卷积交织器盲识别方法

解 辉* 王丰华 黄知涛

(国防科技大学电子信息系统复杂电磁环境效应国家重点实验室 长沙 410073)

摘 要:交织技术通过对传输数据进行重排,克服通信过程中的突发错误,增强数据传输的可靠性,在数字通信系统中具有重要作用。该文针对非合作信号处理中卷积交织器的盲识别问题,提出一种基于交织序列一阶累积量的识别方法。该方法利用帧同步码和码字序列的一阶累积量检测数据序列中的交织周期。根据帧同步码累积量之间的分布关系实现交织宽度和交织深度的估计。仿真结果表明该方法能够在较高误码率条件下实现卷积交织器的识别和参数估计,且计算量小于已有的高斯三角化算法。

关键词:信号处理;信息截获;卷积交织器;帧同步码;一阶累积量;盲识别

 中图分类号: TN919; TN911.7
 文献标识码: A
 文章编号: 1009-5896(2013)08-1952-06

 DOI: 10.3724/SP.J.1146.2013.00287

A Method for Blind Recognition of Convolutional Interleaver

Xie Hui Wang Feng-hua Huang Zhi-tao

(State Key Laboratory of Complex Electromagnetic Environment Effects on Electronics and Information System, National University of Defense Technology, Changsha 410073, China)

Abstract: Interleaving technology plays an important role in most digital communication systems. The data stream is sent to an interleaver that rearranges the data in order to cope with burst errors and improve the reliability of data transmission. Considering the reconstruction of convolutional interleaver in a non-cooperative context, and an algorithm based on the first order cumulant of the data stream is proposed in this paper. The period of the data stream is detected according to the cumulants of the Attached Sync Markers (ASM) and code words. The depth and width of the convolutional interleaver are estimated according to the distributions of the ASMs. The simulation results show that the proposed method provides an optimal performance in the noisy environment and has less computational load than the Gaussian elimination algorithm.

Key words: Signal processing; Communication interception; Convolutional interleaver; Attached Sync Marker (ASM); First order cumulant; Blind recognition

1 引言

在现代数字通信中,信道中除了随机干扰,外 来的脉冲干扰、同频道干扰、相位噪声、多径和衰 落信道等都会引起突发错误¹¹。用纠正随机独立差错 的纠错编码技术对这些突发错误的纠正效果并不明 显,为了使纠错编码能够抵抗这些突发错误,在实 际应用中纠错编码往往与数据交织技术相结合。在 发送端对数据进行交织,在接收端进行解交织,使 得信道中的突发错误分散开来,变成独立的随机差 错,从而可以充分发挥纠错编码的作用^[1]。

在非合作信号处理中,利用截获到的通信信号 恢复出对方通信的原始信息,则必须实现对交织器

2013-03-07 收到, 2013-06-13 改回

和纠错编码的识别。纠错编码的盲识别技术已得到 国内外学者的广泛关注^[2-13],但在已有的纠错编码 识别算法中,并没有考虑到交织器的存在,而是直 接对码字进行分析,通常情况下,解交织是在纠错 译码之前,因此在实际中需要将接收数据进行解交 织及去除帧同步码等步骤,才能将完整的码字序列 重构出来,进而实现纠错编码的识别和译码。

在实际应用中,常用的交织器包括分组交织器 和卷积交织器^[14]。从目前公开发表的文献上来看, 主要针对分组交织器进行识别,包括基于矩阵秩准 则^[15]和有限域高斯约旦三角化^[16,17]的识别算法,算 法在无误码情况下可以实现分组交织参数的估计。 但在含误码的情况下,矩阵求秩和高斯三角化的性 能都迅速下降。对于卷积交织器的识别,主要有基 于四重循环搜索^[18]和基于矩阵秩的缺失性的识别方 法^[19,20]。基于四重循环搜索的方法假设交织参数构

国家自然科学基金(601072102)和教育部新世纪优秀人才支持计划 资助课题

^{*}通信作者: 解辉 xiehui2005@gmail.com

造交织器,并对交织序列进行解交织,将解交织后 的数据构造成矩阵形式,从而利用有限域上的高斯 约旦三角化的方法对矩阵进行变换,得到一个下三 角矩阵,根据其列向量的重量判断交织参数是否正 确,算法的不足之处在于需要经过四重循环搜索, 循环次数多,计算量较大。文献[19,20]对数据直接 进行检测,利用数据矩阵秩的缺失性进行交织参数 的识别,相比文献[18],减少了搜索次数,提高了算 法效率。但上述两种算法对于交织序列误码率的要 求较高,一旦交织序列中存在较多的误码,则算法 难以适应。

由此可见,现有的卷积交织器的识别方法存在 计算量较大、误码适应能力较低等不足。在实际通 信系统中,帧同步码序列同样进行卷积交织,例如 ATSC(Advanced Television Systems Committee), DVB-S,DVB-T等通信标准都对此进行了规定^[21]。 为此,本文提出了一种基于交织序列一阶累积量的 卷积交织器盲识别方法,算法将交织序列划分为矩 阵形式,利用矩阵列累积量的分布规律实现交织周 期的判定。同时根据帧同步码累积量之间的分布距 离和差分关系实现交织宽度和交织深度的估计。仿 真实验表明,相比于已有算法,本文算法计算量较 小,且具有较强的误码适应能力。

2 卷积交织器盲识别的数学模型

卷积交织器不按组的固定格式,是一种连续工 作的交织器。一个(*B*, *N*)卷积交织器的工作方式如 图 1 所示。其中 N = BM, *B* 为交织深度, *M* 为交 织宽度。原始信息序列经过分组编码器后得到(*n*, *k*) 线性分组码字, *m* 个线性分组码组成一帧数据,并 添加帧同步码以便接收时进行帧同步, 假设帧同步 码长度为 *s*,则一帧数据长度为mn + s。然后帧数 据进入卷积交织器,在切换开关的作用下帧数据依 次进入 *B* 条支路,周而复始。每个支路的延迟缓存 器依次以 *M* 的倍数增加,输出端用同步的切换开关 从 *B* 条支路轮流取出符号。经过信道传输后,在接 收端则按相反的缓存器排列方式进行解交织,去除 帧同步码,译码最终得到原始信息序列。在实际应 用中,卷积交织器的交织深度和交织宽度的乘积一 般满足 BM = n。

卷积交织器具有如下特性:

(1)在交织器输入处,任何间隔小于 N 个符号的 两个符号,其交织器输出处的最小间隔为 B 个符号;

(2)任何长度 *b* ≤ *B* 的突发错误,在去交织后成为至少间隔为 *N* 个符号的单个错误;



图1卷积交织器结构及工作原理

(3)间隔为 N+1 的周期性单个错误,在去交织 后成为长度为 B 的突发错误;

(4) 在不计入信道时延时,两端的时延为 N(B-1)个符号,而交织器与去交织器的存储器要 求皆为N(B-1)/2个符号。

卷积交织器在不考虑信道时延的情况下,它本 身所产生的设备时延比分组交织器时延小一半,相 应的设备存储量也少一半,因此更广泛应用于实际 数字通信系统中。卷积交织器的盲识别主要是估计 卷积交织器的交织深度及交织宽度,以便构造一个 卷积交织器对数据进行解交织。

3 基于交织序列一阶累积量的卷积交织器 盲识别

首先,将接收到的交织序列 *C*转换成±1形式, V = 2C - 1,然后将 *V*排列成一个 *Q*行 *P*列的矩 阵形式,同时设定算法的初始搜索周期 *P*_{min}和终止 搜索周期 *P*_{max},*P*的取值遍历区间[*P*_{min} *P*_{max}],假设 序列长度为 *L*,则有 *Q* = [*L*/*P*],其中[•]表示向下 取整。在每次 *P*取值后,得到一个矩阵 *H*,计算累 积向量

$$\boldsymbol{S}_{i} = \left| \sum_{j=1}^{Q} \boldsymbol{H}(j, i) \right| / 2, \quad i = 1, 2, \cdots, P$$
(1)

因为帧数据的长度和交织宽度是固定的,所以 交织序列的周期也是固定的,其长度为*B*(*mn*+*s*)。 当 *P* 的取值为*B*(*mn*+*s*)时,帧同步码在交织序列 中的位置也是固定的。因此在矩阵 **H**中,帧同步码 和码字序列分别位于不同的列当中,不失一般性, 假设传输信道为一个无记忆的二元对称信道 (BSC),转移概率记为ε。所以对应帧同步码的列的 累积量服从一个二项分布:

$$p(x) = \begin{pmatrix} Q \\ x - Q/2 \cdot (1 - 4\varepsilon) \end{pmatrix}$$
$$\cdot \varepsilon^{x - Q/2 \cdot (1 - 4\varepsilon)} (1 - \varepsilon)^{Q/2 \cdot (3 - 4\varepsilon) - x}$$
(2)

当累积样本较多时,该二项分布近似一个正态 分布:

$$p(x) = \frac{1}{2\sqrt{\pi Q\varepsilon}} \exp\left(-\frac{(x - Q(1 - 2\varepsilon))^2}{8Q\varepsilon}\right), \ x \in \mathbb{N}$$
 (3)

而其它码字序列对应列的统计量近似服从一个 正态分布:

$$p(x) = \frac{1}{\sqrt{2\pi Q}} \exp\left(-\frac{x^2}{2Q}\right) , \quad x \in \mathbb{N}$$
(4)

若 P 不是交织周期,则帧同步码和码字序列在 H 中的分布不是固定的,因此无法得到帧同步码所 在列的累积量,H 中的所有列的累积量服从式(4)的 正态分布。

当截获数据较少或数据中误码率很大时,会造成两种分布相交,因此需要保证足够的累积数量, 假设帧同步码累积的值为 *a*,根据 4 倍均方根差准 则, $a < Q(1-2\varepsilon) - 8\sqrt{Q\varepsilon}$ 时的概率小于 3×10^{-5} ;假 设编码序列累积的值为 *b*,同样根据 4 倍均方根差 准则, $b > 4\sqrt{Q}$ 时的概率小于 3×10^{-5} ,要使每次两 种序列的累积值能够相互分离,则需要

$$Q(1-2\varepsilon) - 8\sqrt{Q\varepsilon} > 4\sqrt{Q} \tag{5}$$

因此可得到累积次数与数据中误码率的对应关 系,当误码率为0时,最少的累积次数需要达到16 次。表1给出了不同误码率时对应的累积次数。

表 1 不同误码率对应所需的累积次数

误码率	0	0.005	0.010	0.050	0.100
累积次数	16	21	24	42	67

假设 $\varepsilon = 0.02$, Q = 100,则这两种分布如图 2 所示。从图中可以看出,两种不同序列的累积量分 布完全分离,可以得到帧同步码累积量在矩阵中列 的位置,计算其差分如式(6):

$$\begin{aligned} & \boldsymbol{W}_{1} = \boldsymbol{S}_{1} \\ & \boldsymbol{W}_{i} = \left| \boldsymbol{S}_{i} - \boldsymbol{S}_{i-1} \right|, \quad i = 2, \cdots, P \end{aligned}$$
 (6)

如果数据中没有经过交织,或者交织器为分组 交织器,则帧同步码在矩阵 **H**中的位置是连续的, 其统计差分量 **W**中只有两个峰值,分别为帧同步码 的起始位置和终止位置,交织周期为一个传输帧长 度,其值为P = mn + s;如果交织器为卷积交织器, 则帧同步码在 **H**中的位置是离散的,**W**中有2*Bs*个 峰值,交织周期的长度为B(mn + s)。



图 2 帧同步码和码字序列累积量的概率分布

如果交织器为卷积交织器,则记录 S_i 中所有峰值的位置,记为 p_i , $i = 1, 2, \dots, Bs$,计算 p_i 的差分:

 q_i 的大小表示 S中各峰值之间的距离。通常情况下, 传输帧的长度远大于帧同步码的长度,否则会降低 数据的传输效率。因此每个帧同步码的最后一个比 特与下一个帧同步码的第 1 个比特距离远大于同一 个帧同步码内两个相邻比特之间的距离。在一个交 织周期内共有 B个帧同步码,所以在 q中有 B 个较 大的值,通过 q中的峰值个数来确定交织深度 B, 则一个传输帧的长度为 P/B,同时两个相邻的高峰 之间的距离表示一个帧同步码的长度 s,则一帧内的 编码长度为 P/B-s。

q中的峰值大小表示一个帧同步码最后一个比 特与下个帧同步码第1个比特之间的距离,帧同步 码在矩阵 **H**中的分布如图3所示。假设峰值大小为 v,一个帧同步码在移位寄存器中被分割成几块,两 个帧同步码间的数据长度为P/B-s,则有

v+1 = P/B - s - B(B-1)M + B (8) 而当一个帧同步码被交织周期分割为首尾两部分 时,则有

$$v + 1 = P / B - s - B(B - 1)M$$
(9)

综上所述, 交织宽度 M 为

$$M = \left[\frac{P/B - s - v - 1 + B}{B(B - 1)} \right]$$
(10)



图 3 帧同步码在矩阵 H 中的分布

本文算法可归纳总结如下:

输入:接收长度为 L 的交织序列 C;

输出: 传输帧长度, 卷积交织器的交织深度 *B* 和交织宽度 *M*。

算法步骤:

步骤 1 设定初始交织周期 P_{\min} 和终止交织周期 P_{\max} , 令V = 2C - 1, $P = P_{\min}$;

步骤 2 将序列 V 排列成矩阵 H, 其列数为 P, 行数为|L/P|。根据式(1)和式(6)分别计算 S 和 W;

步骤 3 在 W 中检测峰值,如果 W 中没有峰 值,且 P < P_{max},令 P = P + 1,并返回步骤 2;如 果 W 中有两个峰值,表明序列中没有交织器或者为 分组交织器,此时的 P 即为传输帧长度;如果含有 多个峰值,则交织器为卷积交织器,并转入步骤 4;

步骤 4 记录 W中的峰值位置,记为 p,根据 式(7)计算其差分 q;

步骤 5 在 q 中检测峰值个数,其值为交织深度 B,帧同步码的长度 s 为相邻两个峰值之间的距离,根据式(10)计算交织宽度,此时传输帧长度为 P/B;

步骤6 输出传输帧长度, B和 M。

4 算法仿真

对本文中的算法进行仿真分析,编码序列为(63, 45)BCH 码,帧同步码用 16 进制表示为 1ACFFC 1D。每个传输帧内包含 10 个码字,信道转移概率 为 0.02。卷积交织器的交织深度为 7,交织宽度为 9, 交织后截取 500 k 数据进行分析。

设定初始交织周期为 1000,终止交织周期为 10000,对 *P*进行遍历,当 *P*=4634 时,*H*矩阵的 一阶累积量如图 4 所示。

从图 4 中可以看出,帧同步码的累积量远大于 其它码字序列的累积量,记录所有峰值的位置,根 据式(7)计算峰值位置的差分 q,如图 5 所示。

从图 5 中可以看出, q 中一共有 7 个峰值, 所

以卷积交织器的交织深度为 7。相邻峰值的距离为 32, 所以帧同步码的长度为 32 位。峰值大小为 258 和 265, 所以卷积交织器的交织宽度为

$$M = \left[\frac{4634/7 - 32 - 258 - 1 + 7}{7(7 - 1)}\right] = 9 \tag{11}$$

仿真结果表明本文算法能识别出交织序列中采 用的卷积交织器,并估计其参数。

5 算法性能分析

(1)算法误码适应能力分析 在噪声环境下,误 码会极大地影响高斯约旦三角化算法的性能,导致 截获矩阵为一个满秩矩阵,使得算法不能估计出交 织器的各参数。但本文算法利用交织序列的一阶累 积量,使得误码对编码序列的累积量不产生影响, 但对于帧同步码的累积量有一定影响,使得帧同步 码的统计量呈现出一个二项分布。只要数据长度足 够,仍然可以将编码序列累积量和帧同步码累积量 的两种分布进行分离,因此利用本文算法可以实现 误码环境下卷积交织器的识别和参数估计。

将信道的转移概率分别设为 0 到 0.5, 交织数据 采用上节设定的参数,分别利用高斯约旦三角化算 法和本文算法进行蒙特卡洛仿真。两种算法的误码 适应能力如图 6 所示,从图 6 中可以看出,高斯约 旦三角化算法随着信道转移概率的增加,算法性能 急剧下降,当误码超过 0.02 时,算法已经无法正确 估计交织器参数。但本文算法在转移概率为 0.15 时, 仍能 100%地估计出卷积交织器的交织深度和交织 宽度等参数。

(2)算法计算量分析 本文算法计算量集中在 交织序列的一阶累积量上,假设交织序列长度为 *L*, 设定的初始交织周期为 P_{\min} ,正确的交织周期为 *P*。 本文算法需要计算 $P - P_{\min}$ 步,每一步的计算量为 $([L/P_t] - 1)P_t$,其中 P_t 为在每一步中截获矩阵的列 数。则总共的计算量为



$$\begin{pmatrix} \left| L/P_{t} \right| - 1 \end{pmatrix} P_{t} \left(P - P_{\min} \right) \\ \approx \left(L - P_{t}/2 - P_{t} \right) \left(P - P_{\min} \right) \\ = \left(P - P_{\min} \right) L - \frac{3}{4} \left(P^{2} - P_{\min}^{2} \right)$$
(12)

高斯约旦三角化算法同样需要计算 $P - P_{\min}$ 步,但在每一步的计算量为 $([L/P_t]-1)P_t^2$,且 $[L/P_t]$ 必须远大于 P_t ,所以高斯约旦三角化算法需要更多的截获数据。

截获数据与仿真实验中相同,但传输帧中包含的码字个数不同,接收数据长度设为1 Mbit,则两种算法的计算量如图 7 所示。从图中可以看出,本 文算法的计算量远小于高斯约旦三角化算法,且算 法的计算量随着传输帧长度的增加而增加。两种算 法的计算量都与初始周期的设定有关,对于已知某 些先验信息的数据,可将搜索周期设置在小范围内; 而对于未知数据,则需要设置较小的初始交织周期, 则算法的计算量增大,导致算法效率降低,这也是 目前已有算法中难以克服的一个问题。



6 结论

本文针对卷积交织器的检测识别问题,提出了 一种基于交织序列一阶累积量的识别方法。算法利 用帧同步码和码字序列累积量在交织周期情况下不 同的分布概率,对交织序列的交织周期进行估计。 同时利用帧同步码在交织周期下累积量的分布规 律,实现卷积交织器参数的估计。仿真结果表明该 算法可行,易于工程实现,且误码适应能力和计算 量都明显优于已有的高斯约旦三角化算法,在智能 通信、信息截获等领域具有重要的应用价值和意义。

参考文献

- Cao S, Chen J, Damask J N, et al.. Interleaver technology: comparisons and applications requirements [C]. Optical Fiber Communication Conference, Atlanta, USA, 2003: 1–9.
- [2] Bringer J and Chabanne H. Code reverse engineering

problem for identification codes[J]. *IEEE Transactions on Information Theory*, 2012, 58(4): 2406–2412.

- [3] Moosavi R and Larsson E G. A fast scheme for blind identification of channel codes[C]. Global Telecommunications Conference, Linkoping, Sweden, 2011: 1–5.
- [4] Barbier J, Sicot G, and Houcke S. Algebraic approach for the reconstruction of linear and convolutional error correcting codes[J]. International Journal of Applied Mathematics and Computer Science, 2006, 2(3): 113–118.
- [5] Marazin M, Gautier R, and Burel G. Algebraic method for blind recovery of punctured convolutional encoders from an erroneous bitstream[J]. *IET Signal Processing*, 2012, 6(2): 122–131.
- [6] Marazin M, Gautier R, and Burel G. Some interesting dual-code properties of convolutional encoder for standards self-recognition[J]. *IET Communications*, 2012, 6(8): 931–935.
- [7] Xie H, Chai X M, Wang F H, et al. A method for blind identification of rate 1/2 convolutional code based on improved Euclidean algorithm[C] International Conference on Signal Processing Proceedings, Beijing, China, 2012: 1307–1310.
- [8] Marazin M, Gautier R, and Burel G. Blind recovery of k/n rate convolutional encoders in a noisy environment[J]. EURASIP Journal on Wireless Communications and Networking, 2011, DOI: 10. 1186/1687-1499-2011-168.
- [9] Wang J F, Yue Y, and Yao J. Statistical recognition method of binary BCH code[J]. Communications and Network, 2011, 31(3): 17–22.
- [10] Naseri A, Azmoon O, and Fazeli S. Blind recognition algorithm of Turbo codes for communication intelligence systems[J]. International Journal of Computer Science Issues, 2011, 8(6): 68–72.
- [11] Lehmann F, Kazem A, and Salut G. Blind turbo-detection in the presence of phase noise[J]. *IET Communications*, 2009, 3(8): 1343–1353.
- [12] Cluzeau M, Finiasz M, and Tillich J P. Methods for the reconstruction of parallel Turbo codes[C]. International Symposium on Information Theory, Austin, Texas, USA, June, 2010: 2008–2012.
- [13] Liu X B, Koh S N, and Wu X W. 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.
- [14] 徐衡平,赵勋,潘晟,等.确定性交织器的研究[J].通信学报, 2002, 23(3): 96-99.

Xu H P, Zhao X, Pan S, et al. Deterministic interleaver designs for turbo codes[J]. Journal of China Institute of Communications, 2002, 23(3): 96–99.

- [15] Sicot G and Houcke S. Blind detection of interleaver Parameters[J]. Signal Processing, 2009, 89(4): 450–462.
- [16] Sicot G and Houcke S. Blind detection of interleaver parameters[C]. IEEE International Conference on Acoustics, Speech and Signal Processing-Proceedings, Philadelphia, PA, USA, Mar. 2005: 829–832.
- [17] Lu L, Li K H, and Guan Y L. Blind detection of interleaver parameters for non-binary coded data streams[C] IEEE International Conference on Computing, Dresden, 2009: 1–4.
- [18] 甘露,刘宗辉,廖红舒,等. 卷积交织参数的盲估计[J]. 电子 学报, 2011, 39(9): 2173-2177.
 Gan L, Liu Z H, Liao H S, *et al.*. Blind estimation of the parameters of convolutional interleaver[J]. *Acta Electronica Sinica*, 2011, 39(9): 2173-2177.
- [19] Lu L, Li K H, and Guan Y L. Blind identification of convolutional interleaver parameters[C]. International Conference on Information and Communications Security,

Beijing, China, 2009: 1–5.

- [20] Gan L, Li D, Liu Z H, et al. A low complexity algorithm of blind estimation of convolutional interleaver parameters [J]. Science in China Series. F: Information Sciences, 2012, 55(10): 1–9.
- [21] 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,2010:7-8.
 Zhang Y G and Lou C Y. Channel Coding Recognition and Analysis [M]. Beijing: Publishing House of Electronics Industry, 2010: 7-8.
- 解 辉: 男,1983年生,博士生,研究方向为信道编码识别及通 信协议分析技术等.
- 王丰华: 男,1981年生,博士,讲师,研究方向为通信信号侦察 处理和空间信息对抗技术等.
- 黄知涛: 男,1976年生,教授,博士生导师,研究方向为航天侦察与对抗、综合电子战技术等.