

基于模糊等价矩阵的无线传感器聚类分区算法

李 佺 佺 李 成 海 郭 新 鹏

(空军工程大学导弹学院 三原 713800)

摘 要 基于模糊等价矩阵的聚类方法是模糊聚类中一种经典的分析方法。首次将其引入无线传感器节点分区的应用中。该聚类分区算法通过计算节点间的 Euclid 距离、分析其相关性、形成模糊等价矩阵、进行节点分区,实现了对该分区算法的应用设计。分析了其算法时间复杂度,并利用 Matlab 软件完成了算法仿真。仿真结果显示,该算法可以根据其疏密程度的不同很好地将无线传感节点分成不同区域。

关键词 模糊聚类,无线传感器,相关性,模糊等价矩阵

Clustering Subsection Algorithm Based on Fuzzy Equivalence Matrix for Wireless Sensors

LI Quan-quan LI Cheng-hai GUO Xin-peng

(School of Missile Air Force Engineering University, Sanyuan 713800, China)

Abstract The clustering method based on fuzzy equivalence matrix is one of the classical methods of fuzzy clustering. We applied it to assort wireless sensors for the first time. The algorithm analyzes the pertinence of two sensors with the Euclid distance of them, and then acquires the fuzzy equivalence matrix for subsection. We designed the algorithm, analyzed the time complexity of it. And the algorithm was simulated on Matlab. The results of simulation show the wireless sensors are assorted several classes by different distributing density with the algorithm.

Keywords Fuzzy clustering, Wireless sensor, Pertinence, Fuzzy equivalence matrix

1 引言

在无线传感网络中,通过均衡无线传感节点的能量消耗来延长网络寿命成为人们研究的热点,例如 LEACH^[1]、TEEN^[2]、HEED^[3]等算法。这些算法通过更换簇头的方式来均衡节点的能量消耗,这对延长网络寿命起到积极的作用。但是由于簇头的选取具有随机性,在节点分布非均匀的情况下,可能会有簇头节点过于集中或边缘节点成为簇头节点的情况出现。因此,有人提出了通过传感器节点分类实现区域划分的思想。在各个区域选取簇头可以有效地避免出现上述的两个问题。目前,人们研究比较多的是 k -均值聚类分区^[4,5]方法。

k -均值聚类方法也是一种常见的模糊分类方法,其基本原理是:先把数据分成 k 个部分(这里的 k 由用户确定),对每个部分给出一个聚类中心(聚类中心倾向于标记每一类的中心位置)。这样的初始化不一定正确,通过算法反复地修正聚类中心到“正确的”位置。这个迭代是基于最小化一个目标函数,该函数代表任何数据点到相应聚类中心的距离。直到最后每一类的数据点到其相应的聚类中心的距离加权达到最小,数据也就被成功地分成了 k 类。简而言之, k -均值聚类算法就是先找出每类的中心节点,其他节点以到各中心的距离来判断其从属于哪一类。

图 1—图 3 所示为无线传感器节点分布。图 1、图 2 是根据 k -均值聚类分区算法得到的结果,在图中节点被分为两个

区。节点 1、节点 2 离两个聚类中心较远。带箭头的线代表网络路由的路径和方向。在图 3 中,节点 1、节点 2 分别被单独分成一个区域。从图中可以看出,在图 3 分区基础上的路由明显优于图 1、图 2 的。为了实现图 3 所示的节点分区,提出了基于模糊等价矩阵^[6]的无线传感器聚类分区算法。

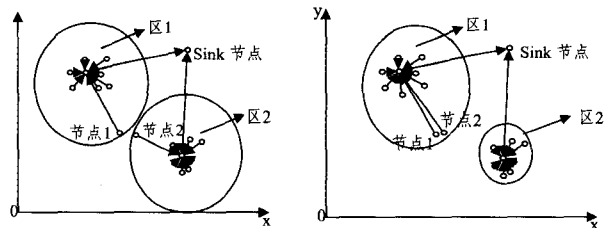


图 1

图 2

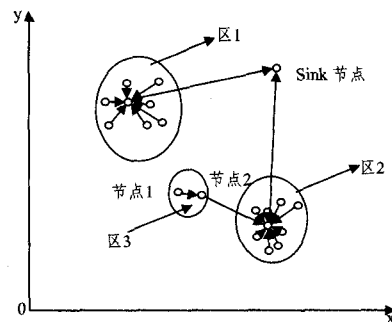


图 3

李佺佺(1990—),男,硕士生,主要研究方向为无线传感器网络、网络信息安全等,E-mail:quanabc12345@126.com;李成海(1966—),男,副教授,主要研究方向为嵌入式系统、无线网络的安全与防护;郭新鹏(1984—),男,讲师,主要研究方向为计算机网络、智能信息处理。

2 算法设计

2.1 算法的基本原理

一个合适的分类应当具备下列 3 个条件:

- (1) 自反性: 任何一个对象必须和自己同类;
- (2) 对称性: 若对象 u 与对象 v 同类, 则 v 与 u 也同类;
- (3) 传递性: 若对象 u 与对象 v 同类, 而 v 与 w 同类, 则 u 与 w 也应同类。

满足上述 3 个条件的关系, 即为一个等价关系。我们的模糊分类算法就是根据模糊等价算法关系进行的。

区域中有 n 个无线传感器节点。所有节点作为分类对象的集合 $\{u_1, u_2, \dots, u_n\}$ 。每个对象 X_i 有两个特性指标 (即反应对象特征的主要指标)——节点的 X, Y 坐标。 n 个对象的所有特性指标构成一个矩阵 U 。

$$U = \begin{bmatrix} u_{11} & u_{12} \\ u_{21} & u_{22} \\ \vdots & \vdots \\ u_{n1} & u_{n2} \end{bmatrix} \quad (1)$$

式中, u_{i1}, u_{i2} 分别为节点 u_i 的 X, Y 坐标。依据坐标之间数据的相关性, 对传感器节点进行分区。具体过程如下。

2.2 数据标准化处理

常用的数据标准化处理方法有: 均值规格化、中心规格化、最大值规格化、极差规格化、对数规格化等。这里采用最大值规格化对矩阵 U 做变换, 得到矩阵 U' , 其变换公式如下:

$$u'_{ij} = \frac{u_{ij}}{\text{MAX}(u_{ij}, \dots, u_{nj})}, i=1, 2, \dots, n; j=1, 2 \quad (2)$$

数据标准化处理之后, $u'_{ij} \in (0, 1]$ 。

2.3 建立模糊相似性矩阵

在有限论域中, 一个模糊关系和一个模糊矩阵建立一一对应关系。设所建立的模糊相似矩阵 $R = (r_{ij})_{n \times n}$, 统计量 r_{ij} 是分类对象间相似程度的定量描述。采用距离法来构造模糊矩阵 R 。距离法利用两对象间的距离来确定其相似程度 r_{ij} , 这是因为距离越大, r_{ij} 就越小。 r_{ij} 值的大小可以由式(3)求得:

$$r_{ij} = 1 - c(d(u'_i, u'_j))^\alpha \quad (3)$$

式中, c, α 是两个适当选取的正数, 使 $r_{ij} \in [0, 1]$; $d(u'_i, u'_j)$ 是两个对象间的 Euclid 距离, $d(u'_i, u'_j) = (\sum_{k=1}^2 (u'_{ik} - u'_{jk})^2)^{\frac{1}{2}}$ 。

2.4 建立模糊等价矩阵

实际上利用标准化值和式(3)所建立的矩阵 R , 只是一个模糊相似矩阵, 通常不满足传递性。为了进行分类, 需要采用矩阵理论中的传递闭包法, 运用“min-max”运算法则进行合成运算, 将 R 改造成模糊等价矩阵 R^* 。

利用平方合成方法计算传递闭包 $tr(R)$, 即 $R \Rightarrow R^2 \Rightarrow R^4 \Rightarrow \dots \Rightarrow R^{2^k} = R^{2^{(k+1)}}$, 其中 $k \leq \lceil \log_2 n \rceil + 1$ 。有 $tr(R) = R^{2^k}$, 即 $R^* = tr(R) = R^{2^k}$ 。模糊矩阵的乘法与普通矩阵的乘法相比较, 其运算过程一样, 只不过是把普通加法改成 \vee (逻辑加, 二者取最大), 把普通乘法改成 \wedge (逻辑乘, 二者取最小)。

2.5 动态聚类

根据上述的模糊等价矩阵 R^* , 利用“ λ -水平截集法”从低到高取不同的 $\lambda (\lambda \in [0, 1])$ 值, 在适当的阈值 λ 上进行截取, 每取一个 λ 值, 对矩阵 R^* 中的元素按式(4)进行一次代换得到矩阵 $R^*(\lambda)$ 。

$$r_{ij}^*(\lambda) = \begin{cases} 0, & r_{ij} < \lambda \\ 1, & r_{ij} \geq \lambda \end{cases} \quad (4)$$

在矩阵 $R^*(\lambda)$ 中将元素相同的行聚成一类, 依次取不同的 λ 值, 可以逐步实现由细分到粗分的聚类。当 λ 取值较小时, 得到粗分的类, 分的类数较少。当 λ 取较大的值时, 得到细分的类, 分的类数较多。

3 算法实例

设有 5 个传感器节点, 其坐标分别为 $u_1(80, 10), u_2(50, 1), u_3(90, 6), u_4(40, 5), u_5(10, 1)$ 。由节点的坐标值形成矩阵 U ,

$$U = \begin{bmatrix} 80 & 10 \\ 50 & 1 \\ 90 & 6 \\ 40 & 5 \\ 10 & 1 \end{bmatrix}$$

① 用式(2)对矩阵 U 进行数据标准化处理, 结果如下:

$$U' = \begin{bmatrix} 0.89 & 1 \\ 0.56 & 0.10 \\ 1 & 0.60 \\ 0.44 & 0.50 \\ 0.11 & 0.10 \end{bmatrix}$$

② 用式(3)和矩阵 U' 建立模糊性矩阵 R , 结果如下:

$$R = \begin{bmatrix} 1 & 0.19 & 0.66 & 0.44 & 0 \\ 0.19 & 1 & 0.44 & 0.65 & 0.61 \\ 0.66 & 0.44 & 1 & 0.52 & 0.13 \\ 0.44 & 0.65 & 0.52 & 1 & 0.56 \\ 0 & 0.61 & 0.13 & 0.56 & 1 \end{bmatrix}$$

③ 利用平方合成方法计算传递闭包 $tr(R)$ (即 R^*)。求得结果如下:

$$R^* = tr(R) = R^4 = \begin{bmatrix} 1 & 0.52 & 0.66 & 0.52 & 0.52 \\ 0.52 & 1 & 0.52 & 0.65 & 0.61 \\ 0.66 & 0.52 & 1 & 0.52 & 0.52 \\ 0.52 & 0.65 & 0.52 & 1 & 0.61 \\ 0.52 & 0.61 & 0.52 & 0.61 & 1 \end{bmatrix}$$

④ 选取适当的 λ 值进行动态分类。 λ 分别取 0.66 和 0.52 两个值, 比较两次不同取值得到的分类结果:

当 $\lambda = 0.66$ 时

$$R^*(\lambda) = \begin{bmatrix} 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

5 个节点被分成了 4 类, 即 $(u_1, u_3), (u_2), (u_4), (u_5)$ 。

当 $\lambda = 0.52$ 时

$$R^*(\lambda) = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \end{bmatrix}$$

(下转第 145 页)

[3] Li B, Xie S S, Keung G Y, et al. An Empirical Study of the Cool-streaming+ System[J]. IEEE Journal on Selected Areas in Communications, Special Issue on Advances in Peer-to-Peer Streaming System, 2007, 25(9): 1-13

[4] Rubenstein D, Sahu S. Can unstructured P2P protocols survive flash crowds? [J]. IEEE Transactions on networking, 2005, 13(3): 501-512

[5] Zhou Y, Chiu J C S. A Simple Model for Analyzing P 2 P Streaming Protocols[J]. IEEE International Conference on Network

Protocols(ICNP), 2007: 226-235

[6] Kumar R, Liu Y, Ross K. Stochastic Fluid Theory for P 2 P Streaming Systems[C]//26th IEEE International Conference on Computer Communications(INFOCOM), May 2007: 919-927

[7] Liu Y. On the Minimum Delay Peer-to-Peer Video Streaming: how Realtime can it be? [C]//Proc. of ACM Multimedia. 2007

[8] Liu F, Li B, Zhong L, et al. Understanding the flash crowd in P2P live video streaming systems [C]//Proc of the International Packet Video Workshop, Seattle, WA, USA, 2009

(上接第 124 页)

5 个节点被分成了 1 类, 即 $(u_1, u_2, u_3, u_4, u_5)$ 。从结果可以看到, λ 取不同值时对分类结果的影响。

4 算法时间复杂度分析

该算法分为 4 个步骤完成, 下面分析每个部分的时间复杂度。

在 2.2 节中由式(2)看出, 实际上是对一个 $n \times 2$ 的矩阵中的个数进行标准化处理, 时间复杂度为 $O(n)$ 。在 2.3 节中, 模糊矩阵 $R_{n \times n}$ 中的每一项 r_{ij} 均通过式(3)来计算, 其时间复杂度为 $O(n^2)$ 。在 2.4 节中, 通过 k 次 $n \times n$ 的矩阵相乘得到模糊等价矩阵 R^* , 其时间复杂度为 $O(n^3)$ 。在 2.5 节中, 将矩阵 R^* 中的每一项与 λ 进行比较, 得到矩阵 $R^*(\lambda)$, 其算法时间复杂度为 $O(n^2)$ 。

综上所述可以得到, 模糊分类算法的时间复杂度为 $O(n^3)$ 。

5 算法仿真

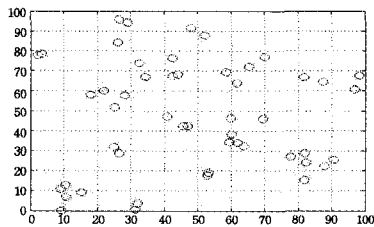


图 4 50 个节点非均匀的分布在 100m * 100m 的区域里

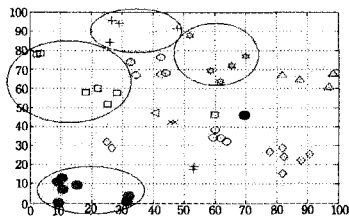


图 5 利用 k -均值聚类算法得到的分区结果, 椭圆标示区域显示边缘节点没有独立分出

利用 MATLAB 进行算法仿真, 在 100m * 100m 的区域中随机放置 50 个无线传感器节点(见图 4), 这些节点非均匀分布。利用 k -均值聚类和基于模糊等价矩阵聚类两种算法进行节点分区仿真, 结果如图 5、图 6 所示: 同颜色、同形状的节

点属于同一类, 两个图中的节点都分成了 14 个类。从图中可以看出, 基于模糊等价矩阵聚类分区算法成功地将边缘节点(离聚类中心较远的节点)分为单独的区域。

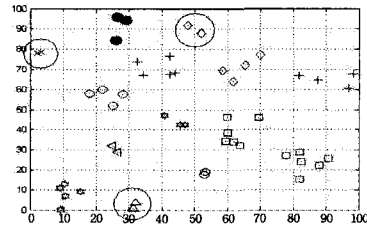


图 6 利用基于模糊等价矩阵聚类的分区结果, 椭圆圈出区域显示边缘节点被独立分出

结束语 本文的创新之处在于, 将模糊集理论中的模糊分类思想用于无线传感器节点分区, 为后续的节能路由算法设计提供了一种新的平台。从算法仿真的实验结果可以看出, 该算法确实可以将区域中的边缘节点分成独立的区域, 弥补了 k -均值聚类分区算法的不足。在分区算法中为无线传感网络路由算法的设计做了铺垫, 如何设计出更节能的路由算法是以后工作的方向和重点。

参考文献

[1] Heinzelman W R, Chandrakasan A, Balakrishnan H. Energy-Efficient communication protocol for wireless micro sensor networks[C]//Proceedings of the Hawaii International Conference on System Sciences, San Francisco: IEEE Computer Society, 2000: 3005-3014

[2] Manjeshwar A, Grawal D P. TEEN: A Protocol for Enhanced Efficiency in Wireless Sensor Networks [C]//Proceeding of the 15th Parallel and Distributed Processing Symp. San Francisco: IEEE Computer Society, 2001: 2009

[3] Younis O, Fahmy S. HEED: A Hybrid Energy-efficient Distributed Clustering Approach for Adhoc Sensor Networks[J]. IEEE Trans. on Mobile Computing, 2004, 3(4): 660

[4] 王佳英, 马洪连, 丁男. 无线传感器网络聚类改进层次路由协议[J]. 计算机工程与应用, 2010, 46(25): 91-93

[5] 贾百川, 刘胜全, 黄蓬国. 基于轮班工作机制的 WSN 多跳聚类协议[J]. 计算机工程与应用, 2010, 46(25): 90-93

[6] 陈水利, 李敬功, 王向公. 模糊集理论及其应用[M]. 2009: 59-122