

基于比特可靠性的 LDPC 码编译码算法^①

冯 军^② 周武旻^③

(中国科学技术大学电子工程与信息科学系个人通信与扩频实验室 合肥 230027)

摘 要 提出了一种基于比特可靠性的低密度奇偶校验(LDPC)码编码算法和一种加权置信传播(BP)译码算法。该编码算法首先利用蒙特卡罗仿真得到 LDPC 码各个比特节点的出错概率,然后用已知信息替换易出错比特节点进行编码;该译码算法根据比特节点可靠性的差异,在译码时为每个比特节点赋予相应的权值,以调整它们对译码的影响程度。仿真表明,新的编译码算法使得系统性能大大提高,同时加快了译码迭代收敛速度。

关键词 低密度码, 比特可靠性, 已知信息, 加权译码

0 引 言

1962 年, Gallager 首次提出低密度奇偶校验(low density parity check, LDPC)码^[1], 证明了它的距离特性, 并给出一种基于概率的迭代译码算法。后来的研究证明 LDPC 码的纠错能力极强, 其性能接近香农极限, 不过由于运算复杂度太大, 逐渐被人们淡忘。1996 年 Mackay 和 Neal 重新开启了对 LDPC 码的研究^[2], 并提出了置信传播(belief propagation, BP)译码算法, Luby 等人基于非正则双向 Tanner 图提出了性能更好的非正则 LDPC 码^[3], Richardson 等人在加性高斯白噪声(AWGN)信道条件下设计出的最优非正则 LDPC 码的性能距香农极限仅 0.0045dB^[4]。

LDPC 码的性能与编码矩阵的结构有很大关系, 由于受到编码矩阵中小环以及列重等诸多因素的影响, LDPC 码的每个比特节点获得的保护程度会有差别, 致使它们在译码过程中提供的概率信息可靠性也不一样。常规的 LDPC 编译码方法并没有考虑这一点, 有的改进方案^[5]虽然利用了列重的非均匀保护特性, 但忽略了小环等因素的影响, 性能提高并不明显。基于此, 本文提出了一种插入已知信息的 LDPC 编码方法, 用于抑制少数易出错比特节点对译码性能的不利影响, 同时又提出了一种加权 BP 译码算法, 该方法通过引入权值来调控各个比特节点的概率信息对译码的影响, 有效地提高了系统性能。

1 比特节点的可靠性差异特性

LDPC 码可以由一个稀疏的校验矩阵来定义, 校

验矩阵的行对应校验节点, 列对应比特节点。每个校验矩阵对应一张 Tanner 图, 在 Tanner 图中可以发现, 所有比特节点都处于不同数量和长度的环当中。对于采用 BP 译码算法的 LDPC 码, 不同长度的环对比特节点有不同的影响。环周长越小, 与此环相关联的比特节点越难被纠错, 因为处于小环中的比特节点与其它节点的连通性变差, 一旦出错将更难在译码中从其它节点获得正确信息进行纠错, 并且自身的错误信息还会在数次迭代之后反作用于自己以及同处于此小环中的其它比特节点。另外, 大量小环的存在还会降低 LDPC 码字的汉明距, 公式

$$d_{\min} = \begin{cases} 1 + \frac{c}{c-2}((c-1)^{(g-2)/4} - 1) & (\text{若 } g/2 \text{ 为奇数}) \\ 1 + \frac{c}{c-2}((c-1)^{(g-2)/4} - 1) + (c-1)^{(g-2)/4} & (\text{若 } g/2 \text{ 为偶数}) \end{cases} \quad (1)$$

给出了正则 LDPC 码的最小环周长(girth)与码字汉明距之间的关系^[6]。其中 c 为校验矩阵列重, g 为校验矩阵 girth 大小, $d_{\min}$ 为码字汉明距。据此关系可以看出, girth 越小使得 LDPC 码字的汉明距也变得越小, 致使 LDPC 码整体性能越差。

导致 LDPC 码各比特节点可靠性差异的另外一个重要因素是校验矩阵的列重不均匀分布^[7], 列重大的比特节点能够在译码过程中从更多校验节点那里获得外信息, 因此更容易被纠错; 列重小的比特节点由于参与的校验方程数少, 译码时能获得的外信

① 新世纪优秀人才计划(NCET)、国家自然科学基金(60496314)和安徽省自然科学基金(070412044)资助项目。

② 男, 1982 年生, 硕士生; 研究方向: 信道编解码技术。

③ 通讯作者, E-mail: wyzhou@ustc.edu.cn
(收稿日期: 2007-09-13)

息也更少,纠错更困难。因此在非正则 LDPC 码中,列重不同的比特节点受到的保护程度也不一样。

如何估计 LDPC 码各比特节点之间的差异性仍是一个问题,由于它受小环、列重等多种因素影响,所以需要综合考虑各项因素。密度进化^[8]和高斯近似^[9]都是分析 LDPC 码的有效工具,通过跟踪译码过程中比特节点和校验节点概率密度分布的变化,不仅能估计译码门限,还能估计译码的迭代收敛性以及每个比特节点的出错概率,遗憾的是这两种方法都假设码长无限长,和实际有限码长情况下的结果有较大差距。联合高斯近似分析方法适于有限码长情况,它是在高斯近似的基础上,将节点向量的概率分布看作联合高斯分布,通过计算节点向量的协方差矩阵以及均值向量,得到节点向量的联合高斯分布概率密度函数,进而估计出每个比特节点在任意次译码迭代后的出错概率。不过当校验矩阵很大时,联合高斯近似分析方法在计算协方差矩阵以及高维向量的概率密度函数时会很复杂,所以本文选择更加简捷的蒙特卡罗仿真方法来近似估计各比特节点的可靠性差异。根据蒙特卡罗仿真方法,通过大数据量的仿真实验来统计每个比特节点的出错概率,出错概率的大小体现了各比特节点可靠性的高低,低出错概率的比特明显具有更高的可靠性。统计表明,蒙特卡罗仿真的结果和采用联合高斯近似分析方法的结果十分接近。

2 插入已知信息的 LDPC 编码方法

LDPC 码校验矩阵中各比特节点之间存在差异,极少数受小环和列重影响严重的易出错比特节点错误率明显偏高,它们的存在严重影响了 LDPC 码的性能。于是,本文提出在易出错比特节点位置用已知信息代替进行编码,接收机也利用该已知信息进行译码。

常规 LDPC 码编码时,先用校验矩阵 H 构造生成矩阵 G ,再用信源序列 u 乘以 G 得到码字 c ,如式

$$c = u \times G = [u, u \times P] \quad (2)$$

所示。在矩阵系统形式下, H 经过高斯变换可表示为 $H = [P^T, I]$, G 表示为 $G = [I, P]$,其中 I 为单位阵。

在插入已知信息的 LDPC 码编码中,先利用蒙特卡罗仿真方法确定原校验矩阵中易出错比特节点的位置。为确保已知信息插在易出错比特节点位置,需要对原校验矩阵作适当的列变换,将易出错比

特节点调整至信源序列的后面,然后用已知信息代替易出错比特节点,并联合信源序列进行编码。在矩阵的系统形式下,码字中的信源序列、已知信息序列和校验序列的位置关系如式

$$\begin{aligned} c &= [u, k] \times G = [u, k] \times [I, P] \\ &= [u, k, (u, k) \times P] \end{aligned} \quad (3)$$

所示,码字最前面的部分对应信源序列 u ,其后是已知信息序列 k ,最后是校验序列 $(u, k) \times P$ 。

由于已知信息序列 k 在编译码端都已知,所以不需要送入传输信道,最终传输数据按照式

$$c' = [u, (u, k) \times P] \quad (4)$$

确定。

译码采用性能最优的 BP 译码算法,先用接收数据和信道信息计算每个比特节点的初始概率信息,然后在比特节点和校验节点之间作信息迭代。与常规 BP 译码过程不同的是,易出错比特节点的初始概率信息需要特殊考虑,由于编码时用于代替易出错比特节点的已知信息并没有被传输,所以接收机没有与之对应的接收数据,在计算易出错比特节点的初始概率信息时,只需要直接利用预先规定的已知信息即可。

为便于描述译码算法,先作必要的符号说明。 R_j 表示参与校验方程 j 的所有比特节点位置的集合, C_i 表示第 i 个比特节点参与的所有校验方程的集合, R_j 中除去比特节点 i 的集合表示为 $R_j \setminus i$, C_i 中除去校验方程 j 的集合表示为 $C_i \setminus j$, $L(c_i)$ 表示第 i 个比特节点的初始概率信息, $L(q_{ij})$ 表示第 i 个比特节点对第 j 个校验方程的概率信息贡献, $L(r_{ji})$ 表示第 j 个校验方程对第 i 个比特节点的概率信息贡献, $L(Q_i)$ 表示译码后第 i 个比特节点新的概率信息,对数域的具体译码过程如下。

第一步:比特节点概率信息初始化:

(1)对于信源比特 u 和校验比特 $(u, k) \times P$,在 AWGN 信道条件下,利用信道信息 σ^2 和接收数据 y_i ,通过公式

$$\begin{aligned} L(q_{ij}) &= L(c_i) = \log \frac{p_r(c_i = 0 | y_i)}{p_r(c_i = 1 | y_i)} \\ &= \log \frac{(1 + e^{-2y_i/\sigma^2})^{-1}}{(1 + e^{+2y_i/\sigma^2})^{-1}} = \frac{2y_i}{\sigma^2} \end{aligned} \quad (5)$$

计算第 i 个比特节点的初始概率信息,并对 $L(q_{ij})$ 初始化。

(2)对于易出错比特节点,直接利用已知信息 k 计算初始概率信息。假设已知信息为 0,根据式(5)的定义,它们的初始概率信息应该是无穷大,仿真时

可以取一个较大的值来表示。

第二步:第 i 个比特节点将得到的概率信息传递给它参与的所有校验节点,假设 $\alpha_{ij} = \text{sign}(L(q_{ij}))$ 表示该比特节点概率信息的符号, $\beta_{ij} = |L(q_{ij})|$ 表示其大小,则传递给第 j 个校验节点的概率信息表示为:

$$L(r_{ji}) = 2 \tanh^{-1} \left(\prod_{i' \in R_j \setminus i} \tanh \left(\frac{1}{2} \beta_{i'j} \right) \right) \times \prod_{i' \in R_j \setminus i} \alpha_{i'j} \quad (6)$$

第三步:第 j 个校验节点将得到的概率信息传递给参与它的所有比特节点,对于第 i 个比特节点表示为:

$$L(q_{ij}) = L(c_i) + \sum_{j \in C_i \setminus i} L(r_{ji}) \quad (7)$$

第四步:计算每个比特节点新的概率似然值,第 i 个比特节点的概率似然值表示为:

$$L(Q_i) = L(c_i) + \sum_{j \in C_i} L(r_{ji}) \quad (8)$$

第五步:判决,第 i 个比特节点判决为:

$$\begin{aligned} \hat{c}_i &= 1 & \text{若 } L(Q_i) < 0 \\ \hat{c}_i &= 0 & \text{若 } L(Q_i) \geq 0 \end{aligned} \quad (9)$$

在译码过程中,如果 $\hat{\mathbf{c}} \times \mathbf{H}^T = 0$ 或者迭代次数超过规定次数就停止译码,否则跳转至第二步,继续迭代过程。

由于易出错比特节点在编码端被已知信息代替,所以在计算译码初始概率信息时,这些原本容易出错的比特节点能够拥有正确的概率信息值,有效避免了它们对译码产生的不利影响。当然这种应用已知信息提高编码性能的方法会在一定程度上降低编码效率,不过通常易出错比特节点数目都占绝对少数,因此这种影响非常微小,这在后面的仿真结果中可以看到。考虑到信道估计需要的导频资源,如果利用导频信息来充当已知信息,那么实际上编码效率是可以不受任何影响的。

3 加权 BP 译码算法

由前述得知,受小环以及列重等因素的影响,LDPC 码不同的比特节点存在不同的出错概率,在迭代译码过程中,出错概率低的比特节点提供的概率信息可靠性高,而出错概率高的比特节点提供的概率信息可靠性则低。常规 BP 译码算法在处理比特节点到校验节点的概率信息传递时,并没有考虑这种不同比特节点提供的概率信息可靠性差异,而是同等对待来自不同比特节点的概率信息。本文提出一种加权 BP 译码算法,在译码过程中为每个比特

节点赋予一个权值,以调整各比特节点的概率信息对译码的影响程度。由于出错概率越高的比特节点提供的概率信息越不可靠,因而获得的权值要能削弱此概率信息的影响,相应出错概率越低的比特节点获得的权值要能增强它提供的概率信息的影响。

图 1 是加权 BP 译码算法的示意图,在计算校验节点概率信息时,从每个比特节点流入的概率信息都被赋予一个权值,若所有权值都为 1,就等效于标准的 BP 译码算法。

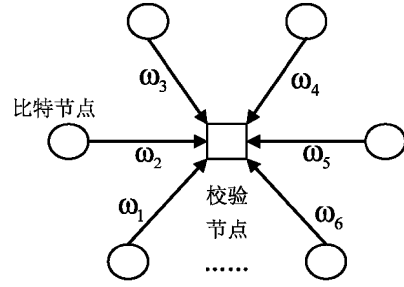


图 1 加权 BP 译码算法示意图

常规 BP 译码过程见第 2 节,这里只详细讨论加权 BP 译码算法如何在标准 BP 译码过程中加入权值。重写前面 BP 译码算法的第二步计算公式如下:

$$L(r_{ji}) = 2 \tanh^{-1} \left(\prod_{i' \in R_j \setminus i} \tanh \left(\frac{1}{2} \beta_{i'j} \right) \right) \times \prod_{i' \in R_j \setminus i} \alpha_{i'j} \quad (10)$$

利用函数 $\phi(x) = -\log \tanh \left(\frac{1}{2} x \right) = \log \left(\frac{e^x + 1}{e^x - 1} \right)$ 的性质 $\phi^{-1}(x) = \phi(x)$ 对(10)作变换,可以将 $L(r_{ji})$ 重新表示为:

$$L(r_{ji}) = \phi \left(\sum_{i' \in R_j \setminus i} \phi(\beta_{i'j}) \right) \times \prod_{i' \in R_j \setminus i} \alpha_{i'j} \quad (11)$$

$\phi(x)$ 的函数特性如图 2 所示。当 $x > 3$ 时,函数值趋于 0; x 趋于 0 时,函数值趋于无穷大。

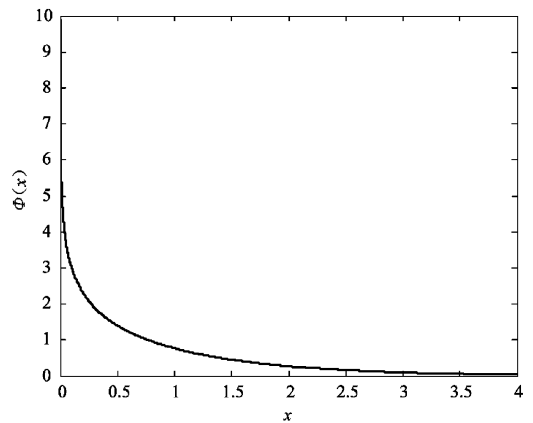


图 2 $\phi(x)$ 在 $x \geq 0$ 的函数特性

根据 $\phi(x)$ 的函数特性,给出的加权方案如式

$$L(r_{ji}) = \phi\left(\sum_{i \in R_j \setminus i} \phi(\omega_i \times \beta_{ij})\right) \times \prod_{i \in R_j \setminus i} \alpha_{ij} \quad (12)$$

所示。

在计算比特节点到校验节点的概率信息 $L(r_{ji})$ 时,为每个比特节点提供的概率信息大小 β_{ij} 附加一个权值 ω_i , 权值的大小由比特节点出错概率的高低决定。根据图 2 中 $\phi(x)$ 的函数特性,出错概率低的比特节点需要一个大的权值以增强其概率信息对译码进程的贡献,出错概率高的比特节点则需要一个小的权值以削弱它提供的概率信息对译码进程的影响。

为了计算权值,先利用蒙特卡罗仿真估计出每个比特节点的出错概率 e_i , 找出所有比特节点出错概率的最小值 $\min\{e_i\}$, 再将该最小值除以各比特节点的出错概率 e_i , 便得到每个比特节点的权值 ω_i , 如式

$$\omega_i = \frac{\min\{e_i\}}{e_i} \quad (13)$$

所示。如果遇到 $\min\{e_i\}$ 为 0 或者接近 0 的情况,可以取多个出错概率很小的比特节点的平均出错概率充当 $\min\{e_i\}$, 并把这些比特节点的权值设置为 1, 以避免其它节点的权值产生畸变。显然,出错概率越大的比特节点获得的权值越小。

4 仿真结果和分析

在 AWGN 信道,二进制相移键控(BPSK)调制方式下,对基于比特可靠性的 LDPC 编译码算法进行了仿真,仿真使用的低密度奇偶校验(LDPC)矩阵基于文献[6]的方法构造,译码最大迭代次数设为 30。

构造码长 625、码率 0.4 的 LDPC 码,校验矩阵中最小环周长为 6。在信噪比为 1.5dB 的条件下,仿真 20 万个数据块,得到该校验矩阵各比特节点的错误图样如图3所示。可以看出,各比特节点的出

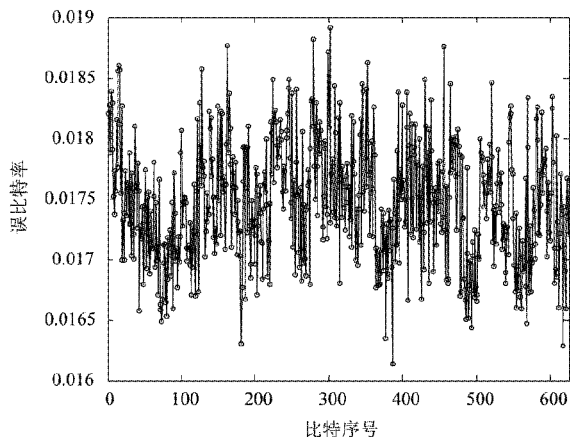


图3 校验矩阵各比特节点错误图样

错概率都不一样,其中少数比特节点的出错概率明显偏高,说明不同比特节点受到的保护程度有差异,可靠性有差别。

利用图 3 中各比特节点出错概率的差异,分别仿真了在编码端插入 25 比特已知信息和在译码端使用加权 BP 译码算法的情况,结果如图 4 和图 5 所示。

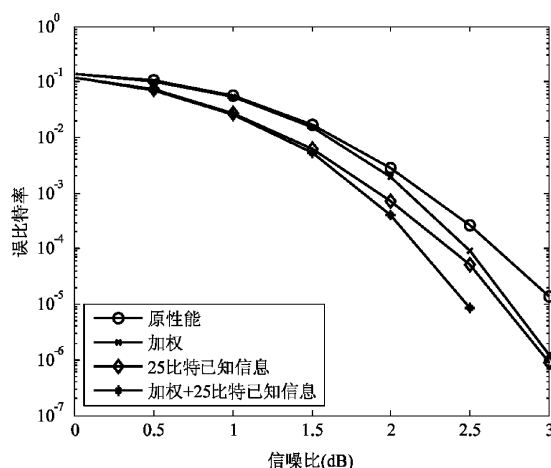


图4 已知信息和加权对译码性能的影响

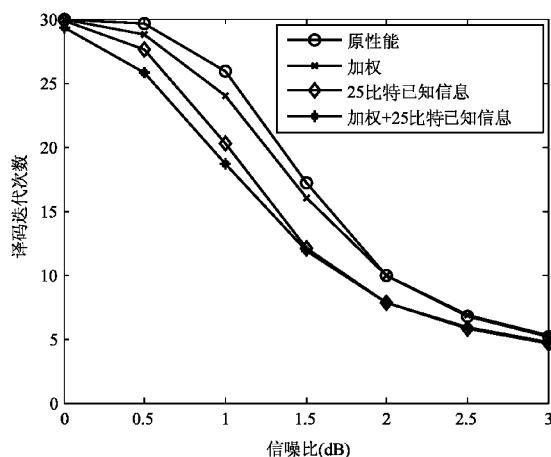


图5 已知信息和加权对迭代次数的影响

在 25 个出错概率明显偏高的易出错比特节点位置插入已知信息,有效抑制了它们对译码的不利影响,使得整体编译码性能有了较大提高,高信噪比区域约有 0.3 ~ 0.4dB 的性能增益,低信噪比区域约有 0.2dB 的性能改善,而编码效率只降低了 0.025。同时,已知信息的应用使得正确译码所需迭代次数也明显减少,信噪比为 1dB 时,平均正确译码迭代次数最大减少了 6 次。

加权 BP 译码算法同样带来了编译码性能的提高以及正确译码迭代次数的减少。在高信噪比区

域,加权处理能够带来约 0.25 ~ 0.3dB 的性能提升,和应用 25 比特已知信息编码方法的结果相差不大,不过在低信噪比区域,加权处理则没有明显效果,这或许与本文采用的简单权值计算方法有关。另外,在低信噪比区域,加权处理使得译码迭代次数有所降低,而高信噪比区域由于正确译码所需迭代次数变少,加权处理并不能进一步减少迭代次数。

图 4 和图 5 还给出了同时在编码端插入 25 比特已知信息并在译码端使用加权 BP 译码算法的仿真结果,相比原性能,新算法带来了最大约 0.6dB 的性能增益,同时让正确译码所需迭代次数降低了最大约 7 次。

5 结 论

基于 LDPC 码各比特节点可靠性不同的特性,本文提出了一种插入已知信息的 LDPC 编码方法,即在少数易出错比特节点位置用已知信息代替进行编码,同时提出一种加权 BP 译码算法,根据每个比特节点出错概率的高低,在译码时为它们提供的概率信息赋予相应的权值,出错概率越低的比特节点由于提供的概率信息更可靠,因此获得的权值越大,这样的加权处理可以有效调整各比特节点对译码进程的影响。这样的编译码算法可大大提高系统性能,并加快译码的迭代收敛速度。

参考文献

[1] Gallager R G. Low density parity check codes. *IRE Trans In-*

form-Theory, 1962, IT-8: 21-28

- [2] MacKay D J C, Neal R M. Near Shannon limit performance of low-density parity-check codes. *Electronics Letters*, 1996, 32: 1645-1646
- [3] Luby M G, Mitzenmacher M, Shokrollahi M A, et al. Improved low-density parity-check codes using irregular graphs and belief propagation. In: *Proceedings of the 1998 International Symposium on Information Theory*, Cambridge, Massachusetts, USA, 1998. 213
- [4] Sae Y C, Forney G D, Richardson T J, et al. On the design of low-density parity-check codes within 0.0045dB of the Shannon limit. *IEEE Communication Letter*, 2001, 2 (5):58-60
- [5] Yang X M, Yuan D F, Ma P M, et al. New research on unequal error protection (UEP) property of irregular LDPC codes. *IEEE Consumer Communications and Networking Conference*, 2004:361-363
- [6] Zhang F, Xu Y, Zhou W Y, et al. High girth LDPC codes construction based on combinatorial design. *IEEE Vehicular Technology Conference*, 2005:591-594
- [7] Richardson T J, Shokrollahi M A, Urbanke R L. Design of capacity-approaching irregular low-density parity-check codes. *IEEE Trans on Info Theory*, 2001, IT-47: 619-637
- [8] Richardson T J, Urbanke R L. The capacity of low-density parity-check codes under message-passing decoding. *IEEE Trans Inform Theory*, 2001, 47(2):599-618
- [9] Sae Y C, Richardson T J, Urbanke R L. Analysis of sum-product decoding of low-density parity-check codes using a Gaussian approximation. *IEEE Trans Inform Theory*, 2001, 47 (2):657-670

Reliability based encoding and decoding algorithms for LDPC codes

Feng Jun, Zhou Wuyang

(Personal Communication Network and Spread Spectrum Lab, University of Science and Technology of China, Hefei 230027)

Abstract

Based on the reliability of bit nodes, the paper proposes a novel encoding structure and a weighted belief propagation (BP) decoding algorithm for low density parity check (LDPC) codes. The former uses the Monte Carlo (MC) simulation to estimate the error probability of each bit node, and then uses prior information to replace the unreliable bits in the parity-check matrix when encoding. The latter, according to each bit node's error probability, gives each bit node a weight to adjust their contribution in the decoding process. The simulation results show that the proposed schemes can improve the coding performance remarkably, and can effectively decrease the iteration number which is needed for correctly decoding.

Key words: LDPC codes, bit reliability, prior information, weighted decoding