

分组马尔可夫叠加传输的神经网络译码

王千帆¹, 毕胜^{2,3}, 陈曾喆⁴, 陈立⁴, 马啸^{2,3}

(1. 中山大学电子与通信工程学院, 广东 广州 510006; 2. 中山大学数据科学与计算机学院, 广东 广州 510006;
3. 中山大学广东省信息安全重点实验室, 广东 广州 510006; 4. 中山大学电子与信息工程学院, 广东 广州 5100063)

摘 要: 研究了分组马尔可夫叠加传输的神经网络 (NN) 译码方案。利用 NN, 实现了不同网络结构、数据表征形式的基本码译码器。在此基础上, 将所实现的基本码译码器嵌入迭代译码机制中, 提出了基于 NN 的分组马尔可夫叠加传输的滑窗译码算法, 并分析了其对应的性能下界。所提出的译码算法提供了一种将 NN 运用到长码译码的解决思路, 即用 NN 替代译码中的部分模块。仿真结果表明, 利用 NN 实现的基本码译码器可以达到最大似然译码性能。基于 NN 的分组马尔可夫叠加传输的滑窗译码算法性能在中高信噪比区域与对应精灵辅助下界贴合, 获得了额外的编码增益。

关键词: 分组马尔可夫叠加传输; 精灵辅助下界; 神经网络; 滑窗译码

中图分类号: TN92

文献标识码: A

doi: 10.11959/j.issn.1000-436x.2020158

Neural network decoding of the block Markov superposition transmission

WANG Qianfan¹, BI Sheng^{2,3}, CHEN Zengzhe⁴, CHEN Li⁴, MA Xiao^{2,3}

1. School of Electronics and Communication Engineering, Sun Yat-sen University, Guangzhou 510006, China

2. School of Data and Computer Science, Sun Yat-sen University, Guangzhou 510006, China

3. Guangdong Key Laboratory of Information Security Technology, Sun Yat-sen University, Guangzhou 510006, China

4. School of Electronics and Information Technology, Sun Yat-sen University, Guangzhou 510006, China

Abstract: A neural network (NN)-based decoding algorithm of block Markov superposition transmission (BMST) was researched. The decoders of the basic code with different network structures and representations of training data were implemented using NN. Integrating the NN-based decoder of the basic code in an iterative manner, a sliding window decoding algorithm was presented. To analyze the bit error rate (BER) performance, the genie-aided (GA) lower bounds were presented. The NN-based decoding algorithm of the BMST provides a possible way to apply NN to decode long codes. That means the part of the conventional decoder could be replaced by the NN. Numerical results show that the NN-based decoder of basic code can achieve the BER performance of the maximum likelihood (ML) decoder. For the BMST codes, BER performance of the NN-based decoding algorithm matches well with the GA lower bound and exhibits an extra coding gain.

Key words: BMST, GA lower bound, NN, sliding window decoding

收稿日期: 2020-03-24; 修回日期: 2020-06-30

通信作者: 马啸, maxiao@mail.sysu.edu.cn

基金项目: 国家自然科学基金资助项目 (No.61971454, No.61671486); 广东省基础与应用基础研究基金资助项目 (No.2020A1515010687); 广东省自然科学基金资助项目 (No.2016A030308008)

Foundation Items: The National Natural Science Foundation of China (No.61971454, No.61671486), The Basic and Applied Basic Research Foundation of Guangdong (No.2020A1515010687), The Natural Science Foundation of Guangdong Province (No.2016A030308008)

1 引言

通信领域是一个相对复杂且成熟的领域，传统的性能优化方法的效果正在逐步减弱，这种现象在物理层尤为明显^[1]。为满足用户对新一代通信的需求，通信技术和人工智能技术相结合成为一种新的研究方向。机器学习被认为是人工智能领域中最能体现“智能”的分支之一，其在处理难以用数学公式描述的问题上具有一定优势，并成功地应用在图像处理、文本处理等领域。

受机器学习在图像去噪和超清分辨等方面研究的启发，基于机器学习来实现或优化旨在去除噪声恢复数据的信道译码引起了越来越多的关注。神经网络(NN, neural network)是机器学习中的重要分支之一，将NN运用到信道译码研究的优势在于NN是不依赖于模型的，即不需要对信道噪声的统计信息做任何假设，便能在训练过程中学习信道的映射关系或提取信道统计信息^[2]。此外，NN在并行性上也具有一定的优势，利用NN可以最大限度地并行化译码过程，从而提高译码器的吞吐量^[3]。目前，基于NN的译码研究已经取得了初步的成果^[4-5]。一般来说，使用NN进行译码可以看作一个分类问题。然而类别的数量随着信息位长度的增长呈指数增长，存在维度诅咒现象^[6]。当信息位长度较大时，使用NN直接译码一般难以获得较好的译码性能，因此不具备实用价值，这也是目前基于NN的译码主要针对中短码的原因。然而即使在中短码的情况下，目前的方案一般也难以达到传统方案最优的译码性能。文献[4]使用NN对码长16的极化码进行译码，性能接近最大似然(ML, maximum likelihood)译码，但仍有一定差距。文献[5]基于NN改进了中短码长的BCH(Bose-Ray-Hocquenghem)码在置信传播译码下的性能，但与传统译码性能相比仍有部分差距。文献[7]提出了一种针对低密度奇偶校验(LDPC, low density parity check)码、极化码的NN联合译码方案，该译码方案可以达到与置信传播译码相近的性能，但所使用的码长仍然是较短的。文献[8]使用NN代替极化码连续消除译码器中的部分子块，但所使用的码长仍然是较短的。通过以上的研究可以发现，基于NN的译码研究主要面临两方面的问题：一方面是短码情况下如何达到传统方案的最优译码性能，另一方面是如何将NN运用到长码的译码中。

分组马尔可夫叠加传输(BMST, block Markov superposition transmission)是近年来提出的一种新的编码传输方案^[9-10]，其编码过程如下。信息序列先由基本码进行编码，然后将编码后的序列进行交织，最后将交织后的序列进行叠加传输，从而达到既能重复传输多次又能保证传输效率的目的。研究者在多个场景下对BMST进行了研究。文献[11]提出了适用于光纤通信场景的BMST-BCH码设计，保证了高码率传输下的高可靠性。文献[12]提出了一种基于BMST通过部分叠加来构造空间耦合LDPC码的一般方法。此外，BMST在加性高斯白噪声(AWGN, additive white Gaussian noise)信道^[9]、湍流信道^[13]和非高斯脉冲信道^[14]上均有逼近香农限的性能，同时研究者也给出了相应的用于预测误码性能的简单下界。BMST具有诸多优点，其由短码出发，构造性能优异的长码为NN运用到长码提供了一种可能，即可以设计针对短码的NN译码器，并将此译码器嵌入长码译码的迭代机制中，从而替代其中部分模块。

文献[15]初步提出了基于NN的BMST译码方案，以汉明(Hamming)码为基本码，利用NN实现了仅比ML译码器差0.5 dB的基本码译码器，并将其嵌入了BMST的迭代译码机制中。本文在文献[15]的基础上改进了基于NN的BMST译码方法，分析了不同网络结构和数据表征形式对译码性能的影响，通过引入独热向量，使基于NN的基本码译码器可以达到ML译码的性能。针对使用独热向量可以达到ML译码性能的原因，本文也给出了相应解释。由于使用独热向量时输出层神经元个数随信息位长度的增长呈指数增长，而输出层神经元个数过多是不利于学习的。针对该问题，本文提出了双热向量，从而极大地减少输出层神经元个数，降低学习复杂度，同时可以达到接近ML译码的性能。最后，本文将所实现的NN基本码译码器嵌入迭代译码机制中，设计了基于NN的分组马尔可夫叠加传输的滑窗译码算法，并分析了其对应的性能下界。仿真结果显示：1)所提NN基本码译码器在使用独热向量表征下可以达到ML译码的性能，在使用降低学习复杂度的双热向量表征下可以达到接近ML译码的性能；2)所提基于NN的分组马尔可夫叠加传输的滑窗译码算法的性能在中高信噪比区域贴合其对应的精灵辅助(GA, genie-aided)下界，并与BMST的传统性能下界贴近，获得了额外的编码增益。

2 预备知识

2.1 神经网络

受到人脑结构的启发，研究者提出了 NN 的概念^[16]。类似于大脑，NN 具有学习、存储和使用经验知识的特性。整个 NN 由输入层、隐含层和输出层构成，如图 1 所示。

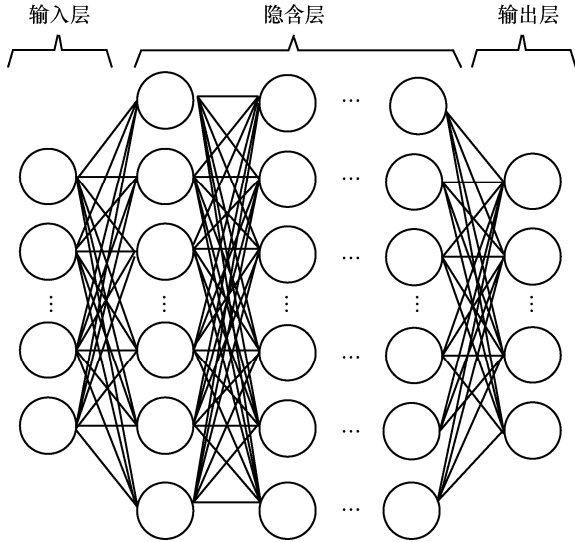


图 1 神经网络结构

图 2 展示了神经元的运算过程。假设前一层有 n 个神经元与本层第 k 个神经元相连，则本层第 k 个神经元的输出 $y^{(k)}$ 如式(1)所示。

$$y^{(k)} = \varphi \left(\sum_{i=0}^{n-1} x^{(i)} w^{(i)} + b^{(k)} \right) \quad (1)$$

其中， $x^{(i)}$ 表示此神经元的第 i 个输入， $w^{(i)}$ 表示第 i 个输入对应的权重， $b^{(k)}$ 表示此神经元对应的偏置， φ 表示激活函数。这里权重和偏置用来描述线性关系，激活函数用来描述非线性关系。

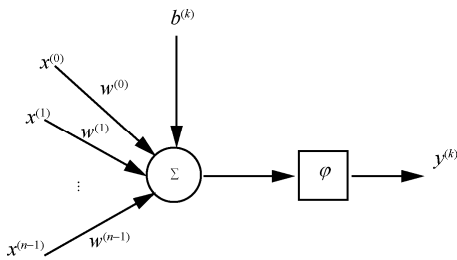


图 2 神经元的运算过程

2.2 分组马尔可夫叠加传输

记忆长度为 m 的 BMST 系统的编码过程如图 3

所示。该系统是由一个基本码编码器、 m 个交织器和 m 个寄存器组成。这里考虑基本码是一个码长为 n 、信息位长度为 k 的二元线性码 $C[n, k]$ 。

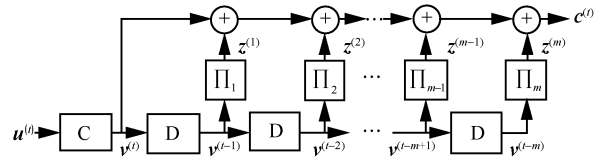


图 3 BMST 系统编码过程

令 \mathbb{F}_2 表示二元域， $\mathbf{u}^{(t)} \in \mathbb{F}_2^k$ ($0 \leq t \leq L-1$) 为待传输的 L 组信源序列，则 BMST 编码的具体过程^[10]如算法 1 所示。

算法 1 BMST 系统的编码算法

初始化 if $t < 0$ ，令 $\mathbf{v}^{(t)} = \mathbf{0} \in \mathbb{F}_2^n$

循环 for $t = 0, 1, \dots, L-1$

1) 编码。将信源序列 $\mathbf{u}^{(t)}$ 按 $C[n, k]$ 的编码算法编码为基本码码字 $\mathbf{v}^{(t)} \in \mathbb{F}_2^n$ 。

2) 交织。对于 $1 \leq i \leq m$ ，用第 i 个交织器 Π_i 将码字 $\mathbf{v}^{(t-i)}$ 交织成 $\mathbf{z}^{(i)}$ 。

3) 叠加。计算 $\mathbf{c}^{(t)} = \sum_{1 \leq i \leq m} \mathbf{z}^{(i)} + \mathbf{v}^{(t)}$ 。

结束 对于 $t = L, L+1, \dots, L+m-1$ ，设置 $\mathbf{u}^{(t)} = \mathbf{0} \in \mathbb{F}_2^k$ ，按照上述步骤“循环”计算 $\mathbf{c}^{(t)}$ 。

BMST 方案的码字 $\mathbf{c}^{(t)}$ 经过调制后发送到噪声信道，接收端获得信号的含噪信号 $\mathbf{y}^{(t)}$ 。本文的调制方式为二进制相移键控 (BPSK, binary phase shift keying)，噪声信道为 AWGN 信道。

BMST 的译码过程采用基于正规图 (normal graph)^[17]的滑窗译码算法，具体译码过程见文献[10]。

3 神经网络译码

3.1 基于神经网络的基本码译码

本节将介绍利用 NN 实现的不同方案的基本码译码器，其中，基本码采用 [7,4] 汉明码，隐含层的激活函数使用 ReLU 函数^[18]，输出层的激活函数使用 sigmoid 函数，具体如式(2)和式(3)所示。

$$\varphi_{\text{ReLU}}(x) = \max(0, x) \quad (2)$$

$$\varphi_{\text{sigmoid}}(x) = \frac{1}{1 + e^{-x}} \quad (3)$$

NN 输入数据为有噪信号，因此，针对 NN 结构和监督数据的表征形式，本文给出了 4 种方案。

方案 1 使用一个输入输出维度均为 7 的 NN，如图 4 所示。NN 训练所用的输入数据为 7 维噪声向量，监督数据为软入软出 (SISO, soft-in soft-out) ML 译码后输出的 7 维概率向量。

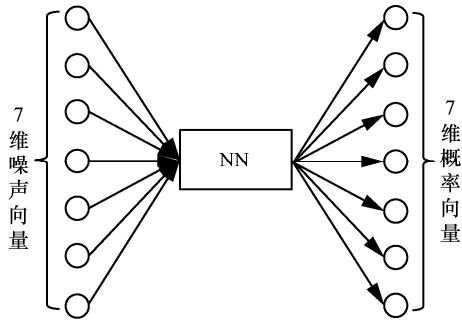


图 4 方案 1 架构

方案 2 架构如图 5 所示，其基本思想是通过缩小 NN 的输出维度来减少每个 NN 的类别数目，以此改进性能。该方案使用 2 个输入维度均为 7 的 NN (NN-1 和 NN-2)，其输出维度分别为 4 和 3。训练所用的输入数据均为 7 维噪声向量，NN-1 所使用的监督数据为与 ML 译码器输出的信息位对应的 4 维概率向量，NN-2 所使用的监督数据为 ML 译码器输出的校验位对应的 3 维概率向量。

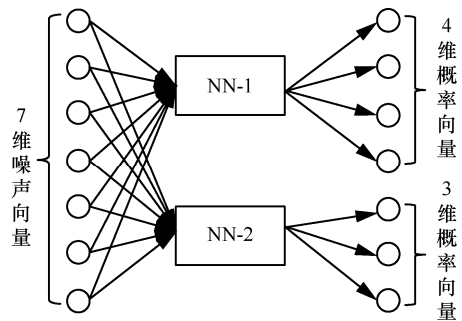


图 5 方案 2 架构

方案 3 如图 6 所示。该方案中的 NN 输入维度为 7，输出维度为 16。训练所用的输入数据为 7 维噪声向量，监督数据为 16 维独热向量表示的信息位。独热向量采用 N 种情况使用一个长度为 N 的向量进行表示，当出现第 i 种情况时 ($1 \leq i \leq N$)，该向量的第 i 个元素值为 1，其他元素值为 0。本文采用的基本码信息位长度为 4，因此 $N=2^4=16$ 。采用独热向量需要更多的神经元来表征不同类别，虽然这增加了输出层神经元数目，但是更易于区分每个类别。

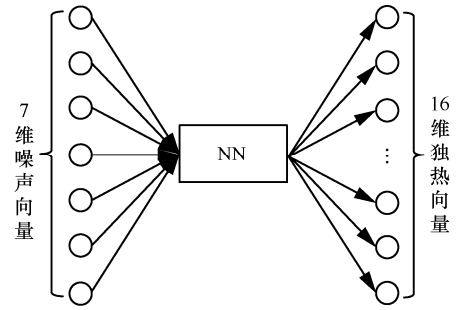


图 6 方案 3 架构

方案 4 架构如图 7 所示。该方案中的 NN 输入维度为 7，输出维度为 8。训练所使用的输入数据为 7 维噪声向量，监督数据为使用 8 维双热向量表示的信息位。监督数据中前 4 维向量是信息位前两位对应的独热向量，后 4 维向量是信息位后两位对应的独热向量。当采用方案 3 中的独热向量来表示信息位时，输出层神经元个数随信息位长度的增长呈指数增长，但输出层神经元个数过多是不利于 NN 训练的。方案 4 可以使输出层神经元个数大幅减少。以 10 bit 信息位为例，采用方案 3，输出层神经元个数为 $2^{10}=1024$ 。若采用方案 4，则输出层神经元个数减少至 $2^5+2^5=64$ 个。

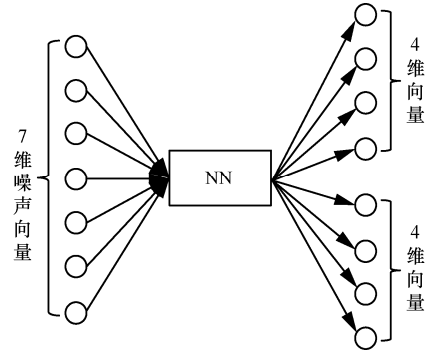


图 7 方案 4 架构

3.2 基于神经网络的 BMST 译码

本节基于 3.1 节的 NN 基本码译码器，针对 BMST 系统提出了基于 NN 的滑窗译码算法，一般使用正规图^[17]来描述该译码算法的信息传递过程。图 8 为 $L=4, m=2$ 的 BMST 基于 NN 的滑窗译码器示意。一般来说，译码窗口 $d \geq m$ ，图 8 中的矩形表示节点，与节点相连的边需要满足一定的信息传递约束。 $\mathbf{Y}^{(t)}, \mathbf{C}^{(t)}, \mathbf{V}^{(t)}$ 和 $\mathbf{U}^{(t)}$ 分别为 $\mathbf{y}^{(t)}, \mathbf{c}^{(t)}, \mathbf{v}^{(t)}$ 和 $\mathbf{u}^{(t)}$ 对应的随机向量，这里 $\mathbf{y}^{(t)}, \mathbf{c}^{(t)}, \mathbf{v}^{(t)}$ 和 $\mathbf{u}^{(t)}$ 分别表示接收序列、发送码字、基本码码字和信源序列； $H(\cdot)$ 表示随机变量的熵。在一个译码层中，包

含 4 类节点，具体如下。

$\boxed{+}$ 节点。相连的变量满足加号约束，即与之相连各边对应变量在二元域上相加和为 0。节点 $\boxed{+}$ 的消息传递算法与 LDPC 码中校验节点上的算法是相似的。

$\boxed{\Pi}$ 节点。对输入信息完成交织或解交织过程。

$\boxed{=}$ 节点。相连的变量满足等号约束，即与之相连各边对应变量取值相等。节点 $\boxed{=}$ 上的消息传递算法与 LDPC 码中变量节点上的算法是相同的。

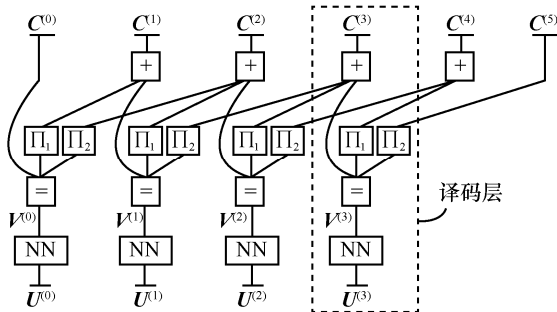


图 8 基于 NN 的 BMST 滑窗译码示意

\boxed{NN} 节点。NN 基本码译码器，完成基本码 SISO 译码过程。

译码窗口为 d 的滑窗译码算法只在包含 $d+1$ 层的子正规图上进行迭代译码。接收向量 $\mathbf{y}^{(t+d)}$ 在时刻 $t+d$ 输入译码器中进行迭代。迭代完成后，基于 NN 的滑窗译码算法输出信源 $\mathbf{u}^{(t)}$ 的估计 $\hat{\mathbf{u}}^{(t)}$ 。详细过程如算法 2 所示。

算法 2 基于 NN 的 BMST 滑窗译码算法

输入 $\mathbf{y}^{(t)}$

输出 $\hat{\mathbf{u}}^{(t)}$

全局初始化 译码器根据接收向量 $\mathbf{y}^{(t)}(0 \leq t \leq d-1)$ 计算 $\mathbf{c}^{(t)}$ 对应的后验概率，正规图上当前层内部其他边上的消息及连接到其他层的所有边上的消息都按照均匀分布进行初始化。设置最大迭代次数 $J_{\max} > 0$ ，阈值 $\sigma > 0$ ，熵率 $H_0(\mathbf{Y}^{(t)}) = 0$ 。

滑窗译码 for $t = 0, 1, \dots, L-1$

1) 局部初始化。当 $t+d \leq L+m-1$ 时，译码器根据接收向量 $\mathbf{y}^{(t+d)}$ 计算 $\mathbf{c}^{(t+d)}$ 对应的后验概率。将第 $t+d$ 层内部其他边上的消息及连接到其他层的所有边上的消息按照均匀分布进行初始化。

2) 迭代。for $J = 1, \dots, J_{\max}$

① 前向递归。对于 $i = 0, 1, \dots, d$ ，正规图上第 $t+i$ 层消息传递的顺序为



② 后向递归。对于 $i = d, d-1, \dots, 0$ ，正规图上第 $t+i$ 层消息传递的顺序为



③ 判决。对 $\mathbf{v}^{(t)}$ 进行硬判决，计算 $\mathbf{Y}^{(t)}$ 的熵率 $H_j(\mathbf{Y}^{(t)})$ ，若满足 $|H_j(\mathbf{Y}^{(t)}) - H_{j-1}(\mathbf{Y}^{(t)})| \leq \sigma$ ，则退出迭代，并输出 $\hat{\mathbf{u}}^{(t)}$ 。

3) 干扰消除。通过更新 $\mathbf{c}^{(t+1)}, \dots, \mathbf{c}^{(t+m)}$ 对应的后验概率来移除所有层上 $\mathbf{v}^{(t)}$ 的影响。

3.3 GA 下界

本节分析得到基于 NN 的 BMST 系统的 GA 下界。文献[9]中给出了采用 BPSK 调制的 BMST 在 AWGN 信道下的 GA 下界，如式(4)所示。

$$F_{\text{BMST}}(\lambda) \geq F_{\text{basic}}(\lambda + 10\log(m+1)) \quad (4)$$

其中， λ 表示 $\frac{E_b}{N_0}$ ， E_b 表示每比特的信号能量， N_0 表示噪声的功率谱密度； $F_{\text{BMST}}(\cdot)$ 和 $F_{\text{basic}}(\cdot)$ 分别表示记忆长度为 m 的 BMST 和基本码的性能函数。在中高信噪比区域，BMST 的最大额外增益为 $10\log(m+1)$ dB。基于 NN 的 BMST 系统同样存在着类似的 GA 下界，如式(5)所示。

$$F_{\text{BMST-NN}}(\lambda) \geq F_{\text{basic-NN}}(\lambda + 10\log(m+1)) \quad (5)$$

其中， $F_{\text{BMST-NN}}(\cdot)$ 和 $F_{\text{basic-NN}}(\cdot)$ 分别为记忆长度为 m 的基于 NN 的 BMST 和基本码的性能函数。对于记忆长度为 m 的 BMST 系统而言，只需将基于 NN 的基本码译码器的误比特率曲线向左平移 $10\log(m+1)$ dB，即可得到对应的 GA 下界。

4 数值结果

本节通过几个例子来分析第 3 节中 4 种不同方案的 NN 基本码译码器和不同基本码译码器嵌入 BMST 系统的性能。BMST 系统中基本码使用汉明码的 2500 重笛卡尔积 $[7,4]^{2500}$ ，其信息位长度和码字长度分别为 $k = 10000$ 和 $n = 17500$ 。在 BMST 系统中，采用的交织器是长度为 17500 的均匀随机交织器^[10]。对于基本码译码器，本文只需要针对原始

的 [7,4] 汉明码训练即可，所产生训练数据的 $\frac{E_b}{N_0}$ 为 6 dB，训练数据量为 10^7 帧，每批样本大小为 256 帧。使用 TensorFlow 开源框架的 ADAM-Optimizer，学习率为 0.1，步长为 0.01。文中用 $(N_0, N_1, \dots, N_{\ell-1})$ 来表示包含 ℓ 层隐含层的 NN，其中第 i 层隐含层的神经元个数为 N_i 。采用算法 2 作为 BMST 的译码算法，设置最大迭代次数 $J_{\max} = 50$ ， $L = 98$ ，译码的熵终止阈值 $\sigma = 10^{-5}$ 。

例 1 图 9 展示了汉明码在方案 1 中不同 NN 结构对译码性能的影响，参数设置与文献[15]相同，本文得到了相似的仿真结果。由仿真结果可知，在一定范围内，增加隐含层神经元的数目可以改善性能。两层隐含层神经元个数分别为 56、28 时（图 9 中表示为 (56,28)），在 BER= 10^{-4} 处，该方案比 ML 译码方案差 1.2 dB。结果还反映再增大神经元数目已无意义，即网络结构(56, 28)已饱和。

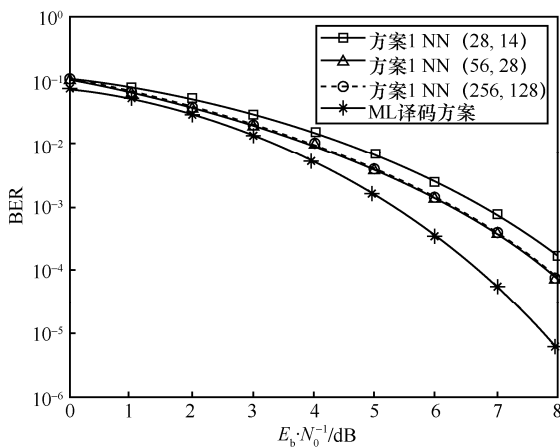


图 9 方案 1 下不同 NN 结构的汉明码译码性能曲线

例 2 图 10 展示了汉明码在方案 2 不同 NN 结构对译码性能的影响。方案 2 通过使用 2 个独立的 NN，缩小 NN 的输出维度并改善性能。由仿真结果可知，相比于方案 1，方案 2 的性能得到了提升。结果还反映在一定范围内，增加隐含层神经元的数目可以改善性能。在 BER= 10^{-4} 处，NN 结构为 (28,14) 的方案 2 性能比 ML 译码方案差 0.5 dB。

例 3 图 11 展示了汉明码在采用方案 3 情况下不同 NN 结构对译码性能的影响。方案 3 使用一个独热向量来表示信息位。由仿真结果可知，所有网络结构的性能曲线均贴合 ML 译码方案的性能曲线。使用没有隐含层的浅层 NN（图 11 中表示为 (none)）同样可以达到 ML 译码方案的性能。在浅

层 NN 结构下，乘法和加法分别需要执行 $7 \times 16 = 112$ 次，比较操作需要执行 15 次，激活函数执行 16 次，这里激活函数可以通过查表获得输出值。而 ML 译码方案分为 2 个步骤，第一步计算 16 个码字的概率，需要进行 96 次乘法；第二步根据 16 个码字的概率，得到译码选择，需要 15 次比较操作。表 1 展示了 2 种译码器的复杂度对比。从表 1 中可以看出，方案 3 NN (none) 译码器相比于 ML 译码方案只是增加了加法运算和较少的乘法运算。

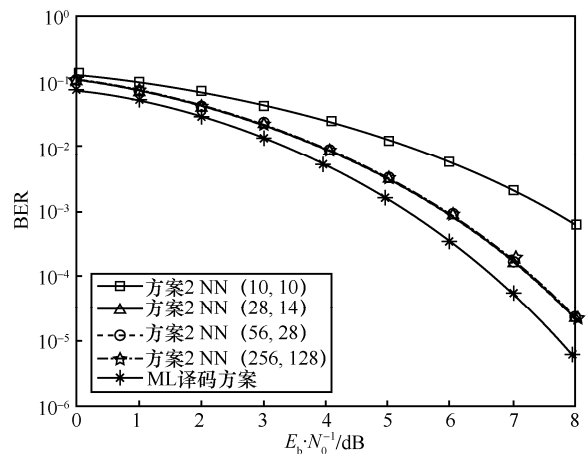


图 10 方案 2 下不同 NN 结构的汉明码译码性能曲线

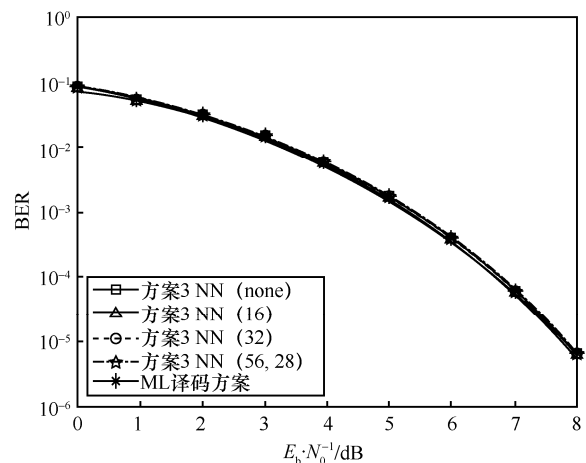


图 11 方案 3 下不同 NN 结构的汉明码译码性能曲线

表 1 2 种译码方案的计算复杂度对比

| 译码方案 | 加法/次 | 乘法/次 | 比较操作/次 |
|----------------|------|------|--------|
| 方案 3 NN (none) | 112 | 112 | 15 |
| ML 译码方案 | 0 | 96 | 15 |

由图 11 还可以看出，方案 3 在没有隐含层的情况下可以达到 ML 译码方案的性能，这是由于无隐含层的方案 3 等效于实现一个 ML 译码方案。以

如下给定权重、偏置的“NN”来解释其等效性。假设接收向量对应的对数似然比作为 NN 的输入，没有隐含层，输出层神经元个数为 16 个，且每个输出层神经元的偏置均为 0，无激活函数。若码表中第 i 个码字的第 j 位为 0，则输入层第 j 个神经元与输出层第 i 个神经元的权重为 +1，反之则为 -1。经过前向传播后，若该 NN 第 b 个输出层神经元的值最大，则码表中第 b 个码字的概率最大。上述给定权重、偏置的“NN”，即 ML 译码方案的网络实现形式。基于此，对于码表已知的任意编码方案的 ML 译码方案，均可以由 NN 来实现。但是使用独热向量情况下，输出层神经元个数会随信息位长度呈指数增长，输出层神经元个数过多是不利于学习的。针对这个问题，本文提出了双热向量，即方案 4，如例 4 所示。

例 4 图 12 展示了汉明码在方案 4 不同 NN 结构对译码性能的影响。方案 4 使用双热向量来表示信息位，从而减小输出层神经元个数，降低学习复杂度。由仿真结果可知，在 $BER=10^{-5}$ 处，网络结构为 (56, 28) 的方案 4 性能比 ML 译码方案差 0.2 dB，达到了近似最优的性能，同时降低了学习复杂度，更具有实用性。

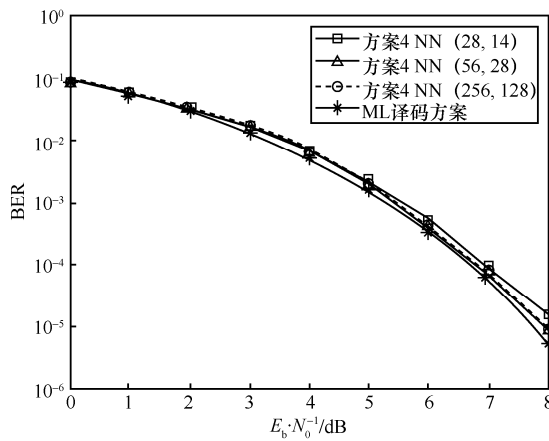


图 12 方案 4 下不同 NN 结构的汉明码译码性能曲线

例 5 图 13 展示了不同方案的 NN 基本码译码器嵌入 BMST 系统中的性能。所有训练方案的隐含层 NN 结构均为 (56, 28)。BMST 系统记忆长度 $m=1$ ，译码窗口 $d=7$ 。由仿真结果可知，使用不同方案基本码译码器的 BMST 系统均获得了额外编码增益。其中，使用方案 3 的 BMST 系统的性能在中高信噪比区域贴合其对应的 GA 下界，并与 BMST 的传统性能下界接近。同时，使用方案 4 的 BMST 系统也

获得了较好的性能，并在中高信噪比区域与方案 3 (56, 28) 的性能下界贴合。

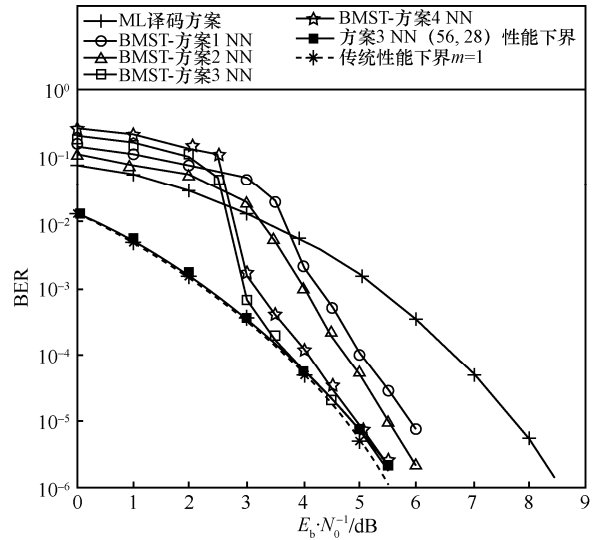


图 13 不同方案的基本码译码器在 BMST 系统中的性能比较

例 6 图 14 展示了记忆长度 m 对基于 NN 的 BMST 系统的性能影响。汉明码采用方案 3 NN (56, 28) 的译码器，BMST 系统中同样采用相同结构的 NN 译码器。BMST 系统记忆长度和译码窗口设置分别为 $[m, d]=[1, 3]$ 和 $[m, d]=[2, 6]$ 。由仿真结果可知，基于 NN 的 BMST 系统增大记忆长度 m 可以提高性能，同时，不同记忆长度 m 的仿真性能在中高信噪比区域都与其对应的下界贴合。

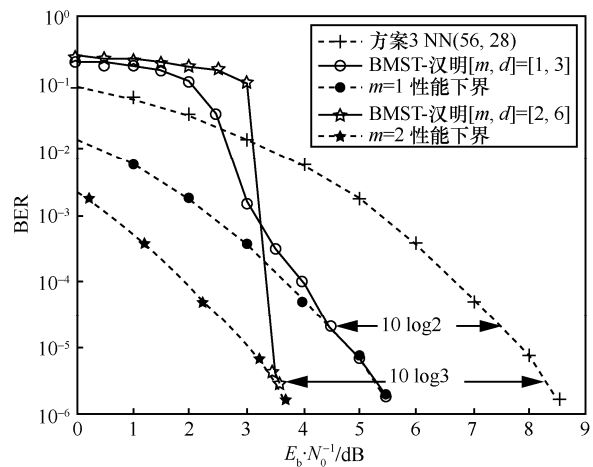


图 14 不同记忆长度的 BMST 系统的性能比较

5 结束语

本文研究了基于 NN 的分组马尔可夫叠加传输的译码方案，给出了不同网络结构和数据表征形式

对译码性能的影响。针对数据表征形式, 提出了缩减 NN 输出维度、使用独热向量表示信息位和使用双热向量表示信息位等方案。其中, 使用独热向量表示信息位的 NN 基本码译码器可以达到 ML 译码的性能。针对独热向量下输出层神经元个数过多、不利于学习的问题, 提出了双热向量, 可以极大减少输出层神经元个数, 同时可以达到接近 ML 译码的性能。在此基础上, 将不同方案的基本码译码器嵌入 BMST 迭代译码机制中, 替换了长码译码中的部分模块, 均获得了额外编码增益。其中采用方案 3 的 BMST 系统的性能贴合其对应的 GA 下界, 并与 BMST 的传统性能下界接近。

参考文献:

- [1] O'SHEA T, HOYDIS J. An introduction to deep learning for the physical layer[J]. IEEE Transactions on Cognitive Communications and Networking, 2017, 3(4): 563-575.
- [2] CAID W R, MEANS R W. Neural network error correcting decoders for block and convolutional codes[C]//1990 IEEE Global Communications Conference. Piscataway: IEEE Press, 1990: 1028-1031.
- [3] WANG X A, WICKER S B. An artificial neural net Viterbi decoder[J]. IEEE Transactions on Communications, 1996, 44(2): 165-171.
- [4] GRUBER T, CAMMERER S, HOYDIS J, et al. On deep learning-based channel decoding[C]//2017 IEEE Annual Conference on Information Sciences and Systems. Piscataway: IEEE Press, 2017: 1-6.
- [5] 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): 119-131.
- [6] BELLMAN R. Dynamic programming[J]. Science, 1966, 153(3731): 34-37.
- [7] WANG Y, ZHANG Z, ZHANG S, et al. A unified deep learning based polar-LDPC decoder for 5G communication systems[C]//2018 10th International Conference on Wireless Communications and Signal Processing. Piscataway: IEEE Press, 2018: 1-6.
- [8] CAMMERER S, GRUBER T, HOYDIS J, et al. Scaling deep learning-based decoding of polar codes via partitioning[C]//2017 IEEE Global Communications Conference. Piscataway: IEEE Press, 2017: 1-6.
- [9] LIANG C, MA X, ZHUANG Q, et al. Spatial coupling of generator matrices: a general approach to design good codes at a target BER[J]. IEEE Transactions on Communications, 2014, 62(12): 4211-4219.
- [10] MA X, LIANG C, HUANG K, et al. Block Markov superposition transmission: construction of big convolutional codes from short codes[J]. IEEE Transactions on Information Theory, 2015, 61(6): 3150-3163.
- [11] CAI S, LIN N, MA X. Block Markov superposition transmission of BCH codes with iterative erasures-and-errors decoders[J]. IEEE Transactions on Communications, 2019, 67(1): 17-27.
- [12] WANG Q, CAI S, LIN W, et al. Spatially coupled LDPC codes via partial superposition[C]//2019 IEEE International Symposium on Information Theory. Piscataway: IEEE Press, 2019: 2614-2618.
- [13] 朱锦顺, 马啸. 户外无线光通信中湍流信道下的分组马尔可夫叠加传输研究[J]. 通信学报, 2017, 38(7): 131-140.
ZHU J S, MA X. Block Markov superposition transmission over turbulence channels in outdoor optical wireless communications[J]. Journal on Communications, 2017, 38(7):131-140.
- [14] 马啸, 吉眉颖, 陈声晓. 分组马尔可夫叠加传输在非高斯脉冲信道上的性能研究[J]. 通信学报, 2019, 40 (3): 109-115.
MA X, JI M Y, CHEN S X. Performance of block Markov superposition transmission over non-Gaussian impulsive channels[J]. Journal on Communications, 2019, 40(3):109-115.
- [15] BI S, WANG Q, CHEN Z, et al. Deep learning-based decoding of block Markov superposition transmission[C]// 2019 11th International Conference on Wireless Communications and Signal Processing. Piscataway: IEEE Press, 2019: 1-5.
- [16] MCCULLOCH W S, PITTS W. A logical calculus of the ideas immanent in nervous activity[J]. The bulletin of Mathematical Biophysics, 1943, 5(4): 115-133.
- [17] FORNEY G D. Codes on graphs: normal realizations[J]. IEEE Transactions on Information Theory, 2001, 47(2): 520-548.
- [18] NAIR V, HINTON G E. Rectified linear units improve restricted Boltzmann machines[C]//2010 International Conference on Machine Learning. Piscataway: IEEE Press, 2010: 807-814.

[作者简介]



王千帆 (1992-), 男, 河南焦作人, 中山大学博士生, 主要研究方向为信道编码及其在无线通信中的应用。



毕胜 (1996-), 男, 湖南常德人, 中山大学硕士生, 主要研究方向为信道编码与机器学习。

陈曾喆 (1996-), 男, 广东汕头人, 中山大学硕士生, 主要研究方向为信道编码中的低时延译码技术。

陈立 (1981-), 男, 广东揭西人, 博士, 中山大学教授、博士生导师, 主要研究方向为信息与编码理论、数据通信等。

马啸 (1968-), 男, 河南焦作人, 博士, 中山大学教授、博士生导师, 主要研究方向为信息与编码理论、编码调制技术、无线通信、光通信等。