

Correlation-Aware Routing Algorithm for Wireless Sensor Networks

DAI Feifei¹, YU Danshi¹, PENG Li^{1,2*}

(1. School of Internet of Things Engineering, Jiangnan University, Wuxi Jiangsu 214122, China;

2. Jiangnan Energy-Aware Research Institute, Wuxi Jiangsu 214122, China)

Abstract: This paper develops energy-efficient data aggregation trees to reduce the energy consumption of wireless sensor networks by using a multi-hop data aggregation approach. First, a game theory framework is proposed to reveal the data correlation between nodes. Subsequently, a cost function is constructed by fully considering the residual energy of sensor nodes, exogenous disturbance and in-network data aggregation. Then, the Correlation-Aware Routing Algorithm (CARA) is employed to obtain the optimal solution, which minimize the total energy of the network. Finally, simulations results show that the CARA can significantly prolong the lifetime of the network.

Key words: wireless sensor networks; data aggregation; correlation-aware; routing; energy-saving

EEACC: 7230

doi:10.3969/j.issn.1004-1699.2013.07.021

无线传感器网络相关感知路由算法

戴菲菲¹, 于丹石¹, 彭力^{1,2*}

(1. 江南大学物联网工程学院, 江苏 无锡 214122; 2. 江南感知能源研究院, 江苏 无锡 214122)

摘要: 通过建立能量有效的数据融合树, 对无线传感器网络中多跳传输的数据进行动态融合, 以减少整个网络的能量消耗。利用竞争游戏算理论框架阐述数据之间的相关性; 综合考虑传感器节点的剩余能量、干扰以及网络间的数据融合, 构造算法的成本函数; 借助相关感知路由算法求解竞争游戏的可能最优解。仿真结果表明相关感知路由算法能够有效地延长整个网络的生命周期。

关键词: 无线传感器网络; 数据融合; 相关感知; 路由; 节能

中图分类号: TP212

文献标识码: A

文章编号: 1004-1699(2013)07-1004-05

无线传感器网络 WSN (Wireless Sensor Networks) 是由许多微小的传感器节点组成的, 它们是随机分布在一个区域内, 各个传感器采集的数据之间存在很大的关联性。比如, 传感器节点测量的是某个区域的温度值, 各节点采集的数据会存在时间和空间的关联。如果将整个网络采集的数据全部进行传输, 会造成网络的带宽拥挤, 造成能量的大量消耗, 甚至造成整个网络的瘫痪。无线传感器网络可以先通过对几个传感器节点数据融合, 再将融合的数据传输到 Sink 节点^[1]。

衡量无线传感器网络路由好坏的最重要的一个指标是节能。在文献[2]利用时间序列数据的时域关联性特征, 针对网内数据流量不均衡分布模式所导致的传输能耗漏斗效应问题, 提出了基于预测模式的能量感知数据路由机制。有些学者则是通过改

进分簇路由均衡整个网络的能量消耗^[3-5]。目前实现无线传感器网络的节能的目标主要是通过以下两个方法: 最小成本路由和最大生命周期路由。MER 是比较经典的最小成本路由, 一般用于减少网络的传输能量^[6]。最大生命周期路由, 主要是均衡整个网络的数据传输流量, 来延长整个网络的生命周期^[7-10]。还有其他的节能路由算法则是不仅将重心放在能源消耗, 还考虑排队延迟、拥塞或最大数据流量等其他网络性能指标^[11-12]。

基于上面文献的分析, 为了减少整个网络的能量消耗, 只考虑部分因素, 还有很多重要的被忽略。本文综合考虑传感器节点的剩余能量、干扰以及网络间的数据融合, 建立数据融合树, 并通过相关感知路由算法求解整个网络的可能最优解, 最终来延长具有冗余数据的传感器网络的生命周期。

1 系统模型

1.1 网络模型

在这里假设传感器网络的模型为 $G=(V,E)$, 假设网络中只有一个 Sink 节点记为 D , V 表示传感器源节点的集合 $V=\{Y_1, Y_2, \dots, Y_N, D\}$ 。 E 表示链路集合。对于任意 (Y_k, Y_j) 链路, 定义 $V_{jk}=1$ 表示链路为通, 为 0 则表示链路不通。

设 $G=(V,E)$ 中的任一点 Y_i 的传输到目标节点 D 的路由为 $S_i=\{Y_0, Y_1, Y_2, \dots, Y_{U-1}, Y_U\}$, 其中 Y_0 表示 Y_i , Y_U 表示 D 。 S_i 的链路集合为 $linkS_i=\{(Y_i, Y_1), (Y_1, Y_2), \dots, (Y_{U-1}, D)\}$, X_i 表示节点 Y_i 所有可能的路由链路集合的集合。

在本文将比特误码率 BER(Bit Error Rate) 考虑在内, 这是衡量链路质量的一个重要指标, 假设一个数据包的长度为 M , 节点能够接受正确的数据包的概率为 $P_c(\gamma)=(1-2BER(\gamma))^M$, $BER(\gamma)$ 与信号的干扰、信噪比 SINR(Signal to Interference and Noise Ratio) γ 有关, 在本文采用 CDMA 传输模式, 符合高斯分布, $BER(\gamma)=0.5\exp(-0.5\gamma)$ 。通过非相干性的频移键控调制, 可得一个 $SINR\gamma^*$ [13]。同时节点 Y_i 到 Y_j 的最小扩频增益为:

$$L_{ij} = \frac{\gamma * (\sum_{k=1, k \neq i, j}^N h_{kj} P_k)}{h_{ij} P_i - \gamma * \sigma^2} \quad (1)$$

式中, 链路增益 $h_{ij}=1/d_{ij}^\eta$, d_{ij} 表示节点 Y_i 到 Y_j 的距离, η 可取值为 2 或 4, σ^2 表示噪声的平方差, P_i 表示节点 Y_i 数据传输时电路能量消耗, 是一个常数。节点传输的比特率 $\Omega_{ij}=W/L_{ij}$, 其中 W 为网络带宽。

在本文的能量消耗问题上, 只考虑能量传输的能耗, 不考虑数据接收以及融合消耗的能量。从节点 Y_i 到 Y_j 传输 1 bit 消耗的能量为

$$E_b^{ij} = \frac{MP_i}{m\Omega_{ij}P_c(\gamma)} \quad (2)$$

其中 M 是数据包的长度, m 是数据包的比特数, 另外假设网络中的节点 Y_i 数据采集的速率为 $\Psi(Y_i)$, 如果节点 Y_i 有子节点(就是邻居节点将数据传输到 Y_i 进行融合), $\Psi_i(Y_i)$ 表示节点平均数据采集速率(bits/Symbol)。

1.2 融合模型

本文融合的策略是单点融合策略, 即某节点是融合节点, 那么它就要融合它的邻居节点(步长为 1)以及自身采集到的信息。 Y_i 表示节点 Y_i 与其邻居节点的集合, 假设 Y_i 一共有 p 个邻居节点, $Y_i=\{Y_i, Y_1, Y_2, \dots, Y_p\}$, 有邻居节点的节点需所有的节

点的数据到融合节点之后才能进行数据融合。

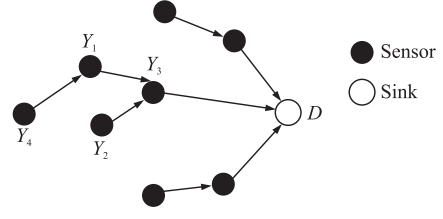


图1 简易融合树示意图

如图 1 所示, 节点 Y_3 首先会收到 Y_2 的数据, 先将收到的数据与自身的数据融合得到 $\Psi_3^{emp}(Y_3)$, 等收到 Y_1 数据后融合得到最终的 $\Psi_3(Y_3)$ 。在文献 [14] 中是将所有的数据收集完再进行融合, 这对于节点的运算和耗能上都有更高的要求。

对于 $\Psi^{emp}(Y_i, Y_j)$ 节点 Y_i 收到它的某一点邻居节点 Y_j 后得到临时的数据采集速率为:

$$\Psi^{emp}(Y_i, Y_j) = \max(\Psi(Y_i), \Psi(Y_j)) + \rho_{ij} \min(\Psi(Y_i), \Psi(Y_j)) \quad (3)$$

式中 ρ_{ij} 表示 Y_i 与 Y_j 的相关性, 一般与两节点之间的距离有关。

2 能量最小化路由框架

在上一节介绍了从节点 Y_i 到 Y_j 传输 1 bit 所消耗的能量, 在这里在介绍一下传输 1 符号消耗的能量。

$$E_s^{ij}(\Psi_i(Y_i)) = E_b^{ij} * \Psi_i(Y_i) = \frac{MP_i}{m\Omega_{ij}P_c(\gamma)} \Psi_i(Y_i) \quad (4)$$

同样的可以得到从节点 Y_i 到 Y_j 符号吞吐率:

$$\zeta_{ij}(\Psi_i(Y_i)) = \frac{W}{L_{ij}} * \frac{1}{\Psi_i(Y_i)} = \frac{\Omega_{ij}}{\Psi_i(Y_i)} \quad (5)$$

所以在整个网络中, 网络的符号吞吐率就是 S_i 的链路集合符号吞吐率的最小值。

$$\lambda_i = \min_{\forall (Y_k, Y_l) \in link(S_i)} \zeta_{kl} \quad (6)$$

整个网络的最小能量消耗为:

$$\min_{\substack{S_i \in X_i, \\ i=1, 2, \dots, N}} \sum_{i=1}^N \sum_{\forall (Y_k, Y_l) \in link(S_i)} E_s^{kl}(\Psi_k(Y_k)) \quad (7)$$

至此得到了整个算法要优化的问题, 如式(7)所示, 得到整个网络的耗能最小值。

上述的优化问题是一个 NP 问题, 所以在本文提出一个竞争游戏算法, 具体在下节介绍。

3 竞争游戏算法

在网络 $G=(V,E)$ 中每一条路由相当于一个竞争 Γ , 每一个节点相当于每一个竞争者, 他们要竞争使用在竞争算法中的共享设备, 并且将竞争游戏 Γ 定义为一个数组: $\Gamma=(G, F, (X_i)_{i \in G}, (u_f)_{f \in F})$ 。其

中 F 表示共享设备的集合为 $F = \{1, 2, \dots, m_f\}$, T 表示以 Sink 节点为根的传播树。

竞争游戏的一员 Y_i , 它可能会被下一跳节点 Y_j 选择。假设同时存在的竞争者(包括自身)共有 N 个, $S_Q = (S_1, S_2, \dots, S_N) = (S_i, S_{-i}), S_i \in X_i$ 。 Y_i 可能存在的竞争者可表示为: $S_{-i} = (S_1, S_2, \dots, S_{i-1}, S_{i+1}, \dots, S_N)$ 。在本文研究的竞争游戏算法中每一个节点就相当于是一个共享设施 f , 选择该设施的成本函数为 w_f 。并定义 $\theta_f(S_i, S_{-i})$ 为选择设施 f 为路由的一个子集。 $\theta_f(S_i, S_{-i}) = \{Y_i | f \in S_i\}$ 。竞争游戏算法的竞争者要选择一条链路 $S_i \in X_i$ 来使链路开销最小。链路开销定义为

$$\delta_i(S_i, S_{-i}) = \sum_{f \in S_i} w_f(\theta_f(S_i, S_{-i}))$$

在竞争游戏中应用函数设为:

$$u_i: u_i(S_i, S_{-i}) = -\delta_i(S_i, S_{-i}) = -\sum_{f \in S_i} w_f(\theta_f(S_i, S_{-i})) \quad (8)$$

由式(8)可看出成本函数 $w_f(\theta_f(S_i, S_{-i}))$ 是影响竞争游戏的性能重要因素。通过下一节提出的相关感知路由算法 CARA (Correlation-Aware Routing Algorithm) 来解决上述问题。

4 相关感知路由算法

在本文提出一个相关感知路由算法(CARA)来求解式(7), 这个方法考虑节点间的相关性进行数据融合, 求出网络的最优解。

对于 CARA 我们定义公共设施 f 的成本函数为:

$$w_f(\Psi_f(\theta_f(S_i, S_{-i}))) = E_b^f[\Psi_f(\theta_f(S_i, S_{-i}))] = E_s^f(\Psi_f(\theta_f(S_i, S_{-i}))) \quad (9)$$

式中 $\Psi_f(\theta_f(S_i, S_{-i}))$ 表示设施 f 数据收集的速率之和, 应用函数为

$$u_i(S_i, S_{-i}) = -\sum_{f \in S_i} [w_f(\Psi_f(\theta_f(S_i, S_{-i}))) - w_f(\Psi_f(\theta_f^i(S_i, S_{-i})))] = -\sum_{f \in S_i} E_b^f(\Psi_f(\theta_f(S_i, S_{-i}))) - \Psi_f(\theta_f^i(S_i, S_{-i})) \quad (10)$$

相应的 $\Psi_f(\theta_f^i(S_i, S_{-i}))$ 表示在节点 Y_i 不被选中的情况下, 设施 f 数据收集的速率之和。

对于竞争游戏的最优解 $P(\cdot)$ 可以定义为: 当 $P: X_1 \times X_2 \times \dots \times X_N, S_i, S'_i \in X_i$ 时, 满足:

$$u_i(S_i, S_{-i}) - u_i(S'_i, S_{-i}) = P(S_i, S_{-i}) - P(S'_i, S_{-i}) \quad (11)$$

解得式(10)的最优解为:

$$P(S_i, S_{-i}) = -\sum_{f=1}^{m_f} w_f(\Psi_f(\theta_f(S_i, S_{-i}))) \quad (12)$$

证明: 假设存在着近似最优解为

$$P(S_i, S_{-i}) = -\sum_{f \in F} w_f(\Psi_f(\theta_f(S_i, S_{-i})))$$

然后

$$\begin{aligned} P(S_i, S_{-i}) &= -\sum_{f \in F} w_f(\Psi_f(\theta_f(S_i, S_{-i}))) = \\ &[-\sum_{f \in S_i \setminus S^*} w_f(\Psi_f(\theta_f(S_i, S_{-i}))) + \\ &[-\sum_{f \in S'_i \setminus S^*} w_f(\Psi_f(\theta_f^i(S_i, S_{-i}))) - \\ &\sum_{f \in S^*} w_f(\Psi_f(\theta_f(S_i, S_{-i}))) + \\ &[-\sum_{f \in F \setminus \{S_i \cup S'_i\}} w_f(\Psi_f(\theta_f^i(S_i, S_{-i})))]] \end{aligned}$$

式中 $S^* = S_i \cap S'_i$, 同理可得 $P(S'_i, S_{-i})$, 因为上式拆分的三四项无论是 S_i 还是 S'_i 值不变, 可得:

$$\begin{aligned} P(S'_i, S_{-i}) &= -\sum_{f \in F} w_f(\Psi_f(\theta_f(S'_i, S_{-i}))) = \\ &[-\sum_{f \in S'_i \setminus S^*} w_f(\Psi_f(\theta_f(S'_i, S_{-i}))) + \\ &[-\sum_{f \in S'_i \setminus S^*} w_f(\Psi_f(\theta_f^i(S'_i, S_{-i}))) - \\ &\sum_{f \in S^*} w_f(\Psi_f(\theta_f(S_i, S_{-i}))) + \\ &[-\sum_{f \in F \setminus \{S_i \cup S'_i\}} w_f(\Psi_f(\theta_f^i(S_i, S_{-i})))]] \end{aligned}$$

所以:

$$\begin{aligned} P(S_i, S_{-i}) - P(S'_i, S_{-i}) &= \\ &[-\sum_{f \in S_i \setminus S^*} w_f(\Psi_f(\theta_f(S_i, S_{-i}))) - \sum_{f \in S'_i \setminus S^*} w_f(\Psi_f(\theta_f^i(S_i, S_{-i})))] - \\ &[-\sum_{f \in S'_i \setminus S^*} w_f(\Psi_f(\theta_f(S'_i, S_{-i}))) - \sum_{f \in S'_i \setminus S^*} w_f(\Psi_f(\theta_f^i(S'_i, S_{-i})))] \\ &\text{又因为:} \\ &u_i(S_i, S_{-i}) - u_i(S'_i, S_{-i}) = \\ &\{-\sum_{f \in S_i} [w_f(\Psi_f(\theta_f(S_i, S_{-i}))) - w_f(\Psi_f(\theta_f^i(S_i, S_{-i})))]\} - \\ &\{-\sum_{f \in S'_i} [w_f(\Psi_f(\theta_f(S'_i, S_{-i}))) - w_f(\Psi_f(\theta_f^i(S'_i, S_{-i})))]\} = \\ &\{-\sum_{f \in S_i \setminus S^*} [w_f(\Psi_f(\theta_f(S_i, S_{-i}))) - w_f(\Psi_f(\theta_f^i(S_i, S_{-i})))]\} - \\ &\{-\sum_{f \in S'_i \setminus S^*} [w_f(\Psi_f(\theta_f(S'_i, S_{-i}))) - w_f(\Psi_f(\theta_f^i(S'_i, S_{-i})))]\} = \\ &[-\sum_{f \in S'_i \setminus S^*} w_f(\Psi_f(\theta_f(S'_i, S_{-i}))) - \sum_{f \in S_i \setminus S^*} w_f(\Psi_f(\theta_f^i(S_i, S_{-i})))] - \\ &[\sum_{f \in S'_i \setminus S^*} w_f(\Psi_f(\theta_f^i(S'_i, S_{-i}))) - \sum_{f \in S_i \setminus S^*} w_f(\Psi_f(\theta_f(S_i, S_{-i})))] \end{aligned}$$

又因为

$$\begin{aligned} -\sum_{f \in S'_i \setminus S^*} w_f(\Psi_f(\theta_f^i(S_i, S_{-i}))) &= -\sum_{f \in S'_i \setminus S^*} w_f(\Psi_f(\theta_f^i(S'_i, S_{-i}))) \\ -\sum_{f \in S'_i \setminus S^*} w_f(\Psi_f(\theta_f^i(S'_i, S_{-i}))) &= -\sum_{f \in S'_i \setminus S^*} w_f(\Psi_f(\theta_f^i(S_i, S_{-i}))) \end{aligned}$$

所以

$$u_i(S_i, S_{-i}) - u_i(S'_i, S_{-i}) = P(S_i, S_{-i}) - P(S'_i, S_{-i})$$

所以 $P(S_i, S_{-i}) = - \sum_{f \in F} w_f(\Psi_f(\theta_f(S_i, S_{-i})))$ 是最优解。

5 仿真与实验

仿真场景:在 40×40 m 的区域内,随机布置 30 个传感器节点,模拟情景如图 2 所示。 ρ_{ij} 的表达式为 $\rho_{ij} = \exp(-d_{ij}^2/c)$,当式中的 c 为 0 时,说明节点 Y_i 与 Y_j 之间没有相关性,相反值越大表示相关性越大。本文的参数选择如表 1 所示。

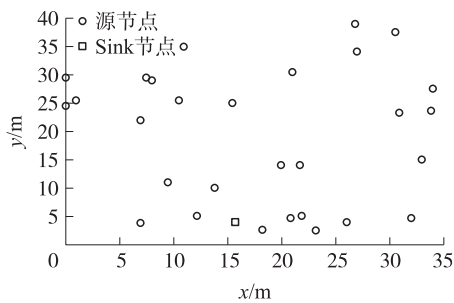


图 2 模拟情景

表 1 相关参数设定

参数	取值
p	2
σ^2	10^{-13}
W	1 MHz
P_i	10^{-2}
γ^*	5 dB
m	80
M	80
$\sum_{i \in m_f} \lambda_i$	100 kpbs

在选定节点数目为 30,相关系数为 1 000 之后,本文所述的 CARA 与经典的 MER、MEGA 算法比较如图 3 所示。在 MER 算法中,路由路径侧重于发现耗能较低的路径,而本文设计的算法则是将重心放在进行数据融合与路由相结合来减少整个网络的总耗能。

图 4 表示 MEGA、MER 以及 CARA 在不同的数据相关系数下与节点数目的关系。一共有三种相关性设置, $c = \{0, 100, 1\ 000\}$, 0 表示没有相关性,100 表示数据之间有弱相关性,1 000 则表示数据之间有强相关性。从图中我们可以看出,在各传感器节点的数据之间存在相关性的时候,本文提出的算法明显优于其他的两种经典算法。例如,在 $c = 100, N = 40$ 时,本文提出的算法与 MER、MEGA 算法比较,能量节省 23.65%、51.89%。随着节点数据相关性的不断变大,能量节省是十分明显的。

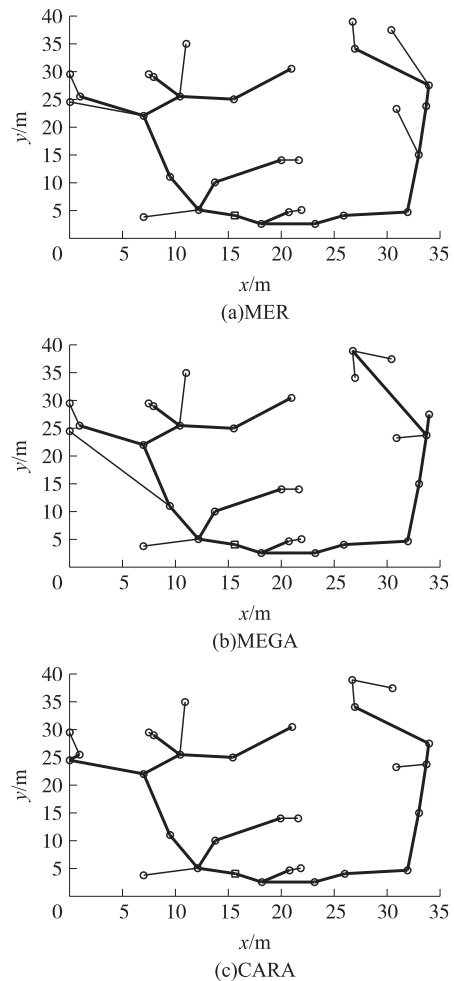


图 3 当 $N=30, c=1\ 000$ 时,三种路由算法比较

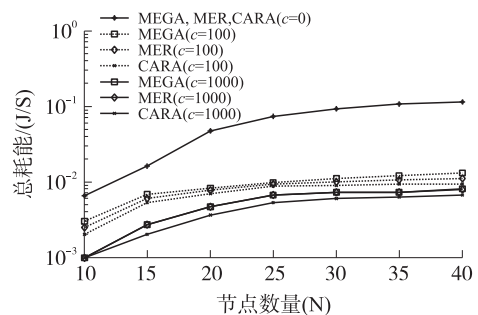


图 4 总耗能与节点数目 N 的关系

图 5 表示节点数目选定之后,节点之间的相关系数取值与网络总能耗的关系。如图所示相关系数的取值从 0 到 1 000,可以看出随着相关系数的不断

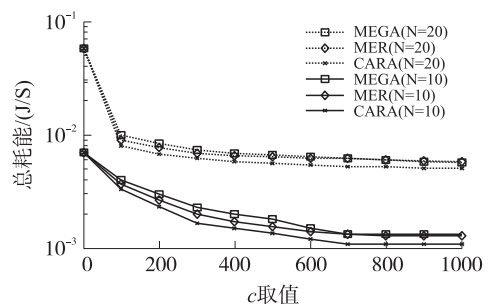


图 5 总耗能与数据相关系数 c 的关系

变大,网络的总能耗不断降低。网络间进行数据融合可以达到很明显的节能效果,例如在 $c=200$ CAR 与 MER、MEGA 比较,能量节省 15.96%、30.2%,在 $c=800$ 时,节省 15.36%、10.87%。这得益于 CARA 算法充分利用节点间的相关性,网络带宽等因素。

6 总结

本文提出 CARA 相关感知路由算法,充分考虑到区域内密集传感器之间数据的相关性,并通过竞争游戏算法找到可能最优解,使整个网络的能耗最低,最后的仿真结果表明,该算法相比现有算法,显著降低了网络的聚合数据率,均衡了节点能耗,有效地延长了网络的生命期。

未来对于 CARA 算法来说,如何将通信安全机制结合起来,进一步完善整个网络的安全。

参考文献:

- [1] 刘铁流,巫咏群. 基于能量优化的无线传感器网络分簇路由算法研究[J]. 传感技术学报,2011,24(5):764-770.
- [2] 黄如,苗澎,陈志华. 基于预测模式蚁群优化的传感网节能路由机制[J]. 传感技术学报,2010,23(5):701-707.
- [3] 彭爱平,郭晓松,蔡伟,等. 基于估计机制的分簇传感器网络数据融合算法[J]. 传感技术学报,2011,24(1):128-133.
- [4] 李建奇,曹斌芳,王立,等. 一种结合 LEACH 和 PEGASIS 协议的 WSN 的路由协议研究[J]. 传感技术学报,2012,25(2):263-267.
- [5] 陈晓娟,王卓,吴洁,等. 一种基于 LEACH 的改进 WSN 路由算法[J]. 传感技术学报,2013,26(1):116-121.
- [6] Stojmenovic I, Lin X. Power Aware Localized Routing in Wireless Networks [J]. IEEE Transactions on Parallel and Distributed Systems, 2001 12(11):1122-1133.
- [7] 谢志鹏,张卿. 无线 Ad Hoc 网络最大生命周期路由算法的诚实机制[J]. 软件学报,2009,20(9):2542-2557.
- [8] 梁俊斌,王建新,李陶深,等. 传感器网络中基于树的最大生命精确数据收集[J]. 软件学报,2010,21(9):2289-2303.
- [9] 叶伟. 延长无线传感器网络生命周期的算法研究[D]. 西安科技大学,2012.
- [10] 张卿,谢志鹏,凌波,等. 一种传感器网络最大化生命周期数据收集算法[J]. 软件学报,2005,16(11):1946-1957.
- [11] Tu L, Hong H, Zhou G. Minimum Cost Routing with a Lifetime Guarantee in Wireless Sensor Networks [C]//IEEE/ACM International Conference on Green Computing and Communications and International Conference on Cyber, Physical and Social Computing, 2010.
- [12] Mottola L, Picco G P. MUSTER: Adaptive Energy-Aware Multi-Sinkrouting in Wireless Sensor Networks [J]. IEEE Transactions on Mobile Computing 2011, 10(12):1694-1709.
- [13] Mahmood H, Comaniciu C. Interference Aware Cooperative Routing for Wireless Ad Hoc Networks [J]. Ad Hoc Networks, 2009, 7(1): 248-263.
- [14] Cristescu R, Lozano B B, Vetterli M. Networked Slepian-Wolf Theory: Theory, Algorithms and Scaling Laws [J]. IEEE Transactions on Information Theory, 2005, 51(12):4057-4073.



戴菲菲(1988-),女,江南大学物联网工程学院,硕士研究生,主要研究方向无线传感器网络数据融合、路由算法研究,daifeifei0229@qq.com;



彭力(1967-),男,江南大学物联网工程学院信息技术系副主任,博士生导师,主要研究方向无线传感器网络, pengli@jiangnan.edu.cn。