

基于模糊聚类的传感网络失效节点检测

柴继贵

(成都电子机械高等专科学校, 四川 成都 610031)

摘 要: 针对当无线传感网中的传感节点分配随机性较强, 节点能量不能得到有效的补充, 一旦节点能量耗尽, 会造成节点失效. 传统检测方法因为无法准确提取节点失效特征, 造成失效节点检测的准确度不高. 为了解决这一问题, 提出一种基于模糊聚类失效节点搜索的检测算法, 通过计算无线传感节点能量消耗的特征模糊聚类概率, 运用模糊概率判定失效节点位置. 实验证明, 该方法能有效的对无线传感网络中失效节点进行检测, 取得了不错的效果.

关键词: 无线传感网; 节点数据; 模糊聚类概率

中图分类号: TP393.08

文献标识码: A

文章编号: 1000-7180(2012)07-0184-04

Based on the Fuzzy Clustering Sensor Failure Detection Network Node

CHAI Ji-gui

(Chengdu Electromechanical College, Chengdu 610031, China)

Abstract: For when the wireless sensor network sensor node random distribution is stronger, the node energy can get effective supplement, once the node energy run out, can cause node failure. Traditional test methods cannot because can't be accurately extract node failure features, cause failure detection of node accuracy is not high. In order to solve the problems, the article proposes a kind of based on fuzzy clustering failure node of the search algorithm, through calculation wireless sensor node energy consumption characteristics of fuzzy clustering probability, applying fuzzy probability judge failure node position. The experimental results prove that the method can effectively of wireless sensor network failure to detect nodes, good results.

Key words: wireless sensor network; node data; fuzzy clustering probability

1 引言

随着现代无线传感网络技术的不断发展, 传感网络技术在很多场得到了极为广泛的应用, 各种现场的数据能够快速的传递到数据终端, 传统的无线传感网络连接的是与外界环境相关的信息, 这些信息通过外界传感装备进行采集, 该项技术给人民的生活带来了极大的便利^[1-3]. 但是, 随着该项技术的深入发展, 也带来了一定的弊端, 无线传感网络中, 节点能量往往不能得到有效的补偿, 在进行远距离的信息传递过程中, 节点常常因为能量耗尽进而失效, 如果节点失效的数量过多, 会对传输信息的能力带来极大的损害. 为了解决以上的问题, 针对无线

传感网络节点失效的检测方法层出不穷. 很多方法都希望利用一种简单有效方式, 对无线传感网络中的失效节点进行准确的定位^[4-6].

传统的无线传感网络失效节点的检测算法中, 多半是通过检测失效节点的特征信息, 但是, 在当前的无线传感网络中, 节点一般根据现场需要采集的信息数量进行部署, 因此, 多半是随机分布的, 这种节点分布的随机性, 造成节点之间的距离无法准确的进行定位, 节点间距离的不定就造成了在信息传输的过程中, 传输距离也是不确定的, 每个节点能量是不同的, 因此, 能够提供的传输距离也不同, 传统的无线传感网络节点中, 节点的能量不能得到有效的补充, 因此, 一旦两个节点之间传输距离较大, 会

造成节点能量的耗尽,形成节点的失效,在不同簇内区域中,如果能量耗尽的节点数量较大,那么在这么区域中,外界环境的信息将不能得到较为有效的传递,影响了无线传感网络的网络通信效率,因此,需要对这种节点进行有效的检测.但是传统的基于特征的检测方法,由于无法获取节点中的统一特征,因此对失效节点的检测准确性一直不高.这也成为一个困扰无线传感网络发展的一个难题.

为了弥补这种缺陷,提出了一种基于特征模糊聚类的失效节点检测算法,通过计算无线传感节点能量消耗的特征模糊聚类概率,运用模糊概率判定失效节点位置,通过数据概率决策的算法,确定是否有失效节点存在^[7-9].实验证明,本文提出的方式能够对无线传感网络失效节点的存在情况进行准确检测,效果明显.

2 节点数据异常检测原理

在当前的无线传感网络的应用中,是通过一个个相互连接的节点,进行数据通信的,节点是连接外部信息与操作中心的一个纽带,因此,对失效节点进行快速的检测定位,及时的进行替换,是十分重要的^[10].传统的方法多半是基于节点特征进行对比,具体方法如下所述:

首先,建立特征提取的空间模型,设为 $Z(y, z)$,根据模型,可以提取节点中的相应特征.

建立模型之后,需要确定,相应的节点是否为待检测节点,需要进行第一次阈值判断,方法如下:

$$\begin{cases} (Z(y, z) + T) / \log T \geq 0.5 & \text{提取 } T \text{ 为检测对象} \\ (Z(y, z) + T) / \log T < 0.5 & \text{不提取 } T \text{ 为检测对象} \end{cases} \quad (1)$$

把待检测节点进行组合,设为 $T = (T_1, T_2, T_3, \dots, T_s)$,其中 T_j 表示第 j 个节点信息, s 代表以共有多少节点需要检测.

在检测到需要检测的节点数据信息之后,需要提取相应的节点特征,并且凭借此特征进行节点信息的检测.提取特征的方法如下:

$$E(l) = \sin x / \sum_{j=2}^s \log T \quad (2)$$

在对待检测的节点进行信息特征提取之后,需要根据信息特征与异常节点阈值进行对比,以判断是否为异常节点:

$$\begin{cases} E(l) \geq \mu & \text{节点数据异常} \\ E(l) < \mu & \text{节点数据正常} \end{cases} \quad (3)$$

通过以上检测原理可以看出,传统的无线传感

网络的节点异常检测方法,是根据提取节点中的某些特征,通过特征对比,完成节点异常检测的.但是在无线传感网络中,节点的分布是随机的,节点采集的信息,也可能完全不同,因此,随着传感距离的增加,节点能量耗尽,这种情况下很难建立稳定的特征识别模型.这就造成,公式1中提取的无线传感网络失效节点的数量较少.一旦发生这种情况,会导致公式2中很难提取到一致的特征,也就没有办法建立准确的特征阈值.这就导致对无线传感网络中的失效节点检测准确性不高.

为了弥补这种缺陷,提出了一种基于特征模糊聚类的失效节点检测算法,对无线传感网络失效节点的存在情况进行准确检测,效果明显.

3 特征模糊聚类概率的节点失效检测

3.1 失效节点数据搜索

因为传统的无线传感网络中,节点的能量不一致,有些能量较大,有些能量较小,因为无法像正常网络那样,建立准确的搜索模型,因也就此无法进行循环搜索、维持高效的搜索效率.因此,本文提出一种基于模糊概率的计算方法,其中最重要的步骤是要建立失效节点与正常节点之间的关联规则模型.在失效节点检测的过程中,很关键的一个步骤就是搜索相关的失效节点信息,运用模糊概率模型的方法表示,即对一个待检测的失效节点信息组,其中包含可能存在的 m 个失效节点信息和 n 个信息组,因此可以组成一个 $m \times n$ 的无线传感失效节点矩阵,用 $B_{m \times n} = \{b_{ij}\}_{m \times n}$ 表示,矩阵中一个点表示的是可能为失效节点的信息,每一列代表一个失效节点集合组成的待检测向量.

$$B_{m \times n} = \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ b_{m1} & b_{m1} & \cdots & b_{mn} \end{bmatrix} \quad (4)$$

由于建立的可能失效节点矩阵模型的维度较高,增加了运算的复杂性,为了避免运算结果不收敛,因此需要进行降维处理,保证结果的正确性.

$$B_{m \times n} = V_{m \times r} \cdot T_{r \times r} \cdot W_{r \times n}^T \quad (5)$$

其中 m 表示需要检测的节点数量, n 表示节点组的数量,由此可得:

$T = \text{diag}(\sigma_1, \sigma_2, \dots, \sigma_r)$, $r \leq \min(m, n)$, σ_i 为失效节点模型中的相关参数.

可以把以上的矩阵进行分解处理:

$$B_k = \sum_{i=1}^k v_i \sigma_i v_i \quad (6)$$

如果 $k \rightarrow \text{rank}(A)$ 时, 可以通过简化运算得到的 B_k 来描述 $B_{m \times n}$. 通过矩阵的降维运算, 建立的待检测失效节点矩阵中, 需要失效节点中的特征模糊权重, 得到需要提取的内容, 就能得到节点失效的主要权重信息, 而其他无关的非无效节点信息被有效的过滤掉, 为下一步的计算打下基础, 提高节点检测的效率. 通过以上方法的计算, 能够对网络中的节点信息进行迭代计算, 保证节点搜索的过程中的准确和高效.

3.2 特征模糊聚类概率的计算

在对包含不同信息的无线传感节点中, 通过获取节点中的信息改变程度, 计算改变的模糊概率. 如果模糊概率与检测前的结果相差不大, 则节点没有发生明显的衰变, 节点性能不发生改变. 节点检测之间具有一定的独立性, 为了完成异常数据的对比, 可以利用以下公式:

$$q(f_k \cdot c_j) = q(w_1 \cdot c_j) + q(w_2 \cdot c_j) + \dots + q(w_n \cdot c_j)$$

其中, $j=1, 2, \dots, p$;

$$P(e_j | b_i) = \frac{p(e_j) p(e_j | \beta_i)}{p(b_i)} \quad j=1, 2, \dots, n$$

式中 $p(b_i)$ 是对异常节点进行初始分类的先验概率; $p(e_j | b_i)$ 是进行节点分类的条件概率. 对关联节点进行检测, $p(e_j)$ 检测结果不是恒定的. 如果节点检测的结果是相互独立的, 之间的概率关系有:

$$\begin{aligned} p(e_j | b_i) &= p(v_1 | b_i) \cdot p(v_2 | b_i) \cdots p(v_m | b_i) \\ &= \prod_{j=1}^m p(v_j | b_i) \quad i=1, 2, \dots, n; \end{aligned}$$

根据全概率的计算得出的结果:

$$p(b_i) = \sum_{j=1}^n p(e_j) p(b_i | e_j)$$

如何计算得到的节点与失效节点的模糊程度值, 是检测失效节点的关键所在, 为了避免误差结果较大, 可以计算模糊概率 $p(b_i | e_j)$ 如下:

$$p(b_i | e_j) = \frac{1 + S(n(b_i) | e_j)}{2 + S_{d_j}} \quad (7)$$

$S(n(b_i) | e_j)$ 为 d_j 类节点分类中出现的失效节点的数目, S_{d_j} 为待检测样本中属于失效节点的类别的数量. S_d 为待检测节点的数量. 由于在失效节点检测中, 失效节点和非失效节点, 为了取得更好的检测结果. 需要运用不同的阈值信息对概率统计进行

节点失效的判断, 保证 $\frac{p(e=sda|\alpha)}{p(e=no-sda|\alpha)} \geq z$, z 为节点失效的阈值. 在节点失效检测中被判定合法节点和失效节点的模糊概率程度, 即

$$p(e=sda|\beta) + p(e=no-sda|\beta) = 1, \text{ 通过 } p(e=sda|\beta) \geq \frac{z}{z+1} \text{ 经过对计算过程进行简化,}$$

当 $\lim_{z \rightarrow 999} \frac{z}{z+1} = 0.99$ 时, 节点失效的计算结果最为准确.

通过以上的计算方法们虽然无法准确的检测到失效节点的一致性信息, 但是能够通过模糊概率的计算得到失效节点的概率, 通过阈值计算出节点的状态, 保证检测的准确性.

4 实验结果及对比

为了验证多本文方法在验证失效节点的效果, 对一个采集外部多样性信息的传感节点网络信息进行节点有效性的检测, 为了营造实验环境, 保证节点检测的高效性, 对节点的传输距离和布置密度不做明确要求, 实验步骤如图 1 所示.

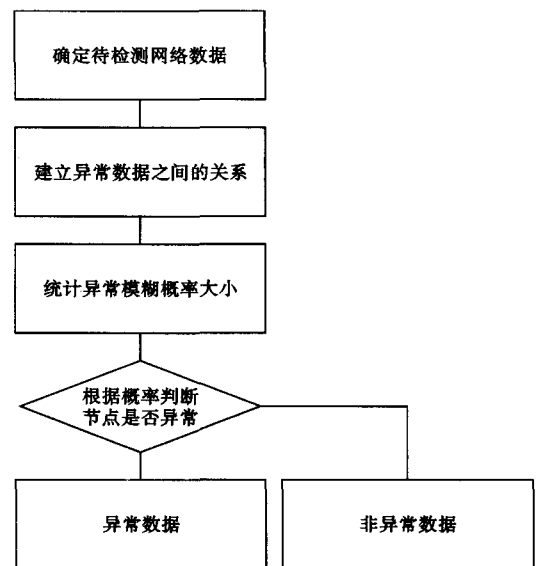


图 1 实验流程示意图

本文选取的节点信息集合中包含 1300 个可能包含无效节点的传感网络, 其中有部分节点是能量耗尽而无效的, 可以通过对异常节点的准确定位统计, 进而不断的增加节点数量, 不断的统计监测结果的准确率, 对两种不同的方法进行统计, 获取的实验数据用图 2 描述.

可以看出, 随着待检测的节点数量的不断增加,

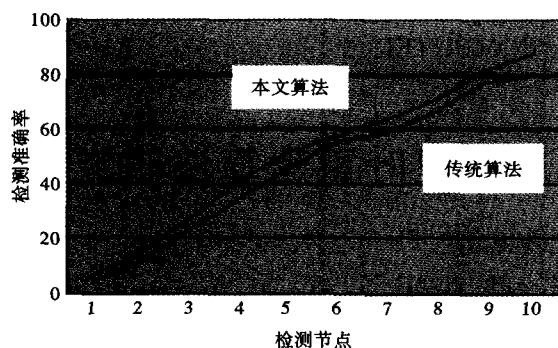


图2 节点数据异常检测曲线示意图

不同方法得到的节点检测准确率也随之增加,符合实际情况。但是,本文算法在相同节点数量的前提下,检测的结果都优于传统的方法,为了进一步说明情况,对平均的检测准确率进行统计。实验数据用表1描述。

表1 两种检测方式准确率表

算法	节点数据异常数目	检测准确度
传统方法	200	53%
本文方法	200	67%

利用公式: $M = (E(l) + \omega_e) \cdot s / T_j \cdot s$ 能够得到分别使用两种方式进行节点检测的漏检率,如表2所示。

表2 检测漏检率对比表

算法	漏检率
传统方法	23.1%
本文方法	16%

本文方法是运用模糊概率的观点对失效的节点进行检测,克服了传统方过于依靠固定特征的弊端,因此本文算法得到的结果优于传统算法。

5 结束语

本文在计算节点数据与异常数据之间的特征模

糊聚类概率的基础上,提出了一种基于特征模糊聚类的失效节点检测算法。通过计算无线传感节点能量消耗的特征模糊聚类概率,运用模糊概率判定失效节点位置,运用数据概率决策的算法,确定是否有失效节点存在。实验证明,使用本文提出的检测方法,取得了很好的效果,有很大的使用价值和应用前景。

参考文献:

- [1] 许荣生,吴海燕,毕学尧. 网络信息安全的关键技术[J]. 计算机世界,2000(18):263-267.
- [2] 金波,林家俊,王行愚. 入侵检测评述[J]. 华东理工大学学报,2000(4):191-197.
- [3] 高翔,苏广文,胡正国. 入侵检测系统中的网络监测[J]. 微电子学与计算机,2002(2):37-39.
- [4] 田俊峰,张喆,赵卫东. 基于误用和异常技术相结合的人侵检测系统的设计与研究[J]. 电子与信息学报,2006(28):2162-2166.
- [5] 程柏良,周洪波,钟林辉. 基于异常与误用的人侵检测系统[J]. 计算机工程与设计,2007(28):3341-3343.
- [6] 姜华,袁晓兵,付耀先,等. 无线传感器网络中信道仿真模型的研究[J]. 计算机仿真,2006(11):129-132.
- [7] 李权,周兴社. 基于 KPCA 的多变量时间序列数据异常检测方法研究[J]. 计算机测量与控制,2011(4):822-825.
- [8] Ball GH, Hall DJ. A novel method of data analysis and pattern classification[M]. USA: search Institute, California, 1965.
- [9] Guha S, Rastogi R, Shim K. ROCK: a robust clustering algorithm for categorical attributes [J]. Information Systems, 2000(25):345-366.
- [10] 袁朝辉,王恽,杨芳. 飞机地面滑跑方向的模糊控制系统仿真研究[J]. 计算机仿真,2011(2):76-79.

作者简介:

柴继贵 男,(1976—),硕士,讲师。研究方向为计算机网络、软件工程。