A GENERAL METHOD FOR CONSTRUCTING ASYMMETRIC CODING SYSTEMS CONNECTED WITH THE SOLUTION OF DIXON' S COMBINATORIAL PROBLEM

R. R. VARSHAMOV

1970

SovietRxiv

Abstract

Full Text

UDC 519.95

CYBERNETICS AND CONTROL THEORY

R. R. VARSHAMOV

A GENERAL METHOD FOR CONSTRUCTING ASYMMETRIC CODING SYSTEMS CONNECTED WITH THE SOLUTION OF DIXON' S COMBINATORIAL PROBLEM

(Presented by Academician V. A. Trapeznikov, 26 XII 1969)

At present, in connection with wide applications, the problem of synthesizing coding systems with an asymmetric character of channel distortions is topical, including distortions of the type of insertions and deletions of symbols in a signal ([1–5]). The mathematical structure of such coding systems has its own particular specificity, distinguishing it from the majority of other systems well studied in the literature. Therefore, for its practical investigation it is necessary to introduce new mathematical ideas. In this note a general method is proposed for constructing efficient (with respect to transmission rate) coding systems that correct any preassigned number of asymmetric errors, connected with the solution of one of the unsolved problems of additive number theory ([6]).

By an r-fold asymmetric distortion of a signal $\gamma = (\gamma_1, \gamma_2, \dots, \gamma_n)$ we shall mean the ordinary addition to it modulo q (where q is the base of the code) of a noise $\varepsilon = (\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n)$ satisfying the following conditions: $|\varepsilon| = r$, where $|\varepsilon| = \sum_{i=1}^{n} \varepsilon_i$, and

$$|\gamma + \varepsilon| = |\gamma| + |\varepsilon|. \tag{1}$$

Let $W_m(f, x, \alpha)$ denote the least nonnegative residue modulo m of the expression $-\alpha + \sum_{i=1}^{n} f(i)x_i$, where $x = (x_1, x_2, \dots, x_n)$ is an arbitrary sequence of length n, each element x_i of which takes the values $0, 1, \dots, q-1$; $f(z)$ is some integer-valued function defined on the given set of natural numbers $1, 2, \dots, n$; and α is any integer.

Consider the set of all possible solutions of the equation $W_m(f, x, \alpha) = 0$. As is easy to understand, in order that this set be a code correcting t and fewer asymmetric errors, in view of (1) it is sufficient that, for distinct (nonempty) sets $U = \{u\}$ of $k \leq t$ (not necessarily distinct) natural numbers $\leq n$, the expressions

$$F(U) = \sum_{i=1}^{k} f(u_i)$$

take values distinct modulo m and not congruent to zero. Thus the problem of synthesizing asymmetric coding systems is in fact reduced to finding a function $f_t(z)$ $(t \geq 1)$ satisfying the indicated requirements. In the case $t = 1$, obviously, $F(U) = f_1(u)$. This suggests that, taking $f_1(z) = z$ and $m > n$, one can construct codes correcting single asymmetric errors. And indeed this is so ([3]); moreover, as was shown ([7]), the codes constructed–

Thus, for $m = n + 1$ and $\alpha = 0$, they are the best among the known codes correcting single errors. For $t = 2$, Zinger [8], using methods of finite projective geometry, succeeded in giving an algorithm that permits computing the values of the function $f_2(z)$ for all n that are powers of a prime number, and $m = n^2+n+1$. However, in the general case, for $t > 2$, the solution of this problem is not known. But since it is closely connected with the problem of synthesizing coding systems that correct multiple asymmetric errors, in what follows, in developing such signal systems, we shall thereby solve this combinatorial problem.

Let $P = (p_1, p_2, \dots, p_s)$ $(1 \leq s \leq n)$ be an arbitrary vector $(p_i \geq 0)$,

$$\sigma_n(x, P) = \sum x_{i_1}^{p_1} x_{i_2}^{p_2} \cdots x_{i_s}^{p_s},$$

where the summation is taken over all s permutations $(i_1, i_2, \dots, i_s)$ of the integers $1, 2, \dots, n$, and let

$$\sigma_n^{(P)}(x) = \prod_{i=1}^{s} \sigma_n(x, p_i),$$

i.e.

$$\sigma_n^{(P)}(x) = \sum_{\delta} x_1^{q_1} x_2^{q_2} \cdots x_n^{q_n}. \tag{2}$$

The summation in (2) is over all representations $(q_1, q_2, \dots, q_n) = P\delta$, where $\delta = (\delta_{i,j})$ is an arbitrary $(0, 1)$-matrix of dimension $s \times n$ satisfying the condition

$$\sum_{j=1}^{n} \delta_{i,j} = 1. \tag{3}$$

Lemma 1. Let δ' be an arbitrary $(0, 1)$-matrix satisfying (3), and let $Q = P\delta'$. Then $\{x, Q\} \subseteq \{x, P\}$, where

$$\{x, P\} = \bigcup_{\delta} x_1^{q_1} x_2^{q_2} \cdots x_n^{q_n}.$$

In what follows we shall denote by $R^{(k)}$ monotone vectors* of dimension k. Then formula (2) can be rewritten in the form

$$\sigma_n^{(P)}(x) = \sum_{k=1}^{s} \sum_{u=1}^{t_k} C_{k,u} \sigma_n(x, R_u^{(k)}), \tag{4}$$

where t_k and $C_{k,u}$ are certain positive integers, and the following fact can be proved.

Lemma 2. Let

$$T_s = \sum_{k=2}^{s-1} t_k \qquad (s > 2)$$

and let W be a T_s-set of symmetric functions

$$\{\sigma_n(x, R_1^{(2)}), \ldots, \sigma_n(x, R_{t_2}^{(2)}), \sigma_n(x, R_1^{(3)}), \ldots, \sigma_n(x, R_{t_3}^{(3)}), \ldots, \sigma_n(x, R_1^{(s-1)}), \ldots, \sigma_n(x, R_{t_{s-1}}^{(s-1)})\},$$

and let

$$M(m) = \{m_1, m_2, \ldots, m_{T_s}\},$$

where

$$m_{T_i+j} = W_i \cup \sigma_n(x, R_j^{(i)}), \quad i = 2, \ldots, s-1, \quad j = 1, 2, \ldots, t_i; \qquad W_i = \bigcup_{u=1}^{T_i} m_u$$

$(W_2 = \emptyset)$ is a T_s-selection of the set W. Then the incidence matrix for the subsets $m_1, m_2, \ldots, m_{T_s}$ of the set W is a nonsingular upper triangular matrix.

Theorem 1. For any positive vector $P = (p_1, p_2, \ldots, p_s)$ there always exist T_s integers $a_{k,u}$ such that

$$\sigma_n^{(P)}(x) + \sum_{k=2}^{s-1} \sum_{u=1}^{t_k} a_{k,u} \sigma_n^{(R_u^{(k)})}(x) + (-1)^s (s-1)! \sigma_n(x, |P|) \equiv \sigma_n(x, P). \tag{5}$$

* We shall say that a vector $R = (r_1, r_2, \dots, r_k)$ is monotone if $r_1 \geq r_2 \geq \cdots \geq r_k$.

Proof. First note that in formula (4) $t_1 = t_s = C_{1,1} = C_{s,1} = 1$, so that

$$\sigma_n^{(P)}(x) - \sigma_n(x, |P|) = \sum_{k=2}^{s-1} \sum_{u=1}^{t_k} C_{k,u} \sigma_n(x, R_u^{(k)}) + \sigma_n(x, P). \tag{6}$$

Substituting successively in (6), in place of the vector P, the expressions $R_u^{(k)}$ $(k = 2, \dots, s-1;\ u = 1, 2, \dots, t_k)$, and taking into account (2) and Lemma 1, we obtain the system of identities:

$$\sigma_n^{(P)}(x) - \sigma_n(x, |R_u^{(k)}|) = \sum_{h=2}^{s-1} \sum_{v=1}^{t_h} C_{h,v}^{u,k} \sigma_n(x, R_v^{(h)}) + \sigma_n(x, R_u^{(k)}). \tag{I}$$

By Lemma 2 it is clear that, by an appropriate choice of suitable coefficients a and by termwise addition of the equations of system (I), one can eliminate from (6) some of the expressions occurring on its right-hand side and thus, taking into account that $|P| = |R_u^{(k)}|$, obtain the identical relation

$$\sigma_n^{(P)}(x) + \sum_{k=2}^{s-1} \sum_{u=1}^{t_k} a_{k,u} \sigma_n^{(R_u^{(k)})}(x) - \left(1 + \sum_{k=2}^{s-1} \sum_{u=1}^{t_k} a_{k,u}\right) \sigma_n(x, |P|) = \sigma_n(x, P). \tag{7}$$

Considering henceforth only binary sequences x for which $|x| = h$, where $h = 1, 2, \dots, s-1$, formula (7), since then $\sigma_n^{(R_u^{(k)})}(x) = h$ and $\sigma_n(x, P) = 0$, can be written in the form

$$h^{s-1} + \sum_{k=1}^{s-2} b_k h^k - 1 - \sum_{k=2}^{s-1} \sum_{u=1}^{t_k} a_{k,u} = 0. \tag{8}$$

The roots of equation (8) are the numbers $1, 2, \dots, s-1$; therefore its constant term is equal to $(-1)^{s-1}(s-1)!$. Substituting the value found for

$$1 + \sum_{k=2}^{s-1} \sum_{u=1}^{t_k} a_{k,u}$$

into formula (7), we obtain (5). Thus Theorem 1 is proved.

Let g be any primitive root of the prime number p, $n < p - 1$, and let $a_0, a_1, \dots, a_{t-1}$ be arbitrary integers.

Theorem 2. The set of all possible solutions of the system

$$W_{p-1}(z, x, a_0) = 0, \qquad W_p(g^{iz} - 1, x, a_i) = 0,$$

$i = 1, 2, \dots, t-1$, is a code correcting t or fewer asymmetric errors.

We give here only the outline of the proof of Theorem 2. Assuming at first that the assertion of the theorem is false and considering formula (5) for $P = (\tau, 1, \dots, 1)$, where $\tau = 1, 2, \dots$, putting $n = t = s$ and taking into account that $s = t < p$, $|R_u^{(k)}| = t - 1 + \tau$, $\sigma_n(x, P) = (t-1)!\sigma_t(x, \tau - 1)x_1 x_2 \dots x_t$, one can obtain the congruence

$$\sigma_V(g^{v-\lambda}, t - 1 + \tau) \equiv \sigma_U(g^{u-\lambda}, t - 1 + \tau) + g^{-\lambda(t-1+\tau)}(k_1 - k_2) \pmod{p},$$

where $\sigma_U(g^u, j) = \sum_{u \in U} g^{ju}$, valid for any positive integers τ, $\lambda \leq n$, and two distinct sets V and U of respectively k_1 and k_2 $(k_1, k_2 \leq t)$ natural numbers. However, a contradiction to this relation is then established, and this completes the proof of Theorem 2.

Theorem 3. The set of all possible solutions of the system $W_p(z^i, x, a_i) = 0$ $(i = 1, \dots, t)$, where the a_i are arbitrary integers and the prime $p > n$, is a code correcting t or fewer asymmetric errors.

Relying substantially on the preceding results and using certain auxiliary means, one can also solve in general form the combinatorial problem discussed above, i.e., for any integer $t \geqslant 2$ analytically construct the functions $f_t(z)$. This problem can also be formulated in a somewhat different form: for the set $0, 1, \dots, m-1$ of residues modulo m, it is required to indicate, if possible, the largest n-subset $b_1^{(t)}, b_2^{(t)}, \dots, b_n^{(t)}$, such that all sums*

$$b_{i_1}^{(t)} + b_{i_2}^{(t)} + \dots + b_{i_k}^{(t)}, \qquad \text{where } k \leqslant t,$$

are nonzero and distinct modulo m.

In the notation adopted, the following is valid.

Theorem 4. *For any integer $t \geqslant 2$, the relation*

$$b_z^{(t)} \equiv f_t(z) \pmod{m},$$

holds, where

$$f_t(z) = z + (2t-1)(p-1)\left[\frac{g^z + z - 1}{2t-1} + \frac{1}{t}\sum_{u=1}^{t-2}(pt)^u(g^{(u+1)z} - 1)\right],$$

$$\sqrt[t]{m} = cp, \qquad c = \begin{cases} (1-1/p)^{1/2}, & \text{if } t = 2, \\ t^{1-3/t}\left[(1-1/p)(2t-1)\right]^{1/t}, & \text{if } t > 2. \end{cases}$$

In conclusion, let us note that the results obtained can also be used in the synthesis of coding systems resistant to failures of the type of insertions and deletions of symbols in signals ([5]).

Computing Center
Academy of Sciences of the Armenian SSR
Yerevan

Received
25 XII 1969

CITED LITERATURE

1. W. H. Kim, C. V. Freiman, Trans. Inform. Theory, IT-5, No. 2, 62 (1959).
2. W. H. Kim, C. V. Freiman, Intern. Symposium Circuit and Information Theory, June, 1959.
3. R. R. Varshamov, T. M. Tenengolts, Avtomatika i telemekh., 26, No. 2, 288 (1965).
4. V. I. Levenshtein, DAN, 163, No. 4, 845 (1965).
5. V. I. Levenshtein, Problems of Information Transmission, 1, issue 1, 1965.
6. P. Erdős, Matem. prosv., issue 6, 315 (1961).
7. R. R. Varshamov, DAN, 164, No. 4, 757 (1965).
8. J. Singer, Trans. Am. Math. Soc., 43, 7, 347 (1938).

* It is not necessary that all elements of a sum be distinct.

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