

一种无线传感器网络中事件区域检测的容错算法

曹冬磊 曹建农 金蓓弘

(中国科学院研究生院 北京 100080)

(中国科学院软件研究所 北京 100080)

(香港理工大学电子计算学系 香港)

摘 要 特殊环境中的事件区域检测是无线传感器网络的一种重要应用. 由于传感器的错误会导致事件区域检测的不准确, 所以相关的容错算法成为近年来的研究热点. 已有研究工作都仅考虑了事件的空间相关性, 通过相邻传感器之间的数据交换实现容错. 文中从事件的空间相关性和时间相关性入手, 提出了一种以局部检测为主的分布式事件区域检测算法. 该算法通过检验传感器本地采样值构成的时间序列与事件随机过程统计特征的符合程度实现容错. 算法分析的结果表明, 该算法可以减少传感器之间的数据交换, 从而有效地利用传感器的能量. 模拟实验表明, 当有 10% 的传感器发生错误时, 该算法可以检测到 93% 的事件区域和 88% 的错误传感器.

关键词 无线传感器网络; 容错算法; 事件区域检测; 高能效

中图法分类号 TP302

A Fault-Tolerant Algorithm for Event Region Detection in Wireless Sensor Networks

CAO Dong-Lei CAO Jian-Nong JIN Bei-Hong

(Graduate University of Chinese Academy of Sciences, Beijing 100080)

(Institute of Software, Chinese Academy of Sciences, Beijing 100080)

(Department of Computing, Hong Kong Polytechnic University, Hong Kong)

Abstract Detecting the region of emergent events is an important application of wireless sensor networks. In recent years, research on fault-tolerant event region detection algorithms becomes a hot topic. By assuming that the occurrence of an event is spatially correlated, previous work distinguish fault and event by exchanging readings among neighboring sensors. Considering that in many cases, an event is both spatially and temporally correlated, this paper proposes a distributed and localized algorithm for fault-tolerant event region detection. Aiming at reducing the network traffic, this algorithm determines a faulty sensor by using statistical hypothesis test for matching the reading sequence of sensors and statistical characters of the event. The analysis shows that the proposed algorithm is more energy-efficient than existing ones. The simulation results show that the algorithm can detect as much as 93% of the event region and 88% of faults, when 10% of sensors are faulty.

Keywords wireless sensor network; fault-tolerance; event region detection; energy-efficient

1 引 言

事件区域检测是无线传感器网络的一种重要应

用. 尤其是对于一些紧急事件, 例如化学物质泄露、火灾等, 往往希望尽快确定事件发生或影响的区域. 无线传感器网络可以快速地覆盖事件可能发生的区域, 并通过传感器之间的协同处理, 检测出事件发生

的确切区域。但是,恶劣的部署环境和传感器本身的不可靠性往往导致传感器产生错误的读数,从而影响整个网络对事件区域的检测。传感器的读数是指传感器的采样值在经过处理之后的输出值。本文对读数和采样值不加区分。

本文关注那些具有特定特征的事件(例如化学物质泄露时高浓度的化学物质、伴随火灾的高温等),当这类事件发生时,相应传感器的采样值就会偏离正常情况的采样值。因此,可以认为当传感器的采样值与正常情况采样值的偏差超过一定阈值时,就意味着发生了事件或者传感器发生了错误。此外,这类事件通常发生在一个连续的区域,称为事件区域,即事件是空间相关的。对应的,没有发生事件的区域称为正常区域。传感器网络通过所有检测到事件的传感器组成的集合来标记事件区域。

传感器的错误可以分为两类:一类是误判错误(positive fault),即当环境处于正常状况时,传感器却报告有事件发生;另一类是失判错误(negative fault),即当有特定的事件发生时,传感器却没有报告事件的发生。

在无线传感器网络中,对于事件区域检测算法的研究主要有两个难点^[4]:其一是错误的传感器会影响整个网络对事件区域的检测;其二是由于传感器只有有限的能量,而消息通信通常消耗了传感器大部分的能量^[1],所以要尽量减少传感器网络中的消息通信。这就要求事件区域检测算法必须是一种分布式和局部的算法。

本文第2节介绍已有的相关工作,并基于与以往不同的假设提出一种新的容错算法;第3节讨论事件区域检测原理,并给出本文使用的符号的定义;第4节给出一种分布式的事件区域检测的容错算法以及性能分析;第5节是模拟实验和实验结果;最后是本文的小结。

2 相关工作

近几年来,在无线传感器网络中,事件区域检测的容错问题逐渐成为研究热点^[2-4,7]。其中,文献[2]最早提出了事件区域检测的一种分布式的局部算法(Bayesian Fault Recognition Algorithms, BFRA)。BFRA假设事件是空间相关的而错误是空间不相关的,而且每个传感器发生错误的概率相同。传感器的读数是0或1。当输出为1时,则认为发生了事件或错误。作为一种局部算法,BFRA仅需要每个传感器

与自己的邻居交换读数(邻居定义为位于该传感器的无线通信半径内的所有传感器),从而获得所有邻居中检测到事件发生的统计概率。BFRA利用这个统计概率描述事件的空间相关性,并连同传感器的错误概率共同作为判断事件发生的条件概率。最后再根据该传感器的读数,利用Bayesian分析的方法,最终决定是否发生了错误或者事件。

文献[3-4]都是基于文献[2]的工作,其中文献[3]改进了文献[2]中理论分析部分的一些错误。文献[4]则改进了文献[2]的算法,考虑了由于传感器误差和错误引起的对事件判断的错误,并且考虑了如何选择合适的邻居数目,以达到容错的同时减少数据交换的目的。文献[2-4]都是基于概率分析的方法,需要假设每个传感器发生错误的概率相同,而且需要传感器进行复杂的运算。文献[7]提出了一种相对简单的非概率的检测算法,通过两轮投票以及少数服从多数原则(majority-voting),确定是否发生错误或者事件。

也有部分研究工作关注于事件边界的检测^[5-6]。事件的边界是指事件区域与正常区域的分界线。如果一个传感器的邻居中既有异常的读数(事件),又有正常的读数,那么就认为该传感器位于事件边界上。已有的事件边界检测算法依然是基于事件是空间相关的假设,同时,采用概率分析的方法。

还有一类工作关注于错误诊断^[8],这类工作关注的是如何确定错误发生的范围或者引发错误的错误源,而不是事件发生的范围。

本文关注的是事件区域检测的高能效的容错算法,也就是考虑如何消耗传感器尽可能少的(计算和通信)资源,在部分传感器发生错误的情况下,实现事件区域的检测。如前所述,已有的相关工作都是基于事件是空间相关而错误是空间不相关的假设,并利用传感器网络在空间上的信息冗余进行容错,但是需要传感器之间频繁地交换读数。

然而,在现实情况中,事件不仅具有空间相关性,而且具有时间相关性(例如火灾发生时地表温度的变化)。也就是说,事件在发生后将持续一段时间,并且事件的特征随时间的变化规律具有某些统计特征。因此,可以将事件特征随时间的变化规律用随机过程来描述,简称为事件过程。并利用传感器采集的数据在时间上的信息冗余进行容错,从而减少传感器之间的数据交换。

此外,已有的相关工作都假设每个传感器发生错误的概率相同,但是现实情况中,有很多因素会导

致传感器发生错误的概率不同,包括:(1)不同批次或工厂制造的传感器发生错误的概率往往不同;(2)在传感器网络部署的过程中,传感器可能会受到与环境相关的损坏,从而导致传感器发生错误的概率不同.例如,利用飞机空中布撒的传感器网络,不同的地面状况会对传感器产生不同程度的损坏;(3)传感器发生错误的概率可能受环境变化的影响.其中,有些因素会导致在实际部署前,无法获知传感器发生错误的概率.

与已有工作不同,本文从事件的空间相关性和时间相关性入手,利用随机过程描述事件的时间相关性,提出了一种基于统计假设检验的事件区域检测的分布式和局部的容错算法.

由于传感器总是周期性地读取数据,因此可以将传感器的读数看作一个按时间次序排列的序列,而位于事件区域中的传感器,其读数序列就是事件过程的一次实现.传感器可以通过统计假设检验的方法(即检验其本地读数序列与事件统计特征的符合程度)检测错误,因此可以减少传感器之间因频繁地交换读数带来的通信开销,从而有效地利用传感器的能量.

本文无需假设传感器发生错误的概率相同,因为传感器可以通过检验本地读数序列来识别错误.

3 事件区域检测原理与术语定义

本文假设有一个大规模的无线传感器网络均匀地覆盖某个感兴趣的区域,用于检测此区域内的某个特定事件.网络中共有 n 个传感器.每个传感器 S_i 可以通过定位设备(例如 GPS)感知自己的位置 x_i , $1 \leq i \leq n$. 传感器之间通过底层机制保证同步的数据采样和消息通信.

传感器每隔 ΔT 时间进行一次采样.当传感器 S_i 的读数超过阈值 R_{th} 时,说明该特定事件在 x_i 发生或者 S_i 发生了错误. R_{th} 的取值通常根据式(1):

$$R_{th}(t) = \frac{Exp_n(t) + Exp_e(t)}{2} \quad (1)$$

其中, $Exp_n(t)$ 是正确的传感器在正常区域中读数的期望函数,在比较稳定的环境(例如自然环境中的温度)中,可以认为 $Exp_n(t)$ 是一个常数; $Exp_e(t)$ 是正确的传感器在事件区域中读数的期望函数. t 是传感器采样的时刻.考虑到事件的特征只有在事件发生后才能被传感器感知,并且事件特征的变化规律与事件发生的具体时刻无关,只与距离事件发生时刻

的时间间隔有关.因此,假设事件在 t_0 时刻发生,并且事件(特征)至少可以持续 T_{th} 时间,则事件过程的期望和方差函数分别满足 $Exp_e(t) = Exp_e(t - t_0) = Exp_e(\tau)$, $Var_e(t) = Var_e(t - t_0) = Var_e(\tau)$, $0 \leq \tau = t - t_0 \leq T_{th}$. 并且定义 $Exp_e(0)$ 为事件刚刚发生时事件特征的期望值, $R_{th}(0)$ 作为判断事件是否发生的阈值.一般情况下, T_{th} 越大对事件的判断就越准确.

将传感器的周期性读数视作一个按照时间次序排列的序列 $\{r(t)\}$, 该读数序列本质上是事件过程的一个样本值.因此,对该序列的理论研究是以对应的随机过程的研究为基础的.虽然有限维分布函数族完全刻画了随机过程的统计特性,但是在实际问题中,要确定随机过程的有限维分布族往往是非常困难的.因此,常转而研究随机过程的某些重要数字特征,如期望函数和方差函数.

考虑到传感器只有有限的计算和存储能力,因此在传感器上只保存 T_{th} 时间内的 $m = T_{th}/\Delta T$ 个期望函数值和方差的平方根函数值.为了计算的简便,假设 ΔT 可以整除 T_{th} , 而且事件过程是一个二阶矩过程,即事件过程的期望函数和方差函数都存在.由于传感器网络通常用于检测特定的事件,因此假设期望函数 $Exp_e(t)$ 和方差平方根函数 $VarSR_e(t)$ 可以在部署前保存在传感器内存中或是在部署后通过汇点(sink node)将 $Exp_e(t)$ 和 $VarSR_e(t)$ 通过消息分发给各个传感器.

基于统计假设检验的事件区域检测原理如下:

(1) 如何识别事件:传感器对事件的检测是基于统计假设检验的方法.如果传感器的读数序列 $\{r(t_j)\}$ 超过事件发生的阈值 $R_{th}(0)$, 并且在之后的 T_{th} 时间内, $\{r(t_j)\}$ 满足式(2)的次数超过 C , 则称 $\{r(t_j)\}$ 满足假设检验条件.式(2)代表传感器读数序列与事件过程统计特征的相关程度,称为假设检验条件.在 T_{th} 时间后,如果该传感器的所有邻居中的多数也检测到事件发生(因为事件是空间相关的),则认为事件发生.从后面的理论分析中可以看到, δ 和 C 的选择会对算法的效果产生影响.

$$\frac{|r(t_j) - Exp_e(t_j)|}{VarSR_e(t_j)} < \delta \quad (2)$$

以上检测机制对位于事件区域边界上的传感器并不适用,因为这类传感器的邻居对事件的检测总是不一致的.由于传感器均匀的分布,所以通常情况下位于事件区域边界上的传感器数目远小于位于事件区域内的传感器数目,因此可以忽略这部分传感

器对事件区域检测的影响。

(2) 如何识别误判错误: 如果传感器的读数序列 $\{r(t_j)\}$ 超过事件发生的阈值 $R_{th}(0)$, 但是不满足假设检验条件, 或者 $\{r(t_j)\}$ 满足假设检验条件, 但是 T_{th} 时间后该传感器的邻居中的多数未检测到事件发生, 则认为该传感器发生了误判错误。

(3) 如何识别失判错误: 如果传感器的读数 $\{r(t_j)\}$ 小于事件发生的阈值 $R_{th}(0)$ 或者 $\{r(t_j)\}$ 不满足假设检验条件, 但 T_{th} 时间后, 该传感器的邻居中的多数检测到了事件, 则认为该传感器发生了失判错误。

本文中使用的符号和术语定义见表 1。

表 1 符号定义	
符号	定义
n	传感器网络中传感器的总数。
N	每个传感器的邻居个数。
T_{th}	事件特征的持续时间。
ΔT	传感器每次相邻采样之间的时间间隔。
m	传感器在 T_{th} 时间内的采样次数。
$r(t)$	传感器在时刻 t 的读数。
t_0	事件发生的时刻。
$R_{th}(t)$	传感器读数的阈值函数。当传感器的读数超过该值时, 认为可能发生了事件或者错误。
$Exp_e(t)$	事件过程的期望值函数。
$VarSR_e(t)$	事件过程方差的平方根函数。
δ	假设检验条件中的阈值。当传感器的读数与事件过程统计特征的差别小于 δ 时, 才认为该读数符合事件过程统计特征。
C	统计假设检验方法中的阈值。只有当传感器读数序列满足式(2)的次数超过 C , 才认为该读数序列符合事件过程。
LN	传感器的状态。代表传感器可能位于正常区域中, 是传感器的初始状态。
LE	传感器的状态。代表传感器可能位于事件区域中。
LF	传感器的状态。代表传感器可能发生了错误。
$Event$	传感器的状态。代表传感器检测到了事件。
$Fault$	传感器的状态。代表传感器发生了错误。

4 事件区域检测的容错算法和性能分析

本节将给出一种基于统计假设检验的区域检测的容错算法。算法中首先通过检验传感器的本地读数序列与事件过程的符合程度, 决定传感器的可能状态, 然后通过邻居传感器之间的信息交换确定传感器的最终状态。在本节的后半部分中, 将给出关于算法的容错能力和能量消耗的分析。

4.1 事件区域检测的容错算法

传感器每隔 ΔT 时间进行一次采样, 同时执行算法 1 中的算法。传感器的初始状态为 LN 。根据第 3 节中介绍的识别事件和错误的原理, 算法分为两

个阶段, 阶段 1 和阶段 2 分别是基于事件时间相关性和空间相关性的检测。

当传感器读数 $r(t)$ 大于 $R_{th}(0)$ 时, 即检测到事件发生, 算法进入阶段 1。如果传感器在 T_{th} 时间内的读数序列 $\{r(t)\}$ 满足假设检验条件, 则认为传感器可能位于事件区域内, 并将传感器的状态标记为 LE ; 否则认为传感器可能发生了错误, 将其状态标记为 LF 。

阶段 1 之后, 状态为 LE 的传感器利用无线信道向所有邻居广播自己的状态, 并且每个传感器收集其邻居发出的广播消息, 之后算法进入阶段 2。在阶段 2 中, 传感器根据邻居的状态, 即传感器网络在空间上的冗余信息, 确定自己的状态。即, 如果传感器发现大多数邻居的状态与自己的状态不同, 则认为自己发生了错误。

算法 1. 事件区域检测的容错算法。

```
//For each sensor;
int eventRegionDetection(double  $\delta$ , int  $C$ )
{
    double  $t, t'$ ;
    int  $status = LN$ ;
    int  $count = 0$ ;
    //detecting the occurrence of the event
    while (every  $\Delta T$  time and  $r(t) < R_{th}(0)$ )
        do nothing...
     $t' = 0$ ;
    //phase one—Upon detecting the occurrence of the event;
    while (every  $\Delta T$  time and  $t' < T_{th}$ )
    {
        if ( $r(t) > R_{th}(t')$ ) {
            //statistical hypothesis test
            if ( $|r(t) - Exp_e(t')| < \delta \cdot VarSR_e(t')$ )
                 $count = count + 1$ ;
        }
         $t' = t' + \Delta T$ ;
    }
    if ( $count > C$ )
         $status = LE$ ;
    else
         $status = LF$ ;
    if ( $status = LE$ ) {
        broadcasting status to all neighbors...
    }
    receiving statuses from neighbors ...
    //phase two—Upon judging the status itself;
    if (more than half of statuses of all the neighbors are  $LE$ ) {
        if ( $status = LN$ )
             $status = Fault$ ;
        else
             $status = Event$ ;
    }
    else {
```

```

if (status ≠ LN)
    status = Fault;
}
reset();
return status;
} // end eventRegionDetection

```

4.2 性能分析

从算法 1 中的算法可以看出,传感器检测到事件的必要条件是传感器的读数序列满足假设检验条件.因此,按照算法阶段 1 进行的事件区域检测可能产生两类错误.第 1 类是将检测到事件的传感器误认为错误的传感器;第 2 类是将错误的传感器误认为检测到事件的传感器.实际应用中,总是希望这两类错误的概率都尽可能的小.

产生第 1 类错误的原因是事件过程的一次实现本身偏离其统计特征,使得传感器读数序列未满足假设检验条件,即在 T_{th} 时间内,式(2)被满足的次数未超过 C 次.因此,传感器发生第 1 类错误的概率如式(3)所示,其中 $\{r(t_j)\}$ 代表事件过程.从式(3)可以看出, δ 越大、 C 越小,则发生第 1 类错误的概率越小.

产生第 2 类错误的原因是错误传感器的读数序列恰好符合事件的统计特征,使得该读数序列满足假设检验条件.因此,传感器发生第 2 类错误的概率如式(4)所示.其中,为了区别事件与错误,将错误传感器的读数记为 $r'(t_j)$, $\{r'(t_j)\}$ 代表错误随机过程.从式(4)可以看出, δ 越小、 C 越大,则发生第 2 类错误的概率越小.

$$\begin{aligned}
 &P\left(\frac{|r(t_{j_1}) - Exp_e(t_{j_1})|}{VarSR_e(t_{j_1})} \geq \delta, \right. \\
 &\quad \frac{|r(t_{j_2}) - Exp_e(t_{j_2})|}{VarSR_e(t_{j_2})} \geq \delta, \\
 &\quad \dots, \\
 &\quad \left. \frac{|r(t_{j_{m-C}}) - Exp_e(t_{j_{m-C}})|}{VarSR_e(t_{j_{m-C}})} \geq \delta\right), \\
 &1 \leq j_1 < \dots < j_{m-C} \leq m
 \end{aligned} \quad (3)$$

$$\begin{aligned}
 &P\left(\frac{|r'(t_{j_1}) - Exp_e(t_{j_1})|}{VarSR_e(t_{j_1})} < \delta, \right. \\
 &\quad \frac{|r'(t_{j_2}) - Exp_e(t_{j_2})|}{VarSR_e(t_{j_2})} < \delta, \\
 &\quad \dots, \\
 &\quad \left. \frac{|r'(t_{j_C}) - Exp_e(t_{j_C})|}{VarSR_e(t_{j_C})} < \delta\right), \\
 &1 \leq j_1 < \dots < j_C \leq m
 \end{aligned} \quad (4)$$

式(3)与式(4)对 δ 和 C 的要求是矛盾的,因此

必须做出平衡.通常,对于事件区域的检测更加重要,所以应该首先调整参数 δ 和 C ,尽量降低第一类错误的概率,然后利用事件空间相关性检验来降低第二类错误的概率.

对于 δ 和 C 的取值,需要根据实际应用的需求来确定.例如,有些事件的发生是受各类复杂因素的影响,因此,可以假设事件过程是一个 Bernoulli 过程,并且任意时刻事件过程的随机变量符合正态分布.于是,可以将事件过程的随机变量 $r(t)$ 化简为符合标准正态分布的随机变量:

$$\begin{aligned}
 p &= P\left(\frac{|r(t_j) - Exp_e(t_j)|}{VarSR_e(t_j)} \geq \delta\right) \\
 &= 1 - P\left(-\delta < \frac{r(t_j) - Exp_e(t_j)}{VarSR_e(t_j)} < \delta\right) \\
 &= 1 - (\Phi(\delta) - \Phi(-\delta)) \\
 &= 2 - 2 \cdot \Phi(\delta), 1 \leq j \leq m
 \end{aligned} \quad (5)$$

其中, $\Phi(\delta)$ 是标准正态分布.如果要求第 1 类错误的概率小于 θ ,即式(3)的概率小于 θ ,则

$$\begin{aligned}
 &P\left(\frac{|r(t_{j_1}) - Exp_e(t_{j_1})|}{VarSR_e(t_{j_1})} \geq \delta, \dots, \right. \\
 &\quad \left. \frac{|r(t_{j_{m-C}}) - Exp_e(t_{j_{m-C}})|}{VarSR_e(t_{j_{m-C}})} \geq \delta\right) = \\
 &\quad \binom{m}{C} \cdot (1-p)^C \cdot p^{m-C} < \theta
 \end{aligned} \quad (6)$$

因此,可以根据 θ 以及式(5)和(6)计算 δ 和 C .

由于传感器通常针对特定的应用设计,所以假设错误传感器的读数取值于区间 $[Exp_n(t), Exp_e(t)]$.并且,假设错误过程是一个宽平稳过程,且任意时刻错误随机过程的随机变量符合区间 $[Exp_n(t), Exp_e(t)]$ 内的均匀分布.于是有

$$P\left(\frac{|r'(t) - Exp_e(t)|}{VarSR_e(t)} < \delta\right) = \frac{\delta \cdot VarSR_e(t)}{Exp_e(t) - Exp_n(t)} \quad (7)$$

从式(7)可以看到,事件的特征越显著(即 $Exp_e(t)$ 与 $Exp_n(t)$ 的差越大),则发生第 2 类错误的概率就越小.当取定 δ 和 C 的值后,可以通过式(4)计算发生第 2 类错误的概率.

已有的事件区域检测的容错算法都是利用相邻传感器之间的数据交换实现容错.所以,传感器网络在 T_{th} 时间内产生的消息量为 $m \cdot n \cdot N$.其中,传感器与其邻居进行数据交换的频率为 $1/\Delta T$,每个传感器的邻居数目是 N , $m = T_{th}/\Delta T \geq 1$.由于传感器的读数为 0 或 1,所以忽略了消息数据包的大小.

本文提出的算法利用传感器本地读数的时间相

关性进行事件区域检测, T_{th} 时间后, 只有位于事件区域中的传感器才需要与其邻居进行数据交换, 所以传感器进行数据交换的频率为 $1/T_{th}$. 传感器网络在 T_{th} 时间内产生的消息量为 $\alpha \cdot n \cdot N$. α 代表事件区域面积与传感器的部署区域面积的比值, $0 < \alpha \leq 1$. 由于传感器只需要交换状态信息, 所以忽略了数据包的大小.

本文提出的算法产生的消息量与已有算法产生的消息量之比为 $\alpha/m \leq 1$. 所以, 当事件特征的持续时间 T_{th} 大于传感器的采样间隔 ΔT 或者事件区域小于传感器网络的部署区域时, 该算法产生的消息量小于已有算法产生的消息量. 与已有算法相比, 该算法只需其 α/m 的消息量, 因此是高效的.

5 实验结果与结果分析

模拟实验中, $n=1024$ 个传感器均匀地部署在 $32a \times 32a$ 的区域中. 假设传感器的无线通信半径为 $\sqrt{2}a$, 则每个传感器有 $N=8$ 个邻居. 由于假设事件具有显著的特征, 所以令 $Exp_n(t)=30$ 为常数.

假设事件过程 $\{r(t)\}$ 和错误随机过程 $\{r'(t)\}$ 分别满足 $r(t) \sim N(\mu, \sigma^2)$ 和 $r'(t) \sim U(Exp_n(t), Exp_e(t))$, 对 $\forall t$, 模拟实验中取 $\mu=100, \sigma=10$.

假设传感器的采样频率为 10Hz, 即 $\Delta T=0.1s$, $T_{th}=1s, m=10$. 如果要求第一类错误的概率小于 5%, 即 $\theta=0.05$. 根据式(5)和(6)计算可得 $p<0.05$ 和 $C>7$; 由于, 当 $\Phi(\delta)>0.975$ 时, $p<0.05$. 查表求得当 δ 约大于 1.96 时, $\Phi(\delta)>0.975$. 所以, 算法中取 $\delta=1.96, C=8$.

图 1 是一次模拟实验的快照, 图中每一个点 $(x, y), 1 \leq x, y \leq 32$, 代表一个传感器. 图中黑色方块表示错误的传感器. 其中, 错误传感器的数目为 103 个, 约占传感器总数的 10%. 定义传感器的错误概率为错误的传感器数目与传感器总数的比值. 如

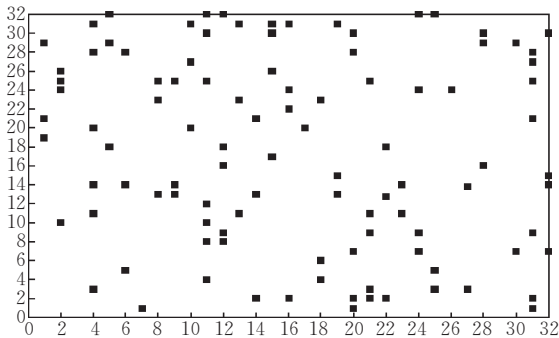


图 1 一次模拟实验的快照

果错误的传感器的最终状态为 Fault, 则认为该传感器识别了错误. 定义传感器的错误识别率为识别了错误的传感器数目与错误的传感器总数的比值.

事件发生在从(17,5)到(27,15)的 $11a \times 11a$ 的区域内. 模拟实验中, 如果位于事件区域中的传感器的最终状态为 Event, 则认为该传感器检测到了事件. 定义事件区域的检测概率为检测到事件的传感器数目与位于事件区域内的传感器总数的比值.

从图 2 中可以看出事件区域的检测概率随着传感器错误概率的增长, 近似呈线性递减. 最佳结果是当传感器错误概率低于 10% 时, 传感器网络可以检测到事件区域内 93% 的事件.

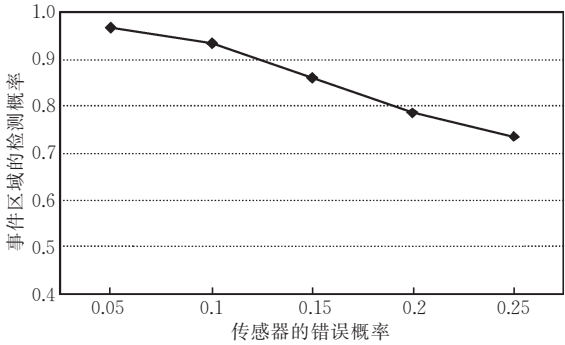


图 2 事件区域的检测概率随传感器的错误概率的变化规律

从图 3 中可以看出当传感器的错误概率增加时, 传感器的错误识别率保持在 90% 附近. 这是由于, 算法可以利用传感器读数的时间相关性识别错误, 所以即使传感器网络中错误传感器的数目增多, 算法依然可以保证 90% 左右的错误识别率.

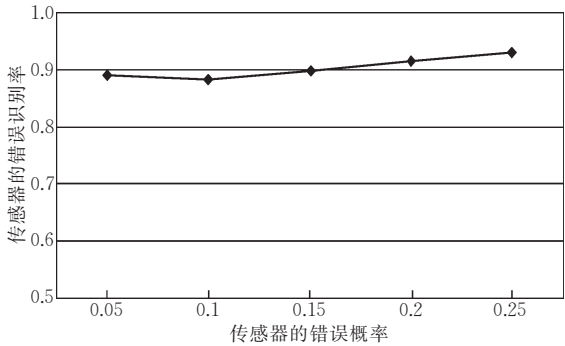


图 3 传感器的错误识别率随传感器的错误概率的变化规律

6 小 结

本文研究无线传感器网络中, 事件区域检测的高能效的容错算法. 已有的研究工作都是基于事件的空间相关性, 利用相邻传感器之间的数据交换实现容错, 但是这种方法会增加传感器网络的通信量.

为了延长传感器网络的寿命,本文考虑事件发生时,传感器的读数不但具有空间相关性,而且具有时间相关性.传感器可以利用本地读数之间的时间相关性实现容错.而且,当传感器网络中发生大规模错误时,本文提出的算法依然可以保证较高的错误识别率.

参 考 文 献

[1] Ian F A, Weilian S et al. A Survey on Sensor Networks. IEEE Communication Magazine, 2002, 40(8): 102-114

[2] Krishnamachari B, Iyengar S. Distributed Bayesian algorithms for fault-tolerant event region detection in wireless sensor networks. IEEE Transactions on Computers, 2004, 53(3): 241-250

[3] Chen Q, Lam K Y, Fan P. Comments on "distributed Bayesian algorithms for fault-tolerant event region detection in wireless sensor networks". IEEE Transactions on Computers, 2005, 54(9): 1182-1183



CAO Dong-Lei, born in 1980, Ph. D. candidate. His research interests include distributed computing, mobile computing and sensor networks.

CAO Jian-Nong, born in 1960, Ph. D. , professor. His

Background

Event region detection is an important application of wireless sensor networks (WSNs), especially when the monitored events are harmful to human beings, such as harmful chemical pollution and forest fire. In these scenarios, wireless sensor networks can be tasked to answer queries about events in the monitored environment. Especially one particular class of queries is important: Determining the event regions in the environment with a distinguishable characteristic.

As sensor nodes in a WSN are small devices with limited energy, they are prone to failures and energy-exhausting. As a result, the two fundamental challenges in event region detection problem are sensor faults and limited energy supply.

To distinguish faults and event, previous work assumed that the event is spatially correlated. Neighboring sensors

[4] Luo X, Dong M, Huang Y. On distributed fault-tolerant detection in wireless sensor networks. IEEE Transactions on Computers, 2006, 55(1): 58-69

[5] Ding M, Chen D et al. Localized fault-tolerant event boundary detection in sensor networks//Proceedings of the Annual IEEE Conference on Computer Communications (INFOCOM). Miami, 2005, 2: 902-913

[6] Li C, Liang C. A fault-tolerant event boundary detection algorithm in Sensor Networks//Proceedings of the IEEE Wireless Communications & Networking Conference (WCNC). Hong Kong, 2007

[7] Chen J, Kher S, Somani A. Distributed fault detection of wireless sensor networks//Proceedings of the Workshop on Dependability Issues in Wireless Ad Hoc Networks and Sensor Networks (DIWANS). Los Angeles, 2006: 65-72

[8] Sheth A, Hartung C, Han R. A decentralized fault diagnosis system for wireless sensor networks//Proceedings of the IEEE International Conference on Mobile Adhoc and Sensor Systems (MASS). Washington, 2005

research interests include parallel and distributed computing, networking, mobile computing, fault tolerance, and distributed programming environments.

JIN Bei-Hong, born in 1967, Ph. D. , professor. Her research interests include distributed computing, software engineering methodology and environment, and mobile computing.

can exchange their readings to determine the event region by using the Bayesian algorithm or majority-voting. But these message exchanges will consume much energy of the sensor nodes and decrease the lifetime of the WSN.

This paper assumes that the event is both spatially and temporally correlated as it is the case in many applications. As a result, the event can be modeled as a stochastic process called event process. In addition, each sensor periodically measures the environment and hence its readings are recognized as a reading sequence. Given the statistical characteristic of the event process, each sensor can locally test its reading sequence against this statistical characteristic to detect faults.