

## 基于 LDPC 码的数字喷泉编码

林广荣<sup>①</sup> 林新荣<sup>②</sup> 依那<sup>①</sup> 金野<sup>①</sup> 梁庆林<sup>①</sup>

<sup>①</sup>(北京大学卫星与无线通信实验室 北京 100871)

<sup>②</sup>(哈尔滨工程大学自动化学院 哈尔滨 150001)

**摘要:** 数字喷泉是近年发展起来的一类基于删除信道、面向数据分组的前向纠错编码技术,在广播类传输中具有明显的效率优势。该文提出了一种基于 LDPC 码的数字喷泉编解码方案,仿真结果表明,与已有的同类数字喷泉编解码算法相比,基于 LDPC 码的数字喷泉编解码算法具有更好的纠错性能和较低的编解码复杂度。

**关键词:** 低密度奇偶校验码; 数字喷泉; 前向纠错编码

**中图分类号:** TN911.22

**文献标识码:** A

**文章编号:** 1009-5896(2008)04-0822-04

## Digital Fountain Base on LDPC Code

Lin Guang-rong<sup>①</sup> Lin Xin-rong<sup>②</sup> Yi Na<sup>①</sup> Jin Ye<sup>①</sup> Liang Qin-lin<sup>①</sup>

<sup>①</sup>(Satellite & Wireless Communication Laboratory, Peking University, Beijing 100871, China)

<sup>②</sup>(College of Automation, Harbin Engineering University, Harbin 150001, China)

**Abstract:** Digital fountain is a packet oriented forward error correction technology for channels with erasures, which has obvious efficiency advantage in broadcast. In the paper, LDPC code is put into practice in digital fountain encoding, and a fast decoding algorithm is presented. The simulation show that the digital fountain scheme base on LDPC code has better error correcting performance and can be implemented with very low complexity to encode and decode.

**Key words:** LDPC(Low Density Parity Check) code; Digital Fountain; forward error correction code

### 1 引言

数字喷泉是近年发展起来的一类前向纠错编码技术,它将带有检错机制的分组交换信道等效为删除信道,将传统纠错编码的处理对象拓展到了数据分组,形成了一种高效可靠的前向纠错机制<sup>[1]</sup>。与传统物理层的 FEC 相比,数字喷泉更加简单灵活,可以针对具体应用的不同需求,实现端到端的可靠传输;与链路层的 ARQ(自动重传请求)机制相比,数字喷泉无需使用反向信道,在广播类传输应用中具有明显的效率优势,拥有良好的发展前景。

最早的数字喷泉方案主要是借鉴了信道编码领域分组码的成果,例如 MDS(最大距离分离)码由于具有最大可能的最小汉明距离,从码距的意义上来说已经达到了最优,应用到数字喷泉领域以后,同样也具有很强的纠错能力;但是 MDS 码的复杂度相当高,编解码运算量近似与码长的平方以及符号域的大小成正比,这使得其很难在数字喷泉领域直接应用<sup>[2]</sup>。

为了解决复杂度的问题,随后又出现了以 Tornado 码为代表的稀疏图码<sup>[3]</sup>。这类码的生成图为稀疏图,大大降低了编解码的运算量,而专门构造的生成图则保证了 Tornado 码仍可以达到较好的纠错性能。

本文将另一种性能优异的稀疏图码 LDPC 码<sup>[4]</sup>用于数字喷泉编码。本文的第 2 节首先从码的构造角度对 Tornado 码和 LDPC 码的纠错性能进行了分析,随后又比较了 MDS 码、Tornado 码和 LDPC 码的编解码复杂度,最后给出了删除信道下基于 LDPC 码的数字喷泉方案的编解码算法。本文的第 3 节给出了多种码率和码长下 LDPC 码与 MDS 码和 Tornado 码的解码仿真曲线,验证了第 2 节的分析结果。

### 2 基于 LDPC 码的数字喷泉编码方案

#### 2.1 两种稀疏图码的构造比较和纠错性能

LDPC(Low Density Parity Check)码是稀疏图码的一种,同时也是一种特殊的线性分组码,它的一致校验矩阵  $H$  中“1”元素的密度很低且没有四元环。通过置信度传播迭代译码算法,LDPC 码可以取得接近香农限的译码效果<sup>[5]</sup>,近年来在许多实际通信系统中被广泛应用。

为了评价 LDPC 码用于数字喷泉编码的性能,下文将其与另一种稀疏图码 Tornado 码进行了对比。在 Tornado 码的构造思想中,为了利用 MDS 码的强纠错能力,同时避免复杂度过大,Tornado 码采用了级联图码的构造形式,将输入的源信息作为第一级二分图的左节点,通过编码得到右节点;再将右节点作为下一级二分图的左节点,持续级联编码;直到当二分图的节点数降低到满足下述条件时,使用 MDS 码进行最后一级编码,此时 MDS 码的码长应保证其复杂度满

足要求。Tornado 码的构造方式使得其最后一级子码拥有较强的纠错能力, 有效地降低了各级二分图右节点的删除概率。

但若与 LDPC 码相比较, Tornado 码相当于将单级二分图拆分为多级较小的子二分图, 所对应各级子码的码率均高于 Tornado 码整体的码率。而稀疏图码的码率会直接影响其解码性能曲线中瀑布区的位置, 因此在信道高误码条件下, 与 Tornado 码相比 LDPC 码能更快进入解码瀑布区。文献[6]给出了独立性假设下估算 Tornado 码瀑布区位置的公式

$$\begin{aligned} \text{per}_{\text{Tornado}} &= \delta \sum_i p_i [1 - (1 - \delta') \rho(1 - y_i |_{l \rightarrow \infty})]^i \quad (1) \\ y_l &= \delta \lambda (1 - (1 - \delta') \rho(1 - y_{l-1})), \quad y_0 = \delta \\ \lambda(x) &= \sum_i \lambda_i x^{i-1}, \quad \rho(x) = \sum_j \rho_j x^{j-1} \\ \lambda_i &= i \left( p_i / \sum_j p_j \right) p_j, \quad \rho_j = j \left( q_j / \sum_i q_i \right) q_i \end{aligned}$$

其中  $\delta$  和  $\delta'$  分别为 Tornado 码单级二分图中左、右节点的删除概率,  $p_i$  为左节点维度取  $i$  的概率,  $q_j$  为右节点维度取  $j$  的概率。由式(1)可得 Tornado 码每级子码的误包率, 进而计算出 Tornado 码整体的误包率。

类似地, 可以求得独立性假设下删除信道中 LDPC 码的误包率公式为

$$\begin{aligned} \text{per}_{\text{LDPC}} &= \delta \sum_i p_i [1 - \rho(1 - y_i |_{l \rightarrow \infty})]^i \quad (2) \\ y_l &= \delta \lambda (1 - \rho(1 - y_{l-1})), \quad y_0 = \delta \\ \lambda(x) &= \sum_i \lambda_i x^{i-1}, \quad \rho(x) = \sum_j \rho_j x^{j-1} \\ \lambda_i &= i \left( p_i / \sum_j p_j \right) p_j, \quad \rho_j = j \left( q_j / \sum_i q_i \right) q_i \end{aligned}$$

其中  $\delta$  为 LDPC 码二分图变量节点的删除概率,  $p_i$  为变量节点维度取  $i$  的概率,  $q_j$  为校验节点维度取  $j$  的概率。由式(1), 式(2)可以验证, 在码率和平均维度均相同时, LDPC 码解码时能更快收敛进入解码瀑布区。以 LDPC 码取 1/2 码率、变量节点平均维度 3.42(具体维度分布详见 3.3 节)为例, 在信道删除概率为 46.2% 时即进入瀑布区(信道删除概率低于 46.2% 时输出误包率即快速收敛到零); 而取相近参数的 Tornado 码(码率 1/2、左节点平均维度 3.44、具体维度分布详见 3.3 节)则需信道删除概率为 36.8% 时才进入瀑布区。

式(1), 式(2)计算的是给定维度分布后, 稀疏图码在码长无限时最优构造所能达到的理想性能。但由于图码码长有限, 实际性能相比该理论值会有一定程度的恶化, 且码长越短其性能偏离理论值越大。可以预期, 在参数相当条件下, Tornado 码由于采用了多级子码的构造形式, 其实际性能与解码瀑布区理论值的偏差会更大。

## 2.2 编解码复杂度

在各类数字喷泉实现方案中, MDS 码尽管具有最大的码距, 但其编解码的复杂度非常高。例如, 专门针对复杂度

进行过优化的  $(n, k)$  RS-Cauchy 码编码复杂度为  $O((n-k)kl)$  [2], 解码复杂度为  $O(xkl) + O(x^2l)$ , 其中  $l$  为包长、 $x$  为被删除包数, 即其编解码复杂度近似与码长的平方及包长成正比。而对于 Tornado 码和 LDPC 码, 若假定该码的平均维度为常数  $d$ , 则编解码的复杂度近似为  $O(ndl)$  [1], 即其编解码复杂度近似与码长及包长成正比。因此, 在码长较大时, MDS 码因编解码复杂度过高, 往往很少被采用; Tornado 码和 LDPC 码具有较低的复杂度, 而在这两种稀疏图码的平均维度相当条件下, Tornado 码由于具有多级串联的复杂结构, 难以并行编解码, 实际的编解码时间会略长。

## 2.3 基于 LDPC 码的数字喷泉的编解码

LDPC 码的构造和编码过程可以用二分图的方式来表示。如图 1, 图中的节点分为两种类型, 左边的节点称为变量节点, 每个变量节点与 LDPC 码的一个码元相对应, 在数字喷泉的编码过程中, 每个数据分组对应一个变量节点; 右边的节点称为校验节点, 与校验方程对应。每个校验节点所相连的变量节点需要满足校验关系, 定义编解码过程中的基本运算为逐比特异或, 则上述变量节点对应的数据分组应满足逐比特异或后和为零。本文讨论的 LDPC 码为系统码, 因此编码输出数据分组中有部分为原始的输入数据分组, 剩余冗余部分的数据分组可以通过校验关系编码得到。

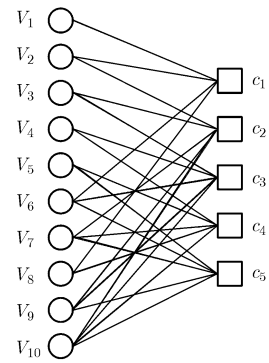


图 1 LDPC 码二分图

完成编码后, 在将新的数据分组通过信道发送前, 需要为每个分组加上校验和。若不考虑校验和的漏检概率, 则接收端可以根据计算比较校验和判定收到的每个数据分组正确与否。因此, 可以将分组交换的逻辑信道等效为删除信道。

相应地, 在接收端, 由于能够确知接收到数据分组的正确与否, 每个变量节点的置信度取值也就退化为一个二值函数, 从而可以将置信度传播算法大幅度简化。具体解码算法如下:

定义  $V_i$ ,  $V_j$  分别为 LDPC 码二分图中变量节点和校验节点的值,  $D_j$  为校验节点的维度,  $d_j$  为校验节点在解码过程中的有效维度(指与其相连但仍未处理过的变量节点的数量),  $C_i$  为变量节点对应数据分组的校验和。

步骤 1 初始化, 令  $V_j = 0$ ,  $d_j = D_j$ ,  $j = 1, 2, \dots, n-k$ ;

令集合  $A = \emptyset$ , 集合  $B = \emptyset$ , 对  $i = 1, 2, \dots, n$ , 如果  $C_i = 0$ , 说明变量节点  $i$  正确接收, 则  $A = A \cup \{i\}$ ; 反之, 如果  $C_i \neq 0$ , 则变量节点  $i$  被删除,  $B = B \cup \{i\}$ 。

步骤2  $\forall a \in A, j = 1, 2, \dots, n - k$ , 如果变量节点  $a$  与校验节点  $j$  相连, 则  $V_j = V_j \oplus V_a, d_j = d_j - 1$ , 其中  $\oplus$  为逐比特异或算符; 如果  $d_j = 1$ , 且校验节点  $j$  的邻点  $b \in B$ , 则  $V_b = V_j, B = B \setminus b, A = A \cup \{b\}$ ; 对  $j$  循环完毕后, 令  $A = A \setminus a$ 。

步骤3 重复步骤2, 直至  $A = \emptyset$ 。

此时若  $B \neq \emptyset$ , 则表明未能完全正确解码, 集合  $B$  中变量节点所对应的数据分组为错误分组。

### 3 LDPC 方案性能仿真及比较

#### 3.1 仿真环境

在仿真中, 设定数据分组的长度为 8000 bit, LDPC 码的二分图使用 PEG 算法进行构造; 假定传输过程中未使用物理层的信道编解码, 信道误码为独立的随机错误; 不考虑数据分组检错时发生漏检的情况。

仿真时主要关注的参数是解码失败概率和解码后的误包率, 其中解码失败概率是指解码失败的次数占总解码次数的比例。在单次解码过程中, 解码结束后只要仍存在任一数据包被删除, 即认为该次解码失败。由于数据文件在传输时, 不能允许出现任何错误, 因此解码失败概率是衡量数字喷泉算法性能的重要指标。

#### 3.2 相同码率、码长的 MDS 码、Tornado 码和 LDPC 码解码失败概率对比

图 2 给出了码长为 1008(Tornado 码码长为 1024)、码率为 3/4 的 MDS 码, Tornado 码和 LDPC 码在不同 BER 下的解码失败概率曲线。其中 LDPC 码的变量节点维度分布为  $0.4x + 0.3x^2 + 0.3x^6$ , 二分图的构造采用 PEG 算法; Tornado 码二分图的左节点维度分布为  $0.5833x^2 + 0.1944x^3 + 0.0972x^4 + 0.0583x^5 + 0.0389x^6 + 0.0278x^7$ , 右节点维度分布及构造参见文献[1]。

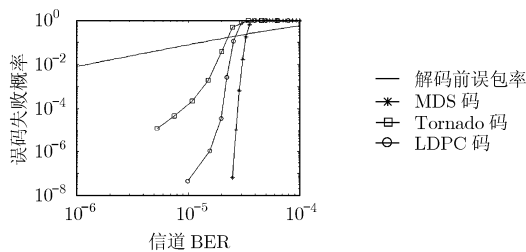


图2 MDS 码, Tornado 码和 LDPC 码 BER-解码失败概率曲线

对于  $(n, k)$  MDS 码而言, 只要接收到的数据分组数不少于  $k$  个, 就能够成功解码。若不能成功解码, 输出误包率则取决于具体的码型, 因此本文在比较其他码型与 MDS 码的性能时, 仅比较解码失败概率。从图中可以看出, 若要使解

码失败概率低于  $10^{-5}$ , Tornado 码、LDPC 码和 MDS 码需要信道 BER 分别低于  $5.0 \times 10^{-6}$ 、 $1.81 \times 10^{-5}$  和  $2.69 \times 10^{-5}$ , 因此, LDPC 码的解码性能略低于 MDS 码, 而优于 Tornado 码。

#### 3.3 相同码率下 Tornado 码和 LDPC 码解码后误包率的对比

图 3 和图 4 是码率分别为 3/4 和 1/2 的 Tornado 码和 LDPC 码在两种码长下解码后误包率随信道 BER 变化的曲线。图 3 码率为 3/4, 其中 Tornado 码 1、2 的码长分别为 1024 和 4096, 平均维度为 3.86(同 3.2 节), LDPC 码 1、2 的码长分别为 1008 和 4032, 平均维度为 3.8; 图 4 码率为 1/2, Tornado 码与 LDPC 码的平均维度分别为 3.44 和 3.42 (LDPC 码的变量节点维度分布为  $0.5x + 0.27x^2 + 0.23x^6$ , Tornado 码的左节点维度分布为  $0.6667x^2 + 0.2222x^3 + 0.1111x^4$ ), Tornado 码 1、2 的码长分别为 1024 和 4096, LDPC 码 1、2 的码长分别为 1008 和 4032。

图 3 和图 4 同时给出了根据式(1)和式(2)求得理想解码瀑布区曲线。可以看出, 在码率与平均维度均相当时, LDPC 码的理想性能界优于 Tornado 码, 其实际仿真结果与理论界的偏差亦比 Tornado 码要小, 均符合 2.1 节中的分析结果。在两种码率下, 随着码长的增加, 两种稀疏图码的解码性能均有较明显的改善, 更加趋近于理想的解码瀑布区曲线。在码率、码长与平均维度均相当时, LDPC 码的解码性能均优于 Tornado 码。

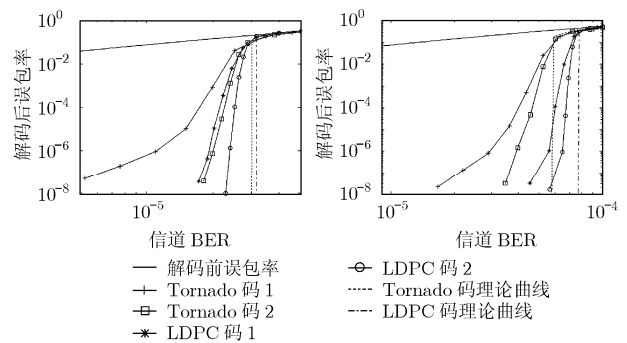


图3 3/4 码率的 Tornado 码和 LDPC 码 BER-误包率曲线

图4 1/2 码率的 Tornado 码和 LDPC 码 BER-误包率曲线

### 4 结束语

综上所述, LDPC 码与 MDS 码相比, 尽管解码性能略低, 但显著降低了编解码的复杂度; 与 Tornado 码相比, 在码率、码长和平均维度相当条件下, LDPC 码的解码失败概率和解码后误包率均优于 Tornado 码。因此, 基于 LDPC 码的数字喷泉编码兼具有良好的纠错性能和较低的编解码复杂度, 是一种优秀的数字喷泉实现方案。

### 参考文献

[1] Byers J W, Luby M, and Mitzenmacher M, et al. A digital

- fountain approach to reliable distribution of bulk data. Proc. ACM SIG-COMM, Vancouver, BC, Canada, 1998: 56-67.
- [2] Blomer J, Kalfane M, and Karpinski M, *et al.* An XOR-based erasure-resilient coding scheme. ICSI, Berkeley, CA, 1995: 48-67.
- [3] Luby M, Mitzenmacher M, and Shokrollahi A, *et al.* Practical loss-resilient codes. Proc. 29th ACM Symposium on Theory of Computing, Texas, 1997: 150-159.
- [4] Gallager R G. Low-density parity-check codes. Cambridge, MA: MIT, 1963.
- [5] Chung S Y, Forney G D, and Richardson T J, *et al.* On the design of low-density parity-check codes within 0.0045 dB of the Shannon limit. *IEEE Commun. Lett.*, 2001, Vol. 5: 58-60.
- [6] Luby M, Mitzenmacher M, and Shokrollahi A. Analysis of random processes via and-or tree evaluation. Proc. 9th Annu. ACM-SIAM Symp. Discrete Algorithms, San Francisco, CA, 1998: 364-373.
- 林广荣: 男, 1980年生, 博士生, 研究方向为信道编码与计算机网络.
- 林新荣: 女, 1976年生, 硕士生, 研究方向为计算机网络与自动控制.
- 依那: 男, 1974年生, 讲师, 研究方向为卫星与无线通信网络.
- 金野: 男, 1966年生, 教授级高工, 博士生导师, 研究方向为卫星与无线通信.
- 梁庆林: 男, 1941年生, 教授, 博士生导师, 研究方向为卫星与无线通信.