



文章编号:1001-9081(2019)08-2359-07

doi:10.11772/j.issn.1001-9081.2019010134

基于遗传算法和模糊 C 均值聚类的 WSN 分簇路由算法

董发志¹, 丁洪伟^{1*}, 杨志军^{1,2}, 熊成彪¹, 张颖婕¹

(1. 云南大学 信息学院, 昆明 650500; 2. 云南省教育科学研究院, 昆明 650223)

(* 通信作者电子邮箱 dhw1964@163.com)

摘要:针对无线传感器网络(WSN)的节点能量有限、生命周期短、吞吐量低等问题,提出一种基于遗传算法(GA)和模糊C均值(FCM)聚类的WSN分簇路由算法GAFCMCR,采取“集中分簇,分布簇头选举”的方式。网络初始化时基站采用由GA优化的FCM聚类算法形成网络分簇。第一轮簇头由距簇中心最近的节点担任;从第二轮开始,簇头的选举由上一轮的簇头负责,选举过程综合考虑候选节点的剩余能量、与基站的距离、与簇内其他节点的平均距离三个因子,并根据网络状态实时调整三个因子的权重。在数据传输阶段,将轮询机制引入簇内通信。仿真结果表明,相同网络环境下,与LEACH算法和基于K-Means的均匀分簇路由(KUCR)算法相比,GAFCMCR将网络生命周期延长了105%和20%。GAFCMCR成簇效果良好,具有良好的能量均衡性和更高的吞吐量。

关键词:无线传感器网络;模糊C均值聚类;遗传算法;均匀分簇;轮询机制

中图分类号: TP393.04 **文献标志码:**A

WSN clustering routing algorithm based on genetic algorithm and fuzzy C-means clustering

DONG Fazhi¹, DING Hongwei^{1*}, YANG Zhijun^{1,2}, XIONG Chengbiao¹, ZHANG Yingjie¹

(1. School of Information Science and Engineering, Yunnan University, Kunming Yunnan 650500, China;

2. Yunnan Academy of Educational Sciences, Kunming Yunnan 650223, China)

Abstract: Aiming at the problems of limited energy of nodes, short life cycle and low throughput of Wireless Sensor Network (WSN), a WSN Clustering Routing algorithm based on Genetic Algorithm (GA) and Fuzzy C-Means (FCM) clustering (GAFCMCR) was proposed, which adopted the method of centralized clustering and distributed cluster head election. Network clustering was performed by the base station using a FCM clustering algorithm optimized by GA during network initialization. The cluster head of the first round was the node closest to the center of the cluster. From the second round, the election of the cluster head was carried out by the cluster head of the previous round. The residual energy of candidate node, the distance from the node to the base station, and the mean distance from the node to other nodes in the cluster were considered in the election process, and the weights of these three factors were real-time adjusted according to network status. In the data transfer phase, the polling mechanism was introduced into intra-cluster communication. The simulation results show that, compared with the LEACH (Low Energy Adaptive Clustering Hierarchy) algorithm and the K-means-based Uniform Clustering Routing (KUCR) algorithm, the life cycle of the network in GAFCMCR is prolonged by 105% and 20% respectively. GAFCMCR has good clustering effect, good energy balance and higher throughput.

Key words: Wireless Sensor Network (WSN); Fuzzy C-Means (FCM) clustering; Genetic Algorithm (GA); uniform clustering; polling mechanism

0 引言

无线传感器网络(Wireless Sensor Network, WSN)由大量廉价的微传感器节点组成,可以部署在监控区域,并通过无线通信方式形成自组织网络。传感器节点可以感知网络,收集和处理数据并将信息传输给监测中心^[1]。近年来WSN广泛应用于监测、跟踪、事件检测、监视和灾害管理等各方面^[2]。WSN的长期有用性主要依赖于节点的生命周期。节点的寿命完全取决于其电池电量,由于这些电池几乎不可更换,所以通过节约电池能量可以实现无线传感器网络生命周期

的延长^[3]。针对这个问题,许多层次型路由协议被提出,通过控制成簇、优化网络的拓扑结构达到降低网络能耗、延长网络生命周期的目的^[4]。

最经典的分簇路由协议是由Heinzelman等^[5]提出的LEACH(Low Energy Adaptive Clustering Hierarchy)算法,它将无线传感器网络划分为若干簇,每个簇中随机选取至少一个簇头,簇内节点将自己收集的数据发送给簇头,簇头进行数据融合之后通过单跳或者多跳将数据发送给基站。采用“轮”的方式将能量损耗均匀分布到各个节点,从而延长了生命周期;但是簇头的选取没有考虑节点的剩余能量,若能量低的节

收稿日期:2019-01-21;修回日期:2019-04-02;录用日期:2019-04-03。 基金项目:国家自然科学基金资助项目(61461053, 61072079)。

作者简介:董发志(1993—),男,云南楚雄人,硕士研究生,主要研究方向:无线传感器网络、轮询多址通信; 丁洪伟(1964—),男,江西于都人,教授,博士,主要研究方向:轮询多址通信、随机多址通信、通信与信息系统; 杨志军(1968—),男,云南保山人,高级工程师,博士,主要研究方向:计算机通信与网络、轮询系统; 熊成彪(1995—),男,云南楚雄人,硕士研究生,主要研究方向:随机多址通信系统; 张颖婕(1993—),女,云南昆明人,硕士研究生,主要研究方向:机器学习。



点被选为簇头，则会加快节点的死亡。Heinzelman 等^[6]在 LEACH 的基础上又提出了 LEACH-C (LEACH-centralized) 算法，由基站集中计算分簇，减少了簇头的计算能耗，但是没有考虑簇头节点的位置。Kumar 等^[7]提出了 U-LEACH (Universal-Low Energy Adaptive Cluster Hierarchy) 算法，利用不均匀成簇解决了 LEACH 协议中的热点区域问题。Bakaraniya 等^[8]提出了基于 K-means 分簇的 K-LEACH (Kmedoids-LEACH protocol) 算法，初始阶段利用 K-means 进行分簇，提升了网络性能，但簇头的选举未考虑节点能量。张雅琼^[9]提出的基于 K-Means 的均匀分簇路由 (Uniform Clustering Routing based on K-Means, KUCR) 算法利用 K-Means 进行分簇，采取一次成簇、多次簇头更新的机制，提升了性能，但簇头选举时并未考虑节点与其他节点的距离关系。严静静等^[10]利用粗糙 C-均值聚类改进了 LEACH 算法，利用粗糙 C-均值聚类进行分簇，保证了簇头节点分布均匀，延长了生命周期，但在簇头选举时未考虑节点与基站的距离。

本文在研究 LEACH、KUCR 算法的基础上，提出了一种基于遗传算法 (Genetic Algorithm, GA) 和模糊 C 均值 (Fuzzy C-Means, FCM) 的 WSN 分簇算法 (Clustering Routing algorithm based on Genetic Algorithm and fuzzy C-means clustering, GAFCMCR)。该算法的主要特点是：在网络初始化阶段，由基站采用 GA 改进的 FCM 算法进行集中式分簇；然后，每个簇内节点根据自身的剩余能量、到基站的距离以及与簇内其他节点的距离三个因子进行簇头的竞选；最后，在数据的传输阶段，簇内通信引入轮询控制，簇间采用载波侦听多路访问 (Carrier Sense Multiple Access, CSMA) 机制。仿真结果表明，相比 LEACH 和 KUCR，GAFCMCR 能够有效延长网络生命周期，降低能耗，提高网络的吞吐量。

1 系统模型

1.1 网络模型

本文对无线传感器网络作出如下假设：网络中有 N 个节点随机分布在面积为 $M \times M$ 的监测区域中，节点部署之后不随时间移动；节点的初始能量相同，可以获取自身的位置和剩余能量，且都可以担任簇头，能够进行计算处理和数据融合，具有唯一地址；基站位于监测区域中心，能量和计算能力无限；所有节点都可以和基站直接通信。

1.2 能耗模型

本文采用与文献[11]相同的一阶无线电能耗模型，节点发送 l bit 的数据消耗的能量包括发射损耗和功率放大损耗两部分，一个节点经过距离 d 发送 l bit 数据的能耗如下：

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

接收消耗的能量和距离无关，接收 l bit 数据消耗的能量如下：

$$E_{RX} = lE_{elec} \quad (2)$$

式中： E_{elec} 为节点每发送或接收 1 bit 数据消耗的能量； d_0 为节点之间传输数据的模型选择的一个阈值，取值为 $d_0 = \sqrt{\varepsilon_{fs}/\varepsilon_{mp}}$ 。当传输距离小于 d_0 时，采用自由空间信道模型；当

传输距离大于 d_0 时，采用多路径衰减模型^[12-14]。其中相应的参数 ε_{fs} 和 ε_{mp} 表示的是上述两种情况下单位功率放大时需要的能量。

2 GAFCMCR

GAFCMCR 采用与 LEACH 相同的按“轮”运行方式，如图 1 所示，分为簇的形成、簇头选举和稳定传输三个过程。在初始化阶段各节点将位置信息发送给基站，由基站采用优化的 FCM 算法进行迭代计算，形成 C 个簇。不同于 LEACH 算法每轮都需要重新分簇，GAFCMCR 的第一轮分簇完成后，以后的轮次只在簇内进行簇头的更新，只在最优簇数发生改变时才重新分簇。这样有的轮次簇结构和簇头都是不变的，减少了频繁分簇和簇头选举带来的能量损耗，能延长生命周期。

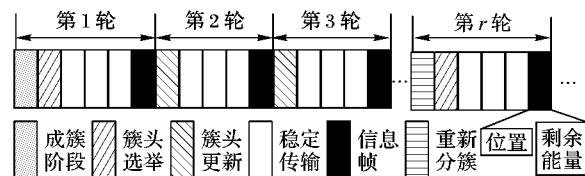


图 1 协议运行轮次
Fig. 1 Protocol running round

2.1 簇的形成

网络初始化时，基站首先为每个节点分配唯一的通信 ID，各节点将自身的位置坐标发送给基站，基站经过计算，将分簇信息广播给节点。广播信息包括簇 ID、节点与基站的距离。

2.1.1 最优簇头数

FCM 算法需指定聚类数目，在无线传感器网络中，为了得到合理的分簇，根据网络特征确定最优簇头数。对于本文的网络模型和能耗模型，基站位于监测区域中心，所以每个节点到基站的距离都小于等于 d_0 ，故簇头和簇内节点运行一轮消耗的能量^[15]为：

$$E_{CH} = l \left[E_{elec} \left(\frac{N}{c} - 1 \right) + E_{da} \frac{N}{c} + E_{elec} + \varepsilon_{fs} d_{toBS}^2 \right] \quad (3)$$

$$E_{noCH} = l(E_{elec} + \varepsilon_{fs} d_{toCH}^2) \quad (4)$$

式中： c 为簇头数， E_{da} 为融合 1 bit 数据消耗的能量。则网络的总能耗为：

$$E_{total} = l[2NE_{elec} + NE_{da} + c\varepsilon_{fs} d_{toBS}^4 + N\varepsilon_{fs} d_{toCH}^2] \quad (5)$$

要使得总能耗最小，则

$$\frac{\partial E_{total}}{\partial c} = 0 \quad (6)$$

d_{toBS} 为节点到基站的距离，计算如下：

$$E[d_{toBS}] = \int_A \sqrt{x^2 + y^2} \frac{1}{A} dA = 0.765 \frac{m}{2} \quad (7)$$

d_{toCH} 为节点到簇头的距离，计算如下：

$$E[d_{toCH}^2] = \iint (x^2 + y^2) \rho(x, y) dx dy = \rho \int_0^{2\pi} \int_0^{\frac{m}{\sqrt{\pi c}}} r^3 dr d\theta = \frac{m^2}{2\pi c} \quad (8)$$

综上可得最优簇头数：

$$c = \sqrt{\frac{N}{2\pi}} \frac{m}{d_{toBS}} = \sqrt{\frac{N}{2\pi}} \frac{2}{0.765} \quad (9)$$



2.1.2 模糊 C 均值聚类

FCM 算法首先是由 Ruspini^[16] 提出来的, 目前在机器学习中应用广泛, 将 FCM 应用到无线器网络的分簇中, 能够将网络中的节点按照紧密程度分为固定的几个簇, 有效减少了由于簇结构不均匀带来的能量损耗。FCM 算法的目标函数为:

$$J_m = \sum_{i=1}^N \sum_{j=1}^c (\mu_{ij})^m (d_{ij})^2; m \in [1, \infty) \quad (10)$$

式中: $d_{ij} = \|x_i - c_j\|$, x_i 为第 i 个样本, c_j 为第 j 个聚类中心; μ_{ij} 是样本 x_i 对第 j 个聚类中心的隶属度; m 是模糊因子。

μ_{ij} 的更新公式为:

$$\mu_{ij} = \frac{1}{\sum_{k=1}^c \left(\frac{d_{ik}}{d_{jk}}\right)^{\frac{2}{m-1}}} \quad (11)$$

聚类中心更新的公式为:

$$c_j = \frac{\sum_{i=1}^N (\mu_{ij})^m x_i}{\sum_{i=1}^N (\mu_{ij})^m} \quad (12)$$

算法终止条件为: $\|\mathbf{U}(k+1) - \mathbf{U}(k)\| \leq \varepsilon$ 。

FCM 算法流程如图 2 所示。

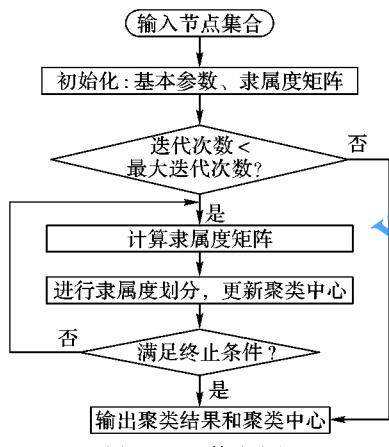


Fig. 2 Flow chart of FCM algorithm

2.1.3 遗传算法优化初始聚类中心

FCM 算法对初始聚类中心敏感, 容易收敛于局部最优解, 为了克服该缺点, 将 GA 应用于 FCM 算法的优化计算中, 由 GA 得到初始聚类中心, 再使用标准的 FCM 算法得到最终的分类结果。

遗传算法源于生物进化论和遗传学, 是一种模仿自然界适者生存机制的自适应搜索算法^[17-18]。GA 优化 FCM 算法的步骤^[19-20] 为:

步骤 1 参数初始化, 设置种群规模 N , 交叉概率 p_c , 变异概率 p_m , 终止进化的迭代次数 T 以及初始聚类数目 c 。采用二进制编码, 根据聚类中心坐标 (x_i, y_i) , 将其编码为基因串 $\alpha = \{\alpha_1 \alpha_2 \cdots \alpha_i \cdots \alpha_q\}$, 其中, $q = c \times 2$, 在 α 中每两个值代表一个聚类中心的坐标。

步骤 2 设置迭代次数 $t = 0$, 并随机选取初始种群。

步骤 3 计算个体适应值, 其适应函数 $f = 1/J_m$ 。

步骤 4 进行选择、交叉、变异, 获得新一代的种群, 判断新种群是否满足停止迭代准则, $t = t + 1$ 。

步骤 5 如果 t 大于最大迭代次数, 则停止迭代, 输出聚类个数和聚类中心, 否则返回步骤 2 继续迭代。

GA 优化 FCM 算法的流程如图 3 所示。

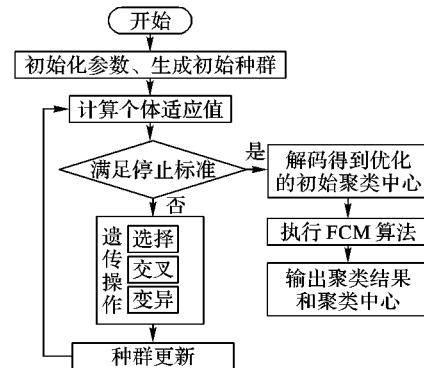


图 3 GA 优化 FCM 算法的流程

Fig. 3 Flow chart of FCM algorithm optimized by GA

GA 优化的 FCM 算法的伪代码如下:

输入 $X = \{x_1, x_2, \dots, x_N\}^T$ (样本集合), c (聚类数目), m (模糊因子), N (种群规模), p_c (交叉概率), 变异概率 p_m (交叉概率), T (迭代次数);

输出 聚类结果, 聚类中心。

```

begin
    t <- 0;                                // 初始化种群
    Initialize Pop(t);
    计算初始化种群个体适应值  $f = 1/J_m$ ;
    while  $t < T$  do
        GA-Operation Pop(t);                // 遗传算法运算
        计算个体适应值  $f = 1/J_m$ ;
         $t++$ ;
    end while
    bestcenter;                                // 最优初始聚类中心
    repeat
        for  $i = 1$  to  $n$  do
            for  $j = 1$  to  $c$  do
                根据式(11)计算价值函数
                根据式(12)计算隶属度矩阵  $U$ 
                根据式(13)更新聚类中心
            end for
        end for
        until
         $\|\mathbf{U}(k+1) - \mathbf{U}(k)\| \leq \varepsilon$ 
        return 聚类结果和聚类中心
    end

```

2.1.4 动态分簇

随着网络的运行, 存活节点数不断减少, 网络特征发生了变化, 初始化阶段确定的最优簇数不再适合当前网络状态。所以 GAFCMCR 在每一轮运行结束时, 基站会根据式(9)计算最优簇数, 当最优簇数改变时, 就重新分簇, 直到簇数为 1。动态的分簇过程优化了性能, 延长了网络的生命周期。

2.2 簇头的选举

为了减少簇头的能耗, 避免节点过早死亡从而导致网络失效, 本文算法的簇头选举综合考虑节点的剩余能量、到基站



的距离以及与簇内其他节点的距离关系等三个因素。第一轮分簇完成后,基站广播分簇结果时将每一簇离簇中心最近的节点指定为簇头。

第二轮开始,簇头的选举综合考虑剩余能量、与基站的距离及它与簇内其他节点的距离等三个因素。

剩余能量是簇头选举的重要参考,GAFCMCR 采用节点的剩余能量与所属簇内所有节点平均剩余能量的比值作为参考标准,如式(13) :

$$f_1(i) = \frac{E_{\text{res}}(i)}{\frac{1}{n} \sum_{j=1}^n E_{\text{res}}(j)} \quad (13)$$

式中: n 为节点 i 所属簇的节点总数。 f_1 越大说明节点 i 的剩余能量越高。

式(14)表示候选节点到基站的距离因素:

$$f_2(i) = \frac{\frac{1}{n} \sum_{j=1}^n d_{\text{tobS}}(i)}{d_{\text{tobS}}(i)} \quad (14)$$

式中, $d_{\text{tobS}}(i)$ 为节点 i 到基站的距离。 f_2 越大,说明节点距离基站越近,簇头到基站传输数据的距离越短,能耗越低。

式(15)表示候选节点与簇内其他节点的距离关系:

$$f_3(i) = \frac{n}{\sum_{j=1}^n d(i, j)} \quad (15)$$

式中, $d(i, j)$ 为节点 i 到节点 j 的距离。 f_3 越大,说明相对于其他节点,候选节点能够使簇内通信的距离更小。

每一轮的最后一帧,节点将自身位置和剩余能量发送给簇头,簇头根据式(16)重新计算每个节点的参量 p 。节点剩余能量多、距离基站近且与其他节点紧密程度较好的节点能够成为候选簇头。

$$p(i) = \alpha f_1(i) + \beta f_2(i) + \gamma f_3(i) \quad (16)$$

其中, α, β, γ 为权重参数,用来调整三个因子的重要程度。初始化阶段, $\alpha = \beta = \gamma = 1/3$; 随着网络的运行,节点能量不断下降,成为影响网络生命周期的重要因子,所以权重应该随着网络的运行而变化。将 α, β, γ 按照下式进行更新:

$$\alpha = E_{\text{init}}/E_{\text{res}}(i) \quad (17)$$

$$\beta = \gamma = (1 - \alpha)/2 \quad (18)$$

运行过程中, α 动态增大,剩余能量大的节点当选簇头的概率更高,使得能量更加均衡。

簇头节点选取参量最大的节点 i 与自身参量进行比较:

$$p_{\text{CH}} \leq \lambda p_{\max}(i), \lambda \in (0, 1) \quad (19)$$

如果式(19)成立,则簇头通知节点 i 当选为新簇头,并交换成员信息,节点 i 广播成为簇头信息;若不成立,则本轮次不进行簇头的更新。由于不是每轮都需要更新簇头,从而节省了能耗,降低了时延。式(19)中, λ 为网络系数,影响网络的更新速度;取值太小,则个别节点会长时间担任簇头而导致能量消耗大,缩短了节点生命周期;取值太大,则需要频繁更新簇头,也导致能耗增加。

2.3 数据传输

LEACH 及其大多数改进算法的簇内通信机制均采用时分多址(Time Division Multiple Access, TDMA)机制^[21-22], 簇

头为每个成员节点划分一个时隙,节点在固定的时隙内传输数据。但在无线传感器网络中,并不是所有节点都一直处于活跃状态,有些节点在一定时期处于休眠状态,当需要采集数据时再唤醒。这就会出现空闲的时隙,增加了能量的浪费。因此根据文献[23],将轮询控制引入无线传感器网络,作为 GAFCMCR 的簇内通信介质访问控制(Medium Access Control, MAC)机制,采用由簇头节点依次轮询成员节点的方式来传输信息,簇头节点建立轮询表来记录所需要轮询的节点。轮询表记录的是轮询顺序和节点 ID 的对应关系,而且轮询表并不是一一对应,一个节点 ID 可以对应多个序号,反映了该节点具有较高优先级,在一个轮询周期内被多次访问。当节点能量耗尽或者处于休眠状态时,将此节点从轮询列表中移除,充分利用了时隙,又可以有效避免常见协议因碰撞带来的能量损耗。

簇头节点完成一个轮询周期后进行数据的融合,采用 CSMA 机制单跳将数据发送至基站。

GAFCMCR 的伪代码如下:

```

初始化网络,设定相关参数;
for round = 1 to maxround do
    if round == 0 then
        执行 GA 优化的 FCM 算法,得到结果和聚类中心;
        CH ← 离聚类中心最近的节点
    else
        for i = 1 to n do
            根据式(17)计算每个节点的参量 p(i)
        end for
        if p_CH <= λp_max(i) then
            CH ← p_max(i) 对应节点
        end if
        if 最优簇数发生改变 && 最优簇数大于 1 then
            执行 GA 优化的 FCM 算法,更新各参数
        end if
    end if
end for

```

2.4 算法复杂度分析

假设网络中有 N 个节点,每轮选取 K 个簇头,下面对算法的复杂度进行分析:

不同于其他轮,第一轮的簇头选举直接由基站指定,当分簇完成后,基站广播 N 条消息,消息中包含了哪些节点是簇头和各节点所属簇,各节点不用发送请求直接入簇。所以算法第一轮的复杂度为: $O(N)$ 。

从第二轮开始,每轮选举 K 个簇头,当前簇头发送一条消息与新簇头进行成员信息的交换,共发送 K 条消息;新的簇头向成员发送消息,告知其他节点自己成为簇头,此过程共产生 $N - K$ 条消息。因此每轮的总消息数为: $K + N - K = N$ 。所以算法每轮的复杂度为: $O(N)$ 。

3 仿真与分析

3.1 仿真环境

本文使用 Matlab 2014Rb 软件进行无线传感器网络的仿真。在理想的信道条件下,忽略随机因素的干扰和信号冲突等影响,监测在 $100 \text{ m} \times 100 \text{ m}$ 传感器网络区域内随机分布的 100 个节点。节点的能量损耗主要考虑数据包和控制包的发



送和接收能量损失,以及数据融合能量。仿真所用到的具体参数见表1。遗传算法的参数设置为:种群规模 $N = N \times 10 \times 2$, 交叉概率 $p_c = 0.7$, 变异概率 $p_m = 0.01$, 终止进化的迭代次数 $T = 100$; FCM 算法终止条件阈值 $\varepsilon = 10^{-6}$ 。

表1 仿真参数

Tab. 1 Simulation parameters

参数	取值
监测区域	100 m × 100 m
节点个数	100
基站坐标	(50, 50)
数据长度	4 000 bit
广播包大小	100 bit
E_{init}	0.5 J
E_{elec}	50 nJ/bit
E_{da}	50 nJ/(bit · packet ⁻¹)
ε_{fs}	10 pJ/(bit · m ⁻²)
ε_{mp}	0.0013 pJ/(bit · m ⁻⁴)
λ	0.6

3.2 GA 改进 FCM 的必要性分析

FCM 算法和 GA 改进的 FCM 算法应用于无线传感器网络同一环境的成簇结果如图4所示。从图中能够直观地看出,GA 改进后的 FCM 算法成簇结果更加均匀。FCM 算法成簇后有几个簇的成员到簇头的距离过远,这势必会增加簇内通信的能量,而改进的 FCM 则使得簇间的通信距离比较均匀,能够减少能耗。说明用 GA 改进 FCM 算法形成网络分簇能够改善成簇效果。

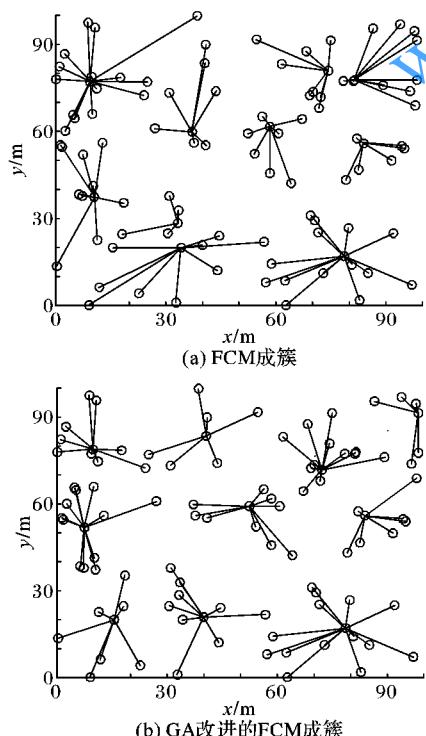


图4 FCM与改进FCM成簇对比

Fig. 4 FCM and improved FCM clustering comparison

3.3 成簇结果

GAFCMCR 与 KUCR、LEACH 算法的成簇对比如图5 所示,可以看出,LEACH 算法由于簇头选举的随机性,成簇不理

想,簇头分布不合理,簇的数量和大小不均匀,导致了能耗的增加;KUCR 算法采用 K-means 聚类算法,根据节点位置和节点 ID 计算形成分簇结果,分簇比较均匀,但是由于 K-means 算法对于初始聚类中心敏感,容易陷入局部最优,成簇存在“极大簇”“极小簇”,造成了能耗失衡。GAFCMCR 使用 GA 优化的 FCM 聚类算法进行分簇,使得成簇相比其他两种算法更加均匀,簇头分布位置更加合理。

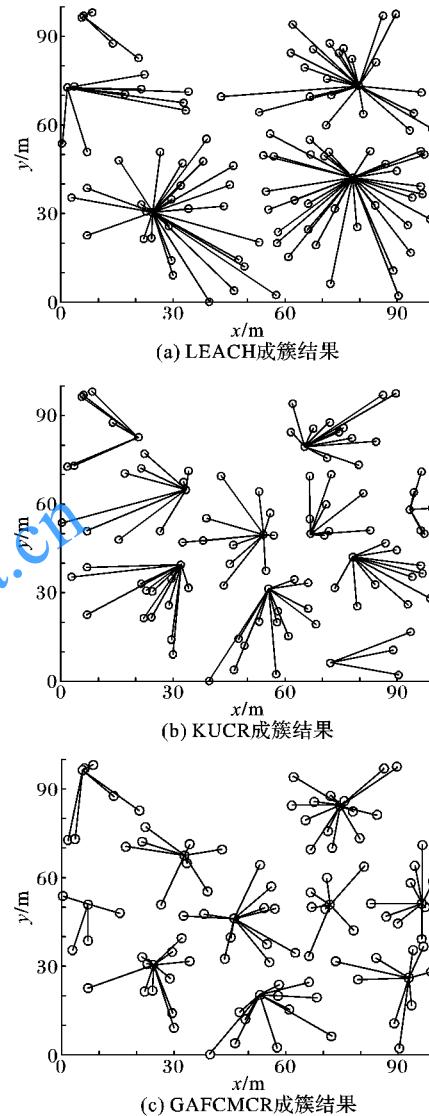


图5 成簇结果对比

Fig. 5 clustering result comparison

3.4 网络生命周期

将网络生命周期定义为网络开始运行到出现第一个死亡节点期间的运行轮数,此定义下 GAFCMCR 算法的生命周期为 2 238 轮,比 LEACH 和 KUCR 分别延长了 105% 和 20%。也有研究将生命周期定义为网络运行开始到网络中全部节点都死亡所运行的轮数或者网络运行截止到一半节点死亡的运行轮数。表2 统计了不同比例节点死亡运行到的轮数,从表中可以看出,网络运行的整个阶段 GAFCMCR 算法的生命周期都有优势。图6 是各算法生命周期对比,从图中可以看出,相比 LEACH 和 KUCR,GAFCMCR 有效提升了生命周期,而且 GAFCMCR 的曲线的斜率较大,说明能量均衡性好。



GAFCMCR 具有优势的原因是改进了聚类算法、成簇方式以及簇头的选举方式。

表 2 不同死亡节点占比出现的轮数

Tab. 2 Number of rounds of different death node ratios

死亡节点占比/%	LEACH	KUCR	GAFCMCR
1	1067	1874	2238
10	1144	2030	2314
50	1261	2243	2441
80	1360	2313	2487
100	1525	2423	2500

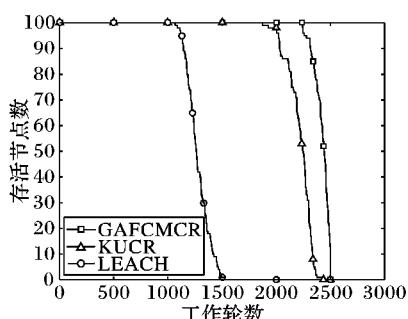


Fig. 6 Life cycle comparison

3.5 网络能耗

三种算法的网络剩余能量对比如图 7 所示，在相同初始能量下，GAFCMCR 的网络剩余能量始终高于另外两种算法。这表明 GAFCMCR 能够有效降低能耗，主要原因是：GAFCMCR 使用 GA 改进的 FCM 算法进行分簇，不存在极大、极小簇，簇头的选举过程综合考虑了剩余能量、节点与基站的距离以及与其他节点的紧密程度，距离基站近、与簇内其他节点近的节点大概率当选簇头，降低了簇头传输数据和簇间通信的能耗。

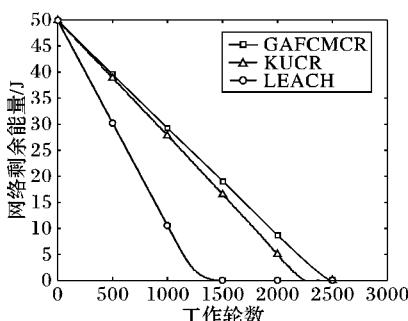


Fig. 7 Network residual energy comparison

3.6 网络吞吐量

传感网络节点部署的主要目的就是收集监测区域的各种数据，因此网络的吞吐量是一个重要的性能指标。本文用基站接收到总的数据量来衡量网络的吞吐量。GAFCMCR 和其他两种算法的吞吐量对比如图 8 所示，从图中可以看出：相比 LEACH 和 KUCR，由于 GAFCMCR 的生命周期较长，所以基站接收到了更多的数据；而且 GAFCMCR 的吞吐量从网络运行开始就高于其他算法，这得益于数据传输阶段引入了轮询机制，充分地利用了时隙。

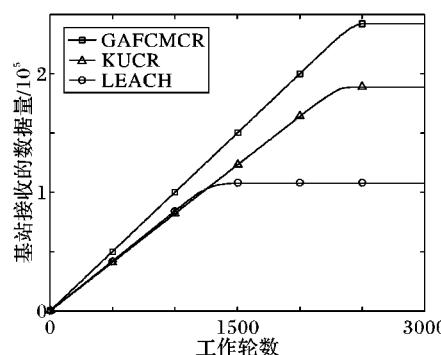


Fig. 8 Throughput comparison

4 结语

为了延长无线传感器网络的生命周期、降低能耗和提高吞吐量，本文提出了基于 GA 和 FCM 的 WSN 分簇算法 GAFCMCR。在成簇阶段，利用 FCM 算法实现分簇，并针对 FCM 算法对初始值敏感的问题，使用 GA 优化初始聚类中心后再进行聚类；在簇头选举阶段，综合考虑节点的状态，使得选出的簇头更合理；在数据传输阶段，将轮询引入簇内通信，提高了网络的吞吐量。实验结果表明，GAFCMCR 能够有效降低能耗，提高能量的均衡性，明显延长网络生命期。

参考文献 (References)

- [1] 徐晶晶, 张欣慧, 许必宵, 等. 无线传感器网络分簇算法综述 [J]. 计算机科学, 2017, 44(2): 31–37. (XU J J, ZHANG X H, XU B X, et al. Overview of clustering algorithms for wireless sensor networks [J]. Computer Science, 2017, 44 (2): 31 –37.)
- [2] SINGH S P, SHARMA S C. Cluster based routing algorithms for wireless sensor networks [J]. IJETI International Journal of Engineering & Technology Innovations, 2014, 1(4): 1–8.
- [3] LIU F, WANG Y, LIN M, et al. A distributed routing algorithm for data collection in low-duty-cycle wireless sensor networks [J]. IEEE Internet of Things Journal, 2017, 4(5): 1420–1433.
- [4] GULERIA K, VERMA A K. Comprehensive review for energy efficient hierarchical routing protocols on wireless sensor networks [J]. Wireless Networks, 2019, 25(3): 1159–1183.
- [5] HEINZELMAN W R, CHANDRAKASAN A, BALAKRISHNAN H. Energy-efficient communication protocol for wireless microsensor networks [C]// Proceedings of the 33rd Annual Hawaii International Conference on System Sciences. Piscataway, NJ: IEEE, 2000: 223–233.
- [6] 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.
- [7] KUMAR N, BHUTANI P, MISHRA P. U-LEACH: a novel routing protocol for heterogeneous wireless sensor networks [C]// Proceedings of the 2012 International Conference on Communication, Information and Computing Technology. Washington, DC: IEEE Computer Society, 2012: 1–4.
- [8] BAKARANIYA P, MEHTA S. K-LEACH: an improved LEACH protocol for lifetime improvement in WSN [J]. International Journal



- of Engineering Trends and Technology, 2013, 4(5): 1521 – 1526.
- [9] 张雅琼. 基于 K-Means 的无线传感网均匀分簇路由算法研究 [J]. 控制工程, 2015, 22(6): 1181 – 1185. (ZHANG Y Q. Research on uniform clustering routing algorithm based on K-Means in wireless sensor networks [J]. Control Engineering of China, 2015, 22(6): 1181 – 1185.)
- [10] 严静静, 张腾飞. 基于粗糙 C – 均值聚类的能量均衡 LEACH 算法 [J]. 计算机工程, 2016, 42(12): 21 – 25. (YAN J J, ZHANG T F. Energy-balanced LEACH algorithm based on rough C-Means clustering [J]. Computer Engineering, 2016, 42(12): 21 – 25.)
- [11] 陶志勇, 王和章. 基于新型聚类的无线传感器网络非均匀分层路由协议 [J]. 计算机科学, 2018, 45(3): 115 – 123. (TAO Z Y, WANG H Z. Non-uniform hierarchical routing protocol based on new clustering for wireless sensor network [J]. Computer Science, 2018, 45(3): 115 – 123.)
- [12] 戴志强, 严承, 武正江. 一种新的无线传感器网络非均匀分簇头算法——PUDCH 算法 [J]. 传感技术学报, 2016, 29(12): 1912 – 1918. (DAI Z Q, YAN C, WU Z J. New uneven double cluster head clustering algorithm for WSN——PUDCH algorithm [J]. Chinese Journal of Sensors and Actuators, 2016, 29(12): 1912 – 1918.)
- [13] 李安超, 陈桂芬. 基于奇偶轮成簇和双簇首的非均匀分簇协议 [J]. 计算机应用, 2018, 38(7): 171 – 176. (LI A C, CHEN G F. Uneven clustering protocol based on odd-even round clustering and double cluster head [J]. Journal of Computer Applications, 2018, 38(7): 1995 – 2000.)
- [14] 戚攀, 包开阳, 马晶源. 基于模糊 C 均值聚类及群体智能的 WSN 分层路由算法 [J]. 计算机应用, 2018, 38(7): 1974 – 1980, 1988. (QI P, BAO K Y, MA X Y. WSN hierarchical routing algorithm based on fuzzy C-means clustering and swarm intelligence [J]. Journal of Computer Applications, 2018, 38(7): 1974 – 1980, 1988.)
- [15] 孙增友, 周池. 基于能量和距离的 WSN 自适应分簇算法 [J]. 东北电力大学学报, 2016, 36(1): 82 – 86. (SUN Z Y, ZHOU C. Adaptive clustering algorithm based on energy and distance [J]. Journal of Northeast Dianli University, 2016, 36(1): 82 – 86.)
- [16] RUSPINI E H. Numerical methods for fuzzy clustering [J]. Information Sciences, 1970, 2(3): 319 – 350.
- [17] KONG L, PAN J-S, SNÁŠEL V, et al. An energy-aware routing protocol for wireless sensor network based on genetic algorithm [J]. Telecommunication Systems, 2018, 67(3): 451 – 463.
- [18] BAROUDI U, BIN-YAHYA M, ALSHAMMARI M, et al. Ticket-based QoS routing optimization using genetic algorithm for WSN applications in smart grid [J]. Journal of Ambient Intelligence and Humanized Computing, 2019, 10(4): 1325 – 1338.
- [19] 黄白梅. 基于 GA 优化的核模糊 C 均值聚类算法的研究 [D]. 武汉: 武汉科技大学, 2013: 23 – 29. (HUANG B M. Researches in kernel-based fuzzy C-means clustering algorithm based on GA optimization [D]. Wuhan: Wuhan University of Science and Technology, 2013: 23 – 29.)
- [20] ZHANG L, LU W, LIU X, et al. A global clustering approach using hybrid optimization for incomplete data based on interval reconstruction of missing value [J]. International Journal of Intelligent Systems, 2016, 31(4): 297 – 313.
- [21] AL-BAZ A, EL-SAYED A. A new algorithm for cluster head selection in LEACH protocol for wireless sensor networks [J]. International Journal of Communication Systems, 2018, 31(1): 1 – 13.
- [22] 吴标, 余剑, 易仁杰. 基于节点剩余能量的分时分簇 LEACH 改进算法 [J]. 火力与指挥控制, 2016, 41(10): 84 – 88. (WU B, YU J, YI R J. A time-division clustering algorithm of LEACH based on residual energy of sensor nodes [J]. Fire Control and Command Control, 2016, 41(10): 84 – 88.)
- [23] 刘龙军, 丁洪伟, 柳虔林, 等. 基于 FPGA WSN 轮询接入控制协议的研究 [J]. 通信学报, 2016, 37(10): 181 – 187. (LIU L J, DING H W, LIU Q L, et al. Design of polling access control protocol in WSN based on FPGA [J]. Journal on Communications, 2016, 37(10): 181 – 187.)

This work is partially supported by the National Natural Science Foundation of China (61461053, 61072079).

DONG Fazhi, born in 1993, M. S. candidate. His research interests include wireless sensor network, polling multiple access communication .

DING Hongwei, born in 1964, Ph. D., professor. His research interests include polling multiple access communication, random multiple access communication, communication and information system.

YANG Zhijun, born in 1968, Ph. D., senior engineer. His research interests include computer communication and network, polling system.

XIONG Chengbiao, born in 1995, M. S. candidate. His research interests include random multiple access communication system.

ZHANG Yingjie, born in 1993, M. S. candidate. Her research interests include machine learning.