有限长不规则 LDPC 码的构造和编码的优化

胡军锋[①②]　　张海林[①]
[①](西安电子科技大学综合业务网理论及关键技术国家重点实验室　西安　710071)
[②](中国电子科技集团第二十研究所　西安　710068)

摘　要：该文分析了影响有限长低密度校验(LDPC)码性能的主要因素，在此基础上从度分布参数的优选为起点，结合改进的循序边增长(PEG)算法构造出初步的校验矩阵，提出一种实用的编码优化算法对该校验矩阵进一步优化，最终得到错误平底低且编码复杂度准线性的有限长不规则 LDPC 码。该优化方法可以容易地推广到一般的信道条件下。

关键词：有限长 LDPC 码；错误平底；边循序增长(PEG)算法
中图分类号：TN911.22　　**文献标识码**：A　　**文章编号**：1009-5896(2008)03-0600-04

Optimization of Construction and Encoding for Finite Length Irregular LDPC Codes

Hu Jun-feng[①②]　　Zhang Hai-lin[①]
[①](*State Key Laboratory of Integrate Services Network*, *Xidian University*, *Xi'an* 710071, *China*)
[②](*CETC No.* 20 *Research Institute*, *Xi'an* 710068, *China*)

Abstract: The main influence for the performance of finite length Low-Density Parity-Check(LDPC) codes is analyzed. According to optimum choice for degree distribution, a check matrix is constructed with improved Progressive-Edge-Growth (PEG) algorithm. A practical efficient encoding algorithm is proposed to optimize the check matrix. A finite length irregular LDPC code with low error-floor performance and approximate linear encoding complexity is obtained. This optimization method can be easily extended to general communication channels.

Key words: Finite length Low-Density Parity-Check (LDPC) codes; Error-floor; Improved Progressive- Edge-Growth (PEG) algorithm

1　引言

低密度校验(LDPC)码因其优异的纠错性能和高效的迭代译码方法而受到越来越多的关注。Richardson 等人采用密度进化算法并结合线性规划的优化方法，找出了几乎达到香农容量的渐近无限长不规则 LDPC 码的度分布参数[1]。另外，文献[2]利用 LDPC 码的校验矩阵密度很低这一特点，提出了准线性复杂度的高效编码方法。

实际应用中的码子都是有限长度的，这时由于诸多因素的影响，使得有限长 LDPC 码的性能有所下降，尤其是在高信噪比区会出现一定的错误平底。但是，如果码子构造适当，便可以把错误平底降到很低，从而大大改善有限长 LDPC 码的性能。本文简单分析了影响有限长 LDPC 码性能的主要因素，在此基础上对有限长不规则 LDPC 码的构造和编码进行优化，得到错误平底低且编码复杂度准线性的码子。

2　影响有限长 LDPC 码性能的主要因素

2.1 有关定义

由于 LDPC 码的校验矩阵与 Tanner 图是一一对应的，在后文中将混合使用这两个概念。为了描述方便，接下来引入几个定义。

定义 1(度分布对)　码长为 N ，度分布对为 (λ,ρ) 的 LDPC 码集用 $C_N(\lambda,\rho)$ 表示。其中 λ ， ρ 定义如下[1]：

$$\lambda(x)=\sum_{i=1}^{d_v^{\max}}\lambda_i x^{i-1}\text{，}\quad \rho(x)=\sum_{i=1}^{d_c^{\max}}\rho_i x^{i-1} \tag{1}$$

这里的 $\lambda_i(\rho_i)$ 是 Tanner 图中连接到度为 i 的变量节点(校验节点)的边在总边数中所占的比例， $d_v^{\max}(d_c^{\max})$ 是变量节点(校验节点)的最大度数。

结合文献[3]引入另外几个概念：Stopping set，用 S 表示；大小为 d 的 S 用 Sd 表示；外信息度，用 EMD 表示；环的近似外信息度，用 ACE 表示。

2.1 影响有限长 LDPC 码性能的因素分析

接下来分析影响有限长 LDPC 码性能的几个主要因素。其一，LDPC 码是一种线性分组码，与一般的线性分组码一样，其纠错能力受到最小码距的限制，所以最小码距是影响 LDPC 码性能的主要因素之一。其二，度分布对是影响 LDPC 码性能的另一个主要因素。以上两个因素对有限长和渐近无限长 LDPC 码都会产生影响。其三，实际应用中的有限长

2006-08-11 收到，2007-03-21 改回

LDPC 码，其 Tanner 图中不可避免地会存在环，BP 译码经过有限次迭代后的迭代信息之间就不再独立，从而导致迭代译码只是一种次优译码方法。所以，Tanner 图中的环(尤其是短环)也是影响有限长 LDPC 码的因素。其四，根据文献[4]，Sd 对有限长 LDPC 码的性能影响至关重要。文献[3]还具体分析了 Sd 与最小码距以及环之间的密切关系，并指出最小 Sd 越大，LDPC 码子的错误平底越低。除了以上因素外，码重分布、环在 Tanner 图中的位置关系、Sd 的分布特征等因素对 LDPC 码的纠错性能和迭代收敛性能也有着不同程度的影响，至于这些因素如何具体影响 LDPC 码的性能，目前还是一个开放性的问题。

3 有限长不规则 LDPC 码的优化构造

文献[5]对有限长 LDPC 码的度分布对做了初步的约束优化，给出了低错误平底的度分布参数。有了度分布对 (λ,ρ)，只是得到了平均性能很好的码集合 $C(\lambda,\rho)$，还需要进一步从 $C(\lambda,\rho)$ 中找到好性能的码。借助 Tanner 图，以最小环长和最小 Sd 最大化为目标构造 LDPC 码是方便的。事实上，这样也间接地提高了最小码距[3]。文献[6] 提出局部环长最大化的 PEG 算法，文献[7]结合文献[3]中的思想将 EMD 和 ACE 引入 PEG 算法得到改进的 PEG 算法(IPEG)。利用 IPEG 算法构造出的 LDPC 码的最小环长(girth)较大，而且最小 Sd 也较大。所以这里用 IPEG 算法来构造有限长 LDPC 码的校验矩阵。

首先根据应用条件确定编码速率 R，由 R 从文献[1]或文献[5]的表中选择合适的变量节点度分布 $\lambda(x)$。接下来计算变量节点度分布序列 Dv。记变量节点的平均度为 $\tilde{d}_v$，则[1]

$$\tilde{d}_v=\frac{1}{\sum_i\left(\lambda_i/i\right)} \tag{2}$$

记度为 i 的变量节点数在变量节点总数中所占的比例为 $\tilde{\lambda}_i$，则

$$\tilde{\lambda}_i=\frac{1}{N}\left(\frac{\lambda_i N\tilde{d}_v}{i}\right)=\frac{\lambda_i\tilde{d}_v}{i} \tag{3}$$

记度等于 i 的变量节点的个数为 n_i，则

$$n_i=\tilde{\lambda}_i N \tag{4}$$

这样便可得到度分布序列[6] $Dv\triangleq\{dv_0,dv_1,\cdots,dv_{N-1}\}$，其中 dv_i 是变量节点 v_i 的度，$0\le i\le N-1$，序列以不减的顺序排列，即 $dv_0\le dv_1\le\cdots\le dv_{N-1}$。求得 Dv，便可根据 IPEG 算法构造出校验矩阵，记为 $\boldsymbol{H}_0$。$\boldsymbol{H}_0$ 便是从 $C_N(\lambda,\rho)$ 码集中找到的性能较好的 LDPC 码的校验矩阵。

4 编码的优化

结合 $\boldsymbol{H}_0$ 的特点，对编码过程加以优化，不但能使编码复杂度降到准线性，还可以进一步提升码子的误信息比特率(BER)性能。

一般 LDPC 码的编码复杂度为 $O(N^2)$，对于行满秩的校验矩阵，文献[2]提出一种高效的编码方法，通过行列置换，把 $\boldsymbol{H}_0$ 变换成如图 1 所示的近似下三角形式，然后进行编码。这种方法的编码复杂度只有 $O(N+g^2)$，对于纠错性能好的 LDPC 码而言，$g\ll N$，所以编码复杂度是准线性的。与近似下三角形式的校验矩阵相一致，把一个 N 长的 LDPC 码子分成三部分，表示为 $x=(s,p_1,p_2)$，其中 s 是信息比特部分，长度为 $N-M$，p_1、p_2 是校验部分，长度分别为 g 和 $M-g$。

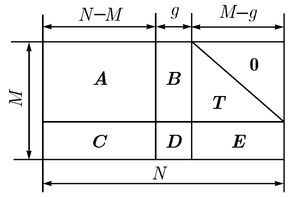

图 1 校验矩阵近似下三角化

在线编码之前要对校验矩阵预处理，记图 1 中的近似下三角矩阵为 $\boldsymbol{H}_1$，即

$$\boldsymbol{H}_1=\begin{pmatrix}\boldsymbol{A} & \boldsymbol{B} & \boldsymbol{T}\\ \boldsymbol{C} & \boldsymbol{D} & \boldsymbol{E}\end{pmatrix} \tag{5}$$

对 $\boldsymbol{H}_1$ 实施初等行变换，得到 $\boldsymbol{H}_2$

$$\boldsymbol{H}_2=\begin{pmatrix}\boldsymbol{A} & \boldsymbol{B} & \boldsymbol{T}\\ -\boldsymbol{E}\boldsymbol{T}^{-1}\boldsymbol{A}+\boldsymbol{C} & -\boldsymbol{E}\boldsymbol{T}^{-1}\boldsymbol{B}+\boldsymbol{D} & \boldsymbol{0}\end{pmatrix} \tag{6}$$

记 $\boldsymbol{f}=-\boldsymbol{E}\boldsymbol{T}^{-1}\boldsymbol{B}+\boldsymbol{D}$。如果 $\boldsymbol{f}$ 是奇异矩阵，则要对 $\boldsymbol{H}_1$ 的列进行调整，以保证 $\boldsymbol{f}$ 可逆。

用 IPEG 算法构造出的 $\boldsymbol{H}_0$ 实际上都是行满秩的。另外，由于变量节点的度越大，与之对应的码子比特的置信度越高，所以尽可能把度数最高的 $N-M$ 个变量节点分配给信息比特 s，这样可以使码子的误信息比特率降低，从而进一步提高 LDPC 码的性能。还有一个目标就是尽可能使得 g 最小，以降低编码复杂度。基于以上策略，引入与文献[2]相同的表示符号 known 和 $\widetilde{\boldsymbol{A}}$，下面给出一种实用的算法。

步骤 1(初始化)：

将 $\boldsymbol{H}_0$ 的列进行逆序排列，得到 $\widetilde{\boldsymbol{H}}_0$，令 known $=\widetilde{\boldsymbol{H}}_0$ 的前 $N-M$ 列，令 $\widetilde{\boldsymbol{A}}=\widetilde{\boldsymbol{H}}_0$ 的后 M 列；

步骤 2(构造 $\boldsymbol{H}_1$)：

while $\widetilde{\boldsymbol{A}}$ 的列大于 1

if $\widetilde{\boldsymbol{A}}$ 中存在行重为 1 的行

用文献[2]给出的方法实施对角扩展，并将扩展后余下的子矩阵重新赋值给 $\widetilde{\boldsymbol{A}}$；

else

在 $\widetilde{\boldsymbol{A}}$ 中找出行重最小(大于 0)的行(如果有多行满足条件，任意选一行)，记该行重为 i，把该行中前 $i-1$ 个“1”所在的列从 $\widetilde{\boldsymbol{A}}$ 中删除并移至 known 部分(放置在 known 的末尾)。将 $\widetilde{\boldsymbol{A}}$ 删除了 $i-1$ 个列以后的子矩阵重新赋

值给 $\widetilde{\boldsymbol{A}}$；

end

end

步骤 3($\boldsymbol{H}_1 \to \boldsymbol{H}_2$) 对 $\boldsymbol{H}_1$ 实施初等行变换，将 $\boldsymbol{E}$ 所在的块变成零矩阵，得到 $\boldsymbol{H}_2$ 和 $\boldsymbol{f}$；

步骤 4(列调整或停止)：

if $\boldsymbol{f}$ 非奇异

令 $\boldsymbol{H} = \boldsymbol{H}_1$，停止；

else

对 $\boldsymbol{H}_1$ 进行列调整，使 $\boldsymbol{f}$ 非奇异，将经过列调整以后的 $\boldsymbol{H}_1$ 重新赋值给 $\boldsymbol{H}_1$，令 $\boldsymbol{H} = \boldsymbol{H}_1$；

end

其中步骤 4 中的列调整规则如下：记 $\boldsymbol{H}_2$ 中的 $-\boldsymbol{E}\boldsymbol{T}^{-1}\boldsymbol{A}+\boldsymbol{C}$ 块构成的子矩阵为 $\boldsymbol{\beta}$，通过对 $\boldsymbol{f}$ 化行阶梯型判断 $\boldsymbol{f}$ 是否奇异，如果 $\boldsymbol{f}$ 奇异，就对 $\boldsymbol{\beta}$ 也化行阶梯型，由 $\boldsymbol{f}$ 和 $\boldsymbol{\beta}$ 的行阶梯型可以容易地看出 $\boldsymbol{f}$ 和 $\boldsymbol{\beta}$ 中的那些列对调可以使 $\boldsymbol{f}$ 非奇异，如果在 $\boldsymbol{\beta}$ 中有多列可供选择，选择列号最大的列，然后将 $\boldsymbol{H}_1$ 中对应的列对调。这样便保证了度数高的变量节点尽量不被调换到校验比特部分。

得到的 $\boldsymbol{H}$ 便是最终的校验矩阵。对 $\boldsymbol{H}$ 实施步骤 3，得到非奇异的 $\boldsymbol{f}$，对增广矩阵 $(\boldsymbol{f}\,|\,\boldsymbol{I})$ 实施初等行变换可以求得 $\boldsymbol{f}^{-1}$。以上算法中的所有操作都是在 GF(2) 域进行。至此便得到了 $\boldsymbol{A}$，$\boldsymbol{B}$，$\boldsymbol{T}$，$\boldsymbol{C}$，$\boldsymbol{E}$，$\boldsymbol{f}^{-1}$，直接用文献[2]中表 1 和表 2 的算法完成在线编码。

5 性能仿真分析

5.1 仿真参数

AWGN 信道，BPSK 调制。所有的码长均为 $N = 1008$，编码速率均是 $R = 0.5$。经过约束优化的度分布参数选自文献[5]的表 1($E = 0.19$)，记为 $\lambda_1(x)$，即 $\lambda_1(x) = 0.2068x + 0.6808x^2 + 0.1025x^3 + 0.1067x^4$；没经过约束优化的度分布参数选自文献[1]的表 1($d_v^{\max} = 4$)，记为 $\lambda_2(x)$，即 $\lambda_2(x) = \ \backslash\ 0.38354x + 0.04237x^2 + 0.57409x^3$。其余参数和校验矩阵的构造方法如表 1 所示。译码采用以对数似然比(LLR)作为软信息的 BP 算法[8]，译码最大迭代次数选取 80(与文献[4]相同)。

5.2 仿真结果分析

仿真结果如图 2 所示。由图可见，用本文提出的优化方法构造出的 LDPC 码的性能不但比随机构造的码要好很多，而且比经过优化的规则 LDPC 码也要好。另外，正如文献[7]所说，IPEG 方法构造出的 LDPC 码和 PEG 方法构造出的 LDPC 码相比，低信噪比区的性能几乎一样，但错误平底却要比后者低。通过比较 Optimized dv=4 和 IPEG,，可以看出经过本文提出的编码算法进一步优化，得到的 LDPC 码的性能比 IPEG 有所提高。如果对变量节点的度分布参数加以约束优化(Optimized E=0.19)，比特信噪比为 2.5dB 时的误比特率接近 10^{-7} 量级，而且在该量级没有出现错误平底的趋势。在编码复杂度方面，Optimized E=0.19 和 Optimized dv=4 的 g 值分别为 26 和 1，都远小于码长 $N = 1008$，所以复杂度接近线性。

表 1 仿真中所用 LDPC 码的构造方法及其参数

($N = 1008, R = 0.5$)

LDPC 码表示	度分布参数	构造方法
Optimized E=0.19	$\lambda_1(x)$	本文的综合优化方法
Optimized dv=4	$\lambda_2(x)$	本文的综合优化方法
IPEG	$\lambda_2(x)$	改进的 PEG 方法
PEG	$\lambda_2(x)$	标准的 PEG 方法
Mackay regular(3,6)	规则的(3，6)	Machay 通过优化找到的规则 LDPC 码[9]
Random	$\lambda_2(x)$	随机构造，满足行满秩

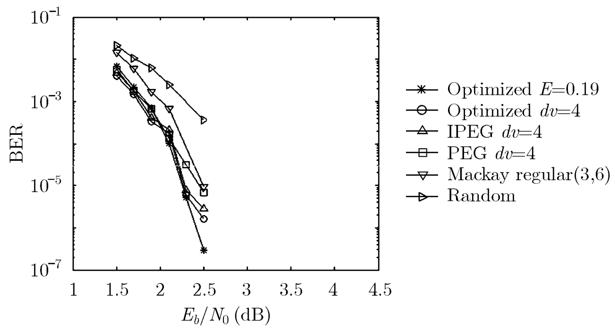

图 2 LDPC 码的误信息比特率性能比较

6 结束语

LDPC 码已被证明是几乎达到香农极限的纠错码，有广阔的应用前景。但是当码长比较短时，如果不精心设计，性能会急剧恶化。利用本文提出的有限长不规则 LDPC 码的优化构造和编码方法，可以构造出性能优良而且编码复杂度准线性的好码。文中虽然以二进制输入 AWGN 信道为模型，但是该优化策略和方法可以方便地推广到衰落、MIMO 等更多的信道条件下。

参 考 文 献

[1] Richardson T, Shokrollahi M, and Urbanke R. Design of capacity-approaching irregular low-density parity-check codes. *IEEE Trans. on Inform. Theory*, 2001, 47(2): 619–637.

[2] Richardson T and Urbanke R. Efficient encoding of low-density parity-check codes. *IEEE Trans. on Inf. Theory*, 2001, 47(2): 638–656.

[3] Tian T, Jones C, Villasenor J D, and Wesel R D. Selective avoidance of cycles in irregular LDPC code construction. *IEEE Trans. on Commun.*, 2004, 52(8): 1242–1247.

[4] Di C, Proietti D, Telatar I E, Richardson T J, and Urbanke R L. Finite-length analysis of low-density parity-check codes on the binary erasure channel. *IEEE Trans. on Info. Theory*, 2002, 48(6): 1570–1579.

[5] Johnson S J and Weller S R. Constraining LDPC degree distributions for improved error floor performance. *IEEE Commun. Letters*, 2006, 10(4): 103–105.

[6] Hu X Y, Eleftheriou E, and Arnold D M. Regular and irregular progressive edge-growth Tanner graphs. *IEEE Trans.on Info. Theory*, 2005, 51(1): 386–398.

[7] Xiao H and Banihashemi A H. Improved Progressive-Edge-Growth (PEG) construction of irregular LDPC codes. *IEEE Commun. Letters,* 2004, 8(9): 715–717.

[8] Chen J, Dholakia A, Eleftheriou E, Fossorier M P C, and Hu X. Reduced-complexity decoding of LDPC codes. *IEEE Trans. on Commun.*, 2005, 53(8): 1288–1299.

[9] Mackay D J C. [Online], http://www.inference.phy.cam.ac.uk/ mackay/codes/data.html, 2001.

胡军锋： 男，1976年生，博士生，工程师，研究方向为MIMO通信系统、LDPC码、通信信号处理.

张海林： 男，1963 年生，教授，博士生导师，主要研究方向包括多媒体通信、宽带无线通信、通信信号处理、应急通信等.