

# 基于 BP 神经网络的 SCL 译码研究<sup>①</sup>

卢丽金, 李世宝

(中国石油大学(华东)计算机与通信工程学院, 青岛 266580)

通讯作者: 李世宝, E-mail: Lishibao@upc.edu.cn

**摘要:** 现存极化码译码算法仍然遭受非常高的复杂度. 针对此问题, 提出一种基于 BP 神经网络的 SCL 译码算法, 该算法通过离线收集数据来搭建并训练一个合适的 BP 神经网络; 借助已完成训练的 BP 神经网络, 通过在线操作来寻找列表大小  $L$  的最优初始值; 在此基础上, 通过设计一种改进的 SCL 译码算法来降低复杂度. 实验结果表明, 与现存算法相比, 新算法在低信噪比下能够显著降低平均译码复杂度.

**关键词:** 极化码; BP 神经网络; SCL 译码; 列表大小; 复杂度

引用格式: 卢丽金, 李世宝. 基于 BP 神经网络的 SCL 译码研究. 计算机系统应用, 2018, 27(12): 246-250. <http://www.c-s-a.org.cn/1003-3254/6677.html>

## Research on SCL Decoding Based on BP Neural Network

LU Li-Jin, LI Shi-Bao

(College of Computer & Communication Engineering, China University of Petroleum, Qingdao 266580, China)

**Abstract:** Existing decoding algorithms for polar codes still suffer from very high complexity. To solve this problem, an SCL decoding algorithm based on BP neural network is proposed. In offline, the algorithm builds and trains an appropriate BP neural network by collecting data. With the trained BP neural network, the optimal initial value of list size  $L$  is found through on-line operation. On this basis, the complexity is reduced by designing an improved SCL decoding algorithm. Experimental results show that compared with existing algorithms, the proposed algorithm can significantly reduce the average decoding complexity at low SNR.

**Key words:** polar codes; BP neural network; SCL decoding; list size; complexity

Arikan 首次提出极化码 (Polar Code), 且极化码是人类已知的第一种能够被严格证明达到信道容量的信道编码方法<sup>[1]</sup>. 在串行抵消 (SC) 译码算法下, 极化码的复杂度低、译码结构简单. 然而, SC 译码算法在码长为有限长的配置下, 纠错性能不理想.

为了提高极化码的性能, 引入置信度传播<sup>[2,3]</sup>和线性规划<sup>[4]</sup>算法, 但, 这两种算法只适用于二进制输入删除信道 (BEC). 随后, I. Tal 和 A. Vardy 等提出串行抵消列表 (SCL) 算法<sup>[5,6]</sup>, 其性能非常接近最大似然 (ML) 的纠错性能. 为了进一步提高极化码的纠错能力, 几种

解码方法相继被提出<sup>[7-9]</sup>. 其中, 文献[8]通过级联循环冗余校验 (CRC) 码, SCL 译码能够获得比 Turbo 码和 LDPC 码更好的性能, 并记为 CA-SCL. 然而, 其复杂度与列表大小  $L$  呈正比关系. 为了降低复杂度, 提出一种自适应的串行抵消列表 (AD-SCL) 译码<sup>[10]</sup>, 其复杂度不会随着列表大小  $L$  的增大而增加. 但是, 在低信噪比下, AD-SCL 会呈现明显的高复杂度现象. AD-SCL 总是把  $L$  的初始值设置为 1. 当  $L=1$  译码失败时, AD-SCL 会将  $L$  更新为  $2L$ , 并返回根节点, 重新进行译码, 直至  $L=L_{\max}$ . 在信噪比较低的情况下, 经过基于

① 基金项目: 中央高校基本科研业务费专项资金 (18CX06042A)

Foundation item: Fundamental Research Funds for the Central Universities (18CX06042A)

收稿时间: 2018-05-09; 修改时间: 2018-06-04; 采用时间: 2018-06-14; csa 在线出版时间: 2018-12-03

$L=1$ 的译码后,往往还需要更新 $L$ 值.这样无疑是增加了复杂度.值得考虑的是,由小到大的更新 $L$ 能够带来明显的低平均译码复杂度,这一优势值得借鉴.如果 $L$ 的初始值不固定为1,那么,就能够解决现存的复杂度问题.因而,寻找 $L$ 的最优初始值是本文的研究关键.若 $L$ 的最优初始值的大小有 $l$ 种可能,则相当于 $L$ 的最优初始值有 $l$ 种类别.在某一实验配置下, $L$ 的最优初始值到底属于哪一种类别,本文需要处理此问题.若在线阶段采用传统的追踪记录与概率统计方式来分类,需要进行大量的数据存储和计算,将会大幅度增加复杂度,且此处理方式不适用于离线操作. BP神经网络可以用作分类器<sup>[11,12]</sup>并实现快速、精确分类.若BP神经网络的训练和分类都是在线操作,也会大幅度增加复杂度,但可以离线训练BP神经网络,在线时将数据输入到已完成训练的BP神经网络模型中,输出一个类别,且不会增加复杂度.因此,本文确定采用BP神经网络来寻找 $L$ 的最优初始值.

本文的主要工作是介绍极化编码、极化译码与BP神经网络的相关原理,然后离线构建并训练BP神经网络以及在线寻找 $L$ 的最优初始值,最后提出一种基于BP神经网络的SCL译码算法.

## 1 相关原理

### 1.1 极化编码与极化译码

#### 1.1.1 极化编码

信道极化将一个二进制输入无记忆信道的一组独立的时隙看作一组相互独立的信道,通过信道分割、信道合并操作引入相关性.当参与信道极化的信道(时隙)数足够多时,所得到的极化信道的信道容量会出现极化现象,即一部分信道的容量将会趋于1、其余趋于0.在信道极化的基础上,只需要在一部分容量趋于1的信道上传输信息比特,而在剩下的容量趋于1的信道以及容量趋于0的信道上传输对收发端都已知的固定比特.

用符号 $W: X \rightarrow Y$ 来表示一个B-DMC信道,其中 $X$ 和 $Y$ 分别表示该信道的输入、输出符号集合, $X = \{0, 1\}$ .信道转移函数被定义为 $W(y|x)$ ,其中 $x \in X, y \in Y$ .那么,对称信道容量表示为

$$I(W) = \sum \sum \frac{1}{2} \beta \log \left( \frac{2\beta}{\beta_0 + \beta_1} \right) \quad (1)$$

其中, $\beta = W(y|x), \beta_0 = W(y|0), \beta_1 = W(y|1)$ .通过对信道

$W$ 的 $N$ 次占用(即信道 $W$ 的 $N$ 个可用时隙),能够得到 $N$ 个独立的具有相同信道特性的B-DMC信道,然后对这 $N$ 个信道进行信道变换,在各个独立的信道之间引入相关性,将会得到信道转移概率函数

$$W_N^{(i)}(y_1^N, u_1^{i-1} | u_i) = \sum_{u_{i+1}^N \in X^{N-1}} \frac{1}{2^{N-1}} W_N(y_1^N | u_1^N) \quad (2)$$

其中, $W_N(y_1^N | u_1^N) = W(y_1|x_1) \cdots W(y_N|x_N)$ .

令 $v_1^N = u_1^N F^{\otimes n}$ ,其中 $n = \log_2 N$ , $F = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$ , $F^{\otimes n} = F \otimes F^{\otimes(n-1)}$ , $\otimes$ 表示 $n$ 次克罗内克乘,以及 $u_1^N = (u_1, u_2, \dots, u_N)$ 由信息比特与固定比特混合而得到.通过对 $v_1^N$ 进行比特反序重排,最终得到编码比特序列 $x_1^N = (x_1, x_2, \dots, x_N)$ .

#### 1.1.2 极化译码

当谈及SC译码时,似然比(LLR)<sup>[13]</sup>和路径度量尤为重要.一般地,LLR的计算如下

$$L(y_i) = \frac{2y_i}{\sigma^2} \quad (3)$$

其中, $y_i$ 是通过AWGN信道获得的接收信号.LLR的递归计算方式为

$$L_N^{(2i-1)}(y_1^N, u_1^{2i-2}) = L_{N/2}^{(i)}(y_1^{N/2}, u_{1,o}^{2i-2} \oplus u_{1,e}^{2i-2}) \odot \Phi \quad (4)$$

以及

$$L_N^{(2i)}(y_1^N, u_1^{2i-1}) = L_{N/2}^{(i)}(y_1^{N/2}, u_{1,o}^{2i-2} \oplus u_{1,e}^{2i-2}) + \Phi \quad (5)$$

其中, $\Phi = L_{N/2}^{(i)}(y_{N/2+1}^N, u_{1,e}^{2i-2})$ , $a_1 \odot a_2 = \log \frac{1 + e^{(a_1+a_2)}}{e^{a_1} + e^{a_2}}$ .

与传统的SC译码一样,SCL译码仍然是从译码树的根节点开始,逐层依次向叶子节点层进行路径搜索.然而,与传统的SC译码不同的是,SCL译码用路径度量来估计每一条路径的可靠性;此外,SCL的每一层译码将会把 $L$ 条最有可能的路径保存下来.那么, $\hat{u}_{1,l}^{(i)}(l=1,2,\dots,L)$ 的路径度量的计算公式为

$$PM(\hat{u}_{1,l}^{(i)}) = PM(\hat{u}_{1,l}^{(i-1)}) + c_{i,l} \cdot \left| L_{N,l}^{(i)} \right| \quad (6)$$

其中, $PM(\hat{u}_{1,l}^{(0)}) = 0$ .当 $\hat{u}_i$ 为信息比特或正确固定比特且 $(1-2\hat{u}_i) = \text{sign}(L_N^{(i)}(y_1^N, \hat{u}_1^{i-1}))$ 时, $c_{i,l} = 0$ ;当 $\hat{u}_i$ 为信息比特或正确固定比特且 $(1-2\hat{u}_i) \neq \text{sign}(L_N^{(i)}(y_1^N, \hat{u}_1^{i-1}))$ 时, $c_{i,l} = 1$ ;当 $\hat{u}_i$ 为固定比特且取值错误时, $PM(\hat{u}_{1,l}^{(i)}) = -\infty$ .

### 1.2 BP神经网络

因强大的学习能力,神经网络吸引着众多学者<sup>[14]</sup>.BP神经网络是神经网络中的一种,本质上是一种前馈

神经网络. BP神经网络主要涉及三部分: 准备样本、搭建结构及训练网络.

在准备样本阶段, 主要是整理训练集和测试集. 在搭建结构阶段, 主要是设置参数. 在训练网络阶段, 基于监督学习, 首先计算输出层和隐藏层的输出值; 然后计算输出层和隐藏层的误差项; 在此基础上, 通过设定一个学习率来设置参数更新的大小, 并不断迭代调整权值和偏置项, 最终完成训练. 特别强调的是, 偏置项可以是固定的, 也可以设置成同权值相关的.

## 2 基于 BP 神经网络的 SCL 译码算法

由于极化码能够达到信道容量, 因此它的译码技术受到了极大关注. 近年来, 各种译码方法应运而生, 如 SC, SCL, CA-SCL, AD-SCL 等. 为了降低低信噪比下的平均译码复杂度, 综合考虑性能及 AD-SCL 对  $L$  进行由小到大更新时所带来的低平均译码复杂度, 若  $L$  的初始值不固定为 1, 那么, 就能够解决低信噪比下的复杂度问题. 因而, 寻找  $L$  的最优初始值是本文的研究关键.

假设令  $L_{\max} = 32$ , 则  $L$  的最优初始值有 6 种情况: 1, 2, 4, 8, 16, 32. 当寻找  $L$  的最优初始值时, 可以看作是简单的分类问题. 在某一信噪比和某一似然比下对应着哪个  $L$  值, 则相对应的  $L$  值即为寻找的  $L$  的最优初始值. 若用普通的计算来寻找  $L$  的最优初始值, 需要大量的数据存储和概率统计, 从而导致复杂度呈上升趋势. 而 BP 神经网络具有任意复杂的模式分类能力和优良的多维函数映射能力, 可以离线操作和准确处理分类问题, 因此, 其技术适用于寻找  $L$  的最优初始值.

本文研究主要分 2 个阶段来进行: 离线阶段和在线阶段, 具体的系统框图如图 1 所示. 观察图 1, 离线阶段主要是构建并训练 BP 神经网络; 在线阶段首先是寻找  $L$  的最优初始值, 其次是完成译码. 对于图 1 的整个系统框图, 首先执行离线操作, 完成虚框中的各项任务, 包括收集数据、构建 BP 神经网络及训练 BP 神经网络, 得到一个训练好的 BP 神经网络模型; 然后启动在线操作, 从极化编码开始, 将码字发送到信道中, 得到接收信号, 并根据接收信号来计算似然比, 一旦完成似然比的计算, 系统会激活已完成训练的 BP 神经网络模型, 将似然比及与该似然比相对应的信噪比输入到 BP 神经网络模型中, 得到一个值, 即为  $L$  的最优初始值, 最后执行译码操作.

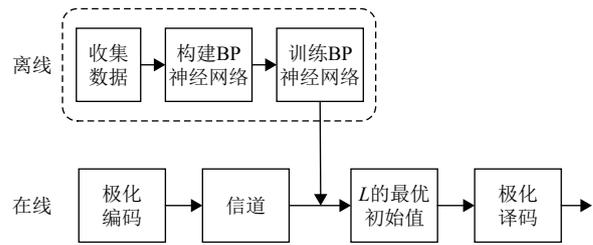


图 1 基于 BP 神经网络的极化码算法系统框图

### 2.1 构建与训练 BP 神经网络

综合考虑网络训练复杂度及极化码译码复杂度, 确定在离线阶段完成 BP 神经网络的构建与训练. 需要注意的是, 此处理方式不会影响极化码译码复杂度和性能.

#### 2.1.1 样本数据的准备及预处理

采用 BP 神经网络方法建模的首要前提条件是有足够多典型性好和精度高的样本. 为此, 建模的关键是收集并选取样本数据; 在此基础上, 随机选取其中的 75% 的数据作为训练样本, 剩余的 25% 数据作为测试样本; 最后, 通过预处理将样本数据转换为数值数据. 本文采用的预处理手段是归一化处理, 其线性转换算法的形式如下:

$$z_r = \frac{s_r - s_{\min}}{s_{\max} - s_{\min}} \quad (7)$$

其中,  $s_{\min}$  为输入向量  $\mathbf{S}$  的最小值,  $s_{\max}$  为输入向量  $\mathbf{S}$  的最大值,  $z_r$  是节点  $r$  的输出值.

#### 2.1.2 神经网络的训练

基于监督学习, 利用反向传播算法来训练网络. 假设每个训练样本为  $(\mathbf{S}, g)$ , 向量  $\mathbf{S}$  是训练样本的特征,  $g$  是样本的目标值.

本文取网络所有输出层节点的误差平方和作为目标函数:

$$E_n = \frac{1}{2} \sum_{r \in \text{outputs}} (g_r - z_r)^2 \quad (8)$$

其中,  $E_n$  表示样本  $n$  的误差. 在此基础上, 运用随机梯度下降优化方法对目标函数进行优化. 经过优化之后, 能分别得到输出层的误差项、隐藏层的误差项以及权重的更新方法. 对于输出层节点  $r$ , 误差项的计算为

$$\delta_r = z_r(1 - z_r)(g_r - z_r) \quad (9)$$

其中,  $\delta_r$  是节点  $r$  的误差项,  $z_r$  是节点  $r$  的输出值,  $g_r$  是样本对应于节点  $r$  的目标值. 对于隐藏层节点, 误差项的计算为

$$\delta_r = a_r(1 - a_r) \sum_{k \in \text{outputs}} \omega_{kr} \delta_k \quad (10)$$

其中,  $a_r$  是节点  $r$  的输出值,  $\omega_{kr}$  是节点  $r$  到它的下一层节点  $k$  的连接权重,  $\delta_k$  是节点  $r$  的下一层节点  $k$  的误差项. 由此, 更新每个连接上的权重:

$$\omega_{jr} \leftarrow \omega_{jr} + \eta \delta_j s_{jr} \quad (11)$$

其中,  $\omega_{jr}$  是节点  $r$  到节点  $j$  的权重,  $\eta$  是一个学习速率常数,  $\delta_j$  是节点  $j$  的误差项,  $s_{jr}$  是节点  $r$  传递给节点  $j$  的输入. 由于偏置项的输入值永远为 1, 故偏置项的计算为

$$\omega_{rb} \leftarrow \omega_{rb} + \eta \delta_r \quad (12)$$

根据上述训练规则, 不断修正  $\omega_{jr}$  与  $\omega_{rb}$ , 则可完成训练.

## 2.2 基于 BP 神经网络的 SCL 译码算法

为了降低平均译码复杂度, 本文提出了一种基于 BP 神经网络的 SCL (BNN-SCL) 译码算法. BNN-SCL 译码算法包括两个阶段: 离线构建并训练 BP 神经网络模型阶段与在线译码阶段. 在第 2.1 节中, 已经完成 BP 神经网络模型的构建与训练, 对应地, 即已完成图 1 系统框图中虚框的操作. 因此, 在本节中主要介绍的是在线阶段的译码操作, 即从图 1 的得到接收信号开始介绍. 根据接收信号来计算似然比, 一旦完成似然比的计算, 系统会激活已完成训练的 BP 神经网络模型, 将似然比及与该似然比相对应的信噪比输入到 BP 神经网络模型中, 得到一个值, 即为  $L$  的最优初始值. 此时, 译码端将会进行初始化, 将  $L$  初始化为  $L$  的最优初始值. BNN-SCL 算法与现存的译码算法不同,  $L$  的初始值的大小不固定为 1.

为详细了解译码端进行初始化之后的操作, 本文给出具体的后续操作步骤, 即本文提出的降低平均译码复杂度的 SCL 译码算法在译码端的具体步骤如下:

- 1) 对  $L$  进行初始化, 将  $L$  初始化为  $L$  的最优初始值;
- 2) 执行 SCL 译码算法, 并对其译码候选路径进行 CRC 校验;
- 3) 如果有一条或多于一条的候选路径通过 CRC 校验, 则输出一条最有可能的路径, 并退出译码; 否则, 跳到 4).
- 4) 将  $L$  更新为  $2L$ , 并判断更新之后的  $L$  是否大于  $L_{\max}$ . 若不大于  $L_{\max}$ , 则跳回 2); 否则, 退出译码.

为辅助理解, 图 2 给出了基于 BP 神经网络的 SCL 译码算法在译码端的流程图.

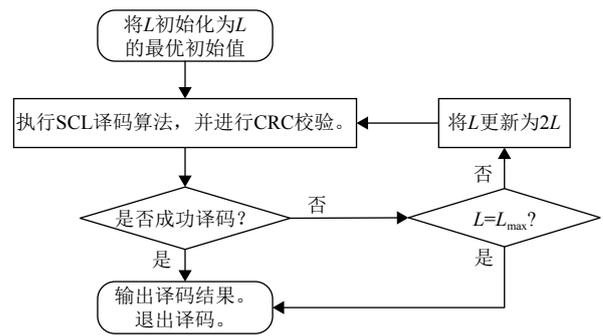


图 2 BNN-SCL 译码算法在译码端的流程图

## 3 实验结果与分析

本文的仿真实验包括 2 个部分: 离线部分和在线部分. 现将各部分主要涉及的参数、配置列出. 离线部分, 将输入层、隐藏层及输出层的节点数分别设置为 2、300、6, 采用全连接方式来搭建网络, 并将激活函数设置为 sigmoid 函数, 网络训练的目标误差设为 0.01, 显示中间结果的周期设为 50, 最大迭代次数设为 500, 学习率设为 0.01. 在线部分的参数、配置如下所示.

根据实验的数值结果来分析新提出算法的误比特率 (BER) 和复杂度. 在本文的仿真实验中, 所涉及的译码方案的码长和码率分别配置为  $N=1024$ ,  $R=0.5$ . 根据文献[15]构造极化码, 极化码级联一个长度为 16 比特的 CRC.

图 3 给出不同译码算法下的 BER 性能比较, 其中列表译码的列表大小分别配置为 4 和 32. 对于 AD-SCL 和 BNN-SCL 译码算法, 均将最大搜索宽度设置为  $L_{\max} = 32$ . 通过已完成训练的 BP 神经网络来寻找的  $L$  的最优初始值被运用于仿真实验中. 从图 3 中可以看出, CA-SCL、AD-SCL 及 BNN-SCL 的曲线是重叠的. 与 CA-SCL 和 AD-SCL 类似, 经过适当配置, BNN-SCL 也能够获得非常接近 ML 译码的性能.

图 4 给出不同译码算法下的平均复杂度比较, 其中列表译码的列表大小为 32. 对于 AD-SCL 和 BNN-SCL 译码算法, 均将最大搜索宽度设置为  $L_{\max} = 32$ . 如图 4 所示, 在高信噪比下, BNN-SCL 译码的复杂度非常接近 SC 译码的复杂度. 此外, 与 AD-SCL 译码算法相比, 在 0~1.25 dB 下, BNN-SCL 能够显著降低平均复杂度, 例如, 在  $SNR = 0.75$  时, BNN-SCL 能够降低约 26.1% 的复杂度.

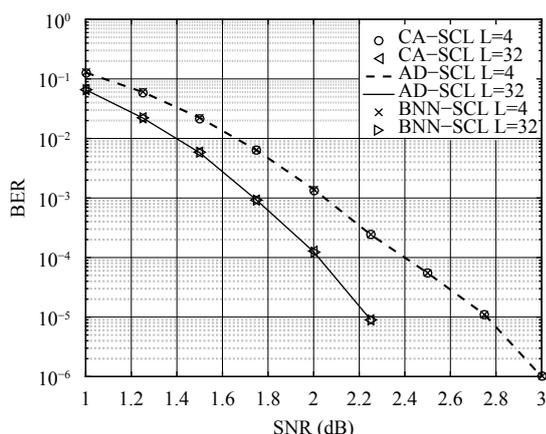


图3 不同译码算法下的 BER 性能比较

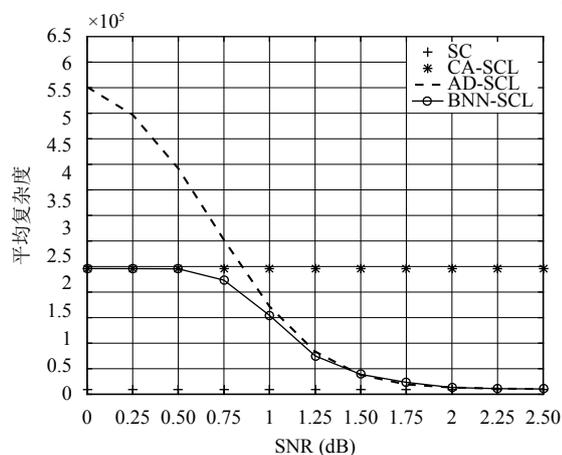


图4 不同译码算法下的平均复杂度比较

#### 4 结语

为了降低复杂度,提出一种基于BP神经网络的SCL译码算法.综合考虑现存算法的优点,通过在离线阶段完成训练的BP神经网络来寻找 $L$ 的最优初始值,以改变 $L$ 的初始值总是设置为1的现状;在此基础上,设计一种改进的SCL译码算法,最终降低复杂度.实验结果表明,与AD-SCL译码算法相比,在0~1.25 dB下,BNN-SCL能够显著降低平均复杂度且保持译码性能不变.

#### 参考文献

- Arıkan E. Channel polarization: A method for constructing capacity-achieving codes for symmetric binary-input memoryless channels. *IEEE Transactions on Information Theory*, 2009, 55(7): 3051–3073. [doi: 10.1109/TIT.2009.2021379]
- Yuan B, Parhi KK. Early stopping criteria for energy-efficient low-latency belief-propagation polar code decoders. *IEEE Transactions on Signal Processing*, 2014, 62(24): 6496–6506. [doi: 10.1109/TSP.2014.2366712]
- Guo J, Qin MH, Fàbregas AGI, *et al.* Enhanced belief propagation decoding of polar codes through concatenation. *Proceedings of 2014 IEEE International Symposium on Information Theory*. Honolulu, HI, USA, 2014. 2987–2991.
- Goela N, Korada SB, Gastpar M. On LP decoding of polar codes. *Proceedings of 2010 IEEE Information Theory Workshop*. Dublin, Ireland, 2010. 1–5.
- Tal I, Vardy A. List decoding of polar codes. *Proceedings of 2011 IEEE International Symposium on Information Theory*. St. Petersburg, Russia, 2011. 1–5.
- Tal I, Vardy A. List decoding of polar codes. *IEEE Transactions on Information Theory*, 2015, 61(5): 2213–2226. [doi: 10.1109/TIT.2015.2410251]
- 陈凯. 极化编码理论与实用方案研究[博士学位论文]. 北京: 北京邮电大学, 2014.
- Niu K, Chen K. CRC-aided decoding of polar codes. *IEEE Communications Letters*, 2012, 16(10): 1668–1671. [doi: 10.1109/LCOMM.2012.090312.121501]
- Li SB, Lu LJ, Deng YQ, *et al.* A reused-public-path successive cancellation list decoding for polar codes with CRC. *IEEE Communications Letters*, 2017, 21(12): 2566–2569. [doi: 10.1109/LCOMM.2017.2756642]
- Li B, Shen H, Tse D. An adaptive successive cancellation list decoder for polar codes with cyclic redundancy check. *IEEE Communications Letters*, 2012, 16(12): 2044–2047. [doi: 10.1109/LCOMM.2012.111612.121898]
- Haykin S. *Neural Networks and Learning Machines*. 3rd ed. Upper Saddle River: Pearson, 2008.
- 高鹏毅. BP神经网络分类器优化技术研究[博士学位论文]. 武汉: 华中科技大学, 2012.
- Balatsoukas-Stimming A, Parizi MB, Burg A. LLR-based successive cancellation list decoding of polar codes. *IEEE Transactions on Signal Processing*, 2015, 63(19): 5165–5179. [doi: 10.1109/TSP.2015.2439211]
- 焦李成, 杨淑媛, 刘芳, 等. 神经网络七十年: 回顾与展望. *计算机学报*, 2016, 39(8): 1697–1716.
- Tal I, Vardy A. How to construct polar codes. *IEEE Transactions on Information Theory*, 2013, 59(10): 6562–6582. [doi: 10.1109/TIT.2013.2272694]