询发出或事件发生后再探测节点间的路径。这些路由策略不适用于要求快速响应查询的实时系统中。

在文[7]中主要是从数据存储的角度出发, 讨论数据如何从底层传到上层, 并且如何通过压缩、解压缩的反复过程传到查询层。对于感兴趣的数据的查询, 则根据用户对精度的要求自顶向下层递进地获取数据。MEADOWS<sup>[8]</sup>构建了一个软件框架, 用于建模、仿真和数据分析, 其中将无线传感器网络中的各个因素次划分层次结构, 指出各部分模型的具体功能, 并在路由策略上使用了 LEACH<sup>[9]</sup>方法和洪泛式方法进行了比较, 说明利用部分汇聚和网内计算可以节约能量。但是这些系统也都没有提供对实时查询的支持。

在 TiNA<sup>[2]</sup>聚集策略中, 允许用户指定一个满足要求的近似答案的容忍度, 减少各个节点的数据传输量, 从而达到能够节约能量的目的。一些新的方法也从这个角度出发, 对即将产生的数据建立数学模型、运用机器学习等手段进行预测或过滤<sup>[4,5]</sup>, 对于结果偏差在容忍范围之内不再予以纠正, 只对于超出容忍范围的结果进行重新计算。

上述工作都是基于节约能量的角度分别从语言定义、路由算法、查询处理和查询优化以及查询结果质量评价等方面采用了不同的技术, 但是都未对无线传感器网络的实时性这一重要应用要求提出明确支持。因而我们要从实时性出发, 分别研究无线传感器网络中的查询技术。

### 3 系统结构

无线传感器网络由分布的传感器节点构成(如图 2a 所示), 这些传感器节点包括具有感应能力的感应单元、数据处理单元和通信模块, 它们可以通过传感器探测感知周围环境的温度、湿度、声音等信号, 利用红外等无线通信介质, 采用“多跳”的路由方式相互通信和数据传输, 最终目的是将数据传给用户(基站)。

根据查询的执行过程、数据的获得和传输的过程, 我们将无线传感器网络划分为三个逻辑层次(图 2b), 即传感器数据库管理系统层、传感器网络层和传感器节点层。这与 MEADOWS<sup>[8]</sup>的方法相似, 但 MEADOWS 的目的在于计算能量的耗费, 我们层次划分的依据是无线传感器网络各部分充当的角色和所实现功能。传感器节点层和传感器网络层都是由传感器节点构成。位于底层的传感器节点感应各种信

号, 比如温度、湿度和亮度等, 并通过模数转换元件将信号转换为数据。节点内部有通信元件, 可以在节点之间进行通信和数据传递。每个传感器节点的传输能力是有限的, 距离根节点较远的节点不能直接与之通信和数据传递, 要通过处于二者之间的很多个中间节点传递数据。我们将这些中间节点划分为第二个逻辑层——传感器网络层, 它除了本身可以感知数据之外, 另一个重要功能就是负责节点之间的通信、数据的传输以及对部分数据进行网内处理。一般地, 数据库管理系统层往往由服务器担当, 其主要功能是发出查询、生成查询计划、进行查询优化、对查询结果进行处理、存储历史数据等。在此基础上, 可安装用户程序, 实现实时应用。这样划分的层次及各层的功能可类比基于 TCP/IP 的网络框架模型(图 2c), 分别对应有线网络中的物理层、网络层和应用层。

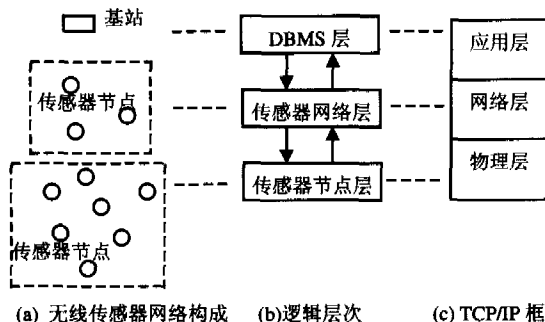


图2 无线传感器网络构成、逻辑层次、TCP/IP 框架

### 4 关键技术

在我们所设计的支持实时查询的无线传感器网络系统中, 数据模型、查询语言模型、路由策略以及查询策略和查询优化策略等都是需要研究的关键技术。

#### 4.1 数据模型

无线传感器网络的目的是要用传感器节点感应其周围环境, 获得数据, 使用户可以做出有效的判断, 得出结论以便采取相应的措施, 因而数据是系统中最重要的元素。

在实时系统中, 时间是影响结果正确性的重要因素。数据的正确性和数据所具有的价值不完全取决于数据值本身, 还要取决于采集数据本身的时间和用户获得数据的时间。对于支持实时查询的传感器网络系统, 采集和使用数据的时间成为系统在功能实现和结果正确性判定方面的必要因素和考虑条件, 因而传统的数据模型已经不能满足本系统的需要, 应使时间成为数据的固有属性。

综合无线传感器网络中数据的实时性、分布式、流动性、多元性的特征, 数据结构可以表示为  $\langle id, a, d, t \rangle$ , 其中  $id$  表示传感器编码,  $a$  表示属性向量  $\langle a_1, \dots, a_n \rangle$ ,  $d$  表示  $a$  对应的数据值向量  $\langle d_1, \dots, d_n \rangle$ ,  $t$  表示传感器节点获得数据的时间。其语义可表示为在  $t$  时刻  $id$  号传感器采集的  $a_i$  属性的数据值为  $d_i$ 。

以数据为中心的传感器网络系统, 可以看作是数据库系统。同一型号的传感器节点产生的数据, 可以理解为一个关系表。由不同类型传感器构成的网络, 就成为有不同关系表构成的数据库, 其上的操作主要是检索。

#### 4.2 查询语言模型

设计和应用无线传感器网络系统的目的是为用户提供其所关心的数据和相关信息。用户可以发出查询, 并最终得到结果的过程也同样要用查询语言来描述。必须要指明查询和

时间因素的相关性,因而我们在通用的查询语言之上加入描述时间因子的语句。

时间性主要是针对两个对象:查询和数据。与查询相关的是查询任务的截止期,即执行查询本身的时间约束。用户要在截止期内得到查询结果,否则查询结果将被视为无效。数据的时间性是指与数据本身时间属性的相关性,比如查询要指明所要求的数据的时间介于某段时间、早于或者晚于某个时间点。为此,在类 SQL 语句中要加入相应的时间条件语句,如图 3 所示,IN ts 指在 ts 时间段内获得结果,BETWEEN tb,te 指出所查询数据的时间范围介于 tb 和 te 之间,BEFORE tp 和 AFTER tp 分别表示查询的数据对象在 tp 时间点之前和之后 tn 时间段。

```
SELECT {agg(expr),attrs}
FROM sensors
WHERE {selpreds}
GROUP BY{attrs}
HAVING {havingpreds}
EPOCH DURATION I
[IN ts]
[BETWEEN tb,te]
[BEFORE tp,tn]
[AFTER tp,tn]
```

图 3 支持实时查询的无线传感器网络中查询语句

#### 4.3 路由策略

无线传感器网络的重要特点之一就是节点之间的协同工作,保证用户发出的查询能够到达相应的节点,各节点采集的数据或部分处理结果能够返回根节点,提供给用户。于是路由策略就成为关键技术之一,不仅要保证数据传输的准确性,而且保证数据传输的快速性。

为实现快速在节点间建立连接的同时尽可能节约能量的需求,我们设计了适用于平层结构无线传感器网络的路由策略。核心思想是由服务器存储全局路由表,传感器节点只维护少量信息。目前,很多研究者<sup>[1,3,4,8,10]</sup>在无线传感器网络的设计和实现上都遵循了最短路径构造路由树的原则。我们扩展这种最小路由生成树,对于网络中的所有节点采用最小生成树的方法建立节点间的父子关系的同时获得备选父节点信息,每个节点存储父节点、备选父节点、子节点信息,并将整个网络的路由表存在服务器端。对于发出的单个查询,可倒序搜索服务器中的路由表,即自底向上查找目标节点的编号,找到目标节点后,向上搜索该节点的父节点,依次重复,从而找出节点到服务器之间的完整路径。利用标签记录路由信息,然后从根节点(服务器端)出发,在传递数据的过程中同时传递标签,根据标签找到所查找的节点。对于未知查询,用户只需按照节点中的信息查找父节点而传递数据。当首选父节点发生拥塞,转而选择备选父节点,查询下推过程中,某个中继节点得到感知数据,即可计算结果。

为了保证节点之间的通信有效,每隔一段时间节点就要进行周期的路由信息广播,尤其当某个节点发生变化时,要向所有的节点发送消息,每个节点都要查看自己的路由表,及时修改路由信息。在本系统中,采用了数据信息携带传感器状态信息和路由信息的策略,当传感器节点向其父节点或是子节点传递数据时,计算所传递数据包和传感器状态数据的量。如果其和小于于一帧,则将两者放在一帧中,传递给相关节点。获得数据包的节点在得到相应计算数据的同时获得了相邻节点的状态信息,保证了路由表的正确性,并且避免了对路由信息的单独传递,节约了能量。

#### 4.4 查询处理和查询优化

无线传感器网络的本质可以看作是动态的分布式数据库。与分布式数据库相类似,数据存储在不同的节点,每个节点数据可以被复制存储在不同的节点(场地),由统一的数据库管理系统发出查询并获得结果。动态是指每个节点存储的数据是随时间变化的,节点不断产生新数据的同时也有一些数据由于过期而被删除。另外,因为查询会引起数据的传递,所以被分配于某个节点的数据也会不断产生变化。采集数据的频率、传递数据的时间、查询执行场地等问题都是在查询处理和查询优化中要解决的问题。

在支持实时查询的传感器网络中,查询的响应速度是需要考虑的重要因素。因而,我们提出双向递交查询机制,提高查询速度,减少查询延迟。所谓双向递交的查询机制是指底层基于触发式规则采集数据,根据查询应用,对数据的采集采取预先向上传递数据包的方法;而上层则是根据应用需求采取自顶向下的递进查询(如图 4 所示)。这种思想是由于查询的响应速度包含两方面:一是指发出的查询传递到每个节点的速度,另一方面是指将原数据或是计算结果传递给服务器的速度。因而,从两方面同时推进,势必会提高查询速度,减小延迟时间。

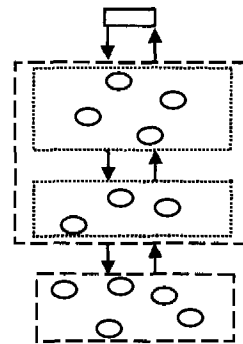


图 4 双向递交机制示意图

随着嵌入式技术的发展,具有多测度的无线传感器节点被广泛应用。于是应用对象从单一的长期查询转向了通用的多样和变化的查询,并且数据要具有时间效应。事实上,在传感器节点中,增加了定时采样的功能模块,采用的时间粒度对于不同的系统根据实际应用采用不同的值。通常可以根据已有的应用的最大公约数来定义采样的最小时间粒度。不断有新的查询加入时,则重新计算并修改采样频率,例如每 2 秒采样一次。而事实上,在底端进行数据采样的时候并不知道上层应用的具体情况,只是将数据上传到其父节点,父节点再将数据传给它父节点。而上层应用可以从根节点向下访问其子节点,直到在某个节点可以从其子节点得到数据,然后判断是否是满足时间要求的。如果能满足要求,则根据查询计算得到结果,然后将结果向上传递给该节点的父节点,直至将计算结果传递到根节点;如果不能满足时间要求,则丢弃数据,继续向下查找下一个数据,而下层的节点继续向上传递数据。

另外,与传统的数据库查询优化相比,不仅要考虑操作符的执行次序、选择度大小对执行效率的影响,而且要考虑一个重要的因素,就是数据具有时间特性,同时要解决网内计算执行场地和时间的选择问题。设计的查询优化策略的核心思想是通过计算和调整来决定何时何地(节点)执行操作最为合理,既能满足时间要求,又能使能量消耗最小,因而对于查询优化方法还要做相应的改进。

总结 与一般无线传感器网络相比,支持实时查询的无

线传感器网络系统更注重数据的时间有效性,因而从这点考虑出发,我们提出了加入时间因子的数据模型、查询语言模型、支持相应快速的路由策略,以及双向递交的查询方式,因而本系统的主要优势主要体现在以下几点:

1)数据模型能够充分体现无线传感器网络的数据特征,不但提供了原数据的位置信息,而且提供了时间属性,可以满足用户对数据时间正确性的判定,从而满足应用的实时性要求。

2)加入了时间因子的查询语言,可以满足用户对查询请求的时间约束,生成合适的查询计划,判定合理的执行时间和查询结果的正确性。

3)扩展的路由策略有效地支持查询路径的确定,数据传递的同时,数据包中包含网络状态信息,保证了数据传递的无阻塞。由此减少了为维持正确路由而周期发布的系统状态信息,因而节约了能量。

4)底层通过主动触发采集的数据和顶端发布的查询双向传递,加快数据传输速度。根据查询的时间要求,决定查询的执行策略,保证数据的正确性。

5)屏蔽了查询的特定性,对于非确定性查询具有很好的适应性,满足对通用查询的要求。

## 参考文献

1 Madden S, Franklin M J. Fjording the stream: An architecture for queries over streaming sensor data. In: Proc. of International Con-

- ference on Data Engineering, 2002. 555~566
- 2 Madden S, Franklin J, Hellerstein J M. The design of an in-acquisitional query processor for sensor networks. In: Proc. of SIGMOD, 2003. 491~502
- 3 Madden S, Franklin J, Hellerstein J M. TAG: a tiny aggregation service for ad-hoc sensor networks. In: Proc of Symposium on Operating Systems Design and Implementation, 2002. 131~146
- 4 Deshpande A, Guestrin C, Madden S R, et al. Model-Driven Data Acquisition in Sensor Networks. In: Proc of VLDB, 2004
- 5 Kumar V, Cooper B F, Navathe S B. Predictive Filtering: A Learning-Based Approach to Data Stream Filtering. In: Proc of the First Workshop on Data Management for Sensor Networks (DMSN), 2004. 17~23
- 6 Intanagonwiwat C, Govindan R, Extrin D, et al. Directed diffusion for wireless sensor networking. IEEE Transactions on Networking, 2003, 11(1): 2~16
- 7 Ganesan D, Greenstein B, Perelyubskiy D, et al. An Evaluation of Multi-resolution Storage for Sensor Networks. Proc. of International Conference on Embedded Networked Sensor Systems, 2003
- 8 Luo Q, Ni L M, He Bingsheng, et al. MEADOWS: Modeling, Emulation, and Analysis of Data of Wireless Sensor Networks. In: Proc. of the First Workshop on Data Management for Sensor Networks (DMSN), 2004. 58~67
- 9 Bonnet P, Gehrke J, Seshadri P. Towards sensor database systems. Conference on Mobile Data Management, 2001
- 10 Yao Yong, Gehrke J. Query processing for sensor networks. Proc of Conference in Innovative Data Systems Research (ICDR), 2003
- 11 Braginsky D, Estrin D. Rumor Routing Algorithm for Sensor Networks. Proceedings of the First ACM International Workshop on Wireless Sensor Networks and Applications, 2002
- 12 Trigoni N, Yao Y, Demers A, et al. WaveScheduling: Energy-Efficient Data Dissemination for Sensor Networks. Internet Draft 2004

(上接第 32 页)

将被丢弃。因此,消息不是总能到达预期的簇。随着节点数目的增加,找到路径的可能性也增加,被及时警告的节点的比例就增加了。

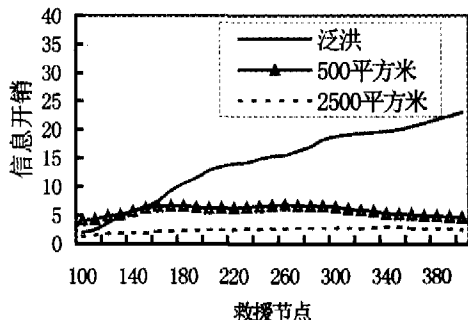


图 5 泛洪和 ACBCM 信息开销

### 3.2.2 组播成功率

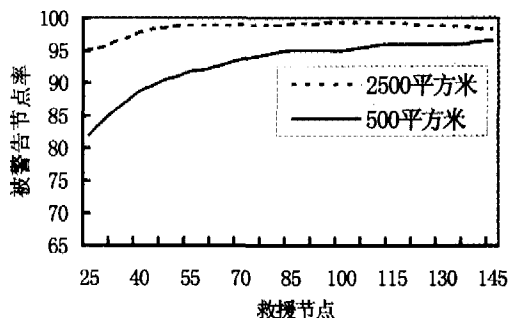


图 6 ACBCM 协议的组播成功率

图 6 显示的是组播的成功率。可见在簇规模较小时,组播成功率会导致信息开销较大。这是因为对应于一个拉请求的回应比较多,并且维护簇首的开销也较大。其次,信息开销

随着救援节点的增加变化不大,基本上趋于直线,这表明 ACBCM 协议是可以延伸到较大规模的网络中的。与泛洪算法比起来,ACBCM 算法的信息开销在大多数情况下要小得多,并且泛洪的信息开销随节点数的增加而增加。只有在少于 140 个救援节点的情况下,ACBCM 协议的开销才略比泛洪算法大,这是因为在这种情况下节点频繁地加入和离开一个簇,存在消息的开销。而较多的开销是由于许多簇是空的,邻居簇得负责那些对空簇的簇首的拉请求做出回应,这就导致簇首维护的开销较高。

**结论** 现有的基于内容的组播路由协议相对较少。本文提出的 ACBCM 协议使用了簇结构,利用推拉的机制发布信息。组播组是根据组播数据的内容来确定接收组播消息的节点集合。实验表明,本协议路由开销较小,网络的可扩展性较好,且当划分的簇较大时,性能更好,具有较低的花费和较高的组播成功率。显然,本协议既保证了所有需要接收组播消息的节点能够成功地接收数据,又有效地限制了组播的范围,从而降低了路由开销,提高了信道利用率。

## 参考文献

- 1 Tzung S C, Yuh S C, Tsai Hua-Wen. HMP: A Hierarchy-Based Multicast Protocol for Wireless Mobile Ad-Hoc Networks [A]. In: The 9<sup>th</sup> IEEE International Conference on Networks [C], 2001. 10~12
- 2 Sajama S, Zygmunt J H. ITAMAR: Independent-Tree Ad hoc Multicast Routing [A]. In: Proc. of IEEE VTS 54<sup>th</sup> [C], 2001. 600~604
- 3 Tomochika O, Jaime B K, Tatsuya S. Bandwidth Efficient Multicast Routing For Multihop Ad-hoc Wireless Networks [A]. In: Proceedings of IEEE INFOCOM [C], 2001. 1182~1191
- 4 Holger F, Jörg W, Michael K, et al. CBF: Contention-Based Forwarding for Mobile Ad-Hoc Networks [J]. Elsevier Science, 2003, 1(4): 351~369
- 5 王海涛. 分簇结构在 Ad Hoc 网络中的应用综述 [J]. 重庆邮电学院学报, 2003, 15(4): 92~98
- 6 盛敏, 李建东, 史琰. 应用于 Ad Hoc 网络中的密度自适应泛洪广播策略 [J]. 电子学报, 2004, 32(7): 1191~1194