

## 基于布谷鸟优化 $K$ 均值的 WSN 分簇路由算法

朱开磊, 孙爱晶

(西安邮电大学通信与信息工程学院, 陕西 西安 710121)

**摘要:** 为了延长无线传感器网络生命周期, 提出了一种基于布谷鸟优化  $K$  均值 ( $K$ -means) 的无线传感器网络分簇路由算法。分簇阶段, 使用布谷鸟算法选出初始聚类中心, 使  $K$ -means 算法的分簇结果更均匀, 以均衡节点能耗; 簇首选举综合考虑节点的剩余能量、与聚类中心的距离因素, 并依据节点剩余能量动态调整权重, 使选举的簇首更合理; 数据通信阶段, 为了进一步均衡簇首负载, 综合考虑中继节点剩余能量及其负载、簇首路由能耗因素, 结合布谷鸟算法为簇首规划路由。仿真结果表明, 提出的算法在能耗均衡性方面比 LEACH- $K$ 、LEACH-improve 和 DTK-means 算法更优, 以首节点死亡作为网络生命周期, 网络寿命分别提高了 173%、21%、6%, 所提算法有效延长了网络生命周期。

**关键词:** 无线传感器网络; 布谷鸟算法;  $K$  均值聚类; 分簇均匀; 能耗均衡

**中图分类号:** TP393

**文献标志码:** A

**doi:** 10.11959/j.issn.2096-3750.2022.00257

## WSN clustering routing algorithm based on Cuckoo Search algorithm optimized $K$ -means

ZHU Kailei, SUN Aijing

School of Communication and Information Engineering, Xi'an University of Posts and Telecommunications, Xi'an 710121, China

**Abstract:** In order to extend the lifetime of wireless sensor network (WSN), a clustering routing algorithm for WSN based on Cuckoo Search (CS) algorithm optimized  $K$ -means was presented. In the clustering stage, the initial cluster centers were selected by CS algorithm, which make the clustering results of the  $K$ -means algorithm more uniform to balance node energy consumption. The remaining energy of the node, the distance from the center of the cluster were comprehensively considered in the cluster election, and the weight according to the remaining energy of the node was dynamically adjusted. In the data communication stage, in order to further balance the load of the cluster head, the remaining energy of the relay node and its load, and the cluster head routing energy consumption were comprehensively considered, CS algorithm was combined to plan routing for the cluster head. The simulation results show that the proposed algorithm is better than LEACH- $K$ , LEACH-improve and DTK-means in terms of energy consumption balance. With the death of the first node as the life cycle of the network, the network lifespan was increased by 173%, 21%, and 6% respectively. The proposed algorithm effectively extending the network life cycle.

**Key words:** wireless sensor network, cuckoo search algorithm,  $K$ -means clustering, even clustering, balanced energy consumption

收稿日期: 2021-07-27; 修回日期: 2021-12-29

通信作者: 朱开磊, zhukailei1995@163.com

基金项目: 陕西省科技成果推广项目 (No.2018CG-007); 陕西省创新人才推进计划-物联网科技创新团队项目 (No.2019TD-028)

**Foundation Items:** The Shaanxi Province Science and Technology Achievement Promotion Project (No.2018CG-007), The Shaanxi Province Innovative Talent Promotion Plan-The Internet of Things Technology Innovation Team Project (No. 2019TD-028)

## 0 引言

无线传感器网络 (WSN, wireless sensor network) 由大量廉价、功耗低的微型传感器节点组成<sup>[1]</sup>。传感器节点通过无线通信方式形成自组织网络, 可以感知监测信息, 收集处理数据并将信息发送至汇聚中心或者基站<sup>[2]</sup>。由于传感器节点的体积小, 所以其自身能量有限, 传感器节点通常部署在恶劣的环境中, 随着网络的运行, 网络节点逐渐死亡, 逐一更换电池是不切实际的, 因此可以通过降低节点能量的消耗延长无线传感器网络的生命周期<sup>[3]</sup>。

针对如何提高网络生命周期问题, 许多基于簇结构的层次性网络路由协议被提出<sup>[4]</sup>。文献[5]提出了低功耗自适应集簇分层型 (LEACH, low energy adaptive clustering hierarchy) 协议, 协议中以循环的方式随机选择簇首, 使各个节点都有机会成为簇首来均衡网络能耗。但是簇首选择没有考虑其剩余能量, 能量低的节点一旦被选为簇首, 就会很快死亡。此外, 簇首的选择没有考虑与基站和簇内其他节点的距离, 增大了簇首的数据传输能耗。针对 LEACH 协议存在的缺点, 文献[6]提出了混合低功耗自适应集簇分层型 (H-LEACH, hybrid-low energy adaptive clustering hierarchy) 协议, H-LEACH 协议将剩余能量加入阈值计算公式, 并比较候选节点剩余能量与平均剩余能量的大小, 来选举簇首。算法考虑了簇首的能量, 缺点是未考虑簇首与其他节点位置, 导致非簇首节点能量消耗过快。文献[7]提出了一种能量高效的非均匀分簇 (EEUC, energy-efficient uneven clustering) 路由协议, 利用非均匀分簇解决多跳模式下簇首的能耗不均衡问题。文献[8]提出了 LEACH-K (LEACH-K, routing protocol LEACH-K using K-means algorithm in wireless sensor network) 路由协议, 该协议先用 K-means 算法进行分簇, 然后采用 LEACH 协议中的簇首选举方式来选举簇首, 根据最近距离原则将 K-means 算法得到的分簇加入最近簇首中。该协议虽然降低了 LEACH 协议出现极小、极大簇的可能性, 但由于簇首选举没有考虑能量、距离因素, 簇首选举仍不合理。文献[9]提出了 LEACH-improve (LEACH-improve, improved LEACH protocol algorithm for WSN based on energy balance and high efficiency) 协议, 该协议将剩余能量因子、间距因子、密度因子融入传统 LEACH 阈值计算公式, 使选举的簇首能量充足、

位置适当, 网络生命周期得以提高, 但是总体上簇首与其他节点的距离仍不均衡且易出现分簇不均匀的现象。文献[10]提出了基于 K-Means 的 WSN 动态信任度双簇头选取 (DTK-means, dual cluster head selection algorithm based on K-means and dynamic trust degree in WSN) 算法, DTK-means 首先通过 K-means 算法进行节点的分簇, 然后利用节点剩余能量、距离等因素作为动态信任度选取双簇头, 延长了网络生命周期, 但未对簇首规划路由以降低簇首能耗、未考虑 K-means 算法对初始聚类中心敏感问题, 导致分簇结果不均匀<sup>[11]</sup>。

针对以上文献出现的不足之处, 在已有研究的基础上, 本文提出了基于布谷鸟优化 K 均值的 WSN 分簇路由 (CS-K, cuckoo search-K-means) 算法。为了避免出现 LEACH、LEACH-improve 算法中分簇不均匀的现象, 本文采用 K-means 算法进行分簇。此外, 为了解决 K-means 算法对初始聚类中心敏感而陷入局部最优的问题, 本文算法使用布谷鸟 (CS, cuckoo search) 算法优化 K-means 算法的初始聚类中心, 使节点分簇更均匀; 为了选举合理的簇首, 均衡簇内节点能耗, 簇首选举综合考虑节点剩余能量、与聚类中心的距离, 同时采用动态的权系数更新以适应网络不同阶段对每个因子的不同需求; 数据传输阶段, 为了减轻簇首因长距离传输数据消耗的能量, 基于簇首路由能耗、中继节点剩余能量和中继节点负载设计簇首路由寻优目标函数, 结合 CS 算法为簇首规划路由。仿真实验证明, 由于分簇均匀稳定, 簇首选举合理, 簇首路由适当, 网络的生命周期得以延长。

## 1 系统模型

### 1.1 网络模型

为了测试提出的算法对无线传感器网络的影响, 对网络模型做以下设定。

- 1) 监测区域的形状为  $M \times M$  的正方形, 传感器节点在监测区域内随机不均匀分布, 布设后将不再移动。
- 2) 监测区域内的所有传感器节点初始能量相同, 处理能力和通信能力相同, 基站的处理能力和能量不受限制。
- 3) 传感器节点可以感知自己的剩余能量, 具有融合冗余数据的能力。
- 4) 传感器节点发射功率可调节。

## 1.2 能耗模型

本文采用的网络能耗模型与文献[12]模型相同。在数据传输过程中，节点的能量消耗主要包含发送数据消耗的能量与功率放大电路消耗的能量两部分。节点发送  $m$  位数据至距离为  $d$  的节点消耗的能量为

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

其中， $E_{\text{elec}}$  表示每传输 1 bit 的数据需要消耗的能量。 $d_0$  为节点之间传输数据的模型选择的一个阈值， $d_0 = \sqrt{\epsilon_{\text{fs}}/\epsilon_{\text{mp}}}$ 。当传输距离  $d < d_0$ ，采用自由空间信道模型；当传输距离  $d \geq d_0$ ，采用多路径衰减模型。 $\epsilon_{\text{fs}}$ 、 $\epsilon_{\text{mp}}$  分别为自由空间信道模型和多路径衰减模型的功放因子参数。接收节点接收融合  $m$  bit 数据消耗能量为

$$E_{\text{rx}} = mE_{\text{elec}} + mE_{\text{da}} \quad (2)$$

其中， $E_{\text{da}}$  表示融合 1 bit 数据消耗的能量。

## 2 布谷鸟算法优化 $K$ -means 算法

$K$ -means 算法的良好成簇特性使其广泛应用在多种实际场景中，包括无线传感器网络<sup>[13]</sup>。 $K$ -means 算法是基于距离或者相似度的聚类算法，通过最小化其目标函数来得到最优分簇， $K$ -means 算法的目标函数  $E$  为

$$E = \sum_{j=1}^k \sum_{x \in C_j} \text{dist}(x - C_j)^2 \quad (3)$$

其中， $\text{dist}(\cdot)$  表示欧氏距离， $C_j$  表示第  $j$  个中心， $k$  表示簇首数。

利用  $K$ -means 算法簇内样本相似度高、簇内成员紧密度好等优点，将其应用到无线传感器网络分簇中，可以得到比较均匀的分簇，减少了由分簇不均匀引起的能量损耗。但是  $K$ -means 算法对初始聚类中心敏感，且容易陷入局部最优，为了解决这个问题，利用 CS 算法<sup>[14]</sup>良好的全局搜索能力，将 CS 算法应用到  $K$ -means 算法的优化中，先通过 CS 算法优化  $K$ -means 的初始聚类中心，再执行  $K$ -means 算法得到最优分簇及聚类中心。

布谷鸟以随机或者类似随机飞行的方式寻找适合自己产卵的鸟巢的位置，为了简化、模拟布谷鸟繁殖方式，为 CS 算法假设 3 个理想状态。

1) 每只布谷鸟一次只产一个卵，并随机选择鸟巢孵化。

2) 具有高质量的鸟巢将被保留到下一代。

3) 能够利用的鸟巢数量固定，鸟巢主人会以  $P_a$  的概率发现外来鸟蛋，并将其推出鸟巢或者重新建立新鸟巢。

在以上 3 个理想状态的基础上，布谷鸟寻找下一鸟巢的位置更新计算式为

$$\mathbf{X}_k^{t+1} = \mathbf{X}_k^t + \alpha \otimes \text{Levy}(\beta) \quad (4)$$

其中， $\mathbf{X}_k^t$  表示第  $t$  代第  $k$  个解， $\alpha$  为步长因子，用来控制随机搜索范围， $\otimes$  表示点对点乘法， $\alpha \otimes \text{Levy}(\beta)$  是莱维飞行步长。步长因子  $\alpha$  与莱维随机数  $\text{Levy}(\beta)$  采用文献[15]的计算式

$$\alpha = \alpha_0 \cdot (\mathbf{X}_k^t - \mathbf{X}_{\text{best}}) \quad (5)$$

$$\text{Levy}(\beta) \sim \frac{\phi \cdot \mu}{|\nu|^{\frac{1}{\beta}}} \quad (6)$$

$$\phi = \left( \frac{\Gamma(1 + \beta) \cdot \sin(\pi \times \beta/2)}{\Gamma\left(\left(\frac{1 + \beta}{2}\right) \cdot \beta \cdot 2^{(\beta-1)/2}\right)} \right)^{\frac{1}{\beta}} \quad (7)$$

其中， $\mu$  和  $\nu$  服从标准正态分布； $\beta = 1.5$ ； $\mathbf{X}_{\text{best}}$  表示当前最优解； $\alpha_0 = 0.001 \cdot r_{\text{max}} \cdot \exp(-r/r_{\text{max}})$ ， $r$ 、 $r_{\text{max}}$  分别表示当前迭代次数与最大迭代次数， $\alpha_0$  的引入是为了提高算法的局部搜索性能。

鸟巢位置更新后，生成随机数  $r$ ，与  $P_a$  比较大小，若  $r > P_a$ ，宿主鸟以概率  $P_a$  发现巢中为外来鸟蛋，淘汰该鸟蛋，并在附近以式(8)方式随机更新一次鸟窝位置，否则位置不变。

$$\mathbf{X}_k^{t+1} = \mathbf{X}_k^t + \nu \cdot (\mathbf{X}_j^t - \mathbf{X}_h^t) \quad (8)$$

其中， $\mathbf{X}_j^t$  和  $\mathbf{X}_h^t$  表示第  $t$  代的两个随机解， $\nu$  是服从[0,1]上的均匀分布。

CS 算法优化  $K$ -means 算法的步骤如下。

**步骤 1** 随机初始化鸟巢位置，设置种群规模  $c$ 、最大迭代次数等参数。根据式(3)计算所有鸟巢的适应度。

**步骤 2** 利用式(4)对每个鸟巢的位置和状态进行更新。

**步骤 3** 根据式(3)计算更新位置后鸟巢的新适应度值，若新的适应度值更优（本文中适应度值越小越优），则替换旧的鸟巢位置。

**步骤 4** 比较随机数  $r$  与  $P_a$  的大小关系，文中  $P_a=0.25$ ，若  $r > P_a$ ，根据式(8)更新一次鸟巢位置，否则位置不变。

**步骤 5** 计算各鸟巢适应度，输出当前最优鸟巢位置。

**步骤 6** 当满足最大迭代次数或者适应度函数收敛，进行下一步，否则返回到步骤 2。

**步骤 7** 将 CS 算法迭代的最优鸟巢位置解码为  $K$ -means 的初始聚类中心。

**步骤 8** 执行  $K$ -means 算法。

**步骤 9** 输出最优聚类中心及聚类结果。

CS 优化  $K$ -means 算法流程如图 1 所示。

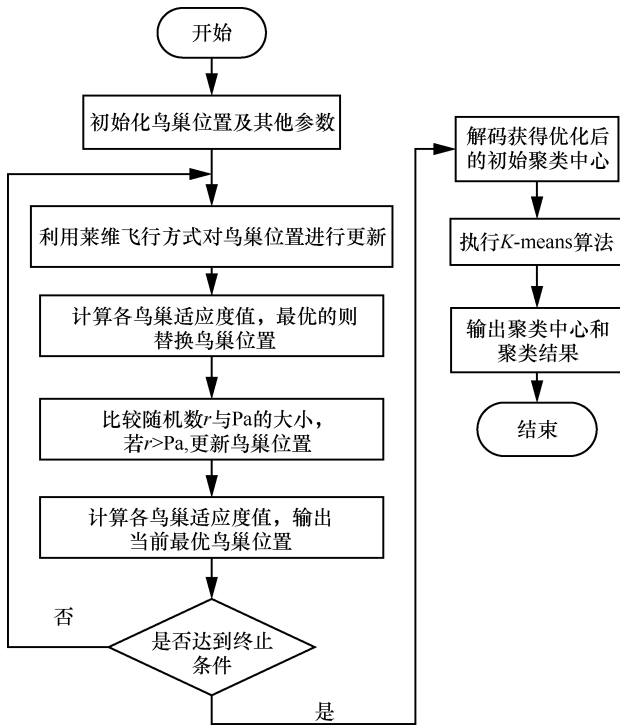


图 1 CS 优化  $K$ -means 算法流程

### 3 CS-K 算法

CS-K 算法以轮为单位，每轮分为成簇阶段、动态簇首选举阶段和数据传输阶段。成簇阶段，采用基于布谷鸟算法优化的  $K$ -means 算法对网络进行分簇；动态簇首选举阶段，综合考虑节点能量、距离因素选举簇首；数据传输阶段，首先根据簇首到基站距离判断是否需要多跳，然后基于簇首路由能耗、中继节点剩余能量和中继节点负载利用 CS 算法为簇首规划路由。

#### 3.1 成簇阶段

合理的簇首数有利于形成均匀的网络分簇，且

进一步均衡网络能耗。为了降低无线传感器网络消耗的能量，本文采用与文献[16]相同的最优簇首数，计算式为

$$k = \sqrt{\frac{N}{2\pi}} \frac{M}{d_{\text{toBS}}} \quad (9)$$

其中， $N$  表示监测区域内存活节点数目， $M$  表示监测区域的边长， $d_{\text{toBS}}$  表示簇首到基站的平均距离。

在使用 CS 算法优化  $K$ -means 之前，基于网络最优簇首数对鸟巢进行编码，编码格式为  $\{x_1, y_1, \dots, x_k, y_k\}$ ，其中， $x_i, y_i$  表示聚类中心在监测区域中的位置，可表示为

$$\begin{cases} x_i = x_{\min} + (x_{\max} - x_{\min}) \cdot \text{rand}() \\ y_i = y_{\min} + (y_{\max} - y_{\min}) \cdot \text{rand}() \end{cases} \quad (10)$$

其中， $x_{\min}, y_{\min}, x_{\max}, y_{\max}$  分别表示区域边界， $\text{rand}()$  为 0 到 1 之间的随机数。

首先使用 CS 算法对聚类中心的位置进行优化，然后利用 CS 算法最后的迭代结果作为  $K$ -means 算法的初始聚类中心， $K$ -means 算法继续执行，直至得到最优聚类中心及分簇，选择距离最优聚类中心最近的节点作为簇首，簇首为簇内节点分配 TDMA 时隙，簇内节点只在分配给自己的时隙到来时向簇首发送数据，其余时间保持休眠以节省能量。

随着网络的运行，网络中存活节点减少，由式(9)可知网络最优簇首数发生改变，当最优簇首数发生改变时需要重新分簇，所以分簇是一个动态过程。

#### 3.2 动态簇首选举阶段

无线传感器网络分簇完成后，需要选举簇首，簇首承担簇内数据的接收、融合、转发任务，能量消耗较簇内其他节点多，故选择一个剩余能量相对较高、距离簇内其他节点较近的候选节点作为簇首可以减缓节点死亡速度、均衡网络能量负载、提高网络生命周期。因此，综合考虑以上两个因素构造适应度函数在簇内选择最优簇首。

剩余能量是一个簇首选举的重要依据，故 CS-K 算法采用簇内节点的剩余能量与簇内所有节点的平均剩余能量比值作为参考标准，如式(11)所示。

$$f_1(i) = \frac{E_{\text{res}}(i)}{\frac{1}{n} \sum_{j=1}^n E_{\text{res}}(j)} \quad (11)$$

其中， $n$  为簇  $i$  中节点总数； $E_{\text{res}}(i)$  为对应节点  $i$  的剩余能量，其值越大表明节点  $i$  的剩余能量越多，被

选为簇首的概率越大。

候选簇首与簇内其他节点的距离对节省传输能耗至关重要，候选节点与簇质心的距离关系为

$$f_2(i) = \frac{d_{\max} - d(x_i, \text{cen}_j)}{d_{\max} - d_{\min}} \quad (12)$$

其中， $d(x_i, \text{cen}_j)$  表示簇  $j$  中的节点  $i$  到该簇中心  $\text{cen}_j$  的距离， $d_{\max}$  表示簇  $j$  内节点到簇中心的最大距离， $d_{\min}$  表示簇  $j$  内节点到簇中心的最小距离。 $f_2(i)$  值越大，表示候选节点  $i$  与簇中心越近，与簇内其他节点的距离就越小，所以簇内传输能耗就越小。

综合考虑以上两个因素，构造最优簇首选举函数为

$$f(i) = a(i) \cdot f_1(i) + b(i) \cdot f_2(i) \quad (13)$$

前期节点能量充沛，设定  $a(i)$ 、 $b(i)$  均为  $1/2$ 。后期节点能量下降，剩余能量成为提高网络生命周期的关键因素，故调节两个因素的所占比重，引入  $a(i)$ 、 $b(i)$  动态权重系数，表示为

$$a(i) = \frac{E_{\text{res}}(i)}{E_{\text{init}}} \quad (14)$$

$$b(i) = 1 - a(i) \quad (15)$$

其中， $E_{\text{init}}$  表示节点初始能量； $E_{\text{res}}(i)$  表示节点  $i$  当前剩余能量；随着网络运行，节点剩余能量越大， $a(i)$  的值越大，剩余能量高的节点被选为簇首的概率变大，降低了簇首过早透支的可能性，整个网络的负载会更加均衡。为了避免因频繁选举簇首消耗能量，簇首选举采用式(16)的方式计算

$$f_{CH} \leq \lambda f_{\max}(i) \quad \lambda \in (0,1) \quad (16)$$

其中， $\lambda$  为网络系数<sup>[16]</sup>。若式(16)成立，上一任簇首通知拥有最大适应度值的节点  $i$  当选为新簇首，并交换簇内成员信息；否则不进行簇首的更新。

簇首确定后，簇内以单跳形式进行通信，并为簇成员分配 TDMA 时隙。簇成员在其对应时隙内向簇首发送数据，其他时间保持休眠状态，以降低节点的能耗。

### 3.3 数据传输阶段

簇首接收融合完簇内数据后，要向基站发送数据，随着数据传输距离的增大，数据传输能耗按指

数级增加<sup>[17]</sup>，为了减少簇首因长距离传输数据消耗的能量、均衡簇首负载，对簇首进行路由规划。

#### 3.3.1 结合布谷鸟算法的簇首路由算法

簇首在充当路由中继节点时，会因接收融合上一跳中继节点的数据而加剧自身能量消耗。为了能够均衡簇首负载，考虑 CS 算法拥有良好的全局搜索能力特点，本文采用 CS 算法在监测区域内的所有节点中为簇首搜索路由中继节点，以降低簇首被选为中继节点的概率。

在利用 CS 算法为簇首规划路由之前，首先需要确认中继节点个数，以便初始化鸟巢位置维度及 CS 算法的目标函数设计。为了能够进一步减轻簇首能耗，根据文献<sup>[18]</sup>中的最佳传输距离  $d'$ 、簇首  $i$  到基站的距离  $d_{i-\text{bs}}$ ，计算簇首经几跳将数据发送至基站，其中， $\text{round}(\cdot)$  为四舍五入。簇首  $i$  的路由跳数  $\text{hop}^i$  如式(17)所示。

$$\text{hop}^i = \text{round}\left(\frac{d_{i-\text{bs}}}{d'}\right) + 1 \quad (17)$$

根据式(17)可知簇首  $i$  需要  $\text{hop}^i - 1$  个中继节点。采用 CS 算法对簇首  $i$  进行路由规划时，需要对鸟巢进行位置编码，基于簇首  $i$  的路由跳数  $\text{hop}^i$ ，鸟巢位置可编码为

$$\mathbf{X} = \{x_{\text{ch}_i}, y_{\text{ch}_i}, \text{re}_x_1, \text{re}_y_1, \dots, \text{re}_x_{\text{hop}^i-1}, \text{re}_y_{\text{hop}^i-1}, x_{\text{bs}}, y_{\text{bs}}\} \quad (18)$$

其中， $x_{\text{ch}_i}$ 、 $y_{\text{ch}_i}$  表示簇首  $i$  位置， $\text{re}_x_{\text{hop}^i-1}$ 、 $\text{re}_y_{\text{hop}^i-1}$  表示簇首  $i$  的第  $\text{hop}^i - 1$  个候选中继节点位置，分别在监测区域内的节点中随机选取， $x_{\text{bs}}$ 、 $y_{\text{bs}}$  表示基站位置。按照以上相同方式对簇首  $i$  的其他鸟巢位置进行编码，直至全部数量的鸟巢位置编码完成，鸟巢数量在实验仿真部分已给出。

鸟巢位置初始化后，CS 算法仅对候选中继节点迭代更新，簇首和基站位置固定不变，CS 算法输出的最优鸟巢位置即为簇首路由。

结合布谷鸟的簇首路由算法步骤可表示为

**步骤 1** 在监测区域内，根据式(18)对鸟巢位置进行编码，并初始化布谷鸟算法其他参数。依据式(22)计算每个鸟巢的适应度值。

**步骤 2** 利用式(4)对每个鸟巢的位置和状态进行更新。

**步骤 3** 根据式(22)计算更新位置后鸟巢的新适应度值，若新的适应度值更优，则替换旧的

鸟巢位置。

**步骤 4** 比较随机数  $r$  与  $Pa$  的大小关系, 若  $r > Pa$ , 根据式(8)更新一次鸟巢位置, 否则位置不变。

**步骤 5** 计算各鸟巢适应度, 输出当前最优鸟巢位置。

**步骤 6** 当满足最大迭代次数或者适应度函数收敛, 进行下一步, 否则返回到步骤 2。

**步骤 7** 将 CS 算法迭代的最优鸟巢位置解码为最优簇首路由。

最优鸟巢位置解码为最优簇首路由的形式见表 1。

簇首坐标	簇首中继节点坐标	基站坐标
$(x_{ch}, y_{ch})$	$(re_{x_1}, re_{y_1}) \rightarrow \dots \rightarrow (re_{x_{hop-1}}, re_{y_{hop-1}})$	$(x_{bs}, y_{bs})$

簇首中继节点坐标为 CS 算法迭代的最优簇首中继节点, 相对式(18)数值已经发生更新, 即得到了最优簇首路由, 最优簇首路由路径的形式如表 1 中各坐标及箭头所示。其他簇首路由的求解与上述算法步骤相同。

此外, 从结合布谷鸟的簇首路由算法的步骤 1 中可以看出, CS 算法在为簇首规划路由时, 需要一个标准评价路由的优劣, 以得到最优簇首路由, 故需要为 CS 算法设计一个簇首路由寻优目标函数。

### 3.3.2 簇首路由寻优目标函数设计

簇首路由规划的目标是均衡簇首能耗, 如何设计一个合理的路由寻优目标函数至关重要。结合簇首数据传输方式, 簇首最优中继节点有以下几点特征。

- 1) 中继节点距离簇首及基站距离适中, 以保证该条路径的数据传输消耗能量越少。
- 2) 中继节点的剩余能量能保证数据传输的稳定。
- 3) 中继节点的选举不能局限于簇首之中, 以减轻簇首负载。
- 4) 中继节点应降低被过度开发的可能性。

根据以上的要求, 设定路由能量消耗、中继节点剩余能量、中继节点负载作为簇首路由寻优目标函数。

路由能量消耗包含簇首向下一跳节点发送  $m$  bit 数据消耗的能量  $E_{ch}$ 、中继节点接收融合来自上一跳节点的  $m$  bit 数据以及发送  $m$  bit 数据至下一跳中继节点或者基站消耗的能量  $E_{relaynode}$ 。

$$E_{ch} = mE_{elec} + m\epsilon_{fs}d_{ch\_j}^2 \quad (19)$$

$$E_{relaynode} = 2 \cdot mE_{elec} + m \cdot E_{da} + m\epsilon_{fs}d_{nexthop}^2 \quad (20)$$

$$E_{total} = E_{ch} + (hop - 1) \cdot E_{relaynode} \quad (21)$$

其中,  $d_{ch\_j}$ 、 $d_{nexthop}$  分别表示簇首到其下一跳中继节点  $j$  的距离, 中继节点到下一跳的距离。 $d_{ch\_j}$ 、 $d_{nexthop}$  值越小, 簇首到基站路由的总能耗开销  $E_{total}$  越小; 此外, 中继节点拥有充足的剩余能量是路由数据稳定传输的关键; 为了降低中继节点被过度开发的可能性, 计算其充当中继节点的次数作为负载  $load$ , 负载越小, 路由能耗越均衡。综上, 将以上 3 个因素构造成簇首路由寻优目标函数  $F$ 。

$$F = \chi \frac{E_{all\_relay\_min\_ave}}{E_{all\_relay\_ave}} + \eta \frac{E_{total}}{E_{total\_max}} + \delta \cdot load \quad (22)$$

其中,  $E_{all\_relay\_ave}$  表示前向路径上所有中继节点剩余能量平均值;  $E_{all\_relay\_min\_ave}$  表示最小  $E_{all\_relay\_ave}$ , 即所有鸟巢对应的  $E_{all\_relay\_ave}$  最小值;  $E_{total\_max}$  表示所有鸟巢对应的  $E_{total}$  最大值;  $\chi$ 、 $\eta$ 、 $\delta$  为权重系数,  $\chi + \eta + \delta = 1$ , 由于负载  $load$  的值通常大于前一项, 故设置  $\chi + \eta > \delta$ 。由目标函数可知, 当中继节点剩余平均能量越大, 路由能耗开销越小, 中继节点的负载  $load$  越小, 目标函数  $F$  的值就越小, 结合 CS 算法得到的簇首路由就越合理。

## 4 实验仿真及分析

采用 MATLAB R2016a 软件对本文提出的 CS-K 算法和已有的 LEACH-K、LEACH-improve、DTK-means 算法进行仿真, 从分簇效果、存活节点数、网络剩余能量进行性能分析评估。定义 CS 算法鸟巢数量为 25, 迭代次数为 300。其他参数设置见表 2。

### 4.1 分簇效果对比分析

在同一无线传感器网络环境中, LEACH-K、K-means 算法和 CS-K 算法的分簇结果分别如图 2、图 3、图 4 所示。

对比图 2、图 4 的分簇效果可以明显看出, 图 2 中 LEACH-K 算法出现了极小、极大簇, 且簇首与簇内其他节点距离不均匀, 加速了远离簇首节点的能耗, 这主要由于 LEACH-K 算法的簇首选举阶段未考虑节点位置因素, 且未对 K-means 算法优化。

表 2 部分仿真参数

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
最佳传输距离 $d$	32 m
$E_{init}$	5 J
$E_{elec}$	50 nJ/bit
$E_{da}$	5 nJ/(bit · packet <sup>-1</sup> )
$\epsilon_{fs}$	10 pJ/(bit · m <sup>-2</sup> )
$\epsilon_{mp}$	0.001 3 pJ/(bit · m <sup>-4</sup> )
$\chi, \eta, \delta, \lambda$	0.4, 0.4, 0.2, 0.8

LEACH- $K$ 、 $K$ -means 算法明显改善了网络分簇效果，均衡了节点能量消耗。

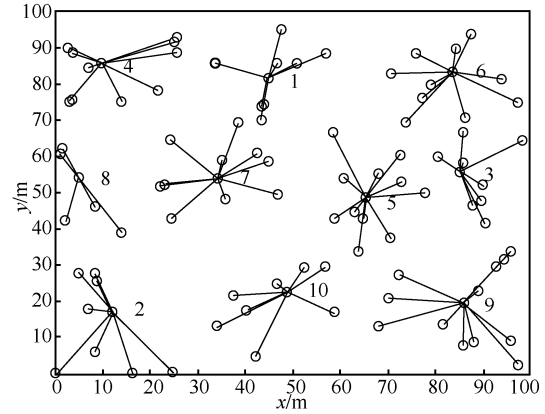


图 4 CS 优化  $K$ -means 分簇

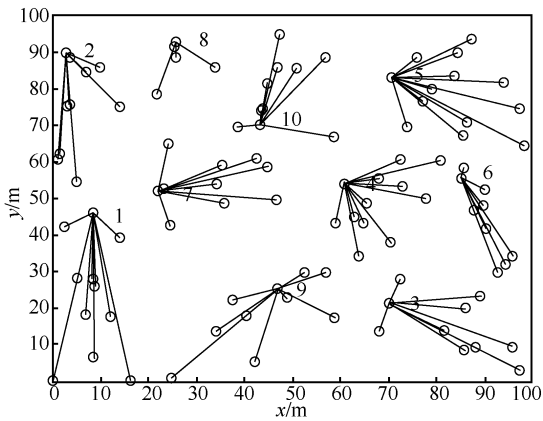


图 2 LEACH- $K$  算法分簇

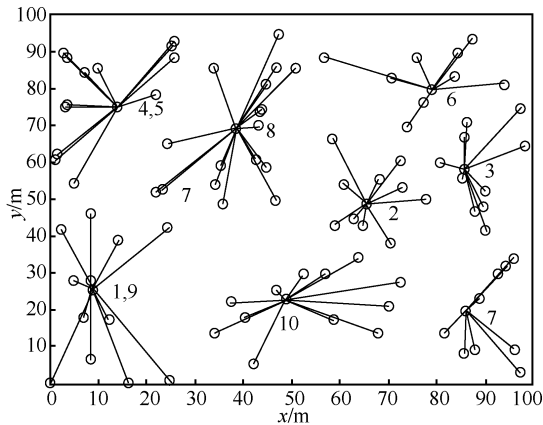


图 3  $K$ -means 算法分簇

对比图 3、图 4 发现，图 3 中的簇 4 和簇 5、簇 1 和簇 9 发生了重叠，增大了分簇规模，进而加剧了簇首能耗。鉴于  $K$ -means 算法对初始聚类中心敏感而引起的分簇不均匀，图 4 中的 CS- $K$  算法分簇利用 CS 算法优化  $K$ -means 的初始聚类中心，得到更均匀的分簇。综上分析，表明 CS- $K$  算法较

### 4.2 网络生命周期

定义无线传感器网络开始运行到首节点死亡之间的轮数为其生命周期。实验中将 CS- $K$  算法分别与 LEACH- $K$ 、LEACH-improve 和 DTK-means 算法进行节点存活数量变化的仿真比较，图 5 为仿真结果。

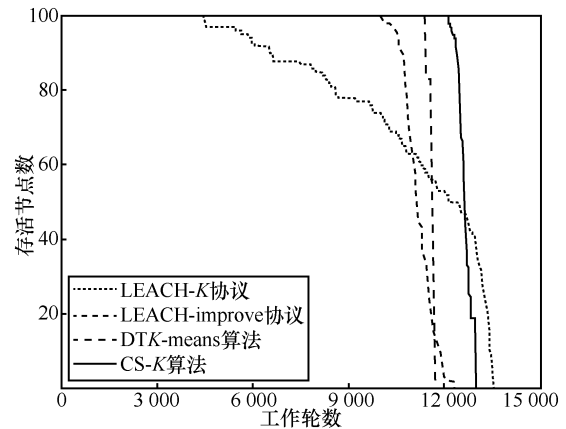


图 5 网络生命周期对比

可以看出 CS- $K$  算法首节点死亡时间出现在第 12 100 轮左右，较 LEACH- $K$ 、LEACH-improve 和 DTK-means 算法分别提高 173%、21%、6%。由于 CS- $K$  算法通过 CS 算法优化  $K$ -means，相较于 DTK-means 算法得到均匀、紧密的分簇效果，使簇内通信开销均衡；簇首选举阶段考虑了节点剩余能量、与聚类中心位置因素，使剩余能量充足且与簇内其他节点距离均匀的节点更有可能成为簇首；同时，在数据通信阶段，使用 CS 算法为簇首规划路由，综合考虑中继节点剩余能量、负载及簇首路由能耗，使得到的中继节点更具合理性。因此，CS- $K$

算法在能量均衡消耗方面较 LEACH-K、LEACH-improve 和 DTK-means 算法有所提高,从而延长了整个网络的生命周期。

### 4.3 网络能耗

无线传感器网络生命周期主要由节点的剩余能量决定,且当节点剩余能量越多时,整个网络维持时间越久。LEACH-K、LEACH-improve、DTK-means 算法和 CS-K 算法的节点剩余能量对比如图 6 所示。

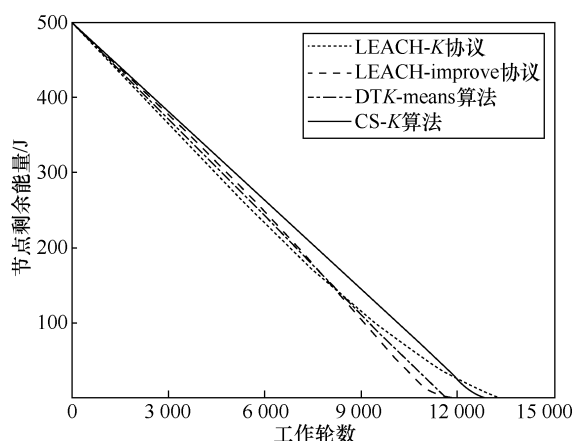


图6 节点剩余能量对比

在大部分工作轮数内,CS-K 算法的节点剩余能量高于 LEACH-K、LEACH-improve 和 DTK-means 算法,说明 CS-K 算法使网络能耗更均衡。这得益于 CS 算法优化后的  $K$ -means 算法分簇均匀,降低了局部“热点”的可能性,使簇内节点能耗均衡;簇首选举综合考虑了节点位置、剩余能量,使选举出的簇首更具合理性;此外,使用 CS 算法为簇首规划路由进一步均衡了簇首负载,从而延长了网络生命周期。

### 4.4 算法复杂度分析

衡量算法计算量的一种标准是时间复杂度,本节主要分析 CS-K 算法的时间复杂度。第一阶段,CS 算法迭代次数为  $iter$ ,种群规模为  $c$ ,种群维度为  $2k$ ,所以 CS 算法的时间复杂度为  $O(iter \cdot c \cdot k)$ 。 $K$ -means 算法的迭代次数为  $iter$ ,由于簇首数为  $k$ ,考虑节点的纵横坐标,故数据维度设为  $2k$ ,那么  $K$ -means 算法的时间复杂度为  $O(iter \cdot k)$ ,所以该阶段的时间复杂度为  $O(iter \cdot c \cdot k)$ ;第二阶段,计算  $N$  个节点的适应度值,选举出最优簇首,故簇首选举阶段的时间复杂度为  $O(N)$ ;第三阶段,CS 算法迭代次数为  $iter$ ,种群规模为  $c$ ,簇首数为  $k$ ,所以簇首路由寻优过程的复杂度为  $O(iter \cdot c \cdot k)$ 。综上,本文算法的复杂度为  $O(iter \cdot c \cdot k)$ 。对比算法中的

LEACH-K、LEACH-improve 算法复杂度较低,是因为其算法过程简单,但是也导致了其性能远远不及本文算法。而 DTK-means 算法的复杂度为  $O(iter \cdot k)$ ,本文算法由于在  $K$ -means 算法的基础上进行了优化,和簇首路由的寻优,所以算法在复杂度上稍显劣势,但是在网络生命周期等方面表现良好。总之,本文算法在略微牺牲复杂度的情况下,算法性能有了明显提升。

### 4.5 算法可行性分析

由于传感器网络设备运算能力有限,所以对提出算法进行可行性分析是必要的。本文的无线传感网络分簇路由算法根据网络运行情况主要分为 3 种运算模式。

模式 1: 在网络运行第一轮时由基站接收节点信息并进行网络分簇、选举簇首、簇首路由规划的运算,之后将运算的结果广播至各节点。

模式 2: 之后的轮数内如果基站依据接收信息判断节点未发生死亡,即最优簇首数未发生变化,不重新分簇。簇首依据接收的节点信息在簇内重新选举簇首,之后融合簇内各节点数据,并将选举信息附加在融合数据内,一并传送给基站。当基站检测到簇首发生改变时,重新运算簇首路由算法并将路由信息发出,上一轮簇首在接收到路由信息后根据信息内容交换簇内信息。

模式 3: 之后的轮数内如果基站依据接收的信息判断节点发生死亡,并由式(9)计算最优簇首数,簇首数未发生改变时,进行与模式 2 相同的运算。簇首数发生改变时,重新分簇,此时由基站进行网络分簇、簇首选举和簇首路由规划的运算,前一轮的簇首在接收基站广播的运算信息后,再将接收的信息分发给前一轮簇内成员,所有节点依据接收的信息重新定义分簇路由信息。从上面的分析中可以发现繁重的计算由处理能力无限的基站完成,簇首即节点仅进行簇首的簇内选举,运算复杂度约为  $O(N/k)$ ,其中, $N$  为节点总数, $k$  为簇首数,即簇首仅涉及对簇内节点簇首选举适应度值的运算和选择,所以节点承担的运算是简单和少量的,对于指令内存大小有 128 MB,处理速度为 4 MHz 的传感器节点<sup>[19]</sup>来说,运算能力完全是能够达到的。综上分析,本文提出的算法可运用于实际。

## 5 结束语

为了延长无线传感器网络生命周期,本文提出

了基于布谷鸟优化  $K$  均值的 WSN 分簇路由算法。在分簇阶段, 针对  $K$ -means 算法对初始聚类中心敏感问题, 使用 CS 算法优化  $K$ -means 算法的初始聚类中心, 再由  $K$ -means 算法对网络进行分簇; 簇首选举综合考虑了候选簇首的剩余能量、与聚类中心的距离, 并为各个因子添加权重, 根据影响因素动态调整因子权重, 使选出的簇首更合理; 数据传输阶段, 使用 CS 算法为簇首规划路由, 以均衡簇首负载。实验结果表明, CS- $K$  算法能够有效降低能耗, 延长网络生命周期。但是文中节点均为同构且静止部署, 今后工作将考虑在节点异构、移动的网络环境下进行无线传感器网络分簇路由研究。

### 参考文献:

- [1] SONG Y, GUI C, LU X C, et al. A genetic algorithm for energy-efficient based multipath routing in wireless sensor networks[J]. *Wireless Personal Communications*, 2015, 85(4): 2055-2066.
- [2] 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. Survey of clustering algorithms for wireless sensor networks[J]. *Computer Science*, 2017, 44(2): 31-37.
- [3] 李文峰, 沈连丰, 胡静. 传感器网络簇间通信自适应节能路由算法[J]. *通信学报*, 2012, 33(3): 10-19.  
LI W F, SHEN L F, HU J. Adaptive energy-saving routing algorithm for inter-clustering communications in WSN[J]. *Journal on Communications*, 2012, 33(3): 10-19.
- [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 B, CHANDRAKASAN A P, BALAKRISHNAN H. An application-specific protocol architecture for wireless micro-sensor networks[J]. *IEEE Transactions on Wireless Communications*, 2002, 1(4): 660-670.
- [6] RAZAQUE A, MUDIGULAM S, GAVINI K, et al. H-LEACH: hybrid-low energy adaptive clustering hierarchy for wireless sensor networks[C]//*Proceedings of 2016 IEEE Long Island Systems, Applications and Technology Conference (LISAT)*. Piscataway: IEEE Press, 2016: 1-4.
- [7] 李成法, 陈贵海, 叶懋, 等. 一种基于非均匀分簇的无线传感器网络路由协议[J]. *计算机学报*, 2007, 30(1): 27-36.  
LI C F, CHEN G H, YE M, et al. An uneven cluster-based routing protocol for wireless sensor networks[J]. *Chinese Journal of Computers*, 2007, 30(1): 27-36.
- [8] GANTASSI R, GOUISSEM B B, OTHMEN J B. Routing protocol LEACH-K using K-means algorithm in wireless sensor network[M]//*Advances in Intelligent Systems and Computing*. Cham: Springer International Publishing, 2020: 299-309.
- [9] 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.  
HUANG L X, WANG H, YUAN L Y, et al. Improved LEACH protocol algorithm for WSN based on energy balance and high efficiency[J]. *Journal on Communications*, 2017, 38(S2): 164-169.
- [10] 张顶, 张琳. 基于  $K$ -means 的 WSN 动态信任度双簇头选取算法[J]. *南京邮电大学学报(自然科学版)*, 2020, 40(2): 108-114.  
ZHANG D, ZHANG L. Dual cluster head selection algorithm based on  $K$ -means and dynamic trust degree in WSN[J]. *Journal of Nanjing University of Posts and Telecommunications (Natural Science Edition)*, 2020, 40(2): 108-114.
- [11] ERISOGLU M, CALIS N, SAKALLIOGLU S. A new algorithm for initial cluster centers in  $K$ -means algorithm[J]. *Pattern Recognition Letters*, 2011, 32(14): 1701-1705.
- [12] JANG S, KIM H Y, KIM N U, et al. Energy-efficient clustering scheme with concentric hierarchy[C]//*Proceedings of 2011 IEEE International RF & Microwave Conference*. Piscataway: IEEE Press, 2011: 79-82.
- [13] RAY A, DE D. Energy efficient clustering protocol based on  $K$ -means (EECPK-means)-midpoint algorithm for enhanced network lifetime in wireless sensor network[J]. *IET Wireless Sensor Systems*, 2016, 6(6): 181-191.
- [14] YANG X S, DEB S. Cuckoo search: recent advances and applications[J]. *Neural Computing and Applications*, 2014, 24(1): 169-174.
- [15] YANG X S, DEB S. Engineering optimisation by cuckoo search[J]. *International Journal of Mathematical Modelling and Numerical Optimisation*, 2010, 1(4): 330.
- [16] 董发志, 丁洪伟, 杨志军, 等. 基于遗传算法和模糊  $C$  均值聚类的 WSN 分簇路由算法[J]. *计算机应用*, 2019, 39(8): 2359-2365.  
DONG F Z, DING H W, YANG Z J, et al. WSN clustering routing algorithm based on genetic algorithm and fuzzy C-means clustering[J]. *Journal of Computer Applications*, 2019, 39(8): 2359-2365.
- [17] SHETA A F, SOLAIMAN B. Evolving clustering algorithms for wireless sensor networks with various radiation patterns to reduce energy consumption[C]//*Proceedings of 2015 Science and Information Conference (SAI)*. Piscataway: IEEE Press, 2015: 1037-1045.
- [18] 李道全, 张玉霞, 魏艳婷. 基于聚类分析的能耗均衡无线传感器网络分簇算法[J]. *计算机工程*, 2019, 45(10): 116-121.  
LI D Q, ZHANG Y X, WEI Y T. Energy consumption balanced clustering algorithm for wireless sensor network based on clustering analysis[J]. *Computer Engineering*, 2019, 45(10): 116-121.
- [19] KIMURA N, LATIFI S. A survey on data compression in wireless sensor networks[C]//*Proceedings of International Conference on Information Technology: Coding and Computing (ITCC'05) - Volume II*. Piscataway: IEEE Press, 2005: 8-13.

### [作者简介]



朱开磊 (1995- ), 男, 西安邮电大学硕士生, 主要研究方向为物联网技术及应用、路由协议。



孙爱晶 (1971- ), 女, 西安邮电大学教授, 主要研究方向为物联网技术及应用。