



## 一种基于循环神经网络的极化码BP译码算法

何彦琦<sup>1,2</sup>, 彭大芹<sup>1,2</sup>, 赵雪志<sup>1,2</sup>

(1.重庆邮电大学 通信与信息工程学院,重庆 400065; 2.重庆邮电大学 电子信息与网络工程研究院,重庆 400065)

**摘要:** 置信传播(BP)算法作为极化码最常用的软判决输出译码算法之一,具有并行传输、高吞吐量等优点,但其存在收敛较慢、运算复杂度高等缺陷。提出一种基于循环神经网络的偏移最小和近似置信传播译码算法。通过偏移最小和近似算法替代乘法运算,修改迭代过程中的消息更新策略,并运用改进的循环神经网络架构实现参数共享。仿真结果表明,相比传统BP译码算法,该译码算法在提升误码率(BER)性能的前提下,减少约75%的加法运算且收敛速度大幅提升,相比基于深度神经网络的BP译码算法,该算法在确保BER性能无显著下降的前提下,使用加法运算替代乘法运算,节省了约80%的存储空间开销。

**关键词:** 极化码;置信传播;循环神经网络;偏移最小和;运算复杂度

开放科学(资源服务)标志码(OSID):



**中文引用格式:** 何彦琦, 彭大芹, 赵雪志. 一种基于循环神经网络的极化码BP译码算法[J]. 计算机工程, 2022, 48(1): 197-203.

**英文引用格式:** HE Y Q, PENG D Q, ZHAO X Z. A recurrent neural network based BP decoding algorithm for polar codes[J]. Computer Engineering, 2022, 48(1): 197-203.

## A Recurrent Neural Network Based BP Decoding Algorithm for Polar Codes

HE Yanqi<sup>1,2</sup>, PENG Daqin<sup>1,2</sup>, ZHAO Xuezh<sup>1,2</sup>

(1.School of Communication and Information Engineering, Chongqing University of Posts and Telecommunications, Chongqing 400065, China; 2.Institute of Electronic Information and Network Engineering, Chongqing University of Posts and Telecommunications, Chongqing 400065, China)

**[Abstract]** Belief Propagation (BP) is one of the most commonly used soft decision decoding algorithms for polar codes, which enables parallel transport and displays high throughput. However, BP suffers from slow convergence and high computational complexity. To address the problem, this paper proposes a Recurrent Neural Network (RNN)-based approximate BP decoding algorithm for polar codes with and Offset Min-Sum (OMS). The message update strategy is modified in the iterative process by replacing multiplication operations with the minimum offset and approximation algorithm, and an improved RNN architecture is used to realize parameter sharing. Simulation results show that compared with traditional BP algorithms, the proposed decoding algorithm can reduce addition operations by about 75% and greatly improve the convergence speed and the Bit Error Ratio (BER) performance. Compared with the DNN-BP decoding algorithm, the proposed algorithm uses addition operation to replace multiplication operation, and saves about 80% storage space overhead with no significant decline in BER performance.

**[Key words]** polar codes; Belief Propagation (BP); Recurrent Neural Network (RNN); Offset Min-Sum (OMS); computational complexity

**DOI:** 10.19678/j.issn.1000-3428.0060159

### 0 概述

极化码<sup>[1]</sup>自2009年由ARIKAN提出以来,因其在无记忆信道中被证明可以达到香农极限信道容量的特性,以及编解码的低复杂度而备受关注<sup>[2]</sup>。由于极化码在短块长度下具有良好的误码率(Bit

Error Rate, BER)性能,因此在第五代移动通信技术(5th Generation Mobile Communication Technology, 5G)中被采用为控制信道编码方案。同时,极化码的译码算法得到学者的广泛研究,其中,串行抵消(Successive Cancellation, SC)和置信传播(Belief Propagation, BP)是极化码最常用的两种译码算法。

**基金项目:** 国家自然科学基金“基于SIW圆极化赋形阵列的毫米波可重构切换波束天线研究”(E020B2018023)。

**作者简介:** 何彦琦(1996—),男,硕士研究生,主研方向为极化码译码、SCMA多用户检测;彭大芹(通信作者),正高级工程师;赵雪志,硕士研究生。

**收稿日期:** 2020-12-01

**修回日期:** 2021-01-05

**E-mail:** 501243867@qq.com

相比于SC译码算法<sup>[3]</sup>以及在其基础上改进的串行抵消列表(Successive Cancellation List, SCL)译码算法<sup>[4]</sup>,BP译码算法的并行译码过程具有更高的吞吐量及更低的时延<sup>[5-6]</sup>,因此更能满足5G场景的需求。但随着迭代次数及码长的增加,运算复杂度大幅提高,即使采用最小和近似<sup>[7]</sup>和早期终止准则<sup>[8]</sup>来降低复杂度和冗余迭代次数,BP译码算法在有限的迭代中仍然表现不理想。

近年来,基于神经网络的译码器受到了人们广泛的关注。为改善BP译码算法的性能,国内外研究者对此进行了大量研究并取得了显著的进展。文献[9-10]将BP迭代译码结构运用到神经网络架构中,并通过深度学习优化最小和近似后用于弥补性能损失的缩放因子,大幅降低了迭代次数并提升了误码率性能。对神经网络译码算法的进一步优化主要从增强译码BER性能和降低运算复杂度或时延两个方面进行。在提升BER性能方面,文献[11-12]结合循环冗余校验(Cyclic Redundancy Check, CRC)与BP译码算法两者的因子图及神经网络,取得了接近于CA-SCL译码算法的性能,文献[13]在此基础上将神经网络架构修改为长短期记忆网络(Long Short-Term Memory, LSTM)以平衡译码性能与运算复杂度。在降低运算复杂度及时延方面,文献[14]结合分块方法及深度神经网络改善了长码情况下的译码性能,文献[15]则是通过卷积神经网络(Convolutional Neural Network, CNN)对神经网络架构中的值传递过程进行简化以降低运算复杂度,文献[16]提出一种残差网络降噪译码算法并完成了硬件实现,取得了极低的译码时延。

但以上方法大部分仍存存在2个待解决的关键问题:额外的权值存储开销,随着极化码码长的增加与迭代次数的影响,成倍增长的权重因子导致了相当大的存储空间开销,阻碍了神经网络在实际通信系统中的应用;大量的乘法运算,与使用最小和近似的BP算法相比,深度学习的辅助使得算法具有更好的性能。但以上算法大部分添加了相乘的缩放因子,增加了运算复杂度。

针对神经网络译码器存在的问题,受BCH(Bose-Chaudhuri-Hocquenghem)译码所采用的偏移最小和近似(Offset Min-Sum, OMS)BP译码算法<sup>[17-18]</sup>,以及循环神经网络(Recurrent Neural Network, RNN)译码架构<sup>[19]</sup>的启发,本文对极化码的BP译码算法和深度学习进行研究,提出一种基于循环神经网络的偏移最小和置信传播(RNN-OMS-BP)译码算法。通过循环神经网络结构实现神经网络多次迭代之间的参数共享,并运用偏移最小和近似算法修改迭代过程中的消息更新策略,以实现低运算复杂度、低内存占用的极化码译码算法。

## 1 极化码与BP译码算法

### 1.1 极化码

极化码是一种通过信道极化原理构造的线性分组码,理论上可以达到香农容量,其原理可简述为对于 $N=2^n$  ( $n$ 为任意正整数)的极化码,根据信道可靠性将信道分裂为两部分,一部分信道分裂为信道容量趋于1的无噪声信道,用于传输信息比特,另一部分则分裂为完全噪声信道,在发送端和接收端都设置为已知的冻结比特。通过生成矩阵实现包含 $K$ 位信息比特、 $(N-K)$ 位冻结比特、码率为 $K/N$ 的 $(N, K)$ 极化码,定义如下:

$$\mathbf{x}^N = \mathbf{u}^N \mathbf{G}_N = \mathbf{u}^N \mathbf{F}^{\otimes n} \mathbf{B}_N \quad (1)$$

其中: $n=\lg N$ ;  $\mathbf{B}_N$ 表示比特反序矩阵;  $\mathbf{F}^{\otimes n}$ 表示对 $\mathbf{F} = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$ 的 $n$ 次克罗内克积;长度的原始比特序列 $\mathbf{u}^N$ 通过与生成矩阵 $\mathbf{G}_N$ 相乘即可得到长度为 $N$ 的编码比特序列 $\mathbf{x}^N$ 。

### 1.2 极化码的BP译码算法

置信传播(BP)是一种广泛应用的消息传递译码算法,主要应用于低密度匹配校验码(Low Density Parity Check Code, LDPC)和BCH码。但不同于以上2种信道编码的BP译码过程是基于奇偶校验矩阵进行消息更新的,极化码的BP译码算法的消息更新过程是在因子图上迭代进行的,并且采用类似SC译码算法的连续消除规则<sup>[20]</sup>。该因子图一共由 $n=\lg N$ 个阶段及 $N \times (n+1)$ 个节点组成。一个(8,4)极化码的BP译码因子图如图1所示。

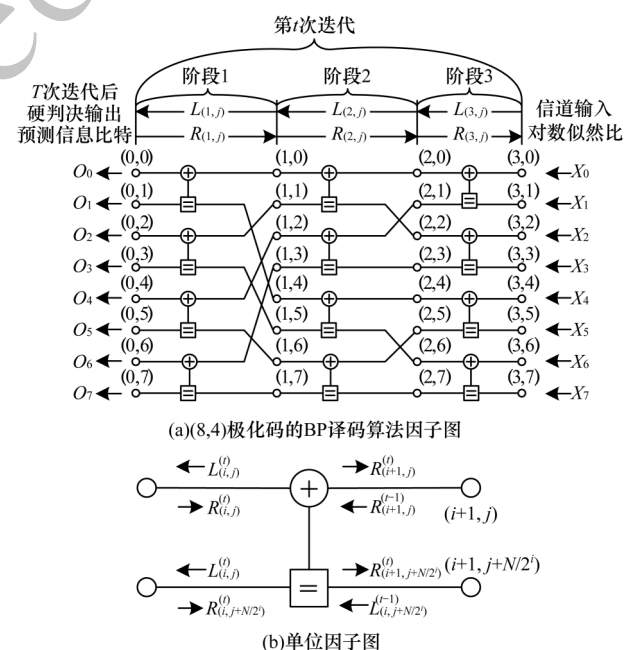


图1 极化码的BP译码算法因子图

Fig.1 Factor graph of BP decoding algorithm for polar code

在图1中, 下标  $i, j$  分别表示阶段数和节点数, 上标  $t$  表示当前迭代次数, 每个节点  $(i, j)$  传递两类对数似然比 (Likelihood Ratio, LLR) 信息: 从左至右传递的右信息对数似然比  $R_{i,j}^{(t)}$  和从右至左传递的左信息对数似然比  $L_{i,j}^{(t)}$ 。用公式表示如下:

$$\begin{cases} L_{i,j}^{(t)} = g(L_{i+1,j}^{(t-1)}, L_{i+1,j+N/2^t}^{(t-1)} + R_{i,j+N/2^t}^{(t-1)}) \\ L_{i,j+N/2^t}^{(t)} = g(R_{i,j}^{(t)}, L_{i+1,j}^{(t-1)}) + L_{i+1,j+N/2}^{(t-1)} \\ R_{i,j}^{(t)} = g(R_{i,j}^{(t-1)}, L_{i+1,j+N/2^t}^{(t-1)} + R_{i,j+N/2^t}^{(t-1)}) \\ R_{i,j+N/2^t}^{(t)} = g(R_{i,j}^{(t)}, L_{i+1,j}^{(t-1)}) + R_{i,j+N/2^t}^{(t-1)} \end{cases} \quad (2)$$

其中: 函数  $g(a, b) = \ln \frac{1+ab}{a+b}$ 。为降低运算复杂度, 该函数一般使用的最小和 (Min-Sum, MS) 函数近似如下:

$$g(a, b) \approx \text{sign}(a) \times \text{sign}(b) \min(|a|, |b|) \quad (3)$$

在迭代开始时, 两类对数似然比信息初始化如下:

$$R_{1,j}^{(1)} = \begin{cases} 0, & j \in \mathcal{A} \\ \infty, & j \in \mathcal{A}^c \end{cases} \quad (4)$$

$$L_{n+1,j}^{(1)} = \ln \frac{p(y_j|x_j=0)}{p(y_j|x_j=1)} \quad (5)$$

其中:  $\mathcal{A}$  表示信息比特集合;  $\mathcal{A}^c$  表示冻结比特集合;  $y_j$  表示从信道中接收到的第  $j$  位信息比特。经过预设的  $T$  次迭代之后, 第  $j$  位估计信息比特可通过对最后一次迭代输出的左信息对数似然比  $L_{1,j}^{(T)}$  进行硬判决, 得到:

$$o_j^N = \begin{cases} 0, & L_{1,j}^{(T)} \geq 0 \\ 1, & L_{1,j}^{(T)} < 0 \end{cases} \quad (6)$$

### 1.3 深度神经网络BP译码算法

最小和近似虽然有效地降低了复杂度, 但也会导致误码率性能损失。文献[10]提出了基于深度神经网络的BP译码算法, 该网络架构可以看作极化码因子图的展开结构, 并在消息更新因子图中设置了可学习的参数, 从而弥补最小和近似带来的性能损失。式(2)被替换如下:

$$\begin{cases} L_{i,j}^{(t)} = Lv_{i,j}^{(t)} \cdot g(L_{i+1,j}^{(t-1)}, L_{i+1,j+N/2^t}^{(t-1)} + R_{i,j+N/2^t}^{(t-1)}) \\ L_{i,j+N/2^t}^{(t)} = Lv_{i,j+N/2^t}^{(t)} \cdot g(R_{i,j}^{(t)}, L_{i+1,j}^{(t-1)}) + L_{i+1,j+N/2}^{(t-1)} \\ R_{i,j}^{(t)} = Rv_{i+1,j}^{(t)} \cdot g(R_{i,j}^{(t-1)}, L_{i+1,j+N/2^t}^{(t-1)} + R_{i,j+N/2^t}^{(t-1)}) \\ R_{i,j+N/2^t}^{(t)} = Rv_{i+1,j+N/2^t}^{(t)} \cdot g(R_{i,j}^{(t)}, L_{i+1,j}^{(t-1)}) + R_{i,j+N/2^t}^{(t-1)} \end{cases} \quad (7)$$

其中:  $Lv_{i,j}^{(t)}$  和  $Rv_{i,j}^{(t)}$  分别表示在第  $t$  次迭代中第  $i$  个阶段、第  $j$  个节点从右至左传递的左信息和从左至右传递的右信息的缩放因子。

在  $T$  次迭代完成后, 使用交叉熵损失函数评估神经网络通过激活函数 (如 sigmoid 函数) 输出的预

测信息比特  $o_j$  与传输信息比特  $u_j$  的误差, 即:

$$L(\mathbf{u}, \mathbf{o}) = -\frac{1}{N} \sum_{j=1}^N u_j \log_a o_j + (1 - u_j) \log_a (1 - o_j) \quad (8)$$

该模型的最终目标是通过反向传播算法计算损失函数梯度, 并使用随机梯度下降 (Stochastic Gradient Descent, SGD) 算法最小化损失函数, 确定最优缩放因子的值。

## 2 基于循环神经网络的偏移最小和近似BP译码算法

### 2.1 偏移最小和近似

DNN-BP 可以更好地解决最小和的近似损失问题, 但是该深度神经网络架构仍然存在2个问题:

1) 引入缩放因子为算法增加了大量的乘法运算, 运算复杂度大幅增加。

2) 存储空间被该深度架构中大量的缩放因子的存储所占用。这2个问题都阻碍了DNN-BP在现实通信系统中的部署。

本文提出一种基于循环神经网络的偏移最小和近似置信传播 (RNN-OMS-BP) 译码算法用于极化码译码, 以实现乘法缩放因子的替代以及不同次迭代之间的权值共享。

使用偏移最小和近似算法<sup>[18]</sup>代替缩放因子最小和近似算法, 式(3)被替换如下:

$$g'(a, b, \beta) = \text{sign}(a) \text{sign}(b) \max(\min(|a|, |b|) - \beta, 0) \quad (9)$$

其中:  $\beta$  为偏移量。

为改善系统性能, 因子图中不同节点应用不同的偏移量, 式(7)更新如下:

$$\begin{cases} L_{i,j}^{(t)} = g'(L_{i+1,j}^{(t-1)}, L_{i+1,j+N/2^t}^{(t-1)} + R_{i,j+N/2^t}^{(t-1)}, \beta_{L_{i,j}^{(t)}}) \\ L_{i,j+N/2^t}^{(t)} = g'(R_{i,j}^{(t)}, L_{i+1,j}^{(t-1)}, \beta_{L_{i,j+N/2^t}^{(t)}}) + L_{i+1,j+N/2}^{(t-1)} \\ R_{i,j}^{(t)} = g'(R_{i,j}^{(t-1)}, L_{i+1,j+N/2^t}^{(t-1)} + R_{i,j+N/2^t}^{(t-1)}, \beta_{R_{i,j}^{(t)}}) \\ R_{i,j+N/2^t}^{(t)} = g'(R_{i,j}^{(t)}, L_{i+1,j}^{(t-1)}, \beta_{R_{i,j+N/2^t}^{(t)}}) + R_{i,j+N/2^t}^{(t-1)} \end{cases} \quad (10)$$

其中:  $\beta_{L_{i,j}^{(t)}}$  和  $\beta_{R_{i,j}^{(t)}}$  分别代表第  $t$  次迭代过程的左信息  $L_{i,j}^{(t)}$  和右信息  $R_{i,j}^{(t)}$  计算过程中添加的偏移量。当偏移量  $\beta=0$  时, 该算法被简化为普通的最小和近似BP译码算法。

### 2.2 循环神经网络架构

使用偏移量替代缩放因子后, 解决了引入大量乘法运算的问题。但大量的可学习偏移量仍然占用了较大的存储空间。因此, 受到文献[19]的启发, 本文基于上述的偏移最小和近似算法提出改进的循环神经网络架构, 代替文献[10]所采用的前馈神经网络架构。以



(4,2)极化码为例的网络架构如图2所示。

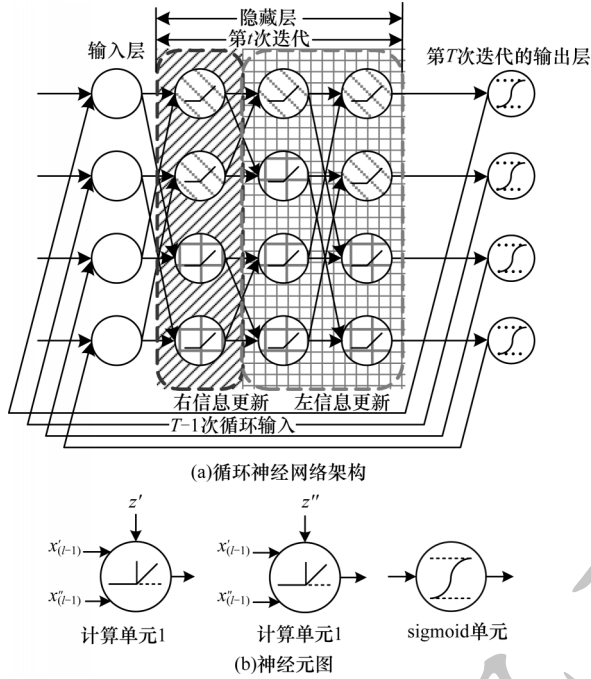


图2 循环神经网络架构及神经元图

Fig.2 Recurrent neural network architecture and neuronal graphs

在图2(a)神经网络结构中,右信息以及左信息传递的过程分别展开为 $(n-1)$ 层和 $n$ 层由 $N$ 个神经元组成的隐藏层,因此,一次完整的迭代转换为 $(2n-1)$ 层隐藏层。本文架构与DNN-BP架构<sup>[10]</sup>的改进主要有以下2个方面:

1)在每层隐藏层的计算单元中因为偏移量最小和近似加入了 $\max$ 函数,因此相当于在隐藏层中的每个计算单元中添加了线性整流函数(Rectified Linear Unit, ReLU)作为激活函数,该函数公式如下:

$$f_{\text{ReLU}}(x) = \max(0, x) \quad (11)$$

2)在每 $(2n-1)$ 层隐藏层的运算完成后,进行循环输入,实现偏移量参数集 $\beta$ 的复用。

图2(b)中计算单元1、计算单元2分别用于承担式(10)中第1、3式及第2、4式与sigmoid函数的计算,分别记为 $f_I(x_{l-1}; \beta_l)^{(0)}$ 、 $f_{II}(x_{l-1}; \beta_l)^{(0)}$ 和 $o(x_L)$ ,式(10)在神经网络中的简化公式如下:

$$\begin{cases} f_I(x_{l-1}; \beta_l)^{(0)} = g'(x'_{l-1}, x''_{l-1} + z', \beta_l^I) \\ f_{II}(x_{l-1}; \beta_l)^{(0)} = g'(x'_{l-1}, z'', \beta_l^{II}) + x''_{l-1} \\ o = o(x_L) = \sigma(x_{L-1}) \end{cases} \quad (12)$$

其中: $x'_{l-1}$ 、 $x''_{l-1}$ 分别表示在第 $l-1$ 层中的两个同类输入信息集; $z'$ 、 $z''$ 分别表示两种运算中不同类的输入信息集; $\beta_l^I$ 、 $\beta_l^{II}$ 分别表示两种运算中的偏移量集。且因为该循环神经网络架构的特性,偏移量 $\beta$ 在同一次迭代中,对于图1(a)中的每一个节点 $(i, j)$ 各不相同,但在不同次迭代之间,同一节点的偏移量 $\beta_{i,j}$ 则是共享的。该网络架构在最后一次迭代完成后使用sigmoid单元进行对数似然比硬判决输出预测比特

$o_j$ ,最终使用合理的优化方法(如AdaGrad、梯度下降等)最小化预测比特 $o_j$ 与传输比特 $u_j$ 之间的损失函数,以获得该网络架构下的最优偏移量集 $\beta_{\text{best}}$ 。

选择合理的损失函数对训练过程的效率和运算复杂度都有明显的帮助,DNN-BP算法采用如式(8)所示的交叉熵损失函数作为损失函数,交叉熵函数无疑是契合于该应用场景的损失函数。但交叉熵损失函数同样也存在不够硬件友好的缺点,在求导计算中需要计算sigmoid函数,而sigmoid激活函数的运算时间较长且更为复杂。因此,本文尝试采用Hinge损失函数对交叉熵损失函数进行替代,该损失函数公式如下:

$$L_{\text{hinge}}(\mathbf{u}, \mathbf{o}) = \frac{1}{N} \sum_{j=1}^N \max(0, 1 - u_j o_j) \quad (13)$$

通过合理的损失函数选择,可有效减少训练复杂度及时间,而几乎未带来可见的性能下降。

### 2.3 译码算法伪代码

算法 RNN-OMS-BP译码算法

输入 模型参数设置(包括码长 $N$ 、码率 $R_c$ 、信噪比范围SNR、batch\_size、学习率 $\eta$ 等),随机生成码字集

输出 预测信息比特集 $\mathbf{o}$ ,损失函数值 $L$ ,0~5 dB下的译码BER

1.初始化:生成用于训练和验证的码字集 $\mathbf{x}$ 为通过高斯白噪声加扰的码字对数似然比序列, $\mathbf{y}$ 为传输比特序列。

for  $i = 1, 2, \dots, \text{SNR}(=5)$  do

Generate:  $\mathbf{x} = \text{llr}$ ,  $\mathbf{y} = \mathbf{u}$

end for

2.RNN-OMS-BP算法:将占位符(placeholder)带入以RNN架构展开的因子图中,计算出所有加入偏移量 $\beta$ 后的左信息 $L_{i,j}^I$ 和右信息 $R_{i,j}^I$ 。

for  $i = 4, 3, \dots, 0$  do

for  $j = 16, 15, \dots, 0$  do

计算 $L_{i,j}^I$ 和 $R_{i,j}^I$ ,按照图2 RNN架构及式(10)

end for

end for

3.设置损失函数loss及优化器optimizer进行训练及验证过程,将生成的码字序列 $\mathbf{x}$ 及传输比特序列 $\mathbf{y}$ 输入到占位符,运行会话输出预测比特序列 $\mathbf{o}$ 及损失函数值 $L$ :

$\mathbf{x} = L_I^T$

loss = reduce\_mean(式(8)或式(13))

for  $i = 1, 2, \dots, \text{epochs} = 25$  do

$\mathbf{o}, L = \text{sess.run}(\mathbf{x}, \mathbf{y}, \text{optimizer}, \text{loss})$

end for

4.进行测试,将最优偏移量集 $\beta_{\text{best}}$ 应用到网络架构中,输入生成的测试码字序列 $\mathbf{x}_{\text{test}}$ 及相应的传输比特序列 $\mathbf{y}_{\text{test}}$ ,输出测试预测比特序列 $\mathbf{o}_{\text{test}}$ 及相应的损失函数值 $L_{\text{test}}$ :

$\mathbf{o}_{\text{test}}, L_{\text{test}} = \text{sess.run}(\mathbf{x}_{\text{test}}, \mathbf{y}_{\text{test}}, \text{loss})$

最终计算出BER以评估译码算法性能。

### 3 仿真结果与分析

为评估本文所提出的RNN-OMS-BP译码算法的性能,分别通过仿真将其与原始BP译码算法、DNN-BP译码算法在误码率(Bit Error Rate, BER)、

收敛速度及运算复杂度方面进行对比。表1所示为仿真中所用的实验参数。

表1 仿真参数设置			
Table 1 Simulation parameters setup			
设置选项	参数值	设置选项	参数值
测试平台	TensorFlow	训练集大小	$7.5 \times 10^4$
信道编码	极化码	测试集大小	$3 \times 10^3$
码长	16,64,128	优化算法	AdaGrad
码率	1/2	损失函数	交叉熵或Hinge
调制方式	BPSK	学习率	0.001
信噪比范围	1~5 dB	Epochs	25

3.1 BER性能分析

第一个实验是模拟在码长及码率相同的情况下,不同迭代次数下原始BP、DNN-BP与本文提出的RNN-OMS-BP译码算法的BER性能,观测迭代次数对神经网络译码器的架构设计和参数的调整有重要意义。仿真结果如图3所示(以(16,8)极化码为例)。

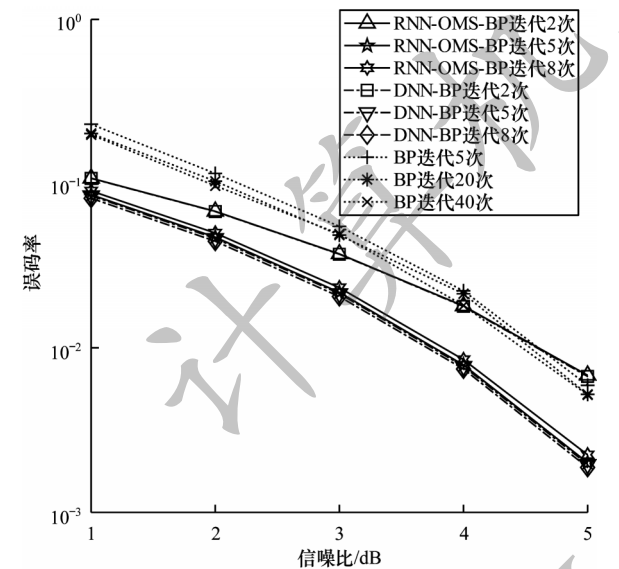


图3 不同迭代次数下3种译码算法BER性能

Fig.3 BER performance of three decoding algorithms under different iteration times

从图3可以看出,对于(16,8)极化码,DNN-BP和RNN-OMS-BP译码算法只需约5次迭代即可收敛,就能获得良好的性能。另一方面,传统的BP译码算法需要超过40次迭代才能达到与神经网络译码器迭代2次相近的BER性能。因此,传统BP算法消耗的计算资源较多,能耗高,时延较长。此外,RNN-OMS-BP与DNN-BP两种算法具有几乎相同的误码率,但RNN-OMS-BP算法计算复杂度与存储空间占用减少,更适合部署在实际通信系统中。

3.2 不同码长下的BER性能分析

信道编码码长对系统BER性能有显著影响,因此本文对(16,8)、(64,32)和(128,64)3种不同码长的RNN-OMS-BP译码算法、DNN-BP译码算法以及

传统BP译码算法BER性能进行仿真对比,其中,传统BP译码算法迭代次数为40,RNN-OMS-BP译码算法迭代次数为5,仿真结果如图4所示。

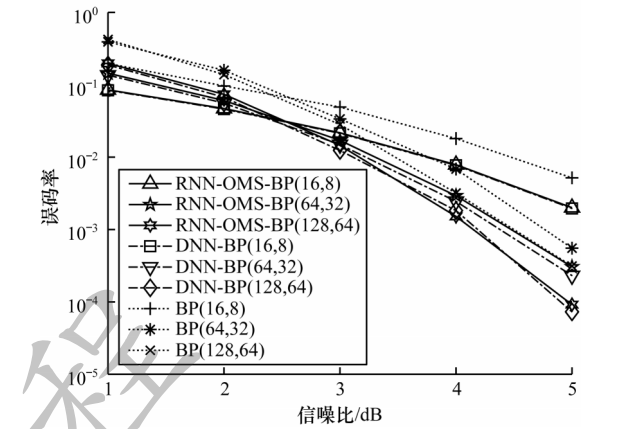


图4 不同码长下3种译码算法BER性能

Fig.4 BER performance of three algorithms under different code lengths

从图4可以看出,在(16,8)的短码长情况下,相比于传统BP译码算法,RNN-OMS-BP算法的BER性能得到明显改善,与DNN-BP算法的BER性能差距非常小;在码长为(64,32)和(128,64)的情况下,与传统BP译码算法相比,RNN-OMS-BP亦可以实现更低的误码率,且优势在高信噪比区域更加明显,与短码长情况类似,相较于DNN-BP译码算法BER性能下降非常微小。仿真结果与预测情况一致,即提出的RNN-OMS-BP算法在长短码情况下相较于传统BP算法均有显著改善,相较于DNN-BP算法,性能下降相当微小。

3.3 偏移量值的分布

本节以(64,32)极化码为例,对本文提出的RNN-OMS-BP译码算法中偏移量 $\beta$ 的值在训练中的变化过程进行仿真分析,该变化过程如图5所示(以(64,32)极化码为例)。

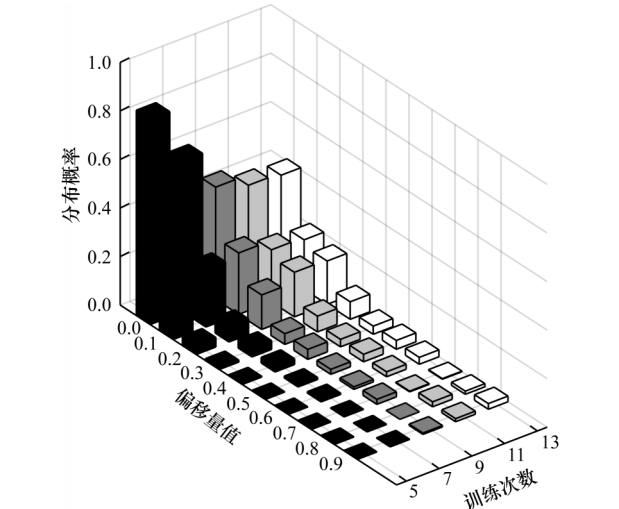


图5 偏移量值的分布直方图

Fig.5 Histogram of distribution of offset values

从图5可以看出,当训练次数即Epoch≈10时,偏移量值的更新过程基本收敛,所有值的分布近似于以0为中心的半正态分布形态,偏移量 $\beta$ 的值主要集中在0.0~0.4之间,因此本文在训练过程开始时将偏移量 $\beta$ 的初始值设置为0。

3.4 运算复杂度分析

表2的运算复杂度分析以(16,8)极化码为例,列举了在传统BP译码算法迭代 $T_{BP}=40$ 次、DNN-BP译码算法及本文提出的RNN-OMS-BP译码算法迭代 $T=5$ 次的情况下,三者分别在加法运算、乘法运算及存储空间占用3个方面的运算复杂度。

表2 运算复杂度分析			
Table 2 Analysis of computational complexity			
运算类型	BP	DNN-BP	RNN-OMS-BP
加法	$2T_{BP}N\text{lb } N\sim 120$ (100%)	$2TN\text{lb } N\sim 640$ (12.5%)	$4TN\text{lb } N\sim 1\ 280$ (25%)
乘法	0	$2TN\text{lb } N\sim 640$ (100%)	0
存储空间	0	$2TN\text{lb } N\sim 640$ (100%)	$2N\text{lb } N\sim 160$ (20%)

由表2可知,相比于传统BP译码算法,两种神经网络结构的BP译码算法将迭代次数减少80%以上,并且有效降低了最小和近似带来的BER性能损失。相比于DNN-BP译码算法,RNN-OMS-BP译码算法在同样经过5次迭代的前提下,以增加12.5%的加法运算量为代价,替代了所有的乘法运算。同时通过循环神经网络结构在迭代次数为 $T=5$ 时,减少了80%的存储空间占用。

除译码过程的运算复杂度分析外,本文还通过仿真测试了训练过程采用不同损失函数下的BER性能,如图6所示(以(64,32)极化码为例)。

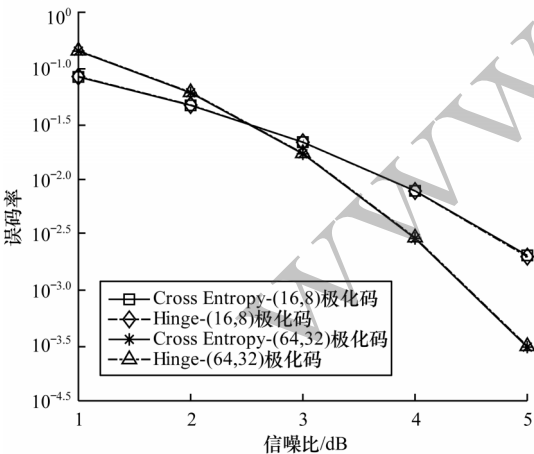


图6 不同损失函数下的BER性能

Fig.6 BER performance under different loss functions

从图6可以看出,使用Hinge损失函数代替交叉熵函数对性能的影响几乎可以忽略不计,该优化方案对于码长极长或需要使用运算性能较弱的设备进

行在线训练的情况,具有一定的优化效果。

4 结束语

本文提出一种基于RNN-OMS-BP的极化码译码算法。将偏移最小和近似算法替代乘法运算,采用改进的循环神经网络架构实现参数共享,有效减少内存开销,并通过离线学习减少BP译码算法的迭代次数。仿真结果表明,该译码算法在误码率性能和运算复杂度之间取得了良好的平衡,对于BP译码算法的硬件实现和实际应用更加友好。下一步将结合具有软判决输出的多用户检测算法,设计基于循环神经网络架构的联合迭代检测译码算法<sup>[21-22]</sup>,以获得更显著的系统性能增益。

参考文献

[1] ARIKAN E. Channel polarization: a method for constructing capacity-achieving codes[C]//Proceedings of 2008 IEEE International Symposium on Information Theory. Washington D. S., USA: IEEE Press, 2008: 143-156.

[2] ARIKAN E. Channel polarization: a method for constructing Capacity-achieving codes for symmetric binary-input memoryless channels [J]. IEEE Transactions on Information Theory, 2009, 55(7): 3051-3073.

[3] ALAMDAR-YAZDI A, KSCHISCHANG F R. A simplified successive-cancellation decoder for polar codes[J]. IEEE Communications Letters, 2013, 17(12): 2360-2363.

[4] 刘亚军,李世宝,刘建航,等. 一种低时延极化码列表连续删除译码算法[J]. 计算机工程, 2018, 44(3): 78-81. LIU Y J, LI S B, LIU J H, et al. A low-latency successive cancellation list decoding algorithm for polar codes[J]. Computer Engineering, 2018, 44(3): 78-81. (in Chinese)

[5] FAYYAZ U U, BARRY J R. Polar codes for partial response channels[C]//Proceedings of IEEE International Conference on Communications. Washington D. S., USA: IEEE Press, 2013: 4337-4341.

[6] YANG J, ZHANG C, ZHOU H, et al. Pipelined belief propagation polar decoders[C]//Proceedings of 2016 IEEE International Symposium on Circuits and Systems. Washington D. S., USA: IEEE Press, 2016: 352-367.

[7] YUAN B, PARHI K K. Architecture optimizations for BP polar decoders [C]//Proceedings of IEEE International Conference on Acoustics. Washington D. S., USA: IEEE Press, 2013: 653-669.

[8] YUAN B, PARHI K K. Early stopping criteria for energy-efficient low-latency belief-propagation polar code decoders [J]. IEEE Transactions on Signal Processing, 2014, 62(24): 6496-6506.

[9] NACHMANI E, MARCIANO E, LUGOSCH L, et al. Deep learning methods for improved decoding of linear codes[J]. IEEE Journal of Selected Topics in Signal Processing, 2018, 12(1): 585-597.

[10] XU W, WU Z, UENG Y L, et al. Improved polar decoder based on deep learning[C]//Proceedings of 2017 IEEE International Workshop on Signal Processing Systems.



- Washington D. S. , USA ;IEEE Press ,2017;235-246.
- [11] DOAN N, HASHEMI S A, MAMBOU E N, et al. Neural belief propagation decoding of CRC-Polar concatenated codes [EB/OL]. [2020-11-01]. <https://arxiv.org/pdf/1811.00124.pdf>.
- [12] WANG X, ZHENG Z, LI J, et al. Belief propagation bit-strengthening decoder for polar codes [J]. IEEE Communications Letters, 2019, 23(11): 1958-1961.
- [13] CHEN C H, TENG C F, WU A Y. Low-complexity LSTM-assisted bit-flipping algorithm for successive cancellation list polar decoder [C]//Proceedings of 2020 IEEE International Conference on Acoustics, Speech and Signal Processing. Washington D. S. , USA ;IEEE Press, 2020: 1451-1467.
- [14] WEN C, XIONG J, GUI L, et al. A BP-NN decoding algorithm for polar codes [C]//Proceedings of the 11th International Conference on Wireless Communications and Signal Processing. Washington D. S. , USA ;IEEE Press, 2019: 325-331.
- [15] TENG C F, HO K S, WU C H, et al. Convolutional neural network-aided bit-flipping for belief propagation decoding of polar codes [C]//Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing. Washington D. S. , USA ;IEEE Press, 2019: 656-667.
- [16] CAO Z, ZHU H, ZHAO Y, et al. Learning to denoise and decode: a novel residual neural network decoder for polar codes [C]//Proceedings of the 92nd IEEE Vehicular Technology Conference. Washington D. S. , USA ;IEEE Press, 2019: 764-778.
- [17] XU W, TAN X, BEERY Y, et al. Deep learning-aided belief propagation decoder for polar codes [J]. IEEE Journal on Emerging and Selected Topics in Circuits and Systems, 2020, 55(9): 145-156.
- [18] LUGOSCH L, GROSS W J. Neural offset min-sum decoding [C]//Proceedings of IEEE International Symposium on Information Theory. Washington D. S. , USA ;IEEE Press, 2017: 1361-1365.
- [19] TENG C F, CHEN H, HO K S, et al. Low-complexity recurrent neural network-based polar decoder with weight quantization Mechanism [C]//Proceedings of 2019 IEEE International Conference on Acoustics, Speech and Signal Processing. Washington D. S. , USA ;IEEE Press, 2019: 667-675.
- [20] WODIANY I, POP A. Low-precision neural network decoding of polar codes [C]//Proceedings of the 20th IEEE International Workshop on Signal Processing Advances in Wireless Communications. Washington D. S. , USA ;IEEE Press, 2019: 256-268.
- [21] PAN Z, LI E, ZHANG L, et al. Design and optimization of joint iterative detection and decoding receiver for uplink polar coded SCMA system [J]. IEEE Access, 2018, 7: 11365-11376.
- [22] SUN F, NIU K, DONG C. Deep learning based joint detection and decoding of non-orthogonal multiple access systems [C]//Proceedings of 2018 IEEE GLOBECOM Workshops. Washington D. S. , USA ;IEEE Press, 2018: 367-378.