CYBERNETICS AND CONTROL THEORY

R. E. Krichevskii

1970

SovietRxiv

Abstract

Full Text

UDC 519.92

CYBERNETICS AND CONTROL THEORY

R. E. Krichevskii

ON THE NUMBER OF ERRORS CORRECTED BY A REED-MULLER CODE

(Presented by Academician S. L. Sobolev, 15 VIII 1969)

1. The well-known Reed-Muller codes were constructed by D. E. Muller in ([1]). A binary Reed-Muller code of order r with block length $n = 2^m$ (an (m, r)-code) has $C_m^0 + C_m^1 + \cdots + C_m^r$ information positions and corrects any combination of $2^{m-r-1} - 1$ errors, i.e., spheres of such radius with centers at code points do not intersect.* The code points of the (m, r)-code are the sets of values assumed by Zhegalkin polynomials (polynomials modulo 2) of degree not exceeding r in m variables on all possible 2^m combinations of values of the variables.

I. S. Reed ([2]) found an easily implementable decoding algorithm for such codes. It was noted that it also permits correction of some combinations containing more than $2^{m-r-1} - 1$ errors; however, estimates of the number of such combinations were not carried out. In the same place I. S. Reed speaks of attempts to prove that, when decoded by his method, these codes will satisfy Shannon' s fundamental theorem ([3]) for a noisy channel.

In the present note it is shown that, as $m \to \infty$, a Reed-Muller (m, r)-code corrects almost any combination of

$$T_1 = \frac{n}{2}\left[1 - \left(\frac{a \cdot 2^r \log n}{n}\right)^{2^{-(r+1)}}\right]$$

errors (a is a certain absolute constant). However, already the probability of correcting

$$T_2 = \frac{n}{2}\left[1 - \left(\frac{2^r}{n}\right)^{2^{-(r+1)}}\right]$$

errors is less than one. In other words, spheres of radius T_1 with centers at code points almost do not intersect. If, however, the code points are surrounded by spheres of radius T_2, then a non-vanishing fraction of the points of each sphere

will be decoded by Reed' s algorithm into the centers of other spheres. To formulate the result more precisely, denote by $S_f(n,r)$ the set of n-dimensional binary vectors which, under decoding, pass into the vector f belonging to the (m,r)-code. Denote by $\pi(n,r,T)$ the ratio of the number of vectors from $S_f(n,r)$ that differ from f in T coordinates to the total number C_n^T of vectors that differ from f in T coordinates. From the properties of Reed' s algorithm it follows that $\pi(n,r,T)$ does not depend on f. To estimate $\pi(n,r,T)$, it is sufficient to consider the set $S_O(n,r)$, $O=(0,0,\ldots,0)$. The main result is that for $T \leq T_1$, $\pi(n,r,T) \to 1$, while for $T \geq T_2$, $\pi(n,r,T) < 1-\varepsilon$, $\varepsilon > 0$. As is easy to see, the number T_1 of errors corrected by a Reed–Muller (m,r)-code almost always exceeds the number $2^{m-r-1}-1$ of errors guaranteed to be corrected. However, by studying the number T_2, starting from which a nonzero fraction of errors is no longer corrected, we shall show that the decoding system described by I. S. Reed does not satisfy Shannon' s fundamental theorem.

The estimates obtained make it possible to find relations between the block length, redundancy, and reliability of coding when Reed–Muller codes are used in a BSC. Relations of this kind were studied by the author in ([4]) for a noiseless channel.

* The distance between two n-dimensional binary vectors is equal to the number of noncoinciding coordinates (Hamming distance).

2. Let us combine the coordinates of n-dimensional vectors into $2^{m-\rho}$ groups of 2^ρ coordinates in each, $0 \leq \rho \leq r$. Denote by $p(n,\rho,T,S_+)$ the ratio of the number of n-dimensional vectors with T ones, in which S_+ groups contain an even number of ones, to C_n^T.

Lemma 1. The following estimates hold for $\pi(n,r,T)$:

$$1-\sum_{\rho=0}^{r} C_m^\rho \sum_{S_+=0}^{2^{m-\rho-1}-1} p(n,\rho,T,S_+) \leq \pi(n,r,T) \leq \sum_{S_+=2^{m-r-1}}^{2^{m-r}} p(n,r,T,S_+).$$

These estimates follow from Reed' s algorithm ([2]).

Introduce some notation.

$$T/n=\tau, \qquad 1-\tau=\tau_1, \qquad 2^r=\chi, \qquad S_-=n/\chi-S_+,$$

$$P_+(\varphi)=\sum_{j \text{ even}} C_\chi^j (\tau e^{i\varphi})^j \tau_1^{\chi-j} = \frac{1}{2}\left[(\tau e^{i\varphi}+\tau_1)^\chi + (\tau e^{i\varphi}-\tau_1)^\chi\right],$$

$$P_+ = P_+(0),$$

$$P_-(\varphi) = \sum_{j \text{ odd}} C_\chi^j (\tau e^{i\varphi})^j \tau_1^{\chi-j} = \frac{1}{2}\left[(\tau e^{i\varphi} + \tau_1)^\chi - (\tau e^{i\varphi} - \tau_1)^\chi\right],$$

$$P_- = P_-(0),$$

$$\lambda_+(\varphi) = \frac{P_+(\varphi)}{P_+}, \qquad \lambda_-(\varphi) = \frac{P_-(\varphi)}{P_-}.$$

Lemma 2. As $n \to \infty$ and $T \to \infty$, the inequality holds

$$\text{a)} \quad P(n, r, T, S_+) \le \sqrt{2\pi T \tau_1}\, C_{n/\chi}^{S_+} P_+^{S_+} P_-^{S_-} (1 + o(1)).$$

There is also the more precise estimate

$$\text{b)} \quad P(n, r, T, S_+) = \sqrt{\frac{T\tau_1}{2\pi}}\, C_{n/\chi}^{S_+} P_+^{S_+} P_-^{S_-} \int_{-\pi/2}^{\pi/2} [\lambda_+(\varphi)]^{S_+} [\lambda_-(\varphi)]^{S_-} e^{-iT\varphi}\, d\varphi \times$$

$$\times (1 + o(1)).$$

Proof. By virtue of its definition,

$$p(n, r, T, S_+) = \frac{1}{C_n^T} \sum C_\chi^{j_1} C_\chi^{j_2} \cdots C_\chi^{j_{n/\chi}}, \tag{1}$$

where the sum is taken over all sets of indices $j_1, \ldots, j_{n/\chi}$, with S_+ of these indices even and

$$\sum_{k=1}^{n/\chi} j_k = T. \tag{2}$$

We may further carry out the summation over sets in which the first S_+ indices are even, and multiply the result by $C_{n/\chi}^{S_+}$. Applying Stirling' s formula, we obtain from (1):

$$p(n, r, T, S_+) = \sqrt{2\pi T \tau_1}\, C_{n/\chi}^{S_+} \sum (C_\chi^{j_1} \tau^{j_1} \tau_1^{\chi - j_1}) \cdots (C_\chi^{j_{n/\chi}} \tau^{j_{n/\chi}} \tau_1^{\chi - j_{n/\chi}}), \tag{3}$$

where the summation is over sets of indices satisfying (2), in which the first S_+ indices are even. Hence the first assertion of the lemma follows. Notice now that the sum in (3) is the Fourier coefficient of the function

$$[P_+(\varphi)]^{S_+}[P_-(\varphi)]^{S_-}$$

and, as such, is equal to

$$\frac{1}{2\pi}\int_{-\pi/2}^{\pi/2}[P_+(\varphi)]^{S_+}\times$$

$$\times[P_-(\varphi)]^{S_-}e^{-iT\varphi}\,d\varphi$$

(integration may be over any interval of length 2π). This gives us the second assertion of the lemma.

Lemma 3. Let $n\to\infty$, $\tau\leqslant\frac{1}{2}-\frac{1}{2}(\alpha\varkappa/n)^{1/2\varkappa}$. Then

$$p(n,r,T,S_+)\leqslant 2C_{n/\varkappa}^{S_+}P_+^{S_+}P_-^{S_-}(1+o(1)).$$

Proof. For the proof it is necessary to estimate the integral appearing in Lemma 2. We shall use the idea of the Laplace method for large numbers ([5]). Choose ε so that integration over the set of φ, $\pi/2\geqslant|\varphi|\geqslant\varepsilon$, gives a small contribution to the integral, while the integral over the set of φ, $|\varphi|\leqslant\varepsilon$, is determined by two terms of the Taylor expansion of the integrand. Analogous estimates are carried out for the interval $-3/2\pi\leqslant\varphi\leqslant-1/2\pi$. For $\varkappa=O(\log n)$ one may put

$$\varepsilon=(n^{1/3}\varkappa^{2/3}\log n)^{-1},$$

while for $\varkappa/\log n\to\infty$ one may put

$$\varepsilon=\left(\frac{\varkappa}{n}\right)^{1/3}\left(\log\frac{n}{\varkappa}\right)^{-2}.$$

Theorem. Let $n\to\infty$, $\varkappa\log^2 n/n\to 0$. Then:

a) if

$$T\leqslant T_1=\frac{n}{2}\left[1-\left(\frac{a\varkappa\log n}{n}\right)^{1/2\varkappa}\right],$$

where a is a certain absolute constant;

b) for any $a>0$ there is an $\varepsilon>0$ such that, for

$$T \geqslant T_2 = \frac{n}{2}\left[1 - \left(\frac{a\varkappa}{n}\right)^{1/2\varkappa}\right]$$

and sufficiently large n, $\pi(n, r, T) < 1 - \varepsilon$.

Remark. The presence of the factor $\log n$ with $a\varkappa/n$ accounts for the discrepancy between T_1 and T_2. This factor can be reduced, for example, to a quantity of order $\varkappa \log\log n$ when $\varkappa$ does not grow too rapidly. We shall not strive here to minimize this factor, so as not to complicate the formulation.

Proof. a) It is known that for $T \leqslant n/2\varkappa$, $\pi(n, r, T) = 1$. Consequently, it is enough to consider $T \geqslant n/2\varkappa$. By the condition, $n/\varkappa \to \infty$, and Lemma 2 is applicable. Substituting into the lower estimate of Lemma 1 estimate a) of Lemma 2 and choosing the constant a in the formula for T_1 sufficiently large, we obtain assertion a).

b) We shall use the upper estimate of Lemma 1, estimate b) of Lemma 2, and also the fact that $p(n, r, T, S_+) = 0$ if the numbers S_+ and T have opposite parity. We have

$$\pi(n, r, T) \leqslant \sum_{S_+ = n/2\varkappa}^{n/\varkappa}{}' p(n, r, T, S_+) \leqslant 2 \sum_{S_+ = n/2\varkappa}^{n/\varkappa}{}' C_{n/\varkappa}^{S_+} P_+^{S_+} P_-^{S_-}(1 + o(1)),$$

where the prime at the summation sign means that S_+ runs only through values of the same parity as T. To the terms on the right-hand side we apply the local limit theorem ([6]). Then the usual arguments convince us of the validity of the integral theorem here as well. The standard deviation is

$$\sigma = \sqrt{(n/\varkappa)P_+P_-} \sim \frac{1}{2}\sqrt{\frac{n}{\varkappa}},$$

$$\left(\frac{n}{2\varkappa} - \frac{n}{\varkappa}P_+\right)/\sigma \geqslant -\sqrt{\frac{n}{\varkappa}}\left(1 - 2\frac{T_2}{n}\right)^{\varkappa} = -\sqrt{a}$$

for $T \geqslant T_2$. Thus,

$$\pi(n, r, T) \leqslant \frac{1}{2\pi}\int_{-\sqrt{a}}^{\infty} e^{-x^2/2}\,dx < 1 - \varepsilon, \qquad \varepsilon > 0.$$

Assertion b) is proved.

The following corollaries follow from the theorem. Consider the (m, r)-Reed-Muller code; T_1 is the boundary up to which almost all errors are corrected;

T_2 is the boundary beginning with which errors are not corrected with nonzero probability, $m \to \infty$.

Corollary 1. If $2^r = o(\log n)$ and, in particular, if $r = \text{const}$, then for every $\varepsilon > 0$ there exists an n so large that $T_1 \geq \frac{1}{2} n(1-\varepsilon)$.

Corollary 2. If $2^r = \gamma \log n$, where $\gamma = \text{const}$, then for every $\varepsilon > 0$ there exists an n so large that

$$T_1 \geq \frac{1}{2} n \left(1 - e^{-1/2\gamma} - \varepsilon\right), \qquad T_2 \leq \frac{1}{2} n \left(1 - e^{-1/2\gamma} + \varepsilon\right).$$

Corollary 3. If $\lim\limits_{m\to\infty} 2^{1-m} m^2 = 0$, then

$$T_1 \sim 2^{m-r-2}(m-r), \qquad T_2 \sim 2^{m-r-2}(m-r-\log m).$$

Corollary 4. In transmission over a binary symmetric channel with distortion probability p, the (m, r)-code ensures, as $m \to \infty$, an arbitrarily small decoding-error probability for any p, if $2^r = o(\log n)$, and for $p < \frac{1}{2}(1 - e^{1/2\gamma})$, if $2^r = \gamma \log n$.

Corollary 5. If the number of information symbols of the Reed-Muller code is proportional to the block length n (nonzero transmission rate over a BSC), then the decoding error does not tend to zero. Thus, the Reed decoding procedure does not satisfy Shannon' s fundamental theorem (3).

Institute of Mathematics
Siberian Branch of the Academy of Sciences of the USSR
Novosibirsk

Received
7 VIII 1969

CITED LITERATURE

1. D. E. Muller, I. R. E. Trans., EC-3 (1954).
2. I. S. Reed, *Kiberneticheskii sbornik*, vol. 1, IL, 1960.
3. K. Shannon, *Sbornik. Raboty po teorii informatsii i kibernetike*, IL, 1963.
4. R. E. Krichevskii, DAN, 171, No. 1 (1966).
5. G. Pólya, G. Szegő, *Problems and Theorems in Analysis*, Part I, Moscow, 1956, pp. 105-111.
6. W. Feller, *An Introduction to Probability Theory and Its Applications*, Moscow, Chapter 7, 1964.

Note: Figure translations are in progress. See original paper for figures.

Source: Math-Net.Ru and CyberLeninka. Machine translation. Verify with the original.