

一种基于谱聚类算法的分簇路由协议

芮啟建¹ 韩冰青¹
RUI Qijian HAN Bingqing

摘要

针对经典 LEACH 协议随机选择簇首导致的极大簇和极小簇、簇首分布不均,从而导致网络生命周期较短的问题,提出一种基于谱聚类的分簇路由协议。在网络初始化时期,在基站计算确定最优簇首数目后,将谱聚类算法应用到节点的分簇中;分簇完成后,节点的分簇保持不变;簇首选择阶段,根据节点剩余能量因子、节点到基站的距离因子来竞争簇首;数据传输阶段,簇内节点通过单跳的方式将数据传送给簇首,簇首将数据融合后通过单跳的方式传送给基站。通过网络仿真实验,比较了网络生存周期、节点剩余总能量和基站接受数据包数量等参数,结果表明:相较于 LEACH 协议,新协议均衡了网络中节点能耗,延长了网络生存期。

关键词

LEACH 协议;分簇路由协议;谱聚类;簇首数;簇首选举

doi: 10.3969/j.issn.1672-9528.2024.03.009

0 引言

无线传感网络是由能够感知湿度、温度等信息的传感器节点和汇聚传感器节点数据包的基站组成的数据收集的网络系统^[1],传感器节点通过传感装置获取相应环境的相应信息后,通过无线通信方式将数据发送到基站,再由基站发送到数据中心对数据进行分析。与传统网络相比,无线传感网络中节点的部署具有随机性,节点的电池能量有限且难以进行电池充电和电池更换,因而节点的能耗问题是无线传感网络研究的重点^[2]。因此,充分利用节点能量,延长网络生存时间是无线传感网络主要的设计目标。

1 相关工作

无线传感器网络路由协议可以从不同的方面对路由协议进行分类。根据网络管理的体系结构,可以将路由协议分为平面路由协议和分簇路由协议。平面路由协议在小范围检测的场景下有着很大优势,但随着网络区域的不断变大,平面路由协议在传输时延和传输能耗方面显露出巨大的劣势^[3]。因此,分簇路由协议成为无线传感网络研究的重点。LEACH^[4-5]路由协议作为最早的分簇路由协议,通过循环的方式随机选择簇首,将整个网络的能量负载平均分配到每个传感器节点中,从而达到了降低网络能耗、提高网络整体生存时间的目的。文献[6]通过引入节点间的间距因子、节点剩余能量因子和节点密度因子来改进簇首选举公式,并动态调整因子权重来选择更优的簇首。文献[7-8]使用 K-means

算法对节点进行分簇,并分别使用了蚁群聚类算法和遗传算法优化了 K-means 对初始聚类中心的选择。文献[9]在选取簇首以及节点成簇阶段使用了 DBSCAN 的聚类算法,基于密度的方式来对节点成簇进行优化。文献[10]基于竞争机制选取簇头,通过模糊推理的方式来进行簇头的选取。文献[11]提出使用蝗虫优化算法确定簇头个数,并识别簇头,一旦簇首被识别,普通节点就被分配到他们最近的对应簇首。文献[12]提出一种基于人工蜂群算法的能量高效分簇路由协议。通过定义簇头的能量因子、位置因子和向心率因子设计适应度函数,选出最优簇首。文献[13]根据五个因素选取簇头,这些因素分别是剩余能量、位置适当性、节点度、紧密度和基站距离,同时采用最小化间隔距离控制簇头在网络中均匀分布。

针对经典 LEACH 协议的随机选择簇首导致的极大簇和极小簇、簇首分布不均,从而导致网络生命周期较短的问题,本文提出了一种基于谱聚类算法的分簇路由协议(简称 SCCR 协议)。

2 网络模型和能耗模型

2.1 网络模型

本协议的网络模型设置如下:(1)所有传感器节点的初始能量都是相同并且有限的,节点被随机部署在正方形区域中;(2)所有传感器节点都是同构的,具备数据融合的功能,数据融合后,数据包的大小是固定的;(3)传感器节点在传输数据时运行,在不传输数据时可以选择休眠;(4)根据信号的强度,接收节点可以近似计算出发送节点的距离;(5)

1. 南京审计大学 江苏南京 211815

基站具有计算能力且能量是无限的，基站被部署在正方形区域的正中心。

2.2 能耗模型

SCCR 协议采用和 LEACH 协议相同的无线电通信能量消耗模型。能量消耗主要来自数据发送与数据接收，数据发送能耗依据发送节点与接收节点之间的距离进行确定，当距离小于阈值 d_0 时，计算能耗采用自由空间模型；当距离大于等于阈值 d_0 时，计算能耗采用多径衰减模型。节点发送数据的能量消耗公式为：

$$E_{tx}(l, d) = E_{tx-elec}(l) + E_{tx-amp}(l, d) = \begin{cases} lE_{elec} + l\epsilon_{fs}d^2, & d < d_0 \\ lE_{elec} + l\epsilon_{mp}d^4, & d \geq d_0 \end{cases} \quad (1)$$

数据接收能耗主要依据消息的比特数，节点接收数据能量的消耗公式为：

$$E_{rx}(l) = E_{rx-elec}(l) = lE_{elec} \quad (2)$$

数据融合的能量主要依据消息的比特数，节点融合数据能量的消耗公式为：

$$E_d = E_{da-elec}(l) = lE_{da} \quad (3)$$

式中： l 为传输数据包比特数， E_{elec} 为接收和发送单位比特能耗系数， ϵ_{fs} 为自由空间模型能耗系数， ϵ_{mp} 为多径衰减模型能耗系数， d 为节点之间的距离， E_{da} 为数据融合能耗系数， $\sqrt{\frac{\epsilon_{fs}}{\epsilon_{mp}}}$ 是决定数据传输采用自由空间模型还是多径衰减模型的距离阈值。

3 SCCR 协议

本协议首先确定最佳分簇数目，然后根据节点位置信息将节点建模成无向图中的顶点，基于最佳分簇数目使用谱聚类算法将节点进行分簇，并将节点的分簇形式固定下来。数据的传输是按照轮的方式进行，每轮数据传输前，节点将节点剩余能量等信息发送给基站，基站以节点剩余能量、节点到基站的距离作为因子，计算得出每一个节点竞争簇首的竞争权重，权重大的节点成为簇首。簇首选定后，普通节点通过单跳的方式将数据传送给簇首，簇首将数据进行融合后，通过单跳的方式将数据传送给基站完成数据的传输。

3.1 确定最优簇首数

最优簇首数根据 LEACH 协议确定，计算最优簇首数公式为：

$$k_{opt} = \sqrt{\frac{N}{2\pi}} \times \sqrt{\frac{\epsilon_{fs}}{\epsilon_{mp}}} \times \frac{M}{d_{toBS}^2} \quad (4)$$

式中： N 为节点个数， M 为方形区域边长， d_{toBS} 为簇头到基站之间的距离（本协议中基站在正方形区域的正中心，因而设定 $1\text{ m} \leq d_{toBS} \leq \sqrt{2}/2 M\text{ m}$ ）。

3.2 基于谱聚类算法进行分簇

网络初始化时期，所有节点将自己的位置信息以及能量信息通过控制数据包发送给基站。基站根据每个节点传送的位置信息，对于方形网络中的每个节点，将其看作无向图上的一个顶点，构造图 $G = (V, E)$ 。

在图 G 中，可以使用邻接矩阵来表示节点之间的距离。在谱聚类中，引入相似度的概念来表征节点之间的距离。 W_{ij} 表示节点 i 与节点 j 之间的相似度，其中 $W_{ij} = W_{ji}$ 。当两个节点之间的距离越近，则相似度越高，节点更倾向于被分到一个簇中；当两个节点之间的距离越远，则相似度越低，节点更倾向于被分到不同的簇中。本协议使用全连接法以及高斯核函数来计算节点之间的相似度，计算点 i 和点 j 之间的相似度公式为：

$$W_{ij} = W_{ji} = \sum_{i=1, j=1}^n \exp \frac{-\|x_i - x_j\|^2}{2\sigma^2} \quad (5)$$

根据节点相似度，类比于数据结构中节点的度定义为与该点直接相连的顶点个数，谱聚类中定义节点的度为该节点与其他节点相似度之和，即 $d_i = \sum_{j=1}^n W_{ij}$ 。

根据已经建模好的无向图，基于前文确定的最优簇首数可以将无向图切割成 k_{opt} 个独立子图，假设每个子图点的集合为 $A_1, A_2, \dots, A_{k_{opt}}$ ，则有 $A_i \cap A_j = \emptyset$ 且 $A_1 \cup A_2 \cup \dots \cup A_{k_{opt}} = V$ 。

对于任意两个子图点集合 $A, B \subset V$ ， $A \cap B = \emptyset$ ，定义 A 和 B 之间的切图权重为：

$$W(A, B) = \sum_{i \in A, j \in B} w_{ij} \quad (6)$$

使用 Ncut 方式对无向图进行割图，则对于 k_{opt} 个子图集合 $A_1, A_2, \dots, A_{k_{opt}}$ ，定义 $vol(A_i)$ 为集合 A_i 中所有节点的度之和，定义切图损失函数为：

$$NCut(A_1, A_2, \dots, A_{k_{opt}}) = \frac{1}{2} \sum_{i=1}^k \frac{W(A_i, \bar{A}_i)}{vol(A_i)} \quad (7)$$

使得切图损失函数 $NCut(A_1, A_2, \dots, A_{k_{opt}})$ 最小化的 $A_1, A_2, \dots, A_{k_{opt}}$ 就是最佳的分簇方式。基站确定好分簇方式后，节点的分簇就固定下来了，后续就不再进行分簇了。通过一次分簇固定下来节点的分簇方式，解决了 LEACH 协议中每轮数据传输前节点进行分布式分簇带来的控制数据包传输的能量消耗问题。

3.3 簇首选择

在网络初始化阶段，节点将位置信息传送给基站，基站基于各个节点位置信息进行分簇并将分簇信息和节点位置保存下来。因而在每轮数据传输过程前，每个节点只需将自身的能量信息发送给基站。

在 LEACH 协议中，每个节点在 $0 \sim 1$ 中选择一个随机数，如果随机数小于阈值 $T(n)$ ，则节点自举为簇首，阈值 $T(n)$ 的计算公式为：

$$T(n) = \begin{cases} \frac{P}{1-P[rmod(1/P)]}, & n \in G \\ 0, & otherwise \end{cases} \quad (8)$$

式中： P 值为网络中簇首节点占总节点数量的百分比，这个值在网络初始化时期设置。

LEACH 协议中，随着网络的运行，没有当选过簇首的节点成为簇首的概率增加。当轮数为 $1/P-1$ 时，当前 $1/P$ 轮没有当选过簇首的节点以概率 1 成为簇首节点。在本文中， P 为 0.1。因而，以 10 轮为一个周期，周期中每个节点有且仅有一次机会当选为簇首。该簇首选举方式没有考虑到节点的位置以及节点的能量信息，会导致特殊位置的节点能量消耗巨大，快速死亡。

针对 LEACH 协议簇首选择出现的问题，提出新的簇首选择方式。在每轮簇首选择过程中，计算出每个节点竞争簇首节点的竞争权重，竞争权重大的节点当选为簇首节点，计算的过程由基站进行。基站为每个节点计算权重的公式为：

$$T(i) = \begin{cases} A \frac{E_i}{E_{avg_c(i)}} + B \frac{d_{avg_c(i)}}{d_i}, & \text{节点存活} \\ 0, & \text{节点死亡} \end{cases} \quad (9)$$

式中： E_i 为节点的剩余能量， $E_{avg_c(i)}$ 为节点所在簇的所有存活节点剩余总能量，可以看到，在其他条件不变的情况下，当节点的剩余能量越大时，节点的竞争权重越高，节点竞争成为簇首的概率则越大。 d_i 为节点到基站的距离， $d_{avg_c(i)}$ 为节点所在簇的所有存活节点到基站的距离总和，可以看到，在其他条件不变的情况下，当节点到基站的距离越小时，节点的竞争权重越高。 A 为剩余能量控制因子、 B 为节点到基站距离控制因子，且 $A+B=1$ 。在网络运行初期，节点能量都较大，可以适当减少剩余能量因子，适当增大距离因子，这样靠近基站的节点竞争权重变大，成为簇首的概率变高；随着网络运行，节点的剩余能量变少，此时适当增大剩余能量因子，适当减少距离因子，这样剩余能量较多的节点竞争权重变大，成为簇首的概率变高。在本协议中，当簇中剩余总能量大于初始簇中总能量的一半时，则这个簇中计算节点权重时取 $A=0.4$ ， $B=0.6$ ；当簇中剩余总能量小于初始簇中总能量的一半时，则这个簇中计算节点权重时取 $A=0.6$ ， $B=0.4$ 。

3.4 数据传输

数据传输按照轮的方式进行，每一轮数据传输时，分为簇内数据传输和簇间数据传输。簇内数据传输时，普通节点根据基站传送控制包中的信息在自己的时隙内通过单跳的方式将数据传送给簇首节点，在非自己时隙时进行节点休眠来节约能量消耗。簇首每收到一个来自普通节点的数据包时，进行一次数据融合，因而每个簇内簇首节点将进行簇内节点数减 1 次数据融合操作。簇间数据传输在簇内数据完成后进行，簇首节点使用单跳的方式传输数据包到基站，基站接收

完所有簇首节点的数据包后，代表着一轮数据传输的完成。

4 仿真实验以及性能分析

4.1 参数设置

本文使用 MATLABR2021b 对 LEACH 和 SCCR 协议进行了仿真对比，参数设置见表 1。

表 1 仿真参数设置

参数	取值
正方形区域边长	200 m
网络中节点个数	50 个
节点初始能量	0.5 J
基站坐标	(100 m, 100 m)
数据包大小	4000 bit
控制包大小	100 bit
ϵ_{fs}	10 pJ/(bit·m ²)
ϵ_{mp}	0.0013 pJ/(bit·m ⁴)
E_{da}	5 nJ/bit
E_{elec}	50 nJ/bit
k_{opt}	5 个
σ	0.1

4.2 结果分析

图 1 给出了 LEACH 协议与 SCCR 协议的网络生存周期比较，其中网络运行轮数代表时间。LEACH 协议第一个死亡节点出现在 419 轮，SCCR 协议第一个死亡节点出现在 766 轮，SCCR 协议提升了 82%；LEACH 协议中最后一个节点死亡出现在 788 轮，SCCR 协议中最后一个节点死亡出现在 1111 轮，SCCR 协议提升了 40%。以 70% 节点存活作为评估无线传感网络生存周期的指标时，LEACH 协议存活了 637 轮，SCCR 协议存活了 999 轮，SCCR 协议提升了 56%。

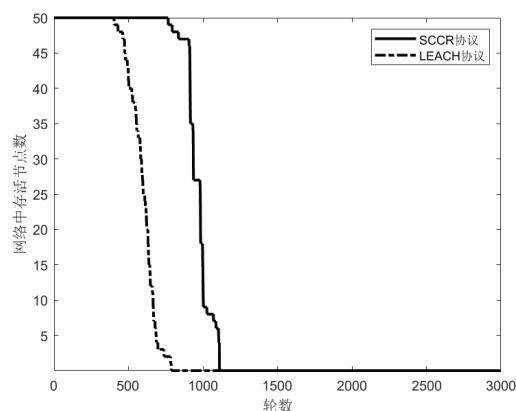


图 1 存活节点对比图

图 2 给出了 LEACH 协议与 SCCR 协议的传输给基站的数据包数量对比，在 LEACH 协议中，基站收到的数据包数量为 3631；在 SCCR 协议中，基站收到的数据包数量为 4952，SCCR 协议提升了 36%。

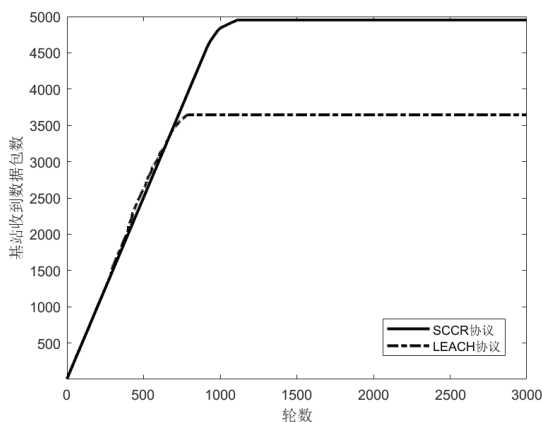


图2 基站接收数据包对比图

图3给出了LEACH协议与SCCR协议网络中所有节点剩余能量的对比,SCCR协议优于LEACH协议,可以看出SCCR协议的设计大幅度减少了网络中节点的能量消耗。

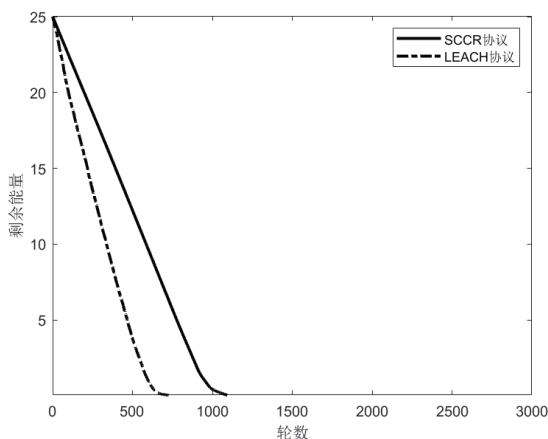


图3 节点剩余能量对比图

基于网络生存周期、基站接收到的数据包和所有节点剩余能量参数的比较,可以看出SCCR协议大幅度延长了网络生存周期,均衡了传感器节点的能量消耗,相较于LEACH协议,更加高效地利用了网络中节点的能量来进行有效数据的传输。

5 结语

本文在研究LEACH协议的基础上,分析了LEACH协议分布式竞选簇首相互发送控制数据包带来的大量能量消耗的缺点,簇头选举公式导致的分簇不均缺点,提出了一种基于谱聚类算法的分簇路由协议。通过一次分簇并固定分簇的方式减少了节点竞选簇首相互传送控制数据包带来的能量消耗,提出新的簇首选举方式并通过引入节点到基站距离因子和节点剩余能量因子,均衡了各个簇中节点的能量消耗。在MATLAB软件下的仿真结果表明,本文提出的SCCR协议相较于LEACH协议,均衡了节点的能量消耗,大幅度延长了网络生命周期。

参考文献:

- [1] 侯佩. 基于蚁群算法的LEACH协议改进研究[J]. 电脑与电信, 2022(6):48-52.
- [2] NAYAK P, REDDY C P. Algorithm for energy consumption minimisation in wireless sensor network[J]. IET communications, 2020, 14(8):1301-1310.
- [3] 时浩. 基于LEACH的分簇优化及多跳传输算法[J]. 物联网技术, 2023, 13(7):46-50.
- [4] LATIF K, JAVAID N, SAQIB M N, et al. Energy-efficient communication protocol for wireless microsensor networks[J]. International journal of AdHoc & ubiquitous computing, 2016, 21(2):130-139.
- [5] HEINZELMAN W B, CHANDRAKASAN A P, BALAKRISHNAN H. An application-specific protocol architecture for wireless microsensor networks[J]. IEEE transactions on wireless communications, 2002, 1(4):660-670.
- [6] 黄利晓, 王晖, 袁利永, 等. 基于能量均衡高效WSN的LEACH协议改进算法[J]. 通信学报, 2017, 38(S2): 164-169.
- [7] 贾惠丽, 范训礼, 吕艳峰. 能量高效的无线传感器网络分簇路由协议[J]. 传感器与微系统, 2018, 37(8):116-119.
- [8] 胡鑫. 基于遗传K-means算法的无线传感器网络路由协议研究[D]. 秦皇岛:燕山大学, 2020.
- [9] 颜宇鸿. 无线传感器网络能量有效的数据收集机制研究[D]. 南京:南京邮电大学, 2021.
- [10] PAWAN S M, MOHAMMAD N D, BASHIR A. Fuzzy based enhanced cluster head selection (FBECS) for WSN[J]. Journal of King Saud University Science, 2020, 32:390-401.
- [11] RODRÍGUEZ A, PÉREZCISNEROS M, ROSASCARO J C, et al. Robust clustering routing method for wireless sensor networks considering the locust search scheme[J]. Energies, 2021, 14:1-19.
- [12] 王宗山, 丁洪伟, 李佳. WSN中基于人工蜂群算法的分簇路由协议[J]. 微电子学与计算机, 2021, 38(4):74-80.
- [13] KHAN A, KARMAKAR G C, KAMRUZZAMAN J. Energy-efficient fuzzy-logic-based clustering technique for hierarchical routing protocols in wireless sensor networks[J]. Sensors, 2019, 19:561-570.

【作者简介】

芮啟建(1997—), 通信作者(email:1050680272@qq.com), 男, 江苏南京人, 硕士研究生, 研究方向: 无线传感网、无线自组网。

韩冰青(1979—), 男, 博士, 副教授, 研究方向: 无线自组网、智能计算等。

(收稿日期: 2023-12-13)