

文章编号: 2095-4980(2013)05-0792-05

基于码重分布概率方差的循环码识别方法

郑瑞瑞¹, 汪立新^{1,2}

(1. 中国电子科技集团公司 第36研究所, 浙江 嘉兴 314001; 2. 杭州电子科技大学 通信工程学院, 浙江 杭州 310018)

摘要: 提出循环码的一种盲识别方法。在码字起点已知的条件下, 分析循环码码重分布特点, 基于码重分布概率方差识别码长; 根据码重分布统计结果选择最佳的码字, 求其循环移位后码多项式的公因式, 最后通过判决关系识别生成多项式完成循环码的盲识别。仿真实验表明该方法对误码率在 10^{-2} 时的循环码具有较好的识别效果。

关键词: 循环码; 盲识别; 码重; 方差; 生成多项式

中图分类号: TN911

文献标识码: A

doi: 10.11805/TKYDA201305.0792

Recognition method of cyclic codes based on code weight distribution probability variance

ZHENG Rui-rui¹, WANG Li-xin^{1,2}

(1. The thirty-sixth Research Institute of China Electronic Technology Group Corporation, Jiaxing Zhejiang 314001, China;
2. Department of Communication Engineering, Hangzhou Dianzi University, Hangzhou Zhejiang 310018, China)

Abstract: A blind recognition method of cyclic codes is proposed. Under the condition of having known the starting point of the code words, code weight distribution of cyclic codes is analyzed, and code length is recognized based on code weight distribution probability variance. According to the statistical results of code weight distribution, the appropriate code words are chosen, then the common factor of code polynomial after cyclic shift is obtained. Finally, the generated polynomial is recognized by setting a rational judgment relation, thus the blind recognition of cyclic codes is completed. The simulation experiments show that the method has better recognition effect for cyclic codes with BER of 10^{-2} .

Key words: cyclic codes; blind recognition; code weight; variance; generated polynomial

循环码是线性分组码中一类重要的子码, 严格精细的代数结构使其具有编译码电路简单、检错纠错能力强等特点, 循环码也是目前研究最深入, 理论最成熟, 应用最广泛的一类线性分组码^[1]。在协作通信、智能通信、通信对抗等领域, 信道编码盲识别技术至关重要^[2], 在捕获到有关循环码的传输数据后, 要实现对信息的准确提取, 必须对循环码编码方式和编码参数进行盲识别。对循环码的盲识别已有一些文献, 如基于码重分布的系统循环码识别方法^[3]、循环码的盲识别方法^[4]、采用秩统计和码根特征的二进制循环码盲识别方法^[5]、二进制BCH(Bose, Ray-Chaudhuri, Hocquenghem)码的一种盲识别方法^[6]、基于码根信息差熵和码根统计的BCH码识别方法^[7]等。文献[3]对循环码进行了全盲识别, 但仅适用于系统循环码; 文献[4]对较高误码条件下的识别未提出可行的解决方案; 文献[5]要求在误码率 10^{-2} 以下情况下才能进行有效识别。文献[6]和文献[7]虽具有较好的容错性能, 但仅用来识别循环码中的BCH码。本文针对循环码在码字起点已知的前提下, 提出一种基于码重分布概率方差估计码长的方法, 并在此基础上识别生成多项式。仿真结果表明了该方法具有良好的识别效果和较优的容错性能。

1 循环码多项式及性质

1.1 循环码的多项式表达

设 (n, k) 循环码的一个码字 $C = (c_{n-1}c_{n-2} \dots c_1c_0)$, 其对应的多项式表示如下:

$$C(x) = c_{n-1}x^{n-1} + c_{n-2}x^{n-2} + \dots + c_1x + c_0 \quad (1)$$

式中: $C(x)$ 为该循环码码字 C 的码多项式; x 仅为码元位置的标记; c_i 为码多项式系数 ($i = n-1, n-2, \dots, 1, 0$), $c_i \in [0, 1]$ 。由循环码的特征知, 码字 $C_1 = (c_{n-1}c_{n-2}\dots c_1c_0)$, 左移 i 位后的码字 $C_i = (c_{n-i-1}\dots c_0, c_{n-1}\dots c_{n-i})$ 仍为许用码字, 对应的码多项式如下:

$$C_i(x) = c_{n-i-1}x^{n-1} + \dots + c_0x^i + c_{n-1}x^{i-1} + \dots + c_{n-i} \quad (2)$$

1.2 循环码的生成多项式

定理 1^[8] 在一个 (n, k) 循环码中, 存在有且仅有 1 个 $n-k$ 次码多项式 $g(x)$, 称之为循环码的生成多项式。令

$$g(x) = x^{n-k} + g_{n-k-1}x^{n-k-1} + \dots + g_2x^2 + g_1x + 1 \quad (3)$$

式中: 常数项 g_0 等于 1; 最高次数项 x^{n-k} 的系数 g_{n-k} 为 1, 其中 $g_{n-k-1}, g_{n-k-2}, \dots, g_2, g_1 \in [0, 1]$ 。

定理 2^[8] 一个 (n, k) 循环码中所有码多项式都是式(3)中 $g(x)$ 的倍式, 即 (n, k) 循环码所有码字可由 $g(x)$ 生成, 且所有以 $g(x)$ 为倍式的小于等于 $n-1$ 次多项式必定都是码字。令信息位码字多项式:

$$m(x) = m_{k-1}x^{k-1} + m_{k-2}x^{k-2} + \dots + m_1x + m_0 \quad (4)$$

它是小于或等于 $k-1$ 次多项式。因此可得码字为:

$$C(x) = m(x) \cdot g(x) \quad (5)$$

2 码长识别

文献[9]详细阐述了线性分组码码重分布特点, 而循环码是特殊的线性分组码, 其码字内具有更严格的线性约束关系, 因此对于 (n, k) 循环码而言, 若估计码长不等于真实码长时, 可认为码元“1”和“0”在所有码字中出现的可能性相等, 此时码重分布可看作等概率分布, 即不同码重的码字出现的概率为 $1/(n+1)$ 。

定义码重分布概率方差 $D(p)$, 即:

$$D(p) = \frac{1}{N} \{E(p^2) - [E(p)]^2\} = \frac{1}{N} \left\{ \frac{1}{n+1} \sum_{i=0}^n p_i^2 - \left(\frac{1}{n+1} \sum_{i=0}^n p_i \right)^2 \right\} = \frac{1}{N} \left\{ \frac{1}{n+1} \sum_{i=0}^n p_i^2 - \frac{1}{(n+1)^2} \right\} \quad (6)$$

式中: $p = (p_0, p_1, \dots, p_n)$; p_i 为实际码重为 i 的码组出现的概率, $i = 0, 1, \dots, n$; N 是估计码长为 n 时的码组个数。由方差的含义知^[7], 变量相等时方差为零, 即变量未偏离其平均值。因此, 当估计码长 n 不是真实码长时, 码重分布趋于等概率分布, 此时码重分布概率方差最小; 反之, 当估计码长 n 为真实码长时, 码重分布概率方差最大。

假定在实际通信中, 恢复后的循环码序列通过先验知识获取了其起始位置, 即码字起点。码长识别过程如下:

- 1) 以码长估计值 n 为单位, 将获取的长度为 L 的码元序列分成 N 个码字;
- 2) 统计 N 个码字中码重 i 出现的次数及码重 i 的值;
- 3) 根据式(6)计算不同码长估计值 n 对应的码重分布概率方差 $D(p)$;
- 4) 判定 $D(p)$ 最大时对应的码长估计值 n 为真实码长。

3 生成多项式识别

识别出码长 n 后, 为了最大程度排除误码影响, 选择码重 i 出现次数最多的码字, 同时选择重复出现次数最高的码字。由式(2)知, 码字 $C_1 = (c_{n-1}c_{n-2}\dots c_1c_0)$ 及 C_1 左移 i ($i=1, 2, \dots, n-1$) 位后的码多项式, 可以表示成信息多项式与生成多项式的乘积:

$$C_i(x) = m_i(x) \cdot g(x) \quad (7)$$

生成多项式 $g(x)$ 一定是码多项式 $C_1(x), C_2(x), \dots, C_n(x)$ 的公因式, 表示成 $d(x) = \gcd[C_1(x), C_2(x), \dots, C_n(x)]$ ^[10]。考虑若所选码字存在误码, 首先判断 $d(x)$ 是否满足 $g(x)$ 的一般关系式, 若 $d(x)$ 等于 $g(x)$, 则 $d(x)$ 一定能被 N 个码字中无误码多项式整除, 即:

$$C(x) \equiv 0 \pmod{d(x)} \quad (8)$$

并且能被 $d(x)$ 整除的码多项式占 N 个测试码多项式的概率最大。假设在 N 个码字中, 每个错误码元随机分布在不

同码字中, 计算错误码字占总码字 N 的概率 a , 当关系式(7)成立的概率大于 $1-a$ 时, 判定此时求得的 $d(x)$ 即为生成多项式 $g(x)$ 。识别流程如图1所示。 $(t$ 为 $d(x)$ 的末项, M 为 N 个码多项式中能被 $d(x)$ 整除的个数)。

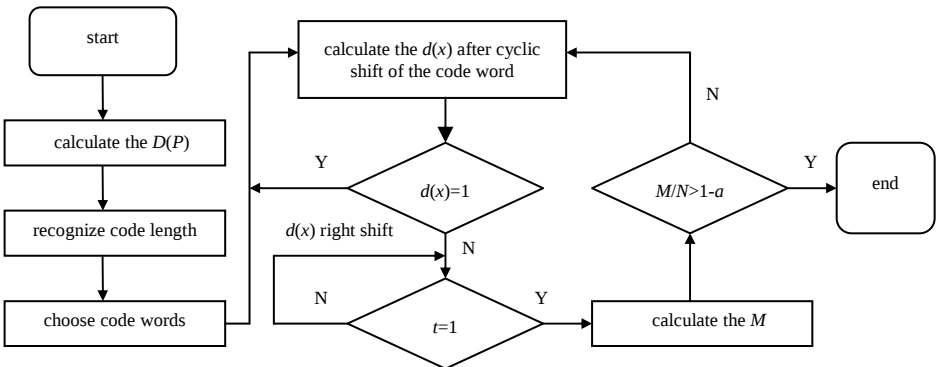


Fig.1 Flow chart of recognition of cyclic code
图1 循环码识别流程图

4 实验仿真

4.1 码长识别性能仿真

本文选用(15,5)、(31,16)循环码来验证提出的识别方案, 并分别对无误码和一定误码的情况进行了性能仿真。参数设定见表 1。

表 1 码长识别仿真中的参数设定
Table1 Parameter settings of the simulation of code length recognition

cyclic code	numbers of code element	numbers of code group	numbers of error code	bit error ratio
(15,5)	15 405	1 027	786	5.1×10^{-2}
(31,16)	15 965	591	421	3.7×10^{-2}

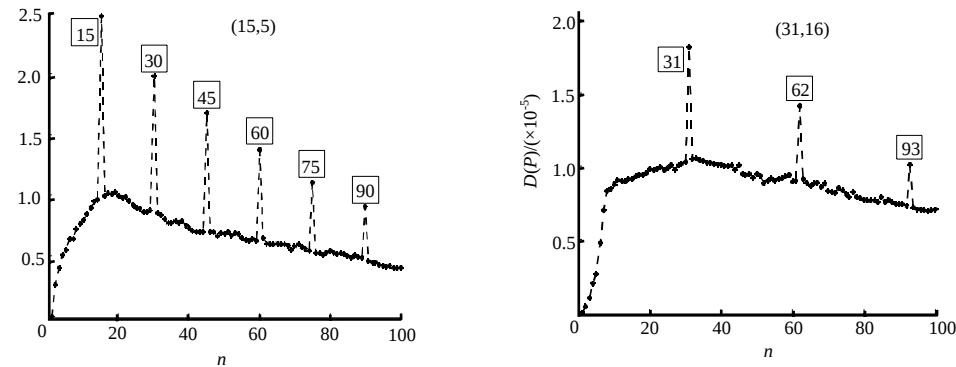


Fig.2 Code length recognition results for no error code
图 2 无误码码长识别结果

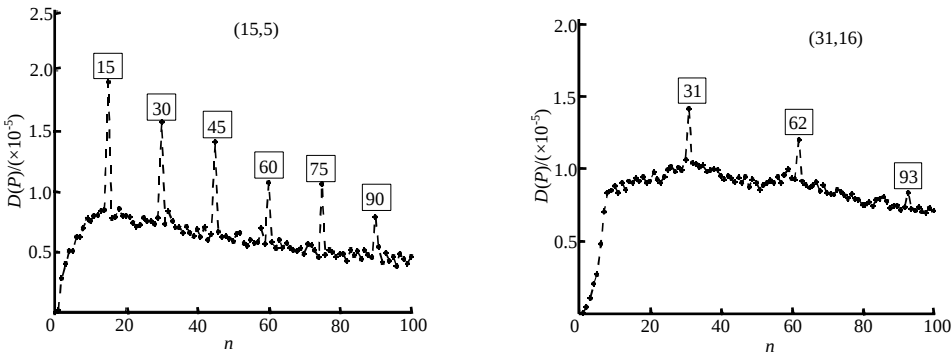


Fig.3 Code length recognition results for high error condition
图3 高误码条件下码长识别结果

无误码情况下码长的识别结果见图2,当遍历码长估计值不是真实码长 n 或 n 的整数倍时,其码重分布概率方差 $D(p)$ 变化平稳;相反,当遍历码长估计值为真实码长 n 或 n 的整数倍时,码组内码元间约束关系明显,分组后得到的码重分布为非等概率分布,其码重概率方差 $D(p)$ 相对于附近值变化剧烈,且出现明显的峰值。

图3是按表1中给出的误码率随机添加一定误码情况下码长的识别结果,当遍历码长估计值为真实码长 n 或 n 的整数倍时,码重分布概率方差 $D(p)$ 受到较明显的影响,其偏离平稳变化轨迹的程度降低,但并未影响真实码长 n 的识别结果,因此利用码重分布概率方差可以识别高误码条件下循环码的码长。

4.2 生成多项式识别验证

在识别出循环码码长的基础上,以误码率为 5.1×10^{-2} 的(15,5)循环码为例,其码重分布统计的结果见图4。

由图4可知,码重 i 为7的码字出现的概率 p 最大,经统计码重为7的码字共441个,其中此5个码字(000111011001010), (010100001110110), (001110110010100), (101000011101100) (011101100101000), 重复出现次数最多,各34次。首先选择码字(000111011001010)设为 C_1 ,来识别生成多项式。码字 C_1 及将 C_1 依次左移 i 位($i=1,2,\dots,14$)得到码多项式分别为: $C_1(x) = x^{11} + x^{10} + x^9 + x^7 + x^6 + x^3 + x$, $C_2(x) = x^{12} + x^{11} + x^{10} + x^8 + x^7 + x^4 + x^2, \dots, C_{15}(x) = x^{10} + x^9 + x^8 + x^6 + x^5 + x^2 + 1$ 。

求解码多项式的公因式得: $d(x) = \gcd[C_1(x), C_2(x), \dots, C_{15}(x)] = x^{10} + x^9 + x^8 + x^6 + x^5 + x^2 + 1$ 。

可知 $d(x)$ 满足定理1中生成多项式 $g(x)$ 的基本条件,否则选择另外4个码字中的码字进行识别。

下面通过设置判决关系式验证 $d(x)$ 是否为生成多项式。假设测试码元中的每个错误码元分布在不同的码字中,由表1计算错误码字的概率为 $786/1027=77\%$ 。统计所有码字的码多项式满足式(8)的概率,若此概率大于 $1-77\%=23\%$,则认为所求 $d(x)$ 即为生成多项式 $g(x)$ 。统计知1027个码多项式满足式(8)的概率为68%,显然,此概率大于23%,因此认为生成多项式 $g(x) = x^{10} + x^9 + x^8 + x^6 + x^5 + x^2 + 1$,信息位个数 k 是5,码率是5/15,识别完成。

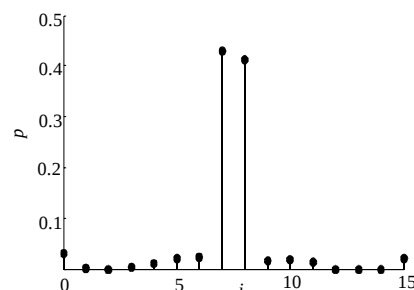


Fig.4 Code weight distribution

图4 码重分布

5 结论

本文在方差含义的基础上,给出了码重分布概率方差的定义,利用码重分布概率方差识别出了循环码的码长,仿真结果表明,在误码率为 5×10^{-2} 时仍能较好地识别出码长;同时根据码重分布统计的结果,选择出最佳的码字,通过求解公因式并设置合理的判决关系式,完成了循环码生成多项式的识别与验证。

参考文献:

- [1] 蒋青,于秀兰,范馨月. 通信原理[M]. 3版. 北京:人民邮电出版社, 2011. (JIANG Qing, YU Xiulan, FAN Xinyue. Communication Theory[M]. 3rd ed. Beijing: The People's Posts and Telecommunications Press, 2011.)
- [2] 柴先明,黄知涛,王丰华,等. 信道编码盲识别问题研究[J]. 通信对抗, 2008,2(101):33-36. (CHAI Xianming, HUANG Zhitao, WANG Fenghua, et al. Channel coding blind identification problem research[J]. Communication Countermeasure, 2008,2(101):33-36.)
- [3] 王磊,胡以华,王勇,等. 基于码重分布的系统循环码识别方法[J]. 计算机工程与应用, 2012,48(7):150-153. (WANG Lei, HU Yihua, WANG Yong, et al. Recognition method of system cyclic code based on code weight distribution[J]. Computer Engineering and Applications, 2012,48(7):150-153.)
- [4] 闫郁翰. 循环码的盲识别方法[J]. 电子科技, 2011(3):112-114. (YANG Yuhan. Blind Recognition Method for the Cyclic code[J]. Electronic Sci. & Tech., 2011(3):112-114.)
- [5] 闻年成,杨晓静. 采用秩统计和码根特征的二进制循环码盲识别方法[J]. 电子信息对抗技术, 2010(6):26-29. (WEN Niancheng, YANG Xiaojing. Blind Recognition of Cyclic Codes Based on Rank Statistic and Codes Roots Characteristic[J]. Electronic Information against Technology, 2010(6):26-29.)
- [6] 王甲峰,岳旻,权友波. 二进制BCH码的一种盲识别方法[J]. 信息与电子工程, 2011,9(5):591-595. (WANG Jiafeng, YUE Yang, QUAN Youbo. Blind recognition method of binary BCH code[J]. Information and Electronic Engineering, 2011,9(5):591-595.)
- [7] 杨晓静,闻年成. 基于码根信息差熵和码根统计的BCH码识别方法[J]. 探测与控制学报, 2010,32(3):69-73. (YANG Xiaojing, WEN Niancheng. Blind recognition method of binary BCH code based on code root information difference entropy and code root statistics[J]. Detection and Control Journal, 2010,32(3):69-73.)

- WEN Niancheng. Recognition Method of BCH Codes Based on Roots Information Dispersion Entropy and Roots Statistic[J]. Journal of Detection & Control, 2010,32(3):69-73.)
- [8] 盛骤,谢式千,潘承毅. 概率与数理统计[M]. 北京:高等教育出版社, 2010. (SHENG Zhou,XIE Shiqian,PAN Chengyi. Probability and Mathematical Statistics[M]. Beijing:Higher Education Press, 2010.)
- [9] 咎俊军,李艳斌. 低码率二进制线性分组码的盲识别[J]. 无线电工程, 2009,39(1):19-24. (ZAN Junjun,LI Yanbin. Blind Recognition of Low Code-rate Binary Linear Block Codes[J]. Radio Engineering, 2009,39(1):19-24.)
- [10] WANG Jiafeng,YUE Yang,YAO Jun, A Method of Blind Recognition of Cyclic Code Generator Polynomial[C]// Wireless Communications Networking and Mobile Computing(WiCOM). Chengdu:[s.n.], 2010:1-4.

作者简介:



郑瑞瑞(1984-), 男, 安徽省亳州市人, 在读硕士研究生, 主要研究方向为信号处理.
email:zrr0209@163.com.

汪立新(1966-), 男, 杭州市人, 研究员, 教授, 主要研究方向为信号处理、软件无线电.