一类新的能够渐进达到 Gilbert-Varshamov 界的 Alternant 子类码

樊继豪[1],陈汉武[1,2]
(1.东南大学计算机科学与工程学院,江苏南京 211189; 2.东南大学计算机网络和信息集成教育部重点实验室,江苏南京 211189)

摘 要: 本文基于 Maximum Distance Separable(MDS)码的 Hamming 重量分布提出一类新的二元 Alternant 子类码.分析表明这类新的子类码包含整个 BCH 码类,并且可以渐进达到 Gilbert-Varshamov(GV)界.

关键词: Alternant 码; BCH 码; Gilbert-Varshamov 界; Hamming 重量分布; Maximum Distance Separable (MDS)码

中图分类号: TN911.22 **文献标识码:** A **文章编号:** 0372-2112 (2015)11-2243-04

电子学报 URL: http://www.ejournal.org.cn **DOI:** 10.3969/j.issn.0372-2112.2015.11.016

A New Subclass of Alternant Codes Can Meet the Gilbert-Varshamov Bound

FAN Ji-hao[1], CHEN Han-wu[1,2]
(1. *School of Computer Science and Engineering College of Software Engineering, Southeast University, Nanjing, Jiangsu* 211189, *China*;
2. *Key Laboratory of Computer Network and Information Integration, Southeast University, Ministry of Education, Nanjing, Jiangsu* 211189, *China*)

Abstract: A new subclass of binary Alternant codes is proposed based on the Hamming weight distribution of Maximum Distance Separable(MDS) codes. It is shown that the new codes include the whole BCH codes subclass and can asymptotically meet the Gilbert-Varshamov(GV) bound.

Key words: Alternant codes; BCH codes; Gilbert-varshamov bound; Hamming weight distribution; Maximum distance separable(MDS) codes

1 引言

Alternant 码是一类范围非常宽广的经典线性纠错码.该类码可以看作是扩展 Reed-Solomon 码(GRS 码)的子域子类码[1].很多好的经典纠错码都属于 Alternant 码的范畴,比如 BCH 码,Goppa 码就是两类非常著名的 Alternant 子类码[1,2].Alternant 码的码长可以在 2 到 q^m(q 为素数幂,m 为正整数)之间自由选择,其维度与最小距离也有好的下界估计,重要的是 Alternant 码包含可以渐进达到 Gilbert-Varshamov(GV)界的好码.Goppa 码则是一类可以渐进达到 GV 界的 Alternant 子类码[3].由文献[1]可知,从渐进的观点来看,长的 BCH 码不是好码,即码率随着相对距离的增加渐进趋近于 0,BCH 码是一类循环码,到目前为止人们尚不知是否存在渐进好的循环码[4].

本文我们首先基于 Reed-Solomon(RS)码的 Hamming 重量分布构造了一类新的 Alternant 子类码:我们通过将 Alternant 码的校验矩阵与 RS 码编码后的非零码字进行交织以得到新的校验矩阵,以此作为 Alternant 子类码的校验矩阵.分析表明,交织后的 Alternant 子类码可以渐进达到 GV 界,并且包含整个 BCH 码类.然后我们进一步将其推广到更为普遍的基于 MDS 码的 Hamming 重量分布来进行构造,结果显示,该构造具有相同的渐近效果.

2 几类经典纠错码介绍

记 p 为素数,q 为 p 的幂次,即 $q=p^r$, $r>0$.记 F_q 表示有 q 个元素的有限域, $GF(q^m)$表示 $GF_{(q)}$的 m 次

收稿日期:2014-06-19;修回日期:2015-06-09;责任编辑:马兰英
基金项目:国家自然科学基金(No.61170321);高等学校博士学科点专项科研基金(No.20110092110024);江苏省普通高校研究生科研创新计划(No.CXZZ13_0105)

扩张域,α 表示有限域 $GF(q^m)$的本原元.

有限域 $GF(q^m)$上码长为 $n=q^m-1$ 的本原 RS 码可以定义为以 $\alpha^l,\alpha^{l+1},\cdots,\alpha^{l+\delta-2}$为根的循环码,其中 l 与 δ 为整数,且 $2\leqslant\delta\leqslant n-1$,记作 $RS(n,\delta)$.其生成多项式为 $g(x)=(x-\alpha^l)(x-\alpha^{l+1})\cdots(x-\alpha^{l+\delta-2})$. $RS(n,\delta)$为 MDS 码,其参数为$[n,k,d]_{q^m}$,其中 $k=n-\delta+1,d=\delta$.如果 $l=1$,那么此时的 $RS(n,\delta)$为狭义本原 RS 码,其校验矩阵可记作:

$$\boldsymbol{H}_{RS(n,\delta)}=\begin{pmatrix}1&\alpha&\cdots&\alpha^{n-1}\\1&\alpha^2&\cdots&\alpha^{2(n-1)}\\\vdots&\vdots&\vdots&\vdots\\1&\alpha^{\delta-1}&\cdots&\alpha^{(\delta-1)(n-1)}\end{pmatrix}\tag{1}$$

GRS 码通过对 RS 码做了进一步推广而得来,其取消了码长与定义集合的限制,是一类码长可以自由变换的纠错码.记 $\boldsymbol{a}=(\alpha_1,\alpha_2,\cdots,\alpha_n)$,其中 α_i 为 $GF(q^m)$中的不同元素;记 $\boldsymbol{v}=(v_1,v_2,\cdots,v_n)$,其中 v_i 为$GF(q^m)$中的非零元素.对于任意的 $1\leqslant k\leqslant n-1$,GRS 码 $GRS_k(\boldsymbol{a},\boldsymbol{v})$定义为:

$$GRS_k(\boldsymbol{a},\boldsymbol{v})=\{v_1F(\alpha_1),v_2F(\alpha_2),\cdots,v_nF(\alpha_n)\mid F(z)\in GF(q^m)[z],\deg F(z)<k\}.$$

$GRS_k(\boldsymbol{a},\boldsymbol{v})$的参数为$[n,k,n-k+1]_{q^m}$,亦是 MDS 码.GRS 码的对偶码仍然为 GRS 码,即 $GRS_k(\boldsymbol{a},\boldsymbol{v})^{\perp}=GRS_{n-k}(\boldsymbol{a},\boldsymbol{y})$,其中 $\boldsymbol{y}=(y_1,y_2,\cdots,y_n)$,并且 $y_i\cdot v_i=1/\prod\limits_{j\neq i}(\alpha_i-\alpha_j)$.GRS 码 $GRS_k(\boldsymbol{a},\boldsymbol{v})$的校验矩阵为:

$$\boldsymbol{H}_{GRS_k(\boldsymbol{a},\boldsymbol{v})}=\begin{pmatrix}y_1&y_2&\cdots&y_n\\y_1\alpha_1&y_2\alpha_2&\cdots&y_n\alpha_n\\\vdots&\vdots&\vdots&\vdots\\y_1\alpha_1^{r-1}&y_2\alpha_2^{r-1}&\cdots&y_n\alpha_n^{r-1}\end{pmatrix}\tag{2}$$

其中 $r=n-k$.

RS 码与 GRS 码都是 MDS 码,对于任意一个参数为$[n,k,d]_{q^m},d=n-k+1$ 的 MDS 码,由文献[1,5]可知其 Hamming 重量分布为:

$$A_w=\binom{n}{w}(q^m-1)\sum_{j=0}^{w-d}(-1)^j\binom{w-1}{j}q^{m(w-d-j)}\tag{3}$$

Alternant 码是 GRS 码的子域子类码,对于上述记号,其定义为 $A_r(\boldsymbol{a},\boldsymbol{y})=GRS_k(\boldsymbol{a},\boldsymbol{v})\mid_{GF(q)}$.因此 Alternant 码 $A_r(\boldsymbol{a},\boldsymbol{y})$与 GRS 码 $GRS_k(\boldsymbol{a},\boldsymbol{v})$有着相同的校验矩阵,即 $\boldsymbol{H}_{A_r(\boldsymbol{a},\boldsymbol{y})}=\boldsymbol{H}_{GRS_k(\boldsymbol{a},\boldsymbol{v})}$.

3 渐进好的 Alternant 子类码

对于通常意义下的 Alternant 码,向量 $\boldsymbol{y}$ 中 y_i 的选择除了零元素之外是完全随机的,这在一定程度上满足香农定理对好码随机性的要求.事实上的确存在好的 Alternant 码能够渐进达到 GV 界(文献[1],第九章).本文我们考虑二元本原 Alternant 码,即我们取 $q=2,n=2^m-1,\alpha_i=\alpha^i,0\leqslant i\leqslant n-1$.那么二元本原 Alternant 码 $A_r(\boldsymbol{a},\boldsymbol{y})$的校验矩阵为

$$\boldsymbol{H}_{A_r(\boldsymbol{a},\boldsymbol{y})}=\begin{pmatrix}y_1&y_2\alpha&\cdots&y_n\alpha^{n-1}\\y_1&y_2\alpha^2&\cdots&y_n\alpha^{2(n-1)}\\\vdots&\vdots&\vdots&\vdots\\y_1&y_2\alpha^r&\cdots&y_n\alpha^{r(n-1)}\end{pmatrix}\tag{4}$$

由文献(1)可知,Alternant 码 $A_r(\boldsymbol{a},\boldsymbol{y})$的维度满足 $k\geq n-mr$,最小距离满足 $d\geq r+1$.需要注意的是,此处 $A_r(\boldsymbol{a},\boldsymbol{y})$校验矩阵的选取与式(2)略有差别.易知,$\boldsymbol{H}_{A_r(\boldsymbol{a},\boldsymbol{y})}=\boldsymbol{H}_{RS(n,r+1)}\cdot\mathrm{diag}(\boldsymbol{y})$,其中 $H_{RS(n,r+1)}$为狭义本原 RS 码 $RS(n,r+1)$对应的校验矩阵,$\mathrm{diag}(\boldsymbol{y})$表示以向量 $\boldsymbol{y}$ 为对角元素的对角矩阵.

我们取 $\boldsymbol{Y}=(\tau_1,\tau_2,\cdots,\tau_n)$为狭义本原 RS 码 $RS(n,\delta)$编码后的码字,并且要求编码后的码字为全非零向量,即 $\tau_i\neq0,1\leqslant i\leqslant n$.那么 $\boldsymbol{Y}$ 的所有可能选择组成了 Alternant 码的一个子类,我们称作 Sub-Alternant 码.对于该子类中的码字,我们记作 $S-A_r(\boldsymbol{a},\boldsymbol{Y})$.具体的,我们有如下定义:

定义 1 对于任意的 $\boldsymbol{Y}=(\tau_1,\tau_2,\cdots,\tau_n)\in RS(n,\delta)$,即 $H_{RS(n,\delta)}\boldsymbol{Y}^{\mathrm{T}}=0$,并且 $\tau_i\neq0,1\leqslant i\leqslant n$.那么 $S-A_r(\boldsymbol{a},\boldsymbol{Y})$定义为:

$$S-A_r(\boldsymbol{a},\boldsymbol{Y})=\{c\in F_2^n\mid H_{A_r(\boldsymbol{a},\boldsymbol{Y})}\boldsymbol{c}^{\mathrm{T}}=0\}$$

其中 $H_{A_r(\boldsymbol{a},\boldsymbol{Y})}$为(4)中选定 $\boldsymbol{Y}$ 后的 Alternant 码 $A_r(\boldsymbol{a},\boldsymbol{Y})$的校验矩阵,而 $H_{RS(n,\delta)}$为(1)中 RS 码 $RS(n,\delta)$的校验矩阵.

例 1 我们取码长为 $n=15$,设计距离 $\delta=3$,有限域 $GF_{(16)}$上的 Reed-Solomon 码 $RS(15,3)$,该码的校验矩阵为

$$\boldsymbol{H}_{RS(15,3)}=\begin{pmatrix}1&\alpha&\cdots&\alpha^{14}\\1&\alpha^2&\cdots&\alpha^{28}\end{pmatrix}$$

其中 α 为 $GF_{(16)}$上的本原元.由式(3)可知定义 1 中非零向量 $\boldsymbol{Y}=(\tau_1,\tau_2,\cdots,\tau_{15})\in RS(15,3)$的数量为 $A_{15}=15\sum\limits_{j=0}^{12}(-1)^j\binom{14}{j}16^{12-j}=1710523009300245$,整个码空间的码元数量为 $16^{13}=4503599627370496$,原 Alternant 码向量 $\boldsymbol{y}$ 的可选择数量为 $15^{16}=6568408355712890625$.任取其中的一个非零向量 $\boldsymbol{Y}$,那么 $S-A_r(\boldsymbol{a},\boldsymbol{Y})$可定义为:

$$S-A_r(\boldsymbol{a},\boldsymbol{Y})=\{\boldsymbol{c}\in F_2^n\mid\boldsymbol{H}_{A_r(\boldsymbol{a},\boldsymbol{Y})}\boldsymbol{c}^{\mathrm{T}}=0\}$$

其中 $H_{A_r(\boldsymbol{a},\boldsymbol{Y})}$为 Alternant 码 $A_r(\boldsymbol{a},\boldsymbol{Y})$的校验矩阵.

比如选取向量 $\boldsymbol{Y}=(1,\alpha^6,\alpha^{12},\alpha^3,\alpha^9,\alpha^{10},\alpha^6,\alpha^5,\alpha^3,\alpha^9,\alpha^5,\alpha^{10},\alpha^{12},\alpha^5,\alpha^{10})$,$r=2$,可得 Alternant 子类码 $S-A_r(\boldsymbol{a},\boldsymbol{Y})$,其参数为$[15,\geq7,\geq3]$,进一步的计算可

得该码的具体参数为[15,7,5],由向量 $\boldsymbol{Y}$ 的选取不难发现,该码不是 BCH 码.

对于定义 1 中的 Alternant 子类码 $S-A_r(\boldsymbol{a},\boldsymbol{Y})$,我们有如下的渐进结果.

定理 1　对于 $\delta/2<r<\min\{\delta,n/2\}$,存在长的 Alternant 子类码 $S-A_r(\boldsymbol{a},\boldsymbol{Y})$可以渐进达到 GV 界.

证明　对于 Alternant 子类码 $S-A_r(\boldsymbol{a},\boldsymbol{Y})$,考虑任意的二元向量 $\boldsymbol{c}=(c_1,c_2,\cdots,c_n)$,设其 Hamming 重量为 t.如果 c 是$S-A_r(\boldsymbol{a},\boldsymbol{Y})$的一个码元,那么由定义 1 知,必有 $\boldsymbol{H}_{A_r(\boldsymbol{a},\boldsymbol{Y})}\boldsymbol{c}^{\mathrm{T}}=0$,也就有

$$\boldsymbol{H}_{RS(n,r+1)}(\tau_1c_1,\tau_2c_2,\cdots,\tau_nc_n)^{\mathrm{T}}=0$$

记向量 $\boldsymbol{c}=(c_1,c_2,\cdots,c_n)$中非零元素的位置集合为$\{c_{i_1},c_{i_2},\cdots,c_{i_t}\}$,$1\leqslant i_1<i_2<\cdots<\cdots i_t\leqslant n$,那么我们有

$$\boldsymbol{H}_{RS(n,r+1)}(\cdots,\tau_{i_1}c_{i_1},\cdots,\tau_{i_t}c_{i_t},\cdots)^{\mathrm{T}}=0$$

其中"…"表示根据具体的需要,零元素相对应的位置.由于 $\boldsymbol{c}$ 是二元向量,这说明

$$\boldsymbol{H}_{RS(n,r+1)}(\cdots,\tau_{i_1},\cdots,\tau_{i_t},\cdots)^{\mathrm{T}}=0$$

如果记

$$B_w^{'}=(2^m-1)\sum_{j=0}^{w-(r+1)}(-1)^j\binom{w-1}{j}2^{m(w-(r+1)-j)}$$

那么以 $\boldsymbol{H}_{RS(n,r+1)}$为校验矩阵的 RS 码的 Hamming 重量分布为 $B_w=\binom{n}{w}B_w^{'}$,那么$(\cdots,\tau_{i_1},\cdots,\tau_{i_t},\cdots)$总的选择数量为 $B_t^{'}$.

根据定义 1,以及 $r<\delta$,我们有 $\boldsymbol{H}_{RS(n,r+1)}(\tau_1,\tau_2,\cdots,\tau_n)^{\mathrm{T}}=0$,那么有

$$\boldsymbol{H}_{RS(n,r+1)}(\cdots,\tau_{j_1},\cdots,\tau_{j_{(n-t)}},\cdots)^{\mathrm{T}}=0$$

其中$(\cdots,\tau_{j_1},\cdots,\tau_{j_{(n-t)}},\cdots)^{\mathrm{T}}=(\tau_1,\tau_2,\cdots,\tau_n)^{\mathrm{T}}-(\cdots,\tau_{i_1}c_{i_1},\cdots,\tau_{i_t}c_{i_t},\cdots)^{\mathrm{T}}$,$1\leqslant j_1<j_2<\cdots<j_{(n-t)}\leqslant n$,"…"表示根据具体的需要,零元素相对应的位置,那么$(\cdots,\tau_{j_1},\cdots,\tau_{j_{(n-t)}},\cdots)$总的选择数量为 $B_{n-t}^{'}$.因此 $\boldsymbol{Y}=(\tau_1,\tau_2,\cdots,\tau_n)$的总的选择数量至多为 $B_t^{'}B_{n-t}^{'}$,注意到 $B_w^{'}\leqslant(2^m-1)^{w-r}$,那么有

$$B_t^{'}B_{n-t}^{'}\leqslant(2^m-1)^{n-2r}$$

因此对于所有 Hamming 重量 $t<\omega$ 的码元,包含这些码元的 Alternant 子类码 $S-A_r(\boldsymbol{a},\boldsymbol{Y})$的选择数至多是

$$\sum_{t=r+1}^{\omega-1}B_t^{'}B_{n-t}^{'}\binom{n}{t}\leqslant(2^m-1)^{n-2r}\sum_{t=r+1}^{\omega-1}\binom{n}{t}$$

另一方面,整个 Alternant 子类码的数量等于向量 $\boldsymbol{Y}$ 的选择数量,即为

$$A_n=(2^m-1)\sum_{j=0}^{n-\delta}(-1)^j\binom{n-1}{j}2^{m(n-\delta-j)}$$

$$\geq(2^m-1)2^{(m(n-\delta))}\left(1-\frac{n-1}{2^m}\right)$$

$$>(2^m-1)^{n-\delta}$$

因此如果

$$(2^m-1)^{n-2r}\sum_{t=r+1}^{\omega-1}\binom{n}{t}<(2^m-1)^{n-\delta}$$

即为

$$\sum_{t=r+1}^{\omega-1}\binom{n}{t}<(2^m-1)^{2r-\delta}\tag{5}$$

对于式(5),利用文献[1]中的二项式估计,取基于 2 的对数,并且取 $n\to\infty$,可将其重写为:

$$H(\zeta)+o(1)<\frac{m(2r-\delta)}{n}+o(1)<\frac{mr}{n}+o(1)\tag{6}$$

其中 $\zeta=\lim_{n\to\infty}\frac{\omega}{n}$,通过选取合适的 n,r 以及 δ,我们可以使得$\frac{mr}{n}$近似等于 $H(\zeta)$,即有对于任意小的 ε,$\frac{mr}{n}=H(\zeta)+\varepsilon$.根据 Alternant 码的维度性质,该码的码率满足:

$$R\geq1-\frac{mr}{n}\geq1-H(\zeta)-\varepsilon\tag{7}$$

因此存在 Sub-Alternant 码能够渐进达到 GV 界.

证毕.

由于所有 MDS 码都具有完全确定的 Hamming 重量分布,因此上述 Sub-Alternant 码中的 $\boldsymbol{Y}$ 也可选为其他非 RS 码的编码后码元,并且码长可以在 2 到 2^m 之间自由变换,即我们可以把 Sub-Alternant 码中对应的 RS 码的校验矩阵选为

$$\boldsymbol{H}_{\boldsymbol{Y}}=\begin{pmatrix}1&1&\cdots&1\\ \alpha_1&\alpha_2&\cdots&\alpha_n\\ \vdots&\vdots&\vdots&\vdots\\ \alpha_1^{\delta-2}&\alpha_2^{\delta-2}&\cdots&\alpha_n^{\delta-2}\end{pmatrix}$$

而将 Sub-Alternant 码的校验矩阵选为

$$\boldsymbol{H}_{A_r(\boldsymbol{a},\boldsymbol{Y})}=\begin{pmatrix}\tau_1&\tau_2&\cdots&\tau_n\\ \tau_1\alpha_1&\tau_2\alpha_2&\cdots&\tau_n\alpha_n\\ \vdots&\vdots&\vdots&\vdots\\ \tau_1\alpha_1^{r-1}&\tau_2\alpha_2^{r-1}&\cdots&\tau_n\alpha_n^{r-1}\end{pmatrix}$$

其中 $2\leqslant r\leqslant n\leqslant2^m$,我们有如下结论:

推论 1　对于码长为 $2\leqslant r\leqslant n\leqslant2^m$ 的 Sub-Alternant 码,当 $n\to\infty$时,存在长的 Sub-Alternant 码可以渐进达到 GV 界.

为了考察 Sub-Alternant 码与 BCH 码之间的关系,易知以式(1)为校验矩阵的 $RS(n,\delta)$码对应的生成矩阵为

$$\boldsymbol{G}_{RS(n,\delta)}=\begin{pmatrix}1&1&\cdots&1\\ 1&\alpha&\cdots&\alpha^{n-1}\\ \vdots&\vdots&\vdots&\vdots\\ 1&\alpha^{n-\delta}&\cdots&\alpha^{(n-\delta)(n-1)}\end{pmatrix}$$

因此当 $\boldsymbol{Y}=(\tau_1,\tau_2,\cdots,\tau_n)$取上述矩阵的行向量时，$S-A_r(\boldsymbol{a},\boldsymbol{Y})$的校验矩阵即为 BCH 码的校验矩阵.因此 BCH 码属于 Sub-Alternant 码的范畴.

4 总结与展望

本文基于 MDS 码的 Hamming 重量分布提出一类新的 Alternant 子类码，这是除了 Goppa 码之外另一类新的可以渐进达到 GV 界的 Alternant 子类码，并且该 Alternant 子类码包含 BCH 码.本文提出的渐进好的 Alternant 子类码的优势在于其码长可以在 2 到 2^m 之间自由变换，而渐进好的二元 Goppa 码其码长限制在 $n=2^m$.另一方面，目前渐进好的纠错码都是存在性的，至于如何具体构造出渐进好的纠错码则需要进一步的研究.

参考文献

[1] MacWilliams F J, Sloane N J A. The Theory of Error-Correcting Codes[M]. Amsterdam: The Netherlands: North-Holland, 1981.1 – 369.

[2] 冯克勤.纠错码的代数理论[M].北京:清华大学出版社，2005.1 – 80.
Keqin Feng. The Algebraic Theory of Error-Correcting Codes [M]. Beijing: Tsinghua University Press, 2005.1 – 80. (in Chinese)

[3] Berlekamp E. Goppa codes[J]. IEEE Transactions on Information Theory, 1973, 19(5): 590 – 592.

[4] Martinez-Perez C, Willems W. Is the class of cyclic codes asymptotically good? [J]. IEEE Transactions on Information Theory, 2006, 52(2): 696 – 700.

[5] Ezerman M F, Grassl M, Sole P. The weights in MDS codes [J]. IEEE Transactions on Information Theory, 2011, 57(1): 392 – 396.

作者简介

樊继豪 男，1987 年 5 月出生，江苏连云港人.2009 年毕业于兰州大学数学与统计学院应用数学专业，现为东南大学计算机科学与工程学院博士研究生，研究方向为差错控制编码与量子纠错码理论.
E-mail: fanjh12@seu.edu.cn

陈汉武(通信作者) 男，1955 年出生于江苏南京.分别于 1981 年、2000 年在东南大学、日本国立山口大学获理学学士和理工学博士学位.现为东南大学教授、博士生导师.主要研究领域为经典信息论，量子信息与量子计算，数理解析.
E-mail: hw_chen@seu.edu.cn