

# 一种基于3容错阵列码的RAID数据布局

万武南<sup>1)</sup> 吴 震<sup>1)</sup> 陈 运<sup>1)</sup> 王晓京<sup>2)</sup>

<sup>1)</sup>(成都信息工程学院网络工程系 成都 610225)

<sup>2)</sup>(中国科学院成都计算机应用研究所 成都 610041)

**摘 要** 在 EVENODD 码的基础上,提出一种新的基于 EEOD 码的 RAID 数据布局,只需要 3 个额外的磁盘保存校验信息,能容许任意 3 个磁盘同时故障,并给出了 EEOD 的代数定义,理论上证明了 EEOD 码的 MDS 性质.从一种新的途径讨论了 EEOD 码的译码过程:用图的回路表示通过“异或”运算得到的校验方程组,把译码过程归结为图回路的叠加,进而校验方程组图中度为偶数的顶点逐步消除.讨论了基于 EEOD 码阵列布局的性能,与其它 RAID 结构相比,容灾能力大幅度提高,编码和译码过程只需要简单的异或运算,但是空间利用率影响非常小,并且 EEOD 具有很好的性能,具有很好的应用前景.

**关键词** EVENODD 码;RAID 结构;阵列码;数据布局;MDS

中图法分类号 TP333

## A Data Placement Based on Toleration on Triple Failures Array Codes in RAID

WAN Wu-Nan<sup>1)</sup> WU Zhen<sup>1)</sup> CHEN Yun<sup>1)</sup> WANG Xiao-Jing<sup>2)</sup>

<sup>1)</sup>(Department of Networks Engineering, Chengdu University of Information Technology, Chengdu 610225)

<sup>2)</sup>(Chengdu Institute of Computer Applications, Chinese Academy of Sciences, Chengdu 610041)

**Abstract** A novel data placement based on EEOD code in RAID has been presented. The EEOD is an extension of the double-erasure-correcting EVENODD, which can tolerate simultaneous failures of three member disks with only three extra disks for parity information. A rigorous theoretical proof of MDS property based on algebraic representation is given. The EEOD code's decoding algorithms is discussed by representing a check group consisting of data and parity units with graphics circuits. The decoding process can be regarded as superimposition of graphics circuits and vertices of even degrees in the graph are eliminated gradually. The encoding and decoding procedure of EEOD code is based on XOR operations, and the decoding complexity of the EEOD code is much lower than those of the existing comparable codes, while decreases the influences on system's throughput and disk capacity. Thus the EEOD code is of practical significance for storage systems.

**Keywords** EVENODD; RAID architectures; array codes; data placement; MDS

收稿日期:2007-05-09;修改稿收到日期:2007-07-25. 本课题得到国家“九七三”重点基础研究发展规划项目基金(2004CB318003)、四川省重点科技攻关项目基金(03GG0326)和成都信息工程学院人才启动基金(KYTZ200706)资助. 万武南,女,1978年生,博士,讲师,主要研究方向为信息安全、编码理论. E-mail: nan\_wwn@hotmail.com. 吴 震,男,1975年生,讲师,主要研究方向为信息安全、3G 通信安全. 陈 运,女,1958年生,教授,主要研究领域为信息安全、3G 通信安全. 王晓京,男,1955年生,研究员,博士生导师,主要研究领域为信息安全、编码理论.

# 1 引 言

随着网络的进一步发展,网络存储已成为互联网信息存储领域的前沿方向<sup>[1]</sup>.可靠性是网络存储系统最重要的指标之一,也是当今信息社会对信息存储的迫切需求<sup>[2]</sup>.1988年加州大学提出的冗余磁盘阵列(Redundant Array of Independent Disks, RAID),目前已成为网络存储的主流技术,它通过数据分布存储、并行访问以及信息冗余等技术,极大地扩大了存储容量,增强了 I/O 请求处理能力,提高了数据可靠性<sup>[3]</sup>.然而由于硬件系统本身的脆弱性和其它各种不确定因素造成系统不可预知的不可用性,数据经常遭受破坏,因此如何为应用系统提供高可靠性是 RAID 技术面临的一个重要难题.早期的 RAID 技术采用奇偶校验码来保障数据的安全性(RAID-3, RAID-5)<sup>[4]</sup>,但随着信息存储自然增长的需要以及多媒体的发展,人们对磁盘阵列容量的需要越来越大.为充分发挥阵列容量、高可靠性的优势,研究磁盘阵列纠双错、纠多错的编码方法已成为 RAID 技术的一个重要的研究领域<sup>[5-13]</sup>.

RAID 结构第一个要考虑的问题就是冗余信息的产生方法(即编码方法).在 RAID 结构中,设计编码需要考虑的几个方面:(1)一般每个存储设备的故障能被存储控制器检测,可预先知道每个磁盘有没有故障,没有发生故障的磁盘上的数据都是正确的,因而,编码技术起的是纠删的作用,它被用于恢复丢失的数据;(2)在 RAID 结构中,考虑故障时总是认为整个磁盘发生故障.这样丢失的多个信息元之间有明确的位置关系,一般相当于编码矩阵中的一列或者一行;(3)RAID 结构中,必须采用确定性编码,因为若采用非确定性编码技术,有可能丢失很少的一部分数据,源数据可能得不到恢复.由于阵列码的二维码字的结构刚好比较符合磁盘阵列的结构;并且编译码只需要异或运算,在相同编码效率下,阵列码按照计算复杂度比一般线性码来得更加有效,一直是 RAID 数据布局的研究热点问题<sup>[5-14]</sup>.在 RAID 结构中,容许单个和两个磁盘同时故障的阵列布局已有大量的研究,如 EVENODD 码<sup>[5]</sup>、X 码<sup>[6]</sup>、B 码<sup>[7]</sup>、RDP 码和 S 码等.容许多个磁盘同时故障,特别是容许 3 个磁盘同时故障的阵列布局也有一些结果<sup>[8-13]</sup>.

Blaum 码<sup>[8]</sup>是一种能够承受多个磁盘故障的 RAID 结构,并且容灾能力随着冗余磁盘数的增加

而增加.但这种结构的主要问题在于解码方法是解多项式环上的一组线性方程,解码算法不易实现,并且解码复杂度高. HoVer 码<sup>[9]</sup>、WEAVER 码<sup>[10]</sup>也是一种能够承受多个磁盘故障的 RAID 结构,这种结构的主要问题在于 HoVer 码、WEAVER 码不是最大可分码(MDS 码,根据编码理论,按照冗余量和恢复能力来说没有达到最佳),冗余盘数目并不是随着磁盘阵列系统总盘数的线性增长,冗余信息量太大、代价昂贵. Tau 在文献<sup>[11]</sup>提出 HDD1 码和 HDD2 码,这种结构要求磁盘的个数为  $N+1$ ,虽然文献中给出了在发生 3 个磁盘故障后恢复用户数据的例  $r$ (例  $r$  中  $N=5$ ),但是文中利用 Hamming 距离进行的正确性证明并不充分,并且 HDD1 和 HDD2 码的解码过程需要做线性方程高斯消元求解,其解码复杂度等于 9(每个信息位的恢复差不多要做大于 9 次的异或运算).另外 Feng 等人在文献<sup>[12-13]</sup>提出的两种新编码,能够承受 3 个和多个磁盘同时故障,这两种编码的校验矩阵分别是由类似于范德蒙矩阵和柯西矩阵来构造的,这两种结构的问题在于编码解码算法不易实现,而且解码复杂度同样需要做线性方程高斯消元求解.

EVENODD 码<sup>[5]</sup>由 Blaum 提出来的一类阵列码,能够同时容许两个磁盘的故障,其编译码实现算法简单,易于软硬件实现,是当前 RAID 系统(RAID-5)广泛使用的一类阵列码.文本在 EVENODD 码的基础上,提出了一种扩展 EVENODD 码——EEOD (extending EVENODD)码,保留了 EVENODD 码具有的好的性能,并能容许任意 3 个磁盘同时故障,与其它 RAID 结构相比,容灾能力大幅度提高,易于在现有的 RAID 结构中扩展,具有很好的应用前景.

## 2 EEOD 码的编码方法

EEOD 码是在 EVENODD 码的基础上扩展出来的,因此在介绍 EEOD 码之前,先简要介绍一下 EVENODD 码.先定义本文下面需要使用的一些符号: $\mathbf{A}_l$ 表示  $l \times l$  大小的方阵,即记为  $\mathbf{A}_l = (a_{i,j})_{l \times l}$ ,其中  $0 \leq i, j \leq l-1$ .  $\mathbf{A}_{l_1 \times l_2}$ 表示  $l_1 \times l_2$  大小的矩阵,即记为  $\mathbf{A}_{l_1 \times l_2} = (a_{i,j})_{l_1 \times l_2}$ ,其中  $0 \leq i \leq l_1-1, 0 \leq j \leq l_2-1$ .  $\langle a \rangle_m$ 表示模  $m$  的运算,即  $x = \langle a \rangle_m = a \pmod{m}$ .

注.本文中的矩阵都是在域  $GF(2)$  上的二元矩阵,  $m$  的值为大于等于 2 的素数.

### 2.1 EVENODD 码

EVENODD 码的码字放在一个  $(m-1) \times (m+2)$

的阵列中,其中原始数据放在前  $m$  列中,最后两列存放冗余校验数据. 根据编码理论,EVENODD 码可以记为  $(m+3,m,3)$  EVENODD 码,二维码字记为  $\mathbf{C}=[c_{i,j}](0\leq i\leq m-1,0\leq j\leq m+1)$ ,  $c_{i,j}$  表示为第  $j$  列第  $i$  行的信息位或校验位,则两列冗余校验

位构造公式如下:

$$c_{u,m}=\bigoplus_{t=0}^{m-1}c_{u,t};$$
$$S_1=\bigoplus_{t=1}^{m-1}c_{m-1-t,t},c_{u,m+1}=S_1\bigoplus\left(\bigoplus_{\substack{t=0\\t\neq u-1}}^{m-1}c_{\langle u-t\rangle_m,t}\right).$$

表 1 给出了  $(7,5,3)$  EVENODD 码的编码实例.

表 1 (7,5,3)EVENODD 码的编码( $S_1=c_{04}+c_{13}+c_{22}+c_{33}$ )

$c_{00}$	$c_{01}$	$c_{02}$	$c_{03}$	$c_{04}$	$c_{00}+c_{01}+c_{02}+c_{03}+c_{04}$	$S_1+c_{00}+c_{14}+c_{23}+c_{32}$
$c_{10}$	$c_{11}$	$c_{12}$	$c_{13}$	$c_{14}$	$c_{10}+c_{11}+c_{12}+c_{13}+c_{14}$	$S_1+c_{01}+c_{10}+c_{24}+c_{33}$
$c_{20}$	$c_{22}$	$c_{22}$	$c_{23}$	$c_{24}$	$c_{20}+c_{21}+c_{22}+c_{23}+c_{24}$	$S_1+c_{02}+c_{11}+c_{20}+c_{34}$
$c_{30}$	$c_{33}$	$c_{32}$	$c_{33}$	$c_{34}$	$c_{30}+c_{31}+c_{32}+c_{33}+c_{34}$	$S_1+c_{03}+c_{12}+c_{22}+c_{30}$

2.2 EEOD 码编码的几何描述

为了能够承受 3 个磁盘同时故障,在 EVENODD 码的基础上进行了扩展,EEOD 码其编码矩阵为  $m+3$  列,行为  $m-1$ ,其中前  $m$  列存放原始数据,后 3 列存放冗余校验数据,则可记为  $(m+3,m,4)$  EEOD 码. 前两列冗余校验列的构造与 EVENODD

完全一样,增加 1 列冗余校验列的信息位则按照下面的公式进行计算:

$$S_2=\bigoplus_{t=1}^{m-1}c_{\langle m-1-2t\rangle_m,t},c_{u,m+1}=S_2\bigoplus\left(\bigoplus_{\substack{t=0\\ \langle u-2t\rangle_m\neq m-1}}^{m-1}c_{\langle u-2t\rangle_m,t}\right).$$

表 2 给出了  $(8,5,4)$  EEOD 码的编码实例.

表 2 (8,5,4)EEOD 码的编码( $S_1=c_{04}+c_{13}+c_{22}+c_{33}$ , $S_2=c_{21}+c_{02}+c_{33}+c_{14}$ )

$c_{00}$	$c_{01}$	$c_{02}$	$c_{03}$	$c_{04}$	$c_{00}+c_{01}+c_{02}+c_{03}+c_{04}$	$S_1+c_{00}+c_{14}+c_{23}+c_{32}$	$S_2+c_{00}+c_{31}+c_{12}+c_{24}$
$c_{10}$	$c_{11}$	$c_{12}$	$c_{13}$	$c_{14}$	$c_{10}+c_{11}+c_{12}+c_{13}+c_{14}$	$S_1+c_{01}+c_{10}+c_{24}+c_{33}$	$S_2+c_{10}+c_{22}+c_{03}+c_{34}$
$c_{20}$	$c_{22}$	$c_{22}$	$c_{23}$	$c_{24}$	$c_{20}+c_{21}+c_{22}+c_{23}+c_{24}$	$S_1+c_{02}+c_{11}+c_{20}+c_{34}$	$S_2+c_{20}+c_{01}+c_{32}+c_{13}$
$c_{30}$	$c_{33}$	$c_{32}$	$c_{33}$	$c_{34}$	$c_{30}+c_{31}+c_{32}+c_{33}+c_{34}$	$S_1+c_{03}+c_{12}+c_{22}+c_{30}$	$S_2+c_{30}+c_{11}+c_{23}+c_{04}$

下面从几何角度来描述 EEOD 码的编码过程,如图 1 所示 EEOD 码的每个二维阵列码字  $\mathbf{C}$  的信息位可看作是平面坐标上的格点,横坐标表示信息位的列号,纵坐标表示信息位的行号. 并且坐标轴的取值为模  $m$  运算,取值范围为  $0,1,\cdots,m-1$ . 如某点坐标为  $(1,2)$ ,该点表示码字  $\mathbf{C}$  的  $c_{1,2}$  信息位,即对应的二维阵列中第 2 列第 3 行的信息位. 其中坐标轴上纵坐标为  $m-1$  的坐标点,其对应的信息位的值全为零,是增加虚拟坐标点,即二维阵列码字  $\mathbf{C}$  增加一行零向量: $c_{m-1,u}=\mathbf{0},0\leq u\leq m-1$ (该行对描述译码过程也非常有用).

每列校验位调节因子  $S_i$  的值是从坐标轴上  $(m-1,0)$  格点开始,沿某固定斜率的直线经过的格点对应的信息位异或运算的值. 校验列第  $i$  个校验位则是从  $(i,0)$  格点开始,沿某固定斜率的直线经过的格点对应的信息位异或运算的值,再与该校验列调节因子  $S_i$  的值进行异或运算的值.

从几何的角度来看,EEOD 码是在 EVENODD 码基础上,增加 1 列其斜率为 2 的校验列,图 1(c) 给出了  $(8,5,4)$  EEOD 码三列校验列编码几何示意图,图中三校验列的每个校验位分别是沿斜率  $(0,1,2)$  经过的格点对应的信息位异或运算的值. 图中黑点的坐标点对应的信息位为零,即  $c_{4,u}=\mathbf{0}(0\leq u\leq m-1)$  (实际并不存在),称为虚拟信息位,白点表示信息位. 虚线表示调节因子  $S$ ,实线表示校验列的每个校验位(与调节因子  $S$  异或之前的值).

2.3 EEOD 阵列码编码的代数定义

为了证明 EEOD 码是一类最小列距离为 4 的 MDS 阵列码,引入代数表示方法来描述 EEOD 码编码过程. 在给出 EEOD 阵列码编码的代数定义之前,预先给出相关矩阵的定义.

定义 1(矩阵  $\mathbf{E}_m$ ). 矩阵  $\mathbf{E}_m$  是  $m\times m$  的矩阵,定义如下:

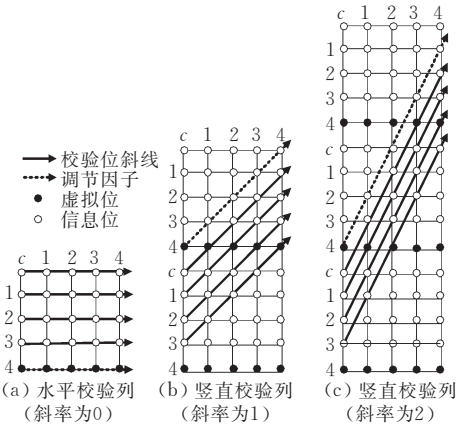


图 1 (8,5,4)EEOD 编码过程

$$E_m = \begin{bmatrix} \mathbf{0}_{m-1} & \mathbf{1} \\ \mathbf{I}_{m-1} & \mathbf{0}_{m-1}^T \end{bmatrix},$$

其中,  $\mathbf{I}_{m-1}$  是  $(m-1) \times (m-1)$  的单位矩阵,  $\mathbf{0}_{m-1}$  是  $1 \times (m-1)$  的零向量,  $\mathbf{0}_{m-1}^T$  是  $(m-1) \times 1$  的零向量.

**引理 1.** 矩阵集  $\{\mathbf{I}_m, \mathbf{E}_m, \mathbf{E}_m^2, \dots, \mathbf{E}_m^{m-1}\}$  构成一个在  $GF(2)$  上的阿贝尔乘法群<sup>[21]</sup>.  $\mathbf{E}_m^a = (e_{i,j})_{m \times m}$  定义如下:

$$e_{i,j} = \begin{cases} 1, & i = \langle j + \mu \rangle_m, \\ 0, & \text{其它} \end{cases},$$

则  $\mathbf{E}_m^m = \mathbf{I}_m$ ,  $\mathbf{E}_m^{-1} = \mathbf{E}_m^{m-1}$ .

证明. 略(参见文献[12]).

**定义 2**( $Q_{m \times (m-1)}, \tilde{Q}_{(m-1) \times m}$  矩阵).  $Q_m, \tilde{Q}_m$  分别是  $m \times (m-1), (m-1) \times m$  矩阵, 定义如下:

$$\tilde{Q}_{(m-1) \times m} = [\mathbf{I}_{m-1} \quad \mathbf{1}_{m-1}^T], \quad Q_{m \times (m-1)} = \begin{bmatrix} \mathbf{I}_{m-1} \\ \mathbf{0}_{m-1} \end{bmatrix},$$

其中,  $\mathbf{I}_{m-1}$  是  $(m-1) \times (m-1)$  的单位矩阵,  $\mathbf{0}_{m-1}$  是  $1 \times (m-1)$  的零向量,  $\mathbf{1}_{m-1}^T$  是  $(m-1) \times 1$  全 1 的向量.

**注.**  $\mathbf{E}_m, Q_{m \times (m-1)}, \tilde{Q}_{(m-1) \times m}$  矩阵分别缩写为  $\mathbf{E}, Q, \tilde{Q}$ .

**定义 3.** (( $m+3, m, 4$ ) EEOD 码的编码过程代数定义) EEOD 码的任意二维阵列码字  $\mathbf{C}$  记为

$$\mathbf{C} = \{\mathbf{c} = (\mathbf{c}_0, \mathbf{c}_1, \dots, \mathbf{c}_{m-1}, \mathbf{c}_m, \mathbf{c}_{m+1}, \mathbf{c}_{m+2})\},$$

其中,  $\mathbf{c}_0, \mathbf{c}_1, \dots, \mathbf{c}_{m-1}$  为  $m$  列信息列向量,  $\mathbf{c}_m, \mathbf{c}_{m+1}, \mathbf{c}_{m+2}$  为 3 列校验列向量,  $\tilde{\mathbf{H}}_3^*$  为校验矩阵, EEOD 码的编码过程代数定义如下:

$$\begin{bmatrix} \mathbf{c}_m \\ \mathbf{c}_{m+1} \\ \mathbf{c}_{m+2} \end{bmatrix} = \mathbf{H}_3^* \begin{bmatrix} \mathbf{c}_0 \\ \mathbf{c}_1 \\ \vdots \\ \mathbf{c}_{m-1} \end{bmatrix},$$

$$\tilde{\mathbf{H}}_3^* = \begin{bmatrix} \mathbf{I}_{m-1} & \mathbf{I}_{m-1} & \mathbf{I}_{m-1} & \cdots & \mathbf{I}_{m-1} \\ \mathbf{I}_{m-1} & \tilde{Q}\mathbf{E}Q & \tilde{Q}\mathbf{E}^2Q & \cdots & \tilde{Q}\mathbf{E}^{m-1}Q \\ \mathbf{I}_{m-1} & \tilde{Q}\mathbf{E}^2Q & \tilde{Q}\mathbf{E}^4Q & \cdots & \tilde{Q}\mathbf{E}^{2(m-1)}Q \end{bmatrix}.$$

根据 EEOD 码编码的代数定义可知, 要证明 EEOD 码具有 MDS 特性, 则就是要证明校验矩阵  $\tilde{\mathbf{H}}_3^*$  的任意子矩阵为可逆矩阵. 下面我们首先给出一个  $(8, 5, 4)$  EEOD 码的实例.

**例 1**  $((8, 5, 4)$  EEOD 码的代数定义编码过程实例). 令  $m=5, p=3$ ,  $(8, 5, 4)$  EEOD 码校验矩阵为  $\tilde{\mathbf{H}}_3^*$ , 则校验列  $\mathbf{c}_5, \mathbf{c}_6, \mathbf{c}_7$  如下计算, 最后得到与表 2 相同的二维阵列码字  $\mathbf{C}$ .

$$\tilde{\mathbf{H}}_3^* = \begin{bmatrix} \mathbf{I}_{m-1} & \mathbf{I}_{m-1} & \mathbf{I}_{m-1} & \mathbf{I}_{m-1} & \mathbf{I}_{m-1} \\ \mathbf{I}_{m-1} & \tilde{Q}\mathbf{E}Q & \tilde{Q}\mathbf{E}^2Q & \tilde{Q}\mathbf{E}^3Q & \tilde{Q}\mathbf{E}^4Q \\ \mathbf{I}_{m-1} & \tilde{Q}\mathbf{E}^2Q & \tilde{Q}\mathbf{E}^4Q & \tilde{Q}\mathbf{E}Q & \tilde{Q}\mathbf{E}^3Q \end{bmatrix},$$

$$\begin{bmatrix} \mathbf{c}_5 \\ \mathbf{c}_6 \\ \mathbf{c}_7 \end{bmatrix} = \tilde{\mathbf{H}}_3^* \begin{bmatrix} \mathbf{c}_0 \\ \mathbf{c}_1 \\ \vdots \\ \mathbf{c}_4 \end{bmatrix}.$$

其中根据定义 1、引理 1 可得矩阵集  $\{\mathbf{I}_5, \mathbf{E}_5, \mathbf{E}_5^2, \mathbf{E}_5^3, \mathbf{E}_5^4\}$ , 由定义 2 可得  $Q, \tilde{Q}$ , 如下所示:

$$\mathbf{I}_5 = \begin{bmatrix} 10000 \\ 01000 \\ 00100 \\ 00010 \\ 00001 \end{bmatrix}, \quad \mathbf{E}_5 = \begin{bmatrix} 00001 \\ 10000 \\ 01000 \\ 00100 \\ 00010 \end{bmatrix}, \quad \mathbf{E}_5^2 = \begin{bmatrix} 00010 \\ 00001 \\ 10000 \\ 01000 \\ 00100 \end{bmatrix},$$

$$\mathbf{E}_5^3 = \begin{bmatrix} 00100 \\ 00010 \\ 00001 \\ 10000 \\ 01000 \end{bmatrix}, \quad \mathbf{E}_5^4 = \begin{bmatrix} 01000 \\ 00100 \\ 00010 \\ 00001 \\ 10000 \end{bmatrix}, \quad Q = \begin{bmatrix} 1000 \\ 0100 \\ 0010 \\ 0001 \\ 0000 \end{bmatrix}, \quad \tilde{Q} = \begin{bmatrix} 10001 \\ 01001 \\ 00101 \\ 00011 \end{bmatrix}$$

因而可得

$$\mathbf{I}_{m-1} = \begin{bmatrix} 1000 \\ 0100 \\ 0010 \\ 0001 \end{bmatrix}, \quad \tilde{Q}\mathbf{E}Q = \begin{bmatrix} 0001 \\ 0101 \\ 0011 \\ 0001 \end{bmatrix}, \quad \tilde{Q}\mathbf{E}^2Q = \begin{bmatrix} 0011 \\ 0010 \\ 1010 \\ 0110 \end{bmatrix},$$

$$\tilde{Q}\mathbf{E}^3Q = \begin{bmatrix} 0110 \\ 0101 \\ 0100 \\ 1100 \end{bmatrix}, \quad \tilde{Q}\mathbf{E}^4Q = \begin{bmatrix} 1100 \\ 1010 \\ 1001 \\ 1000 \end{bmatrix}.$$

### 2.4 EEOD 码 MDS 性质

**引理 2.** 当  $m$  为素数时,  $(m+3, m, 4)$  EEOD 码的最小 Hamming 列距离不等于 3,  $d_{\min} \neq 3$ .

证明. 反证法. 根据编码理论,  $(m+3, m, 4)$  EEOD 码的最小 Hamming 列距离等于 3, 即存在 3 列非零列向量, 设 3 列非零列向量分别为  $0 \leq u_1 < u_2 < u_3 \leq m+2$ , 其余列全为零列向量.

(1) 假设 3 列非零列向量全部是信息列, 那校验列全部为零列向量. 则根据编码规则可得

$$\begin{bmatrix} \mathbf{c}_m \\ \mathbf{c}_{m+1} \\ \mathbf{c}_{m+2} \end{bmatrix} = \begin{bmatrix} \tilde{Q}Q & \tilde{Q}Q & \tilde{Q}Q \\ \tilde{Q}\mathbf{E}^{u_1}Q & \tilde{Q}\mathbf{E}^{u_2}Q & \tilde{Q}\mathbf{E}^{u_3}Q \\ \tilde{Q}\mathbf{E}^{2u_1}Q & \tilde{Q}\mathbf{E}^{2u_2}Q & \tilde{Q}\mathbf{E}^{2u_3}Q \end{bmatrix} \times \begin{bmatrix} \mathbf{c}_{u_1} \\ \mathbf{c}_{u_2} \\ \mathbf{c}_{u_3} \end{bmatrix} = \begin{bmatrix} \mathbf{0}_{m-1} \\ \mathbf{0}_{m-1} \\ \mathbf{0}_{m-1} \end{bmatrix}.$$

根据矩阵的性质可知<sup>[9]</sup>,  $\mathbf{A}\mathbf{X} = \mathbf{0}$  时, 若  $\mathbf{A}$  可逆, 方程没有非零解, 即  $\mathbf{X}$  肯定等于零矩阵. 又因为根据定理 8(见附录)可知,  $(m+3, m, 4)$  EEOD 码校验矩阵任意子矩阵都是可逆矩阵, 所以这与  $\{\mathbf{c}_{u_1}, \mathbf{c}_{u_2}, \mathbf{c}_{u_3}\}$  为非零列相矛盾, 所以假设不成立.

(2) 假设 3 列非零列向量有 1 列是信息列, 2 列是校验列, 其余列为零列向量. 设信息列为  $u_1$  列, 则

根据编码规则可得

$$\begin{bmatrix} \mathbf{c}_m \\ \mathbf{c}_{m+1} \\ \mathbf{c}_{m+2} \end{bmatrix} = \begin{bmatrix} \tilde{Q}Q \\ \tilde{Q}E^{u_1}Q \\ \tilde{Q}E^{2u_1}Q \end{bmatrix} \times [\mathbf{c}_{u_1}] = \begin{bmatrix} \tilde{Q}Q\mathbf{c}_{u_1} \\ \tilde{Q}E^{u_1}Q\mathbf{c}_{u_1} \\ \tilde{Q}E^{2u_1}Q\mathbf{c}_{u_1} \end{bmatrix}.$$

根据假设可知 3 列校验列中,有 1 列是零向量. 而任意矩阵  $\tilde{Q}Q$ 、 $\tilde{Q}E^{u_1}Q$ 、 $\tilde{Q}E^{2u_1}Q$  的秩都是  $m-1$ . 同理根据矩阵的性质可知,  $\mathbf{A}\mathbf{X}=\mathbf{0}$  时,若  $\mathbf{A}$  可逆,方程没有非零解,即  $\mathbf{X}$  肯定等于零. 即  $\mathbf{c}_{u_1}$  为零列向量,但是这与假设相矛盾. 所以假设不成立.

(3) 假设 3 列非零列向量有 2 列是信息列、1 列是校验列,其余列为零列向量. 设非零列信息列为  $u_1, u_2$  列,则根据编码规则可得

$$\begin{bmatrix} \mathbf{c}_m \\ \mathbf{c}_{m+1} \\ \mathbf{c}_{m+2} \end{bmatrix} = \begin{bmatrix} \tilde{Q}Q & \tilde{Q}Q \\ \tilde{Q}E^{u_1}Q & \tilde{Q}E^{u_2}Q \\ \tilde{Q}E^{2u_1}Q & \tilde{Q}E^{2u_2}Q \end{bmatrix} \begin{bmatrix} \mathbf{c}_{u_1} \\ \mathbf{c}_{u_2} \end{bmatrix}.$$

那么  $\mathbf{c}_m, \mathbf{c}_{m+1}, \mathbf{c}_{m+2}$  3 列校验列有两列为零列向量. 则可得下面的公式

$$\begin{bmatrix} \mathbf{c}_m \\ \mathbf{c}_{m+1} \end{bmatrix} = \begin{bmatrix} \tilde{Q}Q & \tilde{Q}Q \\ \tilde{Q}E^{u_1}Q & \tilde{Q}E^{u_2}Q \end{bmatrix} \begin{bmatrix} \mathbf{c}_{u_1} \\ \mathbf{c}_{u_2} \end{bmatrix} = \begin{bmatrix} \mathbf{0}_{m-1} \\ \mathbf{0}_{m-1} \end{bmatrix}$$

或者

$$\begin{bmatrix} \mathbf{c}_m \\ \mathbf{c}_{m+1} \end{bmatrix} = \begin{bmatrix} \tilde{Q}Q & \tilde{Q}Q \\ \tilde{Q}E^{2u_1}Q & \tilde{Q}E^{2u_2}Q \end{bmatrix} \begin{bmatrix} \mathbf{c}_{u_1} \\ \mathbf{c}_{u_2} \end{bmatrix} = \begin{bmatrix} \mathbf{0}_{m-1} \\ \mathbf{0}_{m-1} \end{bmatrix}$$

或者

$$\begin{bmatrix} \mathbf{c}_m \\ \mathbf{c}_{m+1} \end{bmatrix} = \begin{bmatrix} \tilde{Q}E^{u_1}Q & \tilde{Q}E^{u_2}Q \\ \tilde{Q}E^{2u_1}Q & \tilde{Q}E^{2u_2}Q \end{bmatrix} \begin{bmatrix} \mathbf{c}_{u_1} \\ \mathbf{c}_{u_2} \end{bmatrix} = \begin{bmatrix} \mathbf{0}_{m-1} \\ \mathbf{0}_{m-1} \end{bmatrix}.$$

同理根据矩阵的性质可知,  $\mathbf{A}\mathbf{X}=\mathbf{0}$  时,若  $\mathbf{A}$  可逆,方程没有非零解,即  $\mathbf{X}$  肯定等于零. 又因为根据定理 8(见附录)可知,  $(m+3, m, 4)$  EEOD 码任意子矩阵都是可逆矩阵,即  $\{\mathbf{c}_{u_1}, \mathbf{c}_{u_2}\}$  为零列向量,但是这与假设相矛盾,所以假设不成立.

因此根据(1)~(3)可知,  $(m+3, m, 4)$  EEOD 码中不可能出现列的重量为 3 的情况,即最小列距离不可能等于 3. 即  $d_{\min} \neq 3$ . 证毕.

**定理 1.** 当且仅当  $m$  为素数时,  $(m+3, m, 4)$  EEOD 码的最小 Hamming 列距离为 4,  $d_{\min} = 4$ .

**证明.** 根据编码理论,在线性码中,码的最小重量等于码的最小列距离. 要证明  $(m+3, m, 4)$  EEOD 码的最小距离  $d_{\min} = 4$ , 即只要证明  $(m+3, m, 4)$  EEOD 码的最小列重量为 4 即可.

(1) 根据 EVENODD 码<sup>[5]</sup>可知, EVENODD 码的最小列距离为 3. 因此,  $(m+3, m, 4)$  EEOD 码的最小列距  $d_{\min} \geq 3$ .

(2) 根据引理 2,  $(m+3, m, 4)$  EEOD 码的最小列距离不可能等于 3, 即  $d_{\min} \neq 3$ .

因此根据(1), (2), 很容易得到  $(m+3, m, 4)$  EEOD 码的最小列距离  $d_{\min} = 4$ . 证毕.

### 3 EEOD 码译码算法

在前一节可知 EEOD 码具有 MDS 性质, 若预先知道出错列的位置, EEOD 码能够纠正任意小于等于 3 列的错误. EEOD 码是在 EVENODD 码的基础上扩展的, 因此纠正两列删除错的译码算法与 EVENODD 码纠正两列删除错译码算法类似, 在这就不详细讨论了. 在本小节只给出 EEOD 码最复杂的一种译码情况, 出错的三列全部为信息列.

设丢失的信息列分别为  $i, j, k$  三列,  $0 \leq i < j < k \leq m-1$ . 下面给出这种情况的详细译码过程.

首先第 1 步通过  $m, m+1, m+2$  三列校验列计算出调节因子  $S', S_1, S_2$ .

$$S' = \left( \bigoplus_{u=0}^{m-2} c_{u,m} \right), \quad S_1 = S' \oplus \left( \bigoplus_{u=0}^{m-2} c_{u,m+1} \right),$$

$$S_2 = S' \oplus \left( \bigoplus_{u=0}^{m-2} c_{u,m+2} \right).$$

随后则得到只含有  $i, j, k$  三列信息位的校验算子, 分别为  $\tilde{S}^{(0)} = \tilde{S}_0^{(0)}, \tilde{S}_1^{(0)}, \dots, \tilde{S}_{m-1}^{(0)}$ ,  $\tilde{S}^{(1)} = \tilde{S}_0^{(1)}, \tilde{S}_1^{(1)}, \dots, \tilde{S}_{m-1}^{(1)}$ ,  $\tilde{S}^{(2)} = \tilde{S}_0^{(2)}, \tilde{S}_1^{(2)}, \dots, \tilde{S}_{m-1}^{(2)}$ ,  $0 \leq u \leq m-1$ . 计算公式如下:

$$\tilde{S}_u^{(0)} = \bigoplus_{\substack{t=0 \\ t \neq i, j, k}}^m c_{u,t}, \quad \tilde{S}_u^{(1)} = S_1 \oplus c_{u,m+1} \oplus \left( \bigoplus_{\substack{t=0 \\ t \neq i, j, k}}^{m-1} c_{\langle u-t \rangle_m, t} \right),$$

$$\tilde{S}_u^{(2)} = S_2 \oplus c_{u,m+2} \oplus \left( \bigoplus_{\substack{t=0 \\ t \neq i, j, k}}^{m-1} c_{\langle u-2t \rangle_m, t} \right).$$

其译码思想如下: 从解线性方程组的角度, 把校验算子  $\tilde{S}^{(0)}, \tilde{S}^{(1)}, \tilde{S}^{(2)}$  看作一组线性方程组, 其中  $c_{u,i}, c_{u,j}, c_{u,k}$  为未知变量, 译码过程则是解这组线性方程. 而这组方程每个方程中至少含有两个未知变量, 不能直接求解未知变量, 必通过一定的变换, 把原方程组转化为一组只含有  $j$  列未知变量的循环方程组, 记为  $\tilde{S}'_u$  ( $0 \leq u \leq m-2$ ); 并且方程组中每个方程至多只含有两个未知变量, 其中第一个方程  $\tilde{S}'_u$  中的两个未知变量中含有信息位  $c_{m-1,j}$ , 而  $c_{m-1,j} = 0$ , 实际只有一个未知变量. 因此可以通过这组循环方程组依次求解出第  $j$  列的未知变量, 即求解出第  $j$  列的信息位. 则原方程组变换为循环方程组的过程, 称为“消元过程”. 消元过程的目的是依次把第  $i, k$  两列的未知变量消掉, 只剩下第  $j$  列的未知变量. 在给出具体的消元算法之前, 下面先通过一个具体实

例来看 EEOD 码的消元过程。

**例 2** ((8, 5, 4) EEOD 码的消元过程实例). 分两种情况, 若出错的列分别为  $i=1, j=2, k=4$ ,  $j-i \neq k-j$  非对称出错, 则可以通过图 2(a) 所示的方法依次把第 1, 4 列的信息位消掉, 只剩下第 2 列信息位, 并且最终得到至多只含有第 2 列两个未知的变量的一组循环方程组。若出错列分别为  $i=0, j=2, k=4, j-i=k-j$  对称出错, 如图 2(b) 所示。根据图 2(a),  $a \rightarrow b \rightarrow c \rightarrow d \rightarrow a$  称为一条回路。每条回路经过  $i, k$  列的未知变量的度都为 2, 经过  $j$  列的未知变量的度为 1, 一个只含两个未知变量的方程可由两条回路得到。

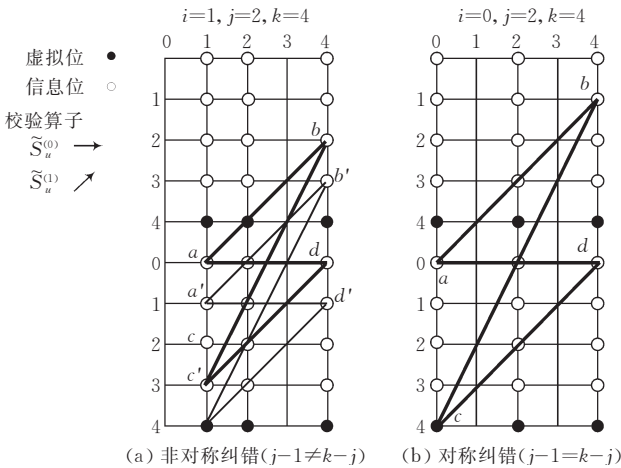


图 2 ((8, 5, 4) EEOD 码译码算法的消元过程

例如  $a \rightarrow b \rightarrow c \rightarrow d \rightarrow a$  回路对应的校验方程的变换为

$$c_{4,2} \oplus c_{1,2} \oplus c_{0,2} \oplus c_{2,2} = \tilde{S}_1^{(1)} \oplus \tilde{S}_0^{(2)} \oplus \tilde{S}_4^{(1)} \oplus \tilde{S}_0^{(0)};$$

$a' \rightarrow b' \rightarrow c' \rightarrow d' \rightarrow a'$  回路则对应的校验方程的变换为

$$c_{0,2} \oplus c_{2,2} \oplus c_{2,2} \oplus c_{3,2} = \tilde{S}_2^{(1)} \oplus \tilde{S}_1^{(2)} \oplus \tilde{S}_0^{(1)} \oplus \tilde{S}_1^{(0)},$$

然后两条回路上的未知变量叠加就可以得到循环方程组的第一个方程:

$$\begin{aligned} \tilde{S}_0' &= c_{4,2} \oplus c_{3,2} \\ &= \tilde{S}_2^{(1)} \oplus \tilde{S}_1^{(2)} \oplus \tilde{S}_0^{(1)} \oplus \tilde{S}_1^{(0)} \oplus \tilde{S}_1^{(1)} \oplus \tilde{S}_0^{(2)} \oplus \tilde{S}_4^{(1)} \oplus \tilde{S}_0^{(0)}. \end{aligned}$$

$$\begin{aligned} &c_{u,j} \oplus d_{\langle u+r \rangle_m, j} \oplus \cdots \oplus d_{\langle u+(l_d-2)r \rangle_m, j} \oplus d_{\langle u+(l_d-1)r \rangle_m, j} \oplus \\ &d_{\langle u+r \rangle_m, j} \oplus d_{\langle u+2r \rangle_m, j} \oplus \cdots \oplus d_{\langle u+(l_d-1)r \rangle_m, j} \oplus c_{\langle u+l_d r \rangle_m, j} \oplus \\ &c_{\langle u+s \rangle_m, j} \oplus d_{\langle u+r+s \rangle_m, j} \oplus \cdots \oplus d_{\langle u+(l_d-2)r+s \rangle_m, j} \oplus d_{\langle u+(l_d-1)r+s \rangle_m, j} \oplus \\ &d_{\langle u+r+s \rangle_m, j} \oplus d_{\langle u+2r+s \rangle_m, j} \oplus \cdots \oplus d_{\langle u+(l_d-1)r+s \rangle_m, j} \oplus c_{\langle u+l_d r+s \rangle_m, j}. \end{aligned}$$

又因为  $\langle s - l_d r \rangle_m = 0$ , 所以  $c_{\langle u+s \rangle_m, j} = c_{\langle u+l_d r \rangle_m, j}$ , 因此

$$\begin{aligned} c_{u,j} + c_{\langle u+2s \rangle_m, j} &= \sum_{v=0}^{l_d-1} (\tilde{S}_{\langle u+vr+j \rangle_m}^{(1)} \oplus \tilde{S}_{\langle u+vr+j+k \rangle_m}^{(2)} \oplus \\ &\tilde{S}_{\langle u+vr+j+r+s \rangle_m}^{(1)} \oplus \tilde{S}_{\langle u+vr+r \rangle_m}^{(0)}), \end{aligned}$$

依照此方法可依次得到至多只含  $j$  列两未知变量的循环方程组:

$$\begin{aligned} \tilde{S}_1' &= c_{3,2} \oplus c_{2,2} \\ &= \tilde{S}_0^{(1)} \oplus \tilde{S}_4^{(2)} \oplus \tilde{S}_3^{(1)} \oplus \tilde{S}_4^{(0)} \oplus \tilde{S}_1^{(1)} \oplus \tilde{S}_0^{(2)} \oplus \tilde{S}_4^{(1)} \oplus \tilde{S}_0^{(0)}, \\ \tilde{S}_2' &= c_{2,2} \oplus c_{1,2} \\ &= \tilde{S}_4^{(1)} \oplus \tilde{S}_3^{(2)} \oplus \tilde{S}_2^{(1)} \oplus \tilde{S}_3^{(0)} \oplus \tilde{S}_0^{(1)} \oplus \tilde{S}_4^{(2)} \oplus \tilde{S}_3^{(1)} \oplus \tilde{S}_4^{(0)}, \\ \tilde{S}_3' &= c_{1,2} \oplus c_{0,2} \\ &= \tilde{S}_3^{(1)} \oplus \tilde{S}_2^{(2)} \oplus \tilde{S}_1^{(1)} \oplus \tilde{S}_2^{(0)} \oplus \tilde{S}_4^{(1)} \oplus \tilde{S}_3^{(2)} \oplus \tilde{S}_2^{(1)} \oplus \tilde{S}_3^{(0)}. \end{aligned}$$

又因为已知  $c_{4,2} = 0$ , 根据上面的循环方程组依次求解出第  $j$  列的信息  $c_{3,2} \rightarrow c_{2,2} \rightarrow c_{1,2} \rightarrow c_{0,2}$ , 再根据 EVENODD 码纠双列译码算法<sup>[5]</sup>, 依次恢复出  $i, k$  两列的信息。

下面给出理论证明: 通过如实例 2 中的消元过程, EEOD 码能够得到一组只含有某一列的未知变量的循环方程组, 并且该循环方程组的每个方程至多只有两个变量, 其中有一个方程只含有一个未知变量。

**引理 3.** 在  $(m+3, m, 4)$  EEOD 码中, 若丢失的数据列为  $0 \leq i < j < k \leq m-1$ , 令  $r = j - i$ ,  $s = k - j$ , 则一定存在  $l_d (1 \leq l_d < m)$ , 满足  $\langle s - l_d r \rangle_m = 0$ , 使得

$$\begin{aligned} c_{u,j} \oplus c_{\langle u+2s \rangle_m, j} &= \sum_{v=0}^{l_d-1} (\tilde{S}_{\langle u+vr+j \rangle_m}^{(1)} \oplus \tilde{S}_{\langle u+vr+j+k \rangle_m}^{(2)} \oplus \\ &\tilde{S}_{\langle u+vr+j+r+s \rangle_m}^{(1)} \oplus \tilde{S}_{\langle u+vr+r \rangle_m}^{(0)}). \end{aligned}$$

证明. 根据校验子计算公式, 对  $0 \leq u \leq m-1$ :

$$\begin{aligned} \tilde{S}_{\langle u+j \rangle_m}^{(1)} &= c_{\langle u+r \rangle_m, i} \oplus c_{u,j} \oplus c_{\langle u-s \rangle_m, k}, \\ \tilde{S}_{\langle u+j+k \rangle_m}^{(2)} &= c_{\langle u+s+2r \rangle_m, i} \oplus c_{\langle u+s \rangle_m, j} \oplus c_{\langle u-s \rangle_m, k}, \\ \tilde{S}_{\langle u+j+r+s \rangle_m}^{(1)} &= c_{\langle u+s+2r \rangle_m, i} \oplus c_{\langle u+s+r \rangle_m, j} \oplus c_{\langle u+r \rangle_m, k}, \\ \tilde{S}_{\langle u+r \rangle_m}^{(0)} &= c_{\langle u+r \rangle_m, i} \oplus c_{\langle u+r \rangle_m, j} \oplus c_{\langle u+r \rangle_m, k}. \end{aligned}$$

因此可得

$$\begin{aligned} c_{u,j} \oplus c_{\langle u+r \rangle_m, j} \oplus c_{\langle u+s \rangle_m, j} \oplus c_{\langle u+s+r \rangle_m, j} &= \\ \tilde{S}_{\langle u+j \rangle_m}^{(1)} \oplus \tilde{S}_{\langle u+j+k \rangle_m}^{(2)} \oplus \tilde{S}_{\langle u+j+r+s \rangle_m}^{(1)} \oplus \tilde{S}_{\langle u+r \rangle_m}^{(0)}. \end{aligned}$$

$$\text{因此 } \sum_{v=0}^{l_d-1} (\tilde{S}_{\langle u+vr+j \rangle_m}^{(1)} \oplus \tilde{S}_{\langle u+vr+j+k \rangle_m}^{(2)} \oplus \tilde{S}_{\langle u+vr+j+r+s \rangle_m}^{(1)} \oplus \tilde{S}_{\langle u+vr+r \rangle_m}^{(0)}) \text{ 等于}$$

$$\begin{aligned} &c_{u,j} \oplus d_{\langle u+r \rangle_m, j} \oplus \cdots \oplus d_{\langle u+(l_d-2)r \rangle_m, j} \oplus d_{\langle u+(l_d-1)r \rangle_m, j} \oplus \\ &d_{\langle u+r \rangle_m, j} \oplus d_{\langle u+2r \rangle_m, j} \oplus \cdots \oplus d_{\langle u+(l_d-1)r \rangle_m, j} \oplus c_{\langle u+l_d r \rangle_m, j} \oplus \\ &c_{\langle u+s \rangle_m, j} \oplus d_{\langle u+r+s \rangle_m, j} \oplus \cdots \oplus d_{\langle u+(l_d-2)r+s \rangle_m, j} \oplus d_{\langle u+(l_d-1)r+s \rangle_m, j} \oplus \\ &d_{\langle u+r+s \rangle_m, j} \oplus d_{\langle u+2r+s \rangle_m, j} \oplus \cdots \oplus d_{\langle u+(l_d-1)r+s \rangle_m, j} \oplus c_{\langle u+l_d r+s \rangle_m, j}. \end{aligned}$$

并且由于  $1 \leq s = k - j \leq m-1$ , 因此  $u \neq \langle u+2s \rangle_m$ ,  $c_{u,j}$  与  $c_{\langle u+2s \rangle_m, j}$  肯定是  $j$  列中两个不同未知变量。

证毕。

注. 引理 3 中的公式给出了  $(m+3, m, 4)$  EEOD



码的原方程组转换位循环方程组的方法的计算公式, 消元之后的循环方程组只含一列的未知变量, 并且每个方程至多两个变量, 其中一个方程只有一个未知变量。

若丢失三列数据信息列  $i, j, k$  中, 若  $r=j-i=k-j=s$ , 则  $l_d=1$ , 则只需要一条回路就能得到循环方程组。如图 2(b) 图所示,  $a \rightarrow b \rightarrow c \rightarrow d \rightarrow a$  一条回路就可以得到只含有  $j$  列的两个未知变量的循环方程组的一个方程。

由于文章篇幅的关系, 在这就不再对三列出错的另外几种情况做讨论了。

## 4 EEOD 码性能分析

在这一节, 我们从存储效率(storage efficiency)、编译码的复杂度(encoding and decoding complexity)、更新复杂度(update complexity) 3 个方面来衡量 EEOD 码的性能。并与其它 Blaum 码、RS 码进行比较。

### 4.1 EEOD 码存储效率和编译码复杂度

编码存储效率  $E$  为一个码字的信息位所占的存储空间与整个码字所占的存储空间之比。对编码来说, 信息位为  $m$ , 最小列距离为 4 的码的最佳存储效率为  $m/(m+3)$ 。根据存储效率  $E$  的定义可得 EEOD 码的存储效率为  $E=m/(m+3)$ , 达到了列距离为 4 的最佳的存储效率。

EEOD 码与其它阵列码一样, 编译码过程只需要异或运算, 因此把每个码字编码总的异或次数与码字信息位的总比特位之比定义为码的编译复杂度。根据第 2 节中 EEOD 码的编码特性可知, EEOD 码的总的异或次数为 3 列校验列的异或次数。设每个信息位为  $b$  bits, 水平校验列需要  $b(m-1)^2$  异或操作; 竖直校验列每列总的需要  $b(m-1)^2 + b(m-2)$  次。EEOD 码整个编码过程需要的总异或次数为  $b(m-1)^2 + 2(b(m-1)^2 + b(m-2))$ , 而 EEOD 码总的信息位比特数为  $(m-1)mb$ , 则 EEOD 的编码复杂度, 即每个比特需要的异或次数为

$$Encoding - eff_{(m+3, m)EEOD} = 3 - \frac{m+1}{(m-1)m}.$$

随着  $m$  值的增大, EEOD 码的每比特信息位需要异或的次数接近 3。

从第 3 节给出的 EEOD 码的译码过程可知, 译码复杂度的计算要比编码复杂度计算复杂。在 EEOD 码译码过程中, 丢失的列全部为信息列的译码

算法是最复杂的一种情况, 因此在本节中只讨论丢失的列全部为信息列的译码复杂度。EEOD 码纠三列信息列译码算法, 第 1 步计算两列垂直校验列的共同调节因子, 计算  $S'$  需要  $m-2$  次异或运算, 则  $S_1$  和  $S_2$  分别需要  $m-1$  次异或运算, 总共需要  $2(m-1) + m-2$  次异或运算, 接下来计算  $\tilde{S}^{(0)}, \tilde{S}^{(1)}, \tilde{S}^{(2)}$  三列校验算子需要  $3(m-3)(m-1) + 2(m-1)$ 。然后采用消元算法恢复第  $j$  列的信息位需要的异或次数为  $(4l_d-1)(m-1) + (m-2)$ , 最后恢复其余两列的信息位需要的异或次数为  $4(m-1)-1$ , EEOD 码纠三列信息列译码过程总的译码大约次数为  $(4l_d-2+3m)(m-1)-3$ 。因此译码复杂度为

$$Decoding - eff_{EEOD} = 3 + \frac{(4l_d-2)(m-1)-3}{m(m-1)}.$$

下面把 EEOD 码与其 Blaum 码以及文献[14]中基于 XOR 的复损码进行比较。EEOD 码的译码复杂度依靠具体丢失信息列的位置来确定。为了便于分析,  $l_d$  取平均值, 因此通过计算译码过程总的异或次数与总信息比特数之比来比较。设码的信息列为  $m$  列, 校验列为 3 列。Blaum 码纠三列信息列译码总的异或次数为  $(3m+21)(m-1)$ [8]。Bloemer 提出的基于 XOR 的复损码译码过程中需要的异或总的次数为  $krL^2$ , 有限域的操作为  $r^2$  (其中  $k, r$  分别表示码的信息位和冗余校验位的长度,  $L$  表示有限域的大小  $GF(2^L)$ )。为了比较方便, 忽略有限域的操作  $r^2$  (实际有限域的计算对译码的速度有影响)[14], 而信息位的比特数为  $kL$ , 因此  $r=3$ , 基于 XOR 的复损码的译码复杂度为  $3L$ , 只与有限域的大小相关了。

从图 3 中 Blaum 码、EEOD 码、RS 码的译码复杂度比较图可以看出。码的长度比较短时(信息列的列数), EEOD 码的每个信息位需要的异或次数最少, 译码性能最好。而 RS 类纠错码译码复杂度最高, 并且随着域的扩大而增加。而随着码长的变大, EEOD 码每个信息位需要的异或次数逐渐接近 4, 而 Blaum 码逐渐接近 3。Blaum 码单从译码复杂度考虑要优于 EEOD 码。但是 Blaum 码的译码算法实现比较复杂, 不易于软件实现。

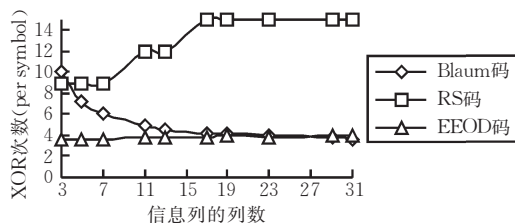


图 3 Blaum 码、RS 码、EEOD 码译码复杂度比较图

## 4.2 EEOD 码更新复杂度

在文献[6]编码的性能分析中指出,阵列纠删码的更新复杂度是衡量码的编码特性的另外的一个重要参数. EEOD 码中,若更新信息位不是调节因子,则更新一个信息位,只需要更新 3 个相应的校验位;若更新的信息位用于计算调节因子,那么更新一个信息位,则相应的应该更新调节因子,而且还需要更新调节因子用于计算的校验列的所有校验位,需要更新  $m$  次. 因此更新一信息位,需要更新校验位的平均次数为  $4 - 3/m$ ,从上式可以知道随着  $m$  值的增大,其更新复杂度接近为 4.

## 5 结 论

独立磁盘冗余阵列(RAID)的一个重要特征是其可靠性. 给出一种有高可靠性、高吞吐量、好的 I/O 性能以及简单的编码和解码算法的磁盘阵列数据布局方案具有重要的实际意义和理论研究价值. 本文提出了 EEOD 码,大幅度提高了 RAID 结构的容灾能力,而需要的额外开销却非常小,并且由于 EVENODD 广泛应用于现有的 RAID 结构中,而 EEOD 是它的扩展码,因此 EEOD 易于在现有的 RAID 结构中进行实现,具有非常好的实用价值.

## 参 考 文 献

[1] Frolund S, Merchant A, Saito Y, Spence S, Veitch A. FAB: Enterprise storage systems on a shoestring//Proceedings of the 9th Workshop on HotOS-IX. Kauai, HI, 2003

[2] Xin Q, Miller E L, Schwarz T J. Reliability mechanisms for very large storage systems//Proceedings of the 20th IEEE/11th NASA Goddard Conference on Mass Storage Systems and Technologies. 2003; 146-156

[3] Patterson D A, Gibson G A, Katz R H. A case for redundant arrays of inexpensive disks(RAID)//Proceedings of the

ACM SIGMOD Conference Proceeding. 1988; 109-116

[4] Hellerstein L, Gibson G A, Karp R M, Patterson D A. Coding techniques for handling failures in large disk arrays. Algorithmica, 1994, 12(3-4): 182-208

[5] Blaum M, Brady J, Bruck J, Menon J. EVENODD: An efficient scheme for tolerating double disk failures in RAID architectures. IEEE Transactions on Computers, 1995, 44(2): 192-202

[6] Xu L, Bruck J. X-code: MDS array codes with optimal encoding. IEEE Transactions on Information Theory, 1999, 45(1): 272-276

[7] Xu L, Bohossian V, Bruck J, Wagner D G. Low density MDS codes and factors of complete graphs. IEEE Transactions on Information Theory, 1999, 45(6): 1817-1826

[8] Blaum M, Bruck J, Vardy A. MDS array codes with independent parity symbols. IEEE Transactions on Information Theory, 1996, 42: 529-542

[9] Hafner J L. HoVer erasure codes for disk arrays. IBM Research Division, Research Report RJ10352 (A0507-015), July, 2005

[10] Hafner J L. WEAVER codes: Highly fault tolerant erasure codes for storage systems//Proceedings of the FAST-2005; 4th Usenix Conference on File and Storage Technologies. December, 2005

[11] Tau Chih-Shing, Wang Tzone-I. Efficient parity placement schemes for tolerating triple disk failures in RAID architectures//Proceedings of the 17th International Conference on Advanced Information Networking and Applications (AINA'03). Xi'an, China, 2003

[12] Feng G-L, Deng R, Bao F, Shen J-C. New efficient MDS array codes for RAID, Part I: Reed solomon like codes for tolerating three disk failures. IEEE Transactions on Computers, 2005, 54(9): 1071-1080

[13] Feng G-L, Deng R, Bao R, Shen J-C. New efficient MDS array codes for RAID, Part II: Rabin-Like Codes for tolerating multiple (4) disk failures. IEEE Transactions on Computers, 2005, 54(12): 1473-1482

[14] Bloemer J M, Kalfane M, Karpinski R. An XOR-based erasure-resilient coding scheme. ICSI: Technical Report ICSI TR-95-048, 1995

## 附录

定义 4(分块矩阵  $\tilde{A}_p$ ). 分块矩阵  $\tilde{A}_p$  是由  $p \times p$  个分块矩阵构成的,每个分块矩阵是一个  $(m-1) \times m$  大小的小矩阵,  $0_{(m-1) \times m}$  是一个  $(m-1) \times m$  大小的零矩阵,  $\tilde{A}_p$  定义如下:

$$\tilde{A}_p = \begin{bmatrix} \tilde{Q} & 0_{(m-1) \times m} & 0_{(m-1) \times m} & \cdots & 0_{(m-1) \times m} \\ 0_{(m-1) \times m} & \tilde{Q} & 0_{(m-1) \times m} & \cdots & 0_{(m-1) \times m} \\ 0_{(m-1) \times m} & 0_{(m-1) \times m} & \tilde{Q} & \cdots & 0_{(m-1) \times m} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0_{(m-1) \times m} & 0_{(m-1) \times m} & 0_{(m-1) \times m} & \cdots & \tilde{Q} \end{bmatrix}.$$

定义 5(分块矩阵  $\tilde{H}_{p \times m}$ ). 分块矩阵  $\tilde{H}_{p \times m}$  是由  $p \times m$  个小矩阵构成,每个小矩阵大小为  $(m-1) \times m$ ,  $\tilde{H}_{p \times m}$  定义如下:

$$\tilde{H}_{p \times m} = \begin{bmatrix} Q & Q & Q & \cdots & Q \\ Q & EQ & E^2 Q & \cdots & E^{m-1} Q \\ Q & E^2 Q & E^4 Q & \cdots & E^{2(m-1)} Q \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ Q & E^{p-1} Q & E^{2(p-1)} Q & \cdots & E^{(p-1)(m-1)} Q \end{bmatrix}.$$



**引理 4.** 矩阵  $\tilde{M}' = \left( \prod_{j=1}^t (I + E^{\mu_j}) \right) Q$  的秩为  $m-1$ ,  $\mu_i \neq 0^{[12]}$ .

**引理 5.** 设  $B$  是一个大小  $m \times n$  的矩阵, 其中  $n \geq m-1$ ,

用  $B_0, B_1, \dots, B_{m-1}$  表示  $B$  的行向量, 于是  $B = \begin{bmatrix} B_0 \\ B_1 \\ \vdots \\ B_{m-1} \end{bmatrix}$ , 若秩

$(B) = m-1$ , 则秩  $(\tilde{Q}B) = m-1$ , 为满秩矩阵.

证明.  $B_0, B_1, \dots, B_{m-1}$  表示  $B$  的行向量, 由矩阵乘法可知,  $\tilde{Q}B$  行向量分别为  $B_0 + B_{m-1}, B_1 + B_{m-1}, \dots, B_{m-1} + B_{m-1}$ . 即  $\tilde{Q}B$  的行向量组可由  $B$  的行向量线性表示.

反证法. 假设矩阵  $\tilde{Q}B$  的秩不是  $m-1$ , 即小于  $m-1$ , 那么一定存在一组不全为零的系数  $k_0, k_1, \dots, k_{m-2}$  使得下式成立:

$$k_0 B_0 + k_1 B_1 + \dots + k_{m-2} B_{m-2} + (k_0 + k_1 + \dots + k_{m-2}) B_{m-1} = 0.$$

又因为矩阵  $B$  的秩为  $m-1$ , 任意  $m-1$  行是线性无关的. 因此要使得不全为零的系数  $k_0, k_1, \dots, k_{m-2}$  使得上式成立, 则必须至少有  $m$  个系数全为 1 才成立. 必须  $k_0 = k_1 = \dots = k_{m-2} = 1$  和  $k_0 + k_1 + \dots + k_{m-2} = 1$ . 又因为  $k_0 = k_1 = \dots = k_{m-2} = 1$ , 使得  $k_0 + k_1 + \dots + k_{m-2} = 0$ , 因此不存在  $m$  个系数全为 1 的情况. 这与假设得到的结论相矛盾, 因此假设不成立, 即证秩  $(\tilde{Q}B) = m-1$ , 为满秩矩阵. 证毕.

**引理 6.** 设分块矩阵  $B$  是一个大小为  $p \times n_1$  的矩阵, 每个小矩阵大小为  $m \times n_2$  的矩阵, 其中  $n_1 \geq m-1, n_2 \geq m-1$ . 用  $B_{i,j}$  来表示小矩阵, 于是  $B$  矩阵如下:

$$B = \begin{bmatrix} B_{0,0} & B_{0,1} & \cdots & B_{0,n_1} \\ B_{1,0} & B_{1,1} & \cdots & B_{1,n_1} \\ \vdots & \vdots & \ddots & \vdots \\ B_{p-1,0} & B_{p-1,1} & \cdots & B_{p-1,n_1} \end{bmatrix}.$$

若秩  $(B_{i,j}) = m-1$ , 秩  $(B) = p(m-1)$ ,  $D = (\tilde{A}_p B)$ , 则秩  $(D) = p(m-1)$ .

证明. 令  $b_{0,0}, b_{1,0}, \dots, b_{m-1,0}, b_{0,1}, b_{1,1}, \dots, b_{m-1,1}, \dots, b_{0,p-1}, b_{1,p-1}, \dots, b_{m-1,p-1}$  表示  $B$  的行向量, 令  $d_{0,0}, d_{1,0}, \dots, d_{m-2,0}, d_{0,1}, d_{1,1}, \dots, d_{m-2,1}, \dots, d_{0,p-1}, d_{1,p-1}, \dots, d_{m-2,p-1}$  表示  $D = (\tilde{A}_p B)$  的行向量, 于是

$$d_{i,j} = b_{i,j} + b_{m-1,j}.$$

因而要证明  $D$  的秩  $(D) = p(m-1)$ , 即要证明  $D$  的行向量线性无关. 即要证明只有  $\forall i, j, k_{i,j} = 0, 0 \leq i \leq m-2, 0 \leq j \leq p-1$ , 才能使得下面式子成立.

$$\begin{aligned} & k_{0,0} d_{0,0} + k_{1,0} d_{1,0} + \dots + k_{m-2,0} d_{m-2,0} + k_{0,1} d_{0,1} + k_{1,1} d_{1,1} + \dots + k_{m-2,1} d_{m-2,1} + \dots + k_{0,p-1} d_{0,p-1} + k_{1,p-1} d_{1,p-1} + \dots + k_{m-2,p-1} d_{m-2,p-1} \\ &= k_{0,0} b_{0,0} + k_{1,0} b_{1,0} + \dots + k_{m-2,0} b_{m-2,0} + (k_{0,0} + k_{1,0} + \dots + k_{m-2,0}) b_{m-1,0} + k_{0,1} b_{0,1} + k_{1,1} b_{1,1} + \dots + k_{m-2,1} b_{m-2,1} + (k_{0,1} + k_{1,1} + \dots + k_{m-2,1}) b_{m-1,1} + \dots + k_{0,p-1} b_{0,p-1} + k_{1,p-1} b_{1,p-1} + \dots + k_{m-2,p-1} b_{m-2,p-1} + (k_{0,p-1} + k_{1,p-1} + \dots + k_{m-2,p-1}) b_{m-1,p-1}. \end{aligned}$$

反证法. 假设矩阵  $D$  的秩  $(D) < p(m-1)$ , 即存在一组

不全为 0 的  $k_{i,j_1}$  使得上式成立, 其中  $0 \leq i_1 \leq m-2, 0 \leq j_1 \leq p-1$ . 又因为秩  $(B) = p(m-1)$ , 分块矩阵  $B$  中任意  $p(m-1)$  个行向量线性无关, 因此假设要成立, 则上式公式中  $k_{i_1,j_1}$  系数至少有  $p(m-1)+1$  个等于 1 才能成立. 因此至少存在一组系数  $k_{0,i} = k_{1,i} = \dots = k_{m-2,i} = 1$ , 而又因为  $m-1$  为偶数, 因此系数不可能存在  $p(m-1)+1$  个系数为 1 的情况, 这与假设推出的结论相矛盾. 因此假设不成立, 因此秩  $(D) = p(m-1)$ . 证毕.

**定理 6**<sup>[12]</sup>. 由  $x_1, x_2, \dots, x_r$  列组成子矩阵  $\tilde{H}_r$  为一个满秩矩阵, 任意  $r$  列是线性独立, 秩  $(\tilde{H}_r) = r(m-1)$ . 则  $x_1, x_2, \dots, x_r, 0 \leq x_i \leq m, \tilde{H}_r$  如下所示:

$$\tilde{H}_r = \begin{bmatrix} E^{x_1} Q & E^{x_2} Q & E^{x_3} Q & \cdots & E^{x_r} Q \\ E^{2x_1} Q & E^{2x_2} Q & E^{2x_3} Q & \cdots & E^{2x_r} Q \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ E^{rx_1} Q & E^{rx_2} Q & E^{rx_3} Q & \cdots & E^{rx_r} Q \end{bmatrix}.$$

**定理 7.** 由  $x_1, x_2, \dots, x_r$  列组成的子矩阵  $\tilde{H}_r^* = \tilde{A}_r \times \tilde{H}_r$  为一个满秩矩阵, 任意  $r$  列是线性独立. 秩  $(\tilde{H}_r^*) = r(m-1)$ .

$$\tilde{H}_r^* = \begin{bmatrix} \tilde{Q} E^{x_1} Q & \tilde{Q} E^{x_2} Q & \tilde{Q} E^{x_3} Q & \cdots & \tilde{Q} E^{x_r} Q \\ \tilde{Q} E^{2x_1} Q & \tilde{Q} E^{2x_2} Q & \tilde{Q} E^{2x_3} Q & \cdots & \tilde{Q} E^{2x_r} Q \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \tilde{Q} E^{rx_1} Q & \tilde{Q} E^{rx_2} Q & \tilde{Q} E^{rx_3} Q & \cdots & \tilde{Q} E^{rx_r} Q \end{bmatrix}.$$

根据定理 6 可知, 矩阵  $\tilde{H}_r$  的秩为  $r(m-1)$ , 再根据引理 5, 秩为  $r(m-1)$  的矩阵左乘一个  $\tilde{A}_r$  矩阵, 其秩仍然为  $r(m-1)$ , 因此矩阵  $\tilde{H}_r^*$  的秩为  $r(m-1)$ , 即为满秩矩阵, 任意  $r$  列线性无关. 证毕.

**定理 8.** 若  $p=3$  时, 令分块矩阵  $\tilde{H}_{3 \times m}^* = \tilde{A}_3 \times \tilde{H}_3$ , 则  $\tilde{H}_{3 \times m}^*$  中, 任意  $r$  阶 ( $1 \leq r \leq 3$ ) 的子方矩阵的秩为  $r(m-1)$ .

证明. 若  $p=3$ , 首先根据定义  $\tilde{H}_{3 \times m}^*$  分块矩阵如下所示:

$$\tilde{H}_{3 \times m}^* = \begin{bmatrix} \tilde{Q} Q & \tilde{Q} Q & \tilde{Q} Q & \cdots & \tilde{Q} Q \\ \tilde{Q} E^{\mu_1} Q & \tilde{Q} E^{\mu_2} Q & \tilde{Q} E^{\mu_3} Q & \cdots & \tilde{Q} E^{\mu_m} Q \\ \tilde{Q} E^{2\mu_1} Q & \tilde{Q} E^{2\mu_2} Q & \tilde{Q} E^{2\mu_3} Q & \cdots & \tilde{Q} E^{2\mu_m} Q \end{bmatrix}.$$

(1) 若  $r=1$ ,  $\tilde{H}_{3 \times m}^*$  1 阶的子矩阵为  $\tilde{Q} E^i Q$ . 由引理 5 可知,  $\tilde{Q}$  右乘以一个秩为  $m-1$  的矩阵, 得到的新的矩阵的秩仍然为  $m-1$ . 根据  $E$  和  $Q$  两矩阵的定义, 很容易可知  $E^i Q$  的秩为  $m-1$ , 因此可以得到子矩阵  $\tilde{Q} E^i Q$  的秩为  $m-1$ .

(2) 若  $r=2$ , 则  $\tilde{H}_{3 \times m}^*$  中 2 阶构成的子矩阵  $\tilde{H}_2$  如下:

$$\tilde{H}_2 = \begin{bmatrix} \tilde{Q} Q & \tilde{Q} Q \\ \tilde{Q} E^{x_1} Q & \tilde{Q} E^{x_2} Q \end{bmatrix} \text{ 或者 } \begin{bmatrix} \tilde{Q} E^{x_1} Q & \tilde{Q} E^{x_2} Q \\ \tilde{Q} E^{2x_1} Q & \tilde{Q} E^{2x_2} Q \end{bmatrix}.$$

根据定理 7, 可知  $\tilde{H}_2$  的秩为  $2(m-1)$ .

$\tilde{H}_{3 \times m}^*$  中的两阶构成的子矩阵  $\tilde{H}_2$  另外一种情况如下:

$$\begin{aligned} \tilde{H}_2 &= \begin{bmatrix} \tilde{Q} Q & \tilde{Q} Q \\ \tilde{Q} E^{2x_1} Q & \tilde{Q} E^{2x_2} Q \end{bmatrix} \\ &= \begin{bmatrix} \tilde{Q} & \mathbf{0}_{(m-1) \times m} \\ \mathbf{0}_{(m-1) \times m} & \tilde{Q} \end{bmatrix} \times \begin{bmatrix} Q & Q \\ E^{2x_1} Q & E^{2x_2} Q \end{bmatrix}. \end{aligned}$$

因此根据引理 6, 要证明上面的子矩阵  $\tilde{H}_2$  的秩为  $2(m-1)$ , 只需要证明下面的矩阵的秩为  $2(m-1)$  即可:

$$\begin{bmatrix} Q & Q \\ E^{2x_1}Q & E^{2x_2}Q \end{bmatrix}.$$

上面矩阵左乘以矩阵 $\begin{bmatrix} I & 0 \\ E^{2x_1} & I \end{bmatrix}$ ,可以得到

$$\begin{bmatrix} Q & Q \\ 0 & (E^{x_1} + E^{x_2})(E^{x_1} + E^{x_2})Q \end{bmatrix}.$$

根据引理 3 可知,  $(E^{x_1} + E^{x_2})(E^{x_1} + E^{x_2})Q$  的秩为  $(m-1)$ , 而  $Q$  的秩也为  $(m-1)$ , 因此可知矩阵的秩为

$2(m-1)$ , 即证  $\tilde{H}_2$  的秩为  $2(m-1)$ .

(3) 若  $r=3$ , 得到 3 阶的子矩阵  $\tilde{H}_3$  如下:

$$\tilde{H}_{3 \times m}^* = \begin{bmatrix} \tilde{Q}Q & \tilde{Q}Q & \tilde{Q}Q \\ \tilde{Q}E^{x_1}Q & \tilde{Q}E^{x_2}Q & \tilde{Q}E^{x_3}Q \\ \tilde{Q}E^{2x_1}Q & \tilde{Q}E^{2x_2}Q & \tilde{Q}E^{2x_3}Q \end{bmatrix}.$$

根据定理 6 可知  $\tilde{H}_3$  的秩为  $3(m-1)$ .

因此通过上面 3 步, 可知  $\tilde{H}_{3 \times m}^*$  中, 任意  $r$  阶  $(1 \leq r \leq 3)$  的子方矩阵的秩为  $r(m-1)$ . 证毕.



**WAN Wu-Nan**, born in 1978, Ph. D., lecturer. Her research interests include information security and codes theory.

**WU Zhen**, born in 1975, lecturer. His research interests include information security and 3G community security.

**CHEN Yun**, born in 1958, professor. Her research interests include information security and 3G community security.

**WANG Xiao-Jing**, born in 1955, professor, Ph. D. supervisor. His research interests include information security and codes theory.

Background

This research is supported by the National Basic Research Program (973 Program) of China under grant No. 2004CB318003 and SiChuan Science and Technology Tackle Key Problem Program under grant No. 03GG0326 and KYTZ200706. As with computing, fault tolerance (or reliability) is increasingly important in storage systems. Some critical data should be available and some services should be provided even when faults occur in storage units. The main purpose of the research is to study to improve high availabili-

ty and high reliability of large scale storage systems, especially on Disaster Tolerance and Fault Tolerance of storage systems based on RAID technology. To improve the availability of user data, the authors propose a class of new binary Maximum Distance Separable (MDS) array codes called EEOD-Code. EEOD-Code is 3-erasure-correcting code which can provide a much longer nonstop functioning time to a distributed storage system.