基于移位等效码字熵率计算的加扰卷积码盲识别

王中方[1,2], 黄伟庆[1,2], 翟留群[1,2], 胡可可[3], 魏 冬[1,2]

[1] 中国科学院大学网络空间安全学院, 北京 中国 100093
[2] 中国科学院信息工程研究所, 北京 中国 100049
[3] 清华大学电子工程系, 北京 中国 100084

摘要 在非合作通信的研究中, 加扰编码码字的盲识别具有关键作用。现有的加扰编码码字盲识别研究主要集中在单一信道编码盲识别或扰码盲识别, 这些方法在实际系统中往往不适用于加扰与编码的级联场景, 且在误码情况下识别效率较低。为解决这一问题, 本文提出了一种面向加扰卷积码级联场景的扰码与卷积码联合盲识别算法, 该算法基于移位等效码字熵率。首先, 本文利用卷积码字加扰后的性质构造移位等效码字, 从而将扰码盲识别问题转化为等效卷积码判决问题。其次, 由于使用传统算法判决移位等效码字过于复杂并且运算资源消耗过高, 本文提出了一种基于信息熵率的卷积码快速判断方法, 并推导出了算法实现所需的相关参数, 实现了低复杂度和高效率的快速联合盲识别。通过仿真实验, 我们证明了本文所提方法能有效地对加扰卷积码进行联合识别。在信道传输误比特率小于 6%的情况下, 扰码识别正确率达到了 84.5%以上, 扰码和卷积码的联合识别率超过了 80%, 展现了良好的抗噪能力。

关键词 扰码; 卷积码; 联合盲识别; 熵率
中图法分类号 TN92 DOI 号 10.19363/J.cnki.cn10-1380/tn.2023.08.08

Blind Recognition of Scrambled Convolutional Code Based on Calculating Shift Equivalent Codeword Entropy Rate

WANG Zhongfang[1,2], HUANG Weiqing[1,2], ZHAI Liuqun[1,2], HU Keke[3] , WEI Dong[1,2]

[1] School of Cyber Security, University of Chinese Academy of Sciences, Beijing 100093, China
[2] Institute of Information Engineering, Chinese Academy of Sciences, Beijing 100049, China
[3] Department of Electronic Engineering, Tsinghua University, Beijing 100084, China

Abstract In the field of non-cooperative communication research, the blind identification of codewords that have undergone scrambling and encoding is of critical significance. Existing studies primarily concentrates on two aspects: the blind identification of individual channel-encoded codewords and the blind identification of scrambled codewords. These methods typically falter in complex scenarios where scrambling and encoding are concatenated, and they also suffer from reduced recognition efficiency in the presence of coding errors. To overcome this limitation, this paper introduces a novel algorithm that enables joint blind identification of both scrambling and convolution codes within concatenated configurations. The foundation of this algorithm lies in the entropy rate of shifted equivalent codewords. To begin with, we leverage the inherent characteristics of convolution codewords after scrambling to construct shifted equivalent codewords. This innovation transforms the challenge of scrambling blind identification into a more manageable problem of determining equivalent convolution codes. Traditional algorithms used for deciding on shifted equivalent codewords are known for their complexity and exorbitant consumption of computational resources. Acknowledging this obstacle, we propose a streamlined decision-making method specifically tailored for convolution codes, which is based on the information entropy rate. The relevant parameters required for the successful implementation of this algorithm are derived with precision, resulting in a solution characterized by both low computational complexity and high operational efficiency. Our simulation experiments provide compelling evidence for the effectiveness of the proposed method in the context of joint identification of scrambled convolution codes. In scenarios where the channel transmission bit error rate is maintained below 6%, our method achieved a scrambling recognition accuracy exceeding 84.5%, and a joint recognition rate for both scrambling and convolution codes that surpassed 80%. These results not only affirm the efficacy of our approach but also demonstrate its robust resistance to noise, underlining its potential applicability across a wide spectrum of non-cooperative communication systems.

通讯作者: 王中方, Email: wangzhongfang@iie.ac.cn。
本课题得到中国科学院 C 类战略性先导科技专项(No. XDC02040300)资助。
收稿日期: 2021-01-06; 修改日期: 2021-02-03; 定稿日期: 2023-08-10

Key words scrambler; convolutional code; joint blind recognition; entropy rate

1 引言

在非合作通信研究领域中, 信道编码盲识别研究一直都是一个热门课题[1-3]。信道编码盲识别是指在先验信息较少甚至完全没有的情况下, 从截获的数据序列中识别出信道编码的类型和编码参数的技术。信道编码盲识别技术在认知通信和信息对抗等领域有着可观的应用前景。例如, 在 AMC(Adaptive Modulation and Coding, 自适应调制编码)中, 为了充分利用信道容量, 发送端根据信道状态不断调整着发射信息的调制样式和信道编码参数, 接收端需要在控制信令未知时, 识别出调制样式与信道编码方式; 在信息对抗中, 只有完成信道编码的分析与识别, 才能正确解码或破译出截获信号携带的信息。因此, 信道编码盲识别有着重要的研究意义与应用价值。

信道编码是通过增加冗余以提高信道可靠性, 使得通信系统具备检错或者纠错的能力。常用的信道编码类型包括 Hamming 码、BCH 码、RS 码、卷积码以及 Turbo 码等。其中, 由于卷积码利用了各组输入数据之间的相关性, 其编码纠错性能一般优于分组码, 在 GSM、TD-SCDMA、LTE 等实际通信系统中有着广泛应用。

实际通信中, 发送信息序列中很可能出现“长连 1”、“长连 0”的情况, 不利于接收端信号提取。因此, 通信系统中引入了扰码技术, 通过对数据进行扰乱处理, 使得原来数据序列具有伪随机特性, 保证接收端可以提取到定时信息。此外, 扰码也可以在一定程度上增强通信保密性。根据加扰方式的不同, 扰码可以分为同步扰码与自同步扰码。在扰码识别研究中, 识别自同步扰码需要分析出 LFSR(Linear Feedback Shift Register, 线性反馈移位寄存器)序列的生成多项式[4-9], 而识别同步扰码除了分析 LFSR 序列的生成多项式, 还需要确定寄存器的初始状态[10-14]。本文主要解决了在卷积码与同步扰码联合使用场景下, 如何仅利用接收数据同时识别扰码生成多项式与卷积编码参数的问题。

目前的研究中, 针对卷积码与扰码联合识别场景, 一般是先对扰码进行识别, 解扰后再对卷积码进行识别。其中, 单独针对卷积码的识别已经有大量的文献进行了研究[15-28]; 而对扰码的分析与识别主要利用了信源序列的 0、1 不平衡的统计特性, 通过遍历生成多项式的集合并建立判决指标来进行分析。在文献[29]中, Cluzeau 提出了使用加扰后的数据构造变量 Z, 然后基于变量 Z 绝对值的统计检验在稀疏多项式集合进行遍历搜寻生成多项式的倍式。在找到生成多项式的两个不同稀疏倍式之后, 计算两个倍式的最大公约式即为同步扰码的生成多项式, 该算法也被称为 Cluzeau 算法。文献[30]在信源序列不平衡度未知时, 推导出了如何通过变量 Z 与重建加扰器需要的比特数目 N 得到不平衡度, 再利用 Cluzeau 算法识别出生成多项式。文献[31]是对 Cluzeau 算法的改进, 解决了当搜索到的第二个稀疏倍式是第一个稀疏倍式的倍式时出现的识别错误问题; 同时考虑了数据经过噪声信道的情况, 证明了 Cluzeau 算法经过改进后依然适用。文献[32-33]主要研究了对接收数据进行相关处理后, 利用 Walsh-Hadamard 变换得到 Walsh 谱进而识别生成多项式。然而, 信源信息通过卷积编码后, 其不平衡度可能会大大降低, 上述扰码识别算法在卷积码与扰码联合使用场景下, 识别结果可靠性将会大大降低。

文献[34]针对“纠错编码-加扰”的场景, 提出了基于校验向量的识别算法。该算法利用了 $\boldsymbol{c}\bullet\boldsymbol{H}^{T}=\boldsymbol{0}$ 的性质, 将接收数据分组后乘以校验向量得到数据序列后再构造变量 Z, 然后与 Cluzeau 算法一样在稀疏多项式集合中进行遍历搜寻生成多项式的倍式。文献[35]在“卷积编码-加扰”场景下, 将扰码生成多项式的识别问题转化为求解由扰码生成多项式系数建立的代价函数问题, 进而得到扰码的生成多项式。然而, 文献[34-35]的方法只能识别事先存储校验向量所对应的卷积码, 限制了其应用范围。文献[36]提出通过接收数据矩阵的零空间识别扰码生成多项式, 解决了信源不平衡度和校验向量未知情况下的盲识别问题, 但是其使用接收矩阵的秩识别卷积码的编码参数的方法, 导致该算法在噪声情况下的识别可靠性不高。

本文针对“卷积编码-加扰”的实际通信场景, 首先利用扰码生成多项式倍式的性质, 通过构建接收数据的移位组合矩阵, 将扰码生成多项式识别问题转化为卷积码判别问题; 然后, 基于卷积码序列熵率小于随机序列熵率的特性, 本文提出了低复杂度的卷积码快速判别方法, 解决了扰码识别问题; 最后, 根据卷积码序列叠加后仍满足卷积码性质的特性, 提出了卷积码参数重构方法, 识别出了卷积码参数。通过在无噪声、有噪声等场景下的仿真分析,

本文所提的联合盲识别算法可以实现同步扰码生成多项式与卷积码参数的同时识别, 且具有较高的鲁棒性, 在误比特率较高的情况下仍然保持较高的识别正确率。

2 系统模型概述

2.1 扰码与卷积码描述

图 1 中, 同步加扰器的加扰过程是将输入序列与 LFSR 序列 $\{s_t\}$ 进行模二相加, 即可得到加扰输出序列。当输入序列为 $\{c_t\}$ 时, 输出序列 $\{y_t\}$ 与 $\{c_t\}$ 的关系满足式(1)。其中, 式(1)等号右边的第二个符号 $\overset{n}{\underset{i=1}{\oplus}}(\bullet)$ 表示 n 个数据进行模二运算。

$$y_t = c_t \oplus s_t = c_t \oplus (\overset{n}{\underset{i=1}{\oplus}} C_i s_{n-i}) \tag{1}$$

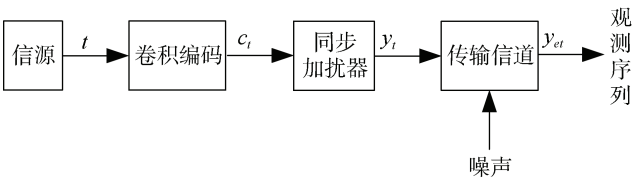

图 1 通信系统模型

Figure 1 Communication system model

同步加扰器的结构如图 2 所示。其中, C_i 表示反馈系数, $C_i \in \{0,1\}$。当 $C_i = 1$ 时, 表示反馈线接通, $C_i = 0$ 则表示反馈线断开。由图 2 可知, 反馈系数 C_i 决定产生的 LFSR 序列 $\{s_t\}$, 其生成多项式表示为 $f(x) = 1 + C_1 x + \cdots + C_n x^n$。本文识别的目标之一就是得到 LFSR 序列的生成多项式 $f(x)$。

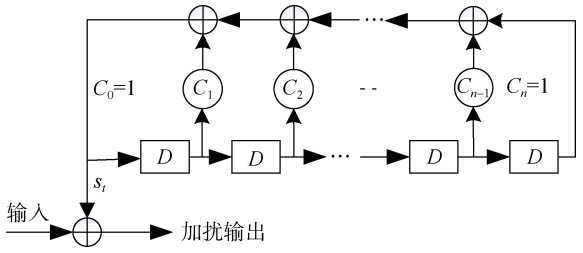

图 2 同步扰码器结构

Figure 2 Structure of a synchronous scrambler

卷积编码是将待编码数据按照特定规则进行运算, 并加入一定的冗余信息, (n,k,K) 卷积编码器结构如图 3 所示。当卷积编码器工作时, 先将输入信息按照 k 位进行分组, 每次输入编码器的 k 位信息比特按照图 3 的规则产生一个 n 位的码组。由于存在 K 级寄存器组, 每级为 k 位移位寄存器, 所以当前时刻产生的一个 n 位码组不仅与此时刻输入的 k 位信息有关, 还与前连续 $(K-1)$ 段的输入信息相关。在实际应用中, 受限于编码和译码的复杂度, 卷积码的码长 n 与信息位长度 k 一般较小。识别卷积编码的参数就是要得到卷积编码中的码字长度 n、信息位长度 k、约束长度 K 以及表示“输入-输出”生成规则的生成矩阵 $\boldsymbol{G}$ 或校验矩阵 $\boldsymbol{H}$。

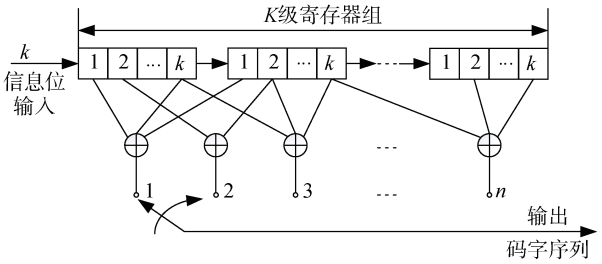

图 3 卷积编码器结构

Figure 3 Structure of convolutional encoder

2.2 加扰卷积码盲识别概述

本小节将简单介绍加扰卷积码盲识别的过程。以 (2,1,3) 卷积码为例, 对应图 3 中的编码过程为 $n=2$, $k=1$, $K=3$。令三位移位寄存器的初始状态全为零状态, (2,1,3) 卷积码中 k 为 1, 即每次只有 1 个比特进入编码器。当第一个输入比特为 1 时, 三个寄存器中的值分别是 100, 第一个输出码字则为 11; 接着, 如果第二个输入比特为 1, 三个寄存器中的值分别是 110, 第二个输出码字为 10; 如果第三个输入比特为 0, 三个寄存器中的值分别是 011, 第三个输出码字为 10。以此类推, 输出的编码序列为 11, 10, 10, …。

若将每一路输出看作输入信息与该路生成矢量的卷积, 则图 4 中的两路输出分别对应的生成矢量为 $\boldsymbol{g}_1 = (101), \boldsymbol{g}_2 = (111)$。将输入信息序列表示为 $\boldsymbol{m} = (m_0, m_1, \cdots, m_n, \cdots)$, 编码器的两路输出分别表示为 $\boldsymbol{c}^{(1)} = (c_1^{(1)}, c_2^{(1)}, \cdots)$, $\boldsymbol{c}^{(2)} = (c_1^{(2)}, c_2^{(2)}, \cdots)$, 即有:

$$\begin{aligned} \boldsymbol{c}^{(1)} &= \boldsymbol{m} \otimes \boldsymbol{g}_1 \\ \boldsymbol{c}^{(2)} &= \boldsymbol{m} \otimes \boldsymbol{g}_2 \end{aligned} \tag{2}$$

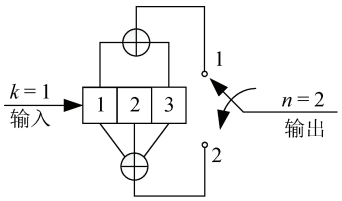

图 4 (2,1,3) 卷积编码结构

Figure 4 Structure of (2,1,3) convolutional encoder

将两个输出序列并/串转换得到一个序列, 即为输出码字序列 $\boldsymbol{c}$ 。

$$\boldsymbol{c}=(c_1^{(1)},c_1^{(2)},c_2^{(1)},c_2^{(2)},c_3^{(1)},c_3^{(2)},\cdots) \tag{3}$$

卷积码除了可以用向量卷积进行表示, 还可以用多项式乘法来表示。令符号 D 表示一个延迟算子, 那么信息序列就可以用多项式表示为:

$$\boldsymbol{m}(D)=m_0+m_1D+m_2D^2+\cdots+m_nD^n+\cdots \tag{4}$$

两个生成矢量分别用多项式可以表示为: $\boldsymbol{g}_1(D)=1+D^2,\boldsymbol{g}_2(D)=1+D+D^2$, 两路输出序列用多项式可以分别表示为:

$$\begin{aligned}\boldsymbol{c}^{(1)}(D)&=\boldsymbol{m}(D)\boldsymbol{g}_1(D)\\ \boldsymbol{c}^{(2)}(D)&=\boldsymbol{m}(D)\boldsymbol{g}_2(D)\end{aligned} \tag{5}$$

输出的码字序列 $\boldsymbol{c}$ 用多项式表示为:

$$\boldsymbol{c}(D)=\boldsymbol{c}^{(1)}(D^2)+D\boldsymbol{c}^{(2)}(D^2) \tag{6}$$

根据图 1 中的系统模型, 构建加扰卷积码识别模型如图 5 所示: (1) 从观测序列 $\{y_{et}\}$ 中解扰出码字序列 $\{c_t\}$; (2) 从解扰码字中识别卷积编码的参数。但是由于数据经过信道编码的处理, 其不平衡度大大降低, 传统算法中利用数据不平衡度扰码识别方法不再适用, 需要设计适应于加扰卷积码字的扰码识别方法。

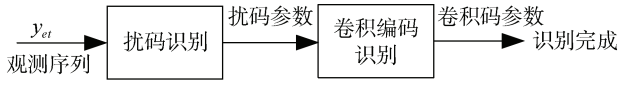

图 5 加扰卷积码盲识别模型

Figure 5 Blind recognition model of scrambled convolutional codes

本文将扰码-卷积码联合识别问题分解为如下两个子问题:

(1) 如何从观测序列 $\{y_{et}\}$ 中实现同步扰码的识别, 也即识别出 LFSR 序列的生成多项式 $f(x)$ 和扰码寄存器的初始状态。由于在得到扰码生成多项式后, 可以通过快速相关攻击等方法确定寄存器的初始状态, 因此本文不再涉及识别同步扰码寄存器初态的内容。

(2) 如何在识别出扰码后, 得到码字序列 $\{c_t\}$ 并识别出卷积编码的参数。

3 扰码与卷积码联合盲识别算法

3.1 基于移位等效码字的扰码盲识别

本小节针对同步扰码的识别, 提出了适应于加扰卷积码的 LFSR 序列生成多项式的识别方法, 将加扰卷积码的扰码识别问题转化为移位等效卷积码判定问题, 降低了不平衡度下降带来的影响。

在传统方法中, 扰码生成多项式的识别基于扰码生成多项式的稀疏多项倍式搜索, 根据搜索得到的两个稀疏多项倍式的最大公约式来确定生成多项式 $f(x)$。当 $f(x)=1+\mathrm{C}_1x+\cdots+\mathrm{C}_nx^n$, 产生的 LFSR 序列 $\{s_t\}$ 为:

$$s_{t+n}=\mathrm{C}_ns_t\oplus\mathrm{C}_{n-1}s_{t+1}\oplus\cdots\oplus\mathrm{C}_1s_{t+n-1} \tag{7}$$

根据式(7)可以看出, 每一时刻产生的加扰输出都与前 n 个时刻的加扰输出有线性关系。式(7)可以改写为:

$$s_{t+n}\oplus\mathrm{C}_ns_t\oplus\mathrm{C}_{n-1}s_{t+1}\oplus\cdots\oplus\mathrm{C}_1s_{t+n-1}=0 \tag{8}$$

因此, 若用于加扰的数据序列 $\{s_t\}$ 满足式(9), 其中, d 为等式左边加扰比特的个数。

$$s_t\oplus\sum_{j=1}^{d-1}s_{t-i_j}=0,0<i_1<i_2<\cdots<i_{d-1} \tag{9}$$

则可以得到, 多项式 $Q(x)=1+\sum_{j=1}^{d-1}x^{i_j}$ 是扰码生成多项式 $f(x)$ 的倍式。因此, 只要得到满足式(9)的等式, 就可以找出 $f(x)$ 的倍式 $Q(x)$ 进而确定 $f(x)$。本节中识别同步扰码的生成多项式就利用到了这个性质。

为了确定同步扰码生成多项式 $f(x)$ 的倍式 $Q(x)$, 首先利用观测数据序列 $\boldsymbol{y_t}=\{y_1,y_2,\cdots,y_n\cdots\}$ 建立接收数据的移位组合矩阵 $\boldsymbol{Y}$。

$$\boldsymbol{Y}=\begin{bmatrix}y_1 & y_2 & y_3 & \cdots & y_n & \cdots\\ y_2 & y_3 & y_4 & \cdots & y_{n+1} & \cdots\\ y_3 & y_4 & y_5 & \cdots & y_{n+2} & \cdots\\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots\end{bmatrix} \tag{10}$$

其中, y_i 是码字比特 c_i 与加扰比特 s_i 的模二和, 即 $y_i=c_i\oplus s_i$。不妨将移位组合矩阵 $\boldsymbol{Y}$ 的第 i 行记作 r_i, 即 $r_i=\begin{bmatrix}y_i & y_{i+1} & \cdots & y_{i+(n-1)} & \cdots\end{bmatrix}$, $r_i\oplus r_j$ 表示第 i 行与第 j 行对应数据进行模二和运算后组成的新序列, 即:

$$r_i\oplus r_j=\begin{bmatrix}y_i\oplus y_j & \cdots & y_{i+(n-1)}\oplus y_{j+(n-1)} & \cdots\end{bmatrix} \tag{11}$$

为了确定扰码生成多项式 $f(x)$, 我们每次取矩阵 $\boldsymbol{Y}$ 的 N 行进行模二和运算得到序列 $\boldsymbol{r}$。如果将取的 N 行数据分别表示为 $\boldsymbol{r}_{i_1},\boldsymbol{r}_{i_2},\cdots,\boldsymbol{r}_{i_N}$, 其中 $\boldsymbol{r}_{i_j}$ 的下标 $i_j(1\leqslant j\leqslant N)$ 表示其为矩阵 $\boldsymbol{Y}$ 的第 i_j 行, 那么序列 $\boldsymbol{r}$ 可以表示为:

$$\boldsymbol{r}=\boldsymbol{r}_{i_1}\oplus\boldsymbol{r}_{i_2}\oplus\cdots\oplus\boldsymbol{r}_{i_N},0<i_1<i_2<\cdots<i_N \tag{12}$$

代入矩阵 $\boldsymbol{Y}$ 的数据, 序列 $\boldsymbol{r}$ 可以改写为:

$$\boldsymbol{r}=\boldsymbol{r}_{i_1}\oplus\boldsymbol{r}_{i_2}\oplus\cdots\oplus\boldsymbol{r}_{i_N}$$

$$=\begin{bmatrix} c_{(1)}\oplus s_{(1)} \\ c_{(2)}\oplus s_{(2)} \\ \vdots \\ c_{(n)}\oplus s_{(n)} \\ \vdots \end{bmatrix}^T \tag{13}$$

其中, $c_{(i)}$ 为 N 个码字比特的模二和, $s_{(i)}$ 为 N 个加扰比特的模二和, 即:

$$\begin{aligned} c_{(i)}&=(c_{i_1+i}\oplus c_{i_2+i}\oplus\cdots\oplus c_{i_N+i}) \\ s_{(i)}&=(s_{i_1+i}\oplus s_{i_2+i}\oplus\cdots\oplus s_{i_N+i}) \end{aligned} \tag{14}$$

当 $s_{(i)}=0$ 时, 即 $s_{i_1+i}\oplus s_{i_2+i}\oplus\cdots\oplus s_{i_N+i}=0$ 时, 序列 $\boldsymbol{r}$ 为 $\begin{bmatrix} c_{(1)} & c_{(2)} & \cdots & c_{(n)} & \cdots \end{bmatrix}$, 是由纯码字数据叠加组成的序列。根据式(9)中的扰码生成多项式的性质可得, $Q(x)=1+\sum_{j=1}^{N-1}x^{i_N-i_{(N-i)}}$ 是加扰生成多项式 $f(x)$ 的倍式。

而当 $s_{(i)}\neq 0$ 时, 序列 $\boldsymbol{r}$ 为码字数据叠加后再经加扰的序列, 此时 $Q(x)=1+\sum_{j=1}^{N-1}x^{i_N-i_{(N-i)}}$ 不是加扰生成多项式 $f(x)$ 的倍式。

因此, 可知如果序列 $\boldsymbol{r}$ 是纯码字移位叠加后的序列, 则多项 $Q(x)$ 是扰码生成多项式 $f(x)$ 的倍式。本文对码字移位叠加后的倍式性质进行归纳总结, 得到如下性质推论:

性质 1: (n,k,K) 卷积码字序列与其经过移位后的卷积码字序列叠加得到的序列为 (n,k,K') 卷积码字序列, 即移位叠加等效的卷积编码输入、输出位数不变, 只有约束长度与生成多项式矩阵发生了改变。移位组合得到的新卷积码与原卷积码的相关参数关系如表 1 所示。

性质 1 证明如下:

首先分析第 2 节中的 (2,1,3) 卷积码序列的移位叠加的结果, 然后将规律推广到一般的卷积编码。现有 (2,1,3) 卷积码的两种不同移位长度的序列, 不妨假设序列 I 为原卷积码序列, 即移位长度为 0; 序列 II 为原卷积码序列移位三个长度得到的序列。两种序列相互叠加后, 即在对应的位置上进行模二和运算, 可以得到序列III。

表 1 移位组合得到的新卷积码与原卷积码关系

Table 1 The relationship between the new convolutional code obtained by the shift combination and the original convolutional code

卷积码相关参数	两者关系
信息位长度 k	相等
码字长度 n	相等
码率 n	相等
约束长度 K	不相等, 由移位长度决定
生成多项式矩阵 G	不相等, 由移位长度决定

序列 I: $y_{11},y_{12},y_{21},y_{22},\cdots,y_{i1},y_{i2},\cdots$

序列 II: $y_{22},y_{31},y_{32},y_{41},\cdots,y_{j1},y_{j2},\cdots$

序列 III: $y_{11}\oplus y_{22},y_{12}\oplus y_{31},y_{21}\oplus y_{32},\cdots$

在序列 III 中, $\{y_{11},y_{21},y_{31},\cdots\}$ 由 $g_1(D)$ 生成, $\{y_{22},y_{32},y_{42},\cdots\}$ 由 $g_2(D)$ 生成, 根据卷积码的生成多项式可得, 序列 III 满足的生成多项式为:

$$\begin{aligned} g_1'(D)&=g_1(D)\oplus g_2(D)\bullet D^{-1} \\ g_2'(D)&=g_2(D)\oplus g_1(D)\bullet D^{-2} \end{aligned} \tag{15}$$

其中, D^{-i} 表示比当前时刻提前了 i 个时刻。将 $g_1(D),g_2(D)$ 代入到式(15)中, 可以得到:

$$\begin{aligned} g_1'(D)&=D^{-1}+D+D^2 \\ g_2'(D)&=D^{-2}+D+D^2 \end{aligned} \tag{16}$$

需要注意的是, 叠加后序列的当前时刻是以最长移位序列中产生第一个码字比特的时刻作为起始时刻的。而在实际的卷积编码中, 某一输入时刻的提前时刻输入可能是不存在的, 因此在不改变编码规则的前提下, 可以将 $g_1'(D),g_2'(D)$ 变换为:

$$\begin{aligned} g_1'(D)&=1+D^2+D^3 \\ g_2'(D)&=1+D^3+D^4 \end{aligned} \tag{17}$$

可以看到: (2,1,3) 卷积码序列的不同长度移位序列进行叠加后得到的新序列, 仍然符合卷积码序列性质。新卷积码序列的码长 n', 信息位 k' 都和原卷积码相同, 只有约束度 K 由 3 变化为 5。

考虑更为一般的 $(n,1,K)$ 卷积码序列, 其生成多项式由式(18)表示, 其中 $g_{i,j}\in\{0,1\}$, $m=K-1$。

$$\begin{aligned} g_1(D)&=g_{10}+g_{11}D+\cdots+g_{1m}D^m \\ g_2(D)&=g_{20}+g_{21}D+\cdots+g_{2m}D^m \\ &\vdots \\ g_n(D)&=g_{n0}+g_{n1}D+\cdots+g_{nm}D^m \end{aligned} \tag{18}$$

令 $(n,1,K)$ 卷积码序列为 $y_{11},y_{12},\cdots,y_{1n},y_{21},y_{22},\cdots$ $y_{2n},\cdots,y_{i1},y_{i2},\cdots,y_{in},\cdots$。该序列经过不同的移位长度 p,q 进行移位后得到序列 IV 与序列 V, 两个序列叠加后得到序列 VI。

序列 IV: $y_{1,(p+1)},y_{1,(p+2)},\cdots,y_{1,n},y_{21},y_{22},\cdots,y_{2n},\cdots$

序列 V: $y_{1,(q+1)},y_{1,(q+2)},\cdots,y_{1,n},y_{21},y_{22},\cdots,y_{2n},\cdots$

序列 VI: $y_{1,(p+1)}\oplus y_{1,(q+1)},y_{1,(p+2)}\oplus y_{1,(q+2)},\cdots$

利用上面对 (2,1,3) 卷积码移位叠加的推导, 同样可以得到序列 VI 的生成多项式为:

$$\begin{aligned}
&g_1'(D)=g_{a_1}(D)\bullet D^{-\left\lfloor\frac{p_1}{n}\right\rfloor}\oplus g_{b_1}(D)\bullet D^{-\left\lfloor\frac{q_1}{n}\right\rfloor}\\
&g_2'(D)=g_{a_2}(D)\bullet D^{-\left\lfloor\frac{p_2}{n}\right\rfloor}\oplus g_{b_2}(D)\bullet D^{-\left\lfloor\frac{q_2}{n}\right\rfloor}\\
&\vdots\\
&g_n'(D)=g_{a_n}(D)\bullet D^{-\left\lfloor\frac{p_n}{n}\right\rfloor}\oplus g_{b_n}(D)\bullet D^{-\left\lfloor\frac{q_n}{n}\right\rfloor}
\end{aligned}\tag{19}$$

其中, $\lfloor\bullet\rfloor$ 表示向下取整运算。参数 p_i,q_i,a_i,b_i 分别表示为:

$$\begin{aligned}
&p_i=p+(i-1),1\leqslant i\leqslant n\\
&q_i=q+(i-1),1\leqslant i\leqslant n\\
&a_i=p_i \bmod n+1,1\leqslant i\leqslant n\\
&b_i=q_i \bmod n+1,1\leqslant i\leqslant n
\end{aligned}\tag{20}$$

其中, 参数 p,q 分别为原码字序列得到移位序列 IV 与移位序列 V 进行的移位长度。令 V_i 为 $\left\lfloor\frac{p_i}{n}\right\rfloor,\left\lfloor\frac{q_i}{n}\right\rfloor$ 中较大的数, 即: $V_i=\max\left(\left\lfloor\frac{p_i}{n}\right\rfloor,\left\lfloor\frac{q_i}{n}\right\rfloor\right)$, 新组合卷积码的生成多项式 $g_1'(D),g_2'(D),\cdots,g_n'(D)$ 可以变换为式(21)。

$$\begin{aligned}
&g_1'(D)=\left(g_{a_1}(D)\bullet D^{-\left\lfloor\frac{p_1}{n}\right\rfloor}\oplus g_{b_1}(D)\bullet D^{-\left\lfloor\frac{q_1}{n}\right\rfloor}\right)\bullet D^{V_1}\\
&g_2'(D)=\left(g_{a_2}(D)\bullet D^{-\left\lfloor\frac{p_2}{n}\right\rfloor}\oplus g_{b_2}(D)\bullet D^{-\left\lfloor\frac{q_2}{n}\right\rfloor}\right)\bullet D^{V_2}\\
&\vdots\\
&g_n'(D)=\left(g_{a_n}(D)\bullet D^{-\left\lfloor\frac{p_n}{n}\right\rfloor}\oplus g_{b_n}(D)\bullet D^{-\left\lfloor\frac{q_n}{n}\right\rfloor}\right)\bullet D^{V_n}
\end{aligned}\tag{21}$$

因此, 对于 $(n,1,K)$ 卷积码序列, 经过移位叠加之后得到的新序列仍然满足卷积码编码规则。同理, 多个 $(n,1,K)$ 卷积码移位序列叠加后仍然为卷积码序列。(n,k,K) 卷积编码可等效于 k 个 $(n,1,K)$ 卷积码的叠加组合, 性质 1 得证。

3.2 基于信息熵率的卷积码序列快速判断

在 3.1 节中, 利用扰码生成多项式倍式的性质, 将扰码生成多项式识别问题转化成了卷积码序列的判断问题。如何快速判断当前序列是否为卷积码序列, 目前鲜有文献进行研究, 为此我们提出了一种基于熵率的卷积码序列快速判断方法。

熵率表示随机序列的熵随着长度增加得到的增长率。随机过程 $\{X_i\}$ 的熵率 $H(\boldsymbol{\chi})$ 定义为:

$$H(\boldsymbol{\chi})=\lim_{n\to\infty}\frac{1}{n}H\left(X_1,X_2,\cdots,X_n\right)\tag{22}$$

结合熵率定义和卷积码性质, 显然卷积码字序列的熵率小于 1, 随机码熵率恒为 1。因此, 可利用观测码字序列的熵率不同, 实现卷积码字序列的判断。然而, 在实际应用中, 可观测数据是有限的, 码字熵率求解是难以实现的。因此引入一个相关的量 $H'(\boldsymbol{\chi})$, 有:

$$H'(\boldsymbol{\chi})=\lim_{n\to\infty}H\left(X_n\mid X_{n-1},X_{n-2},\cdots,X_1\right)\tag{23}$$

根据参考文献[37], 可知:

定理 1[37]: 在平稳随机过程中, 式(22)与式(23)中的极限均存在且相等, 即 $H(\boldsymbol{\chi})=H'(\boldsymbol{\chi})$; 且存在 $H\left(X_n\mid X_{n-1},X_{n-2},\cdots,X_1\right)$ 随着 n 递减且存在极限 $H'(\boldsymbol{\chi})$。

根据定理 1 的结论, 在有限项场景下, 推出 $H\left(X_n\mid X_{n-1},X_{n-2},\cdots,X_1\right)$ 与极限相关量 $H'(\boldsymbol{\chi})$ 的关系, 可得到基于经验熵率实现卷积码字序列的快速判断方法。

推论 1: 给定 $\epsilon>0$、υ 与 L 的情况下, 取码字序列长度大于 L, 码字观测长度为 i 时, 当码字序列的经验熵率 $\hat{H}(c^L)<n$ 时, 可判定该码字序列为卷积码字序列; 否则, 为随机码字序列。

其中, $\epsilon>0$ 为可接受的经验概率偏差, υ 为获取可靠经验概率的可能性, 码字观测长度 i 满足式(24), L' 为样本空间满足式(25), $L'=L-i$。(n,k,K) 为可以进行快速判断的卷积编码参数。

$$i\geqslant\frac{(K-1)k}{n-k},i\in\mathbb{N}^+\tag{24}$$

$$\begin{aligned}
\upsilon&=\Pr\left(p-\epsilon\leqslant\widehat{\Pr}(c^{i+1})\leqslant p+\epsilon\right),p=1/2^{Kk}\\
&=\int_{p-\epsilon}^{p+\epsilon}\frac{1}{\sqrt{2\pi D}}\exp\left(-\frac{(x-p)^2}{D}\right),D=\frac{p(1-p)}{L'}
\end{aligned}\tag{25}$$

推论 1 证明如下:

对于 (n,k,K) 卷积码, 令 m_t 表示在 t 时刻编码器的输入消息, 且消息 m_t 按照概率 $\Pr(m_t)=1/2^k$ 从集合 $\mathcal{M}=\{1,2,\cdots,2^k\}$ 中随机独立选取; 令 $c_t\in\mathcal{C}$ 为 t 时刻时输出的码长为 n 的码字, 其中码簿 $\mathcal{C}$ 为集合 $\{1,2,\cdots,2^n\}$, 且有 $2^n\leqslant 2^{Kk}$; 卷积码的码率为 $R=k/n$。在 t 时刻的输出码字 c_t 依赖于过去 $(K-1)$ 时刻的输出消息, 即卷积编码器满足线性方程:

$$f:(m_t,m_{t-1},...,m_{t-K+1})\in\mathcal{M}^K\to c_t\in\mathcal{C},\forall t\in\mathbb{N}^+ \quad (26)$$

如果将编码器输出作为被观测事件, 那么寄存器状态则可以视作隐藏事件, 则可以将卷积编码输出过程作为一个隐马尔科夫模型。令寄存器在 t 时刻的状态用 $\pi_t=(m_t,m_{t-1},...,m_{t-K+1})$ 来表示, 则按照一阶马尔可夫模型的性质, 任意时刻的状态只与其前一时刻状态有关, 即有:

$$\begin{aligned}\Pr(\pi_t|\pi_{t-1},\pi_{t-2},...,\pi_1)&=\Pr(\pi_t|\pi_{t-1})\\&=\Pr(m_t)\\&=\frac{1}{2^k}\end{aligned} \quad (27)$$

又由于观测码字只与当前隐藏状态有关, 且 $c_t=f(m_t,m_{t-1},...,m_{t-K+1})$, $\mathrm{Imf}(c_t)$=$\mathcal{M}^K$, 则可得:

$$\begin{aligned}\Pr(c_t|\pi_t,\ldots,\pi_1,c_{t-1},\ldots,c_1)&=\Pr(c_t|\pi_t,\ldots,\pi_1)\\&=\Pr(c_t|\pi_t)\\&=1\end{aligned} \quad (28)$$

如果考虑一段码字序列 $(c_1,\cdots,c_N)$, 序列长度为 N, 按照式(22)其熵率定义为:

$$H(\mathcal{C}^N)=\lim_{N\to\infty}\frac{1}{N}H(c_1,c_2,...,c_N) \quad (29)$$

根据隐马尔可夫模型的性质, 可以推出:

$$\begin{aligned}H(\mathcal{C}^N)&\overset{(a)}{=\!=}\lim_{N\to\infty}\frac{1}{N}H(\pi_1,\ldots,\pi_N)\\&\overset{(b)}{=\!=}\lim_{N\to\infty}\frac{1}{N}H(m_{-M+2},\ldots,m_1,\ldots,m_N)\\&\overset{(c)}{=\!=}\lim_{N\to\infty}\frac{1}{N}\sum_{j=-M+2}^{N}H(m_j)\\&=\lim_{N\to\infty}\frac{N+M-1}{N}H(\mathcal{M})\\&=\log(2^k)=k\end{aligned} \quad (30)$$

其中 (a) 成立是由于 c_t 与 π_t 的函数关系满足式(28), (b) 成立是由于重复出现的消息 m_t 不增加信息量, (c) 成立是由于信源输入满足平稳独立同分布。可以看出, 虽然输入消息 m_t 之间相互独立, 但输出码字 c_t 所属的随机过程内随机变量间具有一定的相关性, 导致其熵率 $H(\mathcal{C}^N)=H(\mathcal{M})\leqslant H(\mathcal{C})=n$。

可以将上面这个特性应用到卷积码与随机码的判定上, 如果一个长度为 $n,n\in\mathbb{N}^+$ 的码字序列熵率为 n, 码率 $R=1$, 则该序列为随机码序列, 如果熵率为 $k\leqslant n,k\in\mathbb{N}^+$, 码率 $R<1$, 则为卷积码序列。

实际情况中, 可以通过观测样本获取经验熵:

$$\hat{H}(c_t)=-\widehat{\Pr}(c_t=j)\sum_{j\in\mathcal{C}}\log\widehat{\Pr}(c_t=j),j\in\mathcal{C} \quad (31)$$

其中,

$$\widehat{\Pr}(c_t=j)=\frac{j\text{在样本中出现的个数}}{\text{样本总数}} \quad (32)$$

则根据式(29)中的定义, 一段长度为 N 的观测码字序列 $c^N,c\in\mathcal{C}$, 其经验熵率表示为:

$$\hat{H}(\mathcal{C}^N)=\lim_{N\to\infty}\frac{-\widehat{\Pr}(c^N=j)\sum_{j\in c^N}\log\widehat{\Pr}(c^N=j)}{N} \quad (33)$$

而在实际应用中, 很难通过定义来全局搜索求出经验熵。因此, 确定足够长的序列长度 N, 确定 $c^N=j$ 的遍历范围, 降低经验熵计算资源消耗是必要的。在这里引入一个熵率的相关量 $H'(\mathcal{C}^N)$, 其在平稳过程条件下与熵率是相等的。

$$H(\mathcal{C}^N)=H'(\mathcal{C}^N)=\lim_{N\to\infty}H(X_{N+1}|X_N,\ldots,X_1) \quad (34)$$

于是可以推出:

$$\begin{aligned}H'(\mathcal{C}^N)&=\lim_{N\to\infty}H(c_{N+1}|\pi_N,\ldots,\pi_1)\\&=\lim_{N\to\infty}H(c_{N+1}|m_N,\ldots,m_{-M+2})\\&=\lim_{N\to\infty}H(c_{N+1}|m_N,\ldots,m_{N-M+1})\\&=\lim_{N\to\infty}H(c_{N+1}|\pi_N)\\&\overset{(a)}{=\!=}H(c_{t+1}|\pi_t),\forall t\in\mathbb{N}^+\end{aligned} \quad (35)$$

其中 (a) 成立是根据定义, 知道条件熵与序列长度无关。同时, 根据定义容易得到任意长度为 i 的序列在给定 π_i 条件下的概率为:

$$\begin{aligned}&\Pr(c^i\mid\pi_i)=\Pr(c_i,\ldots,c_1\mid\pi_i)\\&\overset{(a)}{=\!=}\Pr(c_{i-1},\ldots,c_1\mid\pi_i)\\&\overset{(b)}{=\!=}\Pr(\pi_{i-1},\ldots,\pi_1\mid\pi_i)\\&=\Pr(m_{i-M},\ldots,m_{-M+3}\mid m_1,\ldots,m_{-K+1})\\&=\Pr(m_i,\ldots,m_2)\\&=2^{-k(i-1)}\end{aligned}\tag{36}$$

其中, (a) 与 (b) 成立是由于 c_i 由 π_i 决定。于是, 可以得到:

$$\begin{aligned}H(c^i\mid\pi_i)&=H(c^{i-1}\mid\pi_i)\\&=k(i-1)\end{aligned}\tag{37}$$

根据熵值的性质, 可以得到:

$$\begin{aligned}H(\pi_i\mid c^i)&=H(c^i,\pi_i)-H(c^i)\\&=H(c^i\mid\pi_i)+H(\pi_i)-H(c^i)\\&=k(i-1)+Kk-ni\end{aligned}\tag{38}$$

当给定卷积码编码器时, 观测任意一段长度 i 的序列, 只要满足 $H(\pi_i\mid c^i)\leqslant 0$, 说明观测长度大于 i 的序列, 可以确定当前时刻寄存器状态。即在满足条件:

$$i\geqslant\frac{(K-1)k}{n-k},i\in\mathbb{N}^+\tag{39}$$

的情况下, 有唯一对应的 π_N。于是代入到公式(33)中, 可以得到:

$$\begin{aligned}\hat{H}(\mathcal{C}^N)&=H(c_{i+1}\mid c_i,\ldots,c_1)\\&=\frac{H(c_{i+1},\ldots,c_1)}{i+1}\end{aligned}\tag{40}$$

即说明, 观测序列长度 i 满足(39)的情况下, 只用观测 $(i+1)$ 段序列就可以计算出经验熵率。因此, 如果观测方在不知道码本的情况下观测了一段长度为 L 的码字序列 c^L, 且 $L\gg i$, 则可知关于 c^{i+1} 的样本总数有 $(L-i)$ 个, 可以计算出经验概率分布为:

$$\widehat{\Pr}\left(c^{i+1}=j\right)=\frac{j\text{在样本中出现的个数}}{L-i}\tag{41}$$

将计算出的经验概率代入式(33)中, 则可得到:

$$\hat{H}(\mathcal{C}^N)=-\frac{1}{i+1}\Pr(c^{i+1}=j)\sum_{j\in\mathcal{C}^{i+1}}\log\Pr(c^{i+1}=j)\tag{42}$$

然而, 经验概率与真实概率之间存在一定偏差, 那么样本空间 $L'=L-i$ 应当在多大情况下, 才能提供一个比较可靠经验概率。已知信源按固定概率 $1/2^k$ 发送消息, 则可以确定 $\Pr(c^{i+1}=j)=\Pr(\pi_i=j)=1/(2^{Kk})=P_0, j=0,1,2,\ldots,2^{Kk}-1$ 满足均匀分布。令 $l=0,1,\ldots,L'$ 为 $c^{i+1}=j$ 在样本中可能出现的个数, 则可以推测出:

$$\Pr\left(\widehat{\Pr}(c^{i+1}=j)=\frac{l}{L'}\right)=\binom{L'}{l}\left(\frac{1-P_0}{P_0}\right)^{L'-l}\left(\frac{1}{P_0}\right)^{l}\tag{43}$$

则其期望为:

$$\mathbb{E}\left\{\Pr\left(\widehat{\Pr}(c^{i+1}=j)\right)\right\}=P_0\tag{44}$$

式(44)说明其经验概率是无偏的。其方差可以计算为:

$$\mathbb{D}\left\{\Pr\left(\widehat{\Pr}(c^{i+1}=j)\right)\right\}=\frac{P_0(1-P_0)}{L'}\tag{45}$$

式(45)说明, 如果我们想要经验概率与实际概率偏差随 L' 的增大而减小。根据二项分布性质与高斯分布关系, 在 L' 足够大时, 我们可以近似认为经验概率分布服从高斯分布 $\mathcal{N}\left(P_0,\frac{P_0(1-P_0)}{L'}\right)$, 概率密度为:

$$f(x)=\frac{1}{\sqrt{2\pi\frac{P_0(1-P_0)}{L'}}}\exp\left(-\frac{(x-P_0)^2}{\frac{2P_0(1-P_0)}{L'}}\right)\tag{46}$$

令 $\epsilon>0$ 为我们可以接受的经验概率偏差, 令 υ 为获取可靠经验概率的可能性:

$$\begin{aligned}\upsilon&=\Pr\left(P_0-\epsilon\leqslant\widehat{\Pr}(c^{i+1})\leqslant P_0+\epsilon\right)\\&=\int_{p-\epsilon}^{p+\epsilon}f(x)\widehat{\Pr}(c^{i+1})dx\end{aligned}\tag{47}$$

因此, 在给定 $\epsilon>0$ 与 υ 的情况下, 我们可以通过计算查表获取 L' 的最小长度。

3.3 卷积码参数识别与重构

在 3.1 节与 3.2 节中, 我们将扰码生成多项式识别问题转化为判断接收数据序列移位叠加运算后是否为卷积码序列的问题, 随后提出了通过数据序列的熵率来区别卷积码序列与随机序列的卷积码序列快速判断算法, 解决了“卷积编码—扰码”场景下的扰码识别问题。

在成功识别扰码以后, 对于卷积编码参数的识别, 一般解决方法为从观测数据序列中解扰出卷积编码的码字数据, 再使用传统的卷积码参数识别方法得到卷积编码的参数。这样的方法需要对观测序

列执行一次解扰操作才能得到码字数据。在 3.1 节中, 我们已经证明了当正确识别到扰码生成多项式 $f(x)$ 的倍式 $Q(x)$ 时, 接收数据序列移位叠加运算得到的纯码字序列, 也就是移位组合后的卷积码序列。且在性质 1 中, 已经证明了该移位组合后的新卷积码序列与原卷积码序列之间只有约束长度与生成多项式发生了改变, 且两者生成多项式的关系满足式(21)。因此, 可以利用识别扰码时移位组合得到的新卷积码序列得到原卷积码参数, 实现原卷积码参数的重构。

接下来我们介绍如何识别卷积码参数, 在本问题中也即识别移位组合得到的新卷积码参数。对于卷积码参数的识别, 已有大量文献进行了研究。其中, 基于 Walsh-Hadamard 的卷积码参数识别算法容忍噪声的性能要优于大多数识别方法, 因此在这里介绍基于 Walsh-Hadamard 的识别算法。

基于 Walsh-Hadamard 的分析识别方法可以看作是利用 Walsh-Hadamard 变换求解含错方程组, 可以最大化利用观测接收到的数据[38]。使用基于 Walsh-Hadamard 分析法前, 一般需要通过先验知识或者分析来估计卷积码的码率和码组起始位置。通过下面的步骤, 可以估计出要识别的卷积码的码率 $\hat{R}$ 和码组起始位置 $\hat{d}$。

首先, 将 (n,k,K) 卷积码序列数据按照 L_1 行 L_2 列建立分析矩阵 $\boldsymbol{M}$, 其中 $L_1 > L_2, L_2 > nK$。在一定取值范围内遍历列数值 L_2, 记录使得矩阵 $\boldsymbol{M}$ 的秩不等于列数值的 $L_2^{(1)}, L_2^{(2)}, \cdots$, 列数记录值的最大公约数即为码长 $\hat{n}$。在实际应用中, 卷积码的约束长度不大于 8, 码长小于 8。

其次, 对于分析矩阵 $\boldsymbol{M}$, 当列数值 L_2 是码长 $\hat{n}$ 的整数倍时, 记录矩阵在 $\hat{n}$ 种移位情况下, 单位化后左上角单位阵的维数。单位阵维数最小时对应的移位即为卷积码的码组起始位置 $\hat{d}$。

最后, 以上一步得到的码组起始位置作为数据起点, 建立卷积码序列数据的 $\hat{n}\times\hat{n}$ 阶方阵, 基于码组内部的线性相关特性, 该方阵的秩应为信息位长度 $\hat{k}$, 卷积码估计码率 $\hat{R}=\hat{k}/\hat{n}$。需要注意的是, 在含有误码的情况下, 上述估计卷积码参数的结果可能不够准确, 但仍可以采取多次采样计算, 对结果进行统计判决的方式来获取最佳估计值。

以要识别的码为 1/2 码率的 $(2,1,K)$ 卷积码为例, 利用上面得到的码组起点位置为起点, 截取得到接收序列 $\{y_{1,1}, y_{1,2}, y_{2,1}, y_{2,2}, \cdots, y_{n,1}, y_{n,2}, \cdots\}$。不妨, 令校验多项式为 $h_1(D)=h_{1,0}+h_{1,1}D+\cdots+h_{1,m}D^m$, $h_2(D)=h_{2,0}+h_{2,1}D+\cdots+h_{2,m}D^m$, $m=K-1$。根据校验矩阵的性质 $\boldsymbol{c}\bullet\boldsymbol{H}^{\mathrm{T}}=\boldsymbol{0}$, 可以建立式(21)中接收数据矩阵与校验多项式系数构成的矩阵之间的 N 个方程。其中, 矩阵 $\boldsymbol{H}=\begin{bmatrix}h_{1,m}h_{2,m} & \cdots & h_{1,0}h_{0,0}\end{bmatrix}$。

$$\begin{bmatrix} y_{1,1}y_{1,2} & y_{2,1}y_{2,2} & \cdots & y_{m+1,1}y_{m+1,2} \\ y_{2,1}y_{2,2} & y_{3,1}y_{3,2} & \cdots & y_{m+2,1}y_{m+2,2} \\ \vdots & \vdots & \ddots & \vdots \\ y_{N,1}y_{N,2} & y_{N+1,1}y_{N+1,2} & \cdots & y_{m+N,1}y_{m+N,2} \end{bmatrix} \boldsymbol{H}^{\mathrm{T}}=\boldsymbol{0} \quad (48)$$

基于 Walsh-Hadamard 分析法识别卷积码的校验多项式步骤如下:

第一步, 将接收数据矩阵中的每一个行表示为一个十进制数。例如, 将第一行的二进制数据 $y_{1,1}y_{1,2}y_{2,1}y_{2,2}\cdots y_{m+1,1}y_{m+1,2}$ 转化为十进制数 d。一共可以得到 N 个十进制数 $d_1, d_2, \cdots, d_N$。

第二步, 将得到的 N 个十进制数构造为一个新的长度为 2^L 的向量 $\boldsymbol{V}_s$, 即将十进制数对应的向量中的位置设为十进制数出现的次数, 其余设为 0。

第三步, 对向量 $\boldsymbol{V}_s$ 进行 Walsh-Hadamard 变换, 即将向量 $\boldsymbol{V}_s$ 与 $2^L\times 2^L$ 的 Hadamard 矩阵相乘, 得到结果向量 $\boldsymbol{V}_r$。设定置信度 t, 如果结果向量 $\boldsymbol{V}_r$ 中某一元素 e 大于置信度 t, 说明元素 e 位置的二进制向量就是方程组最有可能的解。

第四步, 从上一步中得到的向量包含两个校验多项式 $h_1(D), h_2(D)$, 因此需要将结果向量按奇偶进行分组, 再转换为多项式。一般得到的多项式需要进行化简, 才能得到最终的校验多项式。卷积码的生成矩阵与校验矩阵的关系式为 $G(D)\bullet\boldsymbol{H}^T(D)=0$, 其中 $G(D)=[g_1(D)g_2(D)]$, $H(D)=[h_1(D)h_2(D)]$, 根据此关系式可求出卷积码生成多项式 $g_1(D)$, $g_2(D)$。

对于估计码率 $\hat{R}$ 不等于 1/2 的 (n,k,K) 卷积码, 同样可以上面介绍的识别 $(2,1,K)$ 卷积码的方法来识别。可以将接收数据分为 $\hat{n}$ 路, 每次选择其中的两路数据构造 1/2 码率的子卷积码, 然后按照上面步骤来识别出其中两路的校验多项式, 重复识别操作直至识别所有的校验多项式。

通过基于 Walsh-Hadamard 的卷积码参数识别方法, 可以得到移位组合得到的新卷积码参数。性质 1 中说明了原卷积码与移位组合得到的新卷积码输入信息位长度、码长是相等的, 原卷积码的生成多项式

可以按照式(21)进行重构。

在式(21)中, 利用 3.3 节中基于 Walsh-Hadamard 的卷积码参数识别方法, 我们可以得到多项式组 $g_1'(D),g_2'(D),\cdots,g_n'(D)$。由于 $g_i'(D)$ 与 $g_i(D)$ 都有 K 项, 因此通过式(21)可以列出含有 nK 个未知量 $g_{i,j}$ 的 nK 个等式。求解式(21)中的未知量 $g_{i,j}$, 即可重构原卷积码的生成多项式 $g_i(D)$。

3.4 加扰卷积码盲识别算法流程

在前面三节中, 我们首先将扰码识别问题转化为卷积码序列的判别问题, 提出了卷积码字序列的快速判别方法实现了扰码多项式的识别。然后根据移位组合得到的新卷积码与原卷积码之间的关系, 求解含有生成多项式未知量的方程组重构出原卷积码的参数。

本文提出的加扰卷积码盲识别算法可以概括为: 假定观测序列为 $\{y_{et}\}$

step1 利用 $\{y_{et}\}$ 建立式(10)中移位组合矩阵 $\boldsymbol{Y}$, 记矩阵 $\boldsymbol{Y}$ 的第 i 行为 $\boldsymbol{r}_i$;

step2 For all $(i_1,i_2,\cdots,i_N)$, $0<i_1<i_2<\cdots<i_N<D$

$\boldsymbol{r}\leftarrow\boldsymbol{r}_{i_1}\oplus\boldsymbol{r}_{i_2}\oplus\cdots\oplus\boldsymbol{r}_{i_N}$

$k\leftarrow 1$

If $H(\boldsymbol{r})<H(Random\_Seq)$

If $k\leqslant 2$

存储 $Q_k(x)=1+\sum_{j=1}^{N-1}x^{i_N-i_{(N-j)}}$

存储 $\boldsymbol{r}$ 序列数据

$k\leftarrow k+1$

end

end

step3 计算 $\gcd\left(Q_1(\boldsymbol{x}),Q_2(\boldsymbol{x})\right)=f(x)$, 即 $Q_1(\boldsymbol{x})$ $Q_2(\boldsymbol{x})$ 的最大非平凡公约式为扰码生成多项式 $f(x)$。完成识别扰码生成多项式。

step4 利用存储的 $\boldsymbol{r}$ 序列数据, 使用基于 Walsh-Hadamard 的卷积码参数识别方法, 识别出移位组合卷积码的码长 n、信息位长度 k 以及生成多项式 $g_1'(D),g_2'(D),\cdots,g_n'(D)$。

step5 根据 step4 识别得到的生成多项式 $g_1'(D),g_2'(D),\cdots,g_n'(D)$ 建立式(21)的方程组, 解出未知多项式 $g_1(D),g_2(D),\cdots,g_n(D)$。完成识别卷积编码参数。

本文提出的加扰卷积码盲识别算法处理流程如图 6 所示。其中, 图 6 流程图中观测序列是由通信中接收端接收并解调后包含 0、1 的二进制序列, 接收

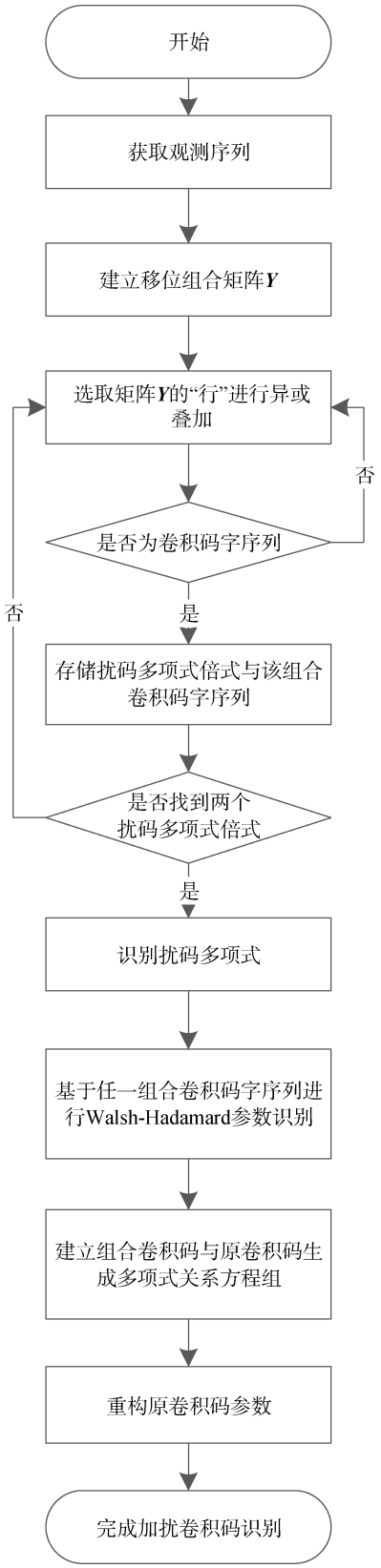

图 6 加扰卷积码盲识别算法处理流程图

Figure 6 Process flow chart of blind recognition algorithm of scrambled convolutional code

的序列由于存在噪声以及信号解调产生一定比例的误码。所有的加减运算、矩阵操作均在二元域上完成, 经过变换后的序列也均为二进制序列。

4 仿真分析

为了验证加扰卷积码盲识别算法, 本部分的仿真实验主要分为两部分: 验证基于经验熵率的卷积码字序列快速判断方法, 分析加扰卷积码盲识别算法的性能。

4.1 卷积码字序列快速判断的仿真

在 3.2 中介绍了基于经验熵率的卷积码序列快速判断方法。为了验证该判断方法, 我们分别测试不同的码字序列在无误码和有误码两种情况下的经验熵率。

首先, 我们仿真测试了卷积码字序列快速判断方法在无误码情况下的表现。取 7 种不同的数据序列, 其中包括 1 种使用 MATLAB 生成的随机序列, 3 种码率 $R=\frac{1}{2}$ 的卷积码序列、2 种码率 $R=\frac{1}{3}$ 的卷积码序列以及 1 种 $R=\frac{2}{3}$ 的卷积码序列。使用的数据长度从 10^3 个码字长度到 10^8 个码字长度, 码字观测长度根据式(39)计算得到。为了对比不同卷积码字序列的经验熵率, 我们对统计得到的经验熵率进行了归一化, 即归一化经验熵率为经验熵率与码字长度 n 的比值。

仿真结果如图 7, 可以看出随机序列的归一化经验熵率在不同数据长度下均接近于 1, 而卷积码字序列的最大归一化经验熵率均小于 1。对比码率相同而约束长度 K 不同的码字序列, 归一化经验熵率随着码字约束长度的增加而增大, 即卷积码字序列的约束长度越大, 其码字序列越接近随机特性; 对比约束长度 K 相同而码率不同的码字序列, 归一化经验熵率随着码率的增加而减小, 即卷积码字序列的码率越大, 其码字序列越不具有随机特性。通过图 7, 我们容易得出, 可以通过计算经验熵率可以判断一个序列是否为卷积码字序列。

其次, 我们仿真测试了码字序列在不同误比特率下的归一化经验熵率。这里固定序列数据长度为 10^7 个码字长度, 设置的误比特率从 0.5%到 7%。仿真结果如图 8 所示, 可以看出: 随着误比特率的增加, 卷积码序列的归一化经验熵率逐渐接近于 1, 也即误比特率的增加使得码字序列越来越接近于随机序列。同时, 可以看出误比特率在小于 6%的情况下, 使用基于经验熵率的卷积码字序列快速判断方法可以卷积码字序列和随机序列。

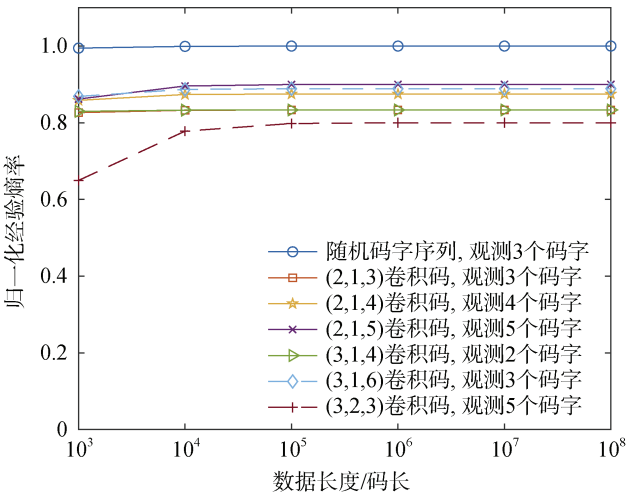

图 7 不同数据长度下码字序列的归一化经验熵率

Figure 7 Normalized empirical entropy rate of codeword sequence under different data length

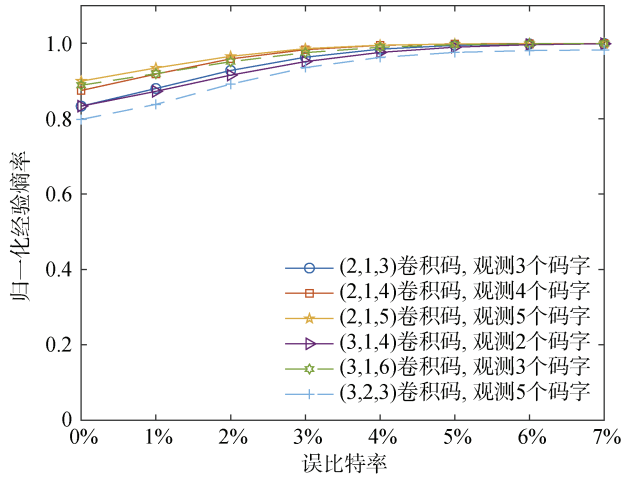

图 8 不同误比特率下码字序列的归一化经验熵率

Figure 8 Normalized empirical entropy rate of codeword sequence under different bit error rates

通过上面的仿真实验, 可以看出基于经验熵率的卷积码序列快速判断方法只需要计算数据序列的经验熵率即可区分随机序列与卷积码字序列。

4.2 加扰卷积码盲识别算法的性能

为了分析本文提出的加扰卷积码盲识别算法的识别性能, 我们设计下面三个仿真实验分别分析扰码识别正确率与加扰卷积码联合识别正确率的对比, 以及不同信道条件、使用不同数据量对识别结果的影响。

仿真实验中使用的卷积编码参数与加扰多项式如表 2 所示。

实验 1 比较在不同信道条件下, 提出的盲识别算法识别加扰多项式的正确率。为了与文献[36]的识

表 2 仿真使用的卷积编码参数与加扰多项式

Table 2 Convoluztional coding parameters and scramble polynomials used in simulation

卷积码	编码生成多项式矩阵	加扰多项式
(3,1,6)	$G_1=\begin{bmatrix}1+x^3+x^4+x^5\\1+x^2+x^4+x^5\\1+x^2+x^3+x^4+x^5\end{bmatrix}$	$f_1(x)=x^7+x^3+1$ $f_2(x)=x^{11}+x^2+1$
(2,1,4)	$G_2=\begin{bmatrix}1+x+x^3\\1+x+x^2+x^3\end{bmatrix}$	$f_1(x)=x^7+x^3+1$ $f_2(x)=x^{11}+x^2+1$

别结果进行对比，选取的卷积码为表2中(3,1,6)卷积码，编码生成多项式矩阵为 G_1，加扰多项式为 $f_1(x)=x^7+x^3+1$，$f_2(x)=x^{11}+x^2+1$，这里固定数据长度为10^7个码字长度。对于每种误比特率情况，进行300次蒙特卡罗模拟，得到的识别结果如图9所示。文献[36]使用数据量为10^6个码字长度，其在较高误比特率下扰码识别性能较差。从图9中可以看出，在误比特率为 6%时，本文算法的扰码识别正确率在85%左右；在相同的误比特率情况下，本文算法的扰码识别正确率高于文献[36]的识别正确率。文献[36]基于接收数据矩阵的零空间实现扰码多项式的识别，这种方法对噪声的容忍度较小，因此扰码识别正确率较低。而本文算法将扰码识别问题转为判断组合后的序列是否为卷积码字序列，识别正确率取决于卷积码序列判断算法是否可以将卷积码字序列与随机序列区分出来，因此噪声情况下的扰码识别正确率较文献[36]高。但是，卷积码序列判断算法依赖于大数量观测数据序列的经验熵率，因此本文使用的数据量要高于文献[36]使用的数据量。

实验 2 对比不同信道条件下，本文算法识别扰码的正确率与加扰卷积码联合识别的正确率。选取的卷积码为表 2 中编码生成多项式矩阵为 G_1 的(3,1,6)卷积码，加扰多项式为 $f_1(x)=x^7+x^3+1$，$f_2(x)=x^{11}+x^2+1$，使用的数据长度为10^7个码字长度。对于每种误比特率情况，进行300次蒙特卡罗模拟，识别结果如图10所示。从图10可以看出，相同误比特率下，扰码识别正确率与加扰卷积码联合识别的正确率相差不大，也即当正确识别出扰码后，基本可以识别出组合卷积码的参数并重构出原卷积码。基于 Walsh-Hadamard 的卷积码参数识别方法容错率较高，这符合图10中的识别结果。

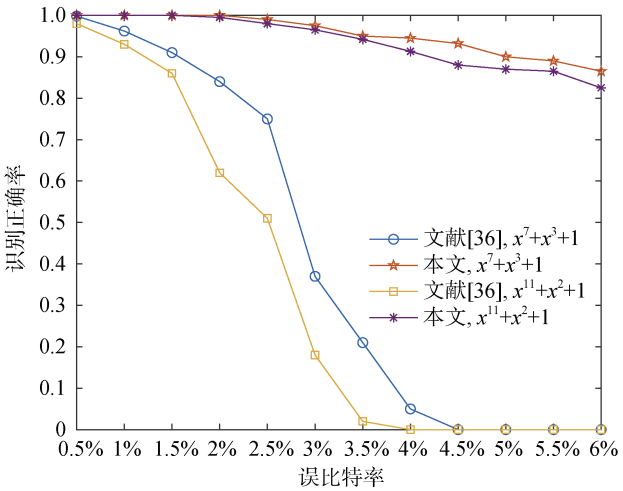

图 9 本文算法与文献[36]算法的扰码识别正确率对比

Figure 9 Comparison of the correct recognition rate of scramble code between the algorithm in this paper and the algorithm in literature [36]

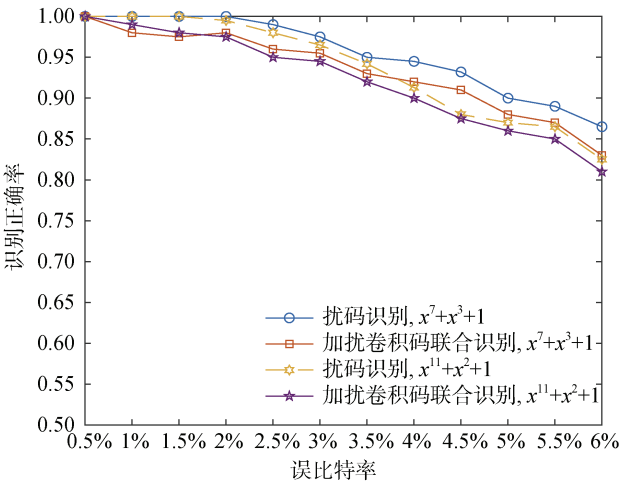

图 10 不同误比特率下扰码识别正确率与加扰卷积码联合识别正确率对比

Figure 10 Comparison of the correct recognition rate of scramble code and the joint recognition rate of scrambled convolutional codes under different bit error rates

除了本文使用的基于 Walsh-Hadamard 的卷积码参数识别方法, 还可以采用基于 BM 的快速合冲法、欧几里得识别法等卷积码参数识别算法。本文提出的加扰卷积码盲识别算法中, 识别出扰码是实现卷积码参数识别与重构的前提, 因此当使用的卷积码参数识别算法抗噪性能较高时, 可以近似认为扰码识别正确率与加扰卷积码联合识别正确率相等。

实验 3 分析使用不同长度的观测数据对扰码识别结果的影响。选取的卷积码为表 2 中编码生成多项式矩阵为 G_1 的 (3,1,6) 卷积码, 编码生成多项式矩阵为 G_2 的 (2,1,4) 卷积码, 采用的加扰多项式为 $f_1(x)=x^7+x^3+1$, $f_2(x)=x^{11}+x^2+1$, 选择没有误码的信道条件, 即信道误比特率为 0。使用的数据长度为 10^3 个码字长度到 10^7 个码字长度, 在每种数据长度情况下进行 300 次蒙特卡罗模拟, 识别结果如图 11 所示。从图 11 中可以看出, 观测数据的长度影响扰码识别的结果, 当观测数据长度大于 10^7 个码字长度后, 扰码识别正确率保持不变; 扰码识别正确率随着观测数据长度的减少而降低。图 11 同样说明了, 本文提出的加扰卷积码盲识别算法性能需要较大数量的观测数据才能有较高的识别正确率。

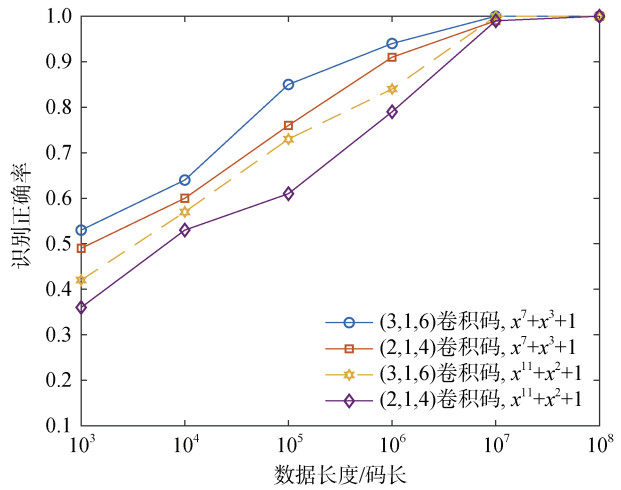

图 11 不同数据长度下扰码识别的正确率

Figure 11 Correct recognition rate of scramble code under different data lengths

从图 9～图 11 可以看出, 采用本文提出的识别算法识别加扰卷积码时, 扰码阶数的高低也影响识别结果。在相同测试条件下, 使用较低阶数扰码时的识别正确率明显高于使用较高阶数扰码的识别正确率。

5 结论

本文主要提出了一种同步扰码与卷积码联合盲识别的方法。该方法首先利用了扰码生成多项式倍式的性质, 将扰码识别问题转化为变换后的序列是否为卷积码的问题; 然后提出并证明了基于经验熵率来快速判断卷积码的方法; 最后, 推导了移位组合后的卷积码与原卷积码参数之间满足的关系。

仿真实验表明, 提出的联合盲识别算法在不使用先验信息的条件下, 可以实现对加扰卷积码进行有效识别。在信道误比特率小于 6% 时, 可以分辨出随机序列和卷积码字序列, 且能正确识别出扰码生成多项式与卷积码的编码参数。

参考文献

[1] ZHANG Y G, LOU C Y. Channel coding recognition and analysis[M].
Beijing: Publishing House of Electronics Industry, 2010.
Zhang Y G, Lou C Y. Channel coding and its identification analysis[M]. Beijing: Publishing House of Electronics Industry, 2010.
(张永光, 楼才义. 信道编码及其识别分析[M]. 北京: 电子工业出版社, 2010.)

[2] Yu P D, Li J, Peng H. A Novel Algorithm for Channel Coding Recognition Using Soft-Decision[J]. *Acta Electronica Sinica*, 2013, 41(2): 301-306.
(于沛东, 李静, 彭华. 一种利用软判决的信道编码识别新算法[J]. *电子学报*, 2013, 41(2): 301-306.)

[3] Ren Y B. *Research on recognition algorithms of channel codes under error conditions*[D]. Beijing: Tsinghua University
(任亚博. 误码条件下信道编码识别研究[D]. 北京: 清华大学)

[4] Liao H S, Yuan Y, Gan L. Novel Blind Recognition Method for Self-Synchronized Scrambler[J]. *Journal on Communications*, 2013, 34(1): 136-143.
(廖红舒, 袁叶, 甘露. 自同步扰码的盲识别方法[J]. *通信学报*, 2013, 34(1): 136-143.)

[5] Huang Z P, Zhou J, Su S J, et al. Order Estimation of Self-Synchronizing Scrambling Polynomial Based on Run Statistic[J]. *Journal of University of Electronic Science and Technology of China*, 2013, 42(4): 541-545.
(黄芝平, 周靖, 苏绍璟, 等. 基于游程统计的自同步扰码多项式阶数估计[J]. *电子科技大学学报*, 2013, 42(4): 541-545.)

[6] Zheng S Q, Janecek A, Li J Z, et al. Dynamic Search in Fireworks Algorithm[C]. *2014 IEEE Congress on Evolutionary Computation*, 2014: 3222-3229.

[7] Lu Q T, Zhang M, Li X H, et al. Self-Synchronized Scrambler Recognition Based on Code Weight Distributing Distance[J]. *Journal of Detection & Control*, 2015, 37(5): 7-13.
(吕全通, 张旻, 李歆昊, 等. 基于码重分布距离的自同步扰码识别方法[J]. *探测与控制学报*, 2015, 37(5): 7-13.)

[8] Ma Y, Zhang L M, Wang H T. Reconstructing Synchronous Scrambler with Robust Detection Capability in the Presence of Noise[J]. *IEEE Transactions on Information Forensics and Security*, 2015, 10(2): 397-408.

[9] Li X H, Zhang M, Han S N, et al. Distinction of Self-Synchronous Scrambled Linear Block Codes Based on Multi-Fractal Spec-

trum[J]. *Journal of Systems Engineering and Electronics*, 2016, 27(5): 968-978.

[10] Luo X Y, Shen L, Lu P Z, et al. Fast Blind Restore of LFSR Sequences with High Error Tolerance[J]. *Signal Processing*, 2004, 20(6): 552-558.
(罗向阳, 沈利, 陆佩忠, 等. 高容错伪随机扰码的快速盲恢复[J]. *信号处理*, 2004, 20(6): 552-558.)

[11] Wu W J, Huang Z P, Tang G L, et al. Fast Recovery of Interfered Scrambling Code Sequence[J]. *Acta Armamentarii*, 2009, 30(8): 1134-1138.
(伍文君, 黄芝平, 唐贵林, 等. 含错扰码序列的快速恢复[J]. *兵工学报*, 2009, 30(8): 1134-1138.)

[12] Shen B, Wang J X. Blind Estimation of the PN Sequence and Information Sequence of a DSSS Signal Based on SVD[J]. *Journal of Electronics & Information Technology*, 2014, 36(9): 2098-2103.
(沈斌, 王建新. 基于奇异值分解的直扩信号伪码序列及信息序列盲估计方法[J]. *电子与信息学报*, 2014, 36(9): 2098-2103.)

[13] Ma Y, Zhang L M, Wang H T. Blind Identification of Frame Synchronization in Scrambled Coding Sequence[J]. *Acta Electronica Sinica*, 2016, 44(9): 2087-2092.
(马钰, 张立民, 王好同. 编码加扰序列的帧同步盲识别[J]. *电子学报*, 2016, 44(9): 2087-2092.)

[14] Chen Z L, Peng H, Gong K X, et al. Scrambler Blind Recognition Method Based on Soft Information[J]. *Journal on Communications*, 2017, 38(3): 174-182.
(陈泽亮, 彭华, 巩克现, 等. 基于软信息的扰码盲识别方法[J]. *通信学报*, 2017, 38(3): 174-182.)

[15] Johansson T, Jönsson F. Improved Fast Correlation Attacks on Stream Ciphers via Convolutional Codes[M]. Advances in Cryptology — EUROCRYPT '99. Berlin, Heidelberg: Springer Berlin Heidelberg, 1999: 347-362.

[16] Liu J, Wang X J, Zhou X Y. Blind Recognition of Convolutional Coding Based on Walsh-Hadamard Transform[J]. *Journal of Electronics & Information Technology*, 2010, 32(4): 884-888.
(刘健, 王晓君, 周希元. 基于 Walsh-Hadamard 变换的卷积码盲识别[J]. *电子与信息学报*, 2010, 32(4): 884-888.)

[17] Xie H, Wang F H, Huang Z T, et al. A Fast Method for Blind Recognition of Convolutional Codes Based on Improved Euclidean Algorithm[J]. *Journal of National University of Defense Technology*, 2012, 34(6): 158-162.
(解辉, 王丰华, 黄知涛, 等. 基于改进欧几里得算法的卷积码快速盲识别算法[J]. *国防科技大学学报*, 2012, 34(6): 158-162.)

[18] Xie H, Wang F H, Huang Z T. A Method for Blind Recognition of Convolutional Interleaver[J]. *Journal of Electronics & Information Technology*, 2013, 35(8): 1952-1957.
(解辉, 王丰华, 黄知涛. 卷积交织器盲识别方法[J]. *电子与信息学报*, 2013, 35(8): 1952-1957.)

[19] Xie H, Wang F H, Huang Z T. Blind Recognition of(N, 1, m) Convolutional Code Based on Maximum Likelihood Detection[J]. *Journal of Electronics & Information Technology*, 2013, 35(7): 1671-1676.
(解辉, 王丰华, 黄知涛. 基于最大似然检测的(n, 1, m)卷积码盲识别方法[J]. *电子与信息学报*, 2013, 35(7): 1671-1676.)

[20] Wang F H, Hui X, Huang Z T. Blind Reconstruction of Convolutional Code Based on Segmented Walsh-Hadamard Transform[J]. *Journal of Systems Engineering and Electronics*, 2014, 25(5): 748-754.

[21] Zhang D, Zhang Y, Yang X J, et al. Blind Recognition Of(N, n-1, m) Convolutional Code Based on Genetic Algorithm[J]. *Fire Control & Command Control*, 2015, 40(9): 31-34, 44.
(张岱, 张玉, 杨晓静, 等. 基于遗传算法的(n, n-1, m)卷积码盲识别[J]. *火力与指挥控制*, 2015, 40(9): 31-34, 44.)

[22] Zhang D, Zhang Y, Yang X J, et al. An Algorithm for Convolutional Codes Recognition Based on Sectionally Extracting Soft-Decision Weighted Walsh Hadamard Transform[J]. *Acta Armamentarii*, 2015, 36(12): 2298-2305.
(张岱, 张玉, 杨晓静, 等. 基于分段抽取软判决加权 Walsh Hadamard 变换的卷积码识别算法[J]. *兵工学报*, 2015, 36(12): 2298-2305.)

[23] Huang L. *Research on blind recognition of interleaver and convolutional encoders based on algebraic theory*[D]. Hefei: University of Science and Technology of China, 2016.
(黄丽. 基于代数结构的交织器与卷积码的盲识别研究[D]. 合肥: 中国科学技术大学, 2016.)

[24] Huang L, Chen W G, Chen E H, et al. Blind Recognition of k/n Rate Convolutional Encoders from Noisy Observation[J]. *Journal of Systems Engineering and Electronics*, 2017, 28(2): 235-243.

[25] Wang W N, Peng H, Ji L. Robust Recognition of Convolutional Codes with Cepstrum and Phase Ambiguity[J]. *Journal of Signal Processing*, 2018, 34(4): 427-438.
(王伟年, 彭华, 冀磊. 倒谱与相位模糊条件下的卷积码高容错识别[J]. *信号处理*, 2018, 34(4): 427-438.)

[26] Yu P D, Peng H, Gong K X, et al. Blind Recognition of Convolutional Codes Based on Least-Square Cost-Function[J]. *Acta Electronica Sinica*, 2018, 46(7): 1545-1552.
(于沛东, 彭华, 巩克现, 等. 基于最小二乘代价函数的卷积码盲识别方法[J]. *电子学报*, 2018, 46(7): 1545-1552.)

[27] Zhang L M, Liu J, Zhong Z G. Blind Recognition of (N, 1, m) Convolutional Codes Based on Modified Walsh-Hadamard Transform[J]. *Journal of Electronics & Information Technology*, 2018, 40(4): 839-845.
(张立民, 刘杰, 钟兆根. 基于改进 Walsh-Hadamard 变换的(n, 1, m)卷积码盲识别[J]. *电子与信息学报*, 2018, 40(4): 839-845.)

[28] Chen Z L, Gong K X, Peng H, et al. Joint Blind Recognition of Packet Interleaver and Convolution Code Based on Soft Information[J]. *Acta Electronica Sinica*, 2018, 46(6): 1454-1460.
(陈泽亮, 巩克现, 彭华, 等. 基于软信息的分组交织和卷积码联合识别[J]. *电子学报*, 2018, 46(6): 1454-1460.)

[29] Cluzeau M. Reconstruction of a Linear Scrambler[J]. *IEEE Transactions on Computers*, 2007, 56(9): 1283-1291.

[30] Liu X B, Koh S N, Wu X W, et al. Investigation on Scrambler Reconstruction with Minimum a Priori Knowledge[C]. *2011 IEEE Global Telecommunications Conference - GLOBECOM*, 2012: 1-5.

[31] Liu X B, Koh S N, Wu X W, et al. Reconstructing a Linear Scrambler with Improved Detection Capability and in the Presence of Noise[J]. *IEEE Transactions on Information Forensics and Security*, 2012, 7(1): 208-218.

[32] Xie H, Wang F H, Huang Z T. Blind Reconstruction of Linear

Scrambler[J]. *Journal of Systems Engineering and Electronics*, 2014, 25(4): 560-565.

[33] Ma Y, Zhang L M. Reconstruction of Scrambler with Real-Time Test[J]. *Journal of Electronics & Information Technology*, 2016, 38(7): 1794-1799.
(马钰, 张立民. 基于实时检测的扰码重建算法[J]. *电子与信息学报*, 2016, 38(7): 1794-1799.)

[34] Liu X B, Koh S N, Chui C C, et al. A Study on Reconstruction of Linear Scrambler Using Dual Words of Channel Encoder[J]. *IEEE Transactions on Information Forensics and Security*, 2013, 8(3): 542-552.

[35] Han S N, Zhang M, Li X H. A Blind Identification Method of Self-Synchronous Scramblers Based on Optimization of Established Cost Function[J]. *Journal of Electronics & Information Technology*, 2018, 40(8): 1971-1977.
(韩树楠, 张旻, 李歆昊. 基于构造代价函数求解的自同步扰码盲识别方法[J]. *电子与信息学报*, 2018, 40(8): 1971-1977.)

[36] Han S N, Zhang M. A Method for Blind Identification of a Scrambler Based on Matrix Analysis[J]. *IEEE Communications Letters*, 2018, 22(11): 2198-2201.

[37] Cover T M, Thomas J A. Elements of Information Theory[M]. New York, USA: John Wiley & Sons, Inc., 1991.

[38] You L, Zhu Z L. The Application of Walsh Function in Resolving of F(2) Equations[J]. *Signal Processing*, 2000, 16(S1): 27-30, 20.
(游凌, 朱中梁. Walsh 函数在解二元域方程组上的应用[J]. *信号处理*, 2000, 16(S1): 27-30, 20.)

王中方 于 2013 年在北京邮电大学信号与信息处理专业获得硕士学位。现在中国科学院大学通信与信息系统专业攻读博士学位。研究领域为通信安全。研究兴趣包括: 信号盲识别、无线通信灵巧干扰。Email: wangzhongfang@iie.ac.cn

黄伟庆 现任中国科学院信息工程研究所第四研究室主任、中国计算机学会信息保密专委会秘书长。研究方向为: 无线通信安全、电磁信号处理、网络安全保密技术、物联网安全、云计算安全技术等。Email: huangweiqing@iie.ac.cn

翟留群 于 2019 年在杭州电子科技大学通信工程专业获得学士学位。现在中国科学院大学计算机技术专业攻读硕士学位。研究领域为无线通信物理层安全、信号处理。Email: zhailiuqun@iie.ac.cn

胡可可 于 2019 年在中国科学院大学在通信与信息系统专业获得工学博士学位。现为清华大学电子工程系博士后。研究领域为信息安全理论、信息处理。研究兴趣包括: 物理层安全、室内定位。Email: hukeke@mail.tsinghua.edu.cn

魏冬 于 2013 年获得北京邮电大学通信与信息系统专业博士学位。现任中国科学院信息工程研究所第四研究室副研究员。研究领域为无线通信物理层安全、调制识别、信号处理。Email: weidong@iie.ac.cn