

传感器网络中基于虚拟坐标的节点调度方案*

李小龙¹, 林亚平^{1,2+}, 易叶青¹, 余建平¹, 卢新国¹

¹(湖南大学 计算机与通信学院, 湖南 长沙 410082)

²(湖南大学 软件学院, 湖南 长沙 410082)

A Node Scheduling Scheme Based on Virtual Coordinate in Sensor Networks

LI Xiao-Long¹, LIN Ya-Ping^{1,2+}, YI Ye-Qing¹, YU Jian-Ping¹, LU Xin-Guo¹

¹(School of Computer and Communication, Hu'nan University, Changsha 410082, China)

²(School of Software, Hu'nan University, Changsha 410082, China)

+ Corresponding author: E-mail: yplin@hnu.com, http://www.hunu.edu.cn

Li XL, Lin YP, Yi YQ, Yu JP, Lu XG. A node scheduling scheme based on virtual coordinate in sensor networks. Journal of Software, 2008,19(8):2089–2101. <http://www.jos.org.cn/1000-9825/19/2089.htm>

Abstract: Firstly a maximum similarity distribution model is proposed. Secondly, a new definition, named virtual coordinate of a node which composes of the minimum hop counts to several special nodes, is introduced in instead of absolute physical coordinate of a node. Based on the theories, a distributed location-independent node scheduling scheme is proposed. The scheme consists of a coverage algorithm and a connection algorithm. The coverage algorithm takes advantage of the nodes' virtual coordinate to divide all nodes into several subsets. Without using location information, this scheduling scheme not only has the ability that sensor nodes in subsets are more uniformly distributed in the target region than other schemes, but also guarantees that all subsets are connective. The simulation results show that the scheme outperforms the randomized node scheduling scheme, on the coverage rate and network lifetime, as well as the number of external nodes when maintaining subsets to be connective.

Key words: sensor network; scheduling; maximum similarity distribution; coverage; connectivity

摘要: 首先,建立了极大相似分布模型;其次,引入节点的虚拟坐标的概念,用来替代节点实际的位置坐标.在此基础上,提出了一种分布式的、与节点位置无关的节点调度方案.方案由一个覆盖算法和一个连通算法组成,方案中的覆盖算法利用虚拟坐标信息进行分组.在不需要节点位置信息的条件下,使各个分组内的传感器节点较为均匀地分布在目标区域.该方案不仅解决了任意分组中的节点充分覆盖目标区域的问题,而且保证了网络的连通.仿真实验表明,该节点调度方案在覆盖率、维持分组连通时额外加入到分组内的节点个数以及网络生存时间等性能上均优于与节点位置无关的节点随机调度协议.

关键词: 传感器网络;调度;极大相似分布;覆盖;连通

中图法分类号: TP393 文献标识码: A

* Supported by the National High-Tech Research and Development Plan of China under Grant No.2006AA01Z227 (国家高技术研究发展计划(863)); the Natural Science Key Foundation of Hu'nan Province of China under Grant No.06JJ20049 (湖南省重点自然科学基金)

Received 2007-01-29; Accepted 2007-04-25

无线传感器网络是由大量计算、通信及存储能力有限的传感器节点组成的特殊 Ad-hoc 网络,在军事和民用领域应用广泛,如战场监视、环境和交通监测、灾难救助等,是目前一个非常活跃的研究领域^[1,2].传感器节点体积较小,节点提供的能量有限,且能量难以补充,如何节约节点能量、最大化网络生命周期是传感器网络研究的中心问题^[3].由于传感器网络密度高,如果每个节点都处于工作状态,则会造成信息冗余,增加网络能量消耗,降低网络通信效率.因此,在满足任意时刻处于活动状态下的节点能够对目标区域进行有效监控等条件下,可把所有节点划分为合适的节点子集,这些节点子集采用轮换活跃/休眠的节点调度方式,达到减少能量消耗和延长网络生存时间的目的.节点调度算法根据调度的特点可分为两种:基于轮次的调度算法^[4-9]和基于分组的调度算法^[10].基于轮次的调度算法要求传感器节点在每一轮的开始执行一次算法,按照某种竞争机制,从所有节点中选择若干个节点作为活动节点,这种算法在传感器网络的生存时间内将执行多次;而基于分组的调度算法在传感器节点被部署之后仅执行 1 次,通过分组将每个传感器节点划分到某一组或者某几组(也就是传感器节点可以属于几个不同的分组),在算法完成之后,依次调度每一组的传感器节点作为活动节点.

目前,在调度算法中的覆盖问题和网络连通问题上已取得了不少研究成果,但是却很少有文献同时兼顾调度算法中的覆盖和网络连通问题,如文献[4-7]仅针对覆盖问题,而文献[8,9]只考虑网络连通问题.此外,现有的定位技术^[11,12]由于硬件成本、能耗以及误差范围的限制,难以保证每个节点获得自身精确的物理位置.当前,已有一些学者开始研究虚拟坐标技术^[13,14],用于解决未知节点位置信息的路由问题.如何在不能获取节点精确位置信息的条件下,设计一个能够充分覆盖目标区域并确保网络连通的节点调度方案,是当前具有挑战性的问题.文献[10]提出了基于分组的、与位置无关的分布式节点随机调度协议.随机调度协议将所有的传感器节点分为 K 组,每个传感器节点随机赋予 $1\sim K$ 的某个值 i ,将该节点分配到第 i 组,并通过让每个节点在所属的分组内建立一条到 Sink 的最短路径来构造连通网络.由于位置非常临近的传感器节点可能分配到同一组或者是某几组,因而导致目标区域内节点分布不均匀,目标区域中某些小区域节点分布过于密集,增加了该区域的通信冲突和数据冗余;而有些小区域节点分布过于稀疏,导致传感器节点不能对目标区域进行有效监控,并且降低了网络的连通性.

基于以上原因,本文从使各个分组内的传感器节点均匀分布在目标区域这一目的来考虑.首先,建立了极大相似分布模型,证明该问题属于 NP-难问题;其次,引入节点的虚拟坐标的概念以替代节点的实际位置坐标,从而消除对传感器节点位置信息的依赖.在此基础上,针对调度算法中的覆盖和网络连通问题,提出了一种分布式的、与节点位置无关的节点调度方案——基于虚拟坐标的节点调度方案.该方案包含两个算法:实施覆盖的覆盖算法和构造连通网络的连通算法.在实施覆盖部分,首先从传感器网络中选取若干个节点作为参照节点,然后每个传感器节点根据到各个参照节点的最小跳数得到自身的虚拟坐标.协议利用节点的虚拟坐标信息将所有节点划分为 K 个不相交的分组,在不需要节点位置信息的条件下,使各个分组内的传感器节点较为均匀地分布在目标区域.本方案通过连通算法将传感器节点中的某一些节点划分到几个不同的分组,确保各个分组的连通.此外,本文还分析了在节点随机分布的条件下,参照节点的个数与节点的通信范围和传感范围之间的关系.算法分析和仿真实验表明,本文提出的方案在覆盖率、维持分组连通时额外加入到分组内的节点个数,以及网络生存时间等性能上均优于与节点位置无关的节点随机调度协议.

1 问题描述

1.1 网络模型

传感器节点往往随机分布在目标区域,即节点在部署之前,预先不知道自己在目标区域内的位置.假设所形成的传感器网络具有如下性质:

- (1) 传感器节点部署密集;
- (2) 传感器节点的感知模型为圆形区域感知模型,节点的通信半径和传感半径均相同;
- (3) 传感器节点的通信能力有限,通信能耗与距离的 n 次方($2 < n < 4$)成正比;
- (4) 所有节点部署后不再移动.

1.2 问题分析

由于传感器节点随机分布在目标区域,因此目标区域中的小区域之间的节点密度分布稠密不均.假设某个较小的区域 ϵ 内分布了 $K \times L (K \geq 1, L \geq 1)$ 个传感器节点(这些节点的覆盖区域可近似地看作相同),如果将所有传感器节点分为不相交的 K 组,为了使 K 组之间的传感器节点分布均匀,这 $K \times L$ 个传感器节点应该平均分配到 K 组.采用随机调度协议^[10],每组恰好包含了 L 个分布在区域 ϵ 内的传感器节点的概率见式(1).

$$P(K) = P(S_1 = L, \dots, S_K = L) = \begin{cases} 1, & K = 1 \\ \prod_{i=K}^1 \binom{L \cdot i}{L} \cdot \frac{1}{i^L} \cdot \left(1 - \frac{1}{i}\right)^{L \cdot i - L}, & K > 1 \end{cases} \quad (1)$$

其中, S_i 为第 i 组中的节点个数, $i=1,2,\dots,K$.由式(1)可知,当 $L=1$ 时,分组数 K 逐渐增大,每组恰好包含了 1 个分布在区域 ϵ 的传感器节点的概率迅速减小.当 $K=3$ 时,概率小于 25%,当 $K=7$ 时,概率几乎为 0.也就是说,存在着一些分组在区域 ϵ 内没有分配到传感器节点,导致在某些时刻处于活动状态下的节点不能有效覆盖区域 ϵ .由于在区域 ϵ 内分布的节点个数为固定值 K ,显然可以推出,存在着一些分组包含了两个或两个以上分布在区域 ϵ 中的节点,导致在某些时刻区域 ϵ 内处于活动状态下的节点分布过于密集,此时,区域 ϵ 内的节点通信冲突较大,获得的传感数据冗余度高.如何在不需要节点位置信息的条件下,对于任意大小的 K 和 $L (K \geq 1, L \geq 1)$,使得这 K 个分组内的传感器节点在目标区域内分布较为均匀,是本文需要解决的关键问题.

1.3 极大相似分布模型

将 N 个传感器节点分为 K 组,如何对 K 组之间分布的均匀程度进行量化是本文考虑的问题之一.在给出具体的数学模型之前,首先给出一些定义.

定义 1. 假设 N 个传感器节点分布在目标区域内,已知传感器节点 i 的位置坐标 $(x_i, y_i), i=1,2,\dots,N$.对于传感器节点 $j (j \in P = \{1,2,\dots,N\})$,到节点 j 距离最短的节点称为节点 j 在节点集 P 中的最佳匹配节点.

定义 2. 假设均包含了 L 个节点的节点集 $A = \{a_1, a_2, \dots, a_L\}$ 和节点集 $B = \{b_1, b_2, \dots, b_L\}$ 分布在同一目标区域, $A \cap B = \emptyset$.对于所有的 A 和 B 之间的匹配,当 $\sum_{i=1}^m |a_i - b'_i|$ 达到最小值时,称该值为节点集 B 到节点集 A 的距离,用 $|A-B|$ 来表示,如式(2)所示.

$$\begin{cases} |A-B| = \min \sum_{i=1}^m |a_i - b'_i| \\ |a_i - b'_i| = \sqrt{(x_i - x'_i)^2 + (y_i - y'_i)^2} \end{cases}, i = 1, 2, \dots, L \quad (2)$$

其中, b'_1, b'_2, \dots, b'_L 各不相同,且 $b'_i \in B, i=1,2,\dots,L$.

定义 3. 假设一个由 $K \times L$ 个传感器节点组成的传感器网络 S 随机分布在目标区域,传感器节点 i 的位置坐标为 $(x_i, y_i), i=1,2,\dots,K \times L$.将所有的传感器节点平均分为 K 组,传感器网络的分布相似度等于第 i 组到第 j 组的距离之和, $i, j \in \{1,2,\dots,K\}$, 且 $i \neq j$.即分布相似度等于式(3).

$$\sum_{i=1}^K \sum_{j=1, j \neq i}^K |X_i - X_j|, \|X_i\| = L, i = 1, 2, \dots, K, K \geq 2, S = \{X_1, X_2, \dots, X_K\} \quad (3)$$

当任意一组包含的传感器节点不同时,可能导致传感器网络的分布相似度增大或者减小.当传感器网络的分布相似度最小时,称这种分配达到了极大相似分布,此时,分布相似度为式(4).

$$\min \sum_{i=1}^K \sum_{j=1, j \neq i}^K |X_i - X_j|, \|X_i\| = L, i = 1, 2, \dots, K, K \geq 2, S = \{X_1, X_2, \dots, X_K\} \quad (4)$$

在证明式(4)的求解是 NP 难问题之前,首先证明式(2)的求解是 NP 难问题.

定理 1. 当节点集 A 和节点集 B 包含了相同个数的传感器节点,且所有节点的位置坐标已知时,节点集 B 到节点集 A 的距离问题是 NP 难问题.

证明:略. □

定理 2. 当 K 为大于等于 2 的整数值时,传感器网络的极大相似分布问题是 NP 难问题.

证明:根据定理 1,定理 2 显然成立. □

2 虚拟坐标和分析

在不能获取节点位置信息的条件下,在文献[13]中,算法根据节点到各个边界节点的最小跳数信息生成节点的虚拟坐标.由于边界节点的个数为 $O(\sqrt{N})$ (N 为网络中的节点个数),显然,这种虚拟坐标生成算法不具有良好的扩展性.在文献[14]中,节点的虚拟坐标由节点分别到 3 个锚节点的最小跳数组成,虽然适合于基于地理位置信息的路由算法,但难以达到覆盖算法的精度要求.本节给出了不同于已有工作的虚拟坐标的定义,并说明了在极大相似分布模型中,节点的虚拟坐标代替位置坐标的合理性.

2.1 参照节点和虚拟坐标的定义

定义 4. 如果传感器网络中存在一类特殊的节点,网络中的所有节点需要依次计算到这类节点中每一个节点的最小跳数,则称这类节点为传感器网络中的参照节点.参照节点到自身的最小跳数定义为 0.

定义 5. 如果传感器节点 A 到各个参照节点 P_1, P_2, \dots, P_m 的最小跳数分别为 N_1, N_2, \dots, N_m , 其中, m 为传感器网络中参照节点的个数,且 $ID_{P_1} < ID_{P_2} < \dots < ID_{P_m}$, 则称节点 A 的虚拟坐标是 (N_1, N_2, \dots, N_m) , 虚拟坐标的维数是 m .

2.2 虚拟坐标及分析

本文提出的是一种与节点位置无关的节点调度方案.在该方案中,传感器实际的位置坐标用虚拟坐标来代替.在本文中,用式(5)计算两个节点之间的距离:

$$|a_i - b'_j| = |N_1 - N'_1| + |N_2 - N'_2| + \dots + |N_m - N'_m| \quad (5)$$

其中, $(N_1, N_2, \dots, N_m), (N'_1, N'_2, \dots, N'_m)$ 分别为 a_i 和 b'_j 的虚拟坐标, m 为参照节点的个数.假设传感器网络中仅有两个参照节点:节点 A 和节点 B , 且 $ID_A < ID_B$, 在目标区域内的每一个点上部署了传感器节点, 节点的通信半径为 R_C , 如图 1 所示, 可以得到节点 C 和节点 D 的虚拟坐标均为 $(4, 4)$, 节点 E 的虚拟坐标为 $(4, 3)$. 从这个例子可以看出, 当参照节点的个数过少时, 式(5)的值不能反映任意两节点之间的实际距离, 如 DE 与 DC . 当参照节点的个数过多时, 由于每个节点需要计算到各个参照节点的最小跳数, 参照节点个数的增多会导致整个网络通信开销增大, 并且虚拟坐标维数的增加将造成计算开销增大. 因此, 需要选择合适个数的参照节点, 使得式(5)的值能够反映任意两节点之间的实际距离.

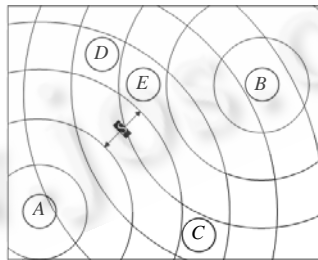


Fig.1 A scenario

图 1 示例

用随机分布的 m 个参照节点生成的虚拟坐标, 代替实际的位置坐标表示任意两个节点之间的距离, 需要满足的条件如下(假设节点 A 、节点 B 和节点 C 为传感器网络中的任意 3 个节点):

条件 1. 当 A, B 之间的距离 a 较大时, $P(|A-B| \geq j)$ 接近于 1 (j 为较大的正整数);

条件 2. 当 A, B 之间的距离 a_1 和 A, C 之间的距离 a_2 都较小时, 如果 $a_1 > a_2$, 则有如下不等式成立:

$$P(|A-B| \geq 1|a_1) \gg P(|A-C| \geq 1|a_2).$$

条件 3. 对于距离为 a_{\min} (a_{\min} 为用户定义的长度单元, 与节点的传感半径成正比) 的 A, B , 可通过 m 个参照节点区分开来, 即 $P(|A-B| \geq j)$ 接近于 1.

下面分析说明选择适当的 m 值能够满足上述条件.

假设任意两个节点 A, B 绝对的位置坐标分别为 $(x_0, y_0), (x_1, y_1)$, 节点之间的距离为 a . 随机选择节点 C 为参照节点, C 的位置坐标为 (x, y) , C 到 A 和 B 的距离之差的绝对值为 i . 在二维平面中, 若 $0 < i < a$, 则根据双曲线的定义, 与两固定点的距离之差为 i 的点的轨迹是一条双曲线; 若 $i=0$ 或 $i=a$, 则点的轨迹为一条直线, 可考虑成一条特殊的双曲线; 若 $i < 0$ 或 $i > a$, 则这样的点在平面内不存在. 距离之差的概率密度函数 $f(i)$ 与 $(x_0, y_0), (x_1, y_1)$ 有关, 也与目标区域的形状有关. 我们假设在二维平面内上述的双曲线的长度均相等, 此时, $f(i)$ 为

$$f(i) = \begin{cases} \frac{1}{a}, & 0 \leq i \leq a \\ 0, & \text{else} \end{cases} \quad (6)$$

当网络中的所有节点记录到参照节点的最小跳数时, 参照节点广播的分组, 以近似于同心圆的方式在网络中逐次传播. 当节点密度较大时, 可以近似地看成参照节点广播的分组以同心圆的方式在网络中逐次传播, 各个相邻的圆半径之差为 S (S 为节点的通信半径). 当参照节点 C 到 A 和 B 的距离之差为 i , 到 A 和 B 的最小跳数分别为 N_A 和 $N_B, i=KS+l$ (K 为自然数, $l < S$) 时, 最小跳数的差值服从如下概率分布:

$$\begin{cases} P(|N_A - N_B| = K + 1) = \frac{2l}{S + l} \\ P(|N_A - N_B| = K) = \frac{S - l}{S + l} \end{cases} \quad (7)$$

根据式(6)和式(7)可以得到, 随机分布的任意参照节点 C 到距离为 a 的任意两节点(假设为节点 A 和节点 B)最小跳数的差值为 0 的概率是

$$P(N_A - N_B = 0) = \begin{cases} \int_0^S \frac{S-x}{S+x} \cdot \frac{1}{a} dx = -\frac{S}{a} + 2S \ln 2 = \frac{S}{a} \ln \frac{4}{e}, & a \geq S \\ \int_0^a \frac{S-x}{S+x} \cdot \frac{1}{a} dx = -1 + \frac{2S}{a} \ln \left(1 + \frac{a}{S} \right), & a < S \end{cases} \quad (8)$$

当 a 和 S 取不同值时, $P(N_A - N_B = 0)$ 对应的值如图 2 所示. 从图 2 可以看出, 随着 a/S 的增大, $P(N_A - N_B = 0)$ 逐渐减小. 当 a/S 较大时(例如 $a/S=5$), $P(N_A - N_B = 0)$ 接近于 0. 也就是说, 当任意两个节点之间的距离 a 与 S 相比之下较大时, 对于一个随机分布的参照节点来说, $P(|N_A - N_B| \geq 1)$ 接近于 1. 当 m 较大时, 显然满足条件 1. 当 a 为 a_{\min} 时, 两个相距为 a_{\min} 的传感器节点具有不同的虚拟坐标, 即能被 m 个参照节点区分开来的概率为式(9):

$$P(|A - B| = |N_1 - N'_1| + |N_2 - N'_2| + \dots + |N_m - N'_m| \geq 1) = 1 - \left(-1 + \frac{2S}{a} \ln \left(1 + \frac{a}{S} \right) \right)^m \quad (9)$$

当两个相距为 a_{\min} 的传感器节点能被若干参照节点区分开来的概率大于 P_0 时, 参照节点的个数至少为 m , 如图 3 所示. 从图中可以看出, 在 $a_{\min}=0.25S$ 的条件下, 当网络中存在着 13 个参照节点时, 任意两个相距为 a_{\min} 的节点具有不同虚拟坐标的概率大于 0.95. 如图 2 和图 3 所示, 虚拟坐标模型能够满足条件 2 和条件 3. 我们根据节点的传感半径、通信半径以及应用要求选择参照节点的个数.

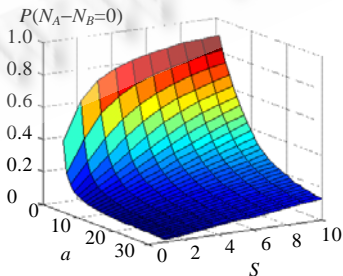


Fig.2 $P(N_A - N_B = 0)$ vs. a and S
图 2 $P(N_A - N_B = 0)$ 与 a, S 之间的关系

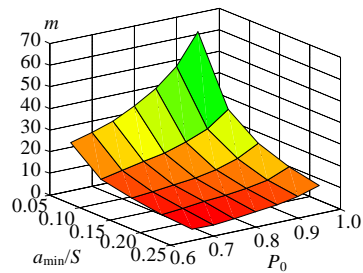


Fig.3 m vs. P_0 and a_{\min}/S
图 3 m 与 $P_0, a_{\min}/S$ 之间的关系

3 基于虚拟坐标的节点调度方案

3.1 覆盖算法

在给出方案中的覆盖算法之前,首先给出极大相似分布模型的一种集中式近似求解算法.

3.1.1 集中式覆盖算法

由定理 2 可知,极大相似分布的求解属于 NP 难问题.下面给出该算法的一般步骤:

初始化. 确定分组数 K 的大小.根据 S, a_{\min} 和 P_0 ,计算出参照节点的个数 m .

步骤 1. 随机选择 m 个传感器节点作为参照节点.

步骤 2. 每一个节点获取自身的虚拟坐标(参照分布式覆盖算法步骤 1.2).

步骤 3. 任意选择一个未分配的节点 i 分配到第 1 组.

步骤 4. 从未分配的节点中选择 i 的最匹配节点分配到第 2 组.

步骤 5. 从未分配的节点中选择 i 的最匹配节点分配到第 3 组.

...

步骤 $2+K$. 从未分配的节点中选择 i 的最匹配节点分配到第 K 组.

重复步骤 3~步骤 $2+K$,直到所有的节点都被分配为止.

3.1.2 基于虚拟坐标的节点调度方案中的覆盖算法

在传感器网络中,节点规模巨大,将所有传感器节点的虚拟坐标信息发送给某一个节点进行集中处理会大量增加整个网络的通信开销,不具有好的扩展性.本文在集中式覆盖算法的基础上,提出了基于虚拟坐标的节点调度方案中的分布式覆盖算法.

初始化. 确定分组数 K 的大小.根据 S, a_{\min} 和 P_0 计算出参照节点的个数 m ,得到 $Q_0=m/N$,其中, N 为网络中的节点个数.

步骤 1. 选择参照节点,得到节点的虚拟坐标.

步骤 1.1. 随机选择 m 个传感器节点作为参照节点.

步骤 1.1.1. 节点随机产生一个 0~1 之间的随机数 q .

步骤 1.1.2. 若 $q < Q_0$,则该节点为网络中的参照节点.

步骤 1.2. 若该节点为参照节点,则节点向邻居节点广播包含了自身 ID 和最小跳数信息的信息,最小跳数的初始值为 0.所有节点将记录到每个参照节点的最小跳数,忽略来自同一参照节点具有较大跳数的消息.然后将跳数值加 1,并转发给邻居节点.通过这种方法,传感器网络中的所有节点能够记录到每个参照节点的最小跳数.根据所有参照节点的 ID 大小顺序,节点将对应的最小跳数值组合起来构成它的虚拟坐标.

在步骤 1.2 中,每个节点至少需要向邻居节点广播 m 个分组.研究表明,节点通信时的能量消耗占节点总能耗的绝大多数,因此,减少传感器节点通信次数对于延长传感器节点的生命周期具有重要意义.步骤 1.2 可以与某些定位算法(如 DV-Hop 算法^[12]和 Amorphous 算法^[11])一并实现.这些定位算法的共同特点是,在其距离向量定位机制中,未知节点需要计算到信标节点的最小跳数.当网络中存在信标节点时,如果将这些节点作为网络中的参照节点,则本文的步骤 1.2 与之稍有不同的是,所有节点(包括未知节点和信标节点)都需要计算到每个信标节点的最小跳数.因此,如果传感器网络的定位技术采用了这种距离向量定位机制,假设信标节点的个数为 m_1 ,那么,当 $m_1 \geq m$ 时,得到节点的虚拟坐标不再需要额外的通信开销;当 $m_1 < m$ 时,每个节点至少可以减少广播 m_1 个分组.

步骤 2. 节点分组过程.

首先由 Sink 在邻近区域随机选择某一个节点作为工作节点,其他节点的初始状态均为准备状态.

A. 节点为工作状态时的算法描述

当节点 S_i 为工作状态时,工作节点首先在邻居区域内广播查询消息 IM(inquiry message).处于候选工作状态

或准备状态的邻居节点收到 S_i 的 IM 后,将自己的位置信息 PI(position information)发送给 S_i ,包括虚拟坐标和 ID 号.此时, S_i 在 T_w 时间内完成如下工作:在收到若干节点的 PI 消息后,从这些节点中选择 $K-1$ 个与自己最匹配的节点.将自己的分组号赋值为 1,把这 $K-1$ 个节点的分组号分别赋值 2,3,...,K,并记录 $K-1$ 个节点的 ID 和对应的分组号,然后把这些包括了分组号的分组 AMGN(announcement message of group number)发送给对应的节点.对于剩余节点, S_i 根据这些节点(不妨假设为 X)到 S_i 之间的距离分别计算这些节点的延迟时间 $T_d=|X-S_i| \times T_0$,并把包含了延迟时间的分组 AMDT(announcement message of delay time)发送给对应的节点.在完成上述工作后,节点 S_i 从工作状态转变成休眠状态.

B. 节点为接收状态时的算法描述

当节点 S_i 为接收状态时,将不再响应工作节点发送的查询请求.在 T_w 时间之内,节点一直等待接收 AMGN 或 AMDT.若节点 S_i 接收到分组号为 x 的 AMGN,则节点将自己的分组号赋值为 x ,之后,节点 S_i 从接收状态转变成休眠状态.若节点 S_i 接收到延迟时间为 T_d 的 AMDT,节点首先判断 T_d 与自身的等待时间 T_c 之间的大小关系:若 $T_c > T_d$,则节点将 T_c 赋值为 T_d ;若 $T_c \leq T_d$,则节点不改变 T_c 的值.之后,节点 S_i 转变成候选工作状态.

C. 节点为候选工作状态时的算法描述

当节点 S_i 为候选工作状态时,若在 T_c 完成之前没有收到查询消息,则节点 S_i 从候选工作状态转变成工作状态;若在 T_c 之前收到了查询消息,则该节点将等待时间延长 T_w ,以等待周围的工作节点完成上述工作过程,并将自身状态转变成接收状态.

D. 节点为准备状态时的算法描述

当节点 S_i 为准备状态时,如果某一时间收到了工作节点的查询请求,则节点 S_i 需要将自身状态信息 PI 发送给工作节点,并将等待时间设置为极大值.之后,节点 S_i 转变成接收状态.节点的状态之间的转换如图 4 所示.

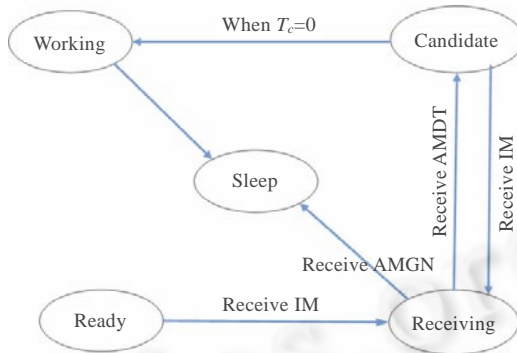


Fig.4 Node status transition graph

图 4 节点状态转换图

3.2 连通算法

通过覆盖算法,所有的传感器节点被划分到不相交的 K 个分组.然而,覆盖算法不能确保各个分组所构成的子网络都是连通的.因此,将有一些节点同时属于几个分组(举例来说,若节点 A 属于第 1、第 3 和第 4 组,当在第 1、第 3 和第 4 组中有一组活跃时,节点 A 均需要保持在活动状态),以确保各个分组是连通的,这是本算法的基本思想.在给出具体的连通算法之前,首先给出一些定义和定理.

定义 6. 若节点 A 同时属于 n 个不同的分组,那么称该节点的组数为 n .我们用 G_A 表示 A 所属的分组, $|G_A|$ 表示 A 的组数.若 $i \in G_A$,称 A 的分组号为 i .

定理 3. 假设目标区域部署了密集传感器节点,基于虚拟坐标的节点调度方案的覆盖算法将所有的传感器节点划分到不相交的 K 个分组时,第 1 分组构成的子网络是连通的,且第 1 分组内任意一个节点与其分配组号的 $K-1$ 个节点也是连通的.

证明:根据上述覆盖算法的算法描述,易证.故省略. □

通过定理 3,我们可以得到传感器网络内节点之间的连通关系,如图 5 所示.从图 5 可以看出,如果第 1 分组内的所有节点一直保持在活动状态,各个分组构成的子网络是连通的,任意时刻的覆盖任务仅需要 N/K 个节点.如果采用这种连通算法,除第 1 分组外,其他分组均额外需要 N/K 个节点构成连通网络.考虑分组内节点之间的连通关系:以图 5 第 2 分组中 $S_{2,i}$ 和 $S_{2,j}$ 为例,假设 $S_{2,i}$ 和 $S_{2,j}$ 互为邻居节点,不需要 $S_{1,i}$ 和 $S_{1,j}$ 属于第 2 分组以确保 $S_{2,i}$ 和 $S_{2,j}$ 是连通的.当新加入节点 $S_{i,n+1}$ 与分组 i 中的其他节点所构成的网络不连通时才启动第 1 分组中的某些节点是一种较好的办法.如果采用这种思想,在传感器网络监测目标区域的过程中,第 1 分组中的某些节点可能持续工作,导致传感器网络过早瘫痪死亡.在考虑分组内新加入节点与原有节点之间连通关系的同时,兼顾节点能耗的均衡性,本文提出了一种分布式连通算法.

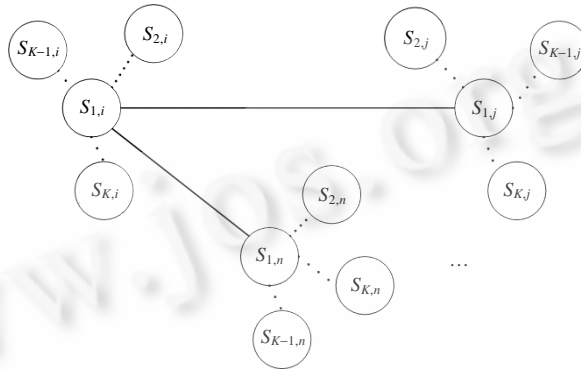


Fig.5 The connectivity relationship between nodes deduced from Theorem 3

图 5 通过定理 3 得到的节点之间连通关系

由定理 3 可以得出,分组 1 所构成的子网络是连通的,因此分组号为 1 的节点不执行方案中的连通算法.在一个节点被某工作节点分配了分组号 i 后,节点开始执行连通算法.不妨假设该节点为 $S_{i,n+1}$,对应的工作节点为 $S_{1,n+1}$,该分组内已有的其他的 n 个节点分别为 $S_{i,1}, S_{i,2}, \dots, S_{i,n}$.下面给出连通算法的实现步骤:

初始化. $S_{i,n+1}$ 的组数为 1, 分组号为 i .

步骤 1. 在 $\{S_{i,1}, S_{i,2}, \dots, S_{i,n}\}$ 中,判断是否存在着一个节点能与 $S_{i,n+1}$ 直接通信.

$S_{i,n+1}$ 向邻居区域广播查询消息:是否存在分组号为 i 的节点,如果存在,则结束;如果不存在,则执行步骤 2.

步骤 2. $S_{i,n+1}$ 和 $\{S_{i,1}, S_{i,2}, \dots, S_{i,n}\}$ 中与自己距离最近的节点建立连接.

步骤 2.1. 查询 $\{S_{i,1}, S_{i,2}, \dots, S_{i,n}\}$ 中哪一个节点与自己距离最近.

步骤 2.1.1. $S_{i,n+1}$ 请求 $S_{1,n+1}$ 广播查询消息.

步骤 2.1.2. $S_{1,n+1}$ 收到请求后向邻居区域广播查询消息:分组号为 1 的节点将自己的位置信息发送给 $S_{1,n+1}$.

步骤 2.1.3. $S_{1,n+1}$ 从这些节点中选择与 $S_{1,n+1}$ 距离最近的节点,不妨假设为 $S_{1,F}$.

步骤 2.1.4. $S_{1,n+1}$ 请求 $S_{1,F}$ 发送 $S_{i,F}$ 的 ID.

步骤 2.1.5. $S_{1,n+1}$ 将 $S_{i,F}$ 的 ID 发送给 $S_{i,n+1}$.

步骤 2.2. $S_{i,n+1}$ 向邻居区域发布查询消息:在已分配了分组号的邻居节点中,哪些节点的邻居节点包含了 $S_{i,F}$.对于包含了 $S_{i,F}$ 的邻居节点,则发送一个包括自身节点组数的消息给 $S_{i,n+1}$.

步骤 2.3. $S_{i,n+1}$ 选择组数最少的节点,如果有两个或者两个以上组数相等的节点,则随机选择一个,假设该节点为 $S_{p,q}$.

步骤 2.4. $S_{i,n+1}$ 向 $S_{p,q}$ 发送包含了分组号 i 的 AMGN,使得 $\|G_{Sp,q}\| = \|G_{Sp,q}\| + 1$, 且 $G_{Sp,q} = G_{Sp,q} + \{i\}$.

定理 4. 假设目标区域内部署了密集传感器节点,在基于虚拟坐标的节点调度方案中的覆盖算法执行过

程中,接收状态下的传感器节点收到分组号为 $x(x \neq 1)$ 的 AMGN 时执行连通算法,在节点调度协议执行完毕后,各个分组构成的子网络均是连通的。

证明:根据定理 3 易证。 □

4 性能评价

为了评估算法性能,我们用 JAVA 语言在 Windows XP 操作系统的平台下设计了基于虚拟坐标的节点调度方案 NSVC(node scheduling scheme based on virtual coordinate)的仿真程序,并实现了文献[10]提出的与位置无关的分布式随机节点调度协议 RSGC(randomized scheduling scheme with guaranteed connectivity).在仿真实验中,本文提出的方案将与与位置无关的分布式随机节点调度协议在如下性能上进行比较:

- (1) 覆盖率.在执行了上述两种协议中的连通算法后,各个分组包含的节点个数并不一致,难以评价两种调度协议在覆盖性能指标上的优劣.而在仅执行覆盖算法之后,各个分组包含的节点个数大致相等(由于边界效应和随机调度等原因,各个分组包含的节点个数不一定完全相同),具有可比性.本文通过仅执行上述两种协议中的覆盖算法、比较各个分组达到的覆盖率来评价这两种协议中覆盖算法的优劣.
- (2) 维持分组连通时增加的节点个数.连通算法致力于在已选择了若干个节点维持覆盖的条件下激活最少的节点,以确保各个分组构成的传感器网络是连通的.如果维持分组连通时分组内新增加的节点个数较大,则网络中将有较多的节点在一个循环调度的工作周期内持续工作或者长时间工作,导致这些节点过早死亡,因此会缩短整个网络的生存时间.
- (3) 网络生存时间.在确保目标区域充分覆盖和任意时刻网络是连通网络的条件下,最大化地延长网络生存时间是传感器网络研究的中心问题.当传感器节点不以基于分组的方式依次轮流工作时,网络生存时间是指从网络开始工作到一定数量的节点已失效的时间长度.当传感器节点采用基于分组的方式时,分组生存时间是指从分组开始工作到分组内一定数量的节点已失效的工作时间长度,网络生存时间是各个处于存活期的分组依次轮流工作时的时间长度.

在传感器网络应用中,传感器节点通常情况下被放置在没有基础结构的地方.因此,在节点随机分布的条件下,各种算法的优劣关系到该算法的实际应用范围.本模拟实验考虑节点随机分布在目标区域的情况.表 1 是该仿真实验中采用的实验参数.

Table 1 Parameters in this experimental simulations

表 1 仿真实验的实验参数

Target area	300m×300m
Number of all nodes	2 000
Communication range	20m~50m
Sense range	10m
a_{\min}	$0.5 \times R_s$
Number of subsets	1~8
Number of reference nodes	8~20
Initial energy of each node	100unit
Energy consumption of working nodes per time	1unit

4.1 覆盖率

由第 2.2 节对虚拟坐标的分析可知,当 a_{\min} 和 P_0 固定时,网络中必需的参照节点的个数与节点的通信半径成正比.为了使任意两个相距为 a_{\min} 的传感器节点具有不同虚拟坐标的概率接近于 1,当节点的通信半径从 20m 依次增加到 50m 时,网络中参照节点的个数也依次从 8 个增长到 20 个.表 2 为模拟实验中选取的节点通信半径与参照节点个数的对应表.

Table 2 Communication range of nodes vs. reference node number

表 2 节点的通信半径与对应的参照节点的个数

Communication radius (m)	Reference node number
20	8
25	10
30	12
35	14
40	16
45	18
50	20

为了验证采用基于虚拟坐标的节点调度方案中的覆盖算法能够获得分布更加均匀的分组,在 300m×300m 目标区域中,随机部署了 2 000 个节点,节点的通信半径和传感半径分别为 20m 和 10m,分组数为 5.图 6(a)和图 6(b)为在一次实验中,在执行两种调度协议中的覆盖算法之后,第 3 组节点工作时的覆盖区域:图 6(a)采用的是基于虚拟坐标的节点调度方案,图 6(b)为与位置无关的节点随机调度协议.

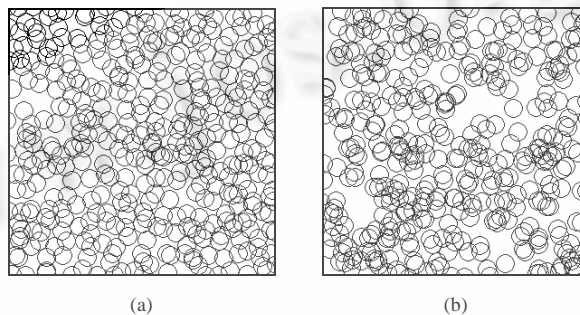


Fig.6 The covered area of nodes in the 3rd subset

图 6 第 3 组节点的覆盖区域

从图 6 中可以直观地看出,采用基于虚拟坐标的节点调度方案中的覆盖算法得到的分组较为均匀地分布在目标区域内,且盲点区域的面积均较小;而与位置无关的节点随机调度协议中的覆盖算法得到的分组分布不均匀,有些小区域内传感器节点分布过于密集,而有一些“盲”区域的面积较大.保证盲区域的面积均较小,对传感器网络的应用具有很强的实际意义.当传感器网络应用于森林火灾监测和军事监视等场合时,如果采用基于虚拟坐标的节点调度方案,即使火灾和敌情出现在目标区域的盲区域,由于盲区域的面积较小,在很短的时间内也能监测到上述事件的发生.而与位置无关的节点随机调度协议由于存在着较大面积的盲区域,因此不能保证这一点.

图 7 为分组个数 $K=5$ 时,上述两种覆盖算法在不同的节点通信半径下获得的覆盖率 CR(coverage rate): maxCR 为 5 个分组依次轮流工作时获得的最大覆盖率,meanCR 为 5 个分组的平均覆盖率,minCR 为 5 个分组依次轮流工作时获得的最小覆盖率,其中,maxCR,meanCR 和 minCR 都是 20 次实验的平均值.从图中可以看出,若采用基于虚拟坐标的节点调度方案中的覆盖算法,则分组获得的最大覆盖率大约为 0.90,而采用与位置无关的节点随机调度协议中的覆盖算法,分组获得的最大覆盖率只有 0.82 左右.前者所有分组获得的平均覆盖率大约为 0.88,而后者所有分组获得的平均覆盖率仅维持在 0.795 左右.并且,前者分组获得的最小覆盖率也要高于后者分组获得的最大覆盖率.

图 8 为节点通信半径 $S=30m$ 时,当分组个数 K 从 1 增加到 8 时,上述两种覆盖算法获得的平均覆盖率.从图中可以看出,当分组个数相等时,前者所有分组获得的平均覆盖率都要大于后者所有分组获得的平均覆盖率.而且我们注意到,前者将 2 000 个节点分为 7 组时获得的平均覆盖率与后者仅分为 5 组时获得的平均覆盖率大致相等,此时,前者平均每个组仅包含了 286 个节点,而后者平均每个组是 400 个节点.

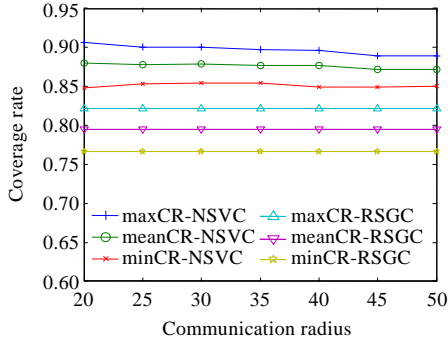


Fig.7 Communication radius vs. coverage rate

图 7 通信半径与覆盖率的关系

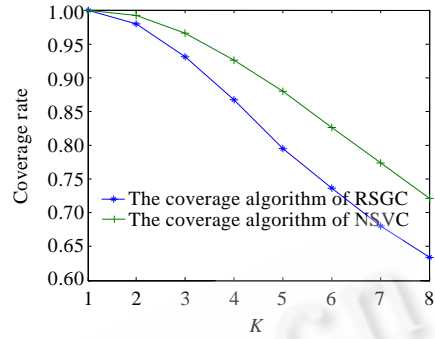
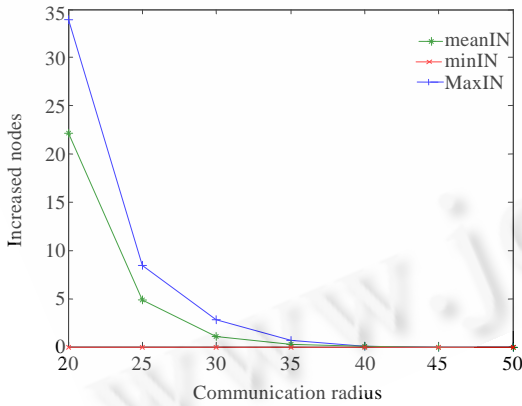


Fig.8 K vs. coverage rate

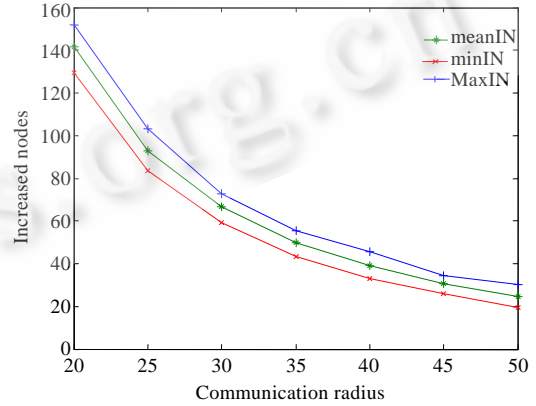
图 8 K与覆盖率之间的关系

4.2 维持分组连通时额外加入到分组内的节点个数

当分组个数 $K=5$ 时,图 9 为节点具有不同通信半径时,执行上述两种协议中的连通算法与没有执行连通算法相比分组内额外增加的节点个数 IN(increased nodes):maxIN 为 5 个分组中分组内增加得最多的节点个数, meanIN 为平均增加的节点个数,minIN 为 5 个分组中分组内增加得最少的节点个数,其中,maxIN,meanIN 和 minIN 均是 20 次实验的平均值.从图中可以看出,若采用基于虚拟坐标的节点调度方案,随着节点通信半径的逐渐增加,维持分组连通时,分组内增加的节点个数迅速递减到 0.并且在相同的节点通信半径的条件下,前者的 maxIN,meanIN 和 minIN 都远远地小于后者的 maxIN,meanIN 和 minIN.这是因为随机调度协议中连通算法的核心思想是,分组内的每个节点建立一条到 Sink 节点的最短路径来维持分组连通:对于存在着很少的上游节点的节点(不妨假设为 A,到 Sink 节点的最小跳数为 n)来说,这些上游节点可能在覆盖算法中被随机分配到其他分组内,而在调度算法中,节点 A 必须将这些上游节点中的某个节点分配到它所在的分组内,才能保持其到第 $n-1$ 跳的连通,因此,在执行连通算法后,所有分组内增加的节点个数也较多.



(a) The connection algorithm of NSVC
(a) NSVC 中的连通算法



(b) The connection algorithm of RSGC
(b) RSGC 中的连通算法

Fig.9 Communication radius vs. increased nodes when executing two connection algorithms above

图 9 当执行上述两种连通算法时,通信半径与增加节点的关系

4.3 网络生存时间

当不考虑节点执行调度协议的能耗时(因为上述两种协议在节点生命周期内都仅执行一次,故可忽略),图 10 给出了分组个数为 5 时,两种调度协议中节点的通信半径与网络生存时间之间的关系.如果不采用调度协议,传感器网络中的所有节点撒播之后都保持工作状态,显然,网络生存时间为 100.而如果不考虑网络的连通,由

于将传感器网络中的所有节点划分为 5 个不相交的分组,此时,网络的生存时间为 500.从图中可以看出,若分组内 1% 的节点失效即认为分组已经死亡,则当节点通信半径从 20m 逐渐增加到 50m 时,若采用与位置无关的节点随机调度协议,则网络的生存时间均为 100,并没有延长网络的生存时间.而如果采用基于虚拟坐标的节点调度方案,则网络生存时间从 211 逐渐增加到 500.若分组内 5% 的节点失效即认为分组已经死亡,则当节点通信半径从 20m 逐渐递增至 50m 时,前者的网络生存时间仅稍微有所增加;而对于后者来说,当节点通信半径大于 25m 时,网络生存时间均为 500.

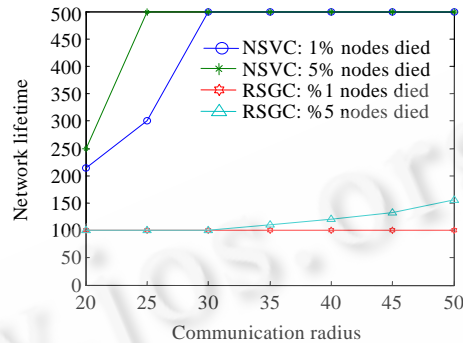


Fig.10 Communication radius vs. lifetime of network

图 10 通信半径与网络生存时间的关系

5 结 论

节点调度是延长网络生存时间的一种有效方法.目前,在调度算法中很少有文献同时考虑覆盖问题和网络连通问题,且现有的基于分组的节点调度算法不能保证各个组内的传感器节点均匀地分布在目标区域.为此,本文提出了一种分布式的、与节点位置无关的节点调度方案——基于虚拟坐标的节点调度方案.本方案中的覆盖算法利用虚拟坐标信息进行分组,在不需要节点位置信息的条件下,使得各个分组内的传感器节点较为均匀地分布在目标区域.本协议不仅解决了任意时刻处于活动状态下的节点充分覆盖目标区域的覆盖问题,而且保证了网络的连通.仿真实验表明了该方案的有效性.并且,方案存在如下具有优势的特征:(1) 该方法仅适用于个体通信能力有限的嵌入式系统,在保证网络连通度的前提下,节点的通信范围越小,需要的参照节点越少,该方案的计算开销和通信开销会进一步地减少;(2) 当目标区域内传感器节点为随机分布时,各个分组依次工作时目标区域内没有被覆盖到的盲区面积均较小.

致谢 在此,我们向曾经对本文提出宝贵建议的审稿专家以及曾参与本文内容讨论的所有老师和同学表示衷心的感谢.

References:

- [1] Ren FY, Huang HN, Lin C. Wireless sensor networks. Journal of Software, 2003,14(7):1282-1291 (in Chinese with English abstract). <http://www.jos.org.cn/1000-9825/14/1282.htm>
- [2] Zhou SW, Lin YP, Zhang JM, Ouyang JC, Lu XG. A wavelet data compression algorithm using ring topology for wireless sensor networks. Journal of Software, 2007,18(3):669-680 (in Chinese with English abstract). <http://www.jos.org.cn/1000-9825/18/669.htm>
- [3] Ren Y, Zhang SD, Zhang HK. Theories and algorithms of coverage control for wireless sensor networks. Journal of Software, 2006,17(3):422-433 (in Chinese with English abstract). <http://www.jos.org.cn/1000-9825/17/422.htm>
- [4] Choudhury RR, Kravets R. Location-Independent coverage in wireless sensor networks. 2004. <http://www.crc.hawaii.edu/~croy/publications.html>
- [5] Ye F, Zhong G, Lu SW, Zhang LX. PEAS: A robust energy conserving protocol for long-lived sensor networks. In: McKinley PK,

- Shatz S, eds. Proc. of the 23rd Int'l Conf. on Distributed Computing Systems. Providence: IEEE CS Press, 2003. 28–37.
- [6] Mao YC, Liu M, Chen LJ, Chen DX, Xie L. A distributed energy-efficient location-independent coverage protocol in wireless sensor networks. *Journal of Computer Research and Development*, 2006,43(2):187–195 (in Chinese with English abstract).
- [7] Zheng R, He GH, Liu X. Location-Free coverage maintenance in wireless sensor networks. Technical Report, UH-CS-05-15, University of Houston, 2005.
- [8] Chen BJ, Jamieson K, Balakrishnan H, Morris R. Span: An energy efficient coordination algorithm for topology maintenance in ad hoc wireless networks. *ACM Wireless Networks*, 2002,8(5):481–494.
- [9] Godfrey PB, Ratajczak D. Naps: Scalable, robust topology management in wireless ad hoc networks. In: Proc. of the 3rd Int'l Symp. on Information Processing in Sensor Networks. Berkeley: ACM Press, 2004. 443–451.
- [10] Liu C, Wu K, Xiao Y, Sun B. Random coverage with guaranteed connectivity: Joint scheduling for wireless sensor networks. *IEEE Trans. on Parallel and Distributed Systems*, 2006,17(6):562–575.
- [11] Nagpal R, Shrobe H, Bachrach J. Organizing a global coordinate system from local information on an ad hoc sensor network. In: Feng Z, Leonidas G, eds. Proc. of the 2nd Int'l Workshop on Information Processing in Sensor Networks. Palo Alto: Springer-Verlag, 2003. 333–348.
- [12] Bulusu N, Heidemann J, Estrin D. GPS-Less low cost outdoor localization for very small devices. *IEEE Personal Communications*, 2000,7(5):28–34.
- [13] Rao A, Ratnasamy S, Papadimitriou C, Shenker S, Stoica I. Geographic routing without location information. In: Proc. of the 9th Annual Int'l Conf. on Mobile Computing and Networking. San Diego: ACM Press, 2003. 96–108.
- [14] Caruso A, Chessa S, De S, Urpi A. GPS free coordinate assignment and routing in wireless sensor networks. In: Johnson DB, ed. Proc. of the 24th Conf. of the IEEE Communication Society. Seattle: IEEE CS Press, 2005. 150–160.

附中文参考文献:

- [1] 任丰原,黄海宁,林闯.无线传感器网络.软件学报,2003,14(7):1282–1291. <http://www.jos.org.cn/1000-9825/14/1282.htm>
- [2] 周四望,林亚平,张建民,欧阳竞成,卢新国.传感器网络中基于环模型的小波数据压缩算法.软件学报,2007,18(3):669–680. <http://www.jos.org.cn/1000-9825/18/669.htm>
- [3] 任彦,张思东,张宏科.无线传感器网络中覆盖控制理论与算法.软件学报,2006,17(3):422–433. <http://www.jos.org.cn/1000-9825/17/422.htm>
- [6] 毛鸞池,刘明,陈力军,陈道蓄,谢立.一种高效节能的与节点位置无关的传感器网络覆盖协议.计算机研究与发展,2006,43(2):187–195.



李小龙(1981—),男,湖南常德人,博士生,主要研究领域为传感器网络覆盖控制,定位技术.



余建平(1979—),男,博士生,主要研究领域为传感器网络数据查询处理.



林亚平(1955—),男,博士,教授,博士生导师,CCF高级会员,主要研究领域为计算机网络,机器学习.



卢新国(1979—),男,博士,主要研究领域为人工智能,计算机网络.



易叶青(1976—),男,博士生,讲师,主要研究领域为传感器网络安全.