May. 2015

DOI: 10. 13718/j. cnki. xdzk. 2015. 05. 016

二进制系统循环码的一个简化查表译码算法®

刘 旭, 包小敏, 武登杰, 李 梅, 袁治华

西南大学 数学与统计学院,重庆 400715

摘要:利用线性循环码的性质和伴随式的重量,给出了一个二进制系统循环码的查表译码算法.本算法具有效率高,信息存储量少的特点.

关 键 词:循环码;查表译码;错误模式;伴随式;Hamming 重量

中图分类号: TN911

文献标志码: A

文章编号: 1673-9868(2015)05-0102-06

本文采用文献[1]中的术语和符号.

域 F 上的一个(n,k) 线性码 \mathcal{C} 是 F 上的长为 n 的向量组成的一个 k 维向量空间. \mathcal{C} 中的向量称为码字. 设 $\mathbf{c} \in \mathcal{C}$, \mathbf{c} 的非零分量的个数称为 \mathbf{c} 的 Hamming 重量,记为 $\mathbf{w}(\mathbf{c})$. \mathcal{C} 中两个码字差的重量称为这两个码字的 Hamming 距离. \mathcal{C} 中任意两个码字的距离组成的集合中的最小值称为 \mathcal{C} 的最小距离,用 d 表示. 此时我们也称 \mathcal{C} 是一个(n,k,d) 码. \mathcal{C} 的纠错能力 $\mathbf{t} = \left\lfloor \frac{d-1}{2} \right\rfloor$. 以 \mathcal{C} 中 \mathbf{k} 个线性无关的码字作成的 $\mathbf{k} \times n$ 阶矩阵 称为 \mathcal{C} 的一个生成矩阵. \mathcal{C} 的对偶码 \mathcal{C}^{\perp} 的生成矩阵称为 \mathcal{C} 的一个一致校验矩阵. 设 \mathbf{H} 是 \mathcal{C} 的一个一致校验矩阵, \mathbf{r} 是一 \mathbf{n} 维向量, \mathbf{r} \mathbf{H}^{T} 称为 \mathbf{r} 的伴随式. \mathbf{r} 是码字当且仅当 \mathbf{r} $\mathbf{H}^{\mathrm{T}} = \mathbf{0}$.

线性码的译码分为 3 步: 首先计算伴随式,其次将伴随式与一个错误模式相联系,最后利用这个错误模式进行纠错. 将伴随式与错误模式相联系的最简单也是最直接的方法就是构造一个伴随式与错误模式的一一对应表,其主要思想是将纠错能力范围内的所有错误模式及对应的伴随式,按一个错误模式和与之对应的伴随式作为一行的格式作成一个译码表. 当收到一个向量 r 后,在这个表中找到其伴随式 r H T ,然后将与之对应的错误模式作为错误向量 e ,输出 r -e . 这种译码方式称为伴随式译码或查表译码. 同其它译码方法相比,伴随式译码具有结构简单,容易实现的优点. 在信息协调协议中伴随式译码算法也得到了应用 [2] . 但当 n 很大时,伴随式译码的译码表会很大,因此占用的存储空间也会很大. 实际上,当纠错能力为 t 时,译码表有 N 行、n + (n-k) = 2n-k 列,其中

$$N = \binom{n}{1} + \binom{n}{2} + \dots + \binom{n}{t}$$

若对任意 $\mathbf{c} = [c_0, \dots, c_{n-2}, c_{n-1}] \in \mathcal{C}$ 都有其右循环移位 $\mathbf{c}^R = [c_{n-1}, c_0, \dots, c_{n-2}] \in \mathcal{C}$,则称 \mathcal{C} 是一个循环码. 作为特殊的线性码,循环码的译码也分为 3 步. 但由于其循环特性,循环码的编码和译码都可通过电路来实现. Meggitt 译码器是循环码译码的一般方法 [3],它主要由 3 部分组成 [4]:伴随式寄存器 (syndrome register)、错误模式检测器 (error — pattern detector)、存放接收向量的缓存器 (buffer register to hold the received vector). 错误模式检测器是通过组合逻辑电路来实现查表的过程,这也是 Meggitt 译码器最复杂的部分. 随着码长 n 和纠错能力 t 的增加,其组合逻辑电路部分会变得很复杂,甚至难以实现. 因此简化译

① 收稿日期: 2014-06-09

基金项目: 国家自然科学基金项目(11471265).

作者简介:刘 旭(1989-),男,河南商丘人,硕士研究生,主要从事编码理论,密码学的研究.

码表就成了简化组合逻辑电路的一个突破口. 一个长为n 的非零向量通过循环可以得到n 个不同的向量,因此在循环码的译码表中,n 个通过循环得到的错误模式可以只列出一个,从而将 N 减少为 $\frac{N}{n}$; 文献 [6-9] 利用个别循环码((23, 12, 7)Golay 码, (15, 5, 7), (31, 16, 7) 二进制 BCH 码和(47, 24, 11) 二进制 QR 码) 的特性和特殊结构,将这几个码的译码表从列出所有可能纠正的错误模式减少成只列出错误出现在信息部分的错误模式. i 个错误出现在信息部分的k 比特中,其可能性有 $\binom{k}{i}$ 种,因此当错误个数不超过 t

且错误只出现在信息部分时,错误总数为 $N' = \sum_{i=1}^{t} {k \choose i}$.

本文也是以减少译码表的行数为目的. 我们利用线性循环码的性质和伴随式的重量,给出了适用于所有二进制循环码的一个新的简化查表译码算法,并从理论上证明了这种算法的正确性.

1 系统循环码的生成矩阵和一致校验矩阵

设 ℓ 是一个(n,k) 循环码. 多项式

$$c(x) = c_0 + c_1 x + \dots + c_{n-1} x^{n-1}$$

称为码字 $\mathbf{c} = [c_0, c_1, \cdots, c_{n-1}] \in \mathcal{C}$ 的多项式. \mathcal{C} 中码字多项式中次数最小的首一多项式称为 \mathcal{C} 的生成多项式. 设 \mathcal{C} 的生成多项式为

$$g(x) = g_0 + g_1 x + \dots + g_{n-k} x^{n-k}$$

则矩阵

$$\mathbf{G}_{1} = \begin{bmatrix} g(x) \\ xg(x) \\ x^{2}g(x) \\ \cdots \\ x^{k-1}g(x) \end{bmatrix} = \begin{bmatrix} g_{0} \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

是 ℓ 的一个生成矩阵. 对 G_1 进行行初等变换,可以得到 ℓ 的如下的生成矩阵:

$$G = [I_b, A_{b \times (n-b)}]$$

这里 I_k 是 k 阶单位阵, $A_{k\times(n-k)}$ 是一个 $k\times(n-k)$ 阶矩阵. 由 G 可得 ℓ 的一个一致校验矩阵:

$$\boldsymbol{H} = [-(\boldsymbol{A}_{k \times (n-k)})^{\mathrm{T}}, \boldsymbol{I}_{n-k}]$$

下面用 G 和 H 分别记二进制(n, k) 循环码 ℓ 的生成矩阵和一致校验矩阵:

$$G = [I_k, A_{k \times (n-k)}], H = [(A_{k \times (n-k)})^T, I_{n-k}]$$

并分别简记为 G = [I, A] 和 $H = [A^T, I]$. 容易看出,信息向量 $m = [m_0, m_1, \dots, m_{k-1}]$ 经过 G 编码得到码字 c,其前 k 个分量恰为 m:

$$c = mG = \lceil mI, mA \rceil = \lceil m, mA \rceil$$

具有这种结构的码称为系统码. 在系统码的情况下,我们将一个二进制 n 维向量的前 k 个分量称为信息部分,后 n-k 个分量称为校验部分.

2 主要定理及证明

下面我们假设 ℓ 是二进制(n, k, d) 循环码, $\mathbf{H} = [\mathbf{A}^T, \mathbf{I}]$ 是 ℓ 的一个一致校验矩阵.

设 e 是错误向量 $e = [e_M, e_P] = [e_0, e_1, \dots, e_{n-1}]$, 其中 $e_M = [e_0, e_1, \dots, e_{k-1}]$ 是信息部分, $e_P = [e_k, e_{k+1}, \dots, e_{n-1}]$ 是校验部分,

$$w(e) \leqslant t, s = eH^{T}$$

下面这个结论在后面的证明中多次用到:

定理 1 若 $\mathbf{H} = [\mathbf{A}^{\mathsf{T}}, \mathbf{I}], \mathbf{e} = [\mathbf{e}_{\mathsf{M}}, \mathbf{0}_{1 \times (n-k)}] \neq \mathbf{0}_{\mathsf{n}}, 则 w(\mathbf{e}\mathbf{H}^{\mathsf{T}}) \geqslant d - w(\mathbf{e}_{\mathsf{M}}).$

证 因为码字的重量至少是 d, $e_M[I,A] = [e_M, e_M A]$ 是码字, 所以

$$w(\mathbf{e}_{M}) + w(\mathbf{e}_{M}\mathbf{A}) \geqslant d$$

而

$$\mathbf{S} = \mathbf{e} \mathbf{H}^{\mathrm{T}} = [\mathbf{e}_{M}, \mathbf{0}_{1 \times (n-k)}] [\mathbf{A}^{\mathrm{T}}, \mathbf{I}]^{\mathrm{T}} = \mathbf{e}_{M} \mathbf{A}$$

故

$$w(\boldsymbol{e}\boldsymbol{H}^{\mathrm{T}}) = w(\boldsymbol{e}_{\boldsymbol{M}}\boldsymbol{A}) \geqslant d - w(\boldsymbol{e}_{\boldsymbol{M}})$$

在纠错范围内,定理2和定理3确定了错误只出现在校验部分的条件.

定理 2 若 $e = [\mathbf{0}_{1 \times k}, e_P]$, 其中 $w(e_P) \leqslant t$, 则

$$w(\mathbf{s}) = w(\mathbf{e}\mathbf{H}^{\mathrm{T}}) \leqslant t$$

证 由 $H = [A^T, I]$ 的结构可得

$$\mathbf{s} = \mathbf{e} \mathbf{H}^{\mathrm{T}} = [\mathbf{0}_{1 \times k}, \mathbf{e}_{P}] [\mathbf{A}^{\mathrm{T}}, \mathbf{I}]^{\mathrm{T}} = \mathbf{e}_{P}$$

故

$$w(\mathbf{s}) = w(\mathbf{e}_P) \leqslant t$$

定理 3 若
$$e = [e_M, e_P]$$
, 其中 $w(e_M) \geqslant 1$, $w(e_M) + w(e_P) \leqslant t$, 则
$$w(s) = w(eH^T) \geqslant t + 1$$

证 此时

$$w(e\mathbf{H}^{\mathrm{T}}) = w([\mathbf{e}_{M}, \mathbf{0}_{1 \times (n-k)}]\mathbf{H}^{\mathrm{T}} + [\mathbf{0}_{1 \times k}, \mathbf{e}_{P}]\mathbf{H}^{\mathrm{T}}) \geqslant (d - w(\mathbf{e}_{M})) - w(\mathbf{e}_{P}) = d - (w(\mathbf{e}_{M}) + w(\mathbf{e}_{P})) \geqslant d - t \geqslant t + 1$$

故结论成立.

由定理2和定理3可得:

推论 1 若 $w(e) \leq t$,则 $w(eH^{T}) \leq t$ 当且仅当 $e = [\mathbf{0}_{1 \times k}, eH^{T}]$.

当 $e = [e_M, e_P]$ 时,下面的定理给出了 e_P 与 e 的伴随式和 e_M 的伴随式之间的一个关系.

定理 4 设 $e = [e_M, e_P]$. 若 $[e_M, \mathbf{0}_{1 \times (n-k)}] \mathbf{H}^T = \mathbf{s}_M, 则$

$$e_P = eH^T - s_M$$

证 因为 $w(e_P) \leq w(e) \leq t$,所以

$$e_P = [\mathbf{0}_{1 \times k}, e_P] \mathbf{H}^{\mathrm{T}} =$$

$$(e - [e_M, \mathbf{0}_{1 \times (n-k)}]) \mathbf{H}^{\mathrm{T}} =$$

$$e \mathbf{H}^{\mathrm{T}} - [e_M, \mathbf{0}_{1 \times (n-k)}] \mathbf{H}^{\mathrm{T}} =$$

$$e \mathbf{H}^{\mathrm{T}} - \mathbf{S}_M$$

结论得证.

设 N' 个信息部分出错的错误向量为:

$$\boldsymbol{a}_1 = [\boldsymbol{e}_M^1, \; \boldsymbol{0}_{1 \times (n-k)}], \; \cdots, \; \boldsymbol{a}_{N'} = [\boldsymbol{e}_M^{N'}, \; \boldsymbol{0}_{1 \times (n-k)}]$$

其对应的伴随式分别为:

$$\mathbf{s}_1 = \mathbf{a}_1 \mathbf{H}^{\mathrm{T}}, \ \cdots, \ \mathbf{s}_{N'} = \mathbf{a}_{N'} \mathbf{H}^{\mathrm{T}}$$

将向量 \mathbf{s}_i 及 \mathbf{e}_M^i 并置在第一行,得到一个 N' 行、n 列的表(表 1),这个表称为信息部分伴随式与错误模式的对应表 MP-SET(Message Part-Syndrome Error Table)(表 1).

表 1 信息部分伴随式与错误模式的对应表

序号	伴随式	错误模式	序号	伴随式	错误模式
1	$\mathbf{s}_1 = [\mathbf{e}_M^1, 0_{1 \times (n-k)}] \mathbf{H}^{\mathrm{T}}$	$oldsymbol{e}_{M}^{1}$	i	$s_i = [e_M^i, 0_{1 \times (n-k)}] H^{\mathrm{T}}$	e_{M}^{i}
2	$\boldsymbol{s}_2 = [\boldsymbol{e}_M^2, \boldsymbol{0}_{1 \times (n-k)}] \boldsymbol{H}^{\mathrm{T}}$	e_{M}^{2}	:	:	:
<u>:</u>	:	:	N'	$oldsymbol{s}_{N'} = ig ig oldsymbol{e}_{M}^{N'}$, $oldsymbol{0}_{1 imes(n-k)} ig oldsymbol{H}^{ ext{T}}$	$oldsymbol{e}_{M}^{N'}$

我们的译码算法就基于 MP-SET. 由于 k < n,所以 $N' < N = \sum_{i=1}^{t} {n \choose i}$,即表 MP-SET 的行数要比前

面提到的 $N \times (2n - k)$ 表的行数小很多,因此要减少很多存储空间.下面的定理是我们的查表译码算法正确性的理论基础.

定理 5 设 $w(e) \leq t$, $e = [e_M, e_P]$, $w(e_M) \geqslant 1$. 若 $eH^T = s$, 则存在唯一一个 $j \in \{1, \dots, N'\}$, 使 得 $w([s - s_i, e_M^i]) \leq t$. 此时一定有 $e = [e_M^i, s - s_i]$.

证 由 MP-SET 的构造可知,存在 $1 \leq j \leq N'$ 使得 $e_M = e_M^i$. 再由定理 4 知 $e_P = s - s_j$,故 $w([s - s_i, e_M^i]) = w(e)$. 下面只需证当 $i \neq j$ 时必有 $w([s - s_i, e_M^i]) \geq t + 1$. 实际上

$$w([\mathbf{s} - \mathbf{s}_{i}, \mathbf{e}_{M}^{i}]) = w([[\mathbf{e}_{M}, \mathbf{e}_{P}]\mathbf{H}^{\mathsf{T}} - \mathbf{a}_{i}\mathbf{H}^{\mathsf{T}}, \mathbf{e}_{M}^{i}]) =$$

$$w([[\mathbf{e}_{M} + \mathbf{e}_{M}^{i}, \mathbf{0}]\mathbf{H}^{\mathsf{T}} + [\mathbf{0}, \mathbf{e}_{P}]\mathbf{H}^{\mathsf{T}}, \mathbf{e}_{M}^{i}]) \geqslant$$

$$w([\mathbf{e}_{M} + \mathbf{e}_{M}^{i}, \mathbf{0}]\mathbf{H}^{\mathsf{T}}) - w([\mathbf{0}, \mathbf{e}_{P}]\mathbf{H}^{\mathsf{T}}) + w(\mathbf{e}_{M}^{i}) \geqslant$$

$$d - (w(\mathbf{e}_{M}) + w(\mathbf{e}_{M}^{i})) - w(\mathbf{e}_{P}) + w(\mathbf{e}_{M}^{i}) =$$

$$d - w(\mathbf{e}) \geqslant d - t \geqslant t + 1$$

3 新的译码算法

设 ℓ 是纠错能力为 t 的 (n, k) 循环码, $g(x) = \sum_{i=0}^{n-k} g_i x^i$ 是它的一个生成多项式.

由g(x)导出 ℓ 的如下的生成矩阵和一致校验矩阵:

$$G = [I_k, A], H = [A^T, I_{n-k}]$$

然后计算 MP-SET. 对于接收到的向量 r, 查表译码算法如下:

SimplifiedTableLookupDecoder(r, H, MP-SET)

$$s \leftarrow rH^{T}$$
 $e \leftarrow [\mathbf{0}_{1 \times k}, s]$
 $i \leftarrow 1$
while $w(e) > t$ and $i \leq N'$

$$do \begin{cases} e \leftarrow [\text{MP-SET}(i, 2), s - \text{MP-SET}(i, 1)] \\ i \leftarrow i + 1 \end{cases}$$
if $w(e) > t$
then return ("failure")
else return $(r + e) \mod 2$

注1 当算法输出"failure"时,表明在 MP-SET 中没有找到对应的错误模式,也就是说错误个数超出了纠错能力,因此算法此时实际上相当于检测到纠错能力范围之外的一个错误.

(15,5)循环码 ℓ的生成多项式为

$$g(x) = 1 + x^2 + x^5 + x^6 + x^8 + x^9 + x^{10}$$

由此得到的生成矩阵

经初等行变换可化为

这时

$$\mathbf{A} = \begin{bmatrix} 1 & 0 & 1 & 0 & 0 & 1 & 1 & 0 & 1 & 1 \\ 1 & 1 & 1 & 1 & 0 & 1 & 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 & 1 & 0 & 1 & 0 & 1 & 1 \\ 1 & 0 & 0 & 1 & 1 & 0 & 1 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 & 1 & 0 & 1 & 1 & 1 \end{bmatrix}$$

其中码 ℓ 的最小距离 d=7,纠错能力 t=3,而 $N'=\binom{5}{1}+\binom{5}{2}+\binom{5}{3}=25$. MP-SET 见表 2.

假设发送的码字是c = [001101110000101],接收到的向量是r = [100101100000101]. 计算伴随式 得 $s = \lceil 11111110000 \rceil$, 其重量大于 3, 进入 while 循环, 在第 7 次得到的错误向量是

 $e = \lceil 10100, s - \lceil 1101110000 \rceil \rceil = \lceil 10100, 00100000000 \rceil = \lceil 1010000100000000 \rceil$

其重量是 3, 故译码输出为 $\mathbf{v} = \mathbf{r} + \mathbf{e} = [001101110000101]$.

与 $N = \sum_{i=1}^{3} {n \choose i} = 575$ 相比,MP-SET 少了 550(575 - 25 = 550) 行.

					表 2	(15	,5) 循	「 环码 F	的 MP	-SET					
	伴随式									错误模式					
1	1	0	1	0	0	1	1	0	1	1	1	0	0	0	0
2	1	1	1	1	0	1	0	1	1	0	0	1	0	0	0
3	0	1	1	1	1	0	1	0	1	1	0	0	1	0	0
4	1	0	0	1	1	0	1	1	1	0	0	0	0	1	0
5	0	1	0	0	1	1	0	1	1	1	0	0	0	0	1
6	0	1	0	1	0	0	1	1	0	1	1	1	0	0	0
7	1	1	0	1	1	1	0	0	0	0	1	0	1	0	0
8	0	0	1	1	1	1	0	1	0	1	1	0	0	1	0
9	1	1	1	0	1	0	1	1	0	0	1	0	0	0	1
10	1	0	O	0	1	1	1	1	0	1	0	1	1	O	0
11	0	1	1	0	1	1	1	O	0	0	0	1	0	1	0
12	1	0	1	1	1	0	0	O	0	1	0	1	0	0	1
13	1	1	1	0	0	0	0	1	0	1	0	0	1	1	0
14	0	0	1	1	0	1	1	1	0	0	0	0	1	0	1
15	1	1	O	1	0	1	1	O	0	1	0	0	0	1	1
16	0	0	1	0	1	0	0	1	1	0	1	1	1	0	0
17	1	1	0	0	1	0	0	O	1	1	1	1	0	1	0
18	0	0	0	1	1	1	1	O	1	0	1	1	0	0	1
19	0	1	0	0	0	1	1	1	1	0	1	0	1	1	0
20	1	0	0	1	0	0	0	1	1	1	1	0	1	0	1
21	0	1	1	1	0	0	0	0	1	0	1	0	0	1	1
22	0	0	0	1	0	1	0	0	1	1	0	1	1	1	0
23	1	1	0	0	0	0	1	0	1	0	0	1	1	0	1
24	0	0	1	0	0	0	1	1	1	1	0	1	0	1	1
25	1	0	1	0	1	1	0	0	1	0	0	0	1	1	1

4 结束语

文献[6-9]中的译码算法正确性是通过程序对所有可能的错误模式进行纠错来验证的. 定理 5 则从理论上保证了本文算法的正确性. 本文给出的译码算法简单明了,容易理解和实现. 除了能纠正所有纠错能力范围内的错误外,还能检测到部分纠错能力范围外的错误. 同传统的查表译码算法中的表的行数 $\sum_{i=1}^{t} \binom{n}{i}$ 相比,本算法只需 $\sum_{i=1}^{t} \binom{k}{i}$ 行,极大地减少了存储量和译码的复杂性.

参考文献:

- 「1] MCELIECE R J. The Theory of Information and Coding(Second Edition) 「M]. 北京: 电子工业出版社, 2002: 140.
- [2] 瞿云云,包小敏,童新安.基于(24,12)扩展 Golay 码的信息协调协议 [J].西南大学学报:自然科学版,2010,32(5): 26-29.
- [3] MEGGITT J E. Error Correcting Codes and Their Implementation for Data Transmission Systems [J]. IRE Trans on Information Theory, 1961, 7(4): 234-244.
- [4] LIN Shu, COSTELLO D J. Error Control Coding [M]. 2nd Edition. Englewood Cliffs: Pearson Prentice Hall, 2004: 156.
- [5] WICKER S B. Error Control Systems for Digital Communication and Storage [M]. Englewood Cliffs: Prentice Hall, 1995: 118.
- [6] CHANG H C, LEE H P, LIN T C, et al. A Weight Method of Decoding the (23, 12, 7) Golay Code Using Reduced Table Lookup [C]// Proceedings of the Sixth International Conference on Communications, Circuits and Systems (ICCCAS'08). Chengdu: UESTC Press, 2008: 1-5.
- [7] CHEN Y H, TRUONG T K, HUANG C H, et al. A Lookup Table Decoding of Systematic (47, 24, 11) Quadratic Residue Code [J]. Information Sciences, 2009, 179: 2470-2477.
- [8] LEE H P, CHANG H C. A Weight Method of Decoding the Binary BCH Code [C]// Proceedings of the Eighth International Conference on Intelligent Systems Design and Applications (ISDA'08). Piscataway: IEEE Press, 2008: 545-549.
- [9] LIN T C, LEE H P, CHANG H C, et al. High Speed Decoding of the Binary (47, 24, 11) Quadratic Residue Code [J]. Information Sciences, 2010, 180; 4060—4068.

A Simplified Table Lookup Decoding Algorithm for Systematic Binary Cyclic Codes

LIU Xu, BAO Xiao-min, WU Deng-jie, LI Mei, YUAN Zhi-hua

School of Mathematics and Statistics, Southwest University, Chongging 400715, China

Abstract: By utilizing the properties of linear cyclic codes and the weights of syndromes, we derive a table lookup decoding algorithm for binary cyclic codes. The algorithm has the advantages of high efficiency and low memory requirement.

Key words: cyclic code; table lookup decoding; error pattern; syndrome; Hamming weight