

不规则 LDPC 码的密度进化方法及其门限值确定

肖娟* 王琳** 邓礼钊***

* (重庆邮电学院编码技术研究所 重庆 400065)

** (厦门大学电子工程系 厦门 361005)

*** (重庆通信管理局 重庆 400042)

摘要: 该文针对不规则 LDPC 码阐述密度进化原理, 比较分析了具体算法的特点与应用。对离散密度进化和高斯近似算法进行了仿真分析, 寻找一些好的分布对。这对促进 LDPC 码以及其它现代高效纠错码的理论分析和应用研究具有重要指导意义。

关键词: 不规则 LDPC 码, 密度进化方法, 门限值, 高斯近似

中图分类号: TN918.1 **文献标识码:** A **文章编号:** 1009-5896(2005)04-0617-04

Density Evolution Method and Threshold Decision for Irregular LDPC Codes

Xiao Juan* Wang Lin** Deng Li-zhao***

* (Institute of Info. and Coding Tech., Chongqing Univ. of Posts & Telecomm., Chongqing 400065, China)

** (Dept of Electronic Engineering, Xiamen Univ., Xiamen 361005, China)

*** (Chongqing Communication Administration, Chongqing 400042, China)

Abstract In this paper, the principle of density evolution combined with decoding process is explained firstly. Two algorithms discretized density evolution and Gaussian approximation to program the evolution proceeding are discussed. Simulation results of some good distribution pairs that are found are presented as well.

Key words Irregular LDPC codes, Density evolution, Threshold, Gaussian approximation

1 引言

传统的编码理论通过距离特性的分析来判定硬判决译码的纠错能力; 而采用软判决的迭代译码的纠错性能是随着迭代次数的增加而渐进体现的, 密度进化方法正是分析这种渐进性能的理论。

不规则 LDPC (Low Density Parity Check) 码是优秀的线性纠错码, 最好的仿真结果显示, 它可以距离香农限只有 0.0045dB^[1]。LDPC 码的纠错性能可以由门限值来描述, 分布对 (λ, ρ) 的门限值越大, 与该码率下的香农限越近^[2,3]。密度进化方法则可以帮助寻找分布对在具体信道下的门限值。不规则 LDPC 码的度的分布对设计目前尚无确定性理论推导方法, 但可以通过密度进化方法实现度的分布对筛选, 从而优化不规则 LDPC 码的设计。

文章安排如下: 第 2 节结合消息传递译码方法分析了不

规则 LDPC 码的密度进化原理; 第 3 节介绍密度进化的具体实现算法: 离散化算法和应用于高斯信道的近似算法, 比较分析了各自的特点和应用; 第 4 节列出了我们通过仿真寻找得到的一些优秀的分布对及其相应的门限值。

2 密度进化与门限值确定

2.1 消息的密度进化

在消息传递算法中, 消息携带概率似然值在二部图的边上来回传递。变量点和校验点接收从各边传来的消息值, 再分别计算发送给各边的新的消息值, 从而实现消息的迭代更新。用 v 代表在第 l 次迭代中从变量点到校验点的 LLR (Log Likelihood Ratio) 消息。消息的迭代更新过程如下^[1,3]:

$$v^{(l)} = \begin{cases} u_0, & l = 0 \\ \sum_{i=0}^{d_v-1} u_i^{(l)}, & l > 0 \end{cases} \quad (1)$$

2003-11-28 收到, 2004-07-09 改回
国家 863 计划项目(2001AA123061)和国家自然科学基金(60272005)资助课题

$$\tanh \frac{u^{(l)}}{2} = \prod_{j=1}^{d_v-1} \tanh \frac{v_j^{j-1}}{2} \quad (2)$$

其中, d_v 和 d_c 是最大的变量点的度和最大的校验点的度(度表示该点连接的其它点的个数)。

译码过程中, v 和 u 是服从一定的概率分布的随机变量, $p_v^l(p_u^l)$ 表示第 l 次迭代中 $v(u)$ 的概率密度函数。假设传输全 0 码字, 则 v 取负值的概率对应了误比特率。当信道质量参数小于门限值时, $p_v^l(p_u^l)$ 具有收敛性: 经过足够多的迭代, $p_v^l(p_u^l)$ 会聚集在 $v(u)$ 等于 $+\infty$ 的点上, 使得误比特率趋近于 0。

根据概率理论, 多个独立的随机变量相加, 其和的概率密度等于原随机变量概率密度的卷积。根据式(1), v 的概率分布变化如下:

$$p_v^l = \begin{cases} p_0, & l=0 \\ p_0 * (\otimes_{i=1}^{d_v-1} p_{u_i}^l), & l>0 \end{cases} \quad (3)$$

其中 p_0 是由信道质量得出的初始概率密度, $*$ 和 \otimes 在此表示卷积运算。由于 $u_i, 1 \leq i < d_v$ 独立同分布, 所以当 $l > 0$ 时, $p_v^l = p_0 * (\otimes_{i=1}^{d_v-1} p_{u_i}^l)$ 。

u 的密度进化相对更复杂。定义运算符 γ : $\gamma(x, y) = 2 \tanh^{-1}[\tanh(x/2) \tanh(y/2)]$, 式(2)则演变为 $u = \gamma(v_1, \gamma(v_2, \dots, \gamma(v_{d_v-2}, v_{d_v-2}) \dots))$ 。

设 $z = \gamma(x, y)$, 则 $p_z(z) = \sum_{(x,y):z=\gamma(x,y)} p_x(x)p_y(y)$, 简记为 $p_z = R(p_x, p_y)$ 。

由于 $v_i, 1 \leq i < d_c$ 是独立同分布的, u 的概率密度可计算如下:

$$p_u^{l+1} = R(p_v^l, R(p_v^l, \dots, R(p_v^l, p_v^l) \dots)) \quad (4)$$

对于不规则 LDPC 码, 度的分布对 (λ, ρ) : $\lambda(x) = \sum_{i=2}^{d_v} \lambda_i x^{i-1}$, $\rho(x) = \sum_{j=2}^{d_c} \rho_j x^{j-1}$; 其中, $\lambda(x)$ 表示变量点的度的分布, $\rho(x)$ 表示校验点的度的分布, $\lambda(1) = 1, \rho(1) = 1$ 。 $\lambda_i(\rho_i)$ 分别表示从度为 i 的变量点(校验点)发出的边占总的边数的比例。根据前面的推导, 可以定义如下两个函数:

$$\lambda'(p) = \sum_{i=2}^{d_v} \lambda_i \otimes^{i-1} p$$

$$\rho'(p) = \sum_{j=2}^{d_c} \rho_j R^{j-1} p$$

结合度的分布对, 不规则 LDPC 码消息的密度进化过程可以总结为

$$p_v^l = p_0 * \lambda'(\rho'(p_v^{l-1}))$$

图 1 复制了 Matlab 仿真时 P_U 和 P_V 的波形, 体现了不规则 LDPC 码在译码过程中, 消息概率密度分布的变化趋势。BIAWGN(Binary Input Additive White Gaussian Noise)信道, $d_v = 30$, 码率 0.5, 信道质量参数 $\sigma = 0.9562$ 。图 1(a)是消息的初始分布, v 为高斯分布, u 为冲击响应; 图 1(b)是 $l=30$ 时, 消息密度概率分布 p_v^{30} 和 p_u^{30} , v 仍保持了高斯分布的波形, 而 u 呈尖脉冲状态。经过有限次迭代, 曲线明显右移, $u(v)$ 取负值的概率趋近于 0。为了在同一张图中展现 P_U 和 P_V 的波形, 对 P_V 的波形进行了线性放大。

注: 本图是波形示意图, 因为 $\int_{-\infty}^{\infty} P_U(u) du = 1$ 和 $\int_{-\infty}^{\infty} P_V(v) dv = 1$, 而 P_V 比较平坦, 所以 P_V 的峰值比 P_U 小很多。为了在同一图内表示两者的波形, 对 P_V 的波形进行了线性放大。

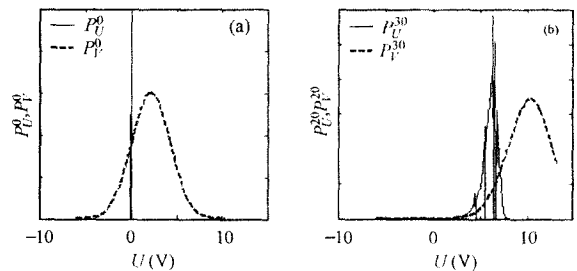


图 1 不规则 LDPC 码的消息的概率密度分布

2.2 门限值的确定

对任一不规则 LDPC 码的分布对 (λ, ρ) , 如果以 σ 表示对称信道的质量参数, 则存在一个最大信道参数 σ^* , 使得当 $\sigma < \sigma^*$ 时, 随着码长和译码迭代次数的增加, 随机构造的码 $C(\lambda, \rho)$ 可以达到任意小的误比特率。该值 (σ^*) 即是分布对为 (λ, ρ) 的 LDPC 码的门限值。

不失一般性, 假设传输全 0 码字。首先选择一个信道参数 σ , 如果多次迭代后 $p_v^l(p_u^l)$ 聚集在 $v(u)$ 等于 $+\infty$ 的点上(误比特率趋近于 0), 则 σ 在门限值之内。逐步增大 σ 的取值, 该分布对 (λ, ρ) 构造的不规则 LDPC 码的门限值 σ^* , 是能使 $p_v^l(p_u^l)$ 成功收敛的 σ 的最大值。

2.3 分布对的优化

在设计分布对时, 我们认为可以根据以下几条原则:

- (1) 很明显, $\lambda(1) = 1, \rho(1) = 1$;
- (2) 校验点的分布相对简单, 不宜太广, 通常选取 1~3 个相邻度。更多地, 在变量点的分布上做变化;
- (3) λ_2 的取值不能超过最大值 λ_2^* [3]。因为不规则 LDPC 码具有波浪效应: 一部分点的约束度高, 使得译码相对快速得多, 可以为其它低约束度点提供相对准确的信息, 最终使

得译码更准确快速。如果低约束度的点过多,将会导致译码失败。

设计出一个分布对以后,利用密度进化原理求得门限值;再设计一个与之相差不太大的新分布对,求其门限值。目标是要寻找门限值大的分布对。

分布对的优化需要很大的计算量,但相比仿真测量误码率与信噪比的曲线图来寻找好的分布对,计算量已经大大减少了。密度进化使得LDPC码设计不再只依靠系统仿真来评估码字性能的好坏,同时可以主动地预先地进行有关编解码理论分析和性能设计,为优化设计LDPC码提供了简便途径。

3 算法实现

密度进化方法在针对具体信道应用时,有各种精确和简化的形式,针对不同的信道模型其算法实现有所差异。

3.1 离散密度进化

为了使得密度进化方法便于计算机仿真,在原计算过程的每一步里都引入离散化函数 $Q(x)$,

$$Q(x) = \begin{cases} \lfloor x/\Delta + 1/2 \rfloor \cdot \Delta, & x \geq \Delta/2 \\ 0, & -\Delta/2 < x < \Delta/2 \\ \lceil x/\Delta + 1/2 \rceil \cdot \Delta, & x \leq -\Delta/2 \end{cases}$$

其中 Δ 为量化间隔, $\lfloor x \rfloor$ 表示下取整, $\lceil x \rceil$ 表示上取整。消息和其概率经过离散化后,消息 v 的卷积计算可以由快速傅里叶变换(FFT)来加速实现。 u 的进化计算量较大,可以预先存储 $\gamma(x,y)$ 的输入输出表,通过查表来节省计算时间。

量化比特越多,离散密度进化的精度越高,复杂度迅速增大。但FFT和查表方式的引入使得计算机可以快速完成运算,具有很强的可操作性。离散密度进化是确定运算,没有任何近似行为,计算得出的门限值与真实值误差非常小。采用14bit量化时,其门限值精度可以达到 10^{-6} ^[1]。并且该算法适用于各种对称信道,例如BSC,BEC,BiLaplace和AWGN信道。

3.2 AWGN信道下的高斯近似

当信道模型为高斯信道时,有如下近似计算方法^[4]。高斯信道下,消息的初始概率服从高斯分布 $N(m, \sigma^2)$,并且 $m = \sigma^2/2$ 。因此,得知均值 m 后可以完整描述概率密度的整个波形。密度进化的过程中, p_v^i 仍然服从高斯分布; p_u^i 虽不是高斯的,但其波形可以近似为高斯分布,其误比特率的变化非常小。高斯近似的基本思想是把消息的密度简化为高斯分布来进行分析,将多维运算转化为求高斯分布均值的一维运算。

m_u 和 m_v 分别代表 u 和 v 的均值。高斯分布叠加之后,其和的均值等于原均值的叠加,故 m_u^i 的进化规则为 $m_u^i = m_0 + (d_v - 1)m_u^{i-1}$ 。其中 $m_0 = 2/\delta^2$ 为初始消息 p_0 的均值;而 $m_u^0 = 0$,因为从任何校验点来的初始消息均值均为0。

m_u^i 可以通过在式(2)两边加数学期望得到:

$$m_u^i = \phi^{-1} \left(1 - \left[1 - \phi \left(m_0 + (d_v - 1)m_u^{i-1} \right) \right]^{d_c - 1} \right)$$

$$\text{其中 } \phi(x) = \begin{cases} 1 - \frac{1}{\sqrt{4\pi x}} \int_{\mathbb{R}} \tanh \frac{u}{2} e^{-\frac{(u-x)^2}{4x}} du, & x > 0 \\ 1, & x = 0 \end{cases}, \text{ 该函数用}$$

于计算 $1 - E \left[\tanh \frac{x}{2} \right]$ 。当 $x < 10$, $\phi(x) \sim e^{\alpha x^\gamma + \beta}$ (其中 $\alpha =$

$-0.4527, \beta = 0.0218, \gamma = 0.86$)。当 $x \geq 10, \phi(x) \sim \sqrt{\frac{\pi}{x}} \exp(-\frac{x}{4})$

$\cdot (1 + \frac{1}{14x} - \frac{3}{2x})$ 。

再引入分布对 (λ, ρ) ,不规则LDPC码的高斯近似算法为

$$m_u^i = \sum_{j=2}^{d_c} \rho_j \phi^{-1} \left(1 - \left[1 - \sum_{i=2}^{d_v} \lambda_i \phi \left(m_0 + (i-1)m_u^{i-1} \right) \right]^{j-1} \right)$$

高斯近似是密度进化的一种近似方法,尽管所获得门限值与离散密度进化获得门限值有差异,但算法复杂度低,是LDPC码的设计分析中常用的方法。它是针对消息密度呈高斯分布的情况,故而只适用于高斯信道。

4 仿真结果

在表1给出的度的分布对中,变量点最大度 $d_v = 20, 30, 50, 80, 100, 200$,码率等于0.5。 σ_{GA} 是通过高斯近似求得的门限值; σ_{exact} 是通过离散密度进化算法求得的门限值,量化比特数为9。对每个 d_v 值我们都构造了很多分布对,表1列出来的,是仿真求得门限值最大的那个分布对。很明显, σ_{GA} 比 σ_{exact} 要略小。

5 结束语

不规则LDPC码是当前高效纠错码研究的一个热点,它的优异性能依赖于好的度的分布对。密度进化能够定量地确定分布对的门限值,为寻找优秀的分布对提供了方向。这对不规则LDPC码的优化设计具有重要的指导意义。密度进化方法也可应用于其他码型,如Turbo码, MN码等的优化设计及分析。

表1 一些好的度的分布对和相应的利用离散密度进化算法和高斯近似求得的门限值

| d_r | 20 | 30 | 50 | 80 | 100 | 200 |
|-----------------|---------|---------|---------|---------|---------|---------|
| λ_2 | 0.22537 | 0.21192 | 0.17230 | 0.17005 | 0.17001 | 0.15640 |
| λ_3 | 0.13905 | 0.14316 | 0.15506 | 0.14294 | 0.16006 | 0.12758 |
| λ_4 | 0.11921 | 0.11539 | 0.11515 | 0.09358 | 0.10395 | 0.10591 |
| λ_6 | 0.10151 | | | | | |
| λ_9 | 0.10020 | 0.00615 | 0.09100 | | | |
| λ_{10} | 0.09627 | 0.09246 | 0.11264 | 0.09528 | 0.03519 | |
| λ_{11} | | 0.10009 | 0.03227 | 0.09184 | | 0.11023 |
| λ_{12} | | | 0.05062 | 0.08191 | 0.10794 | 0.05267 |
| λ_{19} | | 0.00580 | | | | |
| λ_{20} | 0.21839 | 0.14162 | | | | 0.09399 |
| λ_{30} | | 0.18341 | 0.06835 | 0.07057 | 0.08053 | |
| λ_{50} | | | 0.20261 | 0.08274 | 0.08647 | 0.08153 |
| λ_{80} | | | | 0.17109 | 0.06458 | 0.06589 |
| λ_{100} | | | | | 0.19127 | 0.04566 |
| λ_{200} | | | | | | 0.17014 |
| ρ_8 | 0.54768 | | | | | |
| ρ_9 | 0.45232 | 0.67751 | 0.06668 | | | |
| ρ_{10} | | 0.32249 | 0.93332 | 0.28034 | 0.11964 | |
| ρ_{11} | | | | 0.71966 | 0.88036 | 0.09165 |
| ρ_{12} | | | | | | 0.90835 |
| σ_{GA} | 0.9384 | 0.9396 | 0.9431 | 0.9452 | 0.9493 | 0.9519 |
| σ_{exit} | 0.9642 | 0.9681 | 0.9716 | 0.9726 | 0.9735 | 0.9757 |

参考文献

- [1] Chung S Y, Forney J G D, Richardson T, Urbanke R. On the design of low-density parity-check codes within 0.0045 dB of the Shannon limit. *IEEE Commun. Lett.*, 2001, 5(2): 58 – 60.
- [2] Richardson T, Urbanke R. The capacity of low-density parity-check codes under message-passing decoding. *IEEE Trans. on Info. Theory*, 2000, 47(2): 599 – 618.
- [3] Richardson T, Shokrollahi A, Urbanke R. Design of capacity-approaching low-density parity-check codes. *IEEE Trans. on Info. Theory*, 2001, 47(2): 619 – 637.
- [4] Chung S Y, Urbanke R, Richardson T. Gaussian approximation for sum-product decoding of low-density parity-check codes. In *Proc. Int. Symp. Info. Theory*, Sorrento, Italy, Jun, 2000: 318.
- 肖娟: 女, 1981年生, 硕士生, 主要研究方向为移动通信系统、高效纠错编码技术。
- 王琳: 男, 1963年生, 博士, 教授, 主要研究方向为移动通信系统、高效纠错编码技术。