

无线传感器网络中基于压缩感知和 GM(1,1)的异常检测方案

李鹏^① 王建新^① 曹建农^{①②}

^①(中南大学信息科学与工程学院 长沙 410083)

^②(香港理工大学电子计算学系 香港)

摘要: 针对现有的异常事件检测算法准确率低和能量开销较大等问题, 该文提出一种基于压缩感知(CS)和GM(1,1)的异常事件检测方案。首先, 基于分簇的思想将传感器节点的数据进行压缩采样后传输至Sink, 针对传感器网络中数据稀疏度未知的特点, 提出一种基于步长自适应的块稀疏信号重构算法。然后, Sink基于GM(1,1)对节点发生的异常进行预测, 并对节点的工作状态进行自适应调整。仿真实验结果表明, 相比于其它异常检测算法, 该算法的误警率和漏检率较低, 在保证异常事件检测可靠性的同时, 有效地节省了节点能量。

关键词: 无线传感器网络; 异常事件检测; 压缩感知; Grey Model(1,1) (GM(1,1)); 信号重构; 能耗

中图分类号: TP393

文献标识码: A

文章编号: 1009-5896(2015)07-1586-05

DOI: 10.11999/JEIT141219

Abnormal Event Detection Scheme Based on Compressive Sensing and GM (1,1) in Wireless Sensor Networks

Li Peng^① Wang Jian-xin^① Cao Jian-nong^{①②}

^①(School of Information Science and Engineering, Central South University, Changsha 410083, China)

^②(Department of Computing, Hong Kong Polytechnic University, Hong Kong, China)

Abstract: In order to solve the problems of the low accuracy and the high energy cost by the existing abnormal event detection algorithm in Wireless Sensor Networks (WSN), this paper proposes an abnormal event detection algorithm based on Compressive Sensing (CS) and Grey Model(1,1) (GM(1,1)). Firstly, the network is divided into the clusters, and the data are sampled based on compressive sensing and are forwarded to the Sink. According to the characteristics of the unknown data sparsity in WSN, this paper proposes a block-sparse signal reconstruction algorithm based on the adaptive step. Then the abnormal event is predicted based on the GM(1,1) at the Sink node, and the work status of the node is adaptively adjusted. The simulation results show that, compared with the other anomaly detection algorithms, the proposed algorithm has lower probability of false detection and missed detection, and effectively saves the energy of nodes, with assurance the reliability of abnormal event detection at the same time.

Key words: Wireless Sensor Networks (WSN); Anomaly event detection; Compressive Sensing (CS); Grey Model(1,1) (GM(1,1)); Signal reconstruction; Energy consumption

1 引言

异常事件检测^[1,2]是无线传感器网络(Wireless Sensor Networks, WSN)中的一项重要应用, 尤其对于一些紧急事件, 如森林火灾、化学物质泄漏等, 往往希望尽快确定它们发生或影响的区域。然而由于受到网络监测环境变化及节点部件的不可靠性的影响, 节点很容易发生故障或产生错误数据, 影响了事件检测的准确性, 因此对异常事件的检测技术

进行研究具有重要的实际意义。

Xia等人^[3]将异常事件检测问题建模为带权的 l_1 范数最小化问题, 然后采用OMP算法进行迭代求解, 通过权值的自适应修正来发现网络中存在的所有异常事件。Wang等人^[4]提出了一种基于压缩感知(CS)和自回归模型的异常读数检测机制, 该机制能够区分出外部异常数据和内部异常数据。文献[5]针对网络中存在故障节点和噪声的情况, 将网络中存在故障节点下的节点读数恢复问题建模为鲁棒平均值计算问题, 进而提出了一种基于极大似然法的数据融合方法来恢复节点读数, 并为了提高鲁棒计算的收敛性, 提出了一种定点迭代算法。然而以上的

2014-09-17 收到, 2015-03-02 改回, 2015-05-14 网络优先出版
国家自然科学基金重点项目(61232001/F02)和国家自然科学基金面上项目(61173169/F020802)资助课题

*通信作者: 李鹏 lpchs617@csu.edu.cn

异常检测方案还存在不足：(1)在数据重构过程中大多假设数据的稀疏性已知；(2)没有考虑异常事件的局部性，随机对整个网络数据进行压缩采样从而得到观测数据，在造成检测效率下降的同时，也导致了较大的传输能耗。针对以上不足，本文提出了一种基于压缩感知和 GM(1,1) 的异常事件检测方案。该方案能对稀疏度未知的数据进行精确重构，适用于真实的WSN监测环境，在此基础上，提出了一种基于 GM(1,1) 的异常事件检测算法，优化了检测结果，并节省了节点能耗。

2 网络模型和问题描述

2.1 网络模型

本文研究的异常事件检测问题基于以下假设：

(1)所有的传感器节点采用固定的发送功率。每个传感器节点通过定位技术^[6]可知自身的地理位置。(2)设网络中节点数据传输的路径衰减因子为2，拥有 N 个节点的无线传感器网络的总能耗模型为^[7]

$$E_{to}(N) = C_1 + C_2 + C_3 \sum_{i=1}^N d_i^2 \quad (1)$$

其中， C_1, C_2 和 C_3 为常数； d_i 是节点 i 与其接收节点之间的距离。

2.2 问题描述

基于压缩感知进行数据收集的典型过程为：设有一个由 N 个节点和一个固定 Sink 组成的传感器网络负责监测大小为 $N \times N$ 的区域，所有节点的原始数据表示为 $N \times 1$ 的列向量 $\mathbf{x}(n) \in R^N$ 。由于网络数据的时空相关性， \mathbf{x} 在某一变换基 Ψ 上可被稀疏表示为 $\mathbf{x} = \sum_{i=1}^N s_i \Psi_i = \Psi \mathbf{s}$ ，然后采用一个与 Ψ 不相关的测量矩阵 $\phi_{M \times N}$ 对 \mathbf{x} 进行测量，当 Sink 接收到所有节点的测量值后，通过求解 l_1 范数最小化问题就能精确地重构出 \mathbf{x} 。在此基础上，本文研究的问题是：在感知数据稀疏性未知的情况下，如何对节点数据进行精确重构，以及如何在保证异常事件检测准确性的同时，尽可能少地消耗节点能量，从而延长网络寿命？

3 本文方案描述

3.1 总体思路

文中提出了一种基于步长自适应的块稀疏信号重构算法，该算法能在信号稀疏度未知的情况下实现信号的精确重构，然后依据重构结果，节点采用 GM(1,1) 进行异常预测，并通过预测结果来指导下一次节点数据传输，从而在保证异常检测准确性的同时，节省了节点能量，延长了网络生命周期。本文方案的流程如图1所示。

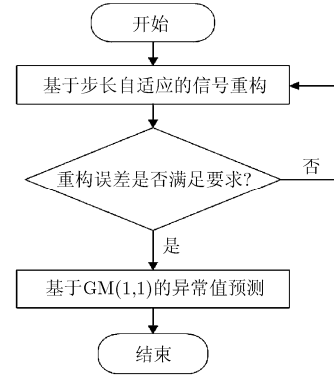


图 1 算法流程图

3.2 数据重构

一般而言，在一定范围内部署的多个传感器节点所测得的信号在用稀疏基进行表示时，其显著系数出现的位置大致相同，因此我们将整个网络中的信号重构问题看作是块稀疏重构^[8]问题。

3.2.1 块稀疏模型 设 \mathbf{x} 是块稀疏信号，它可以表示为

$$\mathbf{x} = \underbrace{[x_1, x_2, \dots, x_k]}_{\mathbf{x}^T[1]} \underbrace{[x_{k+1}, x_{k+2}, \dots, x_{2k}]}_{\mathbf{x}^T[2]} \dots \underbrace{[x_{N-k+1}, x_{N-k+2}, \dots, x_N]}_{\mathbf{x}^T[Z]} \quad (2)$$

其中， $N = Zk, \mathbf{x}[\gamma], (\gamma = 1, \dots, Z)$ 称为一个子块，块稀疏是指大多数子块为零。考虑到异常事件的时空相关性和局部性，我们将同一个簇中的数据看作一个数据块，即信号数据的分块数目对应于网络中的分簇数目，分簇算法则参照文献^[9]的工作。与此同时，测量矩阵 ϕ 也将按此分块：

$$\phi = \underbrace{[\phi_1, \phi_2, \dots, \phi_k]}_{\phi[1]} \underbrace{[\phi_{k+1}, \phi_{k+2}, \dots, \phi_{2k}]}_{\phi[2]} \dots \underbrace{[\phi_{N-k+1}, \phi_{N-k+2}, \dots, \phi_N]}_{\phi[Z]} \quad (3)$$

3.2.2 基于分块的步长自适应信号重构 针对块稀疏信号重构问题，现有的求解算法^[10]复杂度较高且要求信号的块稀疏度作为先验知识，然而真实传感器网络中数据的块稀疏度是未知的。为此，文中提出了一种基于步长自适应的块稀疏信号重构算法，主要思想是算法本身可以在迭代过程中自动调整所选原子数目来重建稀疏度未知的信号。它设置一个可变量步长来代替所选原子数目，并随步长进行增加，对每一个块稀疏度的迭代，算法都会找到信号支撑块的一个子集，并利用回溯思想修正更新上一次找到的支撑块，最后找到信号的整个支撑块，从而达到重构源信号的目的。基于步长自适应信号重构步骤如表1所示。

表 1 基于步长自适应的信号重构

输入:	随机高斯矩阵 $\phi \in R^{M \times N}$, 测量值 y , 分块向量 B , 迭代误差 λ_1 , 分层误差 λ_2 , 步长 sz
输出:	信号的估计值 \hat{x}
步骤 1	初始化残差值 $\Omega_0 = y$, 迭代次数 $k = 1$, 层次 $le = 1$, 索引块集合 $\Pi = \emptyset$, 重构向量 $\hat{x} = 0$;
步骤 2	如 $\ \Omega\ \leq \lambda_1$, 则停止迭代, 利用得到的原子进行信号重构, 否则, 转步骤 3;
步骤 3	在第 k 次迭代时 ($1 \leq k \leq M$), 选择与残差 Ω_{k-1} 最匹配的子空间 Su_k : $Su_k = \arg \max_{sz} (\text{mean}(\phi^T [Su] \Omega_{k-1}))$ 即测量矩阵转置后的每一块与残差相乘后, 取绝对值后再平均, 最后选择最大的 sz 个值的标号赋值给 Su_k (其标号对应 B 中的分块号);
步骤 4	更新索引块集合 ϕ_Π , 其中 $\Pi = \Pi_{k-1} \cup Su_k$;
步骤 5	根据下式重构 \hat{x} , 并更新残差: $\hat{x}_k = \arg \min_{i \in R^N} \ y - \phi_\Pi \hat{x}\ _2$ $\Omega_k = y - \phi_\Pi \hat{x}_k$
步骤 6	如果 $\ \Omega_k - \Omega_0\ \leq \lambda_2$, 则令 $le = le + 1$, $sz = sz \cdot le$. 转步骤 7;
步骤 7	判断步骤 7 中的迭代次数的取值是否大于分块数目, 如果是, 则算法结束, 否则, 令 $\Omega = \Omega_k$, $k = k + 1$, 转步骤 3.

4 异常检测

4.1 构造 GM(1,1)

设表 1 重构得到的任一传感器节点的感知数据为

$$\mathbf{X}^{(0)} = (\mathbf{x}^{(0)}(1), \mathbf{x}^{(0)}(2), \dots, \mathbf{x}^{(0)}(n)) \quad (4)$$

根据灰色理论可得 $X^{(0)}$ 的 1-AGO 序列为 $\mathbf{X}^{(1)} = (\mathbf{x}^{(1)}(1), \mathbf{x}^{(1)}(2), \dots, \mathbf{x}^{(1)}(n))$, 由 $\mathbf{X}^{(1)}$ 构造背景值序列 $\mathbf{Z}^{(1)} = (z^{(1)}(2), z^{(1)}(3), \dots, z^{(1)}(n))$, 其中 $z^{(1)}(k) = 0.5(\mathbf{x}^{(1)}(k) + \mathbf{x}^{(1)}(k-1)), k = 2, 3, \dots, n$, 则 GM(1,1) 可表示为

$$\frac{d\mathbf{x}^{(1)}(t)}{dt} + a\mathbf{x}^{(1)} = u \Leftrightarrow \mathbf{x}^{(0)}(k) + a\mathbf{z}^{(1)}(k) = u \quad (5)$$

使用加权最小二乘法来估计 GM(1,1) 的参数, 即通过赋予不同误差平方和不同权重, 从而加大参数估计的稳健性, 提高异常检测的精度。

令 $s = \sum_{k=2}^n \theta_k [\mathbf{x}^{(0)}(k) + a\mathbf{z}^{(1)}(k) - u]^2$, 其中 $\theta_k, k = 2, 3, \dots, n$ 为误差平方的权重。 θ_k 的设置如下: 设定初始权值, 求出加入这组权的最小二乘估计的均方误差关于权值的表达式, 优化求均方误差的极小值从而得到最优的权值。显然, 在权重 θ_k 确定的情

况下, 使 s 最小的 a, u 应满足

$$\left. \begin{aligned} \frac{\partial s}{\partial a} &= 2 \sum_{k=2}^n \theta_k [\mathbf{x}^{(0)}(k) + a\mathbf{z}^{(1)}(k) - u] \theta_k \mathbf{z}^{(1)}(k) = 0 \\ \frac{\partial s}{\partial u} &= -2 \sum_{k=2}^n \theta_k [\mathbf{x}^{(0)}(k) + a\mathbf{z}^{(1)}(k) - u] \theta_k = 0 \end{aligned} \right\} \quad (6)$$

求解式(6)可得:

$$a = \frac{DEF - G}{H - DF^2}; \quad u = \frac{DF(E - G)}{H - DF^2} \quad (7)$$

其中

$$\left. \begin{aligned} D &= \frac{1}{\sum_{k=2}^n \theta_k^2}, & E &= \sum_{k=2}^n \theta_k^2 \mathbf{x}^{(0)}(k) \\ F &= \sum_{k=2}^n \theta_k^2 \mathbf{z}^{(1)}(k), & H &= \sum_{k=2}^n [\theta_k \mathbf{z}^{(1)}(k)]^2 \\ G &= \sum_{k=2}^n \theta_k^2 \mathbf{x}^{(0)}(k) \mathbf{z}^{(1)}(k) \end{aligned} \right\} \quad (8)$$

综上所述, 可得 $X^{(0)}$ 的预测公式为

$$\hat{\mathbf{x}}^{(0)}(k+1) = (1 - e^a)(\mathbf{x}^{(0)}(1) - u/a) e^{-a(k-1)}, \quad k = 1, 2, \dots \quad (9)$$

4.2 异常检测算法

异常检测的主要思路是: 首先基于表 1 重构得到各个节点的原始数据, 然后采用 GM(1,1) 预测各个节点在下一时间段内可能发生的异常, 进而由 Sink 向节点发送消息来调整节点的工作状态, 从而在保证异常检测准确性的同时, 节省了节点能量。异常事件检测算法如表 2 所示。

表 2 异常事件检测

步骤 1	初始时, 设表 1 重构得到的原始信号序列为 $\mathbf{x}_i, i = 1, 2, \dots, N$;
步骤 2	构造 GM(1,1) 预测模型对 \mathbf{x}_i 进行预测, 得到预测值为 \mathbf{x}_i^p ;
步骤 3	如果 $ \mathbf{x}_i - \mathbf{x}_i^p \geq \delta$, 则 Sink 发送消息 N_s 到普通传感器节点 i , 并调整节点 i 的工作状态以及更新节点所在簇的测量矩阵: $\Phi_{ij} = \begin{cases} G_{M_{ij} \times N_i}, & A \geq \sum_{i=1}^C A_{ij} / C \\ SG_{M_{ij} \times N_i}, & \text{其他} \end{cases}$
	其中 $M_{ij} = M \times \left(A_{ij} / \sum_{i=1}^C A_{ij} \right)$ $N_{ij} = M_{ij} \times N_i \times \left(A_{ij} / \left(\sum_{i=1}^C A_{ij} / C \right) \right)$
步骤 4	重复执行步骤 1~步骤 3, 直到检测得到的异常事件不再变化为止。

在表2中，步骤3中的 C 表示整个网络被划分的簇个数， M_{ij} 表示在第 j 轮迭代中簇 i 的测量值， A_{ij} 表示在第 j 轮迭代中簇 i 中发生的异常事件数目。另外，本文将步骤3中节点的工作状态分为3类^[1]，如表3所示。开始时，传感器节点以状态 s_2 工作一段时间 t_1 后，再以状态 s_1 工作一段时间 t_2 后，然后进入休眠状态 s_3 ，最后Sink根据预测值，发出 N_s 消息通知各个节点对自身的工作状态进行调整：即如果某一节点 j 在随后的几个周期内可能发生异常，则节点 j 的状态调整为： $s_3 \rightarrow s_2 \rightarrow s_1$ ，否则保持 s_3 状态。

表 3 传感器节点工作状态

节点状态	感知设备	通信设备	微处理器	能耗 (mW)	切换时间(ms)
s_1	开	收发	工作	1040	-
s_2	开	收	休眠	200	12
s_3	关	关	休眠	10	50

5 仿真实验

5.1 实验设置

采用Matlab2012进行了仿真实验，考察了本文方案进行异常事件检测的准确性和能量有效性，并与AR&CS^[4]和ISDAD^[3]两种方案进行了对比分析。文中的测试数据来源于Intel Berkeley Research Lab所测的温度值。我们选取了分布在相邻空间内的400个传感器节点在一周内(2014-06-07~2014-06-13)测得的温度值进行实验。仿真参数设置如下：算法1中的迭代误差 λ_1 和分层误差 λ_2 分别取值为0.2和0.4，步长 $sz = M/(2 \log_2(N))$ 。判断阈值 δ 取值为25。选用6层小波对温度信号进行稀疏表示，测量矩阵 ϕ 采用随机高斯矩阵，数据重构采用算法1。使用和文献[7]相同的能量模型，实验结果取50次仿真的平均值。采用漏检率 P_m ，误警率 P_f 和相对误差 RE 进行性能评估：

$$\left. \begin{aligned} P_m &= \frac{N - \|\mathbf{x} - \mathbf{X} - \mathbf{1}\|_{l_0}}{\|\mathbf{x}\|_{l_0}}, P_f = \frac{N - \|\mathbf{x} - \mathbf{X} + \mathbf{1}\|_{l_0}}{N - \|\mathbf{x}\|_{l_0}} \\ RE &= \|\mathbf{x} - \hat{\mathbf{x}}\|_2 / \|\mathbf{x}\|_2 \end{aligned} \right\} (10)$$

其中， \mathbf{x} 表示原始数据， $\hat{\mathbf{x}}$ 是其重构值。 \mathbf{X} 表示检测结果， N 表示节点的数目。

5.2 实验结果

图2首先给出了不同方案的重构性能比较。可以看到，随着测量次数的增加，不同方案的RE都在下降。其中，本文方案的重构误差最低，AR & CS方案次之。这是由于ISDAD在数据重构中认为数据稀疏性不变，而AR & CS考虑了数据稀疏性随时间变化而变化这一特性，在数据重构中利用空间相关性对数据稀疏度进行初始估计，进而利用自回归模型对数据重构结果进行自适应修正，因此取得了比ISDAD更好的性能。而本文方案更进一步，在数据重构过程中不需要将数据稀疏度作为先验知识来进行处理，避免了由于数据稀疏度估计过小或过大所导致的重构精度或效率的下降，因此取得了比其他两种方案更好的性能。

图3给出了不同方案对于异常事件检测的准确性比较结果。可以看到，随着测量次数的增加，不同方案的漏检率和误警率都在明显降低，其中，本文方案的性能最优，AR&CS方案的性能次之，ISDAD方案的性能最差。这是由于ISDAD方案是一种基于迭代的检测方案，检测性能的好坏与传感器节点的初始分布有较大关联，如果容易发生异常区域布置的传感器节点数量较多，则该方案的性能较好，反之即使经过多次迭代，其漏检率和误警率依然较高。AR&CS方案通过在数据重构过程中引入自回归模型来刻画数据的稀疏性变化情况，并定制了规则来区分内部异常(节点故障)和外部异常(异常事件)，提高了异常检测的准确性。而本文方案充分考虑了节点感知数据的时空相关性，将重构得到的数据看作某一时间序列，然后构造GM(1,1)对数据进行

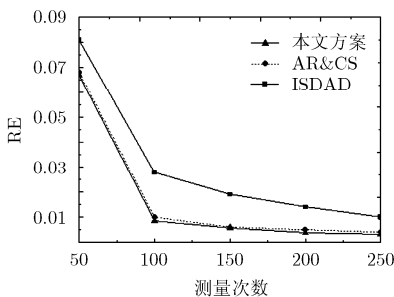


图 2 不同方案的重构误差比较

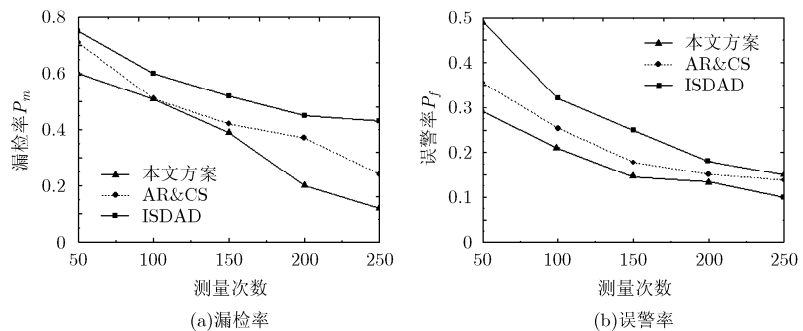


图 3 不同方案的检测结果比较

预测,克服了AR&CS方案存在的不足,因此取得了比其更好的检测结果。

最后,图4给出了不同方案进行异常事件检测的能量有效性比较结果。从图4可以看到,随着SNR的增加,不同方案的能耗都在增加,在相同的SNR条件下,ISDAD方案的能量开销最大,而本文方案的能量开销最小。当SNR从10 dB增加到50 dB时,相比于AR&CS和ISDAD两种方案,本文方案的能耗分别节省了约23.34%和35.25%。仔细分析其原因可知,这主要是因为,本文方案倾向于更多地向“热点”区域收集数据来进行异常事件检测,避免了收集那些不能对异常事件检测做出贡献的节点数据,另外,本文根据异常事件的预测结果,对下一个工作周期内的节点的工作状态进行自适应调整,从而节省了节点能量。

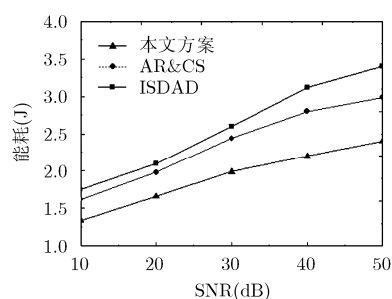


图4 不同方案的能量有效性比较

6 结束语

针对现有异常事件检测方案的不足,本文提出一种基于压缩感知和GM(1,1)的异常事件检测方案。该方案无须知道数据稀疏度和异常事件分布等先验知识,节点只需采样少量的感知数据就能实现异常事件的准确检测。下一步将考虑传感器网络中由于无线信道不可靠存在的丢包现象,研究基于LDPC码的异常事件检测可靠性问题。

参考文献

- [1] 张波, 刘郁林, 王开, 等. 基于概率稀疏随机矩阵的压缩数据收集方法[J]. 电子与信息学报, 2014, 36(6): 1478-1484.
Zhang Bo, Liu Yu-lin, Wang Kai, *et al.*. Compressive data gathering method based on probabilistic sparse random matrices[J]. *Journal of Electronics & Information Technology*, 2014, 36(6): 1478-1484.
- [2] Xie Miao, Hu Jian-kun, Han Song, *et al.*. Scalable hypergrid k-NN-based online anomaly detection in-network aggregation for wireless sensor networks[J]. *IEEE Transactions on Parallel and Distributed Systems*, 2013, 24(8): 1661-1670.
- [3] Xia Yu, Zhao Zhi-feng, and Zhang Hong-gang. Distributed anomaly event detection in wireless networks using compressed sensing[C]. 2011 11th International Symposium on Communications and Information Technologies, Hangzhou, China, 2011: 250-255.
- [4] Wang Jin, Tang Shao-jie, Yin Bao-cai, *et al.*. Distributed compressive sampling for lifetime optimization in dense wireless sensor networks through intelligent compressive sensing[C]. IEEE International Conference on Computer Communications, Orlando, FL, USA, 2012: 603-611.
- [5] Sun Bo, Shan Xue-mei, Wu Kui, *et al.*. Anomaly detection based secure in-network aggregation for wireless sensor networks[J]. *IEEE Systems Journal*, 2013, 7(1): 13-25.
- [6] Vempaty A, Han Y, and Varshney P. Target localization in wireless sensor networks using error correcting codes[J]. *IEEE Transactions on Information Theory*, 2014, 60(1): 697-712.
- [7] Cheng C T, Tse C K, and Lau F C M. A delay-aware data collection network structure for wireless sensor networks[J]. *IEEE Sensors Journal*, 2011, 11(3): 699-710.
- [8] 练秋生, 刘芳, 陈书贞. 基于块 A*正交匹配追踪的多传感器数据联合重构算法[J]. 电子与信息学报, 2013, 35(3): 721-727.
Lian Qiu-sheng, Liu Fang, and Chen Shu-zhen. A joint reconstruction algorithm for multiple sensor data based on block A* orthogonal matching pursuit[J]. *Journal of Electronics & Information Technology*, 2013, 35(3): 721-727.
- [9] 奎晓燕, 张士庚, 王建新. DSCAU: 非均衡负载无线传感器网络的基于支配集的分簇数据收集算法[J]. 高技术通讯, 2012, 22(9): 918-924.
Kui Xiao-yan, Zhang Shi-geng, and Wang Jian-xin. DSCAU: a dominating set based clustering algorithm for data gathering in wireless sensor networks with unbalanced traffic load[J]. *High Technology Letters*, 2012, 22(9): 918-924.
- [10] Eldar Y C, Kuppinger P, and Bolcskei H. Block-sparse signals: uncertainty relations and efficient recovery[J]. *IEEE Transactions on Signal Processing*, 2010, 58(6): 3042-3054.
- [11] Luo R C and Chen O. Mobile sensor node deployment and asynchronous power management for wireless sensor networks[J]. *IEEE Transactions on Industrial Electronics*, 2012, 59(5): 2377-2385.

李鹏: 男, 1983年生, 博士生, 研究方向为无线传感网、压缩感知等。
王建新: 男, 1969年生, 教授, 博士生导师, 研究方向为网络优化、参数计算等。
曹建农: 男, 1963年生, 教授, 博士生导师, “芙蓉学者”, 研究方向为网络优化、生物信息学等。