

一种基于遗传算法与概率转发的分簇协议

陈海南^{1,2} 刘广聪¹ 吴晓鸽² 黄婷婷¹ 李 聪¹

(广东工业大学计算机学院 广州 510006)¹ (广州中国科学院先进技术研究所 广州 511458)²

摘 要 针对以 LEACH 协议为基础的分簇协议存在簇头选择不稳定以及簇间路由能耗高的问题,提出一种结合遗传算法与概率转发准则的新的分簇协议 LEACH-GPF(Genetic and Probabilistic Forwarding algorithm for Energy Balance in LEACH),对分簇算法的簇头选择以及簇头与基站的通信方式进行了优化。将所提协议与 LEACH 协议和 LEACH-C 协议进行了性能对比分析。实验结果表明,这种新的分簇算法相对于 LEACH 和 LEACH-C 协议拥有更好的能耗均衡的能力,且性能稳定。

关键词 无线传感器网络,遗传算法,概率转发,能量均衡,分簇

中图分类号 TP393 **文献标识码** A **DOI** 10.11896/j.issn.1002-137X.2015.3.015

Clustering Protocol Based on Genetic Algorithm and Probabilistic Forwarding

CHEN Hai-nan^{1,2} LIU Guang-cong¹ WU Xiao-ling² HUANG Ting-ting¹ LI Cong¹

(Department of Computer,Guangdong University of Technology,Guangzhou 510006,China)¹
(Guangzhou Institute of Advanced Technology,Chinese Academy of Sciences,Guangzhou 511458,China)²

Abstract Aiming at the insufficient of the wireless sensor networks' cluster protocol based on LEACH,a new cluster protocol based on the genetic algorithm and probabilistic forwarding method was presented. With this new protocol,the course of the cluster header's choosing and the communication method between the header and base station were optimized. To verify the performance of our protocol,the new algorithm was compared with LEACH and LEACH-C protocol. The experimental results show that the new protocol performs better in the energy balance and more stable.

Keywords Wireless sensor network,Genetic algorithm,Probabilistic forwarding,Energy balance,Clustering

1 引言

无线传感器网络(Wireless Sensor Networks,WSNs)是一种多节点、多跳协同网络,如何降低网络能耗、延长网络寿命是一个重要的问题。随着普适计算与智能算法的逐渐成熟,利用群智能优化算法来提高传感器网络的性能已经成为研究热点^[1,2]。其中负载均衡是提高传感器网络性能最直接、最有效的手段^[3-6]。常用的负载均衡方法是采用分簇路由算法对传感网络进行子网划分,利用分层梯度能量耗散代替全局能量耗散,从而实现全网的负载均衡。LEACH(Low-Energy Adaptive Clustering Hierarchy)^[7]是最经典的分簇路由协议,绝大多数的分簇路由协议都是以 LEACH 协议为基础。LEACH 系列协议以极低的计算复杂度和一定的负载均衡能力在无线传感器网络上获得了研究人员的认可。本文针对 LEACH 系列协议,在传统遗传算法的基础上提出利用概率转发的方法解决现有的分簇算法中簇头死亡快、簇间路由能耗高的问题。

2 LEACH 系列协议的分析

LEACH 协议的工作周期主要分为两个阶段,第一个是

簇的建立阶段,第二个是路径的稳态阶段。当这两个阶段结束时立即进入下一轮,通过这两个阶段的不断循环来寻找最优路由,达到平衡全网能量的目的^[8-10]。

建立阶段包括簇头选取和簇头广播两个过程。簇头选择有集中式^[11]和分布式^[7,12]两种方法,分布式簇头选择方法的优点在于速度快、算法复杂度低;缺点在于无法考虑全局负载因素,因此负载均衡能力有限,且分簇结果不稳定。集中式簇头选择方法的优点在于全局能耗影响因子可以直接干预簇头的产生,网络的负载均衡能力较强;缺点在于算法复杂度较高,且通信能耗通常也较高。簇头产生后,簇头通过广播的方式通知邻近节点,节点各自确定所属簇头,分簇过程完成。

稳态阶段,簇头负责簇内节点数据的收集与融合,并将数据传输到基站。簇头与基站的通信,可以采用直接通信或者簇间路由的方式。簇头与基站直接通信时,网络能耗受基站位置影响较大,当基站位置较远时往往造成大量簇头迅速死亡,优点在于簇与簇间不需要交互,能耗较分散。当采用簇间路由的方式时簇头通过协作将数据交付到基站,降低了簇头单次数据通信能耗,但是会导致靠近基站的簇头过早死亡,并出现基站孤立的情况。

到稿日期:2014-04-03 返修日期:2014-06-11 本文受广东省院合作项目(2012B091100266),广东省石化装备故障诊断重点实验室开放基金(GDUPTKLAB201304),广州优秀创业团队项目(7411655926875),广州市中小企业技术创新基金(2013J4400159)资助。

陈海南(1989-),男,硕士生,主要研究方向为无线传感器网络、网络拓扑与数据融合,E-mail:hn.chen@giat.ac.cn;刘广聪(1970-),男,副教授,主要研究方向为无线传感器网络;吴晓鸽(1979-),女,副研究员,主要研究方向为无线传感器网络;黄婷婷(1990-),女,硕士生,主要研究方向为无线传感器网络定位、拓扑结构研究;李 聪(1987-),男,硕士,主要研究方向为无线传感器网络路由协议。

综上,LEACH 系列协议主要存在如下问题:

- (1)如何将簇头选取与全局能耗因子进行有效关联;
- (2)如何确定簇间路由,降低簇头节点的能耗。

3 基于遗传算法与概率转发的能量均衡算法

针对现有路由算法的不足,本文提出基于遗传算法与概率转发的能量均衡算法 LEACH-GPF(Genetic and Probabilistic Forwarding algorithm for Energy Balance in LEACH)。簇头的选择是分簇算法的核心,应该依据簇头的剩余能量、计算能力、与基站的距离、节点密度以及通信代价等因素合理地确定簇头。遗传算法是依据生物进化原理,利用选择、交叉、变异、迁移等方法来对问题进行优化,以找到问题的最优解^[13-15],适用于多维度最优问题的求解。LEACH-GPF 算法采用遗传算法来进行簇头的选取,并通过概率转发函数来确定簇头与基站的通信是采用直接通信还是簇间路由,从而达到整个网络的能耗均衡的目的。

3.1 簇头组基因序列的确定

首先通过最优簇头数量约束条件^[16]确定簇头数量为 k 。基站统计全网剩余能量,并计算得到平均剩余能量 E_{avr} ,计算公式如下:

$$E_{avr} = \sum_{i=1}^N \frac{E_i}{N} \quad (1)$$

其中, E_i 是节点 ID 为 i 的剩余能量, N 是节点总数。如果节点剩余能量大于 E_{avr} , 则将该节点的 ID 序号作为染色体一个基因的编号。最终得到一个由 k 个候选簇头节点 ID 组成的基因序列。

3.2 种群优劣度适应函数

为了实现簇头选择的最优化,我们建立基于剩余能量、与基站距离、通信代价 3 个因素制约的适应度函数。其中通信代价为节点作为簇头后与簇内节点交互所付出的代价,通过与簇内节点的平均距离以及簇内节点比来衡量。

$$f_E = \sum_i^k E_i \quad (2)$$

$$f_D = \sum_i^k |P_{base} - P_i| \quad (3)$$

$$f_C = \sum_i^k (\alpha_1 \times \frac{\sum_j^N |P_j - P_i|}{n} + \alpha_2 \times \frac{n \times k}{N}) \quad (4)$$

$$F = \beta_1 \times f_E + \beta_2 \times \frac{1}{f_D} + \beta_3 \times \frac{1}{f_C} \quad (5)$$

式中, f_E 表示簇头组所含剩余能量, f_D 表示簇头节点到基站的距离, f_C 表示当前簇头组的通信代价。 P_{base} 表示基站位置, P_i, P_j 表示节点 i, d 的位置。 n 表示簇内节点个数, N 表示节点总数。 α_1, α_2 为节点 i 担任簇头时簇内节点的平均距离与簇内节点数量的权重因子,都为大于 1 的实数。 $\beta_1, \beta_2, \beta_3$ 分别为当前簇头组剩余能量、到基站距离、通信代价的权值参数,并且 $\beta_1 + \beta_2 + \beta_3 = 1 (0 < \beta_i < 1, i \in \{1, 2, 3\})$ 。适应度函数与簇头组所含能量成正比,与簇头到基站距离成反比,与簇头组通信代价成反比,因此当适应度函数 F 达到最大值时,种群达到最优,此时的簇头为最优组合。

3.3 簇头进化选择

种群的进化过程就是簇头选择的过程,通过对初始基因序列进行选择、交叉以及变异运算,逐步得到最优良的基因序列。

选择运算:首先依据适应度函数获得最优良的 x 个基因,这 x 个基因代表被初次选中的 x 个节点,并直接将这 x 个基因遗传给下一代,从而确保子代中拥有足够多的最优解,并保持基因序列的优良性;剩下的 $k-x$ 个个体采用赌轮选择法^[17]进行选择。

交叉运算:本文采用随机选择置换的方法执行交叉运算。随机选择一个位置将两个染色体的基因进行交叉互换,若存在相同的编码,则随机换成不属于该染色体的其他基因编码。

变异运算:为了保持种群的多样性,避免陷入局部收敛和不成收敛,采用随机变异的方法,从当前群体中随机选择一个个体作为变异个体,在保持染色体基因各不相同的情况下将其中的部分基因随机地变成其它基因。

3.4 概率转发函数

分簇算法中数据首先要汇聚到簇头,簇头对簇内数据进行融合去噪等处理后,通过簇头与基站间的路由将数据传输到基站。由于簇头节点数量通常只占网络节点数量很小的比例,因此簇头节点的能量消耗比普通节点要快,如何选择簇头与基站的通信方式直接影响到簇头的生命周期。如果采用簇头与基站直接通信的方式,那么将导致距离基站较远的簇过早死亡;而如果采用多跳的方式与基站通信,则会导致距离基站较近的簇过早死亡。为了进一步提高网络能量均衡性能,采用概率转发原则来确定簇头与基站的通信方式。

簇头首先依据自身位置与剩余能量计算一个阈值 Th ,然后在区间 $[0, 1]$ 产生一个随机数 P ,如果 P 大于 Th 则簇头与基站直接通信,如果 P 小于或等于 Th 则进行多跳转发。当簇头节点距离基站距离较远,且能量较低时,应尽可能地选择多跳的方式来降低当前簇头节点的通信能耗;当簇头节点距离基站较近,且能量相对充足时,则应尽可能地采用直接与基站通信的方式来提高数据交付的速率。因此得出, Th 与当前簇头节点到基站距离成正比,与剩余能量成反比。按照经典 LEACH 协议的阈值构造思路,设定 Th 值为:

$$Th = \begin{cases} (D_i - R) \times \frac{E - E_i}{E} \\ D_{max} - R, & D_i > R \\ 0, & D_i \leq R \end{cases} \quad (6)$$

其中, E 为节点的初始能量, D_i 为当前簇头与基站的距离, D_{max} 为网络中节点到基站的最远距离, R 为节点的通信半径。当 D_i 较大且 E_i 偏小时, Th 较大, P 大于 Th 的概率较小,这时可以选择多跳通信,以降低节点能耗,延长网络寿命;当 D_i 较小且 E_i 较大时, Th 较小, P 大于 Th 的概率大,此时节点可以选择直接通信来均衡簇间的能量消耗。当簇头节点到基站的距离小于 R 时, Th 为 0,所有此类簇头节点将采用直接通信的方式与基站进行数据交互,因此靠近基站的簇头节点相对于远离基站的簇头节点能量消耗快,这是多跳传感网络所无法避免的问题。本文提出的利用遗传算法对簇头组合进行最优选择的过程,确保了簇头节点不会长期位于基站附近,从而也从一定程度上均衡了网络的能量消耗。

4 仿真结果及分析

本文使用 MATLAB 仿真,900 个节点随机部署在 $300\text{m} \times 300\text{m}$ 的区域内,基站位于 $(150, 150)$,仿真实验参数如表 1 所示,LEACH-GPF 性能统计如图 1、图 2 所示。

表 1 仿真参数设置

实验参数	参数值
自由空间传输损耗	10^{-11} (J)
多路径衰减传输损耗	1.3×10^{-15} (J)
处理 1 个单位数据 CPU 耗能	5×10^{-9} (J)
发射机耗能	5×10^{-8} (J)
接收机耗能	5×10^{-8} (J)
E	0.25 (J)
D_{max}	87m
α_1, α_2	1.8, 2.4
$\beta_1, \beta_2, \beta_3$	0.42, 0.27, 0.31

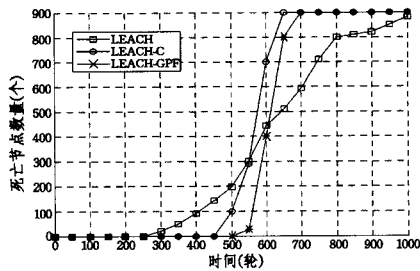


图 1 生命周期

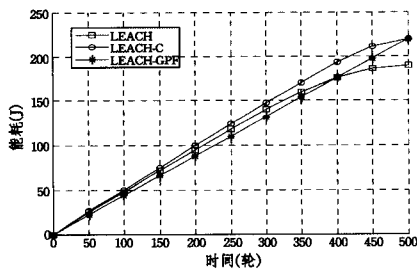


图 2 网络能耗

1) 生命周期对比分析。如图 1 所示, LEACH 算法首先出现第一个死亡节点, 接着是 LEACH-C, 最后是 LEACH-GPF, 其中 LEACH 出现节点死亡的时间明显早于 LEACH-C 与本文算法。另一方面, 随着网络的运行, LEACH-C 与 LEACH-GPF 出现了较大衰减, 当网络死亡时, 节点死亡数量接近 100%, 虽然 LEACH 算法运行的时间比另外两种算法的时间稍长, 但是当网络死亡后, 还存有少量存活节点, 这表明 LEACH 算法过度依赖部分节点, 导致部分节点能耗明显低于其他节点, 并保持存活状态, 但此时网络已经死亡, 网络能量没有得到有效的利用。相比而言, 本文提出的算法能够在较长时间内将网络维持在较高能量水平, 并充分地利用网络的能量, 具有较好的能耗均衡能力, 更适应于实际网络。

2) 网络能耗分析。如图 2 所示, LEACH-GPF 算法在网络的运行周期中, 能量消耗比较稳定, 在初始节点, LEACH-GPF 能耗比其他两种协议略低。随着网络的运行, LEACH 以及 LEACH-C 协议能耗逐渐降低, 最终趋于平缓。这表明 LEACH-GPF 算法整体性能稳定, 到了网络后期网络仍然能够进行通信; 而 LEACH 协议与 LEACH-C 协议由于节点能量消耗不均衡出现了网络空洞, 导致网络失去功能中断数据通信, 这也解释了为什么 LEACH 协议在网络死亡后还存在部分存活节点的现象。以上说明了本文算法具有较高的能耗负载均衡能力。

结束语 本文在遗传算法的基础上结合概率转发方法对 LEACH 协议进行改进。综合考虑了节点剩余能量、到基站

距离以及节点的通信代价 3 种因素来选择最优的簇头组合, 并通过概率转发准则来确定簇头与基站的通信方式, 从簇头层面进一步提高了网络的能耗均衡能力。实验数据表明, LEACH-GPF 算法能够有效延迟第一个死亡节点出现的时间, 并充分利用网络节点的能量, 均衡网络的整体能耗。另一方面, 作为一种集中式分簇算法, LEACH-GPF 算法依然存在依赖网络全局信息的缺陷, 如何降低这种依赖以及算法的复杂度是后续研究的问题。

参考文献

- [1] 沈波, 张世永, 钟亦平. 无线传感器网络分簇路由协议[J]. 软件学报, 2006, 17(7): 1588-1600
- [2] 何世钧, 代岩岩, 周汝雁, 等. 一种基于遗传聚类的无线传感器网络分簇算法[J]. 传感器与微系统, 2006, 31(11)
- [3] 李朋飞, 李志华, 尹熙, 等. 基于能量等级的分簇拓扑控制算法[J]. 计算机科学, 2014, 41(3): 96-99
- [4] 蒋畅江, 石为人, 唐贤伦, 等. 能量均衡的无线传感器网络非均匀分簇路由协议[J]. 软件学报, 2012, 23(5): 1222-1232
- [5] 卢先领, 王莹莹, 王洪斌, 等. 无线传感器网络能量均衡的非均匀分簇算法[J]. 计算机科学, 2013, 40(5)
- [6] Wu Xiao-ling, Wang Yang-yang, Liu Guang-cong, et al. Energy-Efficient Routing Algorithms Based on OVFS Code and Priority in Clustered Wireless Sensor Networks[J]. International Journal of Distributed Sensor Networks, 2013, 2013: 1-8
- [7] Heinzelman W R. Energy-Efficient Communication Protocol for Wireless Microsensor Networks[C]// Proceedings of the 33rd Annual Hawaii International Conference on System Sciences. 2000
- [8] 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)
- [9] Sharma M, Sharma K. An Energy Efficient Extended LEACH (EEE LEACH) [C]// 2012 International Conference on Communication Systems and Network Technologies (CSNT). 2012
- [10] 刘爱东, 卢中武, 刘德浩. 基于 LEACH 的低能耗路由协议研究[J]. 计算机工程与应用, 2012, 48(24): 88-90
- [11] Heinzelman W. Application-Specific protocol architectures for wireless networks [D]. Boston: Massachusetts Institute of Technology, 2000
- [12] Younis 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): 660-669
- [13] 顾云丽, 钱焕延, 徐昕, 等. 基于时延敏感无线传感器网络的最优任播算法[J]. 计算机科学, 2013, 40(1): 85-87, 102
- [14] 韩丽霞. 求解多目标优化问题的新遗传算法[J]. 计算机科学, 2013, 40(6): 64-66, 95
- [15] Wu Yin, Liu Wen-bo. Routing protocol based on genetic algorithm for energy harvesting-wireless sensor networks[J]. IET Wireless Sensor Systems, 2013, 3(2): 112-118
- [16] 李聪, 刘广聪, 史元杰. WSN 中基于能量均衡的路由算法研究[D]. 广州: 广东工业大学, 2013
- [17] Lipowski A. Roulette-wheel selection via stochastic acceptance [J]. Physica A: Statistical Mechanics and its Applications, 2012, 391(6): 2193-2196