

基于内河无线传感网络的分簇路由算法

陈甜甜 江 冰 薛晓清 沙婷婷

(河海大学物联网工程学院 常州 213022)

摘 要 针对内河无线传感网络呈长距离双线性分布特点,提出的 RECR 算法可以有效地改善路由能耗消耗不均的问题。根据节点的分布的特点和节点离汇聚节点的距离来构造簇,根据簇中所有节点的能耗比来选择簇头,根据多汇聚节点和簇头线性分布的特点设计簇间路由。仿真结果表明,RECR 算法显著平衡了网中节点能耗,延长了网络生存时间。

关键词 无线传感网络,分簇,长距离,双线性

中图法分类号 TP393 **文献标识码** A

Clustering Routing Algorithm for Wireless Sensor Network along River

CHEN Tian-tian JIANG Bing XUE Xiao-qing SHA Ting-ting

(College of IOT Engineering, Hohai University, Changzhou 213022, China)

Abstract Aimed at the long-distance and bilinear-form topology of wireless sensor network along river, a new algorithm called RECR which can effectively improve the problem of uneven energy consumption was presented. The algorithm constructs clusters according to the characteristics of node distribution and distance from the sink node, and selects a cluster head from all the nodes in a cluster according to energy consumption ratio. Inter-cluster routing is designed based on the characteristics of multi-sink-node and bilinear-form topology. Simulation results show that RECR algorithm significantly balance the energy consumption of the network and prolong the network lifetime.

Keywords WSN, Clustering, Long distance, Bilinear form

近年来,人们饱受环境污染的恶果,使得环境问题越来越受重视,特别是与人们生产生活息息相关的水污染问题。我国水污染问题严峻,内河污染尤为突出,内河附近企业的偷排行为隐蔽,人力监管困难。因而,构建有效的内河水质监测系统具有重大的现实意义。传统的监测一般采用有线方式,其存在建设维护成本高、环境条件要求严格的问题^[1]。具有分布广、组网简单、监测能力强、系统成本低等特点的无线传感网络提供了一种对内河水质进行有效监测的手段。内河水质监测系统需要周期性地采集水质参数,且内河无线传感网络呈长距离双线性分布,能耗不均,数据冗余问题十分严重,因此,数据收集机制,特别是路由算法^[2]的设计至关重要。

1 相关研究

在无线传感器网络中,不仅要关心单个节点的能量消耗,还需要关心整个网络的能耗均衡,这样才能延长整个网络的生存时间。为此,一些学者提出了很多不同的解决方法,其中分簇^[3,4]是一种比较常用的解决方法。Heinzelman 等人提出了 LEACH(Low Energy Adaptive Clustering Hierarchy) 算法^[5-7],其基本思想是以循环的方式随机选择簇头节点,将整个网络的能量负载均匀分配到网络中的每个节点,从而达到

降低网络能耗、提高网络生存周期的目的。Stephanie 等人提出了 PEGASIS(Power-Efficient Gathering in Sensor Information Systems) 算法^[8-10],其基本思想是从网络中距离汇聚节点最远的节点开始,采用贪婪算法,将网络中的所有传感器节点形成一条链,使得节点在进行数据传输时,只需与它们最近的邻居节点通信即可。节点与 Sink 间的通信是轮流进行的,即每个节点都要当一次链首节点,基于这种轮换通信机制,可使得网络中的能量消耗统一分配到每一个节点上,均衡了网络能耗,有效延长了网络生命周期。

LEACH 算法、PEGASIS 算法和 NGECA 算法都是针对普通的无线传感网络的拓扑结构,而内河无线传感器网络是长距离的双线性拓扑结构,LEACH 算法假设每个簇头都能够直接和汇聚节点进行通信,而在狭长结构的无线传感器网络中这个假设却无法成立,PEGASIS 中虽然每个节点数据传输距离较短,能耗较低,但是在狭长的无线传感器网络中,PEGASIS 协议建立的链路较长,传输延迟大,无法满足环境监测这种实时性要求较高的应用需求,此外链首的负担较重,数据传输极易失效。因此 LEACH 和 PEGASIS 两种协议都不适用于双线性无线传感器网络中。其实,双线型的无线传感网络就是特别的带状无线传感网络。针对带状无线传感网

本文受江苏省水利科技项目(2013091)资助。

陈甜甜(1992-),女,硕士生,主要研究方向为现代通信系统及无线传感网络,E-mail:chentian104@163.com;江冰(1960-),女,教授,硕士生导师,主要研究方向为短距离通信及工业无线网络;薛晓清(1989-),女,硕士生,主要研究方向为现代通信系统及无线传感网络;沙婷婷(1991-),女,硕士生,主要研究方向为现代通信系统及无线传感网络。

络,文献[11]提出了PBCR(Position-based Chain Routing)算法,该算法根据地理位置信息将整个网络划分成若干个大小相等的条形区域,相邻的两个条形区域中的任意两个节点都可以直接通讯,条形区域内的所有节点组成了一个簇,区域内节点周期性地根据剩余能量和到汇聚节点的距离来选取簇头。PBCR算法把节点的剩余能量作为一个参量引入算法,使得选出的簇头更合适担当数据转发任务,簇内能耗更为均衡,但簇头采用多跳方式与汇聚节点通信时,某些簇头会因转发其它簇头的过路数据而负担过重,造成簇间能耗不均。文献[12]提出了CRLDB(Cluster-based Routing in Long Distance Band-type network)协议,该协议主要采用非均匀分簇的思想,引入备选簇首竞争半径的概念和相应的竞争策略,并加入最优簇首个数、节点剩余能量和周围邻居节点个数的簇首选择机制。虽然针对长距离带状网络的系列算法能改善由“热区”效应带来的能量消耗不均衡的问题,但同时在特殊网络和特定环境下缩短了网络生命周期。

本文针对内河水质监测无线传感网络拓扑结构双线型的特点、多个汇聚节点和传感器节点采集的数据的冗余程度,结合PEGASIS算法和PBCR算法的原理,提出了一种适用于内河无线传感网络的能量均衡的RECR(River-based Energy-balance Chain Routing)算法。

2 模型描述

2.1 网络模型

本文的被监测对象是内河河水的水质,传感器节点都安置在河流两岸,传感器节点都为固定节点,内河呈狭长的带状,即节点呈双线型分布,节点信息只能在狭小的带状区域沿着相对确定的路径依次传递。随着离汇聚节点的距离越来越近,节点的负载越重,远离汇聚节点的节点信息可能无法传输给汇聚节点,则网络中需多个汇聚节点。另外,河水的流动性使同岸的相邻节点所采集的数据具有很大的相似性,而对岸的节点采集的数据可能有所不同。根据内河网络的特征,提出其网络模型:

(1)网络所监测的区域为长方形区域,长度为 L ,宽为 W , $L \gg W$;

(2)网络由2个汇聚节点和 N 个普通节点所组成,汇聚节点具有足够的能量、处理能力,普通节点同构,每个普通节点具有相同的初始能量、计算和通信能力,能感知自己的剩余能量,具备数据融合功能,节点部署后其位置静止不变,节点均匀分布在监测区域的长边上;

(3)每个节点都有一个全网唯一的ID,还可通过接收到的信号强度估算出发送节点与自己的近似距离,并能够根据传输距离的长短调整发射功率的大小;

(4)相邻节点采集的数据相关性较大,且数据传输时,无数据丢失现象。

2.2 能耗模型

无线传感网络中,节点的能耗^[13]一般由数据发送能耗、数据接收能耗、数据处理能耗3部分构成。

假设节点X向相距 d 的节点Y传输 k 字节长度的数据,发送端节点X的能量消耗为:

$$E_{Tx}(k, d) = E_{Tx_e}(k) + E_{Tx_a}(k, d) = \begin{cases} kE_e + k\gamma_f d^2, & d < d_0 \\ kE_e + k\gamma_m d^4, & d \geq d_0 \end{cases} \quad (1)$$

式中, E_e 表示发送每比特数据时的能量消耗,节点的数据发送能耗 E_{Tx} 由发射电路损耗 E_{Tx_e} 和功率放大损耗 E_{Tx_a} 两部分构成,其中功率放大损耗根据发送端和接收端之间的距离 d 分别采用不同的无线通信模型,当 d 小于阈值 d_0 ($\sqrt{\gamma_f/\gamma_m}$)时,采用自由空间模型;反之,则采用多路衰减模型。 $\gamma_f d^2$, $\gamma_m d^4$ 分别是这两种功率放大模型发送每比特数据所需的能量。

接受端节点Y的能量消耗 $E_{Ry}(k) = kE_e$,其中 E_e 表示接受每比特数据时的能量消耗。

另外,假设节点Z进行 k 比特数据处理所消耗的能量 $E_{Dz}(k) = kE_d$,其中, E_d 表示融合单位比特数据耗费的能量^[14]。

3 RECR 算法描述

基于内河无线传感网络能量均衡的分簇路由算法由簇的形成、簇头竞选、簇内通信和簇间路由4个阶段组成。

3.1 簇的形成

RECR算法根据地理位置信息和同岸相邻节点采集的水质数据相关性较大的特点将整个网络划分成若干个大小相等的长为 l 、宽为 w ($w = W/2$)的虚拟格,这些虚拟格的位置相邻且不重叠,每个虚拟格的所有节点就组成了一个簇。

要保证簇头形成的链路能进行通信,即虚拟格中的节点能与周围虚拟格的任意节点都能够通信,这就需满足条件 $4w^2 + 4l^2 \leq R^2$,即 $l \leq \sqrt{d_0^2 - 4w^2}/2$, R 表示的是节点通信距离,根据通信距离与距离阈值 d_0 ,采用不同的通信模型以方便研究,这里的 R 就等于 d_0 。另外,由网络划分的虚拟格大小一样可知,监测区域的长 L 能被 l 整除。

RECR算法划分虚拟格的目的是构造簇,一个虚拟格对应一个簇。簇的形成如图1所示。

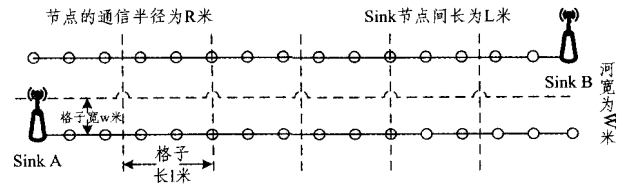


图1 网络分簇图

RECR算法中,簇形成后就不再改变,与LEACH这种每轮循环都要重新构造簇的算法相比,大大减少了构造簇的能量开销。

3.2 簇头竞选

由于簇内节点间的距离均小于距离阈值 d_0 ,根据式(1)可以得出,两节点的通信能耗与节点距离的平方成正比,则簇内通信能耗与簇头到所有簇内节点的距离的平方和成正比,本文将节点的剩余能量与簇内通信能耗之比定义为簇头能耗比,簇头能耗比的计算公式如下:

$$P(c_i, j) = \frac{E_r(j)}{\sum_{k=1}^{N(i)} D^2(j, k)} \quad (2)$$

式中, c_i 表示第 i 簇; $E_r(j)$ 表示第 i 簇中第 j 个节点的剩余能量; $D(j, k)$ 表示第 i 簇中节点 k 到节点 j 的距离; $N(i)$ 为第 i 簇中的节点个数。

各簇选择簇头能耗比最大的节点作为该簇的簇头。因为簇头能耗比最大的节点具有较大的剩余能量,并且与簇内其

他节点通信所需的能耗代价相对最小。网络中各簇选择能量较大且簇内通信代价较小的节点作为正式簇头节点,可以有效地降低簇内能耗。

3.3 簇内通信

各簇选出簇头后,给簇内节点分配固定的 TDMA 时隙,簇内节点在各自的时隙内向簇头发送监测数据。簇头接受簇内节点的传输数据,簇中节点的数据有较大的相似度,簇头将接受的簇内节点的数据和本身采集的数据进行融合,降低了路由节点因传输大量的数据引发的能量损失。

3.4 簇间路由

簇头需要将数据传给汇聚节点,在多汇聚节点的网络中,簇头需要将数据发送给最近的汇聚节点。通过比较节点到附近的汇聚节点的距离,确定节点数据传输方向后,选择下一跳的节点。在数据传输阶段,由簇形成的方式可知,簇头节点也是双线性分布的,也就是说,下一跳节点可以有两个选择,必须综合考虑下一跳节点的剩余能量以及通信代价,并以剩余能量与单位数据通信代价的能耗比作为衡量标准,簇头 i 下一跳节点的能耗比为:

$$P(c_i, c_j) = \frac{E_r(j)}{D^2(c_i, c_j)} \quad (3)$$

式中, $D^2(c_i, c_j)$ 是簇头节点 i 到下跳簇头节点 j 的距离的平方,由于节点间的通信能耗与距离的平方成正比, $D^2(c_i, c_j)$ 即代表节点的通信代价。

RECR 算法还考虑了链路不通的问题,若簇头 i 的数据传输方向上的同等级地位的两个节点都发生故障,导致无法形成到达最近汇聚节点的链路,簇头 i 可反向建立链路,保证数据的及时传输,避免了因链路问题对网络的影响,提高了数据传输的可靠性。

4 仿真与分析

4.1 网络环境仿真

运用 Matlab 平台对 RECR 路由算法进行仿真。取 $200\text{m} \times 10\text{m}$ 长方形区域作为监测区域,两个汇聚节点 Sink A 和 Sink B 的坐标分别是 $(0,0)$ 和 $(200,10)$,100 个传感器节点间隔 4m 均匀分布在长方形区域的长边上。为简化研究,规定节点的最低工作能量 E_{\min} 为 0,当节点的剩余能量 $E_r \leq E_{\min}$ 时,节点视为死亡。

4.2 网络性能仿真

本文在相同的多汇聚节点双线性无线传感网络下对 PEGASIS 算法、PBCR 算法和 RECR 算法进行了性能比较。

首先从网络中剩余节点与时间的关系方面进行研究,其仿真结果如图 2 所示。从图 2 可以很直观地看出,采用 RECR 算法的网络生存时间明显长于其它两个算法,为了更为准确地比较网络的生存时间,本文定义了两个时间节点,一个是网络中出现第一个死亡节点的时间轮数 T_{first} ,另一个是网络的死亡时间轮数 T_{end} ,即 85% 的节点死亡的时间轮数。一般情况下,85% 的节点死亡后,网络中剩余节点的能量将快速耗尽,网络已经无法完成设计的任务。3 种算法下,这两个时间节点的仿真结果如表 1 所列。

在无线传感网络中,从网络开始运行到第一个节点死亡的时间段称为网络运行的稳定期,即稳定期为 T_{first} ,网络的结构不断变化的时间段称为网络运行的不稳定期,即不稳定

期为 $T_{\text{end}} - T_{\text{first}}$ 。从表 1 的数据可以得出,RECR 的稳定期和不稳定期都是最长的,这说明 RECR 算法能很好地适应网络变化,具有很好的稳定性。

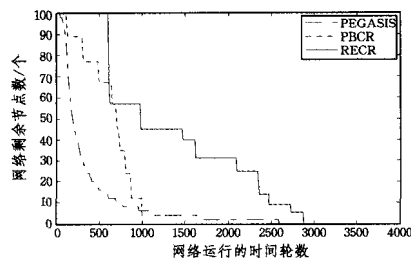


图 2 剩余节点与时间的关系

表 1 网络生存时间轮数比较(单位:轮)

路由算法	T_{first}	T_{end}	$T_{\text{end}} - T_{\text{first}}$
PEGASIS	52	550	498
PBCR	111	866	755
RECR	606	2349	1743

其次,图 3 给出了 3 种算法在网络中的剩余能量与时间的关系。根据 PEGASIS 算法的曲线可知,PEGASIS 能量消耗速度远快于 PBCR 算法和 RECR 算法,且能耗不均。PBCR 算法的曲线类似于一条斜线,能量消耗均衡,但与 RECR 算法相比,网络生存时间较短。在网络运行稳定期,RECR 能量消耗均衡且能耗速率最小;在网络运行不稳定期,RECR 能根据网络拓扑的变化调整能耗速率,保证网络的能量均衡。

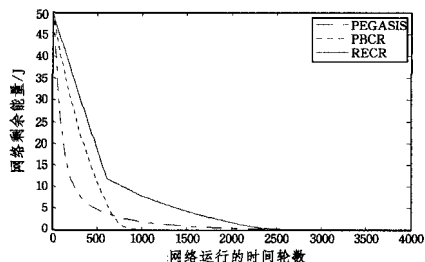


图 3 剩余能量与时间的关系图

结束语 本文在分析内河无线传感网络的拓扑结构和研究一些学者提出的路由算法的基础上,针对多汇聚节点的双线性无线传感网络能量不均、数据冗余的问题,设计了适用于内河水质监测无线传感网络的能量均衡的分簇路由算法,算法根据节点的地理位置、采集数据的相似度等因素构造了大小相等的簇,依据簇头能耗比选取簇头。簇间多跳路由根据簇头也呈双线性分布的特点设计,其不仅注重数据的传输方向、链路的能耗、节点的负载均衡,还注重链路的冗余。仿真结果表明,在多汇聚节点的双线性无线传感网络中,RECR 算法有效均衡了能量负载,明显优化了网络性能,从而延长了网络生存时间。但 RECR 算法还存在一定的不足,越靠近汇聚节点,其转发其他簇的数据而产生的能量消耗越大。解决这个问题可以从节点分布策略方面入手,即使簇离汇聚节点越远,节点分布密度越大,这样,越靠近汇聚节点的簇拥有越多的能量。RECR 算法结合节点分布策略可以使能量更为均衡。

参考文献

- [1] 刘国梅,金秋春,史军勇.一种适合长距离带状无线传感器网络的路由协议[J].科学技术与工程,2013,13(6):1661-1665

- [2] 吴迪,胡钢,倪刚. 无线传感器网络安全路由协议的研究[J]. 传感技术学报, 2008, 21(7): 1195-1201
- [3] 郭书城, 卢昱, 许定根. 基于分簇无线传感器网络的路由算法研究[J]. 通信学报, 2010, 31(8A): 63-69
- [4] 孙利民, 李建中, 陈渝, 等. 无线传感器网络[M]. 北京: 清华大学出版社, 2005: 94-96
- [5] Heinzelman W, Chandrakasan A, Balakrishnan H. Energy-Efficient Communication Protocol for Wireless[C]//Proceedings of the Hawaii Conference on System Sciences Microsensor Networks. 2000, 1: 3005-3014
- [6] 刘睿琼, 齐小刚, 孙正海. 基于节点位置的无线传感器网络分簇路由协议[J]. 电子科技, 2013, 26(3): 130-133
- [7] 李亚男, 徐夫田, 陈金鑫. 基于 LEACH 的 WSNs 分簇优化策略[J]. 传感技术学报, 2014, 27(5): 671-674
- [8] Lindsey S, Raghavendra C. PEGASIS: Power-efficient gathering in sensor information systems[J]. IEEE Aerospace Conference Proceedings, 2002, 3(10): 1125-1130
- [9] 黄飞, 金心宇, 张昱, 等. 基于 GASA 的能耗均衡 WSN 路由协议[J]. 传感技术学报, 2009, 22(4): 586-592
- [10] 张震, 闫连山, 潘炜. 基于 LEACH 和 PEGASIS 的簇头成链可靠路由协议研究[J]. 传感技术报, 2010, 23(8): 1173-1178
- [11] 乔钢柱. 基于无线传感器网络的煤矿安全综合监控系统设计与关键技术研究[D]. 兰州: 兰州理工大学, 2012: 77-83
- [12] 王伟. 长距离带状无线传感器网络路由协议设计[J]. 计算机工程, 2014, 40(3): 132-136
- [13] 胡钢, 谢冬梅, 吴元忠. 无线传感器网络路由协议 LEACH 的研究与改进[J]. 传感技术学报, 2007, 20(6): 1391-1396
- [14] 李鉴, 石鑫, 刘贺平. 煤矿巷道无线传感器网络非均匀分簇数据传送机制[J]. 地球科学-中国地质大学学报, 2013, 38(1): 195-200

(上接第 208 页)

4.2 图像分割实验

最后, 将算法用于图像分割实验中。图像分割^[12,13]就是将图像中各个物体按其自然属性区分开来。由于一般的彩色图像内容比较丰富, 采用无监督的像素值聚类算法对图像进行分割的效果不是很理想, 一般情况下需要人为提供一些先验信息来确定图中哪一块像素区属于这个物体而不属于另一个物体。这里利用本文提出的改进的谱聚类算法来进行图像分割。图 9(a)和图 9(c)是原始图像, 图 9(b)和图 9(d)是图像分割结果。

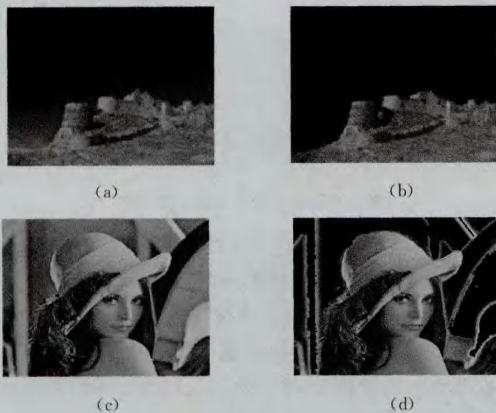


图 9 基于改进的谱聚类的图像分割效果

从图像分割实验可以看出, 本算法在大样本聚类中确实是有效的。

结束语 本文提出的算法对谱聚类算法的改进主要体现在对数据点间的相似度的定义上, 传统的基于路径的谱聚类算法主要是采用 Gauss 函数定义相似度, Gauss 函数中涉及了尺度参数, 尺度参数选择的好坏对聚类结果的好坏影响很大。但是目前又没有确定这个尺度参数的规则, 这一点给传统算法带来了一定的局限。因此我们通过二次规划的方法来鲁棒地确定数样本之间的相似度, 避免了尺度参数选择的不准确性的问题, 同时为了降低了计算路径相似度的时间复杂

度, 构造了最小生成树。最后通过 MNIST 手写数字库和图像分割实验验证了算法的有效性。

参考文献

- [1] 金慧珍, 赵江英, 刘博. 一种基于谱聚类的灰度图像分割法[J]. 计算机系统应用, 2009, 4(4): 74-76
- [2] 张向荣, 蒋晓雪, 焦李成. 基于免疫谱聚类的图像分割[J]. 软件学报, 2010(9): 2196-2205
- [3] 贾建华, 焦李成. 空间一致性约束谱聚类算法用于图像分割[J]. 红外与毫米波学报, 2010(1): 69-75
- [4] 王兴良, 王立宏, 武栓虎. 谱聚类中选取特征向量的动态选择集成方法[J]. 人工智能与模式识别, 2014, 27(5): 452-462
- [5] Andrew Y N, Jordan M, Weiss Y. On Spectral Clustering Analysis and an algorithm[C]//NIPS. 2002
- [6] Chang H, Yeung D Y. Robust path-based spectral clustering[J]. Pattern Recognition, 2008, 41: 191-203
- [7] 潘晓英, 刘芳, 焦李成. 密度敏感的多智能体进化聚类算法[J]. 软件学报, 2010, 21(10): 2420-2431
- [8] 孔万增, 孙昌思, 张建海, 等. 近邻自适应局部尺度的谱聚类算法[J]. 中国图象图形学报, 2012, 17(4): 523-529
- [9] Fischer B, Buhmann J M. Bagging for Path-Based Clustering[J]. IEEE Transaction on Pattern Analysis and Machine Intelligence, 2003, 25(11): 1411-1415
- [10] Fischer B, Buhmann J M. Path-Based Clustering for Grouping Smooth Curves and Texture Segmentation[J]. IEEE Transaction on Pattern Analysis and Machine Intelligence, 2003, 25(4): 513-518
- [11] Wang F, Zhang C. Label Propagation through Linear Neighborhoods[C]// International Conference on Machine Learning. Pittsburgh, 2006
- [12] Ley J D F, van Dam A, Feiner S K, et al. Computer graphics: principles and practice[M]. New York: Addison-Wesley Publishing, 1998
- [13] 章毓晋. 图像处理和分析[M]. 北京: 清华大学出版社, 1999