

• 成电青年学者·信息与通信工程专栏 •



基于重要信息反馈的 BATS 码优化设计

杨娟^{1,2}, 史治平^{1*}, 嵇建波²

(1. 电子科技大学通信抗干扰技术国家级重点实验室 成都 611731; 2. 桂林航天工业学院电子信息与自动化学院 桂林 541004)

【摘要】 BATS 码是一种结合喷泉码和网络编码技术的新型前向纠错码, 能有效保证数据在多跳网络中的可靠传输。在传统 BATS 码编译码方案中, 反馈信息没有得到高效利用, 为提高 BATS 码的译码性能, 提出了一种改进的基于重要信息反馈的 BATS 码编译码算法, 其中, 重要信息包由接收端和发送端共同选择用于下一轮编码; 同时, 对重要信息的编码包在中继节点采用网络编码算法进行再编码, 以保证重要信息在接收端的可靠恢复。仿真结果表明, 该方案相比传统的 BATS 码以及过去用于喷泉码的基于重要信息反馈的编译码方案, 具有更好的译码性能。

关键词 BATS 码; BP 译码算法; 反馈; 文件分发

中图分类号 TN919 **文献标志码** A **doi**:10.12178/1001-0548.2021109

Design of Improved BATS Codes Based on the Important Information Feedback

YANG Juan^{1,2}, SHI Zhiping^{1*}, and JI Jianbo²

(1. National Key Laboratory of Science and Technology on Communication, University of Electronic Science and Technology of China Chengdu 611731;

2. School of Electronic Information and Automation, Guilin University of Aerospace Technology Guilin 541004)

Abstract BATS codes combining fountain codes and network coding are a new class of erasure codes which effectively guarantee the reliable transmission of packets in multi-hop network. However, conventional BATS codes do not make use of the feedback information. For improving the decoding performance of BATS codes, we design an improved BATS codes encoding and decoding algorithm based on the important information. And the important information is selected by the source node combined with the destination node for the next coding process. Furthermore, the coded packets from the important information are recoded at the intermediate nodes for the reliable recovery of the important information at the destination node. Simulation results show that the proposed improved scheme achieves better decoding performance compared with the conventional BATS code and the traditional limited feedback used in fountain code.

Key words BATS codes; BP decoding; feedback; file distribution

在无线通信网络中, 一个重要且基础的工作就是文件分发, 将文件数据包从源节点经中继节点发送至目的节点。数据包在无线通信网络中进行传输时, 可能由于信道噪声、衰落、接入冲突等原因而造成数据丢失现象^[1]。

文献 [2] 提出的数字喷泉码作为一种前向纠错码能很好地解决这一问题。喷泉码是一种发送端源源不断地发送编码包直至接收端能够成功译码的纠错编码技术。将原始文件划分为 k 个数据包, 随机对这 k 个数据包进行编码, 可以得到任意 n 个编码

数据包, 因此喷泉码是任意码率或无码率的。将喷泉码应用于含中继节点的无线通信网络中, 在中继节点采用的是存储转发操作, 而该操作并不是中继节点可采用的最优操作。在多播网络中, 只有在中继节点应用网络编码才能达到网络最大吞吐率^[3]。以两跳的单播删除网络为例, 假设每一条链路的删除概率为 0.2, 如果中继节点对接收数据进行存储转发操作, 网络吞吐率最高为每使用一次网络传输 0.64 包; 而如果中继节点允许对接收数据进行网络编码, 则吞吐率最高可达到每使用一次网络传

收稿日期: 2021-04-21; 修回日期: 2021-06-11

基金项目: 广西自然科学基金 (2020GXNSFBA297018, 2018GXNSFAA281161)

作者简介: 杨娟 (1986-), 女, 博士生, 主要从事网络编码方面的研究。

*通信作者: 史治平, E-mail: szp@uestc.edu.cn

输 0.80 包。BATS 码作为一种适用于多跳删除网络的新型网络编码方案^[4-7], 结合了喷泉码和网络编码算法, 将二者分别作为其外码和内码。BATS 码的外码是 LT 码的矩阵形式, 因此 BATS 码也是一种无速率码, 由发送端不断向接收端发送编码的批次码, 直到接收端译码成功。当批次码的大小 M 等于 1 时, BATS 码的外码即为 LT 码。BATS 码的内码采用随机线性网络编码算法, BATS 码将二者结合起来能保证数据在无线通信网络中可靠传输, 同时渐近最大网络容量。

传统的 BATS 码编译码方案是不依赖于反馈的, 但反馈信道在 BATS 码的应用场景中是可用的, 用于接收端在译码成功后反馈一个 ACK (acknowledge character) 确认信号使发送端不再继续发送编码数据包, 且无论使用与否, 反馈信道都被占用着, 因此应当考虑如何将反馈信道有效利用起来。文献 [8] 针对 LT 码提出当接收端译码过程停止时, 利用反馈信道反馈所有未译码的原始输入数据包包号, 在发送端将它们重新发送以重启接收端译码过程。文献 [9] 在文献 [8] 的基础上提出了一种改进方案, 接收端反馈未译码输入数据包的包号及译码后更新的生成矩阵, 按照未译码数据包的度值从大到小进行重传。然而, 这两种方案都需要反馈大量的信息包, 频带利用率较低, 且会导致反馈拥塞等问题。文献 [10] 针对 LT 码提出了一种基于重要信息反馈的编译码方案, 接收端只反馈度值最大的未译码输入数据包的包号, 发送端接收到后对该数据包进行重传, 在中继节点对其采用转发操作。这种方案在有效减少反馈数据量的同时, 也能使译码性能得到提升。然而, 当网络跳数较多或通信质量较差时, 重要数据包在重传过程中丢失的可能性较大, 从而将影响接收端数据包的恢复。

基于上述原因, 本文方案不仅考虑了如何确保重要信息能够有效传输到接收端, 还研究了译码过程中输入数据包较难完全译码成功的主要原因, 并据此改进了对重要输入数据包的选择方案, 使其能有效增加译码成功次数。在仿真结果中, 本文对比了传统 BATS 码, 基于重要信息反馈的 BATS 码以及本文优化方案, 在结果中论证了本方案的优越性。

1 BATS 码的编译码算法

BATS 码通常应用于包括中继节点的删除网络中, 如图 1 所示为一个多跳的删除网络, s 为源节

点, d 为目的节点, a_1, a_2, \dots, a_n 为中继节点。BATS 码包括内码和外码, 在多跳删除信道中应用 BATS 码, 外码在源节点也即发送端完成, 内码在中继节点进行。具体编码过程描述如下, 发送端对 K 个长度为 L 的原始数据包进行编码, 产生 n 个大小为 M 的批次码。产生第 i 个批次码时, 首先根据预先优化的编码度分布 Ψ 进行采样得到一个度 d_i , 在 K 个原始数据包中随机选取 d_i 个数据包组成一个 $L \times d_i$ 维的矩阵 B_i , 然后经外码编码产生一个包含 M 个编码包的批次码 X_i , 其数学表达形式为:

$$X_i = B_i G_i \quad (1)$$

式中, G_i 为生成矩阵, 是有限域上的一个完全随机矩阵。发送端将批次码一个接一个按顺序发送至中继节点。中继节点对接收到的批次码在同一个批次内进行内码编码, 编码方式采用随机线性网络编码。接收端接收到的第 i 个批次码 Y_i 表示为:

$$Y_i = X_i H_i = B_i G_i H_i \quad (2)$$

式中, H_i 为第 i 个批次码的转移矩阵, 由内码编码过程与信道情况共同决定。

由于计算复杂度低, 接收端通常采用 BP 译码算法进行译码。接收到给定数量的批次码后, 接收端开始译码, 译码时需要计算每一个批次码的度是否等于它的秩, 如果第 i 个批次码的度 d_i 等于它的秩 $\text{rk}(G_i H_i)$, 则采用高斯消元算法对该批次码进行求解, 得到与之关联的 d_i 个原始输入数据包信息, 并利用 d_i 个输入包的信息对与之关联的批次码进行替换和更新, 去除它们对这些批次码的影响。重复上述过程直至找不到满足度与秩相等条件的批次码, BP 译码过程停止。



图 1 多跳删除网络

在传统的 BATS 码编译码方案中, 发送端在第一轮编码时首先发送 N_1 个批次码, 接收端接收到后开始译码, 如译码成功则反馈一个 ACK 译码成功确认信号, 否则不返回反馈信号。在第 i 轮编码前, 发送端先确认是否接收到 ACK 信号, 接收到则停止编码传输, 未接收到则继续对原始数据包进行编码, 产生并发送 N_i 个批次码, 固定编码轮数或重复上述编码过程直至发送端接收到译码成功的 ACK 信号, 编译码过程停止。

2 BATS 码的优化方案设计

2.1 重要信息的选择

由于 BATS 码具有无码率特性, 在未收到 ACK 反馈信号前, 发送端将不断发送批次码, BP 译码器在接收到一定数量的批次码后即可开始译码。批次码的编码包在传输过程中由于信道噪声、衰落等原因会发生丢失现象, 因此在接收端一些批次码由于不满足译码条件而无法进行译码, 当接收端不存在满足译码条件的批次码时, 译码过程停止。要重启译码过程, 需要获得更多的原始输入数据包信息, 将其替换进未译码的批次码中消除更多的关联, 从而产生新的可译批次码。采用传统 BATS 码编码方案, 编码器在进入下一轮编码时, 将随机选取输入数据包进行编码, 如果利用反馈信道获取反馈信息, 编码器可以优先对一些更重要的输入数据包进行编码, 从而加快译码进程。

接收端根据译码后各个未译出的批次码的包头信息可以统计出与之关联的输入数据包的度, 从而找出具有最大度值且未译码成功的输入数据包。由图 2 所示, 圆代表未译出的输入数据包, 它的集合用 $B = \{b_1, b_2, b_3, b_4, b_5, b_6\}$ 来表示, 方块代表不可译的批次码, 它的集合为 $\{Y_1, Y_2, Y_3, Y_4\}$ 。由图中可以看出, 各个未译出的输入数据包的度分别为 $\{1, 0, 4, 0, 3, 2\}$, 如果能将度为 4 的原始数据包 b_3 译出, 那么 Y_1, Y_2, Y_3, Y_4 的度值都会降低 1, 从而增大了产生可译批次码的概率。

文献 [10] 指出, 对于喷泉码, 如果能将度值最大的原始数据包译出, 在有限的反馈信息条件下, 能最大程度地增加各批次码的可译概率。因此本文将未译出且具有最大度值的原始数据包作为重要信息包的一部分, 译码器反馈一个数据包给编码器, 该数据包中包含这个重要信息包的包号。

由图 2 可知, 在未译出的输入数据包中, b_2 和 b_4 未译出的原因是它们并未参与编码, 因此无论是否所有的批次码被译出, b_2 和 b_4 都不能被译出。文献 [11] 指出 BATS 码特别是对于有限码长 BATS 码来说, 输入数据包完全译码成功概率较低, 而造成这个现象的主要原因是由于 BATS 码的编码算法采用的是随机选取输入数据包来进行编码的策略, 有的输入数据包在编码过程中始终没有被选取。基于上述原因, 本文将未参与编码的输入数据包也作为重要信息包的一部分参与下一轮的编码。这一部分输入数据包的包号可以在发送端进行分析和统

计, 不需要接收端进行反馈, 因此也不会由于增大反馈数据量而产生额外代价。所以, 本文将重要信息定义为两部分, 一部分为接收端未译出且具有最大度值的原始数据包, 另一部分为发送端未参与编码的原始数据包, 并将二者作为下一轮编码的重要信息包。

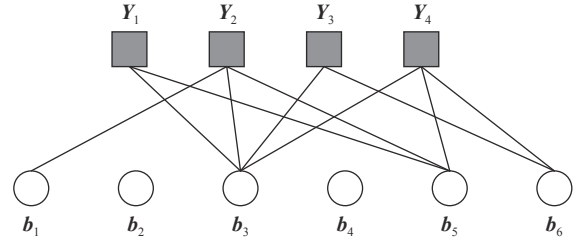


图 2 未译出的输入数据包与不可译批次码之间的二分图

2.2 基于重要信息反馈的 BATS 码优化设计

基于重要信息改进的 BATS 码编译码方案流程如图 3 所示。

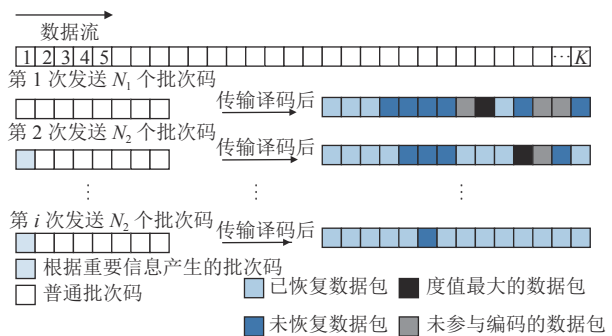


图 3 基于重要信息反馈的 BATS 码优化设计框图

在发送端, BATS 码外码编码器首先根据信道情况确定各轮发送批次码的数量, 将第 i 轮编码器发送的批次码数量设为 N_i 。为方便分析, 本文将发送批次码的数量分为首轮发送和接收反馈后发送两种情况, 发送个数分别为 $\{N_1, N_2\}$ 。在发送端进行的外码编码算法流程可以用下述伪代码来进行描述。

基于反馈和发送端的统计得到的重要信息所改进的 BATS 码外码编码算法

输入: $N_1, N_2, \text{feedback}, B = [b_1, b_2, \dots, b_K]$

输出: $X_i, i = 1, 2, \dots, N_i$

if $i = 1$ do

$N_i \leftarrow N_1$

$X_i = B_i G_i, i = 1, 2, \dots, N_1, B_i \subset B$

else $i \neq 1$ do

$N_i \leftarrow N_2$

if feedback = ACK do

break

```

else feedback = num do
 $X_i = B_i G_i, i = 1, B_i \in [b_{\text{num}}, \bar{B}]$ 
 $X_i = B_i G_i, i = 2, 3, \dots, N_2, B_i \in B$ 
end if
end if

```

接收端在接收到第*i*轮发送的批次码后, 将其与之前接收到且未译出的批次码及已译出的输入数据包进行联合译码, 译码后根据不可译的批次码的包头信息统计出具有最大度值的输入数据包, 并将其包号反馈回发送端。

在编码器的编码过程中, 文献 [10] 对发送端接收到反馈度值最大的数据包包号所对应的信息进行重传后, 中继节点对重传数据包采用的是存储转发操作。如果无线网络的跳数较少或者信道条件较好时, 采用这种方式, 接收端能较好地接收到该数据包。然而, 随着网络跳数的增加或信道环境较差时, 接收端很有可能无法接收到该重要信息包的信息。如考虑一个删除概率为 0.2 的多跳删除网络, 跳数为 *p* 时, 一个数据包在接收端被接收到的概率为 $(0.8)^p$, 跳数越高, 接收到它的概率越低。对一个 *k* 跳的线性网络的最大可达速率进行仿真分析, 如图 4 所示, 如果对重要信息包在中继节点直接进行存储转发, 随着线性网络跳数的增大, 可达速率呈快速下降趋势。如果在中继节点对重要信息包采用随机线性网络编码算法进行编码再转发至下一个节点, 那么可达速率随跳数的增加而下降的速度将变得缓慢。因此, 在信道条件较差时, 也能较好地保证重要信息在接收端具有较高的译码成功率。图 4 中, *M* 代表批次码的大小, *q* 为有限域的大小。

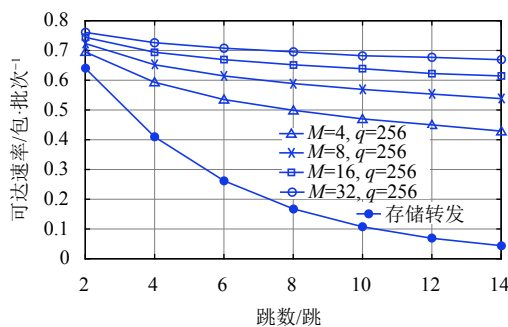


图 4 最大可达速率随跳数的变化趋势

在改进的编译码方案中, 编码器在进行第*i*轮编码时, N_2 个批次码转发至中继节点后, 需要采用随机线性网络编码算法对所有批次码进行内码编码, 再由中继节点转发至下一个中继节点或接收端。需要注意的是, 由重要信息产生的批次码和一

般输入数据包产生的批次码在进行内码编码时, 再编码的批次码长度可以不同^[12-14]。由图 4 所示, 再编码的批次长度越大, 码字的最大可达速率越大, 但同时编码开销也越大。

为减少开销, 提高编码灵活性, 重要信息包对应的批次码长度可随信道环境进行调整。如在信道条件较好时, 由于重要信息包数量较少, 内码编码的长度可采用较小值。改进的基于重要信息反馈的 BATS 码编译码方案的具体操作步骤如下。

- 1) 编码器进行第一轮 BATS 码外码编码, 生成 N_1 个长度为 M_1 的批次码并依次发送至中继节点;
- 2) 中继节点接收到批次码后, 对所有接收到的批次码在同一个批次码内部采用随机线性网络编码算法进行内码编码, 其中, 由重要信息包生成的批次码, 内码编码长度为 M_2 , 其余批次码内码编码长度为 M_1 。完成内码编码后, 将这些批次码按顺序转发至下一个中继节点或接收端, 如果转发至中继节点, 编译码流程转至步骤 2), 如果转发至接收端, 编译码流程转至步骤 3);
- 3) 接收端接收到 N_1 或 N_2 个批次码后, 开始采用 BP 译码算法进行译码, 如果将所有的原始输入数据包译出则表示译码成功, 返回一个 ACK 信号。否则, 译码器开始对所有不可译的批次码进行分析和统计, 找出具有最大度值且未译出的输入数据包, 并将其包号反馈回发送端, 编译码流程转至步骤 4);
- 4) 编码器在进行第*i*轮编码前, 首先分析反馈信息, 如果是 ACK 信号, 编码过程结束, 如果为一个数据包包号, 编码器对所有输入数据包进行分析和统计, 找出未参与编码的数据包, 并将其与反馈包组成集合 *I*, 采用外码编码算法对 *I* 进行编码, 产生第一个批次码, 随后对所有原始数据包进行随机编码产生剩下所需的 $N_2 - 1$ 个批次码, 两种类型的批次码编码长度均为 M_1 , 将它们发送至中继节点, 编译码流程转至步骤 2)。

3 方案性能仿真与分析

在仿真性能分析中, 对传统 BATS 码、基于重要信息反馈 BATS 码以及本文算法的译码性能分别进行了仿真和比较。在仿真模型中, 假设存在理想反馈信道, 反馈距离较短, 反馈时延较小, 且由于反馈包数较少, 为方便分析, 可忽略反馈开销对系统性能造成的影响。仿真参数设置如下: 在内外码编码中, 编码系数在有限域 GF(256) 中进行选择。为方便分析, 批次码内外码编码尺寸统一选择为

$M_1 = M_2 = 8$, 数据包长度为 $L = 2$ 。外码编码度分布按照文献 [15] 中有限长 BATS 码度分布设计方法生成。3 种方案的译码算法均采用 BP 译码算法, 仿真次数设为 100。编码冗余定义为:

$$\varepsilon = \frac{M(N_1 + fN_2)}{K} \quad (3)$$

式中, f 为反馈轮数。译码成功率指的是译码完成后, 已恢复的包数占有所有包数的比例。

不同码长下改进方案性能仿真分析如下: 在以上仿真参数的设置下, 考虑一个两跳的删除信道, 每一条链路的删除概率设为 0.1, 信道秩分布为 $h = [0, 0, 0, 0.0008, 0.0092, 0.0648, 0.2646, 0.4760, 0.1846]$, 设置不同的原始数据长度 $K = [500, 1000]$, 分别对 3 种方案在不同码长下的译码性能进行仿真, 发送的批次码数量设为 $N_1 = 125$, $N_2 = 7$, 仿真结果如图 5 所示。3 种编译码方案的译码成功率随着编码冗余的增加都在上升, 但改进的基于重要信息反馈 BATS 码上升的速度相对其他方法更快。另外, 码长越长, 在相同的编码冗余下译码成功率越高。

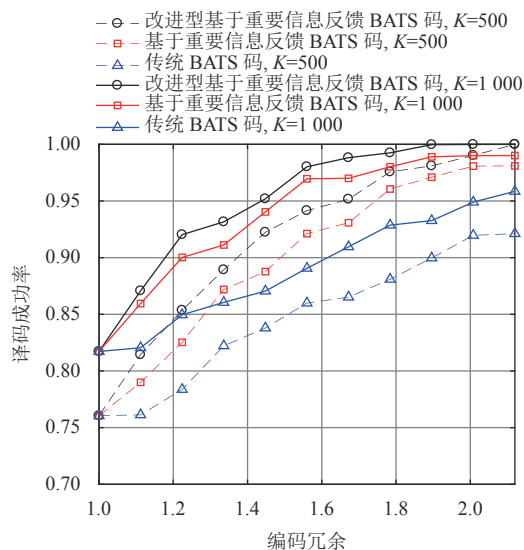


图 5 不同码长下 BATS 码优化方案的性能比较

不同反馈步长下改进方案性能仿真分析如下: 固定码长 $K = 500$, $N_1 = 125$, 设置不同的反馈步长 $N_2 = 3, 7, 14$, 对本文改进的基于重要信息反馈的 BATS 码编译码方案进行了译码性能的仿真, 仿真结果如图 6 所示, 改进算法比传统 BATS 码的性能更好。通过对比 3 组不同反馈步长的译码成功率, 可以发现随步长的减小, 译码成功率快速上升。步长越小, 接收端获得的反馈信息越多, 对重要信息的编码次数也就越多, 译码性能越好。

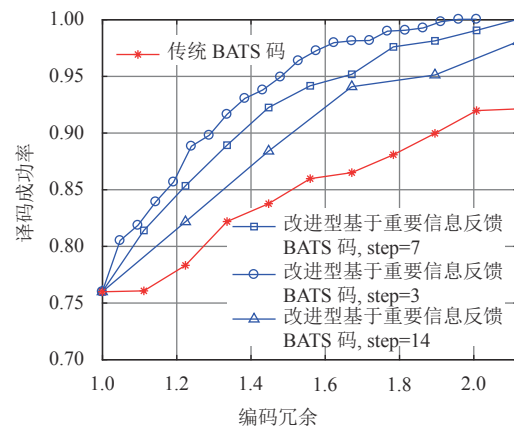


图 6 不同反馈步长下 BATS 码优化方案的性能比较

不同信道条件下改进方案性能仿真分析如下: 固定码长 $K = 500$, 发送的批次码数量设为 $N_1 = 125$, $N_2 = 7$, 编码冗余固定为 $\varepsilon = 2.12$, 图 7 对 3 种编译码方案在多跳删除信道条件下的译码性能进行了仿真。将每一跳链路的删除概率设为 0.1。仿真结果如图 7 所示, 译码成功率随信道跳数的增加在不断下降, 但本文的编译码方案译码成功率随跳数的增加, 下降的速度更为缓慢。

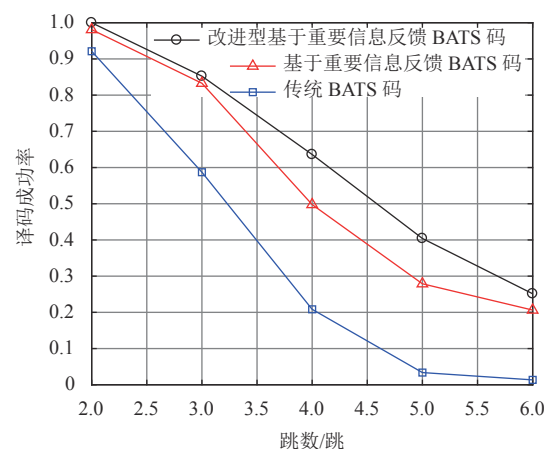


图 7 不同跳数下 BATS 码优化方案的性能比较

表 1 对比了 3 种方案的译码成功次数, 每种方案总的仿真次数为 100, 由表 1 可见, 本文改进方案能够获得更多的译码成功次数。在本文方案中, 每进行一次反馈, 编码器不仅将上一轮未参与编码的输入数据包与反馈数据包作为一个子集合进行外码编码, 同时在经过中继节点时, 还对这部分产生的批次码进行再编码。这样不仅能保证所有数据包都参与编码, 同时使重要信息包在传输过程中得到保护, 因此本文方案能获得更高的完全译码成功概率且随跳数的增加下降的速度最为缓慢。

表 1 3 种编译码方案译码成功次数

跳数/跳	改进方案/次	基于重要 信息反馈/次	传统方案/次
2	92	78	43
3	80	69	29
4	59	42	12
5	37	24	0
6	21	17	0

4 结束语

本文设计了一种改进型的基于重要信息反馈的 BATS 码编译码方案。与传统基于重要信息反馈编译码方案相比, 本文方案更新了重要信息包的内容, 为提高完全译码成功概率, 每次反馈后将之前未参与编码的原始数据包与译码后具有最大度值的原始数据包共同作为重要信息包, 参与到下一次的编码中去。同时, 为了保证重要信息包在多跳删除信道或具有较差信道环境中进行传输时仍能在接收端具有较高的译码恢复率, 本文方案考虑在中继节点对重要信息包产生的批次码采用网络编码算法进行再编码。仿真结果表明, 本文算法与传统 BATS 码以及基于重要信息反馈 BATS 码相比, 具有更高的译码成功率, 较好地实现了设计目标。

参 考 文 献

- [1] DONG X P, ZHANG Y, SONG J. The reliability-enhanced wireless networks through BATS codes[C]//The 18th IEEE International Symposium on Consumer Electronics. Republic of Korea: IEEE, 2014: 1-4.
- [2] LUBY M. LT codes[C]//The 43rd Annual IEEE Symposium on Foundations of Computer Science. Vancouver: IEEE, 2002: 271-282.
- [3] YANG S H, YEUNG R W. Coding for a network coded fountain[C]//2011 IEEE International Symposium on Information Theory Proceedings. St Petersburg: IEEE, 2011: 2647-2651.
- [4] YANG S H, YEUNG R W. Batched sparse codes[J]. *IEEE Transactions on Information Theory*, 2014, 60(9): 5322-5346.
- [5] YANG J, SHI Z P, XIONG J, et al. An improved BP decoding of BATS codes with iterated incremental Gaussian elimination[J]. *IEEE Communications Letters*, 2020, 2(24): 321-324.
- [6] YANG S H, YEUNG R W, SRIKANT R. BATS codes: Theory and practice[M]. [S.l.]: Morgan & Claypool, 2017.
- [7] YANG S H, ZHOU Q Q. Tree analysis of BATS codes[J]. *IEEE Communications Letters*, 2016, 1(20): 37-40.
- [8] 于佳琪. 数字喷泉码及其在广域分集网中的应用研究[D]. 杭州: 浙江大学, 2013.
- [9] YU J Q. Research on digital fountain codes and its application on wide area diversity network[D]. Hangzhou: Zhejiang University, 2013.
- [9] 陈美玲. 带反馈 LT 码的研究与应用[D]. 郑州: 郑州大学, 2015.
- [9] CHEN M L. Research and application of LT code with feedback[D]. Zhengzhou: Zhengzhou University, 2015.
- [10] JIA D, FEI Z S, SHANGGUAN C L, et al. LT codes with limited feedback[C]//2014 IEEE International Conference on Computer and Information Technology. Xi'an: IEEE, 2014: 669-673.
- [11] XIANG M, YI B S, TAN J J. Optimised design and tree analysis of regularised variable-node BATS codes[J]. *Electronics Letters*, 2020, 56(12): 630-633.
- [12] TANG B, YANG S H, YE B L, et al. Near-optimal one-sided scheduling for coded segmented network coding[J]. *IEEE Transactions on Computers*, 2016, 65(3): 929-939.
- [13] YIN H F, YANG S H, ZHOU Q Q, et al. Adaptive recoding for BATS codes[C]//2016 IEEE International Symposium on Information Theory Proceedings. Barcelona: IEEE, 2016: 2349-2353.
- [14] ZHOU Z H, LI C D, YANG S H, et al. Practical inner codes for BATS codes in multi-hop wireless networks[J]. *IEEE Transactions on Vehicular Technology*, 2019, 68(3): 2751-2762.
- [15] YANG S H, NG T, YEUNG R W. Finite-length analysis of BATS Codes[J]. *IEEE Transactions on Information Theory*, 2018, 64(1): 322-348.

编辑 刘飞阳