Mettre un code une source d'informa-
tion de l'alphabet en utilisant un code
spécifique pour correction d'erreurs a
protéger cette information contre le bruit
(noise) n'est pas la solution optimum. Le
fréquence d'occurrence de symboles diffé-
rentes de source diffère enormement.

Un code composite d'une longueur constante
"n" est ce code qui contient dans sa com-
position des groupes différents de poids
constant que possèdent différentes capa-
bilités de correction de fautes avec une
certaine limitation d'une distance minimum
l'un de l'autre. Un code composite optimum
qui convient une source d'information
avec une certaine distribution de
symboles, est ce code qui met au maximum la
probabilité d'un juste decoder a un degré
certain de transmission.

Encoder composite cause en tous les
cas une plus élevée probabilité de décodes
juste et un plus élevé degré de transmis-
sion. Cela est obtenue quand les symboles
de plus grande fréquence d'occurrence sont
encodés avec des groupes de plus grande
capacité de correction de faute et vice
versa. Pour une certaine longueur, il
existe plusieurs possibilités des codes
composites, un deux est être convenable
pour une source spécifique avec une cer-
taine distribution.

Ce travail introduit, un guide général
pour la construction de ces codes, certains
courts codes composés convenable pour l'
alphabet Anglais et la configuration géné-
rale de l'encode et decoder. En plus,
comparaisons faites nous montrer que les
codes composés, ont des meilleures propé-
rités que les plus réputées codes de cor-
rection d'erreurs de la même longueur et
meme de plus les codes de grande longueur
s'ils les codes composés sont proprement
construit. Les résultats obtenus signifie la
modification du système binaire de trans-
misssion qui utilise le code normales pour
correction de fautes a minimiser la perte
dans le temps de transmission et pour ame-
liorer les propriété communication.

Encoding an alphabet information source
by a specific error-correcting code for
protecting this information against 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-cor-
rection capabilities under certain rest-
lritions 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 occur-
cence are encoded by groups of higher error-
correction capability and vice 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 dis-
tribution.

This work introduces a general guide
for construction of these codes, some short
composite codes fitting the English alpha-
bet, and the general configuration of the
encoder and decoder. In addition, compari-
sions 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 modifi-
cation of binary transmission system using
the normal error-correcting codes for
minimization the time waste in transmission
and improving the communication properties.
1. **Basic Idea**

Given a discrete zero-memory information source having "*n*" symbols
\[ S = \begin{pmatrix} S_1 & S_2 & \ldots & S_q \end{pmatrix} \quad \ldots (1) \]

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

\[ p_{av} = \frac{p(S_1) + p(S_2) + \ldots + p(S_q)}{q} = \text{maximum} \ldots (2) \]

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

\[ n > \log_2 q \quad \ldots (3) \]

we proceed as follows:

(a) By computer or by the aid of MacWilliams 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 *i*th group of Hamming weight "*w*" is "*t*" then the adjacent group of correction and/or detection capabilities must have at least the weight

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

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

- If a group of Hamming weight "*w*" and error correction and/or detection capability "*t*", then the following (or preceding) group of no error correction or detection capability must have the weight

\[ w_{i+1} = w_i + t_1 + t_2 + 1 \quad \ldots (5) \]

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

\[ \sum_{i=1}^{n} w_i = \sum_{j=1}^{n} w_j \quad \ldots (6) \]

where

- \( N_1 \) = Number of code groups with no error correction capability
- \( w_i \) = The Hamming weight of the *i*th group,
- \( N_2 \) = Number of code groups having error correction and/or detection capabilities,
- \( N_j \) = Number of code vectors contained in the *j*th 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:

- \( p_{ei} \) (the probability of correct decoding), \( p_{i} \) (the probability of detection of error) and \( p_{ei} \) (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 \( w_j \), \( t_2 \), respectively. The second group encodes symbols of higher importance. Then if

\[ w_j - w_i = t_1 + t_2 + r \quad \ldots (7) \]

If no more than \( t_1 + t_2 \) errors are expected to occur, and \( r = r_1 + r_2 \), then:

\[ p_{ei} = \sum_{i=0}^{t_1} \binom{n}{i_1} (1-p)^{n-i_1} \quad \ldots (8) \]

\[ p_{d_i} = \sum_{i=2}^{t_1+r_2} \binom{n-w_i}{i_2} (1-p)^{n-
\]

\[ p_{d_j} = 1 - p_{ei} + p_{di} \quad \ldots (10) \]

\[ p_{cj} = \sum_{j=0}^{t_2} \binom{n}{j} (1-p)^{n-j} \quad \ldots (11) \]
MAXIMUM RATE-MINIMUM ERROR COMPOSITE CODES
FOR DIRECT-TRANSMISSION BINARY CHANNELS

\[ P_{d_j} = \sum_{j_2=1}^{t_2-r_2} \binom{w_j}{j_2} P_j (1-p)^{n-j_2} \quad \ldots \ldots \ldots \ldots \ldots (12) \]

\[ p_{e_j} = 1 - (p_{d_j} + P_{d_j}) \quad \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ ld
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:
\[
\begin{align*}
w_0 &= 1, & w_7 &= 8, & w_8 &= 7
\end{align*}
\]

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:
\[
\begin{align*}
w_2 &= 7; & k &= 1 \\
w_7 &= 8, & w_8 &= 7; & t &= 3
\end{align*}
\]
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.

<table>
<thead>
<tr>
<th>Letter</th>
<th>Probability</th>
<th>Symbol</th>
<th>Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>Space</td>
<td>0.1374000</td>
<td>*</td>
<td>0.0013611</td>
</tr>
<tr>
<td>E</td>
<td>0.0936188</td>
<td>*</td>
<td>0.0012862</td>
</tr>
<tr>
<td>I</td>
<td>0.0763281</td>
<td>*</td>
<td>0.0032737</td>
</tr>
<tr>
<td>T</td>
<td>0.0412992</td>
<td>*</td>
<td>0.0010114</td>
</tr>
<tr>
<td>G</td>
<td>0.0700299</td>
<td>*</td>
<td>0.0008741</td>
</tr>
<tr>
<td>A</td>
<td>0.0650324</td>
<td>*</td>
<td>0.0008241</td>
</tr>
<tr>
<td>N</td>
<td>0.0649890</td>
<td>*</td>
<td>0.0007117</td>
</tr>
<tr>
<td>R</td>
<td>0.0581428</td>
<td>*</td>
<td>0.0006420</td>
</tr>
<tr>
<td>S</td>
<td>0.0544180</td>
<td>*</td>
<td>0.0006420</td>
</tr>
<tr>
<td>L</td>
<td>0.0382846</td>
<td>*</td>
<td>0.0003495</td>
</tr>
<tr>
<td>D</td>
<td>0.0327115</td>
<td>*</td>
<td>0.0002470</td>
</tr>
<tr>
<td>F</td>
<td>0.0273476</td>
<td>*</td>
<td>0.0002470</td>
</tr>
<tr>
<td>H</td>
<td>0.0296243</td>
<td>*</td>
<td>0.0002996</td>
</tr>
<tr>
<td>U</td>
<td>0.0295499</td>
<td>*</td>
<td>0.0002997</td>
</tr>
<tr>
<td>P</td>
<td>0.0211143</td>
<td>*</td>
<td>0.0002222</td>
</tr>
<tr>
<td>Q</td>
<td>0.0237282</td>
<td>*</td>
<td>0.0002198</td>
</tr>
<tr>
<td>Y</td>
<td>0.0155469</td>
<td>*</td>
<td>0.0001998</td>
</tr>
<tr>
<td>V</td>
<td>0.0112004</td>
<td>*</td>
<td>0.0001498</td>
</tr>
<tr>
<td>-</td>
<td>0.0005460</td>
<td>*</td>
<td>0.0000874</td>
</tr>
<tr>
<td>W</td>
<td>0.0061385</td>
<td>*</td>
<td>0.0000714</td>
</tr>
<tr>
<td>X</td>
<td>0.1556296</td>
<td>*</td>
<td>0.0002499</td>
</tr>
<tr>
<td>K</td>
<td>0.0009270</td>
<td>*</td>
<td>0.0000124</td>
</tr>
<tr>
<td>C</td>
<td>0.0018595</td>
<td>*</td>
<td>0.0000124</td>
</tr>
<tr>
<td>Z</td>
<td>0.0016186</td>
<td>*</td>
<td>0.0000124</td>
</tr>
</tbody>
</table>

(\text{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:
\[
C = \begin{bmatrix}
1 & 1 & 0 & 1 & 1 & 0 & 0 & 0 \\
1 & 1 & 0 & 1 & 0 & 0 & 0 & 0 \\
0 & 1 & 1 & 0 & 0 & 0 & 0 & 1 \\
0 & 1 & 1 & 0 & 0 & 0 & 0 & 1 
\end{bmatrix}
\]

The weight distribution of this code is
\[
w_2 = 28, \quad w_6 = 28; \quad k = 0 \quad \text{and} \quad t = 1
\]

This code can be extended to a \( C(8/80) \) which have the distribution
\[
w_0 = 1, \quad w_4 = 8, \quad w_6 = 28, \quad w_7 = 28, \quad \text{and} \quad w_8 = 1
\]
\[
w_4 = 14 \quad \text{and} \quad t = 1
\]

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
\[
p_{e1} = (1 - p)^n = (1 - p)^6 \quad \text{......(22)}
\]

Let \( n \) stands for the number of ones (zeros) in any of these two groups then the probability of error-detection
\[
P_{d1} = \sum_{i=1}^{n_d} \binom{n_d}{i} p^i (1-p)^{n-i-1}
\]
\[
= \sum_{i=1}^{2} \binom{2}{i} p^i (1-p)^{8-i} \quad \text{......(23)}
\]

and \( p_{e2} = 1 - (p_{d1} + p_{e1}) \quad \text{......(24)}
\]

For the group of weight \( w_4 \)
\[
P_{e2} = \sum_{i=0}^{1} \binom{1}{i} p^i (1-p)^{6-i}
\]
\[
= \sum_{i=0}^{1} \binom{1}{i} p^i (1-p)^{8-i} \quad \text{......(25)}
\]

An error is detected by the \( w_4 \) hamming group if even errors occurs in ones and zeros such that an erroneous code word of the same weight is received. In general if \( n_1 \) ones are present and a detection capability \( d \) then
MAXIMUM RATE-MINIMUM ERROR COMPOSITE CODES
FOR
DIRECT-TRANSMISSION BINARY CHANNELS

\[ P_{d,2} = \frac{3}{2} \left( \frac{n}{i} \right) \left( \frac{n-i}{1} \right) p^{2i(1-p)^n-2i} \]

\[ = \left( \frac{4}{3} \right) p^{2(1-p)^6} \] ........................ (26)

For simplicity of calculations we neglected the higher order 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:

\[ P_{cav} = P_{c2} \sum_{j=1}^{14} P_{s_j} + P_{c1} \sum_{j=1}^{15} P_{s_j} \]

\[ = P_{c2} \sum_{j=1}^{14} P_{s_j} + P_{c1} (1 - \sum_{j=1}^{14} P_{s_j}) \]

where

\[ P_{s_j} = \text{the probability of occurrence of the } j^{th} \text{ symbol according to its order in (Table 1).} \]

In Fig. (1) is plotted \( P_{cav} \) 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.

\[ \text{Fig. (1) Comparison between codes (8/70) and (10,6)} \]

(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

\[ C = \begin{bmatrix}
1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\
0 & 1 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
0 & 1 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 1 & 0 \\
1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1
\end{bmatrix} \]

The following groups are selected from this code
\( w_2 = 6 \), \( w_6 = 6 \), \( w_7 = 2 \), \( w_9 = 1 \)

The zero vector is excluded.

A \( (11,7) \) 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

\[ P_{cav} = \sum_{i=1}^{2} \left( \begin{bmatrix} 1 \\ i_1 \\ i_2 \\ i_3 \end{bmatrix} P^{i_1(1-P)^{i_2}} \sum_{j=1}^{15} P_{S_j} + \sum_{j=1}^{15} P_{S_j} \right) \] ........................ (28)

\[ P_{dav} = \sum_{i=1}^{2} \left( \begin{bmatrix} 1 \\ i_1 \\ i_2 \\ i_3 \end{bmatrix} P^{i_1(1-P)^{i_2}} \right) \] ........................ (29)

\[ P_{eav} = 1 - (P_{cav} + P_{dav}) \] ............................ (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

\[ V_1 \]

\[ V_2 \]

\[ V_3 \]

\[ V_4 \]

Then from the shortened 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.

\[ C = \begin{bmatrix}
1 & 1 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 1 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 \\
0 & 1 & 0 & 1 & 1 & 1 & 0 & 0 & 1 & 0 & 0 \\
0 & 0 & 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0
\end{bmatrix} \]

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 until now.

For this choice

\[ P_{cav} = \sum_{i=1}^{2} \left( \begin{bmatrix} 1 \\ i_1 \\ i_2 \\ i_3 \end{bmatrix} P^{i_1(1-P)^{i_2}} \sum_{j=1}^{15} P_{S_j} \right) + \]

\[ \sum_{i=1}^{2} \left( \begin{bmatrix} 1 \\ i_1 \\ i_2 \\ i_3 \end{bmatrix} P^{i_1(1-P)^{i_2}} \right) \] ........................ (31)
\[ p_{d_{av}} = p(1-p) \left( \frac{10}{11} \right)^{10}, \quad p_{j_{1}} \leq 0 \] (32)

\[ p_{e_{av}} = 1 - (p_{c_{av}} + p_{d_{av}}) \] (33)

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

\[ \text{Fig. (2) Comparison between codes C(11/70), C(11/54) and (11,6)} \]

(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:

\[
\begin{bmatrix}
V_{1} & 0 & 1 & 1 & 1 & 0 & 1 & 1 & 1 & 0 & 1 & 1 \\
V_{2} & 1 & 0 & 1 & 1 & 1 & 0 & 1 & 1 & 1 & 0 & 1 \\
V_{3} & 1 & 1 & 0 & 1 & 1 & 1 & 0 & 1 & 1 & 1 & 0 \\
V_{4} & 0 & 1 & 1 & 0 & 1 & 1 & 1 & 0 & 1 & 1 & 0 \\
V_{5} & 0 & 1 & 1 & 0 & 1 & 1 & 1 & 0 & 1 & 1 & 0 \\
V_{6} & 1 & 0 & 1 & 1 & 0 & 1 & 1 & 1 & 0 & 1 & 1 \\
V_{7} & 1 & 1 & 0 & 1 & 1 & 0 & 1 & 1 & 1 & 0 & 1 \\
\end{bmatrix}
\]

From the shortened Hamming code having the parity check matrix

\[
\begin{bmatrix}
1 & 1 & 1 & 0 & 1 & 0 & 1 & 1 & 1 & 0 & 0 & 0 \\
1 & 1 & 1 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 1 \\
0 & 1 & 0 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 1 & 0 \\
0 & 0 & 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\
\end{bmatrix}
\]

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

\[
w_0 = 1; \quad w_2 = 37. \]

This leads to the composite code C(12/62). Other choice is to encode the first 15 symbols of highest frequency of occurrence by the optimum code (12,4) represented by the parity-check matrix:

\[
\begin{bmatrix}
1 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\
1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\
0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\
\end{bmatrix}
\]

Here is obtained \( w_0 = 3; \quad w_4 = 5; \quad w_5 = 6; \quad w_7 = 1 \).

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:

\[
\begin{bmatrix}
1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 1 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 0 & 1 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
1 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 1 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\
1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\
\end{bmatrix}
\]

The weight distribution of this code is \( w_0 = 1; \quad w_5 = 4; \quad w_6 = 11; \quad w_7 = 12; \quad w_8 = 3; \quad w_9 = 1 \).

According to the weight distribution given by (Table 1) we encode the first 31 symbols by the groups \( w_0 \) up to \( w_9 \).

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 \( p_{c_{av}} \) for the composite code 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.

\[ \text{Fig. (3) A comparison between codes C(12/62), C(13/109) and (14,6)} \]

4. REALIZATION OF COMPOSITE CODES:
(a) The encoder
According to Fig. (4) the encoder can be simply consists of a P.R.O.W., 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.W. which acts
as a code generator.
In each transmission cycle the shift register is cleared, then output of
P.R.O.M. is written into it, finally clock pulses act to transfer 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. m 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 informa-
tion 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 "m" 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 infor-
mation 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.M.s are logically summed, and applied for
correction of the erroneous bits existing in the correction-shift register
through its asynchronous inputs. This can be done after a very small
delay from the last clock pulse. The corrected information begins to
appear at the output of the correction-shift register after 2n clock-period
delay.

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

Fig. (4): Encoder and Decoder of 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 capabi-
licity code vectors will both cooperate
to increase the average probability of
correct decoding and hence decrease the
average probability of error. The utili-
ization of composite codes in practical
transmission systems will lead to error-
free transmission at the highest possi-
able 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-separable codes presen-
ted by Kasami, Zin, Peterson, Turyn and
others seem to be helpful in construction of
composite non binary codes.
Unfortunately they are not applicable in
the binary case.

(c) This work hands over a guide to follow-
ing researches 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
capabilities.

6. REFERENCES:
a) W. WESLEY PETERSON, E.J. WELDON, Jr.
Error - Correcting Codes,
The Massachusetts Institute of
b) MAC WILLIAMS F.J.
A Theorem on the Distribution of Weights
in a Systematic Code,
Bell System Technical Journal, 42, 1963
c) MICHAEL P. LYNCH
Shannon's Mathematical Theory of
Communications and its Implications in
Information Science,
Journal of Society for Information
Science—January 77.