

# 基于度分布合并和可译集优化的 LT 码度分布设计方案

雷维嘉<sup>1</sup>, 张 梦<sup>1,2</sup>, 谢显中<sup>1</sup>

(1. 重庆邮电大学移动通信技术重庆市重点实验室, 重庆 400065; 2. 重庆工商大学融智学院, 重庆 400033)

**摘 要:** 喷泉码的度分布对喷泉码的编译码性能有着十分重要的影响. 目前常用的度分布在源数据码长较长时具有较好的性能, 但在码长较短时, 性能有明显下降. 本文给出一种 LT 码度分布的构造方法. 该方法先对二进制度分布进行调整, 然后将其与鲁棒孤子度分布进行有机结合, 再通过优化可译集合值来进一步优化度分布函数, 得到一种当源数据为短码长时也有较好性能的度分布, 即修正二进制-鲁棒孤子度分布. 仿真结果证明, 采用这种度分布对源数据进行 LT 编码时, 相比较二进制度分布和鲁棒孤子度分布, 其译码性能得到了明显提高, 并且码长越短, 性能提高越明显.

**关键词:** 数字喷泉码; LT 码; 度分布; 可译集合

**中图分类号:** TN911.22

**文献标识码:** A

**文章编号:** 0372-2112 (2015)04-0800-06

**电子学报 URL:** <http://www.ejournal.org.cn>

**DOI:** 10.3969/j.issn.0372-2112.2015.04.025

## A Design Scheme for LT Codes Degree Distribution by Combining Degree Distributions and Optimizing Ripple Size

LEI Wei-jia<sup>1</sup>, ZHANG Meng<sup>1,2</sup>, XIE Xian-zhong<sup>1</sup>

(1. Chongqing Key Lab of Mobile Communications Technology, Chongqing University of Posts and Telecommunications, Chongqing 400065, China;

2. Rongzhi College of Chongqing Technology and Business University, Chongqing 400033, China)

**Abstract:** The degree distribution has a very important influence on the decoding efficiency of fountain codes. The decoding performance is good when the message size is large, but it drops dramatically when the size becomes smaller, which is the inevitable result of using the existing degree distributions to encode. In this paper, we present a new scheme to design the degree distribution. First, we combine the robust soliton degree distribution with the modified binary exponential distribution and further adjust the degree by optimizing the ripple size. By doing this we can get modified binary robust distribution (MBRD), which can also perform well when the source message size is small. Through simulation experiments, we have verified that MBRD greatly improves the encoding and decoding performances. Moreover, compared with the existing degree distributions, the smaller the source message size becomes, the more evidently the performance is improved.

**Key words:** fountain codes; LT codes; degree distribution; ripple size

## 1 引言

喷泉码的概念是由 M Luby 等人<sup>[1]</sup>于 1998 年首次提出, 但是当时并没有给出可实行的喷泉码设计方案. 2002 年 Luby<sup>[2]</sup>提出了第一种现实可行的喷泉码—LT 码. 发送端对传输的信息进行 LT 喷泉编码, 得到的编码信息长度远大于原始信息长度, 接收端只要接收到足够多的编码信息就能将原始信息译出<sup>[3]</sup>. 其编码过程为: 发送端原始数据由  $k$  个数据包组成, 根据某一度分布函数随机产生编码数据包的度  $d$ , 然后从  $k$  个原始数据

包中随机地选择  $d$  个数据包, 再将这  $d$  个数据包进行异或运算, 从而生成一个编码数据包. 因此, 喷泉码的编码度分布对喷泉码的编译码性能有着重要的影响. Luby 在文献<sup>[2]</sup>中提出了两种度分布, 分别为理想孤子度分布和鲁棒孤子度分布. 理想孤子度分布理论上使每一个编码数据包在每一次译码迭代中释放的概率相同<sup>[2]</sup>, 在每次迭代中有且只有一个度为 1 的编码数据包, 完成每一次迭代都有一个源数据包恢复, 然后又有新的度为 1 的编码数据包出现. 然而在实际中, 由于度分布映射的随机性, 很容易造成度 1 编码数据包消失, 从而导致译

码过程无法继续进行.鲁棒孤子度分布(Robust Soliton Distribution, RSD)在理想孤子度分布上做出了改进,设计了一个  $\tau$  函数,并将  $\tau$  函数与理想孤子度分布进行归一化得到度分布. $\tau$  函数主要增加了较大度取值的概率,并引入 2 个参数  $c$  和  $\delta$  来确保译码过程中期望度为 1 的编码数据包的个数.RSD 在码长达到  $10^4$  甚至更高量级时能获得很好的性能,但在源数据码长较短的情况下,RSD 的编码译码效率下降明显<sup>[4]</sup>.

一个好的编码度分布会有一些特点,根据这些特点来设计的喷泉码编码度分布会得到较好的性能.文献[5~7]表明度为 2 的概率值  $p_2$  应为整个度分布函数中的最大值,且应尽量接近 0.5;度为 1 的概率值  $p_1$  要适当大,以保障译码过程的开始及持续进行;还要保证一定的最大度的概率值  $p_{\max}$ ,以确保编码过程能覆盖全部源数据包.因此,这 3 个度的概率值对度分布函数的性能有着重大的影响.在设计度分布时,除了要考虑以上几个概率值的要求外,还需要考虑可译集合的值.可译集合是进行每一步译码时产生的度为 1 的编码数据包的集合<sup>[8]</sup>.当源数据包码长很长时,可译集合的值在译码过程中一般都趋于一个恒值;在源数据包码长较短时,可译集合的值在译码过程中会出现一定的波动.当可译集合的值较小时,即度 1 编码数据包的个数较少,有可能导致译码过程中断,需要得到更多的编码数据包才能进行译码,这样译码开销必然增加.可译集合值跟度分布函数有密切联系,通过优化可译集合值可进一步调整度分布函数.

文献[9]对用鲁棒孤子度分布对短码长源数据进行 LT 编码以及译码的特性进行了详细的分析,并通过改变鲁棒孤子度分布的 2 个参数  $c$  和  $\delta$  的值,来优化其编译码性能.但该文没有对度分布函数进行改进.文献[10]中,度的取值仅为  $\frac{k}{2}$  和  $\frac{k}{2} - 1$ ,  $k$  为源数据包个数,并按照一定的顺序选取源数据包进行编码.当在传输中存在包删除的情况时,其译码效率恶化严重.文献[11]运用马尔科夫链的方法来设计度分布,但只适用于源数据包的个数非常小的情况.目前,对短码长喷泉码编码度分布的研究较少,现有的优化度分布方案仍有进一步改善的空间.

目前利用常用的二进制指数度分布<sup>[12]</sup>(Binary Exponential Distribution, BED)、鲁棒孤子度分布进行 LT 编码时,在短码长时其编码译码效率下降明显.针对于这种情况,本文给出了一种新的编码度分布的设计方案,根据此方案设计出的度分布当源数据为短码长时也有较好性能.首先,对二进制指数度分布进行调整,然后将 BED 与 RSD 归一化合并得到一种度分布.再通过对可译集合值的优化来进一步调整度分布函数中的  $p_1$ 、

$p_2$ 、 $p_{\max}$  的值,得到一种当源数据为短码长时也有较好性能的度分布.

## 2 三种常用度分度

### 2.1 理想孤子度分布

理想孤子度分布理论上使每一个编码数据包在每一次译码迭代中释放的概率相同<sup>[2]</sup>,保证在每次迭代过程中有且只有一个度为 1 的编码数据包出现,完成每一次迭代恢复一个源数据包,同时又有一个度为 1 的编码数据包出现.其度分布函数为:

$$\rho(d) = \begin{cases} 1/k, & d = 1 \\ 1/d(d-1), & d = 2, 3, \dots, k \end{cases} \quad (1)$$

式中, $\rho(d)$  为采用理想孤子度分布进行编码时,编码数据包的度为  $d$  ( $d = 1, 2, 3, \dots, k$ ) 的概率; $k$  为源数据包数量.由其函数表达式可以看出,理想孤子度分布产生的编码数据包中多数的度较小,其中产生度 1 数据包的概率仅为  $1/k$ .在实际中,由于度分布映射的随机性,在译码过程中很容易出现度 1 编码数据包消失;另一方面,产生度值较大编码数据包的概率也较小,喷泉码编码时全覆盖的概率较低,会出现不能完全译出所有源数据包的情况.

### 2.2 鲁棒孤子度分布

鲁棒孤子度分布(RSD)是对理想孤子度分布做出的改进.RSD 在度分布函数中引入了 2 个参数  $c$  和  $\delta$ ,通过  $c$  和  $\delta$  的选择来确保译码过程中期望度为 1 的编码数据包个数  $s$  为:

$$s = c \ln\left(\frac{k}{\delta}\right) \sqrt{k} \quad (2)$$

RSD 设计了一个  $\tau$  函数,目的是增加编码中取较大度值的概率,提高覆盖性. $\tau$  函数为:

$$\tau(d) = \begin{cases} \frac{s}{k} \frac{1}{d}, & d = 1, 2, \dots, (k/s) - 1 \\ \frac{s}{k} \log(s/\delta), & d = k/s \\ 0, & d > k/s \end{cases} \quad (3)$$

将  $\tau$  函数与理想孤子度分布函数进行合并,并归一化就得到鲁棒孤子度分布函数:

$$\mu(d) = \frac{\rho(d) + \tau(d)}{Z} \quad (4)$$

其中,

$$Z = \sum_d (\rho(d) + \tau(d)) \quad (5)$$

式中, $\mu(d)$  表示采用鲁棒孤子度分布进行编码时,编码数据包度为  $d$  的概率; $k$  为信源数据包数量; $\delta$  为译码器未能全恢复源信息的概率; $c$  为 0 和 1 之间的常数.采用鲁棒孤子度分布进行编码时,产生的编码数据包多为度较大的编码数据包,具有较高的覆盖性,但因此

增加了冗余度,造成译码效率的降低,而且产生度 1 和其他小度编码数据包的数量较少,当源数据包码长较短时,容易出现译码中断的情况。

### 2.3 二进制指数度分布

文献[12]中给出了二进制指数度分布(BED):

$$b(d) = \begin{cases} 1/2^d, & d = 1, 2, 3, \dots, k-1 \\ 1/2^{d-1}, & d = k \end{cases} \quad (6)$$

式中,  $b(d)$  表示采用二进制指数度分布进行编码时, 编码数据包度为  $d$  的概率;  $k$  为信源数据包数量. 采用二进制指数进行编码时, 编码数据包度为 1 的概率很大, 可以很好地保证译码的持续性. 但是随着  $d$  值的增加, 取得  $d$  值的概率呈指数减少, 所以产生度值较大的编码数据包的概率很小, 源数据包的充分覆盖得不到保证.

## 3 度分布优化方案

### 3.1 两种度分布的合并

通过分析理想孤子度分布、鲁棒孤子度分布和二进制指数度分布, 可以看出: RSD 产生度较大的编码数据包的概率较大, 但是产生度 1 的编码数据包的概率很小, 译码中断的可能性较高; BED 能够产生足够多的度 1 编码数据包, 能保证译码开始并持续地进行, 然而过多地产生度 1 编码数据包, 会降低全部信源数据包都参与编码的概率; 同时, 按照 BED 产生的编码数据包之间的关联性较小, 度为 2 的编码数据包在译码中度值降为 1 的概率相应也较小, 因而译码的迭代效率较低. 另外, BED 中度 1 的概率  $p_{\text{BED},1} = 0.5$  远远大于度 2 的概率  $p_{\text{BED},2} = 0.25$ , 并不符合一个好的度分布的特点. 因此, 我们在设计度分布时, 首先对 BED 做出调整, 将度 1 的概率值与度 2 的概率值互换, 即改为  $p_{\text{BED},1} = 0.25$ ,  $p_{\text{BED},2} = 0.5$ . 通过这种互换使得改变后的 BED 满足度 2 概率值应是整个度分布函数的最大值的要求. 调整后的 BED 的度 1 概率值仍然过大, 而取较大度的概率过小. 而 RSD 则取较大度的概率较大, 产生度 1 编码数据包的概率较小. 将调整后的 BED 与 RSD 进行合并并归一化, 可结合这两种度分布的优点, 形成一种新的度分布, 即修正二进制-鲁棒孤子度分布 (Modified Binary Robust Distribution, MBRD). 合并时通过调整 BED 与 RSD 的比例系数  $\alpha$  和  $\beta$  来形成最佳的度分布函数. 其概率分布函数如下:

$$m(d) = \frac{\alpha \cdot b(d) + \beta \cdot \mu(d)}{W} \quad (7)$$

$$W = \sum_d (\alpha \cdot b(d) + \beta \cdot \mu(d)) \quad (8)$$

式中,  $m(d)$  表示采用 MBRD 进行编码时, 编码数据包度为  $d$  的概率,  $\mu(d)$  为鲁棒孤子度分布,  $b(d)$  为二进

制指数度分布.

要从理论上推导得到最佳的比例系数  $\alpha$  和  $\beta$  的取值较为困难, 但可以先根据要求推导出它们的取值范围, 然后再利用 Monte-Carlo 方法确定其最佳值. 合并后的度分布中度 1 的概率值主要是由调整后的 BED 的度 1 概率值决定, 而其最大度概率值主要是由 RSD 的最大度概率值决定. 根据概率值的特性, 以及要获得好的度分布的要求,  $\alpha$  和  $\beta$  的取值范围应满足以下约束条件:

$$\begin{cases} p_{\text{MBRD},2} < 1 \\ 0 < p_{\text{MBRD},1} + p_{\text{MBRD},\max} < 0.5 \\ p_{\text{MBRD},1} > p_{\text{RSD},1} \end{cases} \quad (9)$$

其中,

$$p_{\text{MBRD},1} \approx p_{\text{BED},1} \cdot \alpha \quad (10)$$

$$p_{\text{MBRD},\max} \approx p_{\text{RSD},\max} \cdot \beta \quad (11)$$

$$p_{\text{MBRD},2} \approx \frac{p_{\text{RSD},2} \cdot \beta + p_{\text{BED},2} \cdot \alpha}{2} \quad (12)$$

即

$$\begin{cases} p_{\text{RSD},1} < \alpha < \frac{2 \cdot p_{\text{RSD},\max} - 0.5 \cdot p_{\text{RSD},2}}{p_{\text{BED},2} \cdot p_{\text{RSD},\max} - p_{\text{BED},1} \cdot p_{\text{RSD},2}} \\ 0 < \beta < \frac{2 \cdot p_{\text{BED},1} - 0.5 \cdot p_{\text{BED},2}}{p_{\text{RSD},2} \cdot p_{\text{BED},1} - p_{\text{RSD},\max} \cdot p_{\text{BED},2}} \end{cases} \quad (13)$$

$p_{\text{MBRD},1}$ 、 $p_{\text{MBRD},2}$ 、 $p_{\text{MBRD},\max}$  分别表示修正二进制-鲁棒孤子度分布度 1、度 2 和最大度的概率值,  $p_{\text{RSD},1}$ 、 $p_{\text{RSD},2}$ 、 $p_{\text{BED},2}$  分别表示鲁棒孤子度分布度 1、度 2, 以及二进制指数度分布度 2 的概率值.

### 3.2 度分布的进一步优化

为了使度分布在源数据包较短时有更好的编码译码性能, 我们还需要考虑译码过程中可译集合值的波动. 当可译集合的值较小时, 即度 1 编码数据包的个数较少, 这样译码中断的可能性较高, 译码开销增大. 因此, 需要足够大的可译集合值以保证译码过程的持续进行. 文献[8]给出了确定可译集合值的公式:

$$R_{\Omega}(\rho) = (1 + \epsilon) \left( \sum_{d=1}^{\text{dmax}} \Omega_d f(\rho, d) + \frac{k - \rho}{1 + \epsilon} \ln \frac{k - \rho}{k} \right) + O(1) \quad (14)$$

其中,

$$f(\rho, d) = (k - \rho) d \left( \frac{\rho}{k} \right)^{d-1} \quad (15)$$

$$\epsilon = \frac{N - k}{k} \quad (16)$$

式中,  $\rho$  表示成功译码的源数据包个数,  $k$  表示源数据包个数,  $N$  为译码成功时所需要的编码数据包个数,  $\Omega_d$  表示度为  $d$  的概率值,  $\epsilon$  为译码开销.  $O(1)$  为 1 的高阶无穷小量. 由此可见, 可译集合值与度分布有关, 可以通过调整度分布来改善可译集合值. 度分布函数中  $p_1$ 、 $p_2$ 、 $p_{\max}$  的取值对编译码性能的影响最大, 因此我们根

据可译集合值的要求进一步调整  $p_1$ 、 $p_2$ 、 $p_{\max}$  的值,而其他度的概率值保持不变.令  $p_1$ 、 $p_2$ 、 $p_{\max}$  的调整量分别为  $\Delta p_1$ 、 $\Delta p_2$ 、 $\Delta p_{\max}$ ,根据概率之和固定为一的特点,有

$$\Delta p_1 + \Delta p_2 + \Delta p_{\max} = 0 \quad (17)$$

由于这三个调整量间存在这样的约束关系,实际上优化的调整量只有两个,我们选择  $\Delta p_1$ 、 $\Delta p_2$ . 度分布调整后的可译集合值为:

$$\begin{aligned} R_{\Omega_m}(\rho) = & (1 + \epsilon) \left( \Delta p_1 f(\rho, 1) + \Delta p_2 f(\rho, 2) \right. \\ & \left. - (\Delta p_1 + \Delta p_2) f(\rho, d_{\max}) \right) \\ & + \sum_{d=1}^{d_{\max}} \Omega_d f(\rho, d) + \frac{k - \rho}{1 + \epsilon} \ln \frac{k - \rho}{k} \\ & + O(1) \end{aligned} \quad (18)$$

在整个译码过程中,可译集合值的平均值为:

$$\begin{aligned} M = & \frac{1}{k} \sum_{\rho=0}^{k-1} R_{\Omega_m}(\rho) \\ = & \frac{1 + \epsilon}{k} \left( \Delta p_1 \sum_{\rho=0}^{k-1} (f(\rho, 1) - f(\rho, d_{\max})) \right) \\ & + \Delta p_2 \sum_{\rho=0}^{k-1} (f(\rho, 2) - f(\rho, d_{\max})) \\ & + \sum_{\rho=0}^{k-1} \sum_{d=1}^{d_{\max}} \Omega_d f(\rho, d) + \sum_{\rho=0}^{k-1} \frac{k - \rho}{1 + \epsilon} \ln \frac{k - \rho}{k} \end{aligned} \quad (19)$$

方差为:

$$V = \frac{1}{k} \sum_{\rho=0}^{k-1} (R_{\Omega_m}(\rho) - M)^2 \quad (20)$$

改善可译集合值也就是使可译集合值的平均值最大,方差最小,即找到使下式取得最小值的  $\Delta p_1$ 、 $\Delta p_2$  值:

$$\eta = \lambda \times V - M \quad (21)$$

由式(17)以及度分布函数中每个概率值都大于 0 可以得出  $\Delta p_1$ 、 $\Delta p_2$  的约束条件为:

$$\begin{cases} -P_{\text{MBRD},1} \leq \Delta p_1 \leq P_{\text{MBRD},2} + P_{\text{MBRD},\max} \\ -P_{\text{MBRD},2} \leq \Delta p_2 \leq P_{\text{MBRD},1} + P_{\text{MBRD},\max} \\ \Delta p_1 + \Delta p_2 \leq P_{\text{MBRD},\max} \end{cases} \quad (22)$$

可译集值的优化涉及到求解最优值的数学方法,序列二次规划法(Sequential Quadratic Programming, SQP)

是一种求解约束优化问题的有效算法. SQP 的基本思想是:在每一迭代步中通过求解一个二次规划子问题来确立一个下降方向,以减少价值函数来取得步长,重复这些步骤直到求得原问题的解. 详细的求解思路以及步骤可参考文献[13].

我们再把度分布的优化过程归纳如下:

**步骤 1** 将 BED 度分布中度 1、2 的概率值交换;

**步骤 2** 将调整后的 BED 与 RSD 按照一定组成比例进行归一化合并,确定比例系数  $\alpha$  和  $\beta$  的取值;

**步骤 3** 通过优化可译集合值得到  $p_1$ 、 $p_2$ 、 $p_{\max}$  的调整量  $\Delta p_1$ 、 $\Delta p_2$ 、 $\Delta p_{\max}$ ;

**步骤 4** 将进行步 2 后的度分布函数中的  $p_1$ 、 $p_2$ 、 $p_{\max}$  更新为  $\Delta p_1 + p_1$ 、 $\Delta p_2 + p_2$ 、 $\Delta p_{\max} + p_{\max}$ ,得到优化的度分布函数.

## 4 仿真结果及分析

本节对修正二进制-鲁棒孤子度分布进行仿真分析和验证. 首先利用 Monte-Carlo 方法得到比例系数  $\alpha$  和  $\beta$ ,然后给出优化度分布的可译集合值的仿真结果;最后给出优化度分布的译码性能. 文献[9]中给出了鲁棒孤子度分布参数  $c$  的取值范围, $c$  在这范围内取值可以获得较好的性能. 本节中修正二进制-鲁棒孤子度分布的参数选取为: $c = 0.02$ ,  $\delta = 0.01$ ,  $\epsilon = 0.1$ . LT 码的译码采用置信传播算法<sup>[5]</sup>.

### 4.1 比例系数 $\alpha$ 和 $\beta$ 的确定

根据式(9)~(13)计算得到源数据包长度  $k$  取 200、500、1000 时,比例系数  $\alpha$ 、 $\beta$  的取值范围分别为: $0.066 < \alpha < 1.09$ 、 $0 < \beta < 3.29$  ( $k = 200$ ),  $0.042 < \alpha < 1.38$ 、 $0 < \beta < 2.88$  ( $k = 500$ ),  $0.03 < \alpha < 1.53$ 、 $0 < \beta < 2.67$  ( $k = 1000$ ). 然后在该范围内,调整  $\alpha$ 、 $\beta$  的取值,得到不同的度分布函数,用其进行 LT 编码并译码,根据仿真得到译码完成时所需要的编码数据包的数量来确定最优的  $\alpha$ 、 $\beta$  值. 图 1 为信源数据包数量  $k = 200$ 、500、1000 时的仿真结果. 可以看出,在三种  $k$  值下,当  $\alpha$ 、 $\beta$  的取值分别为  $\alpha = 0.5$ 、 $\beta = 2$ ,  $\alpha = 0.3$ 、 $\beta = 2.6$ ,  $\alpha = 0.3$ 、 $\beta = 2.4$  时性能最好.

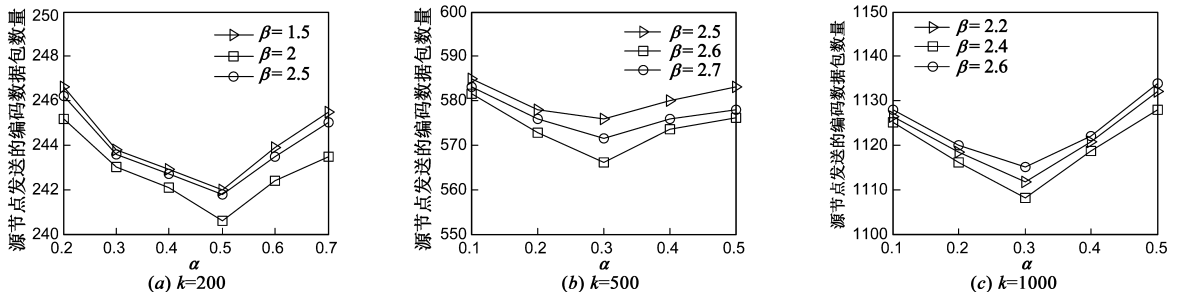


图1 修正二进制-鲁棒孤子度分布:译码成功时所需要的编码数据包数随比例系数 $\alpha$ 和 $\beta$ 的变化情况

## 4.2 可译集合值的仿真

在确定比例系数  $\alpha$  和  $\beta$  后,进一步的工作是调整  $p_1$ 、 $p_2$ 、 $p_{\max}$  的取值.通过求解式(21)、(22)得到  $p_1$ 、 $p_2$ 、 $p_{\max}$  的调整值  $\Delta p_1$ 、 $\Delta p_2$ 、 $\Delta p_3$  如表 1 所示,其中源数据包的长度分别为  $k = 200$ 、 $k = 500$ 、 $k = 1000$ .图 2 给出了  $k = 200$ 、 $k = 500$ 、 $k = 1000$  时,度分布优化后可译集合值在译码过程中的变化情况,作为对比,采用 RSD 度分布

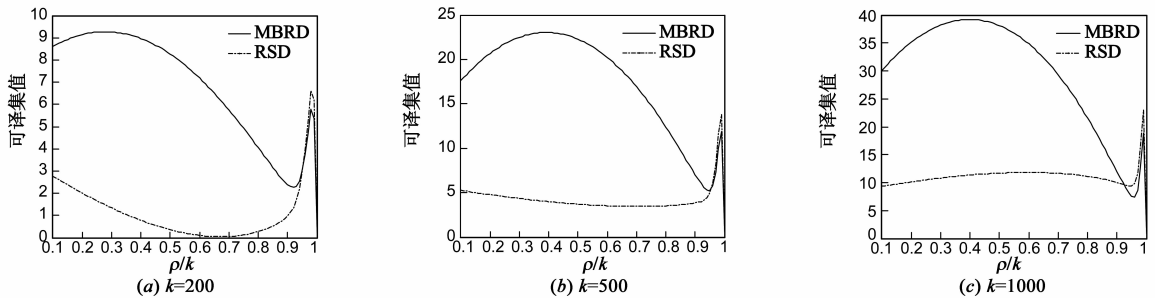


图2 可译集合值随译码成功率  $\rho/k$  的变化情况

表 1  $\Delta p_1$ 、 $\Delta p_2$ 、 $\Delta p_{\max}$  在不同源数据包个数下的取值

$k$	200	500	1000
$\Delta p_1$	-0.0331	-0.0091	-0.0118
$\Delta p_2$	0.0518	0.0535	0.0353
$\Delta p_{\max}$	-0.0187	-0.0444	-0.0235

## 4.3 译码性能的仿真

度分布性能的好坏最终体现在译码效率上.表 2 为分别采用修正二进制-鲁棒孤子度分布、二进制指数度分布和鲁棒孤子度分布进行 LT 编码,译码成功所需要的译码开销的仿真结果.作为对比,鲁棒孤子度分布的参数也同样选取为  $c = 0.02$ ,  $\delta = 0.01$ .从表 2 可以看出,采用二进制指数度分布编码数据包时,所需要译码开销最大,即完成整个译码过程所需要的源数据包个数最多;采用修正二进制-鲁棒孤子度分布编码数据包时,译码开销最小,特别是当源数据码长较短时,修正二进制-鲁棒孤子度分布性能改善更大.

表 2 采用不同度分布进行编码的译码开销

$k$	译码开销 (%)		
	MBRD	BED	RSD
200	19.3	57.5	35
500	11.6	56	20.1
1000	8.5	55.5	15.8

图 3 为当  $k = 500$ 、 $k = 1000$ ,采用 3 种不同度分布进行编码时,译码过程中,成功译出的源数据包比率随接收到的编码数据包数量的变化情况的仿真结果.从图 3 可以看出,在译码器接收到的编码数据包进行译码时,采用二进制指数度分布进行编码,成功译出的源数据包的比率最高.但随着接收到的编码数据包的增加,二进制指数度分布的成功译码源数据包的比率增长较

慢,而修正二进制-鲁棒孤子度分布和鲁棒孤子度分布的成功译码源数据包比率增加迅速.修正二进制-鲁棒孤子度分布的成功译码源数据包比率最先到达 1,说明采用修正二进制-鲁棒孤子度分布进行编码,其译码过程最先完成,反映译码效率最高.

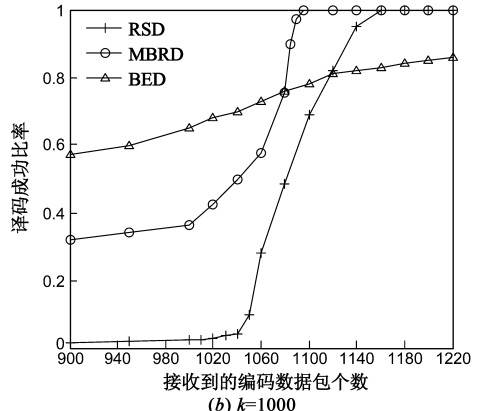
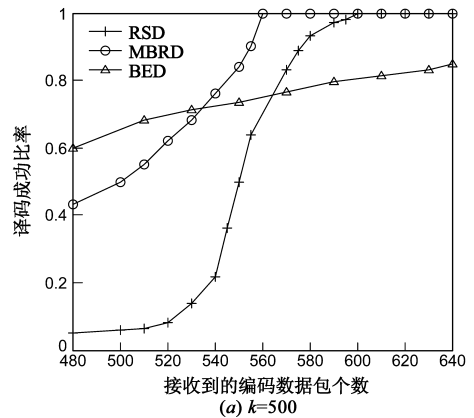


图3  $k=500$ 、 $k=1000$ ,采用 3 种不同度分布进行编码,译码成功比率与接收到的编码数据包的关系

## 5 总结

针对现在常用的 LT 码度分布在源数据码长较短时编译码性能下降明显的情况,本文给出了一种当源数据为短码长时也有较好性能的度分布设计方案.首先,根据度分布特性对二进制度分布进行调整,然后将其与鲁棒孤子度分布进行归一化结合,并引入两个参数来决定这两种度分布的最佳组成比例,再通过对可译集合值的优化来进一步调整度分布函数中度 1、度 2 和最大度的概率值,形成一种当源数据为短码长时也有较好性能的度分布,即修正二进制-鲁棒孤子度分布.仿真结果表明,修正二进制-鲁棒孤子度分布相对二进制指数度分布和鲁棒孤子度分布,其编译码的性能都有提高.特别是当源数据码长较短时,修正二进制-鲁棒孤子度分布的编译码性能提高十分明显.

### 参考文献

- [1] Byers J, Luby M, Mitzenmacher M. A digital fountain approach to reliable distribution of bulk data[J]. ACM Special Interest Group on Data Communication Computer Communication Review, 1998, 28(4): 56 – 67.
- [2] Luby M. LT codes[A]. Proceedings of 43rd Annual IEEE Symposium on Foundations of Computer Science(FOCS)[C]. USA: IEEE, 2002. 271 – 280.
- [3] 雷维嘉, 谢显中, 李广军. 采用数字喷泉码的无线协作中继方案及其性能分析[J]. 电子学报, 2010, 38(1): 228 – 233.  
Lei Wei-jia, Xie Xian-zhong, Li Guang-jun. The scheme and performance of wireless cooperative relay system using digital fountain codes[J]. Acta Electronica Sinica, 2010, 38(1): 228 – 233. (in Chinese)
- [4] 慕建君, 焦晓鹏, 曹训志. 数字喷泉码及其应用的研究进展与展望[J]. 电子学报, 2009, 37(7): 1571 – 1577.  
Mu Jian-jun, Jiao Xiao-peng, Cao Xun-zhi. A survey of digital fountain codes and its application[J]. Acta Electronica Sinica, 2009, 37(7): 1571 – 1577. (in Chinese)
- [5] MacKay D. Fountain codes[J]. IEE Communications Proceedings, 2005, 152(6): 1062 – 1068.
- [6] Etesami O, Shokrollahi A. Raptor codes on binary memoryless symmetric channels[J]. IEEE Transactions on Information Theory, 2006, 52(5): 2033 – 2051.
- [7] Liau A, Yousefi S, Kim I M. Binary soliton-like rateless coding for the Y-network[J]. IEEE Transactions on Communications, 2011, 59(12): 3217 – 3222.

- [8] Maatouk G, Shokrollahi A. Analysis of the second moment of the LT decoder[J]. IEEE Transactions on Information Theory, 2012, 58(5): 2558 – 2569.
- [9] Bodine E A, Cheng M K. Characterization of Luby transform codes with small message size for low-latency decoding[A]. IEEE International Conference on Communications[C]. Beijing: IEEE, 2008. 1195 – 1199.
- [10] Chong Z K, Goi B M, Ohsaki H. Design of short-length message fountain code for erasure channel transmission[A]. Proceedings of IEEE Conference on Sustainable Utilization and Development in Engineering and Technology (STUDENT)[C]. Kuala Lumpur: IEEE, 2012. 239 – 241.
- [11] Hyytia E, Tirronen T, Virtamo J. Optimal degree distribution for LT Codes with small message Length[A]. Proceedings of 26th IEEE International Conference on Computer Communications[C]. USA: IEEE, 2007. 2576 – 2580.
- [12] Agha A K, Kadi N, Stojmenovic I. Fountain codes with XOR of encoded packets for broadcasting and source independent backbone in multi-hop networks using network coding[A]. Proceedings of IEEE 69th Vehicular Technology Conference[C]. Barcelona: IEEE, 2009. 1 – 5.
- [13] 马昌凤. 最优化方法及其 Matlab 程序设计[M]. 北京: 科学出版社, 2010. 211 – 235.

### 作者简介



雷维嘉 男, 1969 年出生, 云南元谋人. 教授、博士. 现主要从事无线通信传输技术的研究.  
E-mail: leiwj@cqupt.edu.cn



张梦 女, 1988 年 10 月出生, 重庆渝中人. 2014 年毕业于重庆邮电大学, 获取硕士学位. 现主要从事信息管理与信息系统专业教学工作, 及信道编码、个人通信等领域的科研工作.  
E-mail: adamengzhang@163.com