

文章编号:1001-9081(2006)08-1772-04

## 传感器网络中基于子网的数据汇聚路由算法

王 雷<sup>1,2,3</sup>, 陈治平<sup>1,2</sup>

(1. 福建工程学院 计算机与信息科学系,福建 福州 350014;

2. 清华大学 计算机科学与技术系,北京 100084;

3. 中国联合通信有限公司 湖南分公司,湖南 长沙 410008)

(wanglei\_hn@hn165.com)

**摘要:**提出了“子网”和“能量核”的概念及一种新的蚂蚁算法,给出了一种传感器网络中基于子网的分布式数据汇聚层次路由算法 DDCHA。与集中式路由算法 AC 和 GITDC 相比,DDCHA 中各节点不需要维护任何全局信息。另外,子网与能量核的构造及蚂蚁算法的正反馈性提高了算法的数据汇聚程度,降低了算法的时间复杂性。理论分析和仿真结果都说明了该算法的有效性。

**关键词:**传感器网络;子网;能量核;位置信息;蚂蚁算法

**中图分类号:** TP393.03    **文献标识码:**A

## Subnet-based data-centric routing algorithm for sensor networks

WANG Lei<sup>1,2</sup>, CHEN Zhi-ping<sup>1,2</sup>

(1. Department of Computer and Information Science, Fujian University of Technology, Fuzhou Fujian 350014, China;

2. Department of Computer Science and Technology, Tsinghua University, Beijing 100084, China;

3. Hunan Branch, China Unicom, Changsha Hunan 410008, China)

**Abstract:** Two new concepts as subnet and energy-core were proposed, and an innovative ant algorithm was designed. Besides, a distributed data-centric hierarchical ant routing algorithm for sensor networks, which is based on subnet and named DDCHA was proposed. Compared with previous centralized algorithms such as AC and GITDC, the proposed algorithm is distributed and requires local location information only. Additionally, the positive feedback ability of ant algorithm and the construction of subnets and energy-core can improve the degree of data aggregation, and reduce the time complexity of the proposed algorithm at the same time. Both theoretical analysis and experiment results show that the algorithm is effective.

**Key words:** sensor networks; subnet; energy-core; location information; ant algorithm

## 0 引言

传感器网络(Sensor Network)中,其传感器节点资源非常有限,因此,传统网络中的路由算法不适合传感器网络,必须针对传感器网络的特性来研究新的路由算法<sup>[1]</sup>。目前提出的传感器网络路由协议主要可分为基于地址的路由协议(Address-Centric, AC)和基于数据的路由协议(Data Centric Protocol, DC)两种。典型的基于 AC 的路由协议有 SAR(Sequential Assignment Routing)<sup>[2]</sup>等。这些基于 AC 的路由协议都没有考虑数据汇聚压缩问题,任意源节点各自将数据独立发送到 Sink。而研究表明,传感器网络中通信开销要远大于数据计算开销<sup>[3]</sup>,因此人们在考虑数据汇聚压缩的基础上提出了 DC 算法思想。代表性的基于 DC 的路由协议有 DD(Directed Diffusion)<sup>[4]</sup>、GITDC(Greedy Incremental Tree Data Centric Protocol)<sup>[5,6]</sup>、LEACH(Low Energy Adaptive Clustering Hierarchy)<sup>[7]</sup> 和 GEAR(Geographic and Energy Aware Routing)<sup>[8]</sup>等。文献[6]证明了基于 DC 的路由算法传输路径经历的跳径跳数较基于 AC 的路由算法要少,而且通过模拟实验表明 GITDC 是目前节能性最优的一种 DC 算法,但它是一种集中式的路由算法。基于分布式思想,LEACH 基于

簇的概念提出了一种分布式层次路由算法,但其要求所有簇头均可直接与 Sink 节点通信,因此不具备通用性<sup>[9]</sup>。而 GEAR 基于传感节点的 GPS 定位功能提出了利用节点的位置信息来进行路由选择的思想,其缺陷是要求所有节点均具有路由功能,因此不具备良好的可扩性。

蚂蚁算法<sup>[10]</sup>是一种分布式的试探算法,在多播路由等方面已得到广泛应用。蚂蚁算法的正反馈特性使得所有蚂蚁趋向于选择同一条路径,这正与 DC 算法中的数据汇聚思想吻合,且蚂蚁算法不需要知道任何全局信息。因此,本文将蚂蚁算法原理和层次 DC 路由思想相结合,提出了“子网”和“能量核”的概念及一种新的蚂蚁算法,并在此基础上给出一种基于子网的分布式数据汇聚层次蚂蚁路由算法 DDCHA(Distributed Data Centric Hierarchical Ant Routing Algorithm based on subnets)。与前述所有算法相比,该算法具有更低的时间复杂性,更好的可扩性、鲁棒性、节能性和可维护性,仿真结果同样表明该算法具有较好的路由性能。

## 1 基本概念

### 1.1 传感器网络中的定向传播模型

在传感器网络中,传感器节点利用传感部件采集被监测

收稿日期:2006-02-13;修订日期:2006-04-11

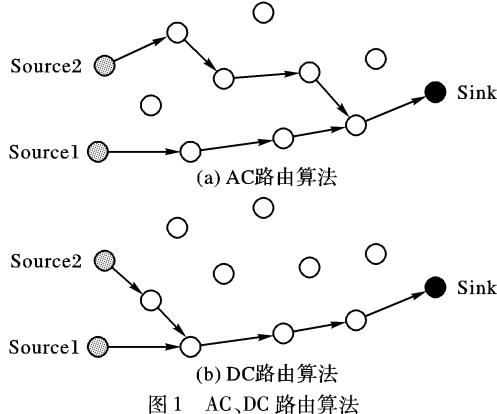
基金项目:福建省青年科技人才创新项目(2005J051);福建省自然科学基金资助项目(A0510024)

作者简介:王雷(1973-),男,湖南长沙人,副教授,博士,主要研究方向:计算机网络、机器学习; 陈治平(1971-),男,湖南益阳人,副教授,博士,主要研究方向:计算机网络、机器学习。

对象的原始数据,经过处理器部件处理后,通过无线网络传输到一个数据汇聚中心节点(Sink),Sink再通过因特网或卫星传输到用户数据处理中心。定向传播模型中,传感器网络中的节点不采用地址作为标识ID,而是以节点可以提供的数据作为寻址依据。即Sink在网络中广播以某种数据格式构成的消息询问它所感兴趣的监测数据,这种消息简称为兴趣。与这种兴趣匹配的节点(称为源节点)响应这种查询(称为事件),并回送监测数据给Sink。由于有多个节点匹配兴趣,因此这些节点都会向Sink回送同样的监测数据。为了节约节点的能量,这种数据传输可以在传输路径上进行汇聚,然后传输给Sink。

### 1.2 基于AC与基于DC的路由算法

基于AC的算法中,各传感节点独立地计算到Sink的最短路径,然后沿最短路径传输数据给Sink。如图1(a)所示,若匹配兴趣的源节点分别为Source1和Source2,则在知道网络全局信息的前提下,Source1和Source2可以分别利用Dijkstra算法计算到Sink的最短路径,然后沿这些路径传输数据到Sink。在图1(a)的情况下,所有路径经历的跳数为9。图1(b)描述了采用DD模型的DC算法,其中Source1和Source2传输的数据在经过汇聚后再统一传送给Sink,显然这种方法可以节省路径的跳数(因此节约节点的能量),对于同样的网络和源节点,DC算法所经历的所有路径的跳数只有6。容易知道,基于数据汇聚的DC算法可类比成传统网络中组播通信的逆过程。因此,与构造最优组播树一样,在每个节点都知道整个网络拓扑的条件下,寻找最优的DC算法是NP难的问题,所以在实际应用中只能用试探算法求解。



### 1.3 蚂蚁算法原理

研究表明,蚂蚁具有找到蚁巢与食物之间最短路径的能力。这种能力是靠其在所经过的路径上留下一种挥发性分泌物pheromone(称为信息素,该物质随着时间的推移会逐渐挥发消失)来实现的。蚂蚁在一条路上前进时,会留下挥发性信息素,后来的蚂蚁选择该路径的概率与当时这条路径上该物质的强度成正比。对于一条路径,选择它的蚂蚁越多,则在该路径上蚂蚁留下的信息素的强度就越大,而强度越大的信息素会吸引更多的蚂蚁,从而形成一种正反馈,通过这种正反馈,蚂蚁最终可以发现最短路径。

## 2 DDCHA

### 2.1 传感器网络中的子网模型

**定义1** 设图 $G = (V, E)$ , $G$ 称为简单连通无向图,当且仅当图 $G$ 满足以下两个条件:

- 1)  $G$ 为无自圈的、连通的无向图;

2)  $G$ 中任意两个节点之间最多有一条边。

假定在传感器网络中各节点具有相同的有效通信距离。称两个节点是相邻的即存在一条通信链路,当且仅当这两个节点在彼此有效通信距离之内。假定相邻节点之间的链路是对称的,则传感器网络的拓扑结构可以看作是一个简单连通无向图 $G = (V, E)$ ,其中 $V$ 为所有节点构成的顶点集合, $E$ 为所有链路构成的边集合。

**定义2** 图 $G$ 的节点集 $C \subseteq V$ 为核,当且仅当节点集 $C$ 满足以下条件: $\forall p \in V \Rightarrow p \in C$ 或 $p$ 为 $C$ 中的某个节点 $q$ 的邻节点。

**定义3** 给定一个图 $G = (V, E)$ ,若图 $G$ 的节点集 $C \subseteq V$ 为满足如下条件的节点集合:由 $C$ 导出的子图是连通图,且 $C$ 是图 $G$ 的一个核,则称 $C$ 为连通核。

**定义4** 任意传感节点 $p$ ,假定其对Sink节点 $R$ 的相对坐标为 $(x, y)$ ,Sink节点 $R$ 及各传感节点具有相同的有效通信半径 $r$ ,则称 $p$ 属于子网 $(m, n)$ ,当且仅当如下公式成立:

$$m = \left\lceil x \left| \frac{r}{\sqrt{2}} \right. \right\rceil \quad (1)$$

$$n = \left\lceil y \left| \frac{r}{\sqrt{2}} \right. \right\rceil \quad (2)$$

其中“ $\lceil$ ”为除法运算符;“ $\lceil \cdot \rceil$ ”为取大于等于的整数运算符。

由上述定义,易证以下定理1和2成立(限于篇幅,证明略):

**定理1** 假定Sink节点 $R$ 及各传感节点具有相同的有效通信半径 $r$ ,则任意传感器网络存在满足公式(1),(2)的唯一子网划分。

**定理2** 同一子网中的各传感节点互为邻节点。

### 2.2 分布式“能量核”自动生成算法DECA

DECA(Distributed Energy-Core Algorithm)操作于定义1中的简单连通无向图 $G$ :

1) 任意节点 $p$ 利用GPS计算自己的地理坐标,然后利用定义4计算出自己所属的子网及自己的剩余能量。

2) 任意节点 $p$ 周期性地和所有邻节点交换如下信息:(1) $p$ 的当前状态 $S_p$ (分为核心、网关或成员三种状态);(2) $p$ 的剩余能量 $E_p$ ;(3) $p$ 所属的子网 $G_p$ ;(4) $p$ 的地理坐标。通过此操作,每个节点可以获知自己邻节点的状态、剩余能量、所属子网及到邻节点的直线距离等信息。

3) 初始时,图 $G$ 中的Sink节点的状态为核心,而任意传感节点的状态均为成员。在每个周期内,任意节点 $p$ 根据自己邻节点的状态 $S_p$ 、能量 $E_p$ 及所属子网 $G_p$ 的信息,按如下规则来计算自身新的状态:

(i) 若 $p$ 所属子网 $G_p$ 内无核心,则 $G_p$ 内的各节点依据剩余能量 $E_p$ 的大小随机选举一个具有最大剩余能量的节点作为本子网的核心;

(ii) 否则,若 $p$ 不为核心或网关,且 $p$ 与其他子网中的节点相邻,则令 $p$ 为网关。

显然,DECA是分布式的,任意节点只需要知道其邻节点的有关信息。

**定理3** 若图 $G = (V, E)$ 是简单连通无向图,则由上述算法DECA得到的节点集 $\Psi = \{p | p \text{ 为核心或网关且 } p \in V\}$ 是图 $G$ 的一个连通核,由于该连通核是基于局部最大能量而自动形成的,因此称为“能量核”。

**证明** 由定理2以及算法DECA的步骤3)的(i)可知,图G中的每个节点要么为核心,要么它的邻节点中至少存在一个核心,即与某个核心相邻,因此集合 $\Psi$ 是图G的一个核。下面用归纳法证明 $\Psi$ 是连通的。

设 $p,q$ 为 $\Psi$ 中的任意两个核心节点,即 $p,q \in \Psi$ ,由于假定传感节点具有相同的有效通信半径,故为表述方便,不妨设均为1个单位长度,因此 $p,q$ 之间的距离即为 $p,q$ 之间的最短路径的长度,记为 $d(p,q)$ 。由于图G是连通的,因此 $d(p,q)$ 为有限整数。

1) 若 $d(p,q) = 1 \Rightarrow p,q$ 相邻 $\Rightarrow \Psi$ 是连通的。

2) 若 $d(p,q) = 2$ ,即存在G中的路径 $(p,r,q)$ 。由于 $d(p,q) = 2$ ,因此 $p,q$ 不相邻,故由定理2可知 $p,q$ 不属于同一子网。又因为 $r$ 为 $p,q$ 的邻节点,由算法DECA的步骤3)的(ii)可知,若 $r$ 不为核心,则 $r$ 一定为网关。故 $r \in \Psi \Rightarrow \Psi$ 是连通的。

3) 若 $d(p,q) = 3$ ,即存在G中的路径 $(p,r_1,r_2,q)$ 。由于 $d(p,r_2) = 2$ ,因此 $p,q$ 不相邻,故由定理2可知 $p,r_2$ 不属于同一子网。又因为 $r_1$ 为 $p,r_2$ 的邻节点,由算法DECA的步骤3)的(ii)可知若 $r_1$ 不为核心,则 $r_1$ 一定为网关。故 $r_1 \in \Psi$ 。又由于 $d(r_1,q) = 2$ ,同理可得 $r_2 \in \Psi \Rightarrow \Psi$ 是连通的。

4) 假设 $d(p,q) = m(m > 3)$ 时 $\Psi$ 是连通的。

5) 对 $d(p,q) = m + 1$ ,即存在G中的路径 $(p,r_1,r_2,\dots,r_m,q)$ 。由 $\Psi$ 是图G的一个核 $\Rightarrow r_2$ 为 $\Psi$ 中的核心或 $r_2$ 与 $\Psi$ 中的某个核心 $r$ 相邻 $\Rightarrow d(p,r) \leq 3$ ,由步骤1)~3)可知节点 $p,r$ 通过 $\Psi$ 可达,而 $d(r,q) \leq m$ ,由归纳假设可知 $r,q$ 通过 $\Psi$ 可达 $\Rightarrow \Psi$ 连通。

综合以上证明步骤可知 $\Psi$ 是连通的。

### 2.3 基于子网的层次蚂蚁路由算法DDCHA

假定通过算法DECA形成的能力核为 $\Psi$ ,任意路径上的信息素浓度以相同速率随时间的推移而挥发,最低浓度为0。设传感器网络中有n个源节点 $S_1, S_2, \dots, S_n$ 发送数据给Sink节点R,假定n个源节点 $S_1, S_2, \dots, S_n$ 分别属于m个子网 $G_1, G_2, \dots, G_m$ ,显然 $m \leq n$ 。假定子网 $G_1, G_2, \dots, G_m$ 的核心分别为 $C_1, C_2, \dots, C_m$ 。则算法DDCHA的具体步骤如下:

1) 初始化网络中每一条边的信息素浓度为0。

2) 将任意核心 $C_d(d \in [1, m])$ 看作一个蚁巢,从每个蚁巢发出一队(K只按顺序编号)蚂蚁,任意蚂蚁x具有相同的生命周期,且包含一个路径域Path来记录其经过的路径,一个目的节点域Destination来记录目的节点标识以及一个禁用列表Prohibit来记录禁止经过的节点集合。初始化时 $x.Path = \{C_d\}$ , $x.Destination = R$ , $x.Prohibit = \emptyset$ 。每只蚂蚁均执行步骤3)。

3) 假定蚂蚁x到达节点A,则:

(i) 若 $A = R$ ,则转步骤5)。

(ii) 若 $A \neq R$ ,且A的相邻核心或网关节点均属于 $x.Path \cup x.Prohibit$ ,则蚂蚁x回退一步到 $x.Path$ 中的上一跳,并更新 $x.Path = x.Path \setminus \{A\}$ ,同时更新 $x.Prohibit = x.Prohibit \cup \{A\}$ 。再对路径 $L_j$ 上的信息素浓度按如下公式进行刷新: $Pher_j = Pher_j - \Delta$ 。其中 $\Delta$ 为信息素浓度增量, $\Delta > 0$ 。然后转步骤3)。

(iii) 否则,设A共有t个不属于 $x.Path \cup x.Prohibit$ 的相邻核心或网关节点 $N_1, N_2, \dots, N_t, L_1, L_2, \dots, L_t$ 表示A与 $N_1, N_2, \dots, N_t$ 之间的边, $Pher_1, Pher_2, \dots, Pher_t$ 表示 $L_1, L_2, \dots, L_t$ 上的信息素浓度。

(a) 若 $R \in \{N_1, N_2, \dots, N_t\}$ ,则蚂蚁x直接选择R为其下一跳,然后更新自己的路径域 $x.Path = x.Path \cup \{R\}$ 再转步骤4)。

(b) 否则,蚂蚁x将任选一个满足条件 $Pher_j = \max\{Pher_1, Pher_2, \dots, Pher_t\}$ 的节点 $N_j(j \in [1, t])$ 作为其下一跳,并更新自己的路径域 $x.Path = x.Path \cup \{N_j\}$ ,再对路径 $L_j$ 上的信息素浓度按如下公式进行刷新: $Pher_j = Pher_j + \Delta$ ,然后转步骤3)。

4) 若x是第一只到达Sink节点R的属于蚁巢 $C_d$ 的蚂蚁,则它将沿其来路返回其所属蚁巢 $C_d$ 。在返回过程中,蚂蚁x调整沿途每条路径L的信息素浓度为 $Pher_L = \lambda * Pher_L$ ,其中 $\lambda$ 为信息素浓度调整参数, $\lambda > 1$ 。

5) 当蚂蚁x返回其源核心 $C_d$ 后,则 $C_d$ 在能量核 $\Psi$ 中找到了一条到Sink节点R的路径 $P_d$ 。属于子网 $G_d$ 的任意源节点p将按以下规则进行数据发送:

(i) 若p不是核心,则p首先将数据发送给 $C_d$ ;

(ii) 路径 $P_d$ 中的各核心节点首先对收到的数据信息进行汇聚处理,然后将汇聚后的数据依次转发到Sink节点R。

从上述算法可知,DDCHA是一种完全分布式的层次路由算法,任意节点均不需要维护任何全局信息。

### 3 算法性能分析与仿真

为了具体分析比较算法DDCHA的路由性能,依据文献[6]进行了模拟实验:首先根据给定的有效通信距离(Communication Radius, CR)在一个 $100 \times 100$ 的区域内利用坐标数据随机生成具有100个节点的连通图,并对每个节点随机生成50个单位以内的能量,然后对传感器网络中两种典型情况进行了模拟:1)情况1,随机抽取10%的节点(1个Sink节点,9个Source节点)进行仿真实验;2)情况2,随机抽取1个Sink节点,再随机选取占整个区域10%大小的子区域(即一个 $10 \times 10$ 的区域)中的节点作为Source节点进行仿真实验。实验数据均为执行算法100次,选取其平均值为最后结果。

图2给出了实验情况1和CR=35时情况2条件下,DDCHA算法中核心节点数占总节点数比例的实验数据。由图2可知,能量核的大小和CR、网络的节点稠密程度成反比,其大小仅为网络节点总数的20%~80%。因此能量核的构造也使得新算法较算法AC和GITDC具有更好的可扩展性。

图3给出了DDCHA、AC和GITDC算法从所有Source节点到Sink节点在情况1、2下时间复杂度的对比数据,其中各算法的时间复杂度计算为其所需的执行时间。由图3可知DDCHA的寻径时间复杂度小于AC和GITDC,并且随CR增大而收敛。通过实验过程分析可知,其原因在于算法AC和GITDC是集中式的,而算法DDCHA只需要在能量核中进行寻径,是分布式的。

图4给出了DDCHA、AC和GITDC算法从所有Source到Sink节点在情况1、2下经历的路径跳数的对比情况。由图4可知:DDCHA的节能性优于AC,当CR较大时DDCHA和GITDC的能耗接近。经分析,这主要是因为DDCHA利用蚂蚁算法的正反馈性使得各源节点尽可能沿相同最短路径到达Sink,使得各源节点发送的数据能在尽可能地被汇聚后统一发送到Sink,该思想和GITDC的数据汇聚思想正好一致。而AC是在没有对数据进行任何汇聚处理的基础上,源节点各自

独立沿最短路径发送数据到 Sink,因此 AC 的数据冗余度较大,需要的路径跳数较多。特别对情况 2,由图 4(b)可知 DDCHA 所需的跳数与 GITDC 所需的跳数基本相同。这主要

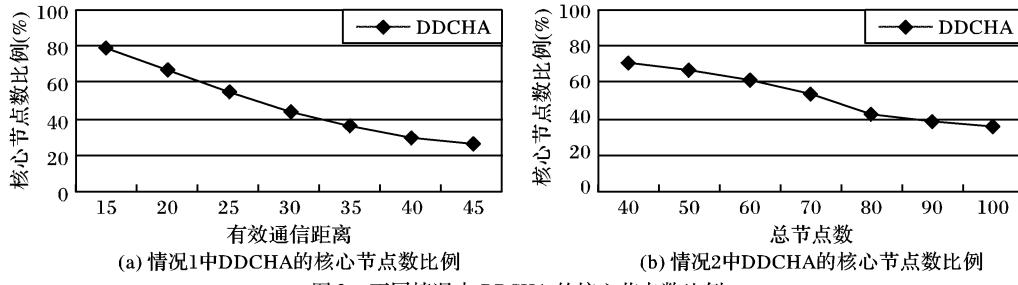


图 2 不同情况下 DDCHA 的核心节点数比例

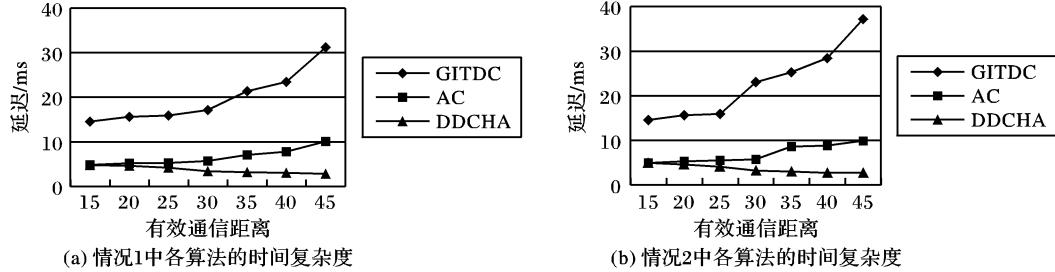


图 3 不同情况下各算法的时间复杂度

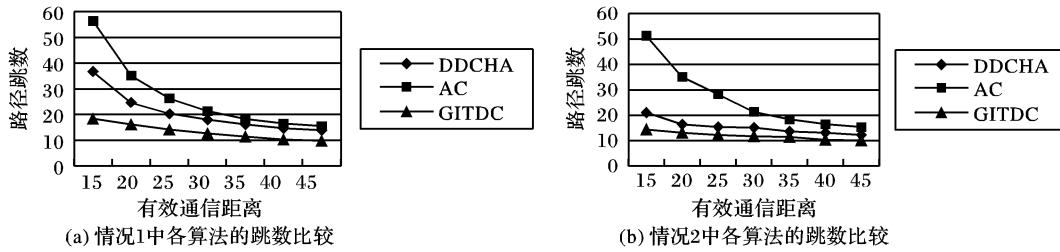


图 4 不同情况下各算法的跳数比较

综上所述可知,与算法 GITDC、AC 相比,算法 DDCHA 具有较好的路由性能,同时具有更好的可扩性和可维护性,是一种适合传感器网络特点的较好的可行路由算法。

## 4 结语

综合考虑传感器网络的节点稠密分布和节点能量有限的特性,结合蚂蚁算法原理和层次 DC 路由思想,提出了“子网”和“能量核”的概念及一种新的蚂蚁算法,给出了一种传感器网络中基于子网的数据汇聚分布式层次路由算法 DDCHA,实验结果表明,与 AC 和以 GITDC 为代表的 DC 算法相比,算法 DDCHA 具有较好的路由性能。传感器网络的研究还有许多开创性的工作值得去做,如传感器网络中的组播、广播以及 QoS 路由算法的问题等,这些也是我们正在进一步深入研究的课题。

## 参考文献:

- [1] SOHRABI K, GAO J, AILAWADHI V, et al. Protocols for self-organization of a wireless sensor network[J]. IEEE Personal Communications, 2000, 7(5): 16 - 27.
- [2] AKYILDIZ IF, SU W, SANKARASUBRAMANIAM Y, et al. A Survey on Sensor Networks[J]. IEEE Communications Magazine, 2002, 40(8): 102 - 114.
- [3] POTTIE G, KAISER W. Wireless Sensor Networks[J]. Communications of the ACM, 2000, 43(5): 51 - 58.
- [4] INTANAGONWIAT C, GOVINDAN R, ESTRIN D. Directed Diffusion: A Scalable and Robust Communication Paradigm for Sensor Networks[A]. Proceedings of ACM MobiCom[C]. Boston, MA, 2000. 56 - 67.
- [5] TAKAHASHI H, MATSUYAMA A. An approximate solution for the steiner problem in graphs[J]. Math, Japonica, 1980, 24(6): 573 - 577.
- [6] KRISHNAMACHARI B, ESTRIN D, WICKER S. Modelling Data-Centric Routing in Wireless Sensor Networks[A]. Proceedings of IEEE Infocom[C]. 2002.
- [7] KULIK J, HEINZELMAN WR, BALAKRISHNAN H. Negotiation Based Protocols for Disseminating Information in Wireless Sensor Networks[J]. ACM Wireless Networks, 2002, 8(2 - 3): 169 - 185.
- [8] YU Y, GOVINDAN R, ESTRIN D. Geographical and Energy Aware Routing: A recursive Data Dissemination Protocol for Wireless Sensor Networks[R]. UCLA Computer Science Department Technical Report UCLA/CSD-TR-01-0023, 2001.
- [9] AKYILDIZ IF, SU W, SANKARASUBRAMANIAM Y, et al. Wireless Sensor Networks: A Survey[J]. IEEE Computer Networks, 2002, 38 (4): 393 - 422.
- [10] DORIGO M, MANIEZZO V, COLORNI A. Ant system: optimization by a colony of cooperating Agents[J]. IEEE Transactions on Systems, Man, and Cybernetics-part B: cybernetics, 1996, 26(1): 29 - 41.
- [11] 林亚平,王雷,陈治平.传感器网络中一种分布式数据汇聚层次路由算法[J].电子学报,2004,32(11): 1801 - 1805.
- [12] 任丰原,黄海宁,林闯.无线传感器网络[J].软件学报,2003,14 (7): 1282 - 1291.