ENUMERATION RESULTS FOR THE CODEWORDS HAVING NO INNER PERIODS IN REED-SOLOMON CODES*

YANG YI-XIAN (杨义先)

(*Beijing University of Posts and Telecommunications, Beijing* 100088, *PRC*)

Received April 12, 1990.

Keywords: Reed-Solomon code, error-correcting code, sequence.

I. Description for the Problem

Let $\boldsymbol{x}=(x_0,\cdots,x_{n-1})$ be a sequence in the finite field $GF(q)$ with length n, $S^i\boldsymbol{x}$ is the i-cyclic shift of $\boldsymbol{x}$, i.e. $S^i\boldsymbol{x}=(x_i,x_{i+1},\cdots,x_{i-1})$ (where $i+1$ means $(i+1)\bmod n$). If there exists a positive integer $0<r\leqslant n$ making $S^r\boldsymbol{x}=\boldsymbol{x}+(u,u,\cdots,u)$ hold for some $u\in GF(q)$, then the r is called one of the generalized periods of this sequence $\boldsymbol{x}$. The least one $r_{\min}$ of such periods is called the minimum generalized period of $\boldsymbol{x}$. In particular, if $r_{\min}=n$ (i.e. the minimum generalized period equals the length of the sequence), then we say that this sequence has no inner period. It is clear that the definitions for the normal period and the generalized period mean the same thing when $u=0$.

Reed-Solomon (R-S) code is one of the most important cyclic codes[1]. Nguyen and Massey found, in 1988, that the R-S codewords with no inner period can be widely used in the construction of sequences with perfect generalized Hamming correlative property. What is the exact number T_0 of R-S codewords having no inner period? This problem remains open by now, although Nguyen has found the lower bound[2] $T_0\geqslant(q-1)q^{k-1}$. By the Fourier transform in $GF(q)$ and the Polya's enumeration formula, the above open problem was solved. In addition, the exact number of random sequences in $GF(q)$ with minimum generalized period n was also shown in Section III.

II. Exact Value for T_0

The exact number T_0 of R-S codewords with no inner period was presented in this section.

Let g be a primitive element in $GF(q)$ ($q=p^m$, p being prime number). The Fourier transform coefficients vector $\boldsymbol{V}=(V_0,V_1,\cdots,V_{N-1})$ of the $N=q-1$ dimensional vector $\boldsymbol{v}=(v_0,v_1,\cdots,v_{N-1})$ in $GF(q)$ was defined as[1]:

* Project supported by the National Natural Science Fund for Youth.

$$V_j=\sum_{i=0}^{N-1} v_i g^{ij} \quad (0\leqslant j\leqslant N-1). \tag{1}$$

The generator matrix G for (N, k, d) R-S code $(N=q-1, d=N-k+1)$ is formulated as:

$$G=\begin{bmatrix} 1 & 1 & 1 & \cdots & 1 \\ 1 & g & g^2 & \cdots & g^{N-1} \\ 1 & g^2 & g^4 & \cdots & g^{2(N-1)} \\ \cdot & \cdot & \cdot & \cdot\ \cdot\ \cdot & \\ 1 & g^{k-1} & g^{2(k-1)} & \cdots & g^{(k-1)(N-1)} \end{bmatrix}. \tag{2}$$

The R-S codeword x was determined by its information vector $a=(a_0, \cdots, a_{k-1})(a_i \in GF(q))$ in the form $x=aG=(x_0, \cdots, x_{N-1})$, where $x_i=\sum_{j=0}^{k-1} a_j g^{ij} (0\leqslant i\leqslant N-1)$. So that $S^r x=(x_r, x_{r+1}, \cdots, x_{r-1})=(y_0, y_1, \cdots, y_{N-1})$ with $y_i=x_{i+r}=\sum_{j=0}^{k-1} a_j g^{(r+i)j}$.

From the definition, we know that r is one of the generalized periods of codeword x iff there exists some $u \in GF(q)$ such that

$$S^r x - x = (u, u, \cdots, u). \tag{3}$$

By carrying out the Fourier transform (1) on both sides of (3), and by orthogonal property, Eq. (3) was changed to

$$(0, \cdots, 0, a_{k-1}(1-g^{(k-1)r}), a_{k-2}(1-g^{(k-2)r}), \cdots, a_1(1-g^r)) = (-u, 0, 0, \cdots, 0). \tag{4}$$

From the inverse property, we know that Eqs. (3) and (4) are equivalent to each other. By comparing the first component on both sides of (4), we have $u=0$. Therefore (4) is also equivalent to

$$(a_{k-1}(1-g^{(k-1)r}), \cdots, a_1(1-g^r)) = (0, 0, \cdots, 0). \tag{5}$$

By now we have proved the following lemma :

Lemma 1. *The R-S codeword x with information vector $(a_0, a_1, \cdots, a_{k-1})$ has r as one of its generalized periods iff $(a_1, a_2, \cdots, a_{k-1}) = (a_1 g^r, a_2 g^{2r}, \cdots, a_{k-1} g^{(k-1)r})$.*

Let $A_r = \{ a=(a_0, \cdots, a_{k-1})$: the R-S codeword $x=aG$ is of generalized period $r \}$ $(1\leqslant r\leqslant N-1)$, then the set D of information vectors corresponding to those R-S codewords having inner periods is $D=A_1 U A_2 U \cdots U A_{N-1}$. The information vector a and the codeword x are uniquely determined by each other. There are q^k codewords in (N, k, d) R-S code, therefore there are $|A_1 U A_2 U \cdots U A_{N-1}|$ R-S codewords having inner periods, i.e. the exact number T_0 of codewords having no inner period is

$$T_0 = q^k - |A_1 U A_2 U \cdots U A_{N-1}|, \tag{6}$$

where $|X|$ means the cardinal number of the set X.

In another aspect, by Lemma 1, A_r can be rewritten as:

$$A_r = \{ a=(a_0,\cdots,a_{k-1}) : (a_1,\cdots,a_{k-1})=(a_1 g^r,\cdots,a_{k-1} g^{(k-1)r}) \}. \tag{7}$$

Because g is a primitive element in $GF(q)$, $g^s=1$ is equivalent to $N|s$. Hence $(a_1,\cdots,a_{k-1})=(a_1 g^r,\cdots,a_{k-1} g^{(k-1)r})$ iff

$$a_m = \begin{cases} 0 & \text{if } N\nmid(mr) \\ \text{arbitrary} & \text{if } N|(mr) \end{cases} \qquad (1\leqslant m\leqslant k-1),$$

while $N|(mr)$ is equivalent to $N/gcd(N,r)\,|m$, hence there are $f(k-1, N/gcd(N,r))$ integers m in the range of $1\leqslant m\leqslant k-1$ such that (mr) can be divided by N, (where and from now on $f(u,v) =: \lfloor u/v \rfloor$, it means the number of integers s in the range $1\leqslant s\leqslant u$ such that s can be divided by v, $\lfloor x \rfloor$ means the integer part of the real number x). From (7), for any $1\leqslant r\leqslant N-1$ we have

$$|A_r| = q^{(1+f(k-1, N/gcd(N,r)))}, \tag{8}$$

where $gcd(\cdot,\cdot)$ means the greatest common divisor function.

By a similar method, we know that for any $1\leqslant s<r\leqslant N-1$ the necessary and sufficient condition for the simultaneous equations

$$\begin{cases} (a_1,\cdots,a_{k-1})=(a_1 g^r, a_2 g^{2r},\cdots,a_{k-1} g^{(k-1)r}). \\ (a_1,\cdots,a_{k-1})=(a_1 g^s, a_2 g^{2s},\cdots,a_{k-1} g^{(k-1)s}) \end{cases}$$

is $$a_m = \begin{cases} \text{arbitrary} & \text{if } N|(mr) \text{ and } N|(ms) \\ 0 & \text{otherwise} \end{cases} \qquad (1\leqslant m\leqslant k-1),$$

while $N|(mr)$ and $N|(ms)$ are equivalent to $N/gcd(N,r)\,|m$ and $N/gcd(N,s)\,|m$, and also to $LCM(N/gcd(N,r),\ N/gcd(N,s))\,|m$, where $LCM(\cdot,\cdot)$ is the least common multiplier function.

Then we know that for any $1\leqslant s<r\leqslant N-1$,

$$|A_r \cap A_s| = q^{(1+f(k-1, LCM(N/gcd(N,r),\ N/gcd(N,s))))}. \tag{9}$$

In general, it can be proved that for any $1\leqslant i_1<i_2<\cdots<i_w\leqslant N-1$

$$|A_{i1} \cap A_{i2} \cap \cdots \cap A_{iw}| = q^{(1+f(k-1, LCM(N/gcd(N,i_1),\cdots,N/gcd(N,i_w))))}. \tag{10}$$

holds.

Finally according to the famous Polya's enumeration formula[5] and Eqs.(8), (9), (10), we have

$$|A_1 U A_2 U \cdots U A_{N-1}| = \sum_{w=1}^{N-1} (-1)^{w-1} \sum_{1\leqslant i_1<i_2<\cdots<i_w\leqslant N-1} |A_{i1} \cap A_{i2} \cap \cdots \cap A_{iw}|$$

$$= \sum_{w=1}^{N-1} (-1)^{w-1} \sum_{1 \leqslant i_1 < i_2 < \cdots < i_w \leqslant N-1} q^{(1+f(k-1, LCM(N/gcd(N, i_1), \cdots, N/gcd(N, i_w))))}.$$

By setting the above equation into (6), the exact value for T_0 is obtained.

$$T_0 = q^k + \sum_{w=1}^{N-1} (-1)^w \sum_{1 \leqslant i_1 < i_2 < \cdots < i_w \leqslant N-1} q^{(1+f(k-1, LCM(N/gcd(N, i_1), \cdots, N/gcd(N, i_w))))},$$

where $f(m, n) =: \lfloor m/n \rfloor$.

III. Enumeration for Random Sequences With Generalized Period n

That an R-S codeword x has no inner period means that both the minimum generalized period and the codeword length of x are equal to each other. Therefore the value of T_0 in the above section equals the number of codewords such that $r_{\min} = N$. In this section we consider a more general problem and show the exact number of random sequences in $GF(q)$ with minimum generalized period n (n is any integer).

Lemma 2. *Let $x = (x_0, \cdots, x_{n-1})$ be an arbitrary sequence in $GF(q)$, $r_{\min}$ is the minimum generalized period of x, then r is one of the generalized periods of x iff $r_{\min} \mid r$.*

Proof. The sufficiency is clear.

Necessity. By counterevidence.

Let $r_{\min} \nmid r$, i.e. there exist some integers m and $0 < e < r_{\min}$ such that $r = mr_{\min} + e$.

Because both of the $r_{\min}$ and r are the generalized periods of x, there exist some $u, v \in GF(q)$ such that $S^r x - x = (u, \cdots u)$ and $S^{r_{\min}} x - x = (v, \cdots, v)$.

$$S^r x - x = S^{mr_{\min} + e} x - x = S^e (S^{mr_{\min}} x) - x = S^e (x + (mv, \cdots mv)) - x = S^e x - x + (mv, \cdots mv).$$

Substituting $S^r x - x = (u, u, \cdots, u)$ into the above equation, we have $S^e x - x = (u - mv, u - mv, \cdots, u - mv)$. This means that e $(0 < e < r_{\min})$ is also one of the generalized periods of x, which is contradictory to the minimum property of $r_{\min}$.

Q.E.D.

Definition. Let x and y be two vectors with lengths m and n respectively. If $(x, x, \cdots) = (y, y, \cdots)$, then x is called the cyclic repeat of y.

Let $g(n)$ be the number of $GF(q)$ random sequences with minimum generalized period of n, $h(n)$ be the number of $GF(q)$ random sequences with generalized period of n.

By Lemma 3, the relationship between $g(n)$ and $h(n)$ can be stated as

$$\sum_{r \mid n} g(r) = h(n).$$

Then from the Mobius inverse formula[4], we have

$$g(n) = \sum_{d \mid n} \mu(d) h(n/d), \tag{11}$$

where $\mu(\cdot)$ is the Mobius function, i.e.

$$\mu(n)=\begin{cases}1 & \text{if } n=1,\\ 0 & \text{if } n \text{ contains a square factor},\\ (-1)^r & \text{if } n \text{ is the production of } r \text{ different primes.}\end{cases}$$

For the computation of $h(n)$, it should be noted that the possible lengths of sequences with generalized period of n are $n, 2n, 3n, rn, \cdots$. In another aspect, it is easy to verify the following 4 results :

(a) Any $GF(q)$ sequence with length n has n as one of its generalized periods. The number of such sequences is $B_1=q^n$.

(b) The sequence $(a_1,\cdots,a_n,a_{n+1},\cdots,a_{2n})$ having generalized period n is not the cyclic repeat of those sequences in (a) iff there exist some $u\in GF(q)$ such that $(a_1,\cdots,a_n,a_{n+1},\cdots,a_{2n})=(a_1,\cdots,a_n,a_1+u,\cdots,a_n+u)$ and $2u=0, u\neq 0, (u\in GF(q))$.

When $gcd(q,2)>1$, $2u=0$ holds for any u. Hence the number of sequences of length $2n$ with generalized period n having no cyclic repeat with the sequences in (a) equals $B_2=(q-1)q^n\varphi(gcd(q,2))$, where

$$\varphi(x)=\begin{cases}0 & \text{if } x=1,\\ 1 & \text{if } x>1.\end{cases}$$

(c) In general, for any $1\leqslant r\leqslant p$ there are $B_r=(q-1)q^n\varphi(gcd(q,r))$ sequences of length rn with generalized period n having no cyclic repeat with the above sequences in (a), (b), etc.

(d) For any sequence x of length $(p+1)n,(p+2)n,\cdots$ if x has n as one of its generalized periods then x is surely the cyclic repeat of some sequences in (a), (b) or (c).

By now, we have known that

$$h(n)=\sum_{r=1}^{p}B_r=q^n+\sum_{r=2}^{p}(q-1)q^n\varphi(gcd(q,r))=q^n+(q-1)q^n(p-\varphi(p))=q^{n+1}.$$

where $\varphi(p)$ is the Euler function[1], for p is prime, $\varphi(p)=p-1$.

Finally by setting the value of $h(n)$ into (11), the number of $GF(q)$ random sequences with generalized period n was derived :

$$g(n)=\sum_{d|n}\mu(d)q^{n/d+1}.$$

Thanks are due to Prof. Hu Zheng-ming and Dr. Lin Xu-duan for the helpful discussions.

REFERENCES

[1] Macwilliams, F. & Solane, N., *The Theory of Error-Correcting Codes*, Part I, North-Holland, New York, 1977.

[2] Nguyen, Q., Families of sequences with optimal generalized hamming correlation properties, *Problems of Control and Information Theory*, **17**(1988), 3 :117.

[3] Blahut, R., *Theory and Practice of Error Control Codes*, Addison-Wesley, London, 1983.

[4] Van Lint, J., *Introduction to Coding Theory*, Springer-Verlag, New York, 1982.

[5] Tucker, A., *Applied Combinatorics* (second edition), John Wiley, 1984.