

用改进的列生成算法优化异构 WSN 生命周期

吕 振, 陆海青

(辽宁工程技术大学 电气与控制工程学院, 辽宁 葫芦岛 125105)

摘 要: 研究了无线传感器网络(WSN)中有路由需求的静止目标实时覆盖问题,重点解决在保持对不同种类、不同需求的目标覆盖情况下,如何通过对路由的调度最优化网络的生命周期.在对问题进行数学建模后,提出改进的列生成算法,它把基于列生成的算法与可取得较好初始可行基加快收敛速度的随机选择算法结合起来,并且改进算法终止条件进一步加速收敛,进而求得最优解.实验探究了网络参数(如初始可行基、通讯半径等)对网络寿命的影响,验证了算法的有效性,该算法可对建立实际的系统提供指导.

关键词: 异构 WSN; 目标覆盖需求; 列生成; 初始可行基; 优化

中图分类号: TP393

文献标识码: A

文章编号: 1000-7180(2014)05-0107-05

Optimization of Heterogeneous WSNs Lifetime with the Modified Column Production Algorithm

LV Zhen, LU Hai-qing

(Research Center for the Electrical Engineering Department, Liaoning Technical University, Huludao 125105, China)

Abstract: The paper studies the real-time stationary target coverage problem with routing constraints for heterogeneous wireless sensor networks (WSN), and focus on how to maximize network lifetime by scheduling routing while fulfilling target coverage situation of different types and different requirements. Mathematical modeling on the problem, the improved column production algorithm is proposed, which combines the column generation based algorithm and a random selection algorithm who can get better basic feasible set to speed up the convergence speed, and improved termination conditions for further accelerating convergence, then get the optimal solution. Experimental results reveal the influence of the parameters on network lifetime and show the effectiveness of the algorithm, which can for establishment of the actual system provide a useful guide.

Key words: heterogeneous WSN; target coverage requirements; column production; basic feasible set; optimization

1 引言

WSN 在环境监测、战场监视、危险区域控制等方面有很大的实用价值.在多目标监控的应用中,由于被监控目标的特性不同,需要不同种类的无线传感器节点,比如视频、温度、湿度等.目标的重要性不同,因此各自的覆盖需求(如覆盖节点的数目、采样率)也各不相同^[1].显然,这样的 WSN 有一个重要的特性:异构性,即不同的目标需被运行在不同采样率下的不同种类不同数目的节点所覆盖,且这些节点初始能量不同^[2-5].

异构网络中,所有的数据由覆盖目标的源节点采集路由到汇聚节点,故路由问题显得比较重要.传感器节点由传感部件和通讯部件组成,分别针对的是覆盖需求和路由需求.在实际应用中,节点能量由干电池供应,通常更换失效电池很困难.因此,通过对目标覆盖与路由需求建立优化模型,提出改进的算法,试图最优化网络的生命周期.

2 问题建模

2.1 基本描述

采用的网络模型包含 n 种不同类型的传感器节

点(总数为 N 个): $S = \bigcup_{i=1}^n S_i$, 其中 S_i 代表第 i 类传感器节点, 且 $|S| = N$. 这 N 个传感节点被用来覆盖 m 个目标: $R = \{r_1, r_2, \dots, r_m\}$. s_0 为汇聚节点(Sink), 负责其他节点的调度和协调, 假定 s_0 有网络中所有节点的位置信息, 并且知道每一类节点的传感与通讯半径. 为定义问题, 先给出几个关键的概念.

(1) 网络连通图(NCG) 为有向图, $NCG = \{V, E\}$, 其中 $V = S \cup \{s_0\}$, $E = \{l_{u,v}\}$. 若节点 $s_u (s_u \in S_i)$ 能以半径 TR_i 与 s_v 通讯, 则节点 s_u 与 s_v 间存在链路 $l_{u,v}$.

(2) 目标覆盖图(TCG) 为二分图, $TCG = \{S, E, R\}$, 其中 $E = \{L_{u,k}\}$. 若节点 $s_u (s_u \in S_i)$ 能以半径 SR_i 覆盖 $r_k (r_k \in R)$, 则节点 s_u 与 r_k 间存在链路 $L_{u,k}$.

(3) 目标 Q 覆盖需求: 在任一时刻 r_k 需被至少 $q_{i,k}$ 个第 i 类节点, 以 $\theta_{i,k}$ 的采样率监控. 为每一个目标 r_k 定义一个覆盖需求向量 $Q_k = (q_{1,k}, \dots, q_{n,k})$ 和数据采样率向量 $\theta_k = (\theta_{1,k}, \dots, \theta_{n,k})$.

(4) 数据采集需求: Sink 必须收集到网络中产生的数据, 即存在路由使得 Sink 可以通过单跳或多跳的方式实时获取每一个活跃的源节点产生的数据.

(5) 网络生命周期: 从网络开始运行的时刻到 Sink 不能完全接收对某一目标 $r_k (r_k \in R)$ 的数据所经历的时间. Sink 不能完全接收对某一目标 $r_k (r_k \in R)$ 的数据的原因可能是缺少足够多的数目某一类节点来满足覆盖要求或找不到数据传输的路由路径.

问题定义: Q 覆盖与路由联合优化问题, 即给出网络的 $NCG = \{V, E\}$ 与 $TCG = \{S, E, R\}$, 最大化网络的生命周期.

2.2 工作组及问题分析

研究的目的是: 在异构 WSN 中, 如何在满足 Q 覆盖需求及数据收集的基础上最优化网络的生命周期. 如果能找到一部分节点, 这些节点在保证 Q 覆盖需求的同时, 也可以通过调节路由把监测到的数据传输到 Sink, 那么称这部分节点为一个工作组. 不属于工作组的节点都应处于休眠状态以减少能耗. 因 WSN 的冗余性, 除这个工作组外, 还可以找到其他工作组轮流工作来延长网络寿命.

将工作组 p 定义为长度为 N 的向量 $p = \{e_u^p, s_u \in S\}$, 其中 e_u^p 表示节点 s_u 单位时间的能耗. 若向量中的某个元素为零, 即相应的节点不在该工作组内.

考虑源节点及无线链路上的数据流: 定义布尔变量 $x_{u,k}^p$ 和 $y_{u,i}$ 分别表示节点 s_u 是否为 p 中正在覆盖目标 r_k 的源节点和是否属于 S_i 类节点. Q 覆盖需求可描述为

$$\sum_{s_u \in S_{i,k}} x_{u,k}^p \geq q_{i,k} \quad (\forall i = 1 \dots n, k = 1 \dots m) \quad (1)$$

为简化, 假定每个节点在一个时刻只能负责监控一个目标, 则源节点的约束条件为

$$\sum_{r_k \in R_u} x_{u,k}^p \leq 1 \quad (2)$$

$$\sum_{r_k \in S \setminus R_u} x_{u,k}^p = 0 \quad (3)$$

其中, $R_u = \{r_k \mid r_k \text{ 可以被 } s_u \text{ 覆盖}, \forall s_u \in S, r_k \in R\}$.

定义 $w_{u,v}^p$ 为 $l_{u,v}$ 在 p 中的流量, 根据流量守恒, 节点 s_u 接收的流量与自身产生的数据流量之和等于发出的数据量. 由此得出边 $l_{u,v}$ 的约束条件:

$$\sum_{l_{u,v} \in E} w_{u,v}^p = \sum_{l_{v,u} \in E} w_{v,u}^p + \sum_{k=1}^m \sum_{i=1}^n x_{u,k}^p \cdot y_{u,i} \cdot \theta_{i,k} \quad (4)$$

又 Sink 收集产生的数据且 Sink 不对数据进行中继, 所以对 s_0 有

$$\sum_{l_{u,0} \in E} w_{u,0}^p = \sum_{s_u \in S} \sum_{k=1}^m \sum_{i=1}^n x_{u,k}^p \cdot y_{u,i} \cdot \theta_{i,k} \quad (5)$$

$$\sum_{l_{0,u} \in E} w_{0,u}^p = 0 \quad (6)$$

在工作组 p 中, 节点 s_u 的能耗 e_u^p 由传感、接收、发送三部分能耗组成. 将第 i 类传感器节点传感、发送、接收的单位数据能耗分别用 e_i^s, e_i^t, e_i^r 来表示. 由此可计算 e_u^p :

$$e_u^p = e_{uS}^p + e_{uT}^p + e_{uR}^p \quad (7)$$

式中, $e_{uS}^p, e_{uT}^p, e_{uR}^p$ 分别为工作组 p 中节点 s_u 总的能耗, 计算如下:

$$e_{uS}^p = \sum_{i=1}^n \sum_{k=1}^m x_{u,k}^p \cdot y_{u,i} \cdot e_i^s \quad (8)$$

$$e_{uT}^p = \sum_{i=1}^n \sum_{l_{u,v} \in E} w_{u,v}^p \cdot y_{u,i} \cdot e_i^t \quad (9)$$

$$e_{uR}^p = \sum_{i=1}^n \sum_{l_{v,u} \in E} w_{v,u}^p \cdot y_{u,i} \cdot e_i^r \quad (10)$$

2.3 问题的数学建模

网络的生命周期可被分为若干长度不同的时间片, 每个时间片上有且只有一个工作组处于激活状态, 不属于该工作组的节点休眠. 因此, 可以把目标覆盖与数据收集联合优化问题转化成经典的时间片分配问题. 重点在于如何在能量有限的情况下, 最大

化时间片长度总和,即

$$\text{Max}(\sum t_p)$$

$$\sum_{p \in P} e_u^p \cdot t_p \leq E_u \quad \forall s_u \in S \quad (11)$$

$$t_p \geq 0 \quad (12)$$

式中, t_p 表示分配给某工作组 p 的时间片长度, P 为所有工作组的集合. 称这个优化问题为原问题. 若给出 NCG 及 TCG, 原问题是一个标准的线性规划问题, 可以用传统算法在 $O(|P|^3)$ 时间内求解^[6-7]. 然而, 联合优化本质上的难度, P 是不太可能在多项式时间内求出的, 且 P 的规模很可能是关于 n, m 指数级的. 因此不能直接去求解原问题得出某个实例的最优解, 为解决该难题, 提出基于列生成的算法来迭代求解.

3 列生产算法

列生成是一种优化框架, 作为算法, 经常被用来计算标准的整形优化问题或对一个不能直接建模的问题求解. 这里, 列相当于工作组, 提出列生成算法通过选择对优化有利的列来简化求解. 它先把原问题分割成一个主问题和一个子问题, 主问题可以使用所有列结合 P 的一个子集得到一个初始可行的解, 即初始可行基 (Basic Feasible Set, BFS). 而 BFS 基于节点能量限定, 对其中的工作组进行时间片分配, 以得到最优的网络生命周期. 随后, 根据主问题中能量限制方程中的最优对偶变量, 使用子问题来确认当前的 BFS 是否最优, 如果是, 就停止迭代循环. 否则, 将求解子问题得到的一个对当前 BFS 最有利的工作模式加入到当前的 BFS 中, 并使用这个扩展后的 BFS 再次求解主问题. 如此迭代循环, 直到当前的 BFS 含有可构建最优解的所有工作组.

3.1 随机选择算法

因基于列生成算法需 BFS 去开始迭代循环, 且 BFS 的选择虽不会影响收敛结果, 但会对收敛速度产生较大影响. 可见, BFS 具有一定的重要性, 于是提出随机选择算法去产生 BFS. 考虑到对网络连通性的假定, 对于选定的需要覆盖目标的节点, 可以使用最短路径法来确保通向汇聚节点的路由^[8-9]. 在选取 BFS 时, 无需考虑能量的因素, 只需要将目标集中在得到满足覆盖需求与路由需求的工作组上去. 随机算法的时间复杂度为 $O(n)$, n 为节点的数目. 显然, BFS 的势越大, 算法就会收敛得越快, 这点也在后面的实验中被证实. 考虑到相比较用于求解迭代优化式的随机算法的时间复杂度较低, 所以在实际的使用中, 应该用随机选择算法产生尽量多的工作

组. 算法步骤:

(1) 输入 NCG、TCG、 Q , 并初始化 $j = 0, P_0 = \Phi$.

(2) 外循环: 若 $j \leq TH$, 将 $U_{i,k}$ 中的所有节点恢复为未染色状态, 令 $k = 1$. 内循环: 若 $k < m$, 则 k 加 1 且判断 $U_{i,k}$ 中有超过 q_k 个未被染色的节点, 若有则从 U_k 中随机选择 q_k 个节点加入 p_j , 否则返回 Φ . 若 $k = m$, 跳出内循环, 对所有被染色的节点调用最短路算法找到 s_0 的路径并且将此路径上所有的点染色形成覆盖模式 p , 若 p 不在当前的 BFS 中, 将其加入, 然后 j 加 1, 否则 j 直接加 1.

(3) 若 $j \leq TH$ 继续循环, 否则输出 P_0 , 算法结束.

3.2 主问题与子问题

假定使用随机生成算法得到初始的可行基 P_0 , 可对原问题重建模为如下主问题:

$$\text{Max}(\sum t_p)$$

$$\sum_{p \in P_0} e_u^p \cdot t_p \leq E_u \quad \forall s_u \in S \quad (13)$$

主问题是一个线性规划问题, 可以用单纯形法求解. 如上所述, 在得到主问题的解以后, 应该验证这个解是否为最优解, 如果不是, 那么就需要找到一个新的工作组加入可行基. 假定 B_u 为约束条件的最优对偶变量, 那么工作组 P 的降低成本为

$$\zeta_p = 1 - \sum_{s_u \in S} B_u \cdot e_u^p \quad (14)$$

需找到最大非负的降低成本 ζ_p^* 的工作组, 将它加入到 BFS, 即 $P_0 = P_0 \cup \zeta_p^*$, 其中 ζ_p^* 可通过求解下面的子问题得到

$$(\text{Sub}) \text{Max}(\zeta_p) \quad S. t(1) \sim (7).$$

若对子问题求解后 $\zeta_p \leq 0$, 则主问题最优, 算法终止; 否则, 就将子问题找到的新的 p 加入 BFS, 再次调用主问题求解.

4 改进终止条件

基于列生成算法的一个主要问题就在于收敛速度. 通过实验, 发现终止条件“找到负的降低成本即终止”在大部分情况下是很有效的; 但在一些特定的实例中, 算法在最后收尾阶段花费的时间较长, 即每次迭代的增长非常小. 为加快算法速度, 提出新的终止条件: 对每次迭代取得的结果做记录, 如果 TH 次循环后, 目标值的提高很小 (如 $< H, H = 1\%$), 则终止算法. 改进后的列生成算法如下:

(1) 输入 NCG、TCG、覆盖需求 Q, TH, H, P_0 .

(2) 令 $h=1, \text{count}=0, T_0=0$. 使用初始可行基 P_0 求解主问题, 得到 T_1, B_1 .

(3) 循环. 若 $(T_h - T_{h-1})/T_h$ 小于 H , count 加 1, 若 $\text{count} \leq TH$, 使用 B_h 求解子问题得到新的列并计算 ζ_p . 若 $\zeta_p \leq 0$ 且 p 在当前的 BFS 中, 返回.

(4) 若 $\zeta_p \leq 0$ 且 p 不在当前的 BFS 中使用新的 BFS 求解主问题, h 加 1, 得到 T_k, B_h , 返回, 若 $\text{count} > TH$, 输出得到 T_k, B_h , 算法结束.

5 实验结果与分析

网络设置: 25 个一类与 25 个二类传感器节点及 3 个目标被随机的布撒在 $100 \text{ m} \times 100 \text{ m}$ 的区域中, 1 个汇聚节点设在区域中心; 节点参数与覆盖需求如表 1、2 所示, 网络设置如图 1 所示. 通过实验探究不同的参数(如 BFS、通讯半径等)对寿命的影响. 实验平台为 Lingo8.0、VC6.0. 为了解单个参数对网络寿命的影响, 采取控制变量法.

表 1 目标覆盖需求

	Q_k	θ_k
目标 1	(0,1)	(0,1)
目标 2	(1,1)	(1,1)
目标 3	(1,0)	(1,0)

表 2 节点参数

节点	N	TR_i	SR_i	E_i	e_i^1	e_i^2	e_i^3
一类	25	20	30	100	1	2	1
二类	25	25	20	150	1	2	1

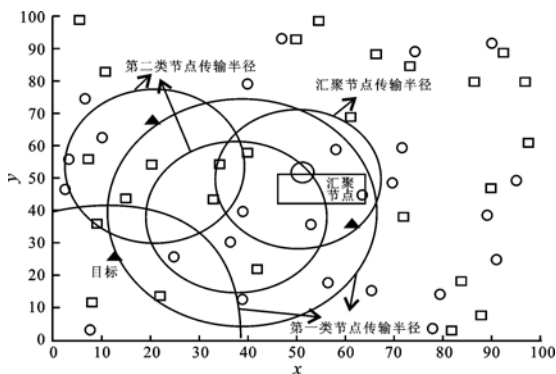


图 1 网络设置示意图

采用随机算法产生了三个初始可行基 (BFS) P_0, P_1, P_2 . 其中, P_0, P_1, P_2 分别包含 5 个、1 个、10 个不同的工作组, 而 P_3 是 P_1 与 P_2 的并集. 如图 2 所示: 不同的 BFS 对收敛结果没有影响, 只影响收敛速度; 收敛速度并不与 BFS 中的工作组数量成比例. 很显然, 由列生产算法本身的性质决定, 不论是哪个 BFS, 列生产算法都迭代收敛到最优解: 100.

P_0, P_2 相对于 P_1 来说收敛得较快, 但比较 P_0 与 P_2 又发现收敛速度并不与工作组个数成正比. 需要注意的是 P_3 的确比单独的 P_0 与 P_2 收敛得要快, 这个结果较符合直觉: 算法的收敛速度与启示工作组个数密切相关.

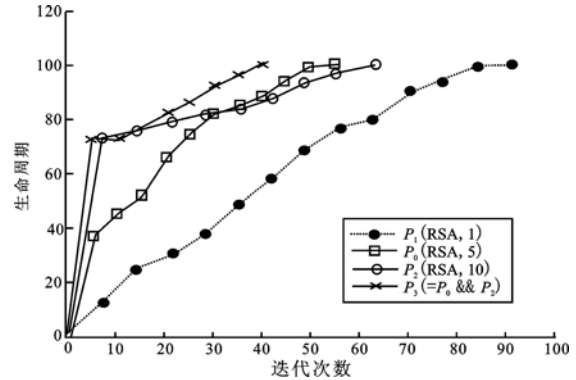


图 2 初始可行基对算法的影响

将第一类节点的通讯半径 TR 以 5 为步长从 15 逐步增加到 40. 通讯模型: $E_t = E_0 \cdot d^\alpha$, 其中 d 是传输距离, α 是信号衰减参数(取 $2 \sim 4$). 为方便考虑, 设定第一类节点通讯半径为 20 时, 传输能耗为 2, 则 $E_0 = 2/20^\alpha$. 如图 3 所示, $TR \leq 25$ 时, 生命周期随 TR 递增. 这是因为尽管源节点的个数与位置都没有变化, 随着 d 变大, 潜在节点间的链路个数增加, 网络上的数据负载被更多的中继节点负担, 带来负载平衡, 从而延长了生命周期. 当 $TR > 25$ 时, 随着 TR 增长生命周期在下降, 且 α 越大生命周期衰减越快. 这是由于, 传输一个数据单元的能耗随 TR 增加而增加, 这样在超过某个平衡点后, 随着 TR 的增加生命周期不但得不到提升反而会缩短. 而很明显, α 越大, 节点传输一个单元数据能耗就越大, 生命周期衰减得越快.

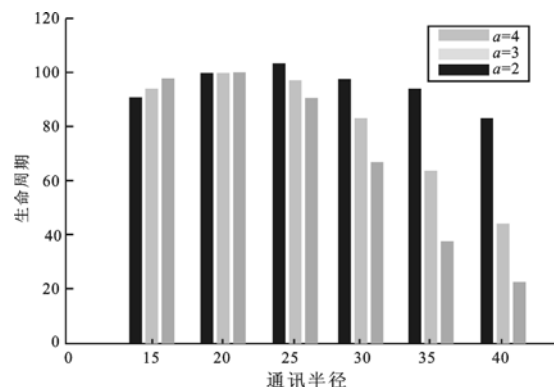


图 3 通讯半径对生命周期的影响

将第一类传感器节点的传感半径 SR 以 5 为步长从 25 逐步增加到 50. 考察线性模型 $e_s = e_0 \cdot d$ 和二次模型 $e_s = e_0 \cdot d^2$, 标准化传输能量: SR 为 30 时能耗

为1,故 $e_0 = 1/30$, $e_0 = 1/900$. 如图4所示,当 $SR \leq 30$ 时,生命周期随 SR 递增, $SR > 30$ 后,它随 SR 增加而减少. 刚开始时,大的 SR 会使更多的源节点可以感知到目标,这样潜在节点的增加自然会导致生命周期延长. 然而,当 SR 超过30后,增加的 SR 使传感能量增加且超过源节点增加对生命周期的影响,故生命周期开始缩短. 二次模型比线性模型衰减得更快. 注意到 SR 在25到50间变化时,生命周期仅在99.5到100间变化,即只变化 SR 对生命周期影响不大. 这是因为网络的瓶颈在于传输的能耗,当随机均匀地布撒节点与目标时,距离汇聚节点一跳的节点会首先死亡. 原因在于它们需要负责将整个网络产生的负载中继到汇聚节点.

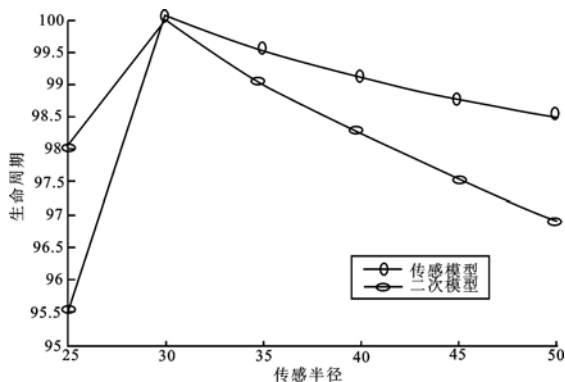


图4 传感半径对生命周期的影响

由上述分析想到:若需布撒一个传感器网络,可以在汇聚节点的周围布撒更多的节点以分担压力,从而延长网络的寿命. 为证实,做如下试验:方法一为向目标周围增加节点,方法二是向汇聚节点周围增加同样数目的节点. 增加节点数目从1到5,每次增加一个. 如图5所示,方法二可以很大程度的提升网络寿命(每个节点平均10%的提升),而对于第一种方法,每增加一个节点延长的生命周期不足1%.

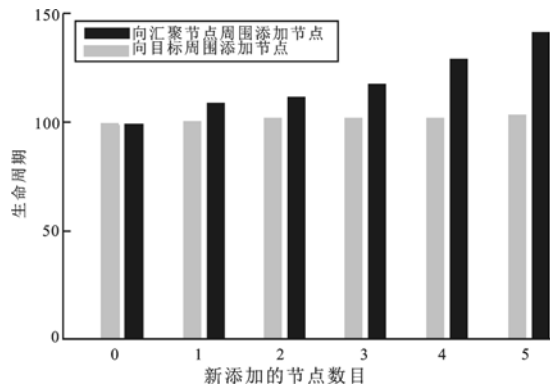


图5 节点数目对生命周期的影响

只调节部分节点也可以较为显著的延长网络寿命. 故而,如果系统中可调节的节点种类越多,使用实验中的调节策略,就越可能延长网络生命周期.

6 结束语

在实际应用中,对数据要求的异构性,WSN 通常是异构的. 文中对目标覆盖问题进行了一定的理论分析,提出改进的列生成算法. 通过实验,探究了不同的参数对网络生命周期的影响,这对实际建立系统具有较大的指导意义. 另外,多参数联合优化问题有待进一步探究.

参考文献:

- [1] 王乐,李德英,李政. 无线传感器网络中最大化生命周期的多属性目标覆盖[J]. 计算机科学, 2009, 36 [10A]:123-125.
- [2] 谷雨. 无线传感器网络目标覆盖的研究[D]. 武汉: 华中科技大学, 2010.
- [3] 唐健华. 无线传感器网络目标优化覆盖的研究[D]. 南京: 南京邮电大学, 2011.
- [4] 韩志杰,吴志斌,王汝传,等. 新的无线传感器网络覆盖控制算法[J]. 通信学报, 2011, 32(10):174-183.
- [5] 潘巨龙,闻育. 无线传感器网络的异构性研究[J]. 航空计算技术, 2007, 37(2):124-126.
- [6] Gu Y, Ji YS, Jie Li, et al. QoS-aware target coverage in wireless sensor networks[J]. Wiley's Journal of Wireless Communications and Mobile Computing, 2009, 9 (12):1645-1659.
- [7] Desaulniers G, Desrosiers J, Solomon MM. Column generation[S]. New York:Springer, 2005.
- [8] Liu H, Wan P, et al. Maximal lifetime scheduling for sensor surveillance systems with k sensors to l target [J]. IEEE Transactions on Parallel and Distributed Systems, 2006, 17(12):1526-1536.
- [9] Zhao Q, Gurusamy M. Lifetime maximization for connected target coverage in wireless sensor networks[J]. IEEE/ACM Transactions on Networking (TON), 2008(16):1378-1391.

作者简介:

吕振 男, (1963—), 硕士, 教授, 硕士生导师. 研究方向为计算机检测技术与应用、智能检测等方面的研究.

陆海青 男, (1988—), 硕士研究生. 研究方向为计算机检测技术及应用.