

# M 序列报文的 MIAC 压缩编码算法

肖 瑶, 甘忠辉, 刘芸江\*

(空军工程大学信息与导航学院, 陕西西安, 710077)

**摘要** M 序列报文是典型短波数据链报文。提出一种针对 M 序列报文的 MIAC 压缩编码算法, 该算法根据 M 序列报文的统计特性, 应用基于字段结构的半动态概率估计模型, 量化低概率符号出现概率的取值且半动态更新, 采用动态的编码区间长度更新机制, 使乘法运算替换为预处理过程中的查表运算, 最大造成 10.04% 的编码效率损失。测试结果表明, 当 M 序列报文数目为 2~9 条时, MIAC 压缩编码算法在损失最多 20% 的压缩率时, 使压缩时间低于原算术编码至少 50%。为今后研究提升数据链系统性能提供一种新的思路。

**关键词** 算术编码; M 序列报文; 概率估计模型; 编码区间

**DOI** 10.3969/j.issn.1009-3516.2014.02.015

**中图分类号** TN919 **文献标志码** A **文章编号** 1009-3516(2014)02-0062-04

## A MIAC Compression Algorithm for M Sequence Message

XIAO Yao, GAN Zhong-hui, LIU Yun-jiang\*

(Information and Navigation College, Air Force Engineering University, Xi'an 710077, China)

**Abstract:** M sequence message is a typical message of HF data link. According to the statistical properties of M sequence message, this paper puts forward a new improved arithmetic coding, i.e. a semi-fixed probability estimative model based on frame structure proposed to quantify the probability of low probability symbols, a dynamic update mechanism of encoding interval length applied to update discretely encoding interval length in real time, and the algorithm caused the coding efficiency loss up to 10.04%. The test results show that when the quantity of M sequence message is from 2 to 9 and MIAC compression coding algorithm loses its compression ratio by 20% at most, compression time of the improved algorithm is lower than the original arithmetic coding at least 50%. And this will provide a new way of thinking for the study of the performance of data link system in the future.

**Key words:** arithmetic coding; M sequence message; probability estimation model; encoding interval

Link-11 数据链主要应用于飞机、舰船以及岸上指挥机构之间的情报交换, 可工作于短波波带, 具备超视距通信能力, 其报文标准采用 STANAG 5511 定义的 M 序列报文<sup>[1]</sup>。随着数据链系统的应

用和发展, 海量数据传输成为其亟待解决的问题。Link-11 数据链采用轮询组网协议, 数据传输速率不高于 2 500 b/s, 极大限制了轮询周期内数据传输量, 数据压缩作为解决这一难题最有效的技术手段,

**收稿日期:** 2013-06-19

**基金项目:** 国家自然科学基金资助项目(61302153)

**作者简介:** 肖 瑶(1989—), 男, 湖南衡阳人, 硕士生, 主要从事航空通信研究. E-mail: a55swa@163.com

\* **通信作者:** 刘芸江(1976—), 男, 四川内江人, 副教授, 博士, 主要从事航空通信研究. E-mail: lyj.afeu@126.com

**引用格式:** 肖瑶, 甘忠辉, 刘芸江. M 序列报文的 MIAC 压缩编码算法[J]. 空军工程大学学报: 自然科学版, 2014, 15(2): 62-65. XIAO Yao, GAN Zhonghui, LIU Yunjiang. A MIAC compression algorithm for M sequence message[J]. Journal of air force engineering university: Natural science edition, 2014, 15(2): 62-65.

越来越受到广泛的关注。

算术编码 (Arithmetic Coding, AC) 是一种接近熵极限的最佳编码<sup>[2-3]</sup>, 其编码器和概率估计模型是完全分离的。在编码器设计方面, 算术编解码过程存在大量的乘除法运算, 导致其实现复杂度巨大, 由此, 关于避免乘法运算的快速算术编码器研究工作大量展开: 移位加编码器<sup>[4]</sup>利用移位、加法运算避免原乘法运算; MQ 编码器<sup>[5]</sup>增加条件交换和概率估计中的贝叶斯学习过程, 应用于 JPEG2000 标准中; M 编码器<sup>[6]</sup>将概率区间大小和低概率符号概率离散化, 用查表运算代替乘法运算, 主要应用于 H.264 标准中。在概率估计模型方面, 主要分为静态模型和动态模型。静态模型中每个符号集的概率分布都预先统计, 编码过程中动态不变, 其应用典型代表是 H.263 标准; 动态模型<sup>[7]</sup>在编码过程中, 根据已编码符号的统计特征, 实时更新概率模型从而捕捉信号的局部特征, 能够提供更好的编码性能。

由于 AC 算法时效性差, 难以在数据链系统中应用, 目前 AC 算法在数据链系统中的应用研究还未见报道。文章试图建立一种针对 M 序列报文的改进型算术编码 (M Sequence Message Improved Arithmetic Coding, MIAC), 实现 M 序列报文有效传输。

## 1 AC 算法基本原理

AC 算法将信源序列概率映射到  $[0, 1)$  之间的一个实数区间, 区间长度对应信源序列概率, 其编码比特对应应该区间中任意一点的二进制数。编码器和概率估计模型 2 个环节完全分离, 使得 AC 算法能够在编码过程中实时更新概率分布, 其编码过程见图 1。

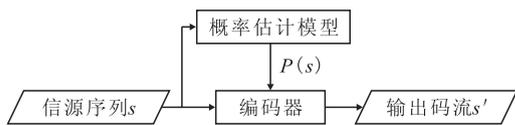


图 1 AC 算法编码器编码过程

Fig. 1 The encoding process of the AC algorithm

图 1 中, 信源序列  $s$  取自符号集  $S = \{a_1, a_2, \dots, a_K\}$ ,  $p_i$  为符号  $a_i$  的概率,  $K$  为符号个数;  $P(s)$  为信源序列  $s$  估计概率, 由概率估计模型决定;  $s'$  为编码后输出码流。

式(1)为各符号的积累概率为  $q(l)$ :

$$q(l) = \sum_{i=1}^l p_i \quad (1)$$

令  $U(n)$ 、 $D(n)$  和  $A(n)$  为第  $n$  步编码区间的

上限、下限和长度, 且  $U(0) = 0$ 、 $D(0) = 1$  和  $A(0) = 1$ , 则第  $n$  步编码过程中符号  $a_l$  的上限、下限和长度更新计算见式(2):

$$\begin{cases} U(n) = D(n) + A(n) \\ D(n) = D(n-1) + A(n-1)q(l-1) \\ A(n) = A(n-1)p_l \end{cases} \quad (2)$$

二进制化  $U(n)$ 、 $D(n)$  分别为  $(b_1^n, b_2^n, \dots, b_M^n)$ 、 $(b_1^d, b_2^d, \dots, b_M^d)$ , 若前  $m$  bit 相同, 则输出; 否则保留所有比特, 继续编码检验下一个符号, 不断递归。由于解码算法即为编码算法的逆过程, 故不深入讨论。

## 2 MIAC 压缩编码算法

数据链系统要求提高数据传输速率, 缩短各种机动目标信息的更新周期, 以便及时显示目标的运动轨迹, 而原 AC 算法运算量大, 实现复杂度高, 难以在高实时性要求的系统中应用。Link-11 数据链使用的每条 M 序列报文都由预定义顺序的 60 bit 组成, 分为 2 帧, 每帧 30 bit, 其中 6 bit 用于纠错编码, 24 bit 按照一定规则用于编码战术信息<sup>[1]</sup>。M 序列报文为典型二元序列, 其符号集  $S = \{0, 1\}$ , 共 16 种类型, 其中空中目标跟踪报文 M.2 见图 2。

17~24 bit	DETAILED INFORMATION
5~16 bit	TRACK NUMBER
1~4 bit	MESSAGE NUMBER
37~48 bit	Y COORDINATE
25~36 bit	X COORDINATE

图 2 M.2 报文的 48 bit 战术信息定义

Fig. 2 The 48 bit tactical information of M.2 message

图 2 中, 不同比特字段有不同含义, 这种字段结构可以用来构建概率估计模型。

### 2.1 半动态概率估计模型

根据式(2), AC 算法每步区间更新存在乘法运算  $A(n-1)p_l$ , 且必须在一个符号周期内完成。在编码过程中, 设低概率符号的出现概率为  $p_{LPS}(n)$ , 且其取值限定在  $Q$  种内, 见式(3):

$$p_{LPS}(n) = \{p_0, p_1, \dots, p_{Q-1}\} \quad (3)$$

式中, 每步区间更新  $p_{LPS}(n)$  的值由概率估计模型决定, 根据 M 序列报文应用的统计特性, 将  $p_{LPS}(n)$  值量化为 32 个递归状态<sup>[6]</sup>, 见式(4):

$$p_i = \sqrt[Q]{0.6323} p_{i-1}, i = 1, 2, \dots, Q \quad (4)$$

式中:  $p_0 = 0.5$ ;  $Q = 32$ 。

令  $N_0^n$  为第  $n$  步区间更新之前符号“0”出现频次,  $N_1^n$  为第  $n$  步区间更新之前符号“1”出现频次, 则  $p_{LPS}(n)$  更新机制采用一种基于字段结构的半动态概率估计模型, 其规则如下:

1) 预存报文编号不同值相应的帧结构,共 16 种类型;

2) 各比特字段内独立统计  $N_0^n, N_1^n$  值,且  $p_{LPS}(n)$  值更新见式(5):

$$p_{LPS}(n) = p_i, \quad \text{if } p_i \leq \frac{\min\{N_0^n, N_1^n\}}{N_0^n + N_1^n} < p_{i-1} \quad (5)$$

式中:若  $\frac{\min\{N_0^n, N_1^n\}}{N_0^n + N_1^n} < p_{31}$ , 则  $p_{LPS}(n) = p_{31}$ ;

3) 含有条件选择关系的比特字段合并统计  $N_0^n, N_1^n$  值,并按规则 2) 更新;

4) 长度小于 4 bit 的字段,分别存储记录,根据多条报文共同统计  $N_0^n, N_1^n$  值,并按式(5)更新。

## 2.2 编码区间长度动态更新机制

$p_{LPS}(n)$  的值根据 M 序列报文的统计特性量化为 32 个离散状态,为避免乘法运算,第  $n$  步编码区间长度  $A(n)$  取值限定在  $W$  种内,见式(6):

$$A(n) \in \{x \mid x = 2^{-i}, i = 1, 2, \dots, W\} \quad (6)$$

式(6)中,  $W$  值的大小需要折中存储复杂度和编码效率,根据不同要求,可设定不同大小的值。码区间长度  $A(n)$  根据字段结构动态分段放大,其值更新规则如下:

1)  $A(n), A(n-1)$  属于同字段结构内,其值更新机制见式(7):

$$A(n) = 2^{-i-1}, \quad \text{if } 2^{-i-1} < A(n-1)p_{LPS}(n) \leq 2^{-i} \quad (7)$$

2)  $A(n), A(n-1)$  属于不同字段结构内,那么将  $A(n-1)$  放大  $2^k$  倍,  $k$  为  $A(n-1)$  所属比特字段长度;

3)  $A(n-1)p_{LPS}(n) \leq 2^{-W}$ , 则  $A(n) = 2^{-W}$ , 且下一步区间更新时,将编码区间长度放大  $2^W$  倍。

通过限定的  $p_{LPS}(n)$  和  $A(n)$  取值,每步区间更新的乘法运算简化为预处理中的查表运算,编码前初始化  $A p_{LPS}$  的  $Q \times W$  表,由于  $A p_{LPS}$  为移位运算,初始化时间可忽略不计。

## 2.3 编码效率分析

半动态概率估计模型下  $p_{LPS}(n)$  取值量化为 32

表 2 不同 M 序列报文数目下 2 种算法的压缩率与压缩时间比较

Fig. 2 The comparison of compression ratio and compression time under different algorithm

M 序列报文数目		2	3	4	5	6	7	8	9
压缩率/%	AC	31.34	31.22	26.86	23.84	22.10	20.62	19.71	18.92
	MIAC	47.48	50.78	42.62	42.96	39.79	40.60	35.10	36.13
压缩时间/ms	AC	12	22	30	30	38	44	50	52
	MIAC	2	2	8	14	18	20	22	26

分析表 2 可以得出:①在压缩率方面,由于采用半动态概率模型以及离散编码区间长度,MIAC 压缩率低于 AC 压缩率 15%~20%。②在压缩时间方面,MIAC 压缩时间优于 AC 压缩时间至少 50%,

个状态,见式(4),其中每个  $p_i$  值对应一段概率区间。根据量化状态  $p_i$  编码,则每个符号的平均码长计算如式(8):

$$L(p) = -p \log p_i - (1-p) \log(1-p_i) \quad (8)$$

信源符号熵计算见式(9):

$$H(p) = -p \log p - (1-p) \log(1-p) \quad (9)$$

根据  $p_{LPS}(n)$  更新机制规则 2),  $H(p_i)$  值代替概率区间  $[p_{i+1}, p_i)$  的平均码长,即在  $p_{i+1}$  处误差最大,编码效率最低。因此,概率区间  $[p_{i+1}, p_i)$  的最低编码效率  $\eta_i$  和平均编码效率  $\bar{\eta}_i$  计算见式(10):

$$\begin{cases} \eta_i = \frac{H(p_{i+1})}{H(p_i)} \times 100\% \\ \bar{\eta}_i = \frac{H(p_{i+1})(p_i - p_{i+1})}{\int_{p_{i+1}}^{p_i} H(p) dp} \times 100\% \end{cases} \quad (10)$$

不同概率区间中,  $\eta_i, \bar{\eta}_i$  值计算部分结果如表 1 所示,采用 32 个离散状态来量化  $p_{LPS}(n)$  最大造成 10.04% 的编码损失,平均编码损失均在 7% 以下。

表 1 不同概率区间内  $\eta_i$  值计算结果

Tab. 1 The calculation results of  $\eta_i$

in different probability interval

$i$	0	1	2	3	4
$H(p_{i+1})$	0.899 6	0.812 5	0.751 2	0.705 2	0.668 8
$\eta_i/\%$	89.96	90.32	92.45	93.87	94.85
$\bar{\eta}_i/\%$	93.04	94.64	95.97	96.79	97.32

## 3 测试结果及分析

测试基于工控机 PC,主频为 2.2 GHz,内存为 1 GB,操作系统为 CentOS5.3。由于数据链系统的高实时性要求,缓冲区不能设置过大,压缩对象一般为 2~9 条 M 序列报文。不同 M 序列报文数目下 AC 算法与 MIAC 算法的压缩率<sup>[2]</sup>与压缩时间测试结果见表 2。

更适合数据链系统实时性要求。MIAC 压缩编码算法通过牺牲一部分压缩率,使压缩时间大幅下降,以满足数据链应用的实时性要求。

## 4 结语

在分析原 AC 算法的基础上,结合 Link11 数据链系统应用特点,提出一种针对 M 序列报文的改进型算术编码 MIAC 算法。MIAC 算法采用半动态概率估计模型和动态编码区间长度更新机制,其编码效率最大损失 10.04%,且测试结果表明,在 M 序列报文数目为 2~9 条时,MIAC 算法在损失最多 20%的压缩率时,使压缩时间低于原算术编码至少 50%,为今后提高数据链系统性能的研究提供一种新的方法。

### 参考文献(References):

- [1] STANAG 5511, Link 11/Link 11B. NATO Standardization Agency.
- [2] 吴乐南. 数据压缩[M]. 北京:电子工业出版社,2012. WU Lenan. Data compression[M]. Beijing: Publishing house of electronics industry,2012.(in Chinese)
- [3] 段黎力,廖晓峰,向涛. 基于 Markov 性质的一阶安全算术编码及应用[J]. 物理学报,2010,59(10): 6743-6751. DUAN Lili, LIAO Xiaofeng, XIANG Tao. Image encryption based on arithmetic coding with order - 1

Markov model[J]. Acta physica sinca,2010,59(10): 6743-6751.(in Chinese)

- [4] Lei S M. Efficient multiplication-free arithmetic codes [J]. IEEE transactions on communications, 1995, 43 (12): 2950-2958.
- [5] 邓海涛,邓家先,邓小梅. 基于 EZW 的图像压缩和树形加密同步算法[J]. 物理学报,2013,62(11): 110701. DENG Haitew, DENG Jiaxian, DENG Xiaomei. Joint compression and tree structure encryption algorithm based on EZW[J]. Acta physica sinca,2013,62(11): 110701.(in Chinese)
- [6] Belyaev E, Turlikov A, Egiastian K, et al. An efficient multiplication-free and look-up table-free adaptive binary arithmetic coder[C]//19th IEEE international conference on ICIP. [S. l.]:IEEE press,2012,701-704.
- [7] 杨文涛,刘卫忠,郑立新,等. 多阶上下文自适应二进制算术编码实现[J]. 华中科技大学学报:自然科学版,2012,35(3): 42-45. YANG Wentao, LIU Weizhong, ZHENG Lixin, et al. Realization of multi-order context-based adaptive binary arithmetic coding[J]. Journal of huazhong university of science and technology; natural science edition, 2012, 35(3): 42-45.(in Chinese)

(编辑:徐楠楠)

(上接第 56 页)

### 参考文献(References):

- [1] Yilmaz, A, Javed, O, Shah M. Object tracking: a survey [J]. ACM computing surveys,2006,38(4):352-355.
- [2] 侯志强,韩崇昭. 视觉跟踪技术综述[J]. 自动化学报,2006,32(4): 603-617. HOU Zhiqiang, HAN Chongzhao. A survey of visual tracking [J]. Acta automatic sinica, 2006, 32(4): 603 -617. (in Chinese)
- [3] Sethi I, Jain R, Finding trajectories of feature points in monocular images [J]. IEEE transactions on pattern analysis and machine intelligence, 1987, 9(1), 56-73.
- [4] Pan P, Schonfeld D. Visual tracking using high-order particle filtering [J]. IEEE signal processing letters, 2011, 8(1): 51-54.
- [5] Ido L. Mean shift trackers with cross-bin metrics [J]. IEEE transactions on pattern analysis and machine in-

telligence, 2012, 34(4), 695-706.

- [6] Comaniciu D, Meer P. Mean shift analysis and applications [C]//Proceedings of the 7th international conference on computer vision. [S.l.]:IEEE press,1999:1197 -1203.
- [7] Comaniciu D, Ramesh V, Meer P. Kernel-based object tracking [J]. IEEE transactions on pattern analysis and machine intelligence, 2003, 25(5), 564-557.
- [8] Adam A, Rivlin E, Shimshoni I. Robust fragments-based tracking using the integral histogram [C]//Proceedings of conference on computer vision and pattern recognition. [S.l.]:IEEE press,2006:798-805.
- [9] Porikli F. Integral histogram: a fast way to extract histograms in cartesian spaces [C]//Proceedings of the conference on computer vision and pattern recognition. [S. l.]:IEEE press,2005:829-836.

(编辑:徐楠楠)