



研究与开发

基于卡尔曼预测与压缩感知的 WSN 中 高效数据收集方法

周颖, 杨丽花, 杨龙祥, 倪梦
(南京邮电大学, 江苏 南京 210003)

摘要: 提出了一种联合双重卡尔曼滤波与混合压缩感知的数据收集方案。在簇内传输时, 簇内成员节点利用卡尔曼滤波预测选择性地向簇头发送数据, 以减少数据传输量; 同时, 簇间传输时, 在多跳传输的过程中利用混合压缩感知技术进行数据压缩, 提高系统的能效。仿真结果表明, 与现有方案相比, 新算法能够在保证汇聚节点重构精度的前提下, 显著地降低数据收集的能耗。

关键词: 无线传感器网络; 卡尔曼预测; 压缩感知; 时空相关性; 能效

中图分类号: TN929

文献标识码: A

doi: 10.11959/j.issn.1000-0801.2019017

Energy-efficient data gathering scheme based on Kalman prediction and compressed sensing

ZHOU Ying, YANG Lihua, YANG Longxiang, NI Meng
Nanjing University of Posts and Telecommunications, Nanjing 210003, China

Abstract: A novel Kalman prediction and compressed sensing based energy-efficient data gathering scheme was proposed. Specially, in the intra-cluster transmission, the cluster members utilized the Kalman prediction to selectively send the data to their cluster heads. In the inter-cluster transmission, the cluster heads leveraged the hybrid compressed sensing to transfer the data to the sink via multi-hop links. Moreover, the communication cost was derived to verify the efficiency of the proposed method. Simulation results show that the proposed method has higher energy efficiency compared with the available schemes, and the sink can obtain measurements with reasonable quality by using the proposed method.

Key words: wireless sensor network, Kalman prediction, compressed sensing, spatial-temporal correlation, energy efficiency

收稿日期: 2018-09-10; 修回日期: 2019-01-09

通信作者: 杨丽花, yanglh@njupt.edu.cn

基金项目: 江苏省高等学校自然科学基金项目 (No.18KJ13510034); 第 11 批中国博士后科学基金资助项目 (No.2018T110530); 国家自然科学基金资助项目 (No.61401232, No.61427801, No.61671251, No.61372124, No.61501254); 江苏省自然科学基金资助项目 (No.BK20140894)

Foundation Items: The Natural Science Foundation of Jiangsu Higher Education Institutions of China (No.18KJ13510034), The 11th Batch of China Postdoctoral Science Fund Special Funding (No.2018T110530), The National Natural Science Foundation of China (No.61401232, No.61427801, No.61671251, No.61372124, No.61501254), The Natural Science Foundation of Jiangsu Province of China (No.BK20140894)

1 引言

近年来,无线传感器网络(wireless sensor network, WSN)^[1]得到了迅猛的发展,同时无线传感器网络作为物联网的重要支撑已引起了众多学者的关注。无线传感器网络在目标监测^[2]、医疗卫生^[3]等领域都有着广泛的应用前景和巨大的商业价值,被认为是未来改变世界的十大技术之一。然而,在无线传感器网络中,节点一般都由电池供电,节点的能量一旦耗尽便意味着节点死亡,也意味着无线传感器网络功能的下降甚至丧失,而节点经常部署在遥远的、无人值守的或很恶劣的环境中,电池很难更换。因此,节点能量受限成为制约无线传感器网络发展的瓶颈^[4]。根据对无线传感器网络能耗的分析,数据传输的能量消耗是节点的主要能耗,故采用高能效的数据收集方案成为延长网络生命期和提高网络性能的关键。

无线传感器网络中,传感器节点分布密集,节点间感知范围相互重叠,同时,传感器节点感知的物理数据大多数随着时间变化缓慢,这使得网络中待收集的原始数据存在时间与空间相关性^[5]。近年来,对利用这两种相关性减少数据量从而减少数据传输的能耗的数据收集方案进行了深入的研究。参考文献[6-9]分别利用数据间的时间或空间相关性来提高数据收集的能效,通过采用预测、压缩感知、主成分分析等数据压缩算法来减少数据的传输量。其中,参考文献[6]联合网络编码与压缩感知来消除数据间的空间冗余,减少数据传输的传输量。参考文献[7]利用主成分分析法基于数据的时间相关性降低所需传输的数据维度,减少传输的数据量。参考文献[8]利用双重预测模型消除数据间的时间冗余,避免不必要的传输。参考文献[9]基于分簇拓扑,利用混合压缩感知,进一步减少簇间传输所需的传输次数,减小能耗。然而,上述方案中,冗余信息量依旧巨大,数据收集的能效依旧不理想。

为进一步提高数据收集的能效,同时利用时间及空间相关性的传输方案得到了研究者的广泛关注^[10-12]。参考文献[10]提出了一种基于时空相关性的联合预测与分簇的传输方案,在该方案中,预测模型及基于相似性的子簇划分方法被分别用来消除数据间的时间冗余及空间冗余。然而,同样不可忽视方案中子簇的频繁更新消耗的能量。参考文献[11]提出一种联合数据预测及压缩的传输框架,该框架利用双重预测模型减少簇内传输的数据量,同时利用主成分分析(PCA)法来减少簇间传输的数据量。参考文献[12]提出一种基于邻居的克罗内克(Kronecker)压缩感知传输方案,该方案中先将时间序列压缩传输,再在多跳传输的过程中进行空间压缩。然而,上述方案在利用数据的时间相关性时需要收集足够的、连续的时间序列后才能执行,因此,不适用于有实时传输需求的无线传感网。

为解决以上问题,本文基于传感网数据间的时空相关性,提出了一种联合双重卡尔曼(Kalman)滤波及混合压缩感知的高能效数据收集方法。该方法基于分簇的无线传感器网络模型,在簇内传输时,各节点利用卡尔曼滤波选择性地将数据传输至簇头,以减少传输数据量。在簇间传输时,各簇头利用混合压缩感知多跳传输,最大化地将数据进行空间压缩,进一步减少传输能耗。本方法在保证重构精度的前提下具有极高的能效。

2 系统模型

如图1所示,本文考虑一个基于分簇的无线传感器网络^[12]。网络中包含 N 个传感器节点,共分为 P 个簇。在 t 时刻网络观测到的数据向量为 $\mathbf{x}(t)=[x_1(t), x_2(t), \dots, x_N(t)]^T$,其中 $x_i(t)$ 为 t 时刻节点 i 观测到的数据值。设 t 时刻节点 i 处观测的真实值为 $s_i(t)$,真实数据为时间离散的递推模型,各节点观测数据时受到了观测噪声的影响,



分别设定状态转移方程和观测方程如下:

$$s_i(t) = F_i(t, t-1)s_i(t-1) + w_i(t) \quad (1)$$

$$x_i(t) = C_i(t)s_i(t) + v_i(t) \quad (2)$$

其中, $F_i(t, t-1)$ 为一步转移矩阵, $C_i(t)$ 为系统观测矩阵, $w_i(t)$ 和 $v_i(t)$ 为均值为零的高斯白噪声, 噪声功率分别为 Q 、 R 。

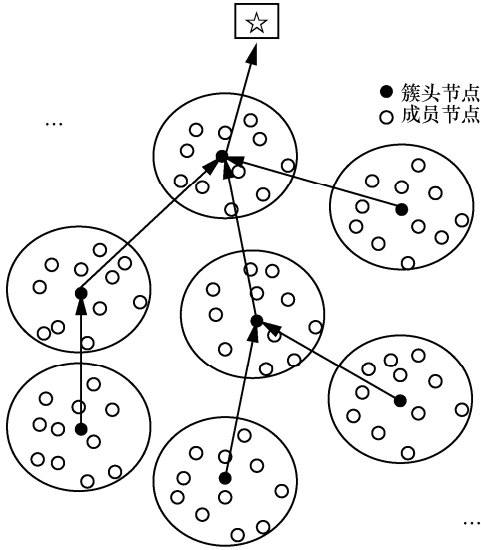


图1 基于分簇的无线传感器网络

向量 $x(t)$ 也可以表示为 $x(t) = [x_1^T(t), x_2^T(t), \dots, x_P^T(t)]^T$, 其中 $x_k(t)$ 表示第 k 个簇的原始观测数据向量。数据收集阶段, 各簇成员节点根据卡尔曼预测模型选择性地将预测不合理的数据发送给其对应簇头, 簇内传输完成后, 任一簇头节点 j 直接或间接获得其所有簇内成员数据, 记为 $y_j(t)$, 则所有簇头获得的数据向量可以表示为:

$$y(t) = [y_1(t), y_2(t), \dots, y_N(t)]^T \quad (3)$$

$$= [y_1^T(t), y_2^T(t), \dots, y_P^T(t)]^T$$

所有簇头通过混合压缩感知完成簇间传输, 汇聚节点得到 M 维编码向量为:

$$z(t) = \Phi y(t) \quad (4)$$

$$= \sum_{k=1}^P \Phi_k y_k(t)$$

其中, $\Phi \in R^{M \times N}$ 为压缩感知观测矩阵, Φ_k 为第 k 个簇中的成员对应的压缩编码系数组成的子

矩阵, $\Phi = [\Phi_1^T, \Phi_2^T, \dots, \Phi_P^T]^T$, P 为网络中簇的数量。汇聚节点利用向量 $z(t)$ 及压缩感知解码算法, 恢复出原始数据向量。

3 基于 Kalman 预测与压缩感知的 WSN 高效数据收集方法

如图2所示, 新联合传输方案主要分为3个阶段, 分别是簇内传输阶段、簇间传输阶段和重构阶段。簇内传输阶段, 各成员节点利用 Kalman 滤波预测当前时刻数据, 并将该预测值与实际观测值比较, 当两者之差小于预设的误差门限时, 认为预测合理, 节点不发送数据, 对应的簇头节点通过相同的历史数据及 Kalman 预测模型预测得到相同的预测值; 反之认为预测不合理, 成员节点对初步预测值进行信息修正后发送数据至簇头。簇间传输阶段, 簇头节点利用混合压缩感知将其簇内所有节点数据多跳传输至汇聚节点。重构阶段, 汇聚节点利用压缩感知解码算法重构获得原始数据。

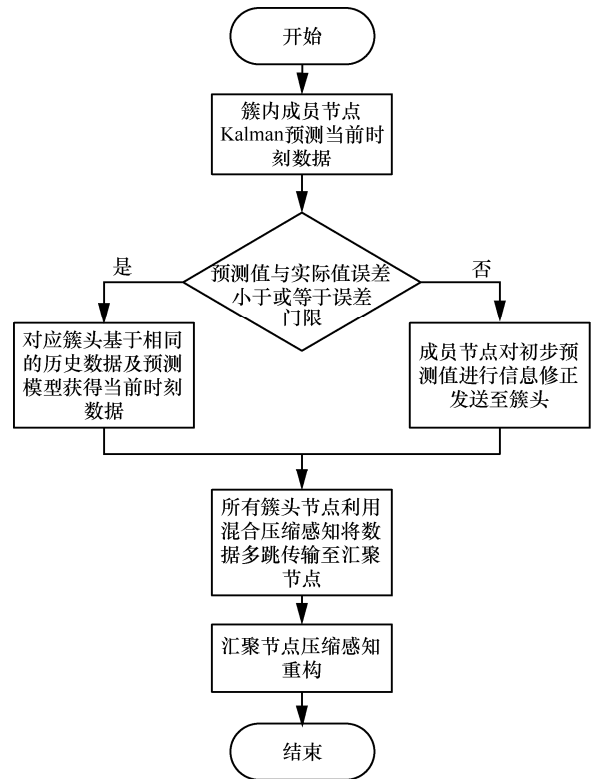


图2 基于 Kalman 预测与混合压缩感知的数据传输流程

3.1 基本原理

在簇内传输阶段,簇内成员节点*i*利用卡尔曼预测根据上一时刻数据预测当前时刻*t*的数据 $s_i'(t|t-1)$,即:

$$s_i'(t|t-1) = F_i(t, t-1)s_i^*(t-1) \quad (5)$$

其中, $s_i^*(t-1)$ 为 $t-1$ 时刻节点保存的数据, $F_i(t, t-1)$ 为一步转移矩阵。预先设定一个发送门限值 ε , 若 t 时刻初步预测值 $s_i'(t|t-1)$ 和实际观测值 $x_i(t)$ 之间的差 $\Delta = |s_i'(t|t-1) - x_i(t)|$ 小于门限 ε , 认为该预测合理, 节点*i*不发送数据, 保存 $s_i'(t|t-1)$, 即:

$$s_i^*(t) = s_i'(t|t-1) \quad (6)$$

节点*i*对应的簇头利用相同的卡尔曼滤波模型, 预测得到相同的 $s_i^*(t)$ 值。反之, 认为预测不合理, 节点*i*利用实际观测值对初步预测值进行修正更新, 得到修正预测值 $s_i'(t|t)$, 保存并发送 $s_i'(t|t)$, 即:

$$s_i'(t|t) = s_i'(t|t-1) + K_i(t)(x_i(t) - C_i(t)s_i'(t|t-1)) \quad (7)$$

$$s_i^*(t) = s_i'(t|t) \quad (8)$$

其中, $K_i(t)$ 为卡尔曼增益, 可由以下迭代获得:

$$P_i(t|t-1) = F_i(t|t-1)P_i(t-1|t-1)F_i^H(t|t-1) + Q \quad (9)$$

$$K_i(t) = \frac{P_i(t|t-1)C_i^H(t)}{C_i(t)P_i(t|t-1)C_i^H(t) + R} \quad (10)$$

$$P_i(t|t) = P_i(t|t-1) - K_i(t)C_i(t)P_i(t|t-1) \quad (11)$$

簇内传输完成时, 簇头节点获得数据向量 $y(t) = [y_1(t), y_2(t), \dots, y_{N_j}(t)]^T$, 其中:

$$y_i(t) = s_i^*(t) = \begin{cases} s_i'(t|t-1), & \Delta < \varepsilon \\ s_i'(t|t), & \Delta \geq \varepsilon \end{cases} \quad (12)$$

簇间传输阶段, 簇头*j*将其簇内数据序列 $y_j(t) = [y_{j,1}(t), y_{j,2}(t), \dots, y_{j,N_j}(t)]^T$ 多跳传输至汇聚节点, 其中, N_j 为第*j*个簇中的节点数目。假设 $J = \{j_1, j_2, \dots, j_H\}$ 为多跳传输链路中其数据能

够到达簇头*j*的簇头的集合, 考虑以下两种情况:

- 若 $\sum_{k=1}^H N_{j_k} + N_j < M$, 其中, M 为汇聚节点用于重构的向量维度, 则簇头*j*直接传输接收到的数据序列及其本身的数据序列, 广播的序列为:

$$y_j^*(t) = [y_{j_1}^T(t), y_{j_2}^T(t), \dots, y_{j_H}^T(t), y_j^T(t)]^T \quad (13)$$

$y_j^*(t)$ 中元素数目为:

$$q_j = \sum_{k=1}^H N_{j_k} + N_j \quad (14)$$

- 若 $\sum_{k=1}^H N_{j_k} + N_j \geq M$, 则簇头*j*将接收到的数据及其本身的数据利用对应的压缩感知编码系数编码为 M 维数据系列后传输, 广播的序列为:

$$y_j^*(t) = \sum_{k=1}^H \Phi_{j_k} y_{j_k} + \Phi_j y_j \quad (15)$$

$y_j^*(t)$ 中元素数目为:

$$q_j = M \quad (16)$$

综合以上两种情况, 在簇间多跳传输的过程中, 当簇头节点所需传输的数据数目小于 M 时, 直接传输; 反之, 簇头节点将数据编码为 M 维后传输。

传输完成后, 汇聚节点整合接收到的数据, 得到 M 维观测数据:

$$z(t) = \Phi y(t) = \sum_{k=1}^p \Phi^k y_k(t) \quad (17)$$

利用压缩感知解码算法重构出原始数据向量。

3.2 能耗分析

在本文中主要考虑传输过程的能量消耗。由图 2 可知, 数据收集的过程主要包括簇内传输和簇间传输, 因此新方法的能耗主要考虑这两部分数据传输的能耗, 即簇内传输能耗和簇间传输能耗。在节点发射功率一定的情况下, 数据传输的次数直接反映了数据传输的能耗。因此, 在本文中, 将用数据传输的次数来表征



数据传输的能耗。

(1) 簇内传输阶段总传输次数 N_1

假设在簇内传输阶段的初步预测值和实际观测值的差 $e_i(t) = (s_i'(t|t-1) - x_i(t))$ 服从 $N(0, \sigma)$ 分布, 对应给定的门限值 ε , 预测不合理需传输的概率 α 为:

$$\alpha = P(|e_i(t)| > \varepsilon) = 2 \left(1 - \Phi \left(\frac{\varepsilon}{\sigma} \right) \right) \quad (18)$$

则簇内传输阶段总传输次数为:

$$N_1 = \alpha NT = 2NT \left(1 - \Phi \left(\frac{\varepsilon}{\sigma} \right) \right) \quad (19)$$

其中, N 为网络中节点个数, 各节点需传输 T 个时刻采集的数据。

(2) 簇间传输阶段总传输次数 N_2

$$N_2 = T \sum_{k=1}^P q_k \quad (20)$$

其中, P 为网络中簇的数量, q_k 为第 k 个簇头传输的数据量, 具体见式 (14) 和式 (16)。

结合式 (19) 和式 (20), 完成数据收集所需的总传输次数为:

$$N_T = N_1 + N_2 = 2NT \left(1 - \Phi \left(\frac{\varepsilon}{\sigma} \right) \right) + T \sum_{k=1}^P q_k \quad (21)$$

4 仿真结果与分析

为验证所提方法的有效性, 在 MATLAB 环境下进行验证。设定网络为随机部署的方式, 网络规模 N 为 500~1500, 簇头节点通过 LEACH 方式产生, 比例为 10%。每个节点有 $T=100$ 个时段内采集的数据待收集。未进行任何压缩的直接传输方案及参考文献[9,11]中的方法用于与所提方法的性能进行对比。

图 3 给出了簇内预测传输时采用 Kalman 预测及 Autoregressive (AR) 预测的传输百分比与设定的误差门限在 $N=1000$ 时的关系。图 4 给出了对应的均方误差 (MSE) 与设定的误差门限的关系。其中传输百分比为预测值与观测值的差超过预设

门限, 认为预测不合理而发送的数据占总的待收集的数据的百分比。传输比率越大, 说明传输的次数越多, 消耗的能量越多。由图 3 可以看出, 采用 Kalman 预测时, 预测合理的概率更高, 传输次数少于采用 AR 预测的情况, 然而, 由图 4 可以看出, 在减少了传输次数的同时, 采用 Kalman 预测并没有牺牲过多的 MSE 性能, 甚至, 在门限大于 0.8 时, 采用 Kalman 预测时 MSE 性能略优于 AR 预测, 这是由于在考虑了观测噪声的情况下, 采用了 Kalman 预测时, 对于发送出去的数据进行了修正, 缓解了观测噪声对环境真实值的影响, 因此所提方法采用 Kalman 预测进行簇内预测传输。

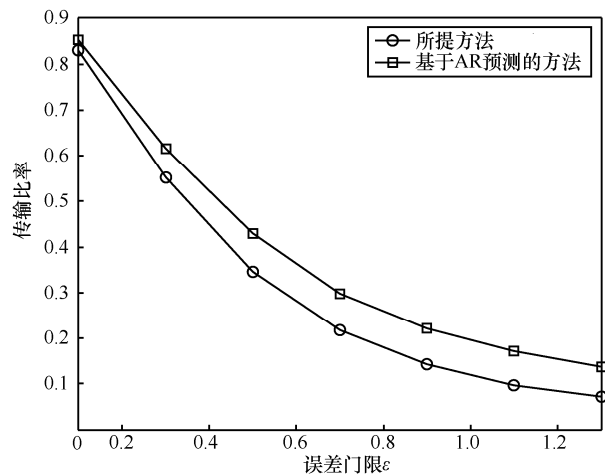


图 3 所提方法与基于 AR 预测的方法的传输百分比与误差门限的关系

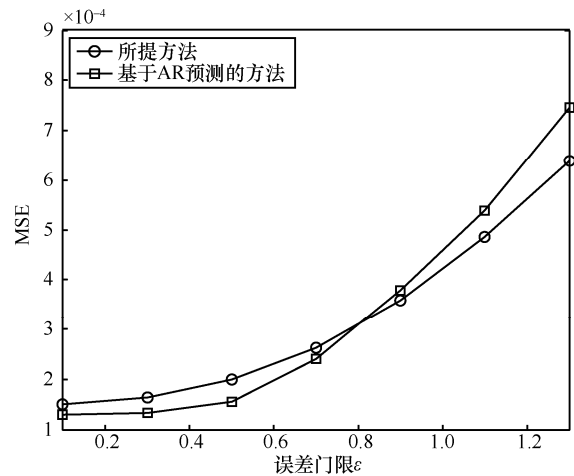


图 4 所提方法与基于 AR 预测的方法的 MSE 与误差门限的关系

图5给出了不同方案的总传输次数与网络规模的关系,其中参考文献[11]及所提方法中的预先设定的误差阈值均为 $\varepsilon=0.5$ 。由图5可以看出,与直接传输相比,所提方法减少了至少50%的传输次数,与参考文献[9]及参考文献[11]中的方法相比,所提方法分别能够减少约25%、43%的传输次数,这极大地减少了传输能耗。这主要是因为所提方法同时利用了数据间的时空相关性在簇内和簇间都对数据进行了压缩传输。

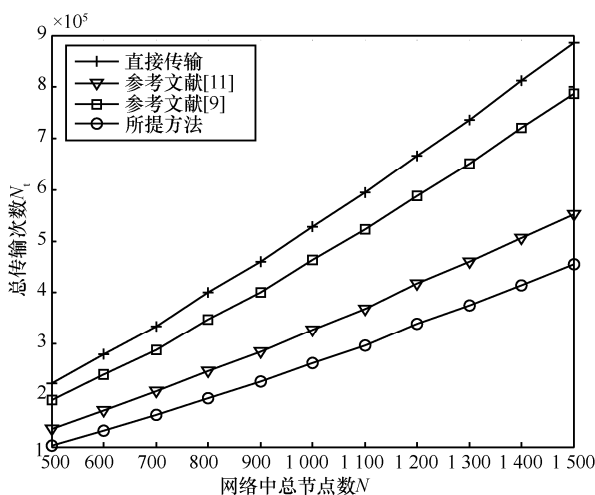


图5 不同方法的总传输次数与网络规模的关系

图6给出了不同方案的MSE性能与网络规模的关系,其中参考文献[11]及所提方法中预先设定的误差阈值均为 $\varepsilon=0.5$ 。

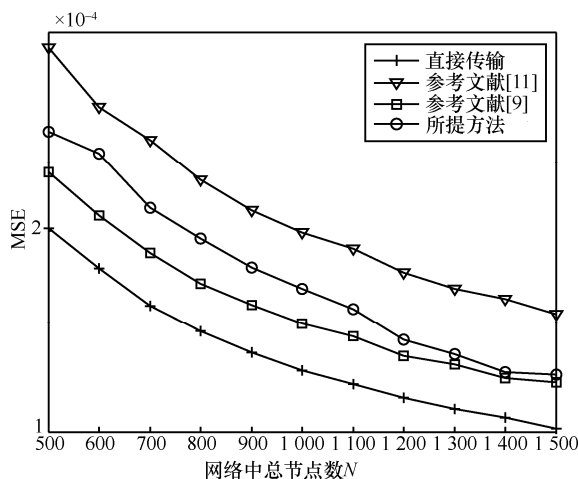


图6 不同方法的MSE性能与网络规模的关系

由图6可以看出,直接传输方法与参考文献[9]方法的MSE性能较优,新方法的MSE性能次之。但对应于图5,直接传输方法与参考文献[9]方法的传输能耗均非常大,接近新方法的两倍,而新方法的MSE性能在网络规模增大的情况下,与以上两种方法的差距逐渐减小。这说明,新方法能够在保证重构精度的前提下,显著地减少数据收集的能耗,其能很好地适用于大规模的无线传感器网络。

5 结束语

本文提出了一种联合卡尔曼滤波及混合压缩感知的高能效数据收集方法。基于无线传感器网络中数据的时空相关性,簇内传输时,成员节点利用卡尔曼预测,仅仅传输预测不合理的数据,减少数据间时间冗余。同时,在簇间传输时,利用混合压缩感知减少节点发送的数据量,消除数据间的空间冗余信息。新方法可以在保证数据重构精度的前提下减小能耗、延长网络生存时间,适用于大规模的无线传感器网络。

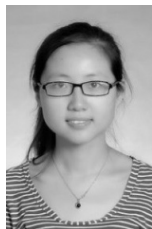
参考文献:

- [1] AKYILDIZ I F, SU W. A survey on sensor networks[J]. IEEE Communications Magazine, 2002, 40(8): 102-114.
- [2] 刘述木, 杨建, 黎远松. WSN中基于自适应预测聚类的多组群目标的跟踪方法[J]. 电信科学, 2016, 32(7): 68-75.
LIU S M, YANG J, LI Y S. Multi-group target tracking method based on adaptive predictive clustering in WSN[J]. Telecommunications Science, 2016, 32(7): 68-75.
- [3] NAVARRO K F, LAWRENCE E. WSN applications in personal healthcare monitoring systems: a heterogeneous framework[C]//Second International Conference on Ehealth, Telemedicine, and Social Medicine, Feb 10-16, 2010, St. Maarten, Netherlands Antilles. Piscataway: IEEE Press, 2010: 77-83.
- [4] 柯炜. 无线传感器网络关键技术及其研究难点[J]. 电信科学, 2005, 21(6): 9-12.
KE W. Key technology and research challenges of wireless sensor network[J]. Telecommunications Science, 2005, 21(6): 9-12.
- [5] DUARTE M F, SARVOTHAM S, BARON D, et al. Distributed compressed sensing of jointly sparse signals[C]//2005 Confer-



- ence Record of the Thirty-Ninth Asilomar Conference on Signals, Systems and Computers, Oct 30-Nov 2, 2005, Pacific Grove, CA, USA. Piscataway: IEEE Press, 2005: 1537-1541.
- [6] NGUYEN N, JONES D L, KRISHNAMURTHY S. Netcompress: coupling network coding and compressed sensing for efficient data communication in wireless sensor networks[C]//2010 IEEE Workshop on Signal Processing Systems, Oct 6-8, 2010, San Francisco, CA, USA. Piscataway: IEEE Press, 2010: 356-361.
- [7] MORELL A, CORREA A, BARCELO M, et al. Data aggregation and principal component analysis in WSN[J]. IEEE Transactions on Wireless Communications, 2016, 15(6): 3908-3919.
- [8] JIANG H, JIN S, WANG C. Prediction or not? An energy-efficient framework for clustering-based data collection in wireless sensor networks[J]. IEEE Transactions on Parallel and Distributed Systems, 2011, 22(6): 1064-1071.
- [9] ARUNRAJA M, MALATHI V. Collective prediction exploiting spatio temporal correlation (CoPeST) for energy efficient wireless sensor networks[J]. Ksii Transactions on Internet & Information Systems, 2015, 9(7): 2488-2511.
- [10] WU M, TAN L, XIONG N. Data prediction, compression, and recovery in clustered wireless sensor networks for environmental monitoring applications[J]. Information Sciences, 2016, 329(C): 800-818.
- [11] QUAN L, XIAO S, XUE X, et al. Neighbor-aided spatial-temporal compressive data gathering in wireless sensor networks[J]. IEEE Communications Letters, 2016, 20(3): 578-581.
- [12] XIE R, JIA X. Transmission-efficient clustering method for wireless sensor networks using compressive sensing[J]. IEEE Transactions on Parallel & Distributed Systems, 2014, 25(3): 806-815.

[作者简介]



周颖(1990-), 女, 南京邮电大学博士生, 主要研究方向为无线传感器网络、无线通信和物联网等。



杨丽花(1984-), 女, 南京邮电大学副教授, 主要研究方向为无线通信、无线传感器网络和网络编码等。



杨龙祥(1966-), 男, 南京邮电大学教授, 主要研究方向为无线通信、移动通信系统和物联网。



倪梦(1993-), 女, 南京邮电大学硕士生, 主要研究方向为无线传感器网络、移动通信和物联网等。