

基于无线传感器网络的分布式信源编码方法

叶金昱, 刘 雨, 张 琳

(北京邮电大学信息理论与教研中心, 北京 100876)

摘 要: 针对在无线传感器网络中节点能量有效性较低的问题, 利用分布式信源编码方法, 去除冗余并压缩节点发送的数据。将分布式信源编码方法与对应网络结构相结合, 提出基于簇的无线传感器网络实现分布式信源编码方法。实验结果表明, 该方法能节省网络中 25%~30% 的能量消耗。

关键词: 无线传感器网络; 基于簇的路由协议; 分布式信源编码

Distributed Source Coding Method Based on Wireless Sensor Networks

YE Qian-yu, LIU Yu, ZHANG Lin

(Research Center of Information Theory & Technology Education, Beijing University of Posts and Telecommunications, Beijing 100876)

【Abstract】In Wireless Sensor Networks(WSN), it is not energy efficient to transfer the redundant information. This paper distributes source coding which is utilized to compress the information, but the method needs the suitable network structure. The cluster based network and distributed source coding method are combined to meet the requirement of energy efficiency of WSN. Simulation results show that there is about 25%-30% less energy consumption than the conventional routing method.

【Key words】 Wireless Sensor Networks(WSN); cluster based routing protocol; distributed source coding

1 概述

无线传感器网络(Wireless Sensor Networks, WSN)具有低成本、低复杂性和易扩展的特点, 目前已在很多领域得到广泛应用^[1]。这种网络由大量低功耗的无线传感器构成网络基本单元, 无线传感器一般通过自组织的方式形成路由并传递采集到的数据信息到一个汇聚节点(sink 节点)。在布网时为增强网络的鲁棒性, 一般采用密集的布网方法, 即通过增加单位区域内传感器节点的数目来增强网络的可靠性。一旦这些传感器节点被散布到监控区域中, 它们往往依靠自身携带的电池进行供电。当网络中大部分的传感器节点电能耗尽失效时, 网络的连通性会大受影响甚至网络将不复存在。因此, 在研究设计与传感器网络相关的协议时, 能量有效性很重要。

目前对 WSN 的研究主要集中在数据汇聚和压缩、路由和目标追踪。在 WSN 中往往通过提高单位面积区域内的节点数量来增强网络的可靠性, 这样在物理位置上比较近的传感器节点所采集的信息必然具有很强的相关性。在传感器节点的能耗上传输所占有的比重很大, 而传输所要消耗的能量又和所要传输的数据包的长度成正比。因此, 如果能在传输之前将要传输的数据进行压缩, 可以在很大程度上降低传感器节点的能量消耗从而延长网络的生存期。

本文的主要工作是将目前普遍使用的无线传感器网络基于簇的路由协议^[2]与正在成为热点的分布式信源编码方法结合起来提高网络的能量有效性。

2 分布式信源编码方法

分布式信源编码方法是由 Slepian 和 Wolf 提出的^[3]。其核心的思想为: 无损压缩时相关信源的独立编码(联合译码器

的复杂度增加)和联合编码同样有效。此后 Wyner 和 Ziv 证明了在有损压缩的情况下该结论的正确性^[4]。图 1 为 Slepian and Wolf 的结论。

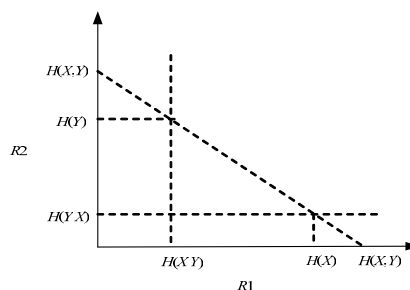


图 1 Slepian 和 Wolf 的结论

由上述分布式信源编码的思想可知, 分布式信源编码可用于压缩具有一定相关性的独立信源上的数据。这些数据由各个独立信源发送到一个统一的解码器上进行联合译码, 这种多对一的结构使分布式信源编码方法在无线传感器网络中应用具有先天的优势。在无线传感器网络中, 传感器节点采集到的数据最后统一发送到一个接收节点(在基于簇的路由协议中, 该节点称为簇头), 这为实现联合译码提供了可能。

基金项目: 国家自然科学基金资助项目(60502036)

作者简介: 叶金昱(1978—), 男, 博士研究生, 主研方向: 无线传感器网络, 分布式信源编码; 刘 雨, 讲师、博士; 张 琳, 副教授、博士

收稿日期: 2008-07-20 **E-mail:** iceburgue@gmail.com

在基于簇的无线传感器网络中，簇内节点的欧氏距都相对较近，因此，簇内节点采集到的数据具有较强的相关性非常适合进行压缩。本文讨论如何将分布式信源编码方法与基于簇的路由协议相结合以提高无线传感器网络的能量有效性。

本文采用的分布式信源构造方法来源于文献[5]，它利用卷积码的校验矩阵所生成的校验子来辅助解码，由于校验子的长度通常要远小于所校验的数据的长度，因此达到了压缩须传送数据的目的。

2.1 问题描述

分布式信源编码系统结构如图 2 所示。设 X 和 Y 为具有相关性的无记忆连续随机变量，分别用独立同分布 (i, i, d) 的序列 $\{X_i\}_{i=1}^{\infty}$ 和 $\{Y_i\}_{i=1}^{\infty}$ 来表示，且 X 和 Y 之间的相关性表现为 Y 为 X 的有噪版本，即 $Y_i = X_i + N_i$ 。其中， N 为独立于 X 的独立同分布的连续高斯随机变量， X 和 N 的均值为 0。方差之比设为相关性(SNR)，即 $C_{SNR} = \sigma_X^2 / \sigma_N^2$ 。

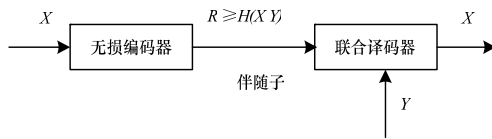


图 2 分布式信源编码系统结构

连续信源的分布式信源编码的目标为：在解码器可知边信息 Y 时，给定 X 的传输速率为每符号 R bit，最佳估计 X 。或在失真 $E[\rho(X, \hat{X})]$ 小于给定值 D 的条件下，最小化 X 的传输速率 R 。编码器实现从信源的输出空间到码字序号集的映射： $R^L \rightarrow \{1, 2, \dots, 2^{LR}\}$ ，译码器实现从码字序号集和相关的 Y 序列的乘积空间到 L 序列的映射： $\{1, 2, \dots, 2^{LR}\} \times R^L \rightarrow R^L$ 。本文主要关注均方误差(MSE)失真： $\rho(x, \hat{x}) = (x - \hat{x})^2$ 。

2.2 构造方法

采用 V 级固定速率(lbV)的 Lloyd-Max 标量量化器，对信源 X 的输出结果 $\{X_i\}_{i=1}^L$ 逐个量化，并将量化后的结果转化为二进制表示，则信源码书为 $S = \nabla^L$ 。

在空间 ∇^L 上进行有记忆的陪集构建。空间 ∇^L 共有 2^{3L} 个不同的序列，须在序列空间 ∇^L 上划分陪集，要求每个陪集中任意 2 个码字序列之间的最小距离越大越好，同时保持陪集之间的对称性。采用基于卷积码和 Trellis 码的标量量化(TCSQ)^[6]进行陪集的划分。

考虑建立在空间 ∇^L 上的系统 Trellis 码，分组长度为 L ，信息速率为 $R_c = 2^L$ bit/符号，共有 2^{2L} 个信息序列，使用码率为 $2/3$ 的卷积码进行编码。为使同一陪集中的码字的码间距最大，且保证不同陪集之间的对称性，采用 Ungerboeck 的四状态 Trellis 码^[7]。取 $(3, 2, 2)$ 系统 Trellis 码的状态机生成的所有序列的集合作为信道码 C ，显然 $C \subset S$ 。信道码 C 的陪集 $C(s)$ 是由所有满足 $H(D)Q^{-1}(z) = s^T$ ， $(z \in \nabla^L)$ 组成，其中 $s \in \{0, 1\}^L$ ， $Q^{-1}(z)$ 为 z 的二进制表示， $H(D)$ 为校验矩阵多项式。则 C 有 2^L 个陪集，每个陪集中包含 2^{2L} 个序列。

编码端需要计算量化后的信源序列对应的陪集 $C(s)$ 的伴随式 s ，其长度为 L bit。设 $x(D)$ 为多项式表示的量化后的信源序列，即 $x(D) = \sum_{i=1}^L x_i D^{i-1}$ ，其中， x_i 表示了 x 的第 i 个元素。编码器发送伴随式序列 $s(D)^T = H(D)x(D)^T$ 到译码器。以 $X = (7, 3, 2, 1, 5)$ 为例，得到的伴随式序列为 $s(1, 0, 1, 1, 0)$ 。

2.3 译码方法

利用伴随式形成对量化空间 ∇^L 的陪集划分，编码器只须将伴随式 s 发送到译码端， R 为伴随式的传输速率。译码器利用修改的 Viterbi 译码算法，结合边信息 Y 序列和伴随式 s 进行译码。即将 Y 作为 Viterbi 译码器的输入信号，在第 k 步时，根据 s_k 修改对应的码序列，即将译出的信道码序列设置为 $\xi = y'_i \oplus a_k$ ，其中， $y'_i \in \nabla$ ，对应于系统 Trellis 码格图上的边标签， $a_k = [0|0|s_k]^T$ 。这样，从 $k=1$ 到 $k=L$ ，在 Viterbi 译码中的每一步都需要重新标记边标签。设 x' 为采用这样的机制通过重新标记标签得到的与 Y 最接近的序列，即为译码器的解码输出序列。

3 分布式信源编码方法与基于簇的路由协议

本文以 LEACH-C 协议^[2]为例介绍分布式信源编码方法的具体实现过程。LEACH-C 运行过程分为 2 个阶段：簇的建立阶段和稳定运行阶段。

(1)簇的建立阶段。*sink* 节点收集传感器节点的位置信息和能量信息。根据这些信息 *sink* 节点将传感器节点按地理位置的远近分成若干的簇并且选出合适的簇头节点，然后将成簇信息以广播的方式发送到网络中。节点收到该信息后根据信息的内容决定其所属簇。簇头节点根据此信息判定其簇头地位。基本簇划分形成后，簇头所辖节点分配 TDMA 时间段。簇内节点在自己的时间段内向簇头传递数据，尽可能避免簇内干扰。簇头之间干扰的避免一般采用码分多址的方式。

(2)簇的稳定运行阶段。在簇的稳定运行阶段，当节点采集到监控信息后在自己的 TDMA 时间段向簇头发送该信息。簇头收集到所有簇内节点的信息后，它将这些信息进行一定的融合然后发送给 *sink* 节点。

在本文方法中，每一次簇内节点采集完信息后并非将完整的信息发送给簇头，而是发送一个压缩的版本，利用上文介绍的内容，此处发送的是该信息的校验子。簇头收到校验子后根据它所采集到的信息进行译码恢复。由于簇内节点的地理位置较近，因此它们所采集到的信息具有很强的相关性，在簇头可以以较高的概率来获得原始的信息。

4 仿真设计与结果分析

本文采用与文献[2]相同的能量消耗模型。但对计算造成的能量消耗进行不同的处理，由于编码简单，因此不计入损耗，但解码比较复杂，为 20 pJ/bit/符号。

实验参数见表 1。

表 1 实验参数

参量	值
仿真区域大小	100 m ²
节点数目	100 个
Sink 节点位置	(50, 175)
簇头个数	5 个
节点初始能量	2 J
无线电消耗能量	50 nJ/bit
传输放大器能量消耗(自由空间)	10 pJ/bit/m ²
传输放大器能量消耗(两路)	0.001 3 pJ/bit/m ⁴
广播数据包大小	25 Byte
TDMA 设定数据包大小	44 Byte
数据包头大小	25 Byte
数据大小(未压缩)	300 Byte
解码能量消耗	20 pJ/bit

根据上文介绍的例子对信源的压缩比为 1:3, 因此, 数据包的大小为 $25+300 \times 1/3=125$ Byte。

图 3、图 4 分别比较了在不同的能量消耗计算方式下分布式信源编码方法对能量消耗的影响。

图 3 显示在某一轮中每个节点的能量消耗, 采用分布式信源编码和不采用分布式信源编码每个节点消耗的能量相差不多, 分布式信源编码方法的影响并不突出。图 4 显示在某一轮时每个节点所累计消耗的能量数, 由于多轮累积的结果信源压缩, 采用分布式信源编码方法可节省 25%~30% 的能量。

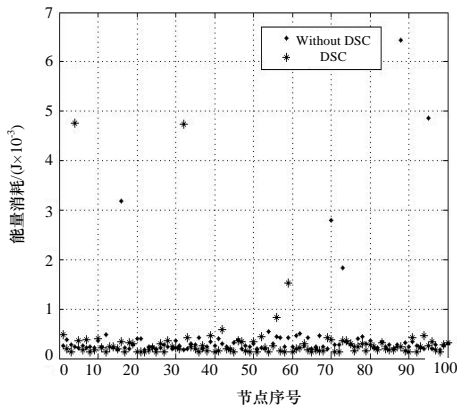


图 3 一轮中每个节点的能量消耗

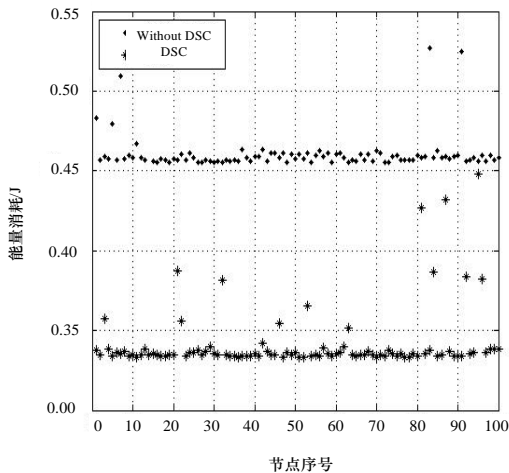


图 4 一轮中每个节点累计的总消耗

图 5 从整体的角度来比较分布式信源编码对能量的影响。可见, 采用分布式信源编码方法延长了近 40% 的网络生存期, 并且节点的死亡时间也比较集中。

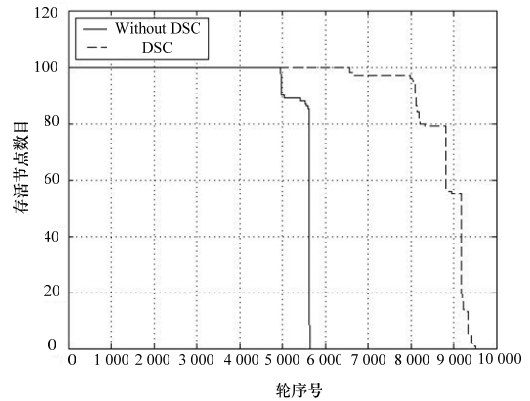


图 5 节点存活数目与轮数关系

5 结束语

本文利用分布式信源编码技术帮助传感器节点压缩它所发送的数据, 从而节省了无线传感器网络协议设计时传输信息的能量消耗。

在基于簇的传感器网络中, 簇内节点在地理位置上相对较近, 它们采集的数据具有很强的相关性。本文根据该特性将分布式信源编码方法与基于簇的路由协议有机结合, 对簇内传输的信息进行压缩, 从而节省了传感器节点的能量消耗。仿真结果显示, 该策略可有效地延长传感器节点的寿命, 进而延长网络的生存期。

参考文献

- [1] Akyildiz I F, Weilian S, Sankarasubramaniam Y, et al. A Survey on Sensor Networks[J]. IEEE Communications Magazine, 2002, 40(8): 102-114.
- [2] Heinzelman W. Application Specific Protocol Architectures for Wireless Networks[D]. Cambridge, MA, USA: Massachusetts Institute of Technology, 2000.
- [3] Slepian D, Wolf J K. Noiseless Coding of Correlated Information Sources[J]. IEEE Transactions on Information Theory, 1973, 19(7): 471- 480.
- [4] Wyner A, Ziv J. The Rate-distortion Function for Source Coding with Side Information at the Decoder[J]. IEEE Transactions on Information Theory, 1976, 22(1): 1-10.
- [5] Pradhan S, Ramchandran K. Distributed Source Coding Using Syndromes(DISCUS): Design and Construction[J]. IEEE Transactions on Information Theory, 2003, 49(3): 626-643.
- [6] Marcellin M W, Fischer T R. Trellis Coded Quantization of Memoryless and Gauss-markov Sources[J]. IEEE Transactions on Communications, 1990, 38(1): 82-93.
- [7] Ungerboeck G. Channel Coding with Multilevel/Phase Signals[J]. IEEE Transactions on Information Theory, 1982, IT-28(1): 55-67.

编辑 金胡考

(上接第 13 页)

- [2] Lee Nam-Kyu, Yang Sung-Bong, Lee Kyoung-Woo. Efficient Parity Placement Schemes for Tolerating up to Two Disk Failures in Disk Arrays[J]. J. of Systems Architecture, 2000, 46(15): 1383-1402.
- [3] Corbett P, English R, Goel A, et al. Row-diagonal Parity for Double Disk Failure Correction[C]//Proc. of FAST'04. [S. l.]: IEEE Press, 2004.
- [4] Xu Lihao, Bruck J. X-Code: MDS Array Codes with Optimal

- Encoding[J]. IEEE Trans. on Information Theory, 1999, 45(1): 272-276.
- [5] Wiebalck A, Breuer P T, Lindenstruth V, et al. Fault-tolerant Distributed Mass Storage for LHC Computing[C]//Proc. of CCGRID'03. [S. l.]: IEEE Press, 2003.

编辑 张帆