SEPTIEME COLLOQUE SUR LE TRAITEMENT DU
SIGNAL ET SES APPLICATIONS

NICE du 28 MAI au 2 JUIN 1979

MAXIMUM RATE-MINIMUM ERROR COMPOSITE CODES FOR DIRECT-TRANSMISSION BINARY CHANNELS

Dr. Eng. Mohi-Eldin Moustafa Rateb.
Dr. Eng. Mohammed Nagi Elyousfi

28 Amin Elshamsy Stréet. Flat No. 10
Hegaz Square - Nozha - Heliopolis, CAIRO, EGYPT.

Work: Military Technical College - CAIRO
Logic-Pulse and Digital Department

RESUME

Mettre un code une source d'information de l'alphabet en utilisant un code specifique pour correction d'erreurs a proteger cette information contre le bruit (noise) n'est pas la solution optimum. La frequence d'occurrence de symboles differentes de source diffire enormement.
Un code composité d'une longueur constante "n" est ce code qui contient dans sa composition des groupes differentes de poids constant que possédent différentes capabilités de correction de fautes avec une certaine limitation d'une distance minimum l'un de l'autre. Un code composité optimum qui convient une source discret d'information avec une certaine distribution de symbols, est ce code qui met au maximum la probabilité d'un juste decoder à un degré certain de transmission.

Encoder composité cause en tous les cas une plus elevé probabilité de décodes juste et un plus élevé degré de transmission. Cela est obtenue quand les symboles de plus grande frequence d'occurrence sont encodés avec des groupes de plus grand capabilité de correction de faute et vice versa. Pour une certaine longueur, il existe plusieurs possibilités des codes composités, un déux est etre convenable pour une source specifique avec une certaine distribution.

Ce travail introduit, un guide general pour la construction de ces codes, certains courts codes composités convenable pour l' alphabet Anglais et le configuration generale de l'encoder et le decoder. En plus, comparaisons faites nous montrer que les codes composités, ont des meilleurs propriétés que les plus repultés codes de correction d'erreurs de la même longueur et même de plus les codes de grande longueur s'ils les codes composités sont proprement construit. Les résultats obtenus signifie la modification du systéme binaire de transmission qui utilize le codes normales pour correction de fautes a minimer la perte dans le temps de transmission et pour ameliorer les proprieté communication.

SUMMARY

Encoding an alphabet information source by a specific error-correcting code for protecting this information againest noise, is not the optimum solution. The frequency of occurrence of different source symbols widely differs. A composite code of a constant length "n" is such a code whose structure contains different constant-weight groups having different error-correction capabilities under certain restrections of minimum distance among them. An optimum composite code fitting a given discrete information source with given distribution of its symbols, is that code which maximizes the probability of correct decoding at a certain transmission rate. Composite encoding leads in all cases to higher correct-decoding probability and higher transmission rate. This is gained when symbols of higher frequency of occurrence are encoded by groups of higher error-correction capability and vise versa. For a given length, there exist many possible composite codes, one of them is the best to fit a specific source with a given distribution.

This work introduces a general guide for construction of these codes, some short composite codes fitting the English alphabet, and the general configuration of the encoder and decoder. In addition, comparisons showed that composite codes have better properties than the best known error-correcting codes of the same length and even of higher length, if they are properly constructed.

The results obtained, imply the modification of binary transmission system using the normal error-correcting codes for minimization the time waste in transmission and improving the communication properties.

MAXIMUM RATE-MINIMUM ERROR COMPOSITE CODES FOR DIRECT-TRANSMISSION BINARY CHANNELS

1. BASIC IDEA:

Given a discrete zero-memory information source having "q" symbols

$$S = \begin{pmatrix} S_1 & S_2 & \cdots & S_q \\ p(S_1) & p(S_2) & \cdots & p(S_q) \end{pmatrix} \quad \ldots(1)$$

Such that $\sum_{i=1}^{q} p(S_i) = 1$

Let these source symbols be encoded by binary code words of constant length "n" and transmitted over a noisy binary symmetric channel having a bit-error probability "p", then there exists a composite error-correcting code of length "n" such that the average probability of correct decoding for all encoded symbols transmitted over this channel

$$p_{c_{av}} = \left[p(S_1)\; p(S_2) \ldots p(S_q)\right] \begin{bmatrix} p_{c_1} \\ p_{c_2} \\ \vdots \\ p_{c_q} \end{bmatrix} = \text{maximum} \ldots(2)$$

This probability is higher than that obtained by encoding the same source symbols by any of the known best error-correcting codes of the same length. In many cases it exceeds it even for codes of greater lengths.

Since direct-transmission channels are our interest, the detection of errors does not provide a remarkable improvement on the transmission system properties. Thus the decisive factor is the correct decoding of received information.

For construction of a composite binary code of length "n" such that

$$n > \log_2 q \quad \ldots\ldots (3)$$

we proceed as follows:

(a) By computer or by the aid of Mac Williams identities the weight distributions of all optimum codes of the same length "n" and for different error-correction capabilities are obtained. This may include the higher efficiency nonlinear block codes.
Groups of constant-weight and of the highest possible number of code vectors are selected such that satisfy the following conditions:

- If the error correction and/or detection capability of the ith group of Hamming weight "w_i" is "t_1" then the adjacent group of correction and/or detection capabilities t_1+t_2 must has at least the weight

$$w_{i+1} = w_i + t_1 + t_2 + 1 \quad \ldots\ldots\ldots\ldots (4)$$

This condition is valid also if the selected group lf lower weight has the higher correction capability.

- If a group of Hamming weight "w_i" and error correction and/or detection capability "t", then the following (or preceeding) group of no error correction or detection capability must have the weight $w_{i\pm 1} = w_i \pm t \pm 1 \ldots\ldots\ldots\ldots (5)$

(b) On the basis of the previous rules we may construct many composite codes of the same length "n" statisfying

$$\sum_{i=1}^{N_1} \binom{n}{w_i} + \sum_{j=1}^{N_2} N_j \geq q \quad \ldots\ldots\ldots(6)$$

where

N_1 = Number of code groups with no error correction capability
w_i = The Hamming weight of the ith group.
N_2 = Number of code groups having error correction and/or detection capabilities.
N_j = Number of code vectors contained in the jth group having correction and/or detection capabilities.

(c) The symbols of higher importance are encoded by the constant-weight groups having the highest correction capabilities and those of the lowest importance by the groups having the lowest capabilities.

(d) To calculate the different probabilities of different obtained code variants.

These probabilities are to be calculated for each code group of a certain weight and certain error correction and/or detection capability according to the new structure. These are: p_{c_i} (the probability of correct decoding), p_{d_i} (the probability of detection of error) and p_{e_i} (the probability of incorrect decoding) for each code group.

Suppose that two adjacent code groups of weights and error-correction capabilities w_i, t_1 and w_j, t_2 respectivly. The second group encodes symbols of higher importance. Then if

$$w_j - w_i = t_1 + t_2 + r + 1 \quad \ldots\ldots\ldots(7)$$

If no more than t_2+r_2 errors are expected to occur, and $r = r_1 + r_2$. then:

$$p_{c_i} = \sum_{i_1=0}^{t_1} \binom{n}{i_1} p^{i_1} (1-p)^{n-i_1} \quad \ldots(8)$$

$$p_{d_i} \geq \sum_{i_2=t_1+1}^{t_1+r_1} \binom{n-w_i}{i_2} p^{i_2} (1-p)^{n-i_2} \ldots(9)$$

$$p_{e_i} = 1 - (p_{c_i} + p_{d_i}) \ldots\ldots\ldots\ldots\ldots (10)$$

$$p_{c_j} = \sum_{j_1=0}^{t_2} \binom{n}{j_1} p^{j_1} (1-p)^{n-j_1} \quad \ldots(11)$$

MAXIMUM RATE-MINIMUM ERROR COMPOSITE CODES
FOR
DIRECT-TRANSMISSION BINARY CHANNELS

$$p_{d_j} \geqslant \sum_{j_2=t_2+1}^{t_2+r_2} \binom{w_j}{j_2} p^{j_2}(1-p)^{n-j_2} \quad \ldots\ldots\ldots(12)$$

$$p_{e_j} = 1 - (p_{c_j} + p_{d_j}) \quad \ldots\ldots\ldots\ldots\ldots\ldots(13)$$

The total detection probability for the two groups can be calculated if the capabilities of the groups of higher weight than w_j and lower weight than w_i are introduced.

This means that maximizing the weight seperation between adjacent groups hands over higher detection capabilities to the entire code structure.

If a certain composite code structure leads to N_1 groups of no error correction or detection capabilities and N_2 groups of different error correction "t_η" and detection "d_ξ" capabilities each consisting of n_j encoded symbols then the average probabilities will be:

$$P_{c_{av}} = \sum_{\beta=1}^{N_1} \sum_{i_\beta=1}^{\binom{n}{i_\beta}} P_{s_{i_\beta}} (1-p)^n + \sum_{\eta=1}^{N_2} \sum_{J_\eta=1}^{r_{n\eta}} P_{s_{J_\eta}} \sum_{r_\eta=0}^{t_\eta} \binom{n}{r_\eta} p^{r_\eta}(1-p)^{n-r_\eta} \quad (14)$$

$$P_{d_{av}} = \sum_{\xi=1}^{N_2} \sum_{J_\xi=1}^{r_\xi} P_{s_{J_\xi}} \sum_{r_\xi=t_\eta+1}^{d_\xi} \binom{n}{r_\xi} p^{r_\xi}(1-p)^{n-r_\xi} \quad (15)$$

for every $\xi = \eta$

$$P_{e_{av}} = 1 - (P_{c_{av}} + P_{d_{av}}) \quad \ldots\ldots\ldots\ldots(16)$$

(e) The optimum composite code among all those code selection is that code which maximizes the probability of correct decoding for the given alphabet source.

3. GENERAL THEORY OF COMPOSITE CODES STRUCTURE:

(a) For a linear block code of length "n" having error-correction capability of "t" or less errors, the following groups of code vectors of weights(w_i) and error-correction capabilities (k) that can suit this code are at least

$$\sum_{i=2\times k+1}^{t-k} w_i \quad \ldots\ldots\ldots\ldots\ldots (17)$$

$$\text{for} \quad t \geqslant 3k + 1 \ldots\ldots\ldots\ldots\ldots (18)$$

or the code groups of no error correction capabilities

$$\sum_{j=0}^{t-1} w_j \quad \ldots\ldots(19)$$

The number of code vectors in each w_j is $\binom{n}{j}$

For any selected $k = k_{max}$ among the permissible groups we can find new groups of error correction capabilities $k-3x$, $x=1,2,3\ldots,k-3$ errors or less, satisfying the restrictions concerning the minimum distance between the adjacent groups.

(b) For a linear block code of length "n", error correction capability "t" errors with an added overall party-check there are at least

$$\sum_{i=2k+1}^{t-k+1} w_i \quad \ldots\ldots\ldots\ldots(20)$$

groups of weights w_i and correction capabilities k for

$$t \geqslant 3k \quad \ldots\ldots\ldots\ldots\ldots\ldots(21)$$

We can construct the composite code in the same manner mentioned before.

In all cases the zero code vector will have the correction capability of the lowest weight group.

In case of large "t" a variety of selections is available. One of them is the best to fit a certain alphabet. In general we may select some groups of all codes used including that having the highest capability "t" to get the best fitting.
By at least mentioned before we mean that we may find at higher weights some code groups of error-correction capabilities less than "t" satisfying the minimum distance restrictions.

For example; the BCH optimum (15,5) code with t=3 and d_{min}=7 has the following weight distribution if an overall parity is added

$w_0 = 1$; $w_8 = 30$; $w_{16} = 1$

according to (20) and (21) we can find only one code group w_3 of error correction capability k=1. By the aid of Mac Williams identity for Hamming code (15,11), we find w_3=35. Adding an overall odd parity, the weight remains the same.
In fact there still remain another group of w_{13}=35 which can be added.

These code vectors can be simply selected as the complements of the "w_3" group for simplification of the encoding and decoding processes. We obtain finally the composite code with the following weight distribution:

w_0=1, w_3=35; w_{13}=35, w_{16}=1
with error-correction capability k = 1 error.

w_8 = 30 with error-correction capability t=3 errors.

MAXIMUM RATE-MINIMUM ERROR COMPOSITE CODES
FOR
DIRECT-TRANSMISSION BINARY CHANNELS

This code can encode an alphabet source consisting of 102 symbols if the sum of probability of occurrence of the first 30 symbols is too large compared with the rest.

We may also construct composite codes consisting of constant weight groups extracted from linear and nonlinear codes e.g. the optimum Peterson code (14,4) with t=3 have the weight distribution

$$w_0 = 1 \;,\; w_7 = 8 \;,\; w_8 = 7$$

We can add a nonlinear group of w_2 consisting of 7 symmetrical code vectors having $d_{min} = 4$ inside the group capable to correct one error, to obtain the composite code with the following distribution:

$$w_2 = 7 \quad ; \quad k = 1$$

$$w_7 = 8 \;,\; w_8 = 7 \;;\; t = 3$$

The new composite code has 22 code vectors.

4. OPTIMUM COMPOSITE CODES FITTING ENGLISH LANGUAGE:

For obtaining the English alphabet distribution I refer to the recent work (c). The result of this work is tabulated in (Table 1) which indicates each symbol (letter) and its probability of occurrence. We must consider that this distribution will differ for different languages and also for different types of information transmitted by the same language. For example, in the commercial information channels, the numbers will have the highest frequency of occurrence.

If this alphabet is to be binary encoded, six information bits are required.

Symbol	Probability	Symbol	Probability
Space	0.1374000	*	0.0013611
E	0.0936188	.	0.0012862
I	0.0763861	1	0.0012737
T	0.0732392	0	0.0010114
O	0.0700299	3	0.0008741
A	0.0650324	,	0.0008241
N	0.0649850	J	0.0007117
R	0.0581418	:	0.0004620
S	0.0544180	(	0.0004620
C	0.0401848	)	0.0004620
L	0.0383991	6	0.0004495
D	0.0274115	4	0.0004370
F	0.0273476	5	0.0004370
H	0.0250624	7	0.0003996
M	0.0250499	9	0.0002997
U	0.0231143	8	0.0002622
P	0.0227522	[	0.0001998
G	0.0155469	]	0.0001998
Y	0.0120004	+	0.0001498
B	0.0084540	?	0.0000874
V	0.0081543	/	0.0000749
-	0.0063186	%	0.0000374
W	0.0053696	;	0.0000249
X	0.0029470	=	0.0000124
K	0.0019855	£	0.0000124
Q	0.0018606		
Z	0.0016108		
2	0.0013861		

(Table 1) Probability distribution of English alphabet.

Let us denote the composite code whose length is "n" bits and which encodes q source symbols by C(n/q) to differentiate between those codes and the error correcting codes (n,k) of total "n" and "k" information bits.

Here we introduce as example a subclass of short composite codes which fit the English alphabet.

(a) The high-rate composite code C(8/70)

The basic code is the code (8,4) with error-correction capability of one error represented by the parity-check matrix

$$C = \begin{bmatrix} 1 & 1 & 0 & 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 & 0 & 1 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

The weight distribution of this code is

$$w_2 = 28 \;,\; w_6 = 28 \quad ; \; k = 0$$
$$w_4 = 14 \quad ; \; t = 1$$

This code can be extended to be C(8/88) which have the distribution

$$w_0 = 1,\; w_1 = 8,\; w_2 = 28,\; w_6 = 28,\; w_7 = 8,\; w_8 = 1 \quad ; \; k = 0$$
$$w_4 = 14 \quad ; \; t = 1$$

This code is suitable for channels where errors more than one have zero probability of occurrence.

For code groups of weights two and six it is valid

$$p_{c_1} = (1-p)^n = (1-p)^8 \quad \ldots\ldots(22)$$

Let n_d stands for the number of ones (zeros) in any of these two groups then the probability of error-detection

$$p_{d_1} = \sum_{i=1}^{n_d} \binom{n_d}{i} p^i (1-p)^{n-i}$$

$$= \sum_{i=1}^{2} \binom{2}{i} p^i (1-p)^{8-i} \quad \ldots\ldots\ldots(23)$$

and $p_{e_1} = 1 - (p_d + p_c) \quad \ldots\ldots(24)$

For the group of weight w_4

$$p_{c_2} = \sum_{i=0}^{1} \binom{n}{i} p^i (1-p^{n-i}$$

$$= \sum_{i=0}^{1} \binom{8}{i} p^i (1-p)^{8-i} \quad \ldots\ldots\ldots(25)$$

An error is detected by the w_4 hamming group if even errors occurs in ones and zeros such that an erronous code word of the same weight is received. In general if n_1 ones are present and a detection capability d then

$$p_{d_2} \geqslant \sum_{i=1}^{d/2} \binom{n_1}{i} \binom{n-n_1}{i} p^{2i}(1-p)^{n-2i}$$

$$= \binom{4}{1}\binom{4}{1} p^2(1-p)^6 \quad \ldots\ldots\ldots\ldots (26)$$

For simplicity of calculations we neglected the higher order even errors. Let us encode the first 14 symbols of the highest frequency of occurrence by the Hamming group and the rest by groups of w_2 and w_7.

Then the average probability of correct decoding will be:

$$p_{cav} = p_{c_2} \sum_{j_1=1}^{14} p_{s_{j_1}} + p_{c_1} \sum_{j_2=15}^{53} p_{s_{j_2}}$$

$$= p_{c_2} \sum_{j_1=1}^{14} p_{s_{j_1}} + p_{c_1}\left(1 - \sum_{j_1=1}^{14} p_{s_{j_1}}\right) \quad \ldots\ldots (27)$$

where

p_{s_j} = the probability of occurrence of the jth symbol according to its order in (Table 1).

In Fig. (1) is plotted $p_{c_{av}}$ for this code and shortened Hamming (10,6) code. The properties are very close in spite of the higher rate obtained by the proposed composite code.

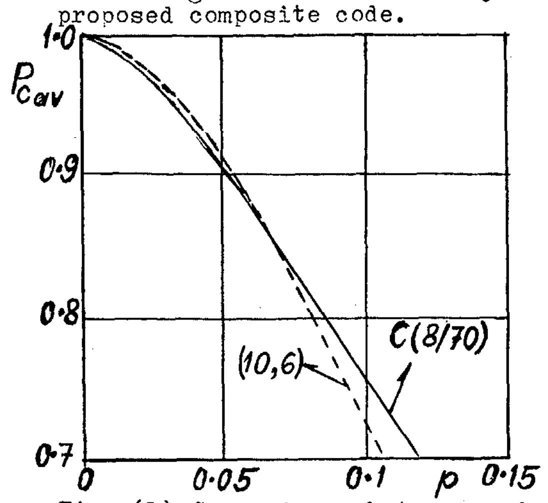


Fig. (1) Comparison between codes C(8/70) and (10,6)

(b) Composite codes of length 11:
Here we have two choices for construction of this code each having its own advantages:

- To use the optimum code (11,4) with correction capability of two errors which is represented by the parity-check matrix

$$C = \begin{bmatrix} 1&0&1&0&1&0&0&0&0&0&0 \\ 0&1&0&1&0&1&0&0&0&0&0 \\ 1&0&0&1&0&0&1&0&0&0&0 \\ 0&1&1&0&0&0&0&1&0&0&0 \\ 1&0&1&1&0&0&0&0&1&0&0 \\ 0&1&1&1&0&0&0&0&0&1&0 \\ 1&1&1&1&0&0&0&0&0&0&1 \end{bmatrix}$$

The following groups are selected from this code

$w_5=6$, $w_6=6$, $w_7=2$, $w_8=1$

The zero vector is excluded.

A $\binom{11}{2}$ or 55 code vectors of w_2 with no error correction capabilities are added.

The resultant is the composite code C(11/70). This code can be extended by including the zero and w_2 vectors to be C(11/81).

For this code it is valid

$$p_{c_{av}} = \sum_{i_1=0}^{2} \binom{11}{i_1} p^{i_1}(1-p)^{11-i_1} \sum_{J_1=1}^{15} p_{s_{J_1}} + \sum_{J_2=16}^{53} p_{s_{J_2}} (1-p)^{11} \quad ..(28)$$

$$p_{d_{av}} = \sum_{i_2=1}^{2} \binom{2}{i_2} p^{i_2}(1-p)^{11-i_2} \sum_{J_2=16}^{53} p_{s_{J_2}} \quad ..(29)$$

$$p_{e_{av}} = 1 - (p_{c_{av}} + p_{d_{av}}) \quad \ldots\ldots\ldots\ldots\ldots\ldots (30)$$

- To encode the first 4 symbols having the highest frequency with the non-linear code of weights 8 and 9 and minimum distance 5. This code is capable to correct two or less errors. This code group is represented by

$$\begin{bmatrix} V_1 \\ V_2 \\ V_3 \\ V_4 \end{bmatrix} = \begin{bmatrix} 0&1&1&1&0&1&1&1&0&1&1 \\ 1&0&1&1&1&0&1&1&1&0&1 \\ 1&1&0&1&1&1&0&1&1&1&0 \\ 1&1&1&0&1&1&1&0&1&1&1 \end{bmatrix}$$

Then from the shortend Hamming code (11,7) with correction capabilities of one error 14 code vector of weight 3 and 25 code vector of weight 4 are selected. The selected code is represented by the parity check matrix.

$$C = \begin{bmatrix} 1&1&1&0&1&0&1&1&0&0&0 \\ 1&1&1&0&0&1&1&0&1&0&0 \\ 0&1&0&1&1&1&1&0&0&1&0 \\ 0&0&1&1&1&1&1&0&0&0&1 \end{bmatrix}$$

Then 11 code vector of no error correction capabilities of w_1 are selected. Those code vectors are used to encode the symbols with the lowest frequency. This selection will lead to the construction of composite code C(11/54) which seems to fit more closely the English alphabet than any code of the same length known uptil now.
For this choice

$$p_{c_{av}} = \sum_{i_1=0}^{2} \binom{11}{i_1} p^{i_1}(1-p)^{11-i_1} \sum_{J_1=1}^{4} p_{s_{J_1}} + \sum_{i_2=0}^{1} \binom{11}{i_2} p^{i_2}(1-p)^{11-i_2} \sum_{J_2=5}^{43} p_{s_{J_2}} + (1-p)^{11} \sum_{J_3=44}^{53} p_{s_{J_3}} \quad (31)$$

MAXIMUM RATE-MINIMUM ERROR COMPOSITE CODES
FOR
DIRECT-TRANSMISSION BINARY CHANNELS

$$p_{d_{av}} = p(1-p)^{10} \sum_{j_3=49}^{53} p_{s_{j_3}} \doteq 0 \qquad \ldots\ldots(32)$$

$$p_{e_{av}} = 1- (p_{c_{av}}+p_{d_{av}}) \qquad \ldots\ldots\ldots\ldots\ldots\ldots(33)$$

In Fig.(2) there are indicated $p_{c_{av}}$ for the codes C(11/70), C(11/54) and the optimum code (11,6).

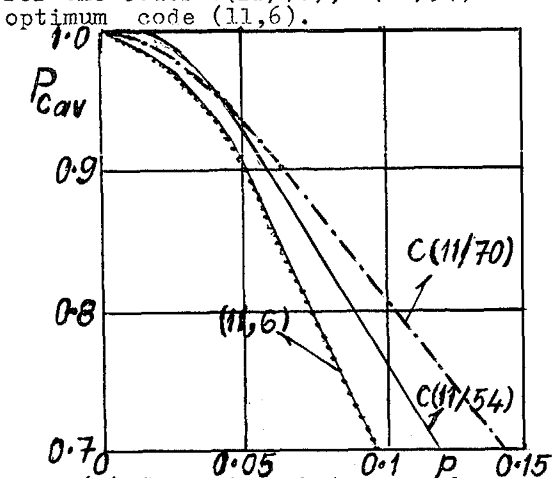


Fig. (2) Comparison between codes C(11/54), C(11/70) and (11,6)

(c) Composite codes of length 12:
A proposed choice is to encode the seven symbols of highest frequency with the non-linear code of error-correction capability of two errors of weights 8 and 9 represented by

$$\begin{bmatrix} V_1\\ V_2\\ V_3\\ V_4\\ V_5\\ V_6\\ V_7 \end{bmatrix} = \begin{bmatrix} 0&1&1&1&0&1&1&1&0&1&1&1\\ 1&0&1&1&1&0&1&1&1&0&1&1\\ 1&1&0&1&1&1&0&1&1&1&0&1\\ 1&1&1&0&1&1&1&0&1&1&1&0\\ 0&1&1&0&1&1&0&1&1&0&1&1\\ 1&0&1&1&0&1&1&0&1&1&0&1\\ 1&1&0&1&1&0&1&1&0&1&1&0 \end{bmatrix}$$

From the shortened Hamming code having the parity check matrix

$$C = \begin{bmatrix} 1&1&1&0&1&0&1&1&1&0&0&0\\ 1&1&1&0&0&1&1&0&0&1&0&0\\ 0&1&0&1&1&1&1&0&0&0&1&0\\ 0&0&1&1&1&1&1&1&0&0&0&1 \end{bmatrix}$$

is selected the following code groups with the corresponding weights and number of code vectors

$w_o=1$; $w_3=17$; $w_4=37$.

This leads to the composite code C(12/62)

Other choice is to encode the first 15 symbols of highest freqneucy of occurrence by the optimum code (12,4) represented by the parity-check matrix

$$C = \begin{bmatrix} 1&1&0&0&1&0&0&0&0&0&0&0\\ 1&0&1&0&0&1&0&0&0&0&0&0\\ 0&1&1&0&0&0&1&0&0&0&0&0\\ 1&1&1&0&0&0&0&1&0&0&0&0\\ 1&0&0&1&0&0&0&0&1&0&0&0\\ 0&1&0&1&0&0&0&0&0&1&0&0\\ 1&1&0&1&0&0&0&0&0&0&1&0\\ 0&0&1&1&0&0&0&0&0&0&0&1 \end{bmatrix}$$

Here is obtained

$w_5= 3$; $w_6=5$; $w_7=6$; $w_8=1$

The last symbols are encoded by code words of weights 0,1,2,11,12 with no error correction or detection capabilities.

(d) Composite codes of length 13.

A proposed choice is to encode the first 30 code words by the optimum code (13,5) represented by the parity-check matrix

$$C= \begin{bmatrix} 1&1&1&0&0&1&0&0&0&0&0&0&0\\ 1&1&0&1&0&0&1&0&0&0&0&0&0\\ 1&0&1&1&0&0&0&1&0&0&0&0&0\\ 1&1&0&0&1&0&0&0&1&0&0&0&0\\ 1&0&1&0&1&0&0&0&0&1&0&0&0\\ 0&1&0&1&1&0&0&0&0&0&1&0&0\\ 0&0&1&1&1&0&0&0&0&0&0&1&0\\ 1&1&1&1&1&0&0&0&0&0&0&0&1 \end{bmatrix}$$

The weight distribution of this code is

$w_o=1$; $w_5=4$; $w_6=11$; $w_7=12$, $w_8=3$;$w_{13}=1$

According to the weight distribution given by (Table 1) we encode the first 31 symbols by the groups w_5 up to w_{13}.

Since the remaining symbols have the probability of occurrence 0.0078911 which is very small, they are encoded by the group of weight 2 which has no error correction capabilities.

This leads to the composite code C(13/109).

Fig. (3) shows the probability p_c for the codes C(12/62), C(13/109) and the optimum code (14,6). It is seen that C(13/109) is better than (14,6) when encoding the English alphabet by both, in addition it has higher transmission rate.

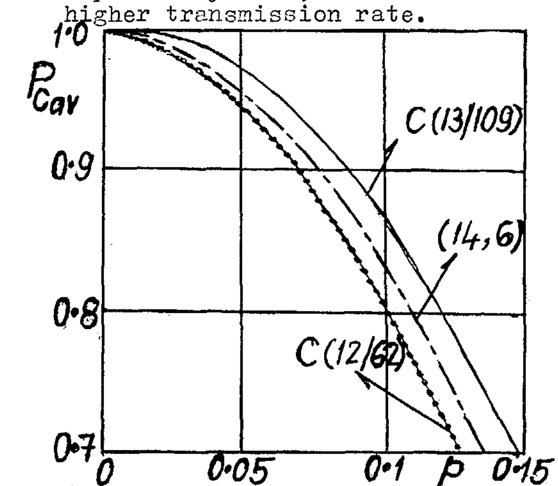


Fig.(3) A comparison between codes C(12/62),C(13/109) and (14,6)

4. REALIZATION OF COMPOSITE CODES:

(a) The encoder

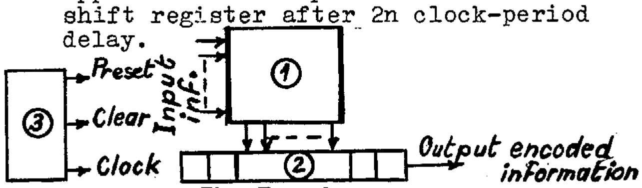
According to Fig. (4) the encoder can be simply consists of a P.R.O.M, a parallel-to-series converter shift register and the associated synchronizer for generation of clear, preset , and clock pulses. The input information stored on a perforated or magnetic tape are applied to a P.R.O.M. which acts

as a code generator.
In each transmission cycle the shift register is clear, then output of P.R.O.M. is written into it, finally clock pulses act to transfere the shift register contents to the modulator input in the form of a series bit-train.

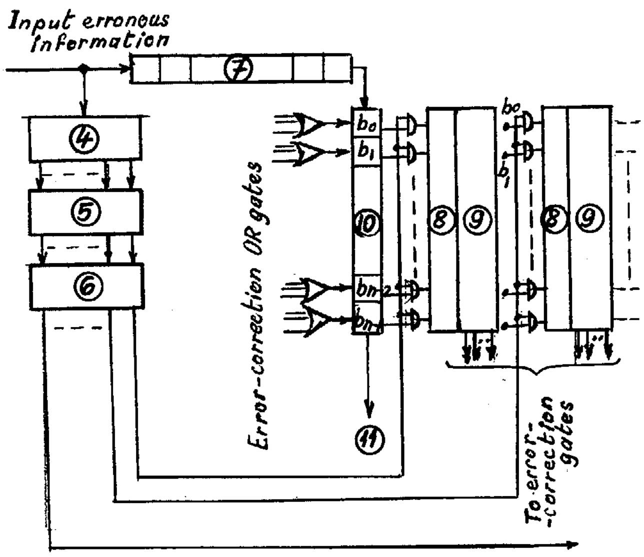
(b) The decoder:
The input information from the channel is applied to a counter mod. n which is cleared before the beginning of each code word and simultaneously to a delay shift register of length n. The result of count is decoded by the weight decoders, then stored in a store consisting of "r" S-R flipflops.
The store acts to keep the result for a new n clock periods until the information is completely settled in the correction n-bit register. In addition, the output of store enables the decoder to which the result of count belongs through a set of "n" two-input and gates.
Each of the r decoders consists of two parts:

- A set of parity checkers to generate the parity-check equations from information and parity checks stored in the correction-shift register.
- An error-pattern location P.R.O.M. which determines the location of error pattern according to syndrome provided by parity checkers.

All outputs of error-location P.R.O.Ms are logically summed, and applied for correction of the erronous bits existing in the correction-shift register through its asynchronous inputs.
This can be done after a very small delay from the last 2nth clock pulse.
The corrected information begins to appear at the output of the correction-shift register after 2n clock-period delay.



The Encoder



The Decoder

1. P.R.O.M. Encoder.
2. Parallel-to-series shift register.
3. Synchronizer.
4. Counter Mod. n
5. Weight decoders.
6. Store.
7. Delay shift register (lengh n).
8. Parity checkers.
9. Error-pattern location P.R.O.M.
10. Correction register.
11. Output corrected information.

Fig. (4): Encoder and Decoderof a composite code.

5. GENERAL CONCLUSIONS:

(a) A proper choice of a composite code will lead to an extensively high increase in the transmission rate and simultaneously an increase in the average probability of correct decoding.
The decrease of code length and encoding the symbols of higher frequency of occurrence by higher correction capability code vectors will both cooperate to increase the average probability of correct decoding and hence decrease the average probability of error. Theutilization of composite codes in practical transmission systems will lead to error-free transmission at the highest possible rate.

(b) The best codes suitable for construction of optimum composite codes are those in which most code vectors are found in the least number of adjacent groups. The maximum-distance-seperable codes presented by Kasami, Lin, Peterson, Turyn and others seem to be helpful in construction of composite non binary codes.
Unfortunately they are not applicable in the binary ease.

(c) This work hands over a guide to following researchs concerned with

- Finding the best composite codes for different code lengths, which have the highest number of code vectors with the highest capabilities.
- Finding a definite code structure with minimum length, having code vectors with maximum possible and variable correction bapabilities.

6. REFERENCES:

a) W. WESELEY PETERSON, E.J. WELDON, Jr.
Error - Correcting Codes.
The Massachusettes Institute of Technology, 1972.

b) MAC WILLIAMS F.J.
A Theorem on the Distribution of Weights in a Systematic Code.
Bell System Technical Journal,42,1963

c) MICHAEL F. LYNCH
Shannon's Mathematical Theory of Communications and its Implications in Information Science.
Journal of Society for Information Science-January 77.