

无线传感器网络中的最大生命期基因路由算法*

唐伟⁺, 郭伟

(电子科技大学 通信抗干扰技术国家级重点实验室, 四川 成都 611731)

Maximum Lifetime Genetic Routing Algorithm in Wireless Sensor Networks

TANG Wei⁺, GUO Wei

(National Key Laboratory of Science and Technology on Communications, University of Electronic Science and Technology of China, Chengdu 611731, China)

+ Corresponding author: E-mail: gauchyler@gmail.com, http://www.ncl.uestc.edu.cn

Tang W, Guo W. Maximum lifetime genetic routing algorithm in wireless sensor networks. *Journal of Software*, 2010,21(7):1646-1656. <http://www.jos.org.cn/1000-9825/3601.htm>

Abstract: Wireless sensor networks (WSNs) consist of low-power and energy-constrained sensor nodes, and a fundamental challenge in the design of such networks is to maximize the network lifetime. In WSNs, data collected by adjacent sensor nodes usually have spatial-temporal correlations, and data aggregation technique is often used as an effective approach to remove data redundancy. Efficient usage of data aggregation technique can significantly reduce the amount of data delivery, lower the cost of overall power consumption of the network, hence increase the network lifetime. This paper studies the optimal data delivery in WSNs that takes advantage of data aggregation and nodal power control, and presents a novel routing algorithm that maximizes the network lifetime. The algorithm uses genetic algorithm (GA) to achieve an optimal selection of aggregation points, and gradient algorithm is also used to further optimize the result. The algorithm balances the power consumption of sensor nodes, and maximizes the network lifetime. Numerical results show that the proposed approach has substantially improved the network lifetime.

Key words: wireless sensor network; network lifetime; routing algorithm; genetic algorithm; gradient algorithm

摘要: 无线传感器网络(wireless sensor networks,简称 WSNs)由一组低功率且能量受限的传感器节点构成,设计此类网络的一个基本挑战便是最大化网络生命期的问题.在 WSNs 中,由于邻近传感器节点所收集的数据之间往往具有时空相关性,多采用数据聚合技术作为去除数据冗余、压缩数据大小的有效手段.合理地应用数据聚合技术,可以有效地减少数据传递量,降低网络能耗,从而延长网络生命期.研究了 WSNs 中结合数据聚合与节点功率控制的优化数据传递技术,提出了一种新的最大化网络生命期的路由算法.该算法采用遗传算法(genetic algorithm,简称 GA)最优化数据聚合点的选择,并采用梯度算法进一步优化结果.该算法均衡节点能耗,并最大化网络生命期.仿真结果表明,该算法极大地提高了网络的生命期.

关键词: 无线传感器网络;网络生命期;路由算法;遗传算法;梯度算法

* Supported by the National Natural Science Foundation of China under Grant No.10577007 (国家自然科学基金); the National Basic Research Program of China under Grant No.2009CB320405 (国家重点基础研究发展计划(973))

Received 2008-09-12; Revised 2008-12-18; Accepted 2009-02-16

中图分类号: TP393

文献标识码: A

由于无线传感器网络(wireless sensor networks,简称 WSNs)^[1]在军事、工业、医疗、运输等领域应用范围的扩大,这种网络逐渐成为近年来研究的热点.无线传感器网络通常由一组静态的传感器节点和一个固定的基站节点构成.传感器节点负责收集所探测范围内的数据,并且通过多跳方式将数据传递给基站;基站则负责接收、处理来自于传感器节点的数据.由于传感器节点往往功率较小而且能量受限,如何提高数据传递的能量效率从而提高网络的生命期,便成为网络设计中的一个关键问题.无线传感器网络的生命期通常被定义为最先因为电池能量耗尽而失效的传感器节点的生命期^[2-4].无线多跳网络的节能机制主要是结合节点功率控制进行路由优化^[5].节点功率控制机制^[6,7]通过动态调整节点的通信功率,在保证网络连通性的情况下,以最小能耗传输数据,从而达到节点本地节能的目的.路由优化^[8-13]旨在为数据传递提供一条或多条从传感器节点到基站的多跳路径,并最大限度地减小数据传递路径的整体能量消耗,以降低网络整体能耗,保证网络具有较长的生命期.除此之外,无线传感器网络还具有其自身的特殊性.由于在该类网络中,根据应用环境,邻近节点所收集的数据之间往往具有时间或空间上的关联性^[14],因而节点在收到邻居节点收集到的数据后,可以根据自身收集的数据,消除邻居数据中的冗余.而这种技术通常被称为数据聚合,其中,进行数据聚合的节点被称为其邻居节点的数据聚合点.由于数据聚合可以减少节点所需要传递的数据量,合理运用该技术可以进一步减少网络能耗,提高网络生命期.但这也为无线传感器网络中节能路由算法的设计增加了相当的难度,成为了本文研究的出发点.

无线传感器网络中,数据聚合可以依照数据压缩的程度分为完美聚合与部分聚合两类.完美聚合是指节点将自身收集到的数据连同所有来自邻居的数据一并压缩为一份数据报文的有损压缩数据聚合方式.部分聚合则是一种无损数据压缩方式,节点根据自己所收集到的数据,将来自邻居的报文进行压缩,除去其中相关冗余的部分后进行转发.在不同的聚合方式与功率控制机制下,路由算法的设计重点和难度也是不同的.在完美聚合方式下,如果不考虑节点功率控制,那么网络最大生命期的问题可以被归结为一个最小 Steiner 生成树的问题^[8],这是一个 NP-Hard 的问题.对此,已有不少文献提出了启发式路由算法,例如 GIT(greedy incremental tree)^[8]在最短跳数路由的基础上,采用贪婪算法连接各数据源,并在路径汇集处对数据进行聚合;Ant-aggregation 算法^[10]通过蚁群算法搜索最优的数据聚合生成树;DDAP(distributed data aggregation protocol)算法^[11]在网络随机选择数据聚合点,并采用地理位置路由进行数据转发.在完美聚合方式下,可以进一步考虑节点功率控制,例如 LEACH (low-energy adaptive clustering hierarchy)算法^[15]等.该算法通过动态分群算法,将节点的数据集中到邻近的群首进行聚合.然而,该算法假定基站在距离网络很远的地方,群首需要通过高功率将数据传送到基站,分群控制开销和能量开销都较大,而且忽略了基站位置对路由性能的影响^[16],一定程度地简化了问题.在部分型数据聚合方式下,同时结合节点功率控制的路由算法设计则更加困难,相关文献还不多见.最早的 MER(minimum energy routing)算法^[12]直接采用了最小功率生成树路由,节点的数据通过其父节点进行聚合,并沿着生成树传递至基站.而 MEGA(minimum energy gathering algorithm)算法^[12]同样基于最小功率生成树,所不同的是,该算法通过编码树选择数据聚合点,采用了有向图中的最短生成树获得问题的解.MLR(maximum lifetime routing)算法^[13]采用地理位置路由,其中,数据被分类为原始数据(节点自身收集的数据)和聚合数据(经过数据聚合后去除了冗余信息的数据),两种数据都按照一定比例被分发到邻居节点,并通过最优化方法对最优比例进行求解以最大化网络生命期.然而,该算法并不适合于具有“局部最大”问题^[16]的网络拓扑,同时假设数据可以被无限细分,因而又增加了数据报头的开销.

总体看来,现有的最大化生命期路由算法还存在如下几个问题:

第一,大部分的算法都只考虑路由路径的能量高效性,容易造成数据传输集中在能量最小的路径上,引起网络拥塞,也容易造成该路径上节点能量的快速消耗.因此,考虑能量负载均衡,并以网络生命期作为目标函数更加合理;

第二,算法考虑的场景都是静态网络,有些算法采用了动态路由,引入了额外的分群、路由维护开销,却很少有算法讨论过维护动态性开销和网络生命期之间的折衷关系.静态路由算法往往更适用于此类网络;

第三,有些算法采用了地理位置路由,既能实现动态路由又节省了路由开销.不过,这类路由协议的问题在于一方面不能适用于所有网络拓扑,具有“局部最大”的问题^[16];另一方面,路由路径受到地理位置路由算法的约束限制,往往不是能耗最优的路径;再者,地理位置路由算法要求节点具有位置测算的能力,例如装配 GPS(global positioning system)设备,增加了网络成本.采用基于生成树的算法既能具有地理位置路由中无路由开销的优点,又能避免其缺点,更适用于静态无线传感器网络.

基于上述分析,本文通过将节点功率控制与部分型数据聚合相结合,提出了一种新的无线传感器网络最大生命期路由算法.该算法以网络生命期作为目标函数进行优化,节点选择各自的数据聚合点进行数据聚合,并采用最小能耗路径作为聚合数据的传递路径.由于数据聚合点的解空间是离散空间而且十分巨大,无法应用传统的基于连续空间的微分学的最优化方法,因而采用了遗传算法(genetic algorithm,简称 GA)^[17]进行全局搜索,并结合梯度算法将解收敛到全局最优,最大化网络生命期.

本文第 1 节提出无线传感器网络的系统模型.第 2 节提出网络生命期的目标函数及优化方法.第 3 节通过仿真实验验证所提出的路由算法.第 4 节总结全文.

1 系统模型

1.1 网络模型

无线传感器网络可以用一个连通的随机无向图 $G(\vec{V}, E)$ 来表示,其中, \vec{V} 表示节点集合,节点 $s \in \vec{V}$ 表示基站, E 表示链接集合,节点 i 和 j 之间的链接由 (i, j) 表示.以 $|V|$ 表示网络中节点的数量, $d(i, j)$ 表示节点 i 和 j 之间的相对距离,每个节点连续地发送数据.为方便起见,令 $V = \vec{V} / \{s\}$.

1.2 能量模型

假设节点具有相同的最大传输功率 p_{\max} ,节点能够在 $[0, p_{\max}]$ 的范围内动态地调整自身的传输功率.我们采用第一阶无线电波能量模型^[18],节点 i 向 j 发送 1bit 数据所需消耗的能量可以表示为

$$\varepsilon_t(i, j) = \varepsilon_{elec} + \varepsilon_{amp} d^\eta(i, j) \quad (1)$$

其中, ε_{elec} 表示驱动发射机或接收机所需要的功率, ε_{amp} 表示发射机功率增益,而 η 表示无线电波传播衰减因子,取值范围通常为 $2 \leq \eta \leq 4$.而节点 j 接收来自 s 的 1bit 数据所需要消耗的能量可以表示为

$$\varepsilon_r = \varepsilon_{elec} \quad (2)$$

同时,记网络中节点 i 的初始能量为 e_i (单位是 J),基站 s 的能量不受限制.

1.3 数据聚合模型

由于传感器网络中,通常情况下节点所收集的数据间在时间和空间上具有关联性,数据聚合技术是通过消除信息间相关联的部分,达到压缩信息并节省数据传输开销的办法.假设节点 i 将原始数据传递给节点 j ,定义数据相关系数 $\rho(i, j) = 1 - H(ij)/H(i)$,其中, $H(i)$ 表示节点 i 发送的原始数据大小,而 $H(ij)$ 表示节点 j 根据其本地信息去掉了该原始数据中与之相关联的信息后所得到的聚合数据的大小.典型的数据相关模型包括反转模型^[12],该模型假设数据间的相关性与节点之间相对距离成反比,即 $\rho(i, j) = \frac{1}{1 + d(i, j)}$;高斯随机场模型^[19]则假设数据间

的相关性与节点之间相对距离呈指数下降趋势,其数据相关系数可以表示为 $\rho(i, j) = \exp[-\alpha d^2(i, j)]$,其中, α 是相关性参数, α 越大,相关性越小,反之,则越大.为了避免循环压缩,我们假定对于任何原始数据只发生一次数据聚合,已经被压缩过的聚合数据不再被压缩.

1.4 路由模型

本文提出一种新的无线传感器路由模型,该模型兼顾了节点功率控制、路径能耗的高效性以及数据聚合技术,借鉴并发展了 MEGA 算法^[12]中分布式数据聚合方式,将数据聚合的任务合理地分配到网络的各个节点上.该模型不同于传统的基于分群的路由算法中将数据聚合的任务集中交给少数群首节点的方式,避免了网络中

少数群首节点成为能耗瓶颈的问题.同时,该算法基于生成树传递数据,不会出现基于地理位置路由中的“局部最大”问题,也不需要类似 GPS 的地理位置定位系统的支持.

如图 1 所示,节点 $i \in V$ 从自己的邻居集合 N_i 中选择一个节点作为数据聚合点,记为 $a_i \in N_i$,并记网络中节点数据聚合点决策向量为 $\vec{A} = \{a_i\}$;决策空间 Φ 为离散笛卡尔乘积空间,即 $\Phi = \prod_{i \in V} N_i$.节点 i 将原始数据传递给数据聚合点 a_i , a_i 负责为节点 i 进行数据聚合.为了降低网络的能耗, a_i 产生的聚合数据通过最小能耗路径传递到基站.其中,最小能耗路径是能量开销最小的数据传递路径,即将拓扑图 G 的各条边 (i,j) 的权重设为发送 1bit 数据所消耗的能量 $\varepsilon_i(i,j)$.然后,采用 Dijkstra 算法或者等价路由算法获得的路径.实际上,所有最小能耗路径构成了一棵以基站为树根的最小功率生成树,记节点 i 在该树中的上一级节点为 u_i .

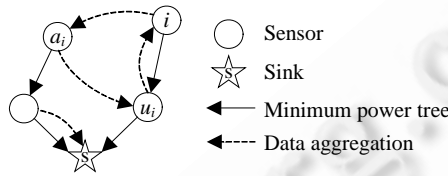


Fig.1 Routing model in WSNs

图 1 无线传感器网络路由模型

在实际应用中,可以通过如下步骤实现:第 1 步,网络中的各节点通过向邻居节点发送训练报文,估计出链路能耗 $\varepsilon_i(i,j)$ 以及数据相关系数 $\rho(i,j)$;第 2 步,由基站发起分布式 Dijkstra 算法获得最小功率生成树,并收集网络中各链路的能耗和数据相关系数;第 3 步,由基站或者外网节点计算得出优化的数据聚合点决策向量;第 4 步,由基站沿生成树将数据聚合点决策向量公布到全网,其中,每个节点只需向其子树中的节点公布对应数据聚合点的选择情况.最后,网络中各节点开始收集数据,并依照优化的数据聚合点决策向量和生成树路径传递数据.

2 无线传感器网络最大生命期路由问题

2.1 问题描述

假设传感器节点的初始能量为 $\vec{E} = \{e_i\}_{i \in V}$ (单位是 J),原始数据产生速率为 $\vec{R} = \{r_i\}_{i \in V}$ (单位是 bps),那么对应于数据聚合点决策向量 $\vec{A} \in \Phi$,节点 i 的聚合数据率 γ_i (单位是 bps) 包括两部分:其一是来自邻居节点的原始数据经过压缩得到;其二是由于节点转发来的聚合数据. γ_i 可以表示为

$$\gamma_i = \sum_{i=a_j, j \in V} r_j [1 - \rho(j, i)] + \sum_{i=u_j, j \in V} \gamma_j \quad (3)$$

其中:以节点 i 为数据聚合点的所有节点 j 的原始数据 r_j 经过数据聚合后变为 $r_j [1 - \rho(j, i)]$,对所有这样的节点求和后成为式中的第 1 项 $\sum_{i=a_j, j \in V} r_j [1 - \rho(j, i)]$;式中的第 2 项表示以节点 i 为父节点的所有节点 j 的聚合数据的求和 $\sum_{i=u_j, j \in V} \gamma_j$.

节点 i 在单位时间内的能量开销比例(相对于节点的能量 e_i) 包括 3 个部分:其一是节点 i 发送自身产生的原始数据的能耗;其二是接收并转发来自邻居节点的原始数据所消耗的能量;最后是接收并转发来自子节点的聚合数据.总体上可以用下式表示:

$$w_i(\vec{A}) = \left\{ r_i \varepsilon_i(i, a_i) + \sum_{i=a_j, j \in V} r_j [\varepsilon_r + (1 - \rho(j, i)) \varepsilon_i(i, u_i)] + \sum_{i=u_j, j \in V} \gamma_j [\varepsilon_r + \varepsilon_i(i, u_i)] \right\} / e_i \quad (4)$$

其中, $r_i \varepsilon_i(i, a_i)$ 表示节点 i 将原始数据发送到其数据聚合点 a_i 的能耗, $\sum_{i=a_j, j \in V} r_j [\varepsilon_r + (1 - \rho(j, i)) \varepsilon_i(i, u_i)]$ 表示选择节

点 i 接收来自邻居节点的原始数据并转发压缩后的聚合数据所消耗的能量,而 $\sum_{i=u_j, j \in V} \gamma_j [\varepsilon_r + \varepsilon_t(i, u_i)]$ 表示节点 i 接收来自其子节点的聚合数据并将其转发给父节点 u_i 的能耗。

于是得到节点 i 的生命期的表达式:

$$t_i(\vec{A}) = 1/w_i(\vec{A}) \quad (5)$$

根据本文上面给出的定义^[2-4],无线传感器网络的生命期是最先因为电池能量耗尽而失效的传感器节点的生命期.问题的目标函数,即网络的生命期可以表示为

$$T_{net}(\vec{A}) = \min_{i \in V} \{t_i(\vec{A})\} \quad (6)$$

因而在所提出的路由模型下,无线传感器网络最大生命期路由问题可以被归纳为一个离散空间中的非线性规划问题:在离散空间 Ω 中求解最优决策向量 \vec{A}^* ,使得目标函数 $T_{net}(\vec{A})$ 在 \vec{A}^* 处取得最大值,最大化的网络生命期以及最优决策向量可以用下式来表示:

$$\left. \begin{aligned} T_{net_max} &= \max_{\vec{A} \in \Omega} \{ \min_{i \in V} \{t_i(\vec{A})\} \} \\ \vec{A}^* &= \arg \max_{\vec{A} \in \Omega} \{ \min_{i \in V} \{t_i(\vec{A})\} \} \end{aligned} \right\} \quad (7)$$

2.2 GA-MLR算法

由于 Φ 是一个离散空间,目标函数 $T_{net}(\vec{A})$ 在该空间中不连续、不可导且非凸函数,所以不能采用传统的基于梯度或者光滑化的最优化方法来寻求 $T_{net}(\vec{A})$ 的最优解.另一方面,由于 WSN 的拓扑图 G 可以视为一个连通的随机单位圆图(connected random unit dist graph),根据文献[20]中节点连接度的概率分布结果,网络中绝大部分节点的邻居数量都不少于 2.由此,通过渐进分析易知,随着网络规模的扩大,离散空间 Ω 的增长不慢于 2 关于节点数量的指数增长速率,即 $|\Phi| = \prod_{i \in V} |N_i| = \Omega(2^{|V|})$,其中, $\Omega(\cdot)$ 表示渐进下界.因此,采用穷举搜索的办法显然也不可行.为此,本文采用现代启发式算法中相对成熟的遗传算法^[18]来求解该问题,并相应地提出了 GA-MLR (genetic algorithm-based maximum lifetime routing)算法.

遗传算法是由达尔文的进化论衍生而来的一种算法,本质上是通过大量备选解的变换和迭代,在解空间中进行并行、动态全局搜索的最优化方法.由于遗传算法具有比较完备的数学模型和理论,在解决很多 NP-Hard 问题上具有良好的性能.

2.2.1 编码方法

由于所考察问题的特殊性,我们没有采用常用的二进制 0-1 编码,而是直接采用决策向量 \vec{A} 作为编码.这是因为决策空间 Φ 是一个笛卡尔乘积空间,决策向量的各分量相互独立,在决策向量间进行交叉和变异运算之后,所得到的仍然是决策空间 Φ 中的决策向量,既不会丧失其物理意义,又简化了变换之后的维护操作.

2.2.2 初始群体的选择

通过将决策向量的每一个分量在对应节点的邻居中进行随机均匀的选择,产生一组在决策空间中均匀随机分布的决策向量.群体的规模不宜太小,不然会导致算法搜索范围狭窄,容易陷入局部最优,降低算法性能.反之,群体的规模也不宜过大,否则将增加算法复杂度,导致算法收敛缓慢.经过实验发现,当群体规模和节点数量 $|V|$ 在一个数量级上时,算法居于较好的性能,能够很快地逼近最优解.

2.2.3 适应度函数和选择方法

定义决策向量的适应度函数 $Fitness(\vec{A}) = T_{net}(\vec{A})$,即直接采用决策向量所对应的网络生命期作为适应度,网络生命期越大,对应的决策向量越适合于继续生存到下一代.

在遗传算法中,选择方式决定了算法收敛的速率.选择方式通常有两种:一种是按比例选择;另一种是基于序数的选择.我们采用基于序数的选择方式,因为该种方式常常具有更好的性能^[17].对适应度函数采取幂定标的方法,其定标可用下式表达:

$$Rank(\vec{A}) = [T_{net}(\vec{A})]^p \quad (8)$$

决策向量被选择的概率与其定标成正比,可以采用蒙特卡罗(Monte Carlo)的方法进行选取.参数 p 与遗传算法的收敛性有着密切的关系, p 越大,选择适应性度高的决策向量的可能性越大,算法收敛越快,但也越难以跳出局部最优.反之,算法收敛越慢,但却相对容易跳出局部最优.实验结果表明, p 取 1.05 时,算法能够取得比较好的效果.在选择过程中,为了更好地保证遗传算法的收敛性,我们还采用了精英模型(elitist model),即如果新一代群体中适应度最高的个体小于上一代中适应度最高的个体,那么将上一代群体中适应度最高的个体保存到新一代群体中.

2.2.4 交叉操作

在遗传算法中,交叉操作与算法的全局覆盖率以及跳出局部最优的能力密切相关,因为交叉操作有利于增加新老两代个体之间的差异性以及同一代群体中个体之间的多样性.当然,差异性和多样性也不是越大越好,否则也容易导致算法不稳定,收敛缓慢.我们采用均匀随机交叉(uniform crossover)的方法,采用该方法能够有效地跳出决策向量的邻域,最大程度地扩大搜索范围,避免算法收敛于局部最优解.同时,经过实验发现,交叉概率 $p_c \in [0.5, 1.0]$ 时,效果比较好.

2.2.5 变异操作

在遗传算法中,变异操作能够在个体的邻域中进行搜索,使得算法具有梯度算法的一些特征,与算法收敛的准确性有密切的关系.我们采用单点变异的方法,即随机选择决策向量的一个分量,从对应节点的邻居中均匀、随机地选择一个节点作为新的数据聚合点.于是,算法能够在决策向量的邻域内搜寻,提高算法收敛的准确性.实验结果表明,变异概率 $p_m \in [0.01, 0.05]$ 时效果较好.

2.2.6 算法收敛性

文献[21]使用马式链模型对采用精英模型的遗传算法的收敛性进行了研究,证明了在精英模型下遗传算法能够收敛到全局最优解.于是得到如下定理:

定理 1. 本文采用的遗传算法能够收敛到全局最优解.

从仿真实验中可以发现,算法在所选择参数下能够在不多于网络规模两倍的迭代次数内收敛,并由此可以分析所用遗传算法的时间复杂度.

2.2.7 算法时间复杂度分析

定理 2. 本文采用的遗传算法的时间复杂度为 $O(|V|^2)$.

证明:注意到,遗传算法中的种群数量为 $O(|V|)$,交叉概率为 p_c ,变异概率为 p_m .因此,算法每次迭代需要 $O(p_c|V|)$ 次交叉操作、 $O(p_m|V|)$ 次变异操作以及 $O(|V|)$ 次选择操作.所以,算法每次迭代需要 $O(|V|)$ 次操作.又注意到,算法的迭代次数为 $2|V|$,因而所用遗传算法总的时间复杂度为 $O(|V|^2)$. \square

2.3 梯度算法

由于遗传算法主要适合于在全局搜索最优解,它在单个点的领域内进行搜索的能力是比较有限的,因而在实际工程应用中往往还要配合贪婪算法进一步优化.为此,在执行遗传算法后又加入了离散空间中的梯度算法以保证决策向量收敛到全局最优解.我们针对目标函数的特殊性,提出了一种离散空间中的梯度定义和相应的梯度算法.

在离散空间 Φ 中,定义决策向量 \bar{A}_0 和 \bar{A}_1 之间的距离 $\|\bar{A}_1 - \bar{A}_0\|$ 为两向量之间的汉明距离(Hamming distance),也就是说,如果在两个决策向量中,有 k 个节点对于数据聚合点的选择不同,那么这两个决策向量之间的距离就是 k .定义决策向量 \bar{A}_0 的邻域:

$$N(\bar{A}_0) = \{\bar{A} \mid \|\bar{A} - \bar{A}_0\| = 1, \bar{A} \in \Phi\} \quad (9)$$

\bar{A}_0 邻域中的决策向量即是让 \bar{A}_0 中的任意一个节点选择不同邻居节点后所得到的新的决策向量,由此定义梯度:

$$\nabla(\bar{A}_1, \bar{A}_0) = [T_{net}(\bar{A}_1) - T_{net}(\bar{A}_0)] / \|\bar{A}_1 - \bar{A}_0\| \quad (10)$$

从梯度的定义中不难发现,对于决策向量 \bar{A}_0 邻域中的决策向量 \bar{A}_1 , $\|\bar{A}_1 - \bar{A}_0\| = 1$.于是,梯度的值等于对应网

络生命期的差值.为了使算法最快收敛,我们在决策向量 \bar{A}_0 邻域中选择这样的决策向量,该向量使得梯度的值达到最大.于是得到下面的最速梯度上升算法:

设初始决策向量为 \bar{A}_0 (该决策向量由遗传算法得到),如果在 \bar{A}_0 的邻域 $N(\bar{A}_0)$ 中存在决策向量 \bar{A}_0^* ,满足 $\nabla(\bar{A}_0^*, \bar{A}_0) > 0$,则记决策向量 $\bar{A}_1 = \arg \max_{\bar{A} \in N(\bar{A}_0)} \{\nabla(\bar{A}, \bar{A}_0)\}$.重复以上步骤.假设当算法进行到第 K 步时,对于任何决策向量 $\bar{A} \in N(\bar{A}_K)$,都有 $\nabla(\bar{A}, \bar{A}_K) \leq 0$,则算法终止,所得到的决策向量 \bar{A}_K 便是所求的最大化网络生命期的决策向量.

定理 3. 本文提出的梯度算法在有限步骤内收敛.

证明:假设定理不成立.令 $T_{net}(k) = T_{net}(\bar{A}_k)$,序列 $\{T_{net}(k)\}$ 随着 k 的增加而单调递增,对任意自然数 K 存在 $k > K$ 满足 $T_{net}(k) > T_{net}(K)$.但是 Φ 是一个有限空间,只需令 $K = |\Phi|$,即得出矛盾.因此,算法必然在有限步骤内收敛. \square

由于之前经过了遗传算法的优化,梯度算法实际上只需要很少的步骤就可以达到收敛.由于目标函数的非线性特性,算法的时间复杂度难以通过理论分析得出.不过,从下面的仿真结果分析中不难看出,算法的时间复杂度也近似为 $O(V^2)$.

3 仿真实验

由于目前同时考虑节点功率控制和部分数据聚合的路由算法还不多见,本文针对现有的 3 种算法进行仿真对比,将 GA-MLR 算法的性能与 MER,MEGA 以及 MLR 算法作了比较.

3.1 场景设置

本文采用了与文献[13]中类似的场景.在 $100\text{m} \times 100\text{m}$ 的矩形区域内随机放置 20~100 个传感器节点,所有节点的最大通信距离为 20m,第一阶无线电波功耗模型参数 $\varepsilon_{elec} = 50\text{nJ/bit}$, $\varepsilon_{amp} = 100\text{pJ/bit/m}^2$,无线传播路径衰减指数 $\eta = 2$.采用高斯随机场模型作为数据互相关模型,其中,参数 α 的取值范围为 $0.001/\text{m}^2$ 和 $0.01/\text{m}^2$ (α 越小,数据互相关程度越低).所有节点的初始能量均为 1kJ,且原始数据产生速率均为 1kb/s.所有仿真结果均为 20 个随机场景结果的平均值,而且所有场景都不出现“局部最大”问题^[16],以保证 MLR 算法能够正确运行.

3.2 仿真结果

图 2 显示了一个典型网络场景中数据聚合的情况,其中,实线代表最小功耗路径上的链路,而带箭头的虚线代表节点对数据聚点的选择情况.从图中可以看出,本文所提出的算法具有如下特点:第一,聚合数据通过最小能量路径传递,以保证数据传递的能量高效性,避免网络整体能量的快速消耗;第二,通过让各节点选择数据聚合点,将数据分配到各条最小能量路径中,均衡路径能耗;第三,由于数据聚合点选择的分散性,网络中不存在类似分群算法的集中式结构,避免出现能耗负载集中在少数节点的情况,有利于能耗负载的均衡;第四,能够适应于各种网络拓扑,不存在基于地理位置路由中对网络拓扑的限制约束,例如相对稀疏的网络等.

图 3 显示了在不同传感器节点数量下网络生命期的对比情况. α 的取值分别为 $0.001/\text{m}^2$ 和 $0.01/\text{m}^2$,以显示不同相关系数对协议的影响.从图中可以看出,随着网络规模的增大,网络生命期呈现逐渐下降的趋势.这是因为,网络的数据负载与节点数量成正比,而节点能量是一定的,造成了网络整体能耗增加,网络生命期下降.GA-MLR 算法的性能明显优于 MLR 算法的性能,而 MEGA 和 MER 算法的性能最差.MEGA 算法由于引入了数据聚合技术,考虑了相关性影响,性能略优于 MER 算法;而 MLR 算法通过数据的分流策略,均衡节点能耗,能够在一定程度上提高网络的生命期.但是,由于 MLR 算法受到地理位置路由以及算法中自由参数 c 对结果精度的影响,数据传递路径难以达到最小能耗路径对网络整体能耗的优化效果,因此优化并不充分.由于 GA-MLR 算法同时考虑了数据的相关性和数据传递路径的节能性,对网络生命期具有很大的提升,因而明显优于其他 3 种算法.GA-MLR 算法中的网络生命期在 $\alpha = 0.001/\text{m}^2$ 时,相对于 MER 和 MEGA 算法具有 56%~251%的提升,相对于 DA-MLR 算法有 30%~65%的提升;在 $\alpha = 0.01/\text{m}^2$ 时,相对于 MER 和 MEGA 算法具有 105%~225%的提升,而相对于 MLR 算法,有 22%~38%的提升.在相同算法下,随着 α 的降低,网络生命期也出现降低的趋势.这是因为,

在相关性较大的情况下,数据聚合技术能够消除大量的相关信息,减少网络的数据负载,降低了网络能耗.当相关性较小时,则主要依靠均衡节点能耗来提高网络生命期.

图 4 显示了基站节点处聚合数据的速率对比情况.我们发现,随着网络规模的扩大,聚合数据率逐渐提高.这是由于,随着节点数量的增加,会出现更多非基站邻居的节点,而这些节点的数据经过邻居进行数据聚合后被传递到基站,引起聚合数据率的提高,而且从图中可以看出,这种增加与节点数量几乎呈线性关系.图中也显示出,随着 α 的增大,数据相关性降低,能够去除的冗余数据量减少,造成了聚合数据率的提高.GA-MLR 算法相对于其他算法能够有效地降低聚合数据率,对减少网络负载、延长网络生命期提供了保障.

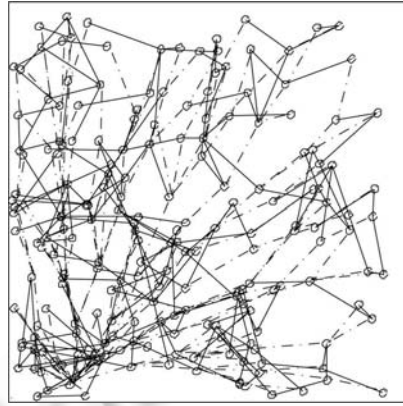


Fig.2 GA-MLR in a typical network

图 2 典型网络中 GA-MLR 算法图示

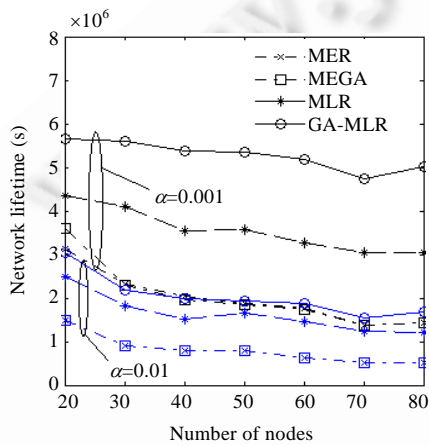


Fig.3 Network lifetime

图 3 网络生命期

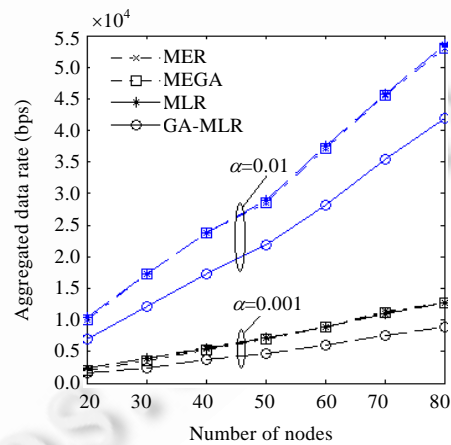


Fig.4 Aggregated data rate

图 4 聚合数据率

为了对比各算法对网络中各节点能耗的负载均衡能力,定义负载均衡系数:

$$LBF = \sqrt{\sum_{i \in V} (t_i - \bar{t})^2 / |V|} \tag{11}$$

其中,节点的平均生命期 $\bar{t} = \sum_{i \in V} t_i / |V|$.由统计学上的知识可知,负载均衡系数越小,网络中各节点的能耗越均衡.

从图 5 不难看出,GA-MLR 算法能够很好地均衡节点能耗,因此可以较好地避免瓶颈节点的出现,防止个别节点快速衰竭,延长网络的生命期.MLR 算法由于采用了分流策略,一定程度地提高了节点能耗的均衡性,但仍远远低于 GA-MLR 算法的性能,这是因为 MLR 算法中地理位置路由对转发节点的约束限制了节点负载均衡能力.

根据文献[22]的分析,基站的邻居节点——这些节点被称为 Band-1 节点——承担着将网络中其他所有节点传送来的原始和聚合数据转发给基站的任务,是网络中负载最大、能耗最高的节点,最有可能成为网络中的瓶颈节点.为此,我们针对 Band-1 节点的负载均衡情况作了对比.如图 6 所示,GA-MLR 算法很好地均衡了 Band-1 节点的负载,节点生命期差别较小.而 MLR 算法由于受到地理位置路由的约束,对 Band-1 节点的负载均衡并不理想,甚至还不如 MEGA 和 MER 等基于生成树的算法的性能.

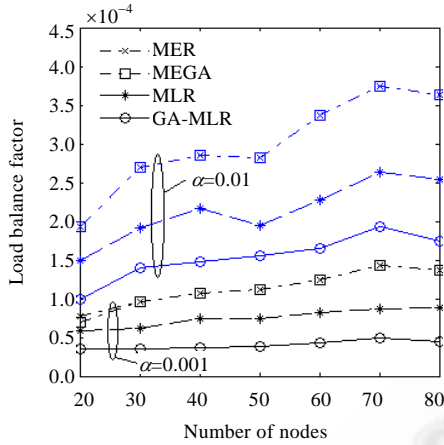


Fig.5 Load balance factor
图 5 网络负载均衡系数

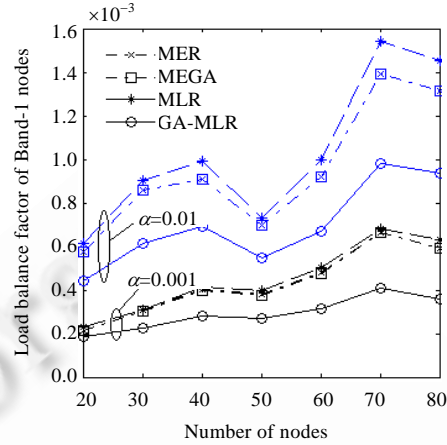
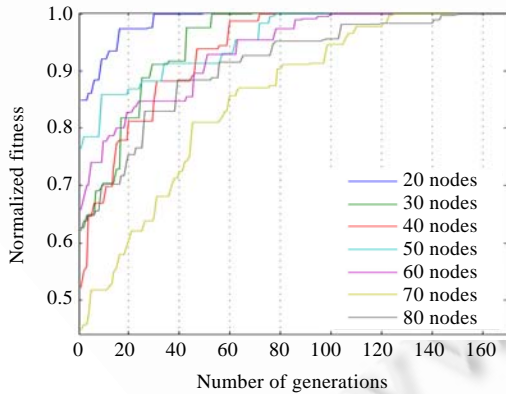
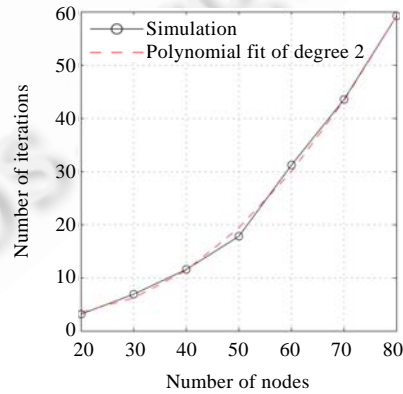


Fig.6 Load balance factor of Band-1 nodes
图 6 Band-1 节点的负载均衡系数

为了研究 GA-MLR 算法的收敛性质,我们定义了一种算法的归一化适应度,即每一代中个体的最优适应度与算法收敛时最优解的适应度的比值,这个比值随着算法迭代次数的增加而逐渐增加,并在算法收敛时变为 1. 图 7(a)显示了 GA-MLR 算法的迭代次数与算法归一化适应度之间的关系.从图中可以看出,随着节点个数的增加,算法收敛所需要的迭代次数也逐渐增加.这是因为,随着网络规模的扩大,决策空间不断增长,算法往往需要经过更多的迭代次数才能达到收敛.总的来说,算法在所考虑的场景和参数下都能够在不多于两倍网络规模的迭代次数内收敛.图 7(b)显示的是梯度算法的收敛速率,其中,实线上的各点是对应节点数量下 20 次仿真结果的平均值,而虚线是对实线点列的二次多项式插值曲线拟合.可以看到,两条曲线是比较吻合的.这也表明了在所考察的场景下,梯度算法的时间复杂度也是 $O(|V|^2)$.



(a) Convergence property of genetic algorithm
(a) 基因算法的收敛特性



(b) Convergence property of gradient algorithm
(b) 梯度算法的收敛特性

Fig.7 Convergence property of GA-MLR

图 7 GA-MLR 算法的收敛特性

4 总 结

本文提出了一种新的无线传感器网络最大生命期的路由算法.该算法综合考虑了节点功率控制、路径优化和无损数据聚合,通过各传感器节点选择数据聚合点来消除数据间的冗余,并且通过最小能耗路径向基站递交聚合数据,同时兼顾了数据聚合的高效性与节点能耗的均衡性.针对问题的目标函数是一类大离散空间中的不连续、不可导且非凸的函数这一特性,本文采用了遗传算法在数据聚合点决策向量空间中对最优决策向量进行全局搜索,并结合所提出的离散空间中的梯度算法得到近似最优决策向量,算法的时间复杂度为 $O(|V|^2)$.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,极大地提高了网络的生命期.

References:

- [1] Chong CY, Kumar SP. Sensor networks: Evolution, opportunities, and challenges. *Proc. of the IEEE*, 2003,91(8):1247–1256. [doi: 10.1109/JPROC.2003.814918]
- [2] Chen Y, Zhao Q. On the lifetime of wireless sensor networks. *IEEE Communications Letters*, 2005,9(11):976–978. [doi: 10.1109/LCOMM.2005.11010]
- [3] Ok C, Mitra P, Lee S, Kumara S. Distributed energy-adaptive routing for wireless sensor networks. In: *Proc. of the IEEE Conf. on Automation Science and Engineering*. Scottsdale: IEEE Robotics and Automation Society, 2007. 905–910.
- [4] Fan Z, Chen YP, Zhou H. An aggregator deployment protocol for energy conservation in wireless sensor networks. In: *Proc. of the IEEE Int'l Conf. on Networking: Sensing and Control*. Sanya: IEEE Systems, Man, and Cybernetics Society, 2008. 1019–1024.
- [5] Pantazis NA, Vergados DD. A survey on power control issues in wireless sensor networks. *IEEE Communications Surveys & Tutorials*, 2007,9(4):86–107. [doi: 10.1109/COMST.2007.4444752]
- [6] Simic L, Berber SM, Sowerby KW. Partner choice and power allocation for energy efficient cooperation in wireless sensor networks. In: *Proc. of the IEEE Int'l Conf. on Communications*. Beijing: IEEE Communications Society, 2008. 4255–4260.
- [7] Panichpapiboon S, Ferrari G, Tonguz OK. Optimal transmit power in wireless sensor networks. *IEEE Trans. on Mobile Computing*, 2006,5(10):1432–1447. [doi: 10.1109/TMC.2006.155]
- [8] Krishnamachari B, Estrin D, Wicker S. The impact of data aggregation in wireless sensor networks. In: *Proc. of the Int'l Conf. on Distributed Computing Systems Workshops*. Vienna: IEEE Computer Society, 2002. 575–578.
- [9] Oh H, Chae K. An energy-efficient sensor routing with low latency, scalability in wireless sensor networks. In: *Proc. of the Int'l Conf. on Multimedia and Ubiquitous Engineering*. Seoul: Science & Engineering Research Support Society, 2007. 147–152.
- [10] Misra R, Mandal C. Ant-Aggregation: Ant colony algorithm for optimal data aggregation in wireless sensor networks. In: *Proc. of the IFIP Int'l Conf. on Wireless and Optical Communications Networks*. Bangalore: IEEE Robotics & Automation Society, 2006.
- [11] Vass D, Vidacs A. Distributed data aggregation with geographical routing in wireless sensor networks. In: *Proc. of the IEEE Int'l Conf. on Pervasive Services*. Istanbul: IEEE Computer Society, 2007. 68–71.
- [12] Rickenbach P, Wattenhofer R. Gathering correlated data in sensor networks. In: *Proc. of the Joint Workshop on Foundations of Mobile Computing*. Philadelphia: ACM SIGMOBILE, 2004. 60–66.
- [13] Hua C, Yum TP. Optimal routing and data aggregation for maximizing lifetime of wireless sensor networks. *IEEE Trans. on Networking*, 2008,16(4):892–903. [doi: 10.1109/TNET.2007.901082]
- [14] Vuran MC, Akan OB. Spatio-Temporal characteristics of point and field sources in wireless sensor networks. In: *Proc. of the IEEE Int'l Conf. on Communications*. Istanbul: IEEE Communications Society, 2006. 234–239.
- [15] Heinzelman W, Chandrakasan A, Balakrishnan H. Energy-Efficient communication protocol for wireless micro-sensor networks. In: *Proc. of the Int'l Conf. on System Science*. Istanbul: IEEE Communications Society, 2000.
- [16] Karp B, Kung HT. GPSR: Greedy perimeter stateless routing for wireless networks. In: *Proc. of the MobiCom*. Boston: ACM SIGMOBILE, 2000. 243–254.
- [17] Lee KY, El-Sharkawi MA. *Modern Heuristic Optimization Techniques: Theory and Applications to Power Systems*. Hoboken: Wiley-IEEE Press, 2008. 25–70.
- [18] Heinzelman W, Chandrakasan A, Balakrishnan H. An application-specific protocol architecture for wireless micro-sensor networks. *IEEE Trans. on Wireless Communications*, 2002,1(4):660–670. [doi: 10.1109/TWC.2002.804190]

- [19] Cristescu R, Lozano BB. Lossy Network correlated data gathering with high-resolution coding. IEEE Trans. on Information Theory, 2005,52(6):2817–2824. [doi: 10.1109/TIT.2006.874536]
- [20] Onat FA, Stojmenovic I, Yanikomeroglu H. Generating random graphs for the simulation of wireless ad hoc, actuator, sensor, and internet networks. Pervasive and Mobile Computing (Elsevier), 2008,4(5):597–615.
- [21] Suzuki J. A Markov chain analysis on simple genetic algorithms. IEEE Trans. on Systems, Man, and Cybernetics, 1995,25(4): 655–650. [doi: 10.1109/21.370197]
- [22] Ammari HM, Das SK. Promoting heterogeneity, mobility, and energy-aware Voronoi diagram in wireless sensor networks. IEEE Trans. on Parallel and Distributed Systems, 2008,19(7):995–1008. [doi: 10.1109/TPDS.2008.31]



唐伟(1980—),男,四川成都人,博士生,主要研究领域为无线多跳网路由算法。



郭伟(1964—),男,教授,博士生导师,主要研究领域为移动通信网,信号与信息处理。

第 1 届中国网络攻防与系统安全会议(NADSS 2010)

征文通知

第 1 届中国网络攻防与系统安全会议(NADSS 2010)将于 2010 年 10 月 30–31 日在北京举行。本次会议将为海内外网络攻防与系统安全专家、研发人员、网络安全爱好者、黑客技术爱好者、企业用户搭建一个开放的交流平台,交流有关网络攻防与系统安全研究与应用的成果和经验,探讨网络攻防与系统安全理论与技术所面临的关键性挑战问题和研究方向,并邀请国内外著名专家到会作专题报告。现诚征网络攻防与系统安全领域各方面的最新研究成果以及有关新应用与新方法的论文、专题讨论、演示等。会议论文集仅为交流之用,无版权声明,每篇论文须有一名作者参加会议并演讲。会议推选高质量论文到《通信学报》(EI 检索)、《中国科学院研究生院学报》等专业期刊发表。

一、征文范围(但不限于)

1、**网络攻防类**:包括漏洞挖掘技术、漏洞利用技术、网络设备漏洞挖掘、Oday 漏洞及其防范、漏洞挖掘与利用成果展示、网络攻防渗透技术、网络渗透测试流程与平台研究、攻防演练平台、木马技术、恶意代码分析与检测、入侵检测、入侵容忍、蜜罐技术等。

2、**系统安全类**:包括操作系统安全、数据库安全、隐私保护、电子取证、电子商务安全、信息隐藏与水印、加密与解密、访问控制、认证与授权、可信计算、信任管理、分布式系统安全、无线安全、应急响应、风险评估等。

3、**攻防与安全新进展类**:包括 Web 2.0 安全、智能手机安全、物联网安全、云安全、信息对抗新技术、黑色产业链最新进展、网络窃密与防御、反计算机犯罪(网络钓鱼、垃圾邮件、电子欺诈、泄密)等。

二、投稿要求

1、来稿内容用中、英文撰写均可,本研讨会可接收 Word 或 PDF 格式的电子投稿。愿意被上述刊物发表的论文须在投稿附件中声明,并且应是作者没有发表过的研究成果。

2、投稿一般不超过 7000 字,通过 E-mail 发送电子稿。另附一个 TXT 文件备注稿件相关信息,包括投稿题目、作者、演讲者、征文范围、单位、通信地址、电子邮件、电话、即时通、是否愿意在《通信学报》或《中国科学院研究生院学报》发表等。

三、重要日期

论文截稿日期:2010 年 08 月 16 日

通知录用通知:2010 年 09 月 30 日

最终稿件日期:2010 年 10 月 16 日

会议日期:2010 年 10 月 30–31 日

四、联系方式

会务组:陈深龙,刘奇旭

联系电话:010-88256218

投稿邮箱:NADSS2010@gmail.com