

基于并行SOM和自适应禁忌优化算法的 负载均衡分簇路由设计

刘珂, 陈中良

(黄淮学院 信息工程学院, 河南 驻马店 463000)

摘要: 为了克服传统分簇路由协议中无法有效缓解“热区”和数据传输可靠性不高的问题, 设计一种基于并行SOM和自适应禁忌优化算法的负载均衡路由协议。首先, 将整个网络区域分为若干子区域, 在每个子区域中自适应的根据节点与基站的距离确定分簇数目, 采用并行SOM并行对网络区域进行分簇, 同时根据能量和位置因素选举簇头; 然后, 以最小化各簇头到基站的最小跳数为目标, 采用改进的禁忌优化算法来寻找各簇头到达基站的最优解和次优解, 构成多条冗余路由, 以增强数据传输的可靠性, 最后, 对基于改进禁忌优化算法的最小跳数路由协议进行了设计和描述。仿真实验表明: 文中方法能有效地提高网络生命周期和提高数据传输的可靠性, 具有很好的负载均衡能力, 是一种适用于WSN的有效分簇路由算法。

关键词: 并行自组织映射; 分簇路由; 禁忌优化算法; 传感器

中图分类号: TP393

文献标识码: A

文章编号: 1001-7119(2015)03-0211-05

DOI:10.13774/j.cnki.kjtb.2015.03.051

Design of Clustering Routing in Wireless Sensor Networks with Load Balance Based on Parallel SOM and Adaptive Tabu Optimal Algorithm

Liu Ke, Chen Zhongliang

(Huanghuai University Henan, Henan Zhumadian 463000, China)

Abstract: In order to conquer the traditional clustering routing protocol having the problems such as hotspot and low efficient data transportation, a clustering routing protocol based on parallel SOM and adaptive tabu optimal algorithm is proposed. Firstly, the network is divided into several areas, and the number of clusters is decided by the distance with the sink node, using the parallel SOM to form the clusters, and the cluster head is selected by the energy and positions factors, then the minimal distance between every cluster head to the sink node is used as the goal, and the improved tabu optimal algorithm is used to construct the optimal solution and second-optimal solution to obtain multi-routing and improve the data transporting feasibility. Finally, the algorithm for the minimal routing based on improved tabu is designed and described. The simulation experiment shows the method in this paper has the longer network life circle, and compared with the other methods, it has good load balance ability, therefore, it is an effective routing algorithm suiting WSN.

Keywords: parallel SOM; clustering routing; tabu optimal algorithm; sensor node

收稿日期: 2014-04-21

基金项目: 河南省科技攻关计划资助项目(122102310474)。

作者简介: 刘珂(1980-), 男, 河南新蔡县人, 硕士, 讲师, 主要从事计算机网络应用方向研究。

E-mail: lndkitl@163.com。

0 引言

无线传感器网络(wireless sensor network, WSN)^[1,2]是一种集成信息采集、处理和传输的通信技术,是由采集和感知数据的传感器节点通过无线自组织的方式组成,目前在环境监控、生物医疗和灾难预警等方面都有着非常广阔的应用前景^[3]。由于WSN通常部署在恶劣的无人能及的区域周期性地对环境监测,而传感器节点是采用电池供电,当传感器节点因为电池耗尽而发生故障时,无法继续完成数据的感知和采集。因此,能源问题是制约传感器网络进一步发展和应用的重要问题,如何在节点能源有限的条件下,设计一个能延长无线传感器网络的生命周期的路由协议是WSN发展的首要问题^[4,5]。

层次型路由协议相对传统的平面型路由协议,具有吞吐量大、能耗低等优点^[6],其中,经典的分簇层次型路由协议LEACH协议(low-energy adaptive clustering hierarchy)^[7]采用数据传输局部控制技术,能满足大范围数据实时采集、传输和分析处理的要求。但由于其簇头选举中未考虑能量因素从而会造成簇内能耗不均衡,同时由于簇头到基站之间的单跳传输机制,使得距离基站较近的节点过早死亡,导致了“热区”的出现。

为了对LEACH协议进行进一步改进,文献[8]设计了一种基于多目标和声搜索的分簇路由算法,以节点与基站距离、各簇之间的距离以及最小化网络总能耗为目标,通过参数自适应的二进制和声搜索算法实现网络的路由优化。文献[9]设计了一种基于粒子群优化的分簇路由协议,根据节点与汇聚节点之间的距离计算竞争半径自适应构造大小不同的簇,在选择簇间的多跳路由时,将能耗较高的簇头作为下一跳节点。文献[10]设计了一种基于权值机制的分簇路由算法,采用基于权值的局部竞选簇首策略,并根据距离信息来构建多个簇,簇成员节点通过链式结构向簇头传递数据,簇头之间采用多跳传输策略向基站发送数据以改善“热区”问题。

上述工作均研究了分簇路由协议,具有重要意义,但是在生成簇时,没有根据距离基站的距离自适应地创建规模大小不同的簇,以最大程度地缓解“热区”问题,并且在建立了簇头到基站的多跳传输路由时,仅传输了一条路径,当路径中

的任何节点发生故障时,不能自适应地调整构成新的路径。为此,文中设计了一种基于SOM算法和禁忌优化算法的无线传感器路由协议,并通过实验证明了文中方法的有效性。

1 网络拓扑结构和能耗模型

文中设计的无线传感器网络部署在一个矩形的监测区域中,每个节点物理结构相同,具有相同的处理、存储和通信能力,节点位置固定且具有唯一的标志符,每个节点都可以根据与目标节点之间的距离自适应地调整通信半径,节点不具备定位设备,但可以根据与其它节点协作提供的信息来对自身位置进行定位,基站位于监测区域外且位置固定。由于文中方法要有效解决热区问题,因此,将网络区域中的传感器节点根据与基站的距离分为大小不同的簇,越靠近基站的节点由于要承担更多的簇间数据转发的任务,因此,规模应更小。文中设计的分簇路由协议拓扑结构如图1所示:

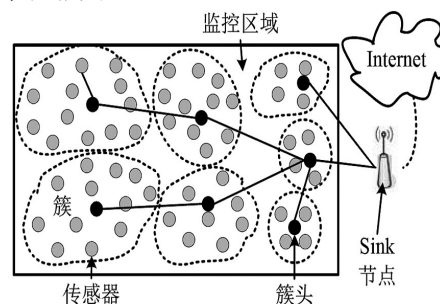


图1 分簇拓扑结构

Fig.1 Cluster topological structure

假设发送节点发送 h 比特数据至接收节点的数据发送耗能 $E_{Tx}(h,d)$,其可以表示为:

$$E_{Tx}(h,d) = \begin{cases} h * (E_{elec} + \varepsilon_{fmp} * d^2), & d < d_0 \\ h * (E_{elec} + \varepsilon_{amp} * d^4), & d \geq d_0 \end{cases} \quad (1)$$

接收节点接收 h 比特数据的数据接收耗能 E_{rec} 表示如下:

$$E_{rec} = h * E_{elec} \quad (2)$$

在式(1)和式(2)中, E_{elec} 为单位比特数据对应的发送/接收能耗, d 为发送节点和接收节点之间的距离, ε_{fmp} 为自由空间模型所对应的功率放大电路功耗系数, ε_{amp} 为多路衰减模型对应的功率放大电路功耗系数, d_0 为距离阈值,可通过公式(3)进行计算:

$$d_0 = \sqrt{\varepsilon_{jnp} / \varepsilon_{amp}} \quad (3)$$

2 基于并行SOM的分簇和簇头选举

2.1 原理描述

为了有效地解决热区问题,将整个网络区域划分为若干子区域,假设网络区域的宽度为 L ,子区域的个数为 n ,每个子区域的宽度可以设置为 L/n ,则可以均匀地将各子区域的宽度设置为 $0 \leq d < L/n$ 、 $L/n \leq d < 2L/n$ 和 $(n-1)L/n \leq d < L$,在 n 个区域中通过并行地运行SOM算法将每个子区域划分为多个簇,为了解决“热区问题”,每个子区域中簇的个数各不相同,靠近Sink节点的簇数目较多,每个簇的节点较少,远离基站的子区域中簇的数目较少,每个簇中节点的数目较多,每个子区域中簇的个数与其边界的关系可以表示为:

$$k = \frac{k_0 N}{d_{\max}} \quad (4)$$

其中, k_0 为初始设定的各子区域的最大簇数, d_{\max} 为当前子区域的宽度的最大值, N 为该区域的节点总数。

在每个区域中运行一个SOM算法进行分簇,因此对于 n 个子区域就运行 n 个SOM算法进行并行分簇,每个区域中的簇的数目可以根据式(4)获取,因此SOM算法的输出神经元的个数即为式(4)所示。

2.2 基于并行SOM的分簇算法

自组织映射网络(self-organizing feature map, SOM)^[10]是由Kohonen提出的一种通过竞争、合作和权值调节来实现无监督训练的机器学习方法,其组成结构如下所示:

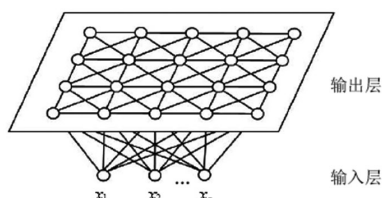


图2 SOM结构图

Fig.2 SOM structure

从图2中可以看出,SOM自组织映射结构图是由输入层和输出层两部分组成,输入层的节点个数为节点坐标维数为2,输出层的节点个数为我们划分的簇的总数,可以根据式(4)获得,则采

用并行SOM算法对子区域进行分簇的过程可以描述为:

算法1 并行SOM分簇算法

初始化:根据网络区域划分的子区域数,初始化 n 个SOM网络,每个输入神经元个数 $n=2$ (节点位置的坐标维度),输出神经元个数 k (分簇的个数)根据式(4)计算,当前迭代次数 t ,最大迭代次数 T ,初始化各SOM网络中的学习速率因子 $\psi(1)$,邻域半径 $N(1)$,以及初始权值 $W_{ij}(1)$;

步骤1:归一化所有样本数据和输入层神经元与输出神经元之间连接的权值 $W_{ij}(1)$,将其变换到 $(0,1)$ 之间;

步骤2:将每个样本数据输入到初始化的SOM网络:计算其与各个输出层神经元之间的距离 $D_{ij}(t)$:

$$D_{ij}(t) = \sum_i^n (w_{ij}(t) - x_i(t))^2 \quad j=1,2,\dots,m \quad (5)$$

从式(5)中可以看出, $D_{ij}(t)$ 是通过计算样本数据与各神经元连接权值之间的欧氏距离来获得。

步骤3:将根据式(5)计算得到的节点与各输出神经元的距离代入式(2),将具有最小 $D_{ij}(t)$ 的神经元作为获胜神经元 j :

$$j = \arg \min_j D_{ij}(t) \quad (6)$$

步骤4:对输出层获胜神经元连接的邻域范围的神经元权值进行调整,如下所示:

$$w_{ij}(t+1) = w_{ij}(t) + \psi(t)(x_i - w_{ij}(t)) \quad (7)$$

步骤5:根据式(8)对输出层获胜神经元的邻域及学习率进行更新:

$$\begin{cases} \psi(t) = \psi(t)(1 - t/T) \\ N(t) = N(1)(1 - t/T) \end{cases} \quad (8)$$

步骤6:判断下列条件之一是否成立:

- (1) 所有的SOM的学习率 $\psi(t)$ 下降为0;
- (2) 所有SOM的当前迭代次数达到最大值

T ;

如果上面条件之一成立,则算法结束,输出每个子区域的每个簇以及对应的簇成员节点;

否则 $t=t+1$,并返回步骤2继续开始迭代。

2.3 簇头选举

当各子区域中的簇形成后,将各簇位于较为中心的位置且能量最大的节点作为簇头,即根据下式选举能量较大的节点作为簇:

$$\Phi_i = \alpha E_i + \beta \frac{1}{\sum_{j \neq i} d} \quad (9)$$

从式(9)中可以看出, Φ_i 值越大其能量就越高,同时距离簇中其他节点的距离就越短,因此将具有最大式(9)值的节点作为簇头,同时当 Φ_i 的值低于某阈值时,不再选取该节点作为簇头,此时,重新计算所有节点的 Φ 值,将具有最大 Φ 值的节点作为新的簇头。

3 基于禁忌优化算法的簇间多跳路由

3.1 基于改进禁忌优化算法的多跳路由算法原理

文中基于改进禁忌优化算法的多跳路由设计的主要思想为:当所有子区域的簇和簇头均已生成后,采用改进的禁忌优化算法来获取各节点到基站的最小跳数路由,将所有的簇头节点按照其所在的子区域与基站的距离,从远到近依次编码,如下所示:

$$(x_1, x_2, \dots, x_m) \quad (10)$$

其中, m 为簇头节点总数目, $x_i = 1$ 表示第 i 个节点为最小跳数路由上的转发节点,适应度函数定义为:

$$f(t) = 1/\text{Count}_{\text{sink}} \quad (11)$$

其中, $\text{Count}_{\text{sink}}$ 为从簇头到基站的最小跳数路由长度。

采用禁忌算法寻求最优原子集的流程可以描述为图3。

3.2 算法设计流程

文中基于改进禁忌优化算法可以描述为:

初始化:当前迭代次数 $t=1$, 禁忌表 l , 禁忌长度 len , 禁忌表的大小 l_s , 邻域大小 l_n , 最大迭代次数 T ;

步骤1:根据式(10)所示编码方式随机初始化初始解,并将初始解赋为当前解;

步骤2:寻求当前解的邻域:

$$\begin{aligned} (x_1, x_2, \dots, x_m) &= (0, 1, 0, 1, 0, 1, 1, \dots, 0) \\ \downarrow \\ (x_1, x_2, \dots, x_m) &= (1, 1, 1, 0, 0, 0, 0, \dots, 1) \end{aligned} \quad (12)$$

根据式(12)求取当前解的 l_n 个邻域解,将这些邻域解作为候选解。

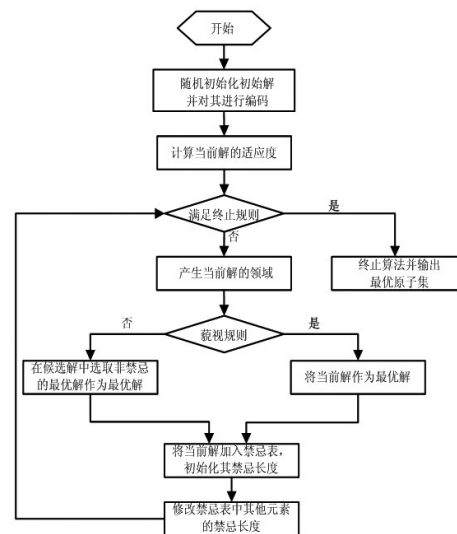


图3 改进禁忌优化算法获取最优解

Fig.3 Improved TS algorithm for getting the optimal solution

步骤3:将所有候选解代入式(11),并计算所有候选解适应度,并按适应度从大到小进行排序,选择具有最优适应度的解作为候选解;

步骤4:判断当前的最优适应度候选解,是否满足禁忌准则:即当前解虽然出现在禁忌表中,但其较以往的任何状态更优:

如果满足,则无视其禁忌属性,选择最优适应度候选解为当前解,并将其加入到禁忌表 l 中,初始化其禁忌长度为最大长度 p ;

否则,将未被禁忌的最优适应度候选解作为当前解加入到禁忌表中,初始化其禁忌长度为最大长度 p ;

步骤5 将禁忌表中的其它禁忌解的禁忌长度-1,此时,将禁忌表中禁忌长度为0的禁忌解移出禁忌表 l ;

步骤6:判断迭代次数是否已经达到最大值 T ;

如果达到最大值,则算法结束,输出历史最优解,将禁忌表中的其他被禁忌的禁忌解作为多个次优解。

否则,当前迭代次数 $t=t+1$,并返回步骤2继续迭代。

输出最优解和多个次优解的目标是构成多条从簇头到基站的冗余路由,当最优路径发生故障后,仍然可以沿着其他次优解所示的路径进行数据传输。

4 仿真实验

在NS2仿真环境下对文中基于并行SOM和改进禁忌优化算法的分簇协议进行仿真和评价。监测区域的无线传感器网络参数设置如下:网络区域为1000 m×1000 m,基站位置(600 m, 550 m),节点个数为800,发送接收数据能量消耗 $E_{elec}=60$ nJ/bit,数据融合能量耗费 $E_{da}=505$ nJ/bit,节点初始能量 $E_0=2$ J,距离阈值 $d_0=80$ m,自由空间模型 $\varepsilon_{fmp}=30$ pJ/(bit·m²),多路衰减模型 $\varepsilon_{amp}=0.0015$ pJ/(bit·m²),数据包大小2000 bit。

分簇时采用的并行SOM算法参数设置如下:SOM网络个数 $n=4$,每个输入神经元个数 $n=2$,输出神经元个数 k (分簇的个数)根据式(4)计算,当前迭代次数 $t=1$,最大迭代次数 $T=100$,初始化各SOM网络中的学习速率因子 $\psi(l)=0.18$,邻域半径 $N(l)=3$,以及初始权值 $W_{ij}(l)=(1,0)$;

最小跳数路由的改进禁忌算法参数设置如下:当前迭代次数 $t=1$,禁忌表 l ,禁忌长度 $len=5$,禁忌表的大小 $l_s=5$,邻域大小 $l_n=8$;

为了验证文中方法的有效性和优越性,将文中方法与经典LEACH协议以及文献[8]所示的基于权值的改进分簇算法进行比较,从网络生命周期和基站接收的信息量两方面比较,结果如下。

4.1 网络生命周期对比

仿真得到的存活节点数随仿真时间变化的曲线如图4所示。

从图4中可以看出,文中方法、经典的层次型分簇协议LEACH协议、和文献[8]方法的网络生命周期分别为280轮、205轮和140轮,

文中方法较其它两种方法对应的网络生命周期分别提高了36.6%和100%,且文中方法出现的第一个死亡节点的时间远比其它方法要晚,同时文中方法的仿真曲线较其它两种方法更为平滑,这说明文中方法负载均衡,能较为平均地利用所有节点,具有很好的负载均衡性能。

4.2 数据传输可靠性对比

图5为三种方法在网络运行了300轮后基站节点收到的信息量的对比。从图5可以明显地看出文中方法基站在仿真期间接收的信息量远大于其它两种方法,约为762 kByte,而文献[8]方法接收到的总信息量为603 kByte,LEACH协议最少为309 kByte,文中方法中基站接收的总信息量较文献[8]方法和LEACH协议分别高26.7%和146.6%,显然,文中方法的数据传输可靠性最高,

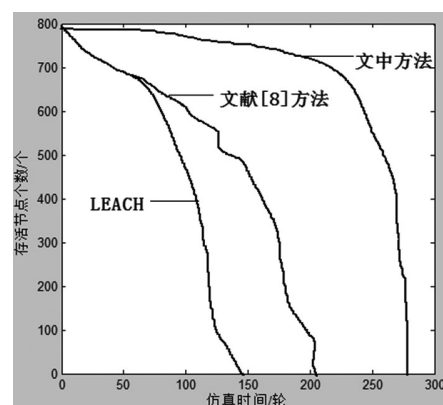


图4 网络生命周期对比

Fig.4 Comparison of network life cycle

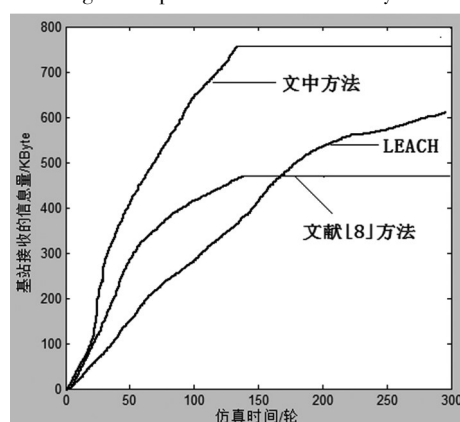


图5 基站接收信息包对比

Fig.5 Comparison of information received from Sink node

基站接收的信息最多,这是因为文中方法不仅能有效地均衡所有节点的能量开销,而且在选举簇头时考虑了能量因素,并设置了多条从簇头到基站的路由,避免“热区”问题,而LEACH协议采用从簇头到基站的单跳路由,同时选举簇头时未考虑能量因素,因此导致大部分的簇头的过早死亡,而文献[8]方法虽然能在一定程度上均衡节点负载,但没有建立簇头到基站的多条冗余路径,因此,一旦路径发生故障,数据传输的可靠性大大下降。

参考文献:

- [1] 钱志鸿,朱爽,王雪.基于分簇机制的ZigBee混合路由能量优化算法[J].计算机学报,2013,3(36):485-492.
- [2] Huang L C, Chang H C, Chen C C, Kuo C C. A ZigBee based monitoring and protection system for building electrical safety. Energy and Buildings, 2011, 43(6): 1418-1426.
- [3] Alemdar H, Ersoy C. Wireless sensor networks for health-

(下转第219页)

表2 潜在威胁样本复杂度较高时实验结果

Table 2 Potential threat sample complexity is higher when the experimental results

实验次数	精确度/%	
	传统方法	本文方法
1	87	90
2	88	89
3	83	91
4	78	92
5	79	92
6	82	91
7	83	90
8	80	92
9	76	88
10	78	89

按照上述实验结果可得,采用本文方法对潜在威胁下的网络数据进行挖掘,可有效处理传统方法无法分析潜在威胁数据关联性的缺陷,本文方法的数据挖掘精确度比传统方法平均高12%左右,具有重要的应用前景。

3 结论

本文提出一种基于区域频繁扩展树的支持向量机回归算法,用于挖掘网络潜在威胁数据,构建潜在威胁下网络数据的区域频繁扩展树,获取潜网络数据间的关联性,依据潜威胁下网络数据的关联性塑造基于支持向量机回归网络潜在威胁数据挖掘模型,模型利用支持向量回归估计模型分析数据回归曲线的平稳度,对比回归预测值同真实值间残差,对潜在威胁数据进行识别,完成潜在威胁下网络数据地准确挖掘。实验结果说明,采用所提方法对不同复杂度潜在威胁下的网络数据进行挖掘,具有较高精确度。

参考文献:

- [1] 周丽娟,石倩,葛学彬等.基于聚类的模糊遗传挖掘算法的研究[J].计算机工程与应用,2010,46(13):118-121.
- [2] 范明,王秉政.一种直接在Trans-树中挖掘频繁模式的新算法[J].计算机科学,2003,8:117-123.
- [3] 张净,孙志挥,等.基于信息论的高维海量数据离群点挖掘[J].计算机科学,2011,38(7):148-161.
- [4] 庄艳,陈继明,徐丹,等.基于 multi-iagents 系统的分布式数据挖掘[J].计算机科学,2007,34(12):163-167.
- [5] 司凤琪,徐治皋.基于自联想神经网络的测量数据自校正检验方法[J].中国电机工程学报,2002,22(6):152-155.
- [6] care: A survey[J].Computer Networks,2010,54(15):2688-2710.
- [7] 蒋鹏,阮斌锋.基于分簇的水下传感器网络覆盖保持路由算法[J].电子学报,2013,41(10):2067-2073.
- [8] Xu H Y.Construction of energy-aware ZigBee network [C]//Proceedings of the 2011 10th International Symposium on Distributed Computing and Applications to Business, Engineering and Science(DCABES 2011).Wuxi, Jiangsu, China, 2011:378-381
- [9] Heinzelman W,Chandrakasan A,Balakrishnan H. Energy-efficient communication protocol for wireless microsensor networks[C]//Proceedings of the 33rd Hawaii International Conference on System Science,2000.
- [10] 乔学工,王哲,王华倩,高绍斌.基于权值的非均匀分簇路由算法[J].2014,1(27):107-112.
- [11] 邹杰,史长琼,姬文燕.基于粒子群优化的非均匀分簇路由算法[J].计算机应用,2012,32(1):131-133.
- [12] 李明,曹晓莉,胡卫军.基于多目标和声搜索的无线传感器网络分簇路由算法[J].仪器仪表学报,2014,1(35):162-168.

(上接第215页)