

有限长低密度奇偶校验码的去环分析^①郭 强^②

(东南大学移动通信国家重点实验室 南京 210096)

摘 要 针对低密度奇偶校验码(LDPC)中环的存在影响 LDPC 码性能的问题,研究了基于校验阵的有限长 LDPC 码的去环和优化,指出搜索随机矩阵的方法比较直观,但计算量是码长和列重乘积的指数函数。给出了搜索并替换的去环方法,该方法能够显著减小短环对译码性能的影响。计算机仿真结果表明,去除长度为 4 和 6 的环能够在误比特率为 10^{-5} 时获得 0.5dB 的编码增益。讨论了去环图的优化和判定准则,计算机仿真结果表明,优选的去环编码方案的性能要好于未经优选的方案。

关键词 低密度奇偶校验码(LDPC), 环, Tanner 图

0 引 言

近年来,使用稀疏矩阵构造的低密度奇偶校验码^[1](LDPC)由于其逼近香农限的出色性能和迭代译码的低复杂性,因而引起了越来越多的注意。在较长的分组长度时,非规则 LDPC 码的性能超过了 Turbo 码^[2]。LDPC 的编码随机集合对应的 Tanner 图无环时,采用置信传播算法得到的性能渐进收敛于最大似然译码的性能^[3]。在码长变短时,经验表明此时给定集合的码字实现之间性能差别很大。但是,研究分组长度较短(1000 比特左右的长度)的 LDPC 编码是有实际意义的。首先,低密度编码的译码和一般的分组码相比要简单一些;其次,低密度码纠错的出发点不同于传统的最小距离纠错,因而在相同的码长时可能获得更好的性能。

从直观上看,LDPC 的译码能否获得很好的性能,取决于迭代时传输信息的良好独立性。如果对应 Tanner 图中某一节点在长度为 $2l$ 的环中,则迭代次数大于 l 时节点收到信息之间的独立性假设将不再成立,这会降低迭代译码的品质甚至导致译码失败。在设计较长的编码时,对环的考虑是两方面的。一方面,经验证明分组长度越长,则长度较短的环所占的比例就越小,危害也就越小;另一方面,在设计校验矩阵时,部分短环已经被排除,如 MacKay 提出

的 Construction 1B 和 Construction 2B 方案首先排除了长度为 4 的环^[4]。本文的目的在于针对较短的分组,分析环的消除办法和对应的性能,并给出相应的仿真结论。

1 直观的方法

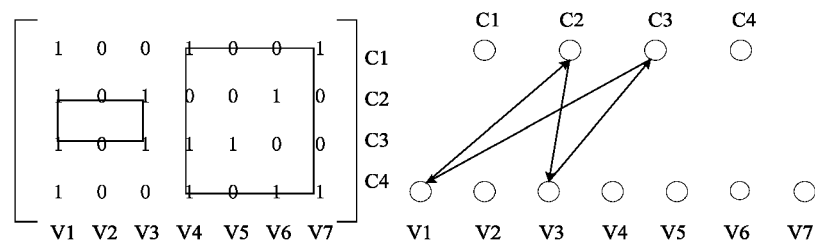
环对于 LDPC 性能的影响主要是环的存在使外部信息的转移条件遭到破坏。也就是从当前节点发出的信息经过迭代以后又返回到当前节点,这样,节点的人信息之间就不是独立的,则迭代的基础就不成立。显然,去环是很有必要的。由于一般是采用随机生成的编码方法,给定 (λ, ρ, n) , 对应的 Tanner 图肯定会含有环,所以设计中只能设法使其中的最小环长(girth)尽可能地大一些。不允许节点间有重复边时,长度为 2 的环是不存在的;长度为 4 和 6 的环较多,长度大于 8 以上的环少一些(针对短码)。所以,去环的对象是长度小于 8 的环。

由于 Tanner 图和校验阵的对称性,可以从校验阵上直观地看到环的存在,如图 1 所示。

可以看出,当校验阵的两列在对应比特上都为 1 时,也即有矩形顶点存在时,产生了长度为 4 的环。对于长度为 6 的环(同以前的说明,不会有奇数长度的环存在),可以归纳为具有特定的形状如图 2 所示。

① 863 计划(2007AA01Z2B1)资助项目。

② 男,1957 生,学士,副教授;研究方向:信号处理、模式识别;联系人, E-mail: qguo@seu.edu.cn
(收稿日期:2008-05-31)



(为了清楚表示为有向的,实际中是双向的)对应校验阵左边的矩形

图 1 校验阵和 Tanner 图上环的对应关系, Tanner 图上的长度为 4 的环

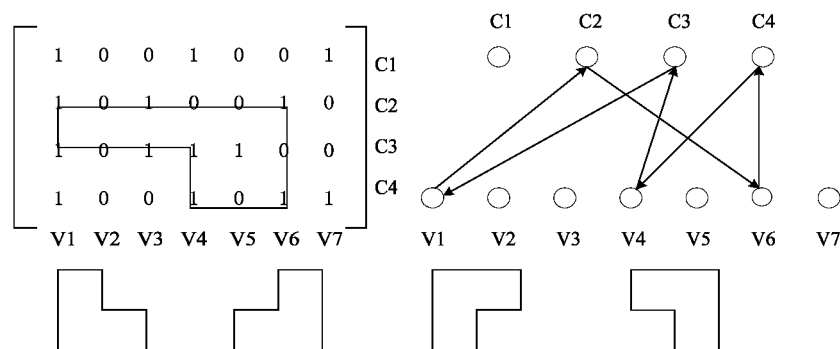


图 2 长度为 6 的环的情形

长度为 4 的环不会包含在长度为 6 的环中,因为某一节点的独立迭代次数只与所在的最短环长有关。在图 2 所示的矩阵中,如果去掉(2,3)和(3,3)位置的 1,不会影响长度为 6 的环。进一步推广,可以得到判定长度为 8 的环。

可以看出,随着环的长度增加,这样直观的方法会越来越复杂,但对于有限长(一般为 1000 比特以下)的 LDPC 码,长度大于 8 的环的数目是很少的,所以不需要再进一步搜索。由于每一步的判断是在与上一判断的垂直方向上进行,所以可以粗略地估计搜索总的计算量为 2^m , 其中 n 为分组的长度, t 为列重。

2 搜索并替换的简单方法

上面的方法是搜索随机生成的矩阵并判定是否具有较大的 girth, 这样搜索到一个合乎要求的矩阵需要较多的判断。下面介绍的方法是检查校验阵中环的存在性并且替换其中的长度较短的环。

给定校验阵,可以得到如下形式的方阵 A 及 A 的平方为:

$$A = \begin{pmatrix} 0 & H \\ H^T & 0 \end{pmatrix} \quad A^2 = \begin{pmatrix} HH^T & 0 \\ 0 & H^T H \end{pmatrix}$$

其中 A^2 的元素 $A_{ij}^2 = \sum_{k=1}^p a_{ik}a_{kj}$ 。

显然,和值表示了节点 v_i 和 v_j 之间长度为 2 的路径的数目,由此得到推论: A^n 的值 A_{ij}^n 表示在节点 v_i 和 v_j 之间长度为 n 的路径的数目。对角元素 A_{ii}^n 表示开始且终止于节点 v_i 的长度为 n 的路径(环)的数目,这其中包含了冗余,进一步判定半程环的存在后即可确定环的存在。

设 v_j 和 v_i 是长度为 n 的环中的对极两点,可以选择 v_i 的邻接点 v_k , 显然有 $A_{ik}^{(n/2)-1} > 0, a_{jk} = 1$ 成立,这样的到一条边 $e = \overline{v_j v_k}$ 。同时得到一个节点集,其中的任一节点 c 到 e 的距离都大于 $(n-1)$, 显然满足 $A_{cj}^{n-1} = A_{ck}^{n-1} = 0$ 。随便选择两个端点都在上述节点集中的一条边 e' , 和边 e 交换端点如图 3 所示,即可去掉长度为 n 的环,同时不产生新的环^[5]。

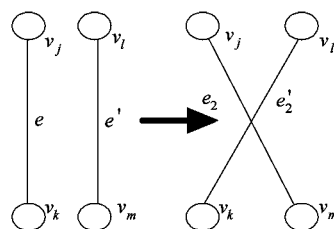


图 3 新旧边的替换

经验表明^[5],一定码长(N_v)的可以去环的最大值 $g_{\max}(N_v)$ 如表 1:

表 1 码长可去环最大值

g	N_v	$g_{\max}(N_v)$
6	40	7.1
8	250	10.3
10	1600	13.5

需要指出的是,随着码长的增加,采用方阵多次乘方的形式,运算量会大大增加而影响去环的效率,所以,去环的目的不是完全消除环的存在,而是减小环的危害,当然,环的长度愈短,危害就越大。从图 4 可以看出,当信噪比 (E_o/N_o) 较小时,去环带来的性能改善不大,而信噪比大于 3dB 时,去环的影响开始显现,但同时去除更大的环带来的增益逐渐减小。

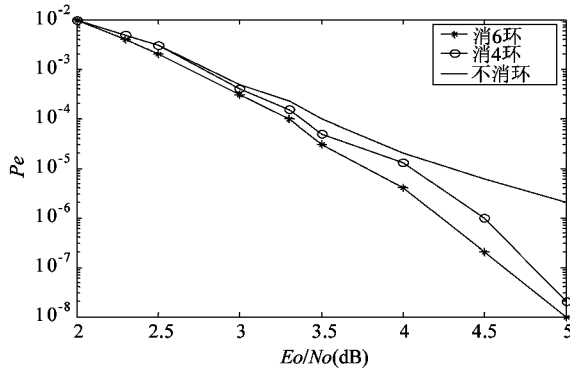


图 4 分组长度 300bit, 码率为 0.5 的 LDPC 码在没有去环, 消去长度为 4 的环, 消去长度为 6 的环所带来的比特错误概率 (P_e) 的变化

3 去环图的优化和判定准则

为衡量一个 Tanner 图对应码字的优劣,可以定义平均最小环分布 (average girth distribution) 为:

$\sum_{k=2}^{l_{\max}/2} g(2k) \cdot 2k$, 其中, $g(2k)$ 为拥有环长 $2k$ 的符号节点的比例, $l_{\max}$ 为最大的环长。搜索图中的环使用树的生长办法,计算量为 $o(n^2)$ 。直观上看,平均最小环分布与迭代译码的次优化(使用的是有限长 LDPC,谈不上最优化)相关^[6]。

使用 Construction 2A 方法^[4],可以得到一组具有相同分组长度和度分布对的码字,但是,这些有限长码字的性能差别很大。此时,可以通过计算每一种码字的平均最小环分布来选择具有较好性能的码字,当然,要选择具有最大指标的码字。图 5 给出了一个借助与上述方法选择码字的例子。图 6 给出了

随机选择和优选的码字的性能比较。注意到性能比图 4 的情形要好一些,这是因为分组的长度增加了。

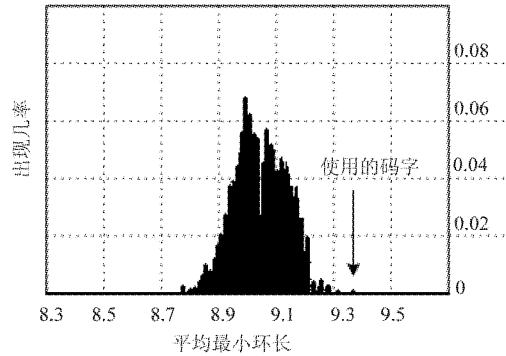


图 5 分组长度为 1268 比特, 信息比特为 456 的随机码集合的平均最小环分布情况, 横轴表示平均最小环的值, 纵轴表示出现的相对几率

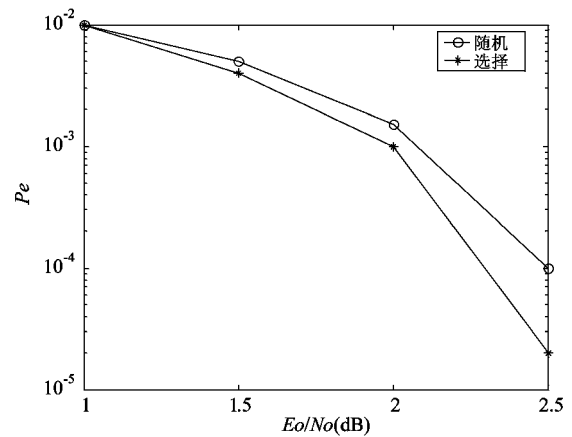


图 6 两种不同方法选择的码字的性能, 优选改进了性能

对于非规则 LDPC 编码, 自然存在对度分布对的优化问题, 但是试验表明^[7], 对于长 LDPC 优化得到的度分布对, 对于分组长度较短的情形并不是合适的。不过值得注意的是如果考虑到不同节点的不等错误保护, 非规则的短码会得到一定的性能改善。

对于 LDPC 的最小距离特性, 已经知道在分组无限长时, 可以取得吉尔波特限, 但是分组长度缩小时, 计算精确的最小距离将变的困难起来。文献[8]证明了构造具有较大 girth 的码并不导向具有较大最小距离的码字。同时, MacKay 指出, LDPC 的优势在于采用近似最优的译码可取得的性能是同等条件下最小距离译码的两倍, 同时, 没有办法通过最小距离特性区分好码^[4]。所以, 分析较短分组的最小距离特性并无决定性的意义。

4 结 论

本文主要讨论了处理校验阵来搜索 Tanner 图中环的方法,其中,直观的方法和平均最小环分布的方法都是建立在随机判定的基础上,而矩阵乘方的方法则是发现并替换。总的说来,这些方法本身具有的高计算量决定了只能使用于有限长的 LDPC 码,而且提出这些方法的初衷也是为了去除长度较短的环,而不是盲目地追逐最大 girth。

但是也看到 LDPC 在有限长时,所表现的性能(见图 5,图 7)不再是 MacKay 所说的“甚好码”,只能归入“好码”的类型。近来业界对于 LDPC 短码的研究重点放在了其改进形式上,例如和 RS 码结合在 xDSL 中的应用,采用多进制的形式,以及以 LDPC 为母码,构造结合自动请求重发(automatic repeat quest, ARQ)方式的高效码率可变码。

参考文献

- [1] Gallager R. Low-density Parity-check Codes. Cambridge MA: MIT Press, 1963. 1-56
- [2] Young S C, Richardson T J, Urbanke R L. Analysis of sum-product decoding of low-density parity-check codes using a Gaussian approximation. *IEEE Trans Information Theory*, 2001, 47(2):657-670
- [3] Richardson T J, Urbanke R L. The capacity of low-density parity-check codes under message-passing decoding. *IEEE Trans Information Theory*, 2001, 47(2):599-618
- [4] MacKay D J C. Good error-correcting codes based on very sparse matrices. *IEEE Trans Information Theory*, 1999, 45(2):399-431
- [5] McGowan J A, Williamson R C. Loop removal from LDPC codes. In: Proceedings of the 2003 IEEE Information Theory Workshop, Paris, France, 2003. 230-233
- [6] Mao Y Y, Banihashemi A H. A heuristic search for good low-density parity-check codes at short block lengths. In: Proceedings of 2001 IEEE International Conference on Communications, Helsinki, Finland, 2001. 41-44
- [7] Xue Y J, Xiang H G. Performance analysis of finite-length LDPC codes. In: Proceedings of the 4th IEEE Workshop on Signal Processing Advances in Wireless Communications, Rome, Italy, 2003. 85-89
- [8] Wei L. Several properties of LDPC codes. *IEEE Trans On Communications*, 2004, 52(5):721-727

Analysis of loop removal from finite-length LDPC codes

Guo Qiang

(National Mobile Communications Research Laboratory, Southeast University, Nanjing 210096)

Abstract

In order to eliminate the negative effect of loops on the performance of low-density parity-check codes (LDPC), the paper studies loop removal and optimization of finite-length LDPC codes based on parity-check matrices, and points out that the overall search is simple but has a computation complexity with the exponential degree of the product of code length and column weight. A method of search and substitution is given, and its computer simulation results indicate that the performance of codes gain can be improved 0.5dB with 10^{-5} BER when girth is more than 6. The optimization rules of loop-removal graph are analysed and the simulation results indicate that the selected coding is better.

Key words: low-density parity-check codes(LDPC), loop, Tanner graph