

无线传感器网络中的分布式信源编码的 仿真研究^①

A Simulation Study on Distributed Source Coding for Wireless Sensor Networks

王 兰 (桂林电子科技大学 信息与通信学院 广西 桂林 541004)

唐智灵 (桂林电子科技大学 电子工程学院 广西 桂林 541004)

摘 要: 分析了无线传感器网络中的分布式编码方法, 基于 Slepian-Wolf 理论和 Wyner-Ziv 理论, 在 Turbo 编码方法的基础上提出了一种 DSC-Turbo 编译码的改进方法。通过 MATLAB 仿真验证了该方法在满足一定可靠性的前提下, 能够有效的压缩信源速率。进而证明了在无线传感器网络中应用分布式信源编码理论可以减少部分节点的传输数据量, 提高无线传感器网络能量效率。

关键词: 无线传感器网络 能量效率 分布式信源编码 相关性 Turbo 码

1 引言

无线传感器网络 (WSN, Wireless Sensor Network) 中大部分节点的能量靠电池提供, 其能量有限, 且在使用过程中不便于更换电池, 所以传感器节点的能量是制约 WSN 工作寿命的瓶颈。此外, 网络需要成千上万的节点以保证可用性和生存能力, 结果是网络的节点密度很高, 会导致发生多个节点采集到相同类型且相关性很强的数据的现象。由于在 WSN 中, 通信模块是能量消耗的主要部分, 节点传输信息比进行计算处理要消耗更多的电能, 例如传输 1 比特信息所需要的电能大约可以执行 3000 条计算指令^[1]。因此, 为减少网络通信流量而利用上述传感器节点数据的相关性, 在满足应用要求的前提下, 对数据进行压缩就成为 WSN 研究中降低能量消耗的一个途径。本文讨论的分布式信源编码(DSC, Distributed Source Coding)是 WSN 中信息压缩的一种主要方法。

2 分布式信源编码的原理

分布式编码理论是在 Slepian 和 Wolf 提出的信息

理论上建立的。DSC 是指对相互之间不能进行通信的多个传感器的相关输出结果分别进行压缩 (独立编码), 并将压缩后的结果发送到同一中心节点进行联合译码的信源编码方法。与传统的编码方法类似, 分布式编码的原理也有两种形式: 无损分布式编码的 Slepian-Wolf 理论^[2]和使用解码端辅助信息的有损分布式编码 Wyner-Ziv 理论^[3]。

2.1 Slepian-Wolf 编码理论

Slepian-Wolf 编码适用于离散信源的分布式信源编码。假定 X 和 Y 是两个互相关的离散无记忆信源, 使用做参考信息来无损的压缩 X 。如果在编码端和解码端都可以得到参考信息 Y , 那么根据香农信息理论我们知道 X 理论上的无损压缩极限是已知 Y 的情况下 X 的条件熵 $H(X|Y)$ 。对于仅可在解码端获得参考信息的情况下, Slepian-Wolf 理论证明了 X 的压缩极限仍然是 $H(X|Y)$, 与在编码端可获得参考信息的情况下所能取得的编码效率是一样的。因此只需知道 X 和 Y 的联合概率分布, 编码器不需要得到参考信息就能够取得和已知参考信息一样的编码效率。根据

① 基金项目:广西科学研究与技术开发计划应用基础专项项目(桂科基 0731019)

收稿时间:2008-08-19

Slepian-Wolf 理论，当下面条件满足时：

$$R_x > H(X|Y), R_y > H(Y|X), R_x + R_y > H(X,Y),$$

互相关的两个信源 X 、 Y 可以分别独立的以 R_x 和 R_y 的编码速率编码，使得在解码端可以以任意小的错误概率来联合解码。

2.2 Wyner-Ziv 编码理论

Winer-Ziv 编码是指对具有相关性的连续信源进行分布式编码。首先需要对信源进行量化，得到量化码字 W ，引入了量化失真。量化后的 W 与边信息 Y 之间仍具有相关性，应用 Winer-Ziv 编码来降低 W 的编码速率。图 1 描述了基本的 Winer-Ziv 编码的框图。

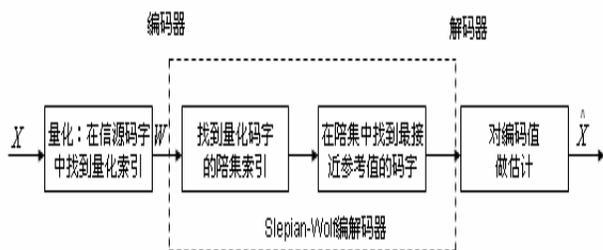


图 1 Winer-Ziv 编解码器

3 编码实现方法

目前大多数的 DSC 方法都源自于信道编码的思想(但最终目的不同：信道编码的目的是纠错，而 DSC 的目的是对信源数据进行有效的压缩)。信道编码的过程是在源数据流中加插一些码元(又称校验位或者冗余码位)，从而达到在接收端进行判错和纠错的目的。在 WSN 中应用 DSC 的基本思想是：将传感器节点数据之间的冗余看成是被加插的码元，通过 DSC 系统对节点数据进行有效的压缩(编码)，达到提高数据传输效率、节约节点能量的目的。

现有的 DSC 方法包括由 Pradhan 和 Ramchandran 提出的使用伴随式的分布式信源编码(DISCUS, Distributed Source Coding Using Syndrome)方法、使用 Turbo 码的分布式编码方法、使用低密度奇偶校验码(LDPC, Low Density Parity Code)的分布式编码方法等。

DISCUS 方案使用网格编码调制的信道码来分割信源符号空间，构建出不同的陪集，发送各个信源符号所在陪集序号作为编码信息发送到信道中。接收端根据收到的陪集索引信息，结合与当前译码信源相关的信源的译码结果，在陪集中找到最佳的译码结果作

为输出。DISCUS 考虑了两个互相关的信源之间的相关性，把一个信源看成另一个信源经过了噪声干扰的结果，利用信道编码的研究成果实现分布式信源编码。

Turbo 码已应用于分布式信源编码中^[4]。在 Turbo 码的设计中，交织器是最重要的部分。通常交织器长度越长，纠错性能越好，但编解码复杂度也随着增加。LDPC 码是一类可以用非常稀疏的校验矩阵或者二分图定义的线性纠错码，目前也被用于分布式编码中^[5]。

本文从 Turbo 码的设计角度出发，提出 Turbo 码的编译码改进结构，实现一种 DSC-Turbo 编译码方案，如图 2 所示。图中可以用虚竖线隔成三部分，分别是编码、网络信道和译码。Turbo 码编码器分成三部分：RSC 编码器，交织器和删余矩阵；译码采用最大后验概率译码(MAP)。下面讨论如何按 DSC 的要求对这几部分进行优化设计，构造 DSC-Turbo 编译码结构。

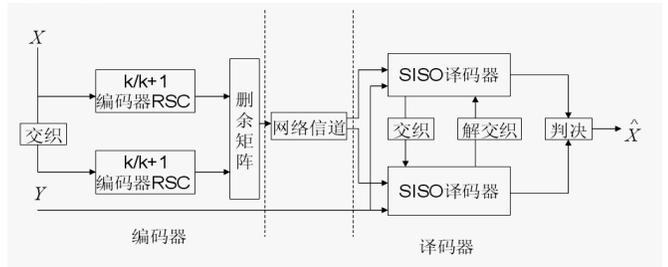


图 2 DSC-Turbo 编译码框图

3.1 编码器的设计

传统的 Turbo 码编码器结构如图 3 所示。信息序列 u 经过交织器，形成一个新的序列 u_1 (长度与内容没变，但比特位置经过重新排列)。 u 和 u_1 分别传送到两个分量编码器(RSC1 和 RSC2)。一般情况下，这两个分量编码器结构相同，生成序列 X^{p1} 和 X^{p2} 。为了提高码率，序列 X^{p1} 与 X^{p2} 需经过删余器，形成校验位序列 X^p 。 X^p 与未编码序列 X^s 经过复用调制，生成了 Turbo 码序列 X 。

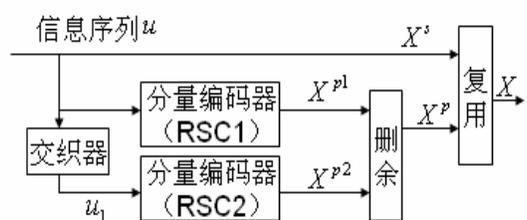


图 3 传统 Turbo 码编码器结构图

与传统的 Turbo 码编码器相比, 在 DSC-Turbo 编码器设计中, 我们将信源用传统的信源编码方法进行编码, 传输到解码端, 并在解码端作为参考信息。将信源 x 送入到两个结构相同的 $k/k+1$ RSC 分量编码器中, 即每个编码周期进入 RSC 的数据为 k 比特, 经过编码后产生 1 比特的校验信息与原始的 k 比特信息组成编码输出。由于边信息存在, 丢弃原始的 k 比特信息, 只需要 1 比特的校验信息。经过删余器, 两个 RSC 交替保留校验比特, 最后得到的码率为 $k/1$, 即将信源速率压缩到原来的 $1/k$ 。

3.2 译码器的设计

典型的移动通信环境中, 所使用的信道码编码方式为 $1/k, (k=2,3,4,...)$ 的 Turbo 码, 对应 1 比特的输入信息, 产生若干比特的校验信息, 同时发送到接收端进行译码, 因此采用二进制译码算法, 每个译码周期的判决输出为一个二进制比特。而在本文提出的 DSC-Turbo 编译码中与此刚好相反, 其码率为 $k/1$, 若干比特位的输入信息进行编码后只产生 1 比特的校验信息作为编码输出发送到译码端, 因此需要采用非二进制译码, 每个译码周期的判决输出为一个非二进制数据, 共 2^k 个备选符号。

由于 DSC 中接收端可以是性能很强且没有限制的基站设备, 因此译码可以使用复杂度比较高但性能很好的软输入软输出(SISO)MAP, LogMAP,SOVA 等 Turbo 码译码算法, 与以往的 DSC-Turbo 方案不同, 本文提出的设计方案中首次采用了 MAP 译码算法。在非二进制 MAP 算法推导中令

$$L(u_k) = p(u_k = i | y_1^N)$$

其中, $i=0,1,2,...,2^k-1$, 根据似然函数的大小关系进行判决, 似然函数值最大时对应的即为判决结果。

4 仿真分析

设信源 x 和边信息 y 的相关性可以用二进制对称信道来建模, 且信道的转移概率为 p , 即

$$H(X|Y) = H(Y|X) = h(p)$$

其中, $h(g)$ 是二进制熵函数。通过改变参数 p 的值来模拟 x 和 y 的相关性, 从而比较不同的条件熵下的性能。

假定虚拟信道为理想信道, 即校验位能够正确接

收。我们以和之间的相关性 $H(X|Y)$ 为横坐标, 以误比特率为纵坐标, 通过改变的值来说明误比特率与信源相关性之间的关系, 通过 MATLAB 对 1200 帧随机信源数据进行编译码仿真得到图 4, 其中 $L=500$ 、1000 表示帧长度, 图 4 中 data(1-4) 四条曲线分别表示未编码、迭代 1 次、迭代 3 次和迭代 10 次四种情况。

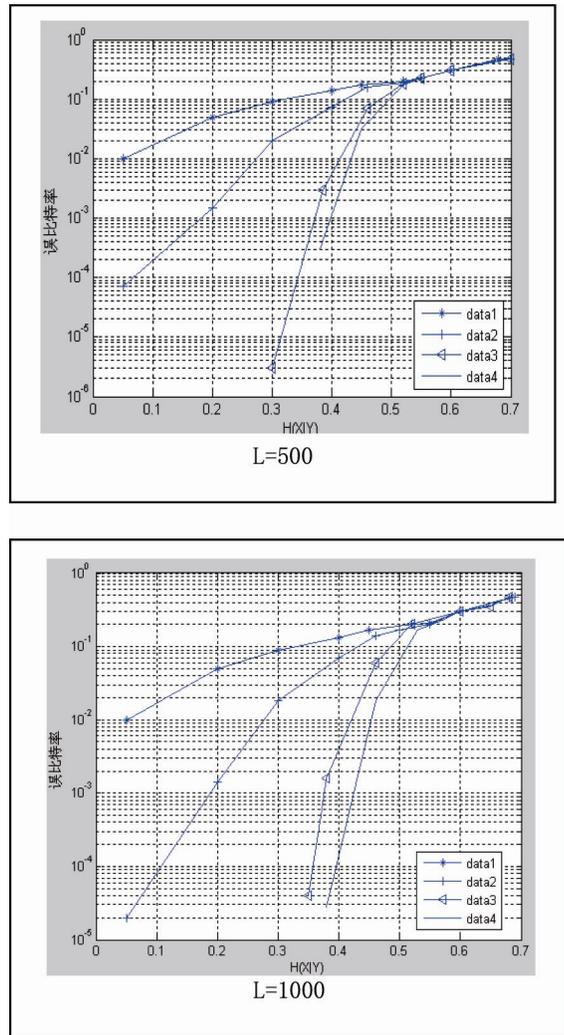


图 4 二进制对称信道误比特率性能

从图 4 可以看出, 相同条件下, 帧长度越大, 误码率性能越好; 迭代译码的迭代次数越多, 性能也越好。帧长度的增加意味着编码复杂度的增加, 迭代次数的增加对应译码复杂度的增加, 为了取得比较好的编译码效果以及尽量减少编译码复杂度, 两者需要适当的选择。当分量器编码速率为 $R_x=1/2$ 时, $H(X|Y) \leq R_x$ 的范围内, 误比特率曲线迅速下降, 当迭

代次数达到 10 次以上时, 可以认为达到了数据传输所要求的误差水平。也就是说, 在满足一定可靠性的条件下, 信源 X 的数据压缩率可以达 $1/k$, 网络能量效率可以提高 $k+1/2k$ 。根据 DSC 理论推导证明, 在 $H(X|Y) \leq R_x$ 时, 是可以实现无差错传输的, 仿真结果表明本文提出的方案已经逼近了这一理论界限。

从仿真结果可以看出, DSC 编译码可以显著减小信源发送的数据量, 在 DSC 译码过程中, 边信息和接收到的校验序列所起到的作用不同, 边信息作为辅助信息, 是必不可少的, 但是允许一定的失真; 而校验信息作为译码过程的主要处理对象, 具有决定性的作用。因此, 目前在 WSN 中应用 DSC 时, 传送相关消息的节点需要进行两个或者三个的合作, 其中一个节点提供边信息, 另一个节点根据 Slepian-Wolf 界或 Wyner-Ziv 界压缩信息。

5 结束语

本文针对无线传感器网络的应用前景以及低复杂度、低能耗、高压缩率的要求, 提出了一个完整的使用 Turbo 码构造 DSC 编译码过程的方案。通过对 Turbo 码分量编码器、交织器、迭代译码三部分的重新设计, 使其能够应用于 DSC 这一特定的场合。从仿真结果来看, 这一编译码方案验证了 DSC 理论推导的结果, 性能上完全可以胜任实际应用的要求。这一方案把复杂度由编码端转移到译码端, 实现了编码器简化的目标。

参考文献

- 1 Pottie GJ, Kaiser WJ. Wireless integrated network sensors. *Communications of the ACM*, 2000,43(5):51-58.
- 2 Slepian D, Wolf JK. Noiseless coding of correlated information sources. *IEEE Trans. Inform. Theory*, 1973, 19(4):471-480.
- 3 Wyner JZ. The rate-distortion function for source coding with side information at the decoder. *IEEE Trans, Inform. Theory*, 1976,22(2):1-10.
- 4 Liveris A, Xiong Z, Georghiadis C. Distributed compression of binary sources using conventional parallel and serial concatenated convolutional codes. *Proceedings of the Data Compression Conference*. Piscataway, USA: IEEE, 2003:193-202.
- 5 Tian T, Garcia-Frias J, Zhong W. Compression of correlated sources using LDPC codes. *Proceedings of the Data Compression Conference*. Piscataway, USA: IEEE, 2003:450.
- 6 Xiong ZX, Liveris AD, Cheng S. Distributed Source Coding for Sensor Networks. Texas A&M University, College Station, 2004.