

改进的块复制不等差错保护喷泉码

邓在辉¹ 同小军¹ 甘良才²

(1. 武汉纺织大学数学与计算机学院, 武汉, 430200; 2. 武汉大学电子信息学院, 武汉, 430072)

摘要: 在多媒体传输、深空通信等应用中, 由于数据的重要性程度不同, 需要喷泉码具有不等差错保护的 特性, 根据数据重要性给予不同程度的保护。对基于块复制的不等差错保护喷泉码的机制进行了 理论分析, 提出了改进的信息符号选取策略, 让度 1 编码符号全部在重要信息比特中选取, 一定数量的 度 2 编码符号分别在重要信息比特和次要信息比特中选取, 从局部上提升对重要信息比特的保护。仿 真结果表明, 相比于现有的基于块复制的不等差错保护喷泉码, 改进方法在进一步提高重要信息比特保 护的同时没有降低对次要信息比特的保护, 且有更好的不等恢复时间性能。

关键词: 喷泉码; LT 码; 不等差错保护; 块复制

中图分类号: TN911.22 **文献标志码:** A

Improvement of Unequal Error Protected Fountain Codes Based on Block Duplication

Deng Zaihui¹, Tong Xiaojun¹, Gan Liangcai²

(1. School of Mathematics and Computer Science, Wuhan Textile University, Wuhan, 430073, China; 2. Electronic Information School, Wuhan University, Wuhan, 430072, China)

Abstract: In several applications such as multimedia transmission and deep space communications, some data may need more protection than the rest because of the different importance of the data. Such applications raise the need for fountain codes with unequal error protection (UEP). The operation mechanism of the unequal error protected fountain codes based on block duplication is analyzed, and an improved selection strategy of information symbols is proposed to improve the protection of more important bits (MIB) in local area. Degree 1 encoded symbols are selected from MIB and a certain number of degree 2 encoded symbols are chosen from MIB and less important bits (LIB), respectively. Simulation results show that compared with the existing method, the proposed scheme perform better on decoding MIB without reducing the decoding of LIB while reserving a better unequal recovery time (URT) property.

Key words: fountain codes; LT codes; unequal error protection; block duplication

引 言

喷泉码由于其能很好适应信道变化和具有低的编译码复杂度而受到理论界和产业界的广泛关注,

主要包括 LT(Luby transform, LT)码、Raptor 码^[1-3]。其最初的实现方案中,所有的数据是同等被保护的,但在很多情况下,比如多媒体通信、深空通信中,部分数据需要更高的可靠性,因此设计具有不等差错保护(Unequal error protected, UEP)的喷泉码成为研究喷泉码的一个重要方面。

文献[4]首次实现了不等差错保护的喷泉码,在编码时对于给定的度分布在生成度值后,选取原始信息符号时增加重要信息比特(More important bits, MIB)的选择概率,从而提高对 MIB 的保护程度。文献[5]通过改变原始信息符号的选择策略来实现喷泉码的不等差错保护,让小的度值编码符号偏向 MIB。文献[6]采用扩展窗方法实现喷泉码的不等差错保护性能,通过窗口包含使优先级越高的信息符号被包含的窗口越多,通过窗口的选择概率和其采用的度分布来增加对优先级高的信息符号的保护。文献[7]通过一个编码二分图来改变编码符号度值分配以及与原始信息符号的映射关系来实现喷泉码不等差错保护,其完全不能利用已有的好度分布。文献[8-9]在不改变标准喷泉码结构的基础上结合视频流的参数来实现喷泉码的不等差错保护,由于其采用了多个编码二分图,会增大总的传输冗余开销。文献[10]通过对信息符号集复制的方法来使不同优先级的信息符号的复制程度不同以实现不等差错保护,实验表明其性能整体上好于文献[4,6]中的算法。本文采用与或树工具对其性能进行分析,并在其基础上通过改变原始信息符号的选择策略,对于度为 1 和 2 的编码符号链接的原始信息符号偏向 MIB,使得对 MIB 有更好的保护能力,而没有损失对次要信息比特(Less important bits, LIB)的保护。

1 LT 码编译码原理

给定 k 个原始信息符号 $\{i_0, \dots, i_{k-1}\}$ 和集合 $\{1, \dots, k\}$ 上的概率分布 $\Omega(x)$, 产生无限长的编码符号 e_m ($m \geq 0$) 的方式如下:(1)根据度分布 $\Omega(x)$ 随机产生一个度值 $d \in \{1, \dots, k\}$; (2)从 $\{1, 2, \dots, k\}$ 个符号中均匀随机选取 d 个符号的序号;(3)将选取的 d 个符号进行异或操作得到编码符号 e_m 。

重复以上步骤即能产生无限的编码符号,原始信息符号和编码符号的关系如图 1 所示,度分布通常采用鲁棒孤子分布^[1]。

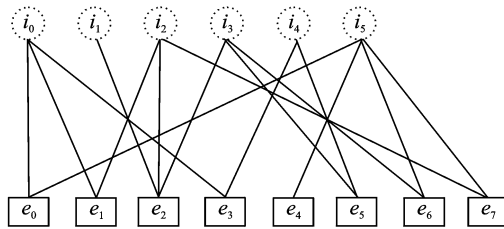


图 1 LT 码示意图

Fig. 1 Graph of LT code

其解码过程采用置信传播(Belief propagation, BP)译码算法,在由信息符号和编码符号形成的二分图中,度为 1 的编码符号能直接恢复出信息符号。(1)找出只有一条边连在原始符号 i_j 的编码符号 e_m , 如果没有这样的符号,即停止译码。令 $i_j = e_m$, 恢复出 i_j 。(2)对于所有的连接到 i_j 且 $x \neq m$ 的编码符号 e_x , 令 $e_x = e_x \oplus i_j$ (\oplus 表示异或操作), 同时去除所有连接到原始符号 i_j 的边。

重复以上步骤即完成 BP 译码算法,从中可看出其译码过程为不断去边的过程,所有原始信息符号被成功解码的概率随着接收到的编码符号的增加而增加。

2 改进的块复制 UEP 喷泉码

传统喷泉码对信息符号的保护是等差错保护,但在许多情形下,如视频传输中往往需要其具有不等

差错保护特性。

2.1 块复制 UEP 喷泉码

对于给定长度的 LT 码可以通过复制信息符号的方法来增加码长^[10],类似于在滑动窗口喷泉码中的对窗口中的信息符号进行编码^[9],假设窗口大小为 w ,每次滑动 s 个符号直到所有信息符号都能被编码。该方法通过直接对原始信息符号进行复制的方法来增加信息符号长度,构成新的原始信息符号,度数和信息符号的选取按照扩展后的信息符号长度进行计算。假设原始信息符号为 $S=i_0 \times \dots \times i_{k-1}$ 包含 k 个信息符号 i_0, \dots, i_{k-1} , $\Omega(x)$ 是长度为 k 的 LT 码的度分布,对原始符号进行复制扩展,假设扩展因子 (Expanding factor, EF) 表示为 EF,则新的原始符号可写成 $\underbrace{S \times S \times \dots \times S}_{EF}$,新的原始符号长度变为 $EF \times k$,其信息符号的索引范围从 0 到 $EF \times k - 1$ 。另外,度分布 $\Omega(x)$ 的范围从 $\{1, \dots, k\}$ 变到 $\{1, \dots, EF \times k\}$,使用标准的 LT 码编码器和新的度分布来产生编码符号。对于鲁棒孤子分布而言,新的度分布的长度由 $EF \times k$ 代替 k 来产生度值。在符号选取上,需要通过模运算得到,如果选取的符号索引为 $j \in \{0, \dots, EF \times k - 1\}$,则在 i_0, i_1, \dots, i_{k-1} 中对应的符号索引为 $j \bmod k$ (\bmod 表示模运算)。

依此思路构造不等差错保护 LT 码的过程为:假设一个信源块 $S=i_0 \times \dots \times i_{k-1}$ 分为 L 个相邻的块 S_1, S_2, \dots, S_L ,第一个块 S_1 包含最重要的信息比特、下一个块 S_2 包含次要的比特、依此类推 S_L 包含相对最不重要的比特。根据一系列复制因子 $RF_i, i=1, \dots, L$ 来对这些分成的子块进行不同程度的保护,形成虚拟长度的信息序列 $\underbrace{S_1 \times S_1 \times \dots \times S_1}_{RF_1} \times \underbrace{S_2 \times S_2 \times \dots \times S_2}_{RF_2} \times \dots \times \underbrace{S_L \times S_L \times \dots \times S_L}_{RF_L}$,该序列的信息符号的索引范围从 1 到 $\sum_{i=1}^L RF_i |S_i|$,相当于增加了码长。这样 LT 码度分布的度数取值从

$\{1, \dots, k\}$ 扩展为 $\{1, \dots, \sum_{i=1}^L RF_i |S_i|\}$,在 LT 编码过程中,根据扩展序列得到新的度分布概率值,对于度值为 d 的编码符号则从扩展后的信息块中选择 d 个信息符号,而在选取符号的索引值 $\{0, \dots, \sum_{i=1}^L RF_i |S_i| - 1\}$ 需要映射到扩展前的原始符号 $\{i_0, \dots, i_{k-1}\}$,其对应的映射规则为

$$l = \begin{cases} j \bmod |S_1| & 0 \leq j \leq RF_1 |S_1| - 1 \\ [(j - RF_1 |S_1|) \bmod |S_2|] + |S_1| & RF_1 |S_1| \leq j \leq RF_1 |S_1| + RF_2 |S_2| - 1 \\ \vdots & \\ [(j - \sum_{i=1}^{L-1} RF_i |S_i|) \bmod |S_L|] + |S_{L-1}| + \dots + |S_1|, & \sum_{i=1}^{L-1} RF_i |S_i| \leq j \leq \sum_{i=1}^L RF_i |S_i| - 1 \end{cases} \quad (1)$$

另外,在编码过程中为了进一步增加码长,可将得到的序列通过设定的扩展因子 EF 进一步扩展,从而进一步让虚拟的码长增加。设扩展因子表示为 EF,原始信息符号的复制和扩展如图 2 所示。假设 $k=6, L=2$,重要信息符号长度 $|S_1|=2$,次要信息符号长度 $|S_2|=4$,虚拟的信息序列的长度为 8,相应的复制因子分别为 $RF_1=2, RF_2=1$ 。相应地,LT 码的原始信息序列长度从 $k=6$ 变为 $k=8$,如果编码符号的度值为 d ,则 d 从 8 个虚拟的符号中均匀选取,如果选择的信息符号的索引为 5 则利用式(1)映射为 $5 \bmod 4 + 2 = 3$ 。

2.2 基于与或树的理论分析

令 $\Omega(x) = \sum_{i=1}^k \Omega_i x^i$ 表示 LT 码编码过程中的度分布多项式, k 个输入符号分 r 个集合: s_1, s_2, \dots, s_r , 每个集合的大小分别为 $\alpha_1 n, \alpha_2 n, \dots, \alpha_r n$, p_j 表示在 LT 码编码过程中集合 s_j 中符号的选择概率, γ_k 表

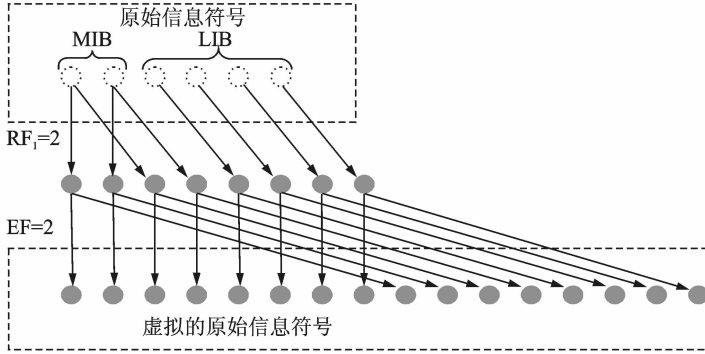


图 2 块复制 UEP 喷泉码的虚拟原始分组构建

Fig. 2 Building virtual source block with UEP fountain codes based on block duplication

示接收到编码符号长度, $y_{t,j}$ 表示集合 s_j 中的原始信息符号在经过 t 次迭代后还没被恢复的概率^[4], 根据与或树分析理论有

$$y_{t,j} = \delta_j (1 - \beta (1 - \sum_{k=1}^r p_k \alpha_k^n y_{t-1,k})) \quad t \geq 1 \tag{2}$$

式中 $y_{0,j} = 1, \beta(x) = \Omega'(x) / \Omega'(1), \mu = \Omega'(1), \delta_j(x) = e^{n p_j \mu \gamma (x-1)}$ 。

由此, 原始信息符号的恢复概率比 $G_{t,i,j} = \frac{y_{t,i}}{y_{t,j}}$ 的计算为

$$G_{t,i,j} = e^{n(p_i - p_j) \mu \beta (1 - \sum_{a=1}^r p_a \alpha_a^n y_{t-1,a})} \tag{3}$$

显然, 在 LT 码编码过程中, 原始信息符号的选择概率能决定其在译码过程中的恢复概率, 只要 $p_i > p_j$, 在 $t \geq 1$ 时, 则集合 s_j 中的符号的译码成功概率就大于 s_j 中符号的译码成功概率, 满足 $G_{t,i,j} > 1$ 。

假设原始信息符号的长度为 k , 其中高优先级信息符号 MIB 长度为 h , 复制因子 RF_1, RF_2 分别为 R_1, R_2 , 扩展因子 EF 为 E_1 , 令 $RF_2 = 1$, 则编码过程中 MIB 中符号的选择概率 p_h , LIB 中符号的选择概率 p_l 分别为式(4)和式(5)

$$p_h = \frac{R_1}{k + (R_1 - 1)h} \tag{4}$$

$$p_l = \frac{1}{k + (R_1 - 1)h} \tag{5}$$

由于 $R_1 > 1$, 于是 $p_h > p_l$, 据式(4)得 $G_{t,h,l} > 1$, 即相同次数译码迭代后, 高优先级信息符号的译码概率要大于低优先级符号的译码概率, 实现喷泉码不等差错保护性能。通过增加高优先级符号的选择概率能从整体上提高喷泉码的不等差错保护性能, 即度数为 1 到截断值的编码符号中有边连接到高优先级符号的概率整体上要高于有边连接到低优先级符号的概率。

2.3 改进的块复制 UEP 喷泉码

文献[10]中将复制扩展的 UEP 和文献[4,6]中不等差错保护的性能进行比较, 文献[10]中的性能在当收到大于 k 的编码符号时表现较好, 但是, 在许多应用环境中经常出现收到的编码符号数小于 k , 如在无线视频实时传输中, 接收到的编码符号往往少于码长 k , 在该情况下希望能尽早地、尽量多地恢复出重要的信息比特。文献[10]通过复制扩展能从整体上改变 MIB 和 LIB 的不等保护程度, 其思想本质上一方面加大了对 MIB 的选择概率, 另一方面增加了原始信息符号的长度, 从而改变了整体上的度分布。在此基础上, 本文从局部上进行改进, 在对信息符号的选取上改变其选取策略, 而不损失整体上

的不等差错保护性能。

度为 1 和 2 的编码符号对整个原始信息符号的译码起至关重要的作用,因为解码过程是从度为 1 的编码符号开始,所有的度为 1 的编码符号直接从高优先级的符号中选取,则在译码过程中,它们能被直接恢复为原始符号,而不需要等到其他的符号先被解码出来,增加了高优先级符号先于低优先级符号被解码出来的可能性,且其所占比例不能太大。并且度数越低的输出符号,其在解码过程中通过去边而成为度为 1 符号的可能性越大。将度为 2 的编码符号的一条边连接到优先级高的比特,一边连接着低优先级的原始符号,这样既保证了对优先级高的符号的优先保护,同时也使得整体的原始符号都有边能连接到,因为喷泉码的译码过程实质是不断去边的过程。改变度为 1 和 2 的原始符号选取策略从局部加大了对高优先级符号的可靠性保护,而剩下的度为 2 和更高度数的符号能开始低优先级符号的解码并继续下去,因此,低优先级符号的解码性能基本没有下降。

令 c 表示编码过程中度为 2 的编码符号索引,其初值为 1, $|S_{\text{MIB}}|$ 为重要信息符号的长度,本文所提出的算法步骤为:(1)根据原始信息符号序列中不等重要的块进行不等复制。(2)对复制后的原始信息符号序列进行扩展,改变信息符号长度。(3)对扩展后的信息符号序列按照新的度分布产生度值 d ,如果度值 d 为 1,则从序列中复制后的重要比特中选取信息符号。如果度值 d 为 2,当 c 小于或等于 $|S_{\text{MIB}}|$,一个符号从 MIB 中选取,另一符号从 LIB 中选取,同时令 $c=c+1$ 。否则,当 c 大于 $|S_{\text{MIB}}|$,在复制扩展后的整个序列中选取信息符号。如果度值 d 大于 2,在复制扩展后的整个序列中选取信息符号。(4)根据式(1)进行选取符号在复制和扩展前的信息序列中进行映射,并对映射的符号进行异或操作产生编码符号,转到第(3)步。

3 实验结果及分析

为了评估本文所提出算法的性能,采用与文献[10]相同的参数进行性能比较,文献[10]中算法记为块复制方法、本文算法记为改进的块复制方法,简称改进方法。假设原始信息符号长度分别为 $k=1\ 000$ 和 $k=5\ 000$,信息符号分为两个优先级块,其中 MIB 块的大小为原始信息符号长度的 10%,鲁棒孤子分布的参数为 $c=0.1, \delta=0.5$,于是,度值为 1 的编码符号个数远小于 MIB 块原始符号个数,由于删除信道上的 LT 码为通用喷泉码,通信信道假设为无损信道^[11]。

3.1 比特错误率的比较

采用 BP 算法对喷泉码进行译码,在迭代过程中,二分图中如果没有了度为 1 的编码符号,而且这时还有原始符号没有被译码出来,则整个译码过程结束,未被译码出来的原始符号被判决为误码。图 3 和图 4 分别反映当 k 为 1 000 和 5 000 时的 MIB 和 LIB 的比特错误率(Bit error rate, BER)随信息冗余长度变化的情况。BER 表示错误符号个数与总符号个数的比值,即 $(k_1 - d)/k_1$, d 表示正确解码出来的符号的个数, k_1 表示相应的原始符号的个数。由于 LT 码的编解码具有随机的特性,仿真采用多次仿真求平均值的方法,仿真次数为 1 000 次。

在 $k=1\ 000$ 时,采用与文献[10]中相同的复制因子 $\text{RF}_1=3, \text{RF}_2=1$,扩展因子 $\text{EF}=4$ 。如图 3 所示,本文提出算法 MIB 的 BER 要明显好于文献[3]的仿真结果,当传输冗余开销为 $t'=0.21$ 时, MIB 的 BER 值低于 10^{-4} ,而文献[10]算法 MIB 的 BER 的值低于 10^{-4} 时对应的头部开销 t' 为 0.25,虽然对于 LIB 的 BER 相对于文献[10]中的算法有所下降,本文算法在传输冗余开销为 0.38 时, LIB 的 BER 低于 10^{-3} ,而文献[10]中此情况下对应的 t' 为 0.37,但是,从图中可以看出,在相同的传输冗余开销时,两种算法对于 LIB 的 BER 而言基本相同,所以说从整体上而言,在基本没有损失 LIB 的性能情况下,很大程

度上提高了 MIB 的 BER 性能。另外从不等恢复时间上看,所提出的算法加大 MIB 和 LIB 的译码不等恢复时间的程度。

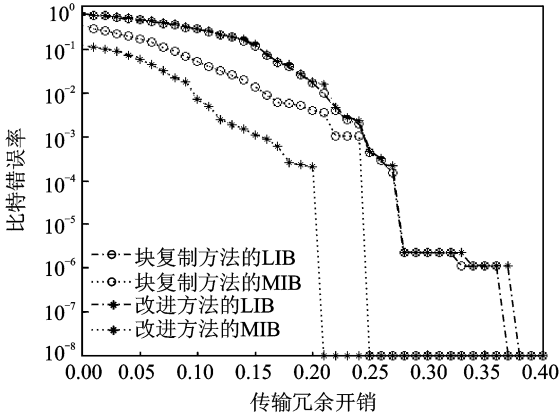


图 3 不同传输冗余开销时的 BER($k=1\ 000$)

Fig. 3 BER versus transmission overhead for improvement approach and method of Ref. [10] ($k=1\ 000$)

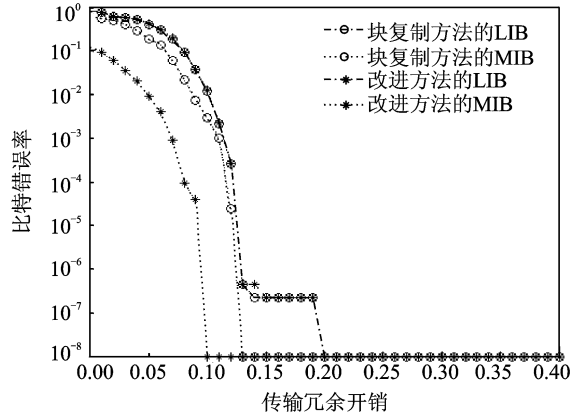


图 4 不同传输冗余开销时的 BER($k=5\ 000$)

Fig. 4 BER versus transmission overhead for improvement approach and method of Ref. [10] ($k=5\ 000$)

在 $k=5\ 000$ 时,采用与文献[10]中相同的复制因子 $RF_1=2, RF_2=1$ 和扩展因子 $EF=4$ 。如图 4 所示,本文提出算法 MIB 的 BER 同样要明显好于文献[10]的仿真结果,当传输冗余开销 t' 为 0.1 时, MIB 的 BER 值低于 10^{-5} ,而文献[10]中块复制算法 MIB 的 BER 值低于 10^{-5} 时对应的传输冗余开销 t' 为 0.13,同样,所提出的算法在整体上没有损失 LIB 的性能,两种方法都在 $t'=0.2$ 时 BER 值低于 10^{-7} 。本文所提出的算法在于不同的码长都能更高程度地保护重要的信息比特。

3.2 高优先级符号译码失败分布比较

为评估不等差错保护方法对重要性数据的保护能力,当收到的编码符号个数少于原始信息符号个数时,通过直方图来仿真分析接收 MIB 数据的失败分布情况,仿真次数为 1 000 次。

在 $k=1\ 000$ 、传输冗余开销为 -0.05 时,高优先级符号译码失败分布如图 5 所示,负号表示接收到的编码符号个数小于原始符号个数。纵坐标表示发生的次数,横坐标表示还有待于译码的高优先级符号个数的百分比,反映在该冗余开销量情况下,没有译码出来的高优先级符号个数所占比重。文献[10]中算法在收到 950 个编码符号时,至少有 95% 以上的高优先级符号没有译出的次数约为 28,而且,占比最多的情况发生在 70%~75%,约为 140。而在本文提出改进算法中最多有 70%~75% 的高优先级符号未译出,且该次数为 4。占比最多的情况发生在 25%~30%,约为 210。比较两种方法所得出的优先级符号译码失败分布,从中可得出提出算法相比文献[10]中算法能明显加快 MIB 的译码,使高优先级符号得到更好的差错保护。

在 $k=5\ 000$ 、传输冗余开销为 -0.05 时,高优先级符号译码失败分布如图 6 所示,文献[10]中算法在收到 4 500 个编码符号时,至少有 95% 以上的高优先级符号没有译出的次数约为 16,而且,占比最多的情况发生在 65%~70%,约为 138。而在提出改进算法中最多有 55%~60% 的高优先级符号未译出,且次数为 3。占比最多的情况发生在 15%~20%,约为 260。从图 6 中可得出在 $k=5\ 000$ 时提出算法相比文献[10]中算法同样能明显加快高优先级符号的译码,得到了更好的不等差错保护性能。

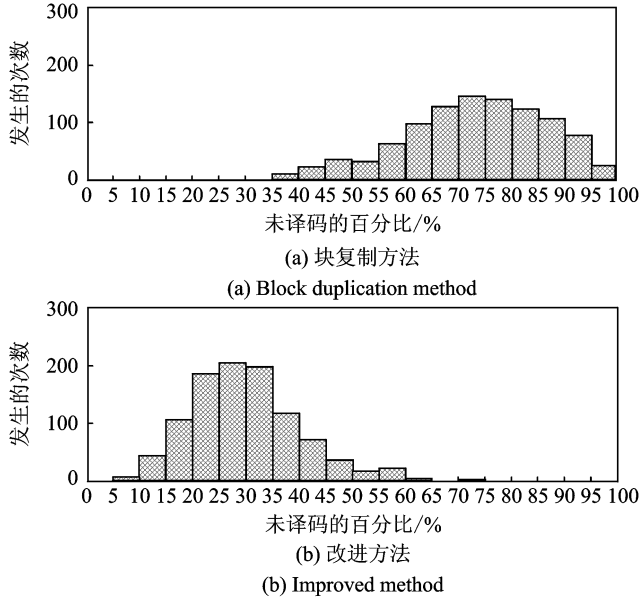


图 5 MIB 译码失败分布 ($k=1000$)

Fig. 5 Histogram of failure distribution for MIB with improvement approach and method of Ref. [10] ($k=1000$)

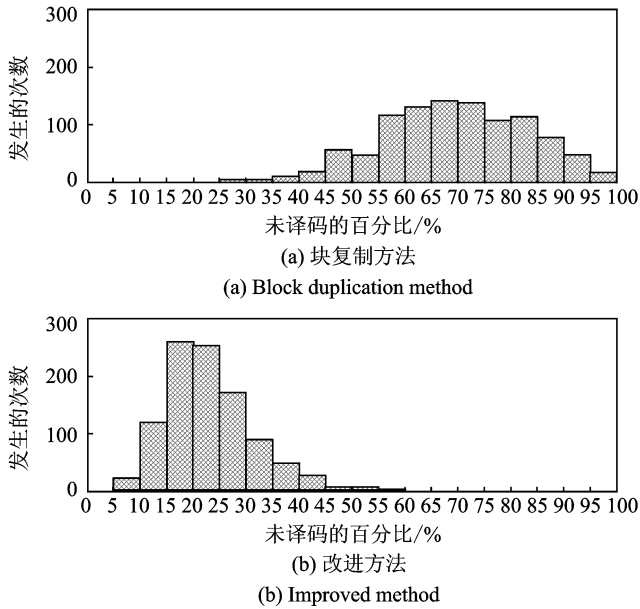


图 6 MIB 译码失败分布 ($k=5000$)

Fig. 6 Histogram of failure distribution for MIB with improvement approach and method of Ref. [10] ($k=5000$)

4 结束语

本文分析了基于块复制的不等差错保护 LT 码的实质,在其基础上提出了改进的不等差错保护的

LT 码。在编码时改变了 MIB 和编码符号的链接策略,不仅能保持从整体上提高 MIB 和 LIB 的差错保护,同时在不损失 LIB 的保护基础上,能更高程度保护 MIB。通过信息符号选择策略从局部增加了对 MIB 的保护力度,同时提升了 MIB 和 LIB 的恢复时间差别。仿真结果表明本文提出的方法要好于文献 [10] 中的方法。

参考文献:

- [1] Luby M. LT codes[C]//Proc 2002 IEEE Symp Foundations of Computer Science(FOCS). Vancouver, Canada;IEEE, 2002: 271-282.
- [2] Mackay D J C. Fountain codes[J]. IEEE Proceedings Communications, 2005, 152(6): 1062-1068.
- [3] Shokrollahi A. Raptor codes[J]. IEEE Transactions on Information Theory, 2006, 52(6):2551-2567.
- [4] Rahnavard N, Vellambi B N, Fekri F. Rateless codes with unequal error protection property[J]. IEEE Transactions on Information Theory, 2007, 53(4):1521-1532.
- [5] Simon S W, Cheng M K. Prioritized LT codes [C]//Proceedings of the 42 Annual Information Sciences and Systems. New Jersey, USA: [s. n.], 2008:568-573.
- [6] Vukobratovic D, Stankovic V, Sejdinovic D, et al. Scalable video multicast using expanding window fountain codes [J]. IEEE Transactions on Multimedia, 2009, 11(6):1094-1104.
- [7] Cao Y, Blostein S D, Chan W Y. Unequal error protection rateless coding design for multimedia multicasting[C]//ISIT 2010. Austin, Texas, USA;IEEE, 2010:2438-2442.
- [8] Tan A S, Aksay A, Akar G B, et al. Rate-distortion optimization for stereoscopic video streaming with unequal error protection[J]. EURASIP Journal on Advances in Signal Processing, 2009, 2009:1-14.
- [9] Cataldi P, Grangetto M, Tillo T, et al. Sliding-window raptor codes for efficient scalable wireless video broadcasting with unequal loss protection[J]. IEEE Transactions on Image Processing, 2010, 19(6):1491-1503.
- [10] Ahmad S, Hanzaoui R, Al-Akaidi M. Unequal error protection using fountain codes with application to video communication [J]. IEEE Transactions on Multimedia, 2011, 13(1):92-101.
- [11] 黄诚, 易本顺. 基于抛物线映射的混沌 LT 编码算法 [J]. 电子与信息学报, 2009, 31(10):2527-2531.
Huang Cheng, Yi Benshun. Chaotic LT encoding algorithm based on parabolic map[J]. Journal of Electronics & Information Technology, 2009, 31(10):2527-2531.

作者简介:



邓在辉 (1976-), 男, 博士, 研究方向: 通信信号处理, 信道编码, 无线网络, E-mail: zaihuideng@126.com.



同小军 (1967-), 男, 博士, 教授, 研究方向: 模式识别与图像处理。



甘良才 (1942-), 男, 教授, 博士生导师, 研究方向: 跳频通信, 数字通信。