

基于最优阶估计与分布式分簇的传感器网络数据压缩方法研究

蒋 鹏* 李胜强

(杭州电子科技大学自动化学院 杭州 310018)

摘 要: 在无线传感器网络的诸多应用中, 被监测区域发生异常情况的概率通常较小, 正常情况下, 同一传感器节点在前后连续时刻所采集的数据具有时间相关性, 处于相邻区域的不同传感器节点在同一时刻所采集的数据具有空间相关性, 发送存在时间、空间冗余的数据至基站必将耗费节点大量的能量。该文提出了基于最优阶估计和分布式分簇的传感器网络数据压缩方法, 利用节点采集数据的时空相关性, 基于最优阶估计在基站处建立相关系数, 经分布式分簇, 节点仅需传送少量数据, 基站根据时空相关性恢复原始数据。仿真结果表明应用该算法, 可以有效减少传感器网络中冗余数据传输量和节点能耗, 进而延长系统寿命。

关键词: 无线传感器网络; 数据压缩; 最优阶估计; 分布式分簇

中图分类号: TP393

文献标识码: A

文章编号: 1009-5896(2011)03-0569-06

DOI: 10.3724/SP.J.1146.2010.00529

Research on Data Compression Algorithm for Wireless Sensor Networks Based on Optimal Order Estimation and Distributed Clustering

Jiang Peng Li Sheng-qiang

(School of Automation, Hangzhou Dianzi University, Hangzhou 310018, China)

Abstract: The possibility of occurring exception is relatively small in most applications of wireless sensor networks. So data obtained in sequent moment by the same node have time correlation, and data obtained in the same time by adjacent nodes have space correlation. A large number of energy of node will be wasted if data which include time and space correlation is transmitted. Therefore, this paper proposed a data compression algorithm for wireless sensor networks based on optimal order estimation and distributed clustering. The algorithm explores the time and space correlation among data obtained by sensors. The correlation parameter can be get based on optimal order estimation. Then all data can be restored based on time and space correlation parameters and only a little necessary data are transmitted by nodes. Because redundancy is decreased when data is transmitted, the average energy cost of node is reduced and the life of the whole wireless sensor networks can be extended.

Key words: Wireless sensor networks; Data compression; Optional order estimation; Distributed clustering

1 引言

无线传感器网络中的数据压缩方法就是在保证一定数据精度的前提下, 寻求一种有效的减少数据传输量, 从而降低节点能耗, 提高整个网络综合性能的方法。近年来, 应用于无线传感器网络的数据压缩方法得到了广泛的研究。Antonios 等人^[1]提出的 SBR(Self-Base Regression)算法, 该算法先将节点连续采集的数据值分割为若干可变长线性数据段, 再基于人工组建的基信号, 对数据段进行编码, 基信号的值随节点采集值动态变化, 算法关键是用

于编码的基信号, 算法性能在很大程度上受采集信号的分段线性程度影响。Song 等人^[2]提出了无线传感器网络在线信息压缩算法, 该算法首先将节点搜集的数据分成不同长度的数据段, 再对每一个数据段分别编码构造成字典, 其中编码过程由 2 级聚合子程序实现, 由于引入寻求次优字典数据段子算法及字典更新编码子算法, 该算法的复杂度略高于 SBR 算法。Rassi 等人^[3]提出的多维时间序列模型挖掘算法, 通过对时间序列数据流最少的扫描, 不仅对同维实时数据实现基于出现频率的模式编码压缩, 还实现了多个物理量即多维实时数据的模式编码压缩, 该算法不足在于需要首先对采集数据进行扫描, 增加了节点的存储负担和能耗, 影响了系统实时性。Pattem 等人^[4]提出的基于数据空间冗余的路径选择压缩算法, 该算法将传感器网络中数据传

2010-05-24 收到, 2010-09-20 改回

国家自然科学基金(60974012), 浙江省自然科学基金(Y1100054), 浙江省科技厅科技计划重点项目(2008C23097)和杭州市科技局科技计划项目(20091133B03)资助课题

*通信作者: 蒋鹏 pjiang@hdu.edu.cn

输时的路由选择与数据压缩综合起来考察,通过平衡相邻节点数据的相关性与所划分簇大小的关系,从而最大限度地消除传感器网络中数据的空间冗余,进而延长节点的平均寿命,该算法要求对监控区域的节点进行精确的坐标设置,这在实际应用中较难实现,同时检测数据的空间相关性很大程度上影响着算法性能,且相关性参数需要大量实际验证估计才能获得。Jia 等人^[5]提出的 GIST(Group-Independent Spanning Tree)算法也是为消除数据空间冗余而设计的,该算法虽然建立了较完整的针对传感器网络的数据传输结构,但对具体应用时需要设置的系统参数探讨不多,也没有关于最小区域数据冗余消除策略的深入探讨。周四望等人^[6]提出了基于环模型的分布式时-空小波数据压缩算法,该算法以小波变换为基本工具,设计出适合任意支撑长度的小波函数的小波变换环模型,能够有效消除环上节点数据的时间空间冗余。Jim 等人^[7]提出的分布式结构树压缩(Distributed Structure Tree Depression, DSTD)算法,即利用节点感知数据中存在的时间、空间相关性,在基站处根据预先求得的相关系数及少量原始数据恢复全部原始数据。

本文提出了最优阶分布式分簇结构树压缩(Cluster Optimal Order Estimation Distributed Structure Tree Depression, COOE-DSTD)算法,即一方面引入最优阶估计,从而界定算法每个周期系统要传输的原始数据组数,既要获得有效的相关性系数,又要避免冗余系数的产生,另一方面对网络进行分簇处理,以簇为单位处理节点数据,不仅可提高基站处理各节点数据的效率,还可增强基站迅速定位产生异常值或出现异常节点的能力,从而有效地提高整个传感器网络对监控区的监控能力。本文的第2节将详细说明 DSTD 算法所基于的分布式编码基本原理,及 COOE-DSTD 算法引入的最优阶估计模型;第3节是 DSTD 及 COOE-DSTD 的算法描述及流程图;第4节分别以节点平均能耗,信噪比以及两者的比值作为评价算法性能优劣的指标,通过具体的仿真算例,进一步分析算法性能;最后是结束语。

2 分布式编码及最优阶估计基本原理

分布式编码理论基础,源于由 Wolf 等人^[8]验证的一个关于同域数据源编码的定理。该定理表述如下:若两个相关随机变量 X, Y 服从任意概率分布 $P(X, Y)$, 当编码器对 X 进行压缩时,无论其对 X 与 Y 相关性是否可知,其压缩性能不变。更一般的表述是,在 Y 与 X 相关性不可知的情况下, X 可用

$H(X|Y)$ 比特位压缩码表示,其中

$$H(x|y) = -\sum_y P_Y(y) \sum_x P_X(x|y) \log P_X(x|y) \quad (1)$$

式(1)中 $H(X|Y)$ 指给定观察值 Y 条件下,变量 X 的不确定性。现举例说明,设变量 X, Y 为两个等概率分布在三位二进制数组成的数据集中,即 X, Y 等概率分属集合 $\{000, 101, 011, 110\}, \{001, 010, 100, 111\}$ 。变量 X, Y 服从 $d(X, Y)$ 不大于 1, 其中, $d(X, Y)$ 为 Hamming 距离函数。当编码器、解码器对 Y 均可知时,可将 X 压缩至两比特位,以表示在 Y 已知的情况下, X 的不确定性信息,若仅有解码器可知 Y 时,仍可将 X 压缩至 2 位。具体操作如下,先将 3 位二进制组成的集合分为 4 个陪集,即 $\text{coset1}=(000, 111), \text{coset2}=(001, 110), \text{coset3}=(010, 101), \text{coset4}=(011, 100)$, 而对 4 个陪集编码仅需 2 位二进制码即可,假设 X 属于陪集 coset3 , 这时只需结合 Y 值及 X, Y 服从 $d(X, Y)$ 不大于 1, 便能对 X 正确解码,而不需要耗费额外比特位用于确定陪集中的 X 。由于传输原始数据需要 $H(Y)$ 位,而传输压缩码需要 $H(X|Y)$ 位,即非对称编码,故将分布式编码方法应用到无线传感器网络的数据压缩时,会遇到节点能耗不均衡问题,在此利用时间段轮转的方式,让不同节点轮流传输原始值。

COOE-DSTD 算法所依据的最优阶估计的基本原理,即对于一个给定的自回归(Auto aggressive, AR)模型,其性能取决于有限样本所在的实际过程,样本容量,采用的估计样本及阶数选取准则。有限样本准则给出了残差能量统计平均及回归估计算法的预测方差的经验估计,这使得用于阶数选择的有限样本准则性能的优劣取决于所采用的估计方法。FSIC(Finite Sample Information Criterion)^[9] 及 CIC(Combined Information Criterion)准则^[10] 作为两个重要的估计准则,均考虑了随模型阶数增加而增大的残余能量方差,现具体说明 FSIC, CIC 两个阶估计准则,

$$\text{FSIC}(g) = \ln\{\text{RET}(g)\} + \prod_{d=0}^g \frac{1+V(d,\bullet)}{1-V(d,\bullet)} - 1 \quad (2)$$

方差系数 $V(d,\bullet)$ 为 AR 模型有限样本准则的重要组成部分, g 指模型阶数,“ \bullet ”代表具体估计方法,本文所用估计方法为 Burg, 属于协相关估计方法,其中 $\text{RET}(g)$ 均值定义为

$$E\{\text{RET}(g)\} \approx \delta_c^2 \prod_{d=0}^g \{1 - V(d,\bullet)\}, \quad g \geq R \quad (3)$$

式(3)中 δ_c^2 指完全稳定随机过程残余方差, R 指实际

过程阶数, 与其它通过对 $\ln\{\text{RET}(g)\}$ 修正产生的准则相比, 当 $g/N > 0.1$ 时, 该准则能够更好地反映 g/N 变化较大部分, 即能够更好地获取模型最优阶, 其中, N 指观测样本容量。然而, 当模型最优阶数较小时, FISC 准则就无法给出较好的阶数判定, 而 CIC 准则很好地解决了这一问题, 该准则融合了适合低阶最优估计的惩罚因子和 FSIC 准则的优点, 即无论模型最优阶数相对于 N 较小还是较大, CIC 准则对最优阶数的选取均不会偏低或偏高, 该准则定义如下^[10]:

$$\text{CIC}(g) = \ln\{\text{RET}(g)\} + \prod_{d=0}^g \frac{1+V(d,\bullet)}{1-V(d,\bullet)} - \min \left\{ 1, 3 \sum_{d=0}^g V(d,\bullet) \right\} \quad (4)$$

对于最优阶数较小的模型, 该准则能够通过惩罚因子 $3 \sum_{d=0}^g V(d,\bullet)$, 避免模型最优阶数被过低估计, 对于最优阶数较大的模型, CIC 准则可自动切换为 FSIC 形式, 从而避免对阶数的过高估计。

3 算法描述和流程图

基于上节所述的分布式编码, 对算法涉及到的结构树, 预测系数及编码指令 i 值的确定作进一步的说明, 结构树首先是以前述 $N/3$ 组原始数据的均值为起点, 以 Δ 为间隔向两端扩展, Δ 大小决定算法精度, 扩展范围由具体应用决定, 对上述以 Δ 为间隔的扩展序列进行奇偶序列分裂, 可得到间距为 2Δ 的两组子序列, 以同样的方式可再对子序列进行分裂, 分裂次数由 i 的位数决定, i 次分裂后, 各子序列相邻节点间距为 $2^i \Delta$, 第 i 层每个子序列对应唯一的 i 位编码, 从而形成 i 层结构树。由于监控数据时空上的相关性, 可通过建立一个预测模型^[7], 由已获得的节点监测值, 预测当前或将来时刻节点监测值。设 T 时刻, j 节点监测值为 X_T^j , 通过预测模型获得其线性拟合值 Y_T^j , 则有

$$Y_T^j = \sum_{k=1}^W \alpha_k X_{T-k}^j + \sum_{p=1}^S \beta_p X_T^p \quad (5)$$

其中 $\sum_{k=1}^W X_{T-k}^j$ 表示同一节点 j 在 T 时刻之前 W 个时刻的监测值, $\sum_{p=1}^S X_T^p$ 表示 T 时刻与节点 j 相邻的 S 个节点的监测值, $\alpha_k (k=1, 2, \dots, W)$ 及 $\beta_p (p=1, 2, \dots, S)$ 分别为同节点时间域、同时刻空间域的预测系数, 式(5)中 W, S 分别表示在同节点时域、同时刻空域监测值数目范围, 要使拟合值 Y_T^j 尽可能逼近 X_T^j , 则有

$$\begin{aligned} E[(N_j)^2] &= E \left[\left(\sum_{k=1}^W \alpha_k X_{T-k}^j + \sum_{p=1}^S \beta_p X_T^p - X_T^j \right)^2 \right] \\ &= E[(X_T^j)^2] - 2 \sum_{k=1}^W \alpha_k E[X_T^j X_{T-k}^j] \\ &\quad - 2 \sum_{p=1}^S \beta_p E[X_T^j X_T^p] + 2 \sum_{k=1}^W \sum_{p=1}^S \alpha_k \beta_p E[X_{T-k}^j X_T^p] \\ &\quad + \sum_{k=1, q=1}^W \alpha_k \alpha_q E[X_{T-k}^j X_{T-q}^j] \\ &\quad + \sum_{p=1, q=1}^S \beta_p \beta_q E[X_T^p X_T^q] \end{aligned} \quad (6)$$

设同时刻不同节点随机监测过程 X_T^h 与 X_T^p ($h, p=1, 2, \dots, S$) 为两两联合宽平稳过程, 则

$$E[(N_j)^2] = c_{x_j x_j}(0) - 2P_j^T \Phi_j + \Phi_j^T C_{rr}^j \Phi_j \quad (7)$$

其中 $P_j = [c_{x_j x_j}(1) \ c_{x_j x_j}(2) \ \dots \ c_{x_j x_j}(W) \ c_{x_j x_1}(0) \ c_{x_j x_2}(0) \ \dots \ c_{x_j x_S}(0)]$, $c_{x_j x_p}(k) = E[X_{T-k}^j X_T^p]$, $\Phi_j = [\alpha_1 \ \alpha_2 \ \dots \ \alpha_W \ \beta_1 \ \beta_2 \ \dots \ \beta_S]^T$, $C_{rr}^j = \begin{bmatrix} C_{x_j x_j} & C_{x_j x_p} \\ C_{x_j x_p}^T & C_{x_p x_p} \end{bmatrix}$ 。

令 $E[(N_j)^2]$ 对 Φ_j 的偏导数为零得

$$\Phi_j = (C_{rr}^j)^{-1} P_j \quad (8)$$

在此通过让 Φ_j 沿着目标函数(即均方误差函数 $E[(N_j)^2]$)梯度的相反方向移动, 实现对 Φ_j 的动态调整, 即

$$\Phi_j^{(t+1)} = \Phi_j^t - \mu \nabla_j^t \quad (9)$$

其中 ∇_j^t 由 $E[(N_j)^2]$ 对 Φ_j 求偏导获得, μ 代表目标函数沿其梯度相反方向的下降系数, 由于目标函数为凸函数, 故目标函数存在最小值, 当选取合适的 μ 时, ($\mu < 2/\lambda_{\max}$, λ_{\max} 代表矩阵 C_{rr}^j 最大特征值), 则式(9)可以收敛到最优解, Φ_j 的更新公式归结为

$$Y_t^j = (\Phi_j^t)^T U_{t,j} \quad (10)$$

$$N_{t,j} = X_t^j - Y_t^j \quad (11)$$

$$\Phi_j^{t+1} = \Phi_j^t + \mu U_{t,j} N_{t,j} \quad (12)$$

至此, 获得了能够随监测值变化而动态调整的更新系数 Φ_j , 从而可以求得边信息 Y_t^j , 因为只有当基站结构树中第 i 层子序列相邻节点间距 $2^{i-1} \Delta$ 大于 $N_{t,j}$ 时, 才可能通过边信息 Y_t^j 估计出 X_t^j , $N_{t,j}$ 的均值为 0, 方差为 $\delta_{N_j}^2$ 。若使 $2^{i-1} \Delta$ 以概率 P 大于 $N_{t,j}$, 根据契比雪夫不等式可推得:

$$i = \frac{1}{2} \log_2 \frac{\delta_{N_j}^2}{(1-P)\Delta^2} + 1 \quad (13)$$

分布式结构树压缩算法和最优阶分布式分簇结构树压缩算法的流程图分别如图 1, 图 2 所示。最优阶分布式分簇结构树压缩算法的实现步骤如下:

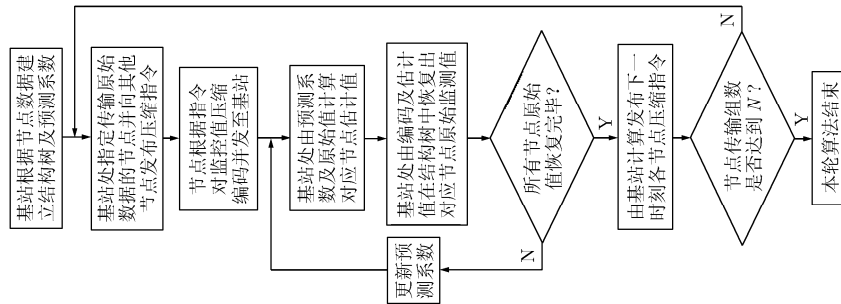


图1 分布式结构树压缩算法流程图

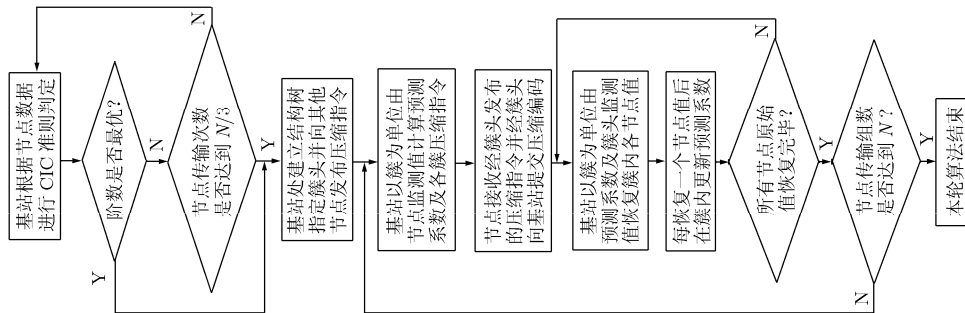


图2 最优阶分布式分簇结构树压缩算法流程图

(1)由监测区节点向基站传输各自监测值,所有节点每传输完一组值,基站处采用CIC准则进行最优阶判定,若所得阶数不是最优,且节点所传数据组数没有超过 $N/3$,节点继续传输监测值,否则由节点所传输的数值获得初始的预测系数 ϕ_m ,并在基站处建立结构树;

(2)基站根据节点采集数据空间相关性,将监控区节点划分为若干簇,以簇为单位,基站根据 i 值计算公式,产生簇内各节点所要传输的二进制位数 i ,同时基站轮流指定簇内节点担当簇头,簇头负责传输自身未经压缩的监测值及簇内其它节点的压缩编码至基站,同时接收来自基站的指令;

(3)簇内各节点经簇头依次获得基站指令 i 值,各节点将其模数转换后的二进制值对 2^i 取模,得到 i 位二进制压缩码,将该压缩码经簇头传至基站,同时,簇头向基站传输其未经压缩的监测值;

(4)基站处以簇为单位由步骤(1)中计算的预测系数 ϕ_m 及簇头所传原始监测值,可得到簇内各节点估计值 Y_r^m (对应第 m 个节点第 r 组监测值的估计值),再由该节点的 i 位压缩码可以在结构树中定位出一个子序列,由于 $2^{i-1}\Delta$ 以概率 P 大于 $N_{r,m}$,故在该子序列中离 Y_r^m 最近的序列值即为节点最终估计值 X_r^m ,同时, X_r^m 又成为估计同一时刻下一节点的已知条件,即在更新 ϕ_m 时要考虑新估计出的数值;

(5)经步骤(4),当各簇所有节点值被恢复完毕,

由基站计算并发布下一时刻各簇内节点需要将各自监测值压缩的位数 i ,当节点监测到的数值组数达到 N ,本轮算法结束,否则进行步骤(2)。

4 仿真算例与分析

与DSTD算法相比,COOE-DSTD算法的优点在于:

(1)由于节点能量主要消耗在传输原始监测信息上,通过引入最优阶估计,减少了节点向基站传输的原始数据的组数,从而降低了节点平均能耗;

(2)将计算集中于时间域和空间域上与待估计节点值更加相关的系列节点值上,剔除了与待估计节点值时空相关性微弱的节点值,从而有效降低了算法所处理数据的维度,提高了算法运算速度及节点数据恢复精度;

(3)当监控区域网络规模较大时,基站对网络节点分簇分层处理后,若监控区域出现异常,基站能够迅速定位异常区所在的簇,进而与监测到异常的节点直接建立连接,与在全区域逐个搜寻相比,可有效降低系统时延,提高系统监控实时性。

4.1 算法性能评价模型

为了合理地评价算法性能,分别将节点平均能耗、信号传输信噪比以及节点平均能耗与信号传输信噪比的比值作为算法性能评判标准。其中节点能耗 E 的计算参考文献[11]中的描述,以StrongARM SA-1100A为例,节点总能耗为

$$E = E_{lp} + E_{lt} + E_{lr} + E_{rt} \quad (14)$$

其中节点运算能耗 $E_{lp} = UCvdd^2$, U 代表节点主芯片执行每条指令的时钟周期数, C 代表切换电容均值, 对于 StrongARM SA-1100A, C 约为 0.67 nf, Vdd 代表电源电压, 节点数据交互能耗 $E_{lt} = E_{elec}k + \varepsilon_{amp}kD^2$, 其中 E_{elec} , ε_{amp} 为恒定参数, k 为发送数据位数, D 代表发送数据的距离, 节点接收数据能耗 $E_{lr} = E_{elec}k$, E_{rt} 表示节点向基站发送数据能耗, 计算公式同 E_{lt} 。峰值信噪比计算模型为

$$SNR = 10 \lg \left[\frac{AM}{\sum_{x=0}^{M-1} (f(x) - \hat{f}(x))^2} \right] \quad (15)$$

式(15)中 A 表示传输的最大数据值, M 表示输入序列数, $f(x)$ 表示输入数据, $\hat{f}(x)$ 表示输出数据。

4.2 仿真实验与分析

在此假设考察的网络节点均匀分布于监测区内, 节点之间距离为 1 m, 节点距基站为 1000 m。图 3 为分别采用 DSTD 算法及 COOE-DSTD 算法处理数据时 SNR 性能比较仿真图, 当网络节点数恒定时, DSTD 算法的 SNR 随算法单个周期内包含节点采集的数据组数的增加而逐渐降低, 而 COOE-DSTD 算法的 SNR 则先逐渐增加, 接着稳定在一个较高的位置小幅波动, 这表明算法改进后, 恢复的节点监测值精度得到了提高。因为 DSTD 算法在恢复当前时刻节点值时, 处于该节点当前时刻之前的监测值, 被纳入相关性计算的数目增多了, 这就引入了较远时刻相关性微弱的监测值, 相当于引入了新的误差, 导致 DSTD 算法的 SNR 下降, 而 COOE-DSTD 算法通过最优阶估计, 在恢复当前时刻节点值时, 将运算收缩到在时间上与当前节点有较强相关性的若干监测值上, 不仅能有效提高恢复精度, 还可以减小运算量。图 4 显示, 随着网络节点数 J 的增加, DSTD 算法及 COOE-DSTD 算法的信噪比 SNR 均有小幅波动, 但后者信噪比始终高于

前者, 其原因同上述分析。

由图 5 可知, 随着算法单个周期包含的节点采集数据次数 N 的增加, DSTD 算法及 COOE-DSTD 算法节点平均能耗(Average Energy Cost, AEC)均有一定的波动, 但后者平均能耗始终低于前者, 原因在于随着算法单个周期包含的节点采集数据次数 N 的增加, DSTD 算法在运行的初始阶段传输的原始数据增多, 这必然要耗费节点大量能量, 而 COOE-DSTD 算法通过最优阶估计, 经 CIC 准则获得最优阶后, 仅传输少量的原始数据, 从而降低了节点平均能耗。图 6 显示, 随着网络节点数 J 的增加, COOE-DSTD 算法节点平均能耗低于 COOE-DSTD 算法, 原因同图 5 分析, 同时可以发现在网络节点数为 120 时, 节点平均能耗出现了一个局部最小值点, 在应用 COOE-DSTD 算法时, 若网络节点规模在 100 左右时, 可选定为 120 个节点, 从而实现节点平均能耗局部最小。

由图 7 及图 8 可知, 将节点平均能耗与信噪比的比值 AEC/SNR 作为算法性能评价指标时, 得到了与前述分析相一致的仿真结果, 即随着算法单个周期包含的节点采集数据次数 N 的增加, DSTD 算法及 COOE-DSTD 算法中均有一定的波动, 但后者始终低于前者; 随着网络节点数 J 的增加, DSTD 算法的 AEC/SNR 始终高于 COOE-DSTD 算法, 即通过增加最优阶估计环节以及对监控区节点进行分簇分层处理, 有效减少了基站及节点的实际运算量, 进而降低了节点平均能耗, 提高了节点监测值的恢复精度。

5 结束语

本文提出了 COOE-DSTD 算法, 与 DSTD 算法相比, 通过引入最优阶估计环节及对监控区节点分布式分簇处理, 可以获取估计模型的最优阶, 不仅减少了算法实际运算量, 提高了基站处数据恢复精度, 而且有效降低了节点平均能耗; 节点对监控值分布式编码, 仅需简单的取模运算, 有利于降低

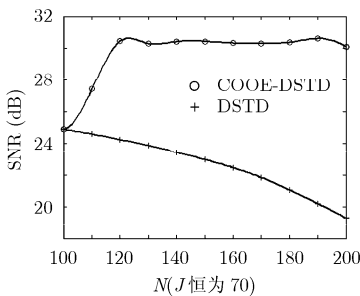


图 3 SNR 随算法周期内采集时刻数 N 变化的关系图

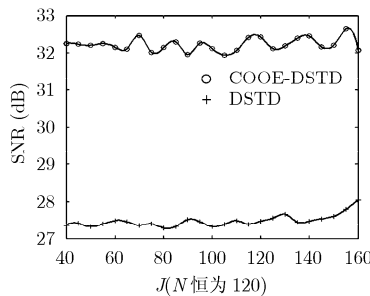


图 4 SNR 随网络节点数 J 变化的关系图

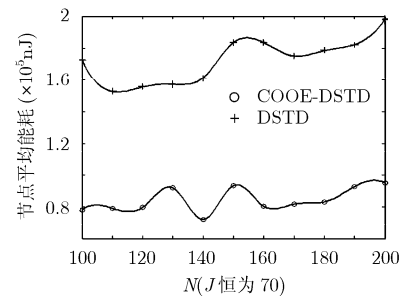


图 5 AEC 随算法周期内采集时刻数 N 变化的关系图

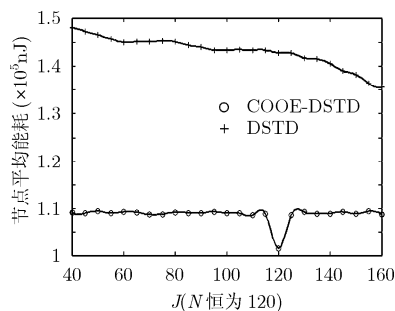


图 6 AEC 随网络节点数 J 变化的关系图

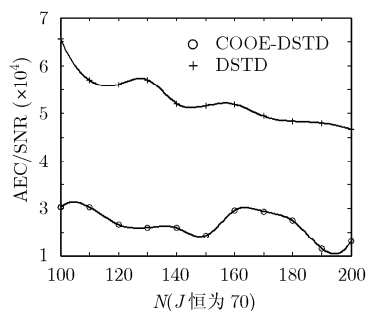


图 7 AEC/SNR 随采集时刻数 N 变化的关系图

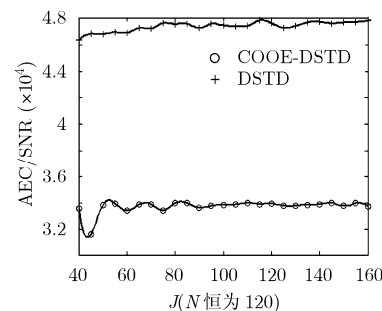


图 8 AEC/SNR 随网络节点数 J 变化的关系图

节点能耗；采用分簇划分监控区节点，从而使基站能够迅速定位节点，有利于减小系统时延及扩展监控网络。然而本文的 COOE-DSTD 算法主要针对 1 维数据，关于如何压缩无线传感网络 2 维数据的相关算法的研究将是我们下一步工作的重点。

参 考 文 献

- [1] Antonios D and Yannis K. Another outlier bites the dust: computing meaningful aggregates in sensor networks[C]. Proceedings of the 2009 IEEE International Conference on Data Engineering, Washington, DC, USA, March, 2009: 988-999.
- [2] Song Lin, Kalogeraki V, Gunopulos D, and Lonardi S. Online information compression in sensor networks[C]. 2006 IEEE International Conference on Communications, Istanbul, Turkey, June 2006, Vol.7: 3371-3376.
- [3] Rassi C and Plantevit M. Mining multidimensional sequential patterns over data streams[C]. 10th International Conference on Data Warehousing and Knowledge Discovery, DaWaK 2008 Turin, Italy, September 2-5, 2008: 263-272.
- [4] Pattem S, Krishnamachari B, and Govindan R. The impact of spatial correlation on routing with compression in wireless sensor networks [J]. *ACM Transactions on Sensor Networks*, 2008(4): 10.1145/1387663. 1387670.
- [5] Jia Lu-jun, Guevara N, Rajmohan R, and Sundaram R. GIST: Group-Independent Spanning Tree for data aggregation in dense sensor networks[C]. Second IEEE International Conference on Distributed Computing in Sensor Systems, DCOSS 2006, San Francisco, CA, United States, 2006: 282-304.
- [6] 周四望, 林亚平, 张建明, 欧阳竟成, 卢新国. 传感器网络中基于环模型的小波数据压缩算法[J]. *软件学报*, 2007, 18(3): 669-680.
- [7] Zhou Si-wang, Lin Ya-ping, Zhang Jian-ming, Ouyang Jing-cheng, and Lu Xin-guo. A wavelet data compression algorithm using ring topology for wireless sensor networks [J]. *Journal of Software*, 2007, 18(3): 669-680.
- [8] Jim Chou, Dragan Petrovic, and Kannan Ramchandran. A distributed and adaptive signal processing approach to exploiting correlation in sensor networks. *Ad hoc Networks*, 2004, 2(4): 387-403.
- [9] Wolf J K, Iyengar A R, and Siegel P H. LDPC codes for the cascaded BSC-BAWGN channel[C]. Allerton 2009. 47th Annual Allerton Conference on Communication, Control, and Computing 2009, Monticello, IL Sept. 30 Oct. 2 2009: 620-627.
- [10] Broersen P M T. The quality of lagged products and autoregressive Yule-Walker Models as autocorrelation estimates [J]. *IEEE Transactions on Instrumentation and Measurement*, 2009, 58(11): 3867-3873.
- [11] Broersen P M T. The removal of spurious spectral peaks from autoregressive models for irregularly sampled data [J]. *IEEE Transactions on Instrumentation and Measurement*, 2010, 59(1): 205-214.
- [12] Wang A and Chandraksan A. Energy-efficient dsps for wireless sensor networks [J]. *IEEE Signal Processing Magazine*, 2002, 19(4): 68-78.

蒋 鹏: 男, 1975 年生, 副教授, 硕士生导师, 研究领域为传感器网络、嵌入式系统及其应用、智能仪表。

李胜强: 男, 1984 年生, 硕士生, 研究领域为无线传感器网络。