

基于可信度和邻居协作的传感器故障检测算法

蔺艳斐¹, 邵苏杰¹, 钟成², 郭少勇¹, 邱雪松¹

(1. 北京邮电大学网络与交换技术国家重点实验室, 北京 100876; 2. 河北省电力公司, 河北 石家庄 050022)

摘 要: 通过可疑节点与邻居节点的历史数据比较, 判断节点是否故障是传感器故障检测的有效方法之一。该方法的关键点在于可疑节点的判断问题以及故障诊断消息发送时机的决策问题。为此, 提出一种基于可信度和邻居协作的传感器故障检测算法。首先建立可信度模型进行节点可疑状态判定; 随后基于邻居协作的思想向邻居节点发送故障诊断请求, 通过等概率时间窗口决定该可疑节点诊断请求的发送时机; 最后依据邻居节点的诊断回复进行故障情况的分类以及状态判断。通过仿真验证, 该算法能够在降低故障诊断次数和数据拥塞率的前提下, 提高故障检出率。

关键词: 无线传感器; 故障检测; 可信度模型; 邻居协作; 等概率时间窗口

中图分类号: TP393

文献标识码: A

Sensor fault detection algorithm based on credibility and neighbor-cooperation

LIN Yan-fei¹, SHAO Su-jie¹, ZHONG Cheng², GUO Shao-yong¹, QIU Xue-song¹

(1. State Key Laboratory of Networking and Switching Technology, Beijing University of Posts and Telecommunications, Beijing 100876, China;

2. Power Grid of Hebei Province, Shijiazhuang 050022, China)

Abstract: To effectively detect whether a node was failure, one of the effective sensor fault detection methods was to compare historical data of suspicious node with its neighbors. The key points of this method were identification of suspicious nodes, as well as decision problem of sending timing for fault diagnosis messages. To this end, a sensor fault detection algorithm was presented based on credibility and neighbor-cooperation. Firstly, a credibility model to determine whether nodes are suspicious was established. Then suspicious nodes send fault diagnosis requests to neighbor nodes based on neighbor-cooperation. The sending timing was determined by equal probability time window. Finally, it could finish classification of fault condition and state judgment based on diagnosis responds of neighbor nodes. Simulation experiments show that this algorithm can achieve higher fault detection rate with less fault diagnosis times and low data congestion probability.

Key words: wireless sensor, fault detection, credibility model, neighbor-cooperation, equal probability time window

1 引言

传感器工作环境中的振动、噪声、信道干扰、火灾等都可能导致传感器错误感知数据甚至永久故障^[1], 为防止产生错误的传感数据, 保证传感器网络的正常运行, 需要及时检测传感器故障。集中式和分布式是传感器故障检测的 2 种主要方法。集中式方法周期性收集传感器测量值和状态信息到中心节

点, 但大量数据通信导致靠近中心节点传感器能耗加快, 缩短传感器网络的生命周期^[2-4]。分布式方法通过与邻居传感器监测数据相比较实现可疑节点的故障检测, 克服了传感器节点间能耗不平衡的问题, 但依然存在邻居节点集合间突发大量数据通信导致拥塞的情况, 影响故障检测的效率和准确性^[5,6]。因此, 基于分布式方式以少量数据通信完成快速可靠的传感器故障检测成为需要解决的问题。

收稿日期: 2016-03-21; **修回日期:** 2016-05-19

基金项目: 国家电网科技项目基金资助项目(No.SGIT0000KJJS1500008)

Foundation Item: The Science and Technology Project Foundation of State Grid (No.SGIT0000KJJS1500008)

针对分布式传感器故障检测方法，目前有大量文献进行了深入研究。文献[7]提出了基于历史数据与邻居协作的故障检测机制，选择邻居节点最多且健康的传感器为根节点，迭代式判断其邻居节点状态，但仅依据单个健康节点的数据进行故障判断并不可靠。文献[8,9]根据邻居节点数据计算故障概率，并通过边界节点数据调整该概率，但仅通过数据差值计算故障概率的方法有待优化。文献[10]提出了基于 k -means 的故障诊断算法，并利用蚁群算法进行优化，在一定程度上提高了故障检测的准确率，但算法的高复杂度并不适用于大规模传感器网络。

文献[11]提出基于分布式散列表的传感器故障识别算法，通过与邻居节点数据的比较实现故障判定，但并没有深入研究其参数的取值方法。文献[12,13]分别与自身历史数据以及可信邻居历史数据相比较实现节点状态判断，但没有涉及诊断请求的发送时机问题，可能导致数据无序反复发送，影响判断的效率和准确性。文献[14]引入时间窗口确定发送诊断请求的时机，但时间窗口的执行方式并不明确，且没有对可疑节点邻居节点的状态做详细分类，故障判断的准确性存在一定缺陷。

基于上述分析，分布式传感器故障检测研究需重点解决 2 个问题：1) 可信的可疑节点判断；2) 诊断请求发送时机。一旦传感器节点数据异常立刻转为可疑节点并向邻居节点发送诊断请求的方式并不合理，数据突变可能是由于环境中确实存在突发事件而导致，如温度传感器监测区域的温度骤升时监测数据会发生突变^[15]。因此需要判断节点是否切实可疑，避免不必要的诊断请求传输开销。由于同一时刻可能存在大量可疑节点，同时发送诊断请求会造成数据拥塞，因此需要避免大量无序重复的数据通信出现。

为此，本文提出一种基于可信度和邻居协作的传感器故障检测算法。首先建立可信度模型判定节点可疑状态；随后基于邻居协作思想发送诊断请求，并通过等概率时间窗口决定发送时机；最后为进一步提高故障诊断的准确性，依据邻居节点的诊断回复对故障情况详细分类并进行状态判断。

2 问题模型

本文所提的基于可信度和邻居协作的传感器故障检测如图 1 所示，由若干个处于健康、可疑、

故障状态的传感器节点组成。当某个传感器节点处于可疑状态时，在以该节点为圆心，半径为 R 的虚线圆所围区域内进行基于邻居协作的故障诊断。

当某个传感器发现当前 t 时刻的数据与 $t-1$ 时刻的数据差距较大时，首先需要基于可信度对其是否为可疑节点进行判断。当可信度模型判断该数据突变确实是由环境中存在的突发事件导致，则不需要启动故障诊断过程；否则立即向以该节点为圆心，半径为 R 的虚线圆区域内的传感器发送故障诊断请求。因此如何根据历史数据建立合理准确的可信度模型是本文需要解决的一个关键问题。

此外，为避免因同一区域内多个传感器同时被判定为可疑状态而导致的大量突发的数据通信，如何在时间轴上区分这些故障诊断请求，建立合理的时间窗口以决策诊断请求发送时机是另一个需要解决的问题。

可疑节点故障诊断过程启动后，邻居传感器节点都发送诊断回复是不必要的。如果邻居节点是可疑的，则其数据并不具有参考性。如图 1 中的可疑节点 S_1 、 S_2 、 S_3 所示，其邻居节点状态可能存在以下 3 种情况：1) 邻居节点中只存在少量或不存在可疑节点和故障节点；2) 邻居节点中存在大量可疑节点，只有少量或者不存在正常节点和故障节点；3) 邻居节点中存在大量故障节点，只有少量或者不存在可疑节点和正常节点。邻居节点需要根据自身状态确定是否发送诊断回复，并结合自身监测数据与收到的诊断请求数据确定诊断回复的具体内容。此时，收到诊断回复后需要进行状态判断的可疑传感器节点并不能立刻完成状态判定，需要对诊断回复数据进行统计分析并按以上 3 种情况进行分类，最终完成状态判定。

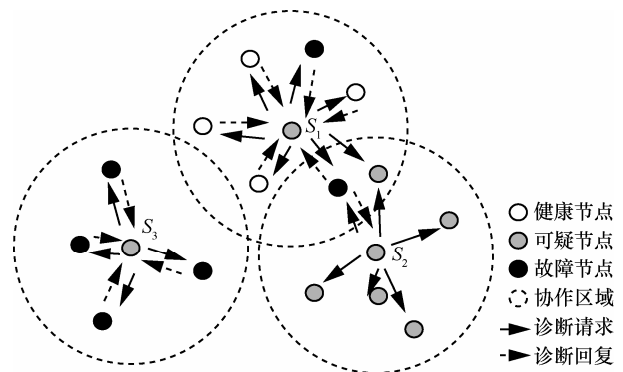


图 1 无线传感器网络故障检测示意图

3 传感器故障检测算法

本文的传感器故障检测算法主要包括 2 个阶段：分布式基于可信度模型的全网可疑节点确定以及基于邻居协作的可疑节点故障状态判定。

3.1 确定可疑节点

3.1.1 可信度模型

判断节点是否可疑，是根据其监测数据是否可信来实现的。通过传感器节点当前监测数据相对于历史数据的变化趋势来确定当前监测数据是否可信。本文通过方差来量化监测数据的变化趋势，数据方差较小时，说明数据比较稳定。

假设传感器节点可以保存 k 个数据，分别为 $t-k, t-k+1, \dots, t-1$ 时刻的数据 d_1, d_2, \dots, d_k 。设

\bar{d}_{t-1} 是这 k 个数据的均值， $\bar{d}_{t-1} = \frac{\sum_{i=1}^k d_i}{k}$ ，依照式(1)

计算这 k 个数据的方差 S_{t-1}^2 。

$$S_{t-1}^2 = \frac{\sum_{i=1}^k (d_i - \bar{d}_{t-1})^2}{k} \quad (1)$$

t 时刻数据为 d_t ，则当前 k 个数据的均值 $\bar{d}_t = \frac{(d_2 + d_3 + \dots + d_k + d_t)}{k}$ 。依照式(2)计算 S_t^2 。

$$S_t^2 = \frac{\sum_{i=2}^k (d_i - \bar{d}_t)^2 + (d_1 - \bar{d}_t)^2}{k} \quad (2)$$

定义参数 $s > 0$ ，当 $|S_t^2 - S_{t-1}^2| < s$ 时，方差变化较小，说明当前数据与历史数据相似。当 $|S_t^2 - S_{t-1}^2| \geq s$ 时，说明当前数据产生突变。

设置可信度 $\theta (0 \leq \theta \leq 1)$ 判断节点数据突变是否由环境中切实存在突发事件导致。置所有节点的 θ 初始值为 1，当 $|S_t^2 - S_{t-1}^2| \geq s$ 时，减小 θ 值；否则增大 θ 值。随着时间的推移，数据逐步更新会使 θ 值回归。

但是仅依靠方差来调整可信度可能会忽略一些传感器故障。假设在 t 时刻，某个传感器节点发生故障，其监测数据异常增大，且后续数据也是偏大的，则 $|S_t^2 - S_{t-1}^2|$ 可能随着时间的推移逐渐减小甚至小于 s ，导致可信度 θ 反而开始增加，该故障无法被监测到。因此需要将方差和均值同时作为调整可

信度的因素，定义参数 d ，当 $|S_t^2 - S_{t-1}^2| < s$ 且 $\bar{d}_t - \bar{d}_{t-1} < d$ 时，可信度 θ 增加；否则可信度 θ 减小。节点可信度模型如式(3)所示。

$$\theta = \begin{cases} 1, & |S_t^2 - S_{t-1}^2| < s \text{ 且 } \bar{d}_t - \bar{d}_{t-1} < d \text{ 且 } \theta + \partial \geq 1 \\ \theta + \partial, & |S_t^2 - S_{t-1}^2| < s \text{ 且 } \bar{d}_t - \bar{d}_{t-1} < d \text{ 且 } \theta + \partial < 1 \\ \theta - \partial, & \text{其他} \end{cases} \quad (3)$$

定理 1 式(3)中的参数 d 和 s 的取值取决于实际故障检测中的 k 值，以及可容忍的 d_t 、 d_1 、 \bar{d}_{t-1} 两两之间的差值，与 d_t 、 d_1 、 \bar{d}_{t-1} 的具体值无关。

证明 经计算得 $\bar{d}_t - \bar{d}_{t-1} = \frac{d_t - d_1}{k}$ ，由式(3)可知 $\bar{d}_t - \bar{d}_{t-1} < d$ ，即 $\frac{d_t - d_1}{k} < d$ ，所以 d 的值取决于实际故障检测中的 k 值以及可容忍的 d_t 、 d_1 之间的差值。

$$\text{由 } \bar{d}_t - \bar{d}_{t-1} = \frac{d_t - d_1}{k}, \text{ 可知 } \bar{d}_t = \bar{d}_{t-1} + \frac{d_t - d_1}{k},$$

代入式(2)可得

$$S_t^2 = \frac{\sum_{i=2}^k \left(d_i - \bar{d}_{t-1} - \frac{d_t - d_1}{k} \right)^2}{k} + \frac{\left(d_1 - \bar{d}_{t-1} - \frac{d_t - d_1}{k} \right)^2}{k}$$

将 $d_i - \bar{d}_{t-1}$ 从平方项中取出并化简可得

$$kS_t^2 = \sum_{i=2}^k (d_i - \bar{d}_{t-1})^2 - \frac{2(d_t - d_1) \sum_{i=2}^k (d_i - \bar{d}_{t-1})}{k} + \frac{(k-1)(d_t - d_1)^2}{k^2} + \left(d_1 - \bar{d}_{t-1} - \frac{d_t - d_1}{k} \right)^2$$

结合式(1)，第一项替换为 $kS_{t-1}^2 - (d_1 - \bar{d}_{t-1})^2$ ，进一步化简可得

$$kS_t^2 = kS_{t-1}^2 - (d_1 - \bar{d}_{t-1})^2 + (d_t - \bar{d}_{t-1})^2 - \frac{(d_t - d_1)^2}{k}$$

因此

$$S_t^2 - S_{t-1}^2 = \frac{k(d_t - \bar{d}_{t-1})^2 - k(d_1 - \bar{d}_{t-1})^2 - (d_t - d_1)^2}{k^2}$$

由式(3)可知 $|S_t^2 - S_{t-1}^2| < s$ ，即

$$\left| \frac{k(d_t - \bar{d}_{t-1})^2 - k(d_1 - \bar{d}_{t-1})^2 - (d_t - d_1)^2}{k^2} \right| < s$$

所以， s 的值取决于实际故障检测中的 k 值以

及可容忍的 d_1 、 d_i 、 \bar{d}_{i-1} 两两之间的差值，与 d_1 、 d_i 、 \bar{d}_{i-1} 的具体值无关。

3.1.2 可信度门限

式(3)中的 θ 并不是无限减小的，引入一个门限值 ε 来作为可疑节点与健康节点可信度的分界，如式(4)所示。

$$\begin{cases} \theta > \varepsilon, & \text{健康节点} \\ \theta \leq \varepsilon, & \text{可疑节点} \end{cases} \quad (4)$$

定理 2 为保证节点有连续 $m(0 < m < k, m \in \mathcal{Z})$ 次的数据异常时，该节点状态一定转变为可疑， ∂ 与 ε 的取值应该满足 $1 - m\partial \leq \varepsilon < 1 - (m - 1)\partial$ 。

证明 由于节点可信度 $\theta \leq 1$ ，依照式(3)，每次的数据异常都会使 θ 减小 ∂ ，当某节点连续出现 m 次的数据异常时 $\theta \leq 1 - m\partial$ ，该节点可判定为可疑，依据式(4)， $\theta \leq \varepsilon$ ，为保证该判定切实成立， $1 - m\partial \leq \varepsilon$ 。为进一步精确 ε ，可知 $\varepsilon < 1 - (m - 1)\partial$ 。因此，当 $1 - m\partial \leq \varepsilon < 1 - (m - 1)\partial$ 时，可以保证节点在连续 m 次数据异常下，节点状态一定为可疑。

当 $\varepsilon < \theta < 1$ 时，不需要 m 次连续异常也能使节点转变为可疑状态。如 $t - 2$ 时刻 $\theta = 1 - (m - 1)\partial$ ，假设 $t - 1$ 时刻数据正常，此时 $\theta = 1 - (m - 2)\partial$ ，如果 t 时刻和 $t + 1$ 时刻数据均异常，则 $\theta = 1 - m\partial$ ，此时该节点为可疑节点。该情况通常发生在节点检测数据在异常和正常之间多次波动时，此时确定节点为可疑节点是必要的。

定理 2 中 m 的值依照实际要求设置。如果一发现异常数据就立即将节点状态转变为可疑并启动故障诊断程序，那么令 $m = 1$ 。 m 取值较大可以降低网络数据传输压力，但会忽略一些瞬时故障。 m 取值较小时，虽然可以提高故障诊断准确率，但会带来大量不必要的数据传输开销。因此本文权衡故障检测的准确率和网络流量 2 个因素，确定 m 的最佳取值。

可疑节点确定过程如图 2 所示。

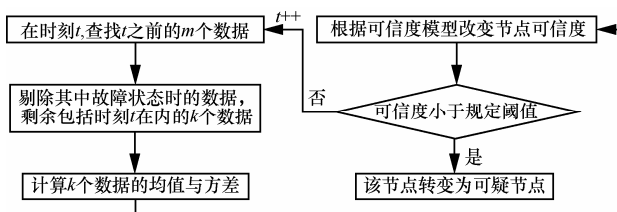


图 2 可疑节点确定流程

3.2 基于邻居协作的故障诊断

确定可疑节点以后，需要通过邻居节点历史数据的比较确定该节点是否为故障节点。关键要解决 3 个问题：1) 故障诊断请求时机决策问题；2) 邻居节点如何回复故障诊断请求；3) 如何根据诊断回复完成可疑节点状态判定。

3.2.1 诊断请求

本节在文献[14]的基础上提出了等概率时间窗口机制来确定可疑节点故障诊断请求的发送时机。

每一个可疑传感器节点设置一个时间窗口，如图 3 所示，单位时间间隔为 l ，长度为 w 个时间间隔， l 和 w 的大小依据传感器节点所监测设备的实时性要求等具体情况设定。可疑节点在每个时间窗口 $t + il (i \in \mathcal{Z}, 0 \leq i < w)$ 向邻居节点发送诊断请求的概率是相等的，为 $\frac{1}{w}$ 。假设 t 时刻有 n 个传感器节点转变为可疑节点，则这些节点在同一时间窗口发送诊断请求的概率为 $\frac{1}{w^n}$ ，每个时间窗口发送诊断请求的传感器节点数目的期望为 $\frac{n}{w}$ ，这 n 个可疑节点发送诊断请求的时间被极大地分散，有效缓解了数据传输压力。

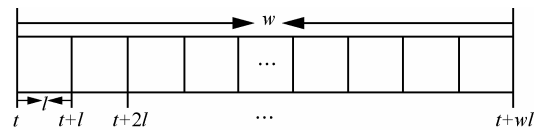


图 3 等概率时间窗口

当可疑节点 i 的时间窗口到达时，其向邻居节点发送诊断请求 $D_{req}(S_i^2, loc, t)$ ，其中， S_i^2 表示节点 i 的 S_i^2 ， loc 为节点 i 的坐标， t 为当前时刻。

3.2.2 诊断回复

可疑节点 i 发送故障诊断请求后，以 i 为圆心， R 为半径的圆形区域内的状态已确定的邻居节点 j 向 i 发送诊断回复。节点 j 首先检测自身状态并根据 loc 计算与 i 的距离 d_{ij} ，如果自身状态为可疑则仅向 i 回复一个 $state_j = -1$ 的消息；否则依照式(5)处理收到的信息。

$$cop = \begin{cases} 1, & \frac{|S_i^2 - S_{ji}^2|}{d_{ij}} < \lambda \\ 0, & \frac{|S_i^2 - S_{ji}^2|}{d_{ij}} \geq \lambda \end{cases} \quad (5)$$

其中, S_{ji}^2 为邻居节点 j 的 S_i^2 , 由于传感器节点间距离越近, 监测数据的相关度相对越高, 因此以 d_{ij} 作为参数, λ 依照实际情况确定。节点 j 发送诊断回复 $D_{req}(cop, state_j, t)$ 给节点 i , $cop=1$ 表示节点 j 认为节点 i 为健康节点; 否则为故障节点。 $state_j \in \{1, 0\}$ 分别表示节点 j 处于健康和故障状态。

基于邻居协作的故障诊断请求发送和诊断回复过程如图4所示, 首先利用等概率时间窗口机制确定发送诊断请求的时间, 之后不断判断时间窗口是否到达, 规定的时间窗口到达后, 向邻居节点发送诊断请求, 收到诊断请求的可疑邻居节点仅回复自身状态信息, 健康节点和故障节点依照式(5)处理接收到的数据, 并发送诊断回复。

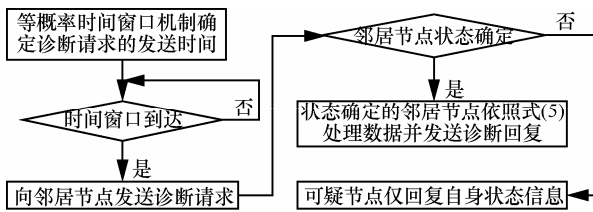


图4 基于邻居协作的故障诊断请求发送和诊断回复过程

3.2.3 状态判定

假设邻居协作区域 D 边缘的传感器发送数据到可疑节点 i 所用的时间为 T 。节点 i 发送诊断请求后等待 $2T$ 时间, 在等待时间内不断对收到的回复消息依照式(6)进行分类统计。

$$\begin{cases} N_0 = N_0 + 1, state_j = 1 \text{ 且 } cop = 0 \\ N_1 = N_1 + 1, state_j = 1 \text{ 且 } cop = 1 \\ F_0 = F_0 + 1, state_j = 0 \text{ 且 } cop = 0 \\ F_1 = F_1 + 1, state_j = 0 \text{ 且 } cop = 1 \\ Q = Q + 1, state_j = -1 \end{cases} \quad (6)$$

其中, N_0 、 N_1 、 F_0 、 F_1 、 Q 表示节点 i 的邻居节点

中各类节点的数目, $state_j$ 和 cop 是诊断回复中包含的信息。假设节点 i 在协作区域的邻居节点总数为 $A, A = N_0 + N_1 + F_0 + F_1 + Q$, 根据各类邻居节点所占比例进行分析, 可能出现的情况分为以下3类。

1) 如图5(a)所示, 可疑节点 i 的邻居协作区域内超过一半的节点为健康节点, 有很少的故障节点和可疑节点。健康节点和故障节点收到 i 的诊断请求后, 向节点 i 发送诊断回复, 可疑节点仅回复自身的状态信息。依照式(6)统计的回复信息满足 $N_0 + N_1 \geq \frac{A}{2}$ 。如果 $N_0 > N_1$, 则节点 i 为故障节点; 如果 $N_0 \leq N_1$, 则节点 i 为正常节点。

2) 如图5(b)所示, 可疑节点 i 的协作区域内故障节点的数量和健康节点的数量都不足一半, 有多个节点为可疑节点。依照式(6)统计的回复信息满足 $N_0 + N_1 < \frac{A}{2}, F_0 + F_1 < \frac{A}{2}$ 。由于可疑节点的数据不具有参考性, 应该重新等待 $2T$ 时间, 部分邻居可疑邻居节点完成状态判断之后, 重新统计数据进行分析。

3) 如图5(c)所示, 可疑节点 i 的协作区域内有超过一半的节点为故障节点。依照式(6)统计的回复信息满足 $F_0 + F_1 \geq \frac{A}{2}$ 。故障节点收到诊断请求消息后, 按照式(6)处理, 并发送诊断回复。如果 $F_0 < F_1$, 说明大多数的故障节点认为节点 i 为正常节点, 即节点 i 与大多数故障节点的数据相近, 此时节点 i 为故障节点。反之, 节点 i 为正常节点。

状态判定流程如图6所示, 可疑节点 i 发送诊断请求以后, 初始化已等待时间为 0, 开始接收并统计邻居节点的诊断回复。等待 $2T$ 时间后, 根据统计的数据分3种情况分别处理。情况1) 和情况3) 直接确定节点 i 的状态并修改节点的可信度, 健

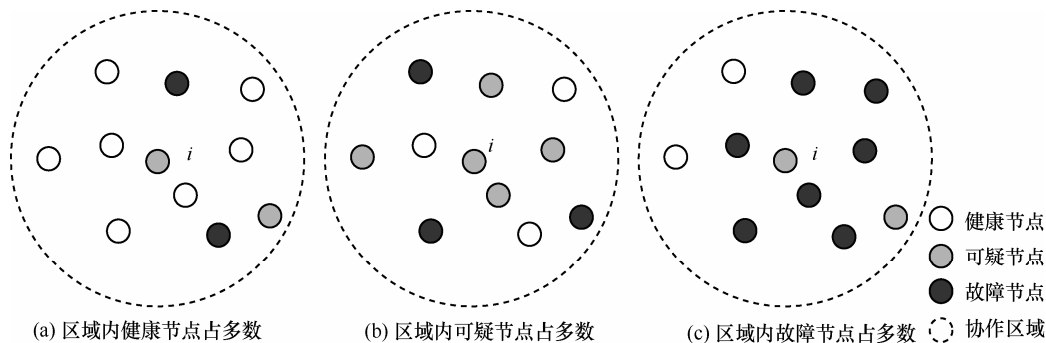


图5 3种故障场景

康节点可信度置为 1，故障节点可信度置为 0。情况 2)则初始化等待时间并重复以上过程直到节点状态确定。

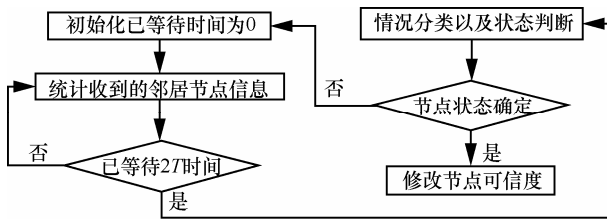


图 6 状态判定流程

4 仿真实验

本文在 100 m×100 m 的区域内分别随机放置 30、60、90 个传感器节点，在这 3 种不同的节点密度下进行仿真实验，确定 m 的最佳取值。并在 90 个传感器节点的情况下，将本文算法与文献[13]、文献[14]中的算法在故障诊断次数和故障检出率 2 个方面加以比较。

1) m 值的确定

图 7 和图 8 分别显示了在 3 种节点密度下，故障诊断次数以及故障检出率随 m 值的变化情况。从图 7 可以看出，故障诊断次数随着 m 的增大而减小，因为 m 值越小，故障诊断所需要的连续异常次数越少，因此故障诊断次数越多。从图 8 可以看出，故障检出率随着 m 的增大而减小，因为 m 值越大，故障诊断所需要的连续异常次数越多，连续异常数据较少的故障可能被忽略。为保证较少的故障诊断次数和较高的故障检出率，在 30、60、90 个传感器节点的情况下 m 的取值分别为 6、5、4，从图 7、图 8 可以看出，此时故障诊断次数较少，且故障检出率为 100%。当 m 的取值分别为 9、9、7 时，故障诊断次数更少，此时故障检出率为 97%以上。

2) 故障诊断的时间分布

为验证等概率时间窗口机制平衡数据流量的有效性，本实验对比了等概率时间窗口的方案与发现可疑节点立即发送故障诊断请求的方案。较多的故障诊断次数会造成较大的数据流量压力，为平衡数据流量和故障检出率，本实验设定节点数目为 90， m 取值为 7，此时故障诊断次数较少，且故障检出率可达 97%以上，时间窗口 w 的取值为 5。各个时间点发送故障诊断请求的传感器节点数目分布情况如图 9 所示，第 5 s、8 s、18 s、27 s、40 s

都有多个传感器需要发送故障诊断请求。不采用时间窗口机制时，这些诊断请求直接在以上时间点被发送，而采用时间窗口机制以后，诊断请求发送的时间被分散到各个时间窗口。如图 9 中第 27 s 的 5 个传感器诊断请求，被分散到第 27 s、28 s、29 s、30 s 发送，分别发送 1、1、2、1 个诊断请求。采用该机制可有效避免同一时间点发送大量诊断请求而导致的数据拥塞。

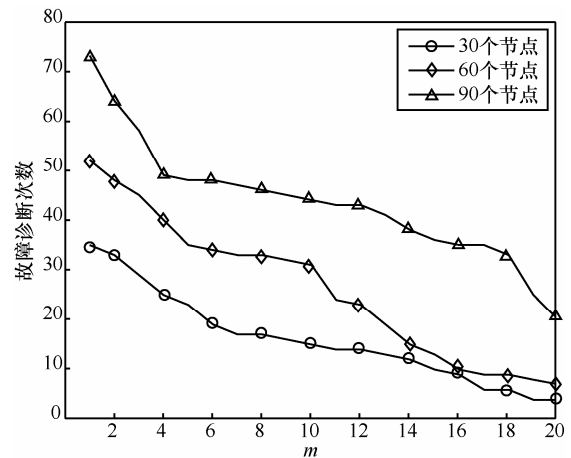


图 7 故障诊断次数随 m 值的变化

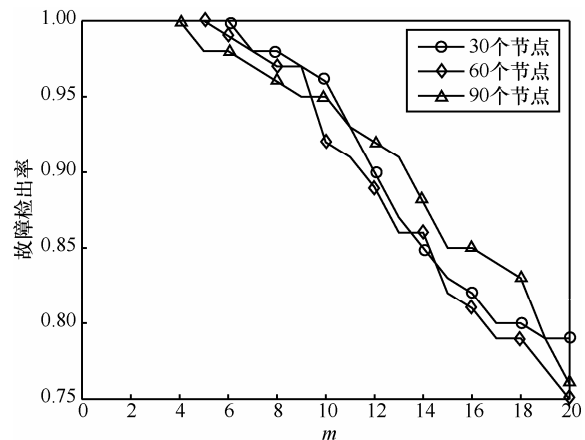


图 8 故障检出率随 m 值的变化

3) 故障检出率

本文从故障诊断次数和故障检出率 2 个方面比较本文算法与文献[13]中基于邻居节点数据分析的 WSN 故障检测方法，以及文献[14]中吞吐量下降和节能的传感器故障诊断机制的性能。从图 10 可以看出，本文算法故障诊断次数最少，文献[13]次之，文献[14]故障诊断次数最多。这是因为本文建立了可信度模型，传感器节点的数据有多次异常后使节点的可信度低于一定阈值

后才会发送诊断请求，从而减少了一些不必要的通信开销。从图 11 可以看出，本文算法的故障检出率最高，都在 97% 以上，文献[14]算法次之，最后是文献[13]算法。并且随着节点数目的增多，本文算法的故障检出率变化较小，而文献[13,14]故障检出率迅速下降。因此本文基于可信度和邻居协作的传感器故障检测算法能够适用于大规模无线传感器网络。

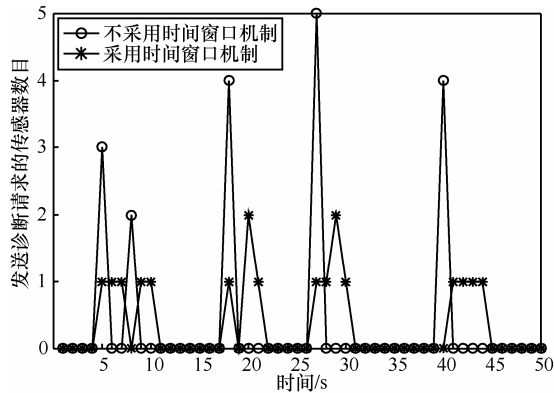


图 9 故障诊断的时间分布

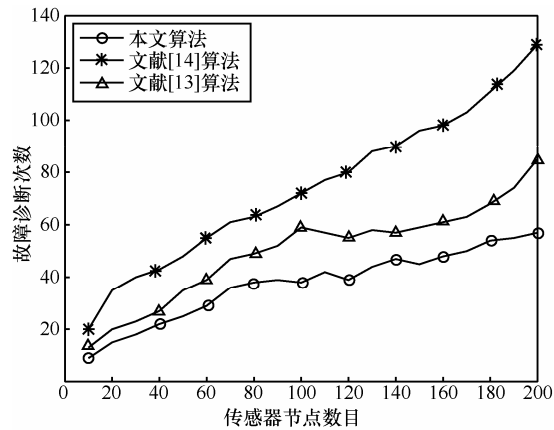


图 10 3 种算法的故障诊断次数

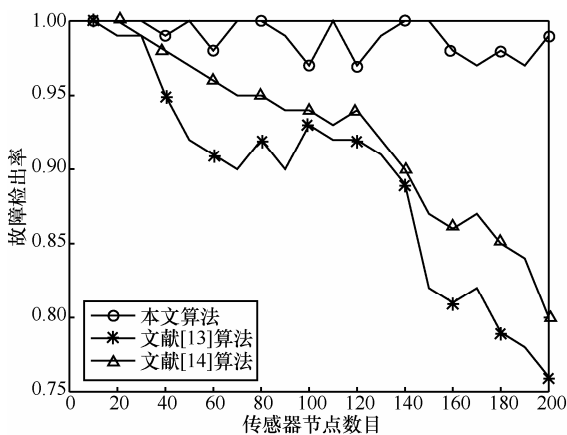


图 11 3 种算法的故障检出率对比

5 仿真结果分析

本文的仿真实验主要与文献[13]中基于邻居节点数据分析的 WSN 故障检测方法，以及文献[14]中吞吐量下降和节能的传感器故障诊断机制，在故障诊断次数和故障检出率 2 个方面进行比较。本文算法能够在降低故障诊断次数和数据拥塞率的前提下，提高故障检出率，主要是由于建立了可信度模型，引入时间窗口机制并基于邻居协作进行可靠的状态判定。

文献[13,14]中都提到了首先进行节点可疑状态判断，文献[13]在计算节点的可信度水平时仅依据历史数据的方差，但是如果发生异常后的监测数据维持在一个错误范围内的平稳值时，其历史数据的方差可能会逐步减小，节点的可信度逐步增加，此种情况下的异常可能会被忽略。文献[14]仅根据历史数据的均值进行节点可信度的增减，但是较大和较小的异常数据同时出现时，会由于 2 种异常数据值的平均导致异常无法发现。而且仅依靠均值进行判断会产生较多不必要的可疑节点。本文结合方差和均值 2 个因素建立可靠的可信度模型，可以避免忽略上述异常，同时有效减少不必要的故障诊断过程。

文献[13]在确定可疑节点以后立即向邻居节点发送该故障诊断请求，虽然诊断请求的发送及时，但是当环境突变等原因造成某一时刻产生大量可疑节点时，这些可疑节点同时立即向邻居节点发送诊断请求可能会造成数据拥塞，拥塞发生以后，这些请求被反复发送造成数据拥塞情况不断加剧，反而影响故障诊断的效率。文献[14]提到将诊断请求的发送时间推迟，但是没有具体介绍如何推迟。本文提出等概率时间窗口机制，当节点的可信度低于一定阈值之后根据该机制确定诊断请求发送的具体时间，可以使诊断请求的发送时间被分散，降低数据拥塞的可能性，减少故障诊断请求的发送次数。

收到邻居节点的诊断回复后，文献[13]在邻居节点中有超过一半的节点认为该节点为健康节点时，判定该节点健康；否则可疑。该方法没有考虑邻居节点的状态，因为可疑邻居节点的回复并不具有参考性。文献[14]根据邻居节点的诊断回复分情况讨论，但是其情况分类并不准确，没有考虑大部分都是故障节点的情况，且在各种情况下的状态判

断也不可靠。因此, 本文通过分析不同邻居节点的诊断回复进行分类统计, 产生 3 种故障情况, 并在不同的情况下, 根据各类回复消息的数目做相应处理, 有效提高了故障检出率。

6 结束语

通过分布式邻居协作完成传感器节点的状态判定是传感器故障检测的有效方法之一, 该方法的关键点在于可疑节点的判断和故障诊断消息发送时机的决策。本文提出了一种基于可信度和邻居协作的传感器故障检测算法, 可以按照可信度模型判定可疑传感器节点, 并按照等概率时间窗口机制确定可疑节点发送诊断请求的时间。仿真实验表明, 该算法能够在降低故障诊断次数和数据拥塞率的前提下, 提高故障检出率, 且适用于大规模无线传感器网络。

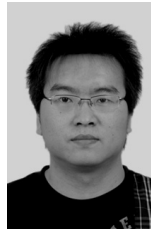
参考文献:

- [1] LU Z Q, WEN Y G. Distributed algorithm for tree-structured data aggregation service placement in smart grid [J]. IEEE Systems Journal, 2014, 8(2): 553-561.
- [2] CHANG C Y, LIN C Y, KUO C H. EBDC: an energy-balanced data collection mechanism using a mobile data collector in WSNs [J]. Sensors, 2012, 12(5): 5850-5871.
- [3] XUE L, KIM D, ZHU Y. Multiple heterogeneous data ferry trajectory planning in wireless sensor networks[C]//IEEE Conference on Computer Communications. Toronto, 2014: 2274-2282.
- [4] LIU X F, CAO J N. Fault tolerant complex event detection in WSNs: a case study in structural health monitoring [J]. IEEE Transactions on Mobile Computing, 2015, 12(14):2502-2515.
- [5] ZHAO M, CHOW T W. Wireless sensor network fault detection via semi-supervised local kernel density estimation[C]//2015 IEEE International Conference on Industrial Technology (ICIT). Seville, 2015:1495-1500.
- [6] RAVINDRA V K, ASHISH B J. A fault tolerant approach to extend network life time of wireless sensor network[C]//2015 IEEE International Conference on Advances in Computing, Communications and Informatics (ICACCI). Kochi, 2015: 993-998.
- [7] 邱雪松, 陈新颜, 杨杨. 历史数据与邻居协作融合的无线传感器故障检测机制[J]. 北京邮电大学学报, 2015, 38(sup): 1-5.
QIU X S, CHEN X Y, YANG Y. Neighbor-coordination in wireless sensor network [J]. Journal of Beijing University of Posts and Telecommunications, 2015, 38(sup): 1-5.
- [8] PAOLA A D, GAGLIO S, RE G. Adaptive distributed outlier detection for WSNs[J]. IEEE Transactions on Cybernetics, 2015, 45(5): 888-899.
- [9] YUAN H, ZHAO X X, YU L Y. A distributed Bayesian algorithm for data fault detection in wireless sensor networks[C]//2015 International Conference on Information Networking (ICOIN). Cambodia, 2015: 63-68.
- [10] YANG Y, LIU Q, GAO Z P. Data clustering-based fault detection in WSNs[C]//7th International Conference on Advanced Computational Intelligence. Fujian, China, 2015:334-339.
- [11] SENTHIL M, SUGASHINI K, ABIRAMI M. Identification and recovery of repaired nodes based on distributed hash table in WSN[C]//IEEE Sponsored 2nd International Conference on Innovations in Information Embedded and Communication Systems ICIIECS'15 Coimbatore. 2015:1-4.
- [12] SAIHI M, BOUSSAID B, ZOUINKHI A. Distributed fault detection based on HMM for wireless sensor networks[C]//4th International Conference on Systems and Control. Sousse, Tunisia, c2015:189-193.
- [13] 黄日茂, 邱雪松, 高志鹏. 无线传感器网络中邻居数据分析的故障检测方法[J]. 北京邮电大学学报, 2011, 34(3):31-34.
HUANG R M, QIU X S, GAO Z P. A neighbor-data analysis method for fault detection in wireless sensor networks [J]. Journal of Beijing University of Posts and Telecommunications, 2011, 34(3):31-34.
- [14] SHARMA K P, SHARMA T P. A throughput descent and energy efficient mechanism for fault detection in WSNs[C]//2015 International Conference on Industrial Instrumentation and Control (ICIC) College of Engineering Pune. India, 2015:28-30.
- [15] AMNA Z, BILAL W, BEENISH A A. A hybrid fault diagnosis architecture for wireless sensor networks[C]//2015 International Conference on Open Source Systems and Technologies. Lahore, 2015:7-15.

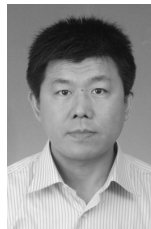
作者简介:



蔺艳斐 (1992-), 女, 山东莱芜人, 北京邮电大学硕士生, 主要研究方向为智能电网、网络与业务管理。



邵苏杰 (1985-), 男, 陕西汉中, 北京邮电大学博士生, 主要研究方向为研究方向为网络管理与智能电网。



钟成 (1970-), 男, 河北石家庄人, 河北省电力公司高级工程师, 主要研究方向为光缆、终端接入网。

郭少勇 (1985-), 男, 河北邢台人, 北京邮电大学博士后, 主要研究方向为网络管理、终端管理与智能电网。

邱雪松 (1973-), 男, 江西上饶人, 北京邮电大学教授、博士生导师, 主要研究方向为网络与业务管理。