ON ERROR CLUSTERING IN DIGITAL COMMUNICATION SYSTEMS

Francis K. Varshney
Electrical and Computer Engineering Dept.
Syracuse University
Syracuse, NY 13210

SUMMARY

This paper examines the error clustering properties of digital communication systems. In the past, the error behavior of digital communication systems has been investigated in terms of their input-output behavior. Actual communication system components and the physical interference processes present have not been taken into account. In this paper, an attempt is made to examine the relationship between the actual communication system components and the error sequence generated by this system. Two measures are defined which try to characterize this relationship quantitatively. Two classes of digital communication systems are considered. The first class considers the error behavior of various modulation formats in the presence of correlated Gaussian noise. The second class considers a communication system employing a receiver with memory over fading channels. The examples compute the measures related to the occurrence of double errors.

I. INTRODUCTION

Characterization of the stochastic behavior of errors in a digital communication system has been an area of intense research activity. These analytical error models find applications in the analysis and performance evaluation of various error control techniques. These studies also provide an insight into the error clustering properties of digital communication systems. The emphasis has been on input-output behavior of the digital communication system. The channel error sequence \( \{e_i\} \) is defined as a binary sequence where \( e_i \) is one if a transmission error occurs and zero otherwise, i.e.,

\[
e_i = \begin{cases} 
0 & \text{if } x_i = y_i \\
1 & \text{if } x_i \neq y_i 
\end{cases}
\]

where \( \{x_i\} \) is the transmitted information sequence and \( \{y_i\} \) is its estimate at the output. Three major groups of input-output models have been proposed. The simplest class of models is the renewal process, and developed the third group of input-output models which characterized the short-term error behavior more accurately.

The common denominator of the analytical models described above and other similar models is that they do not take into account the actual communication system components and the interference processes causing the errors. In this paper, an attempt is made to examine the relationship between the actual communication system blocks and the error sequence generated by communication systems with memory. Some preliminary results on the error clustering properties of these systems are presented. In order to study the error clustering properties of communication systems, some measures are to be developed which would characterize the relationship quantitatively. In the next section, two general measures are defined which characterize the statistical correlation of error events. In the third section, the general digital communication system under consideration is briefly described. Attention is restricted to two classes of digital communication systems and their error clustering properties are examined. In section four, we consider communication systems employing different modulation formats in the presence of correlated noise. In section five, a receiver with memory for fading channels is briefly discussed (a detailed description is available in [5]). Section six examines the error behavior of the receiver with memory as compared to a receiver without memory. Finally, the results are summarized in the last section.

II. MEASURES OF ERROR CLUSTERING

The study of the error clustering properties of a digital communication system involves the statistical characterization of various error patterns. One possible approach is to study the Markov chain associated with the bits or gaps of the error sequence generated by a digital communication system. But this approach does not seem to be very practical and, therefore, other more general measures are needed to characterize the statistical correlation of errors. These measures provide information about the occurrence of error events based on the past errors. The error sequence \( \{e_i\} \) is assumed to be stationary with

\[
P(e_i = 1) = p
\]

where \( p \) is the average error rate. Two measures are considered. The first measure is obtained from statistical considerations and the other derived from the information theoretic point of view. In statistics, the correlation of two random variables is determined by means of the correlation coefficient. A similar measure is defined here using the same concept to describe the dependence of error events \( \{e_i\} \).

Definition 1: The \((n+1)\)-th order generalized correlation coefficient \( \rho_{n+1}(e) \) of the sequence of random variables \( \{e_i\} \) is defined as
where $\mu$ is the average error rate, $E(\cdot)$ represents the expected value and $\sigma^2 = p(1-p)$ is the variance of $e_i$. It can be observed that $\mu = 0$ and the value of $\mu$, $i > 1$ can be computed from (3). The second measure is derived using the information theoretic approach. The dependence of error events is considered and a measure in terms of mutual information is described.

**Definition 2:** The $(n + 1)$-th order error clustering coefficient $B_{n+1}(m_1, \ldots, m_n)$ is defined as

$$B_{n+1}(m_1, \ldots, m_n) = \frac{P(e_k = 1)P(e_{k-1} = 1)\cdots P(e_{k-n+1} = 1)}{P(e_k = 1)P(e_{k-1} = 1)\cdots P(e_{k-n+1} = 1)}$$

$$\mu_{n+1} (m_1, m_2, \ldots, m_n) = k_l P(e_k = 0)P(e_{k-1} = 0)\cdots P(e_{k-n+1} = 0)$$

where $\mu$ is the average error rate, $E(\cdot)$ represents the expected value and $\sigma^2 = p(1-p)$ is the variance of $e_i$. It can be observed that $\mu = 0$ and the value of $\mu$, $i > 1$ can be computed from (3). The second measure is derived using the information theoretic approach. The dependence of error events is considered and a measure in terms of mutual information is described.

**Definition 2:** The $(n + 1)$-th order error clustering coefficient $B_{n+1}(m_1, \ldots, m_n)$ is defined as

$$B_{n+1}(m_1, \ldots, m_n) = \frac{P(e_k = 1)P(e_{k-1} = 1)\cdots P(e_{k-n+1} = 1)}{P(e_k = 1)P(e_{k-1} = 1)\cdots P(e_{k-n+1} = 1)}$$

$$\mu_{n+1} (m_1, m_2, \ldots, m_n) = k_l P(e_k = 0)P(e_{k-1} = 0)\cdots P(e_{k-n+1} = 0)$$

where $\mu$ is the average error rate, $E(\cdot)$ represents the expected value and $\sigma^2 = p(1-p)$ is the variance of $e_i$. It can be observed that $\mu = 0$ and the value of $\mu$, $i > 1$ can be computed from (3). The second measure is derived using the information theoretic approach. The dependence of error events is considered and a measure in terms of mutual information is described.

**Definition 2:** The $(n + 1)$-th order error clustering coefficient $B_{n+1}(m_1, \ldots, m_n)$ is defined as

$$B_{n+1}(m_1, \ldots, m_n) = \frac{P(e_k = 1)P(e_{k-1} = 1)\cdots P(e_{k-n+1} = 1)}{P(e_k = 1)P(e_{k-1} = 1)\cdots P(e_{k-n+1} = 1)}$$

$$\mu_{n+1} (m_1, m_2, \ldots, m_n) = k_l P(e_k = 0)P(e_{k-1} = 0)\cdots P(e_{k-n+1} = 0)$$

where $\mu$ is the average error rate, $E(\cdot)$ represents the expected value and $\sigma^2 = p(1-p)$ is the variance of $e_i$. It can be observed that $\mu = 0$ and the value of $\mu$, $i > 1$ can be computed from (3). The second measure is derived using the information theoretic approach. The dependence of error events is considered and a measure in terms of mutual information is described.

**Definition 2:** The $(n + 1)$-th order error clustering coefficient $B_{n+1}(m_1, \ldots, m_n)$ is defined as

$$B_{n+1}(m_1, \ldots, m_n) = \frac{P(e_k = 1)P(e_{k-1} = 1)\cdots P(e_{k-n+1} = 1)}{P(e_k = 1)P(e_{k-1} = 1)\cdots P(e_{k-n+1} = 1)}$$

$$\mu_{n+1} (m_1, m_2, \ldots, m_n) = k_l P(e_k = 0)P(e_{k-1} = 0)\cdots P(e_{k-n+1} = 0)$$

where $\mu$ is the average error rate, $E(\cdot)$ represents the expected value and $\sigma^2 = p(1-p)$ is the variance of $e_i$. It can be observed that $\mu = 0$ and the value of $\mu$, $i > 1$ can be computed from (3). The second measure is derived using the information theoretic approach. The dependence of error events is considered and a measure in terms of mutual information is described.

**Definition 2:** The $(n + 1)$-th order error clustering coefficient $B_{n+1}(m_1, \ldots, m_n)$ is defined as

$$B_{n+1}(m_1, \ldots, m_n) = \frac{P(e_k = 1)P(e_{k-1} = 1)\cdots P(e_{k-n+1} = 1)}{P(e_k = 1)P(e_{k-1} = 1)\cdots P(e_{k-n+1} = 1)}$$

$$\mu_{n+1} (m_1, m_2, \ldots, m_n) = k_l P(e_k = 0)P(e_{k-1} = 0)\cdots P(e_{k-n+1} = 0)$$

where $\mu$ is the average error rate, $E(\cdot)$ represents the expected value and $\sigma^2 = p(1-p)$ is the variance of $e_i$. It can be observed that $\mu = 0$ and the value of $\mu$, $i > 1$ can be computed from (3). The second measure is derived using the information theoretic approach. The dependence of error events is considered and a measure in terms of mutual information is described.

**Definition 2:** The $(n + 1)$-th order error clustering coefficient $B_{n+1}(m_1, \ldots, m_n)$ is defined as

$$B_{n+1}(m_1, \ldots, m_n) = \frac{P(e_k = 1)P(e_{k-1} = 1)\cdots P(e_{k-n+1} = 1)}{P(e_k = 1)P(e_{k-1} = 1)\cdots P(e_{k-n+1} = 1)}$$

$$\mu_{n+1} (m_1, m_2, \ldots, m_n) = k_l P(e_k = 0)P(e_{k-1} = 0)\cdots P(e_{k-n+1} = 0)$$

where $\mu$ is the average error rate, $E(\cdot)$ represents the expected value and $\sigma^2 = p(1-p)$ is the variance of $e_i$. It can be observed that $\mu = 0$ and the value of $\mu$, $i > 1$ can be computed from (3). The second measure is derived using the information theoretic approach. The dependence of error events is considered and a measure in terms of mutual information is described.

**Definition 2:** The $(n + 1)$-th order error clustering coefficient $B_{n+1}(m_1, \ldots, m_n)$ is defined as

$$B_{n+1}(m_1, \ldots, m_n) = \frac{P(e_k = 1)P(e_{k-1} = 1)\cdots P(e_{k-n+1} = 1)}{P(e_k = 1)P(e_{k-1} = 1)\cdots P(e_{k-n+1} = 1)}$$

$$\mu_{n+1} (m_1, m_2, \ldots, m_n) = k_l P(e_k = 0)P(e_{k-1} = 0)\cdots P(e_{k-n+1} = 0)$$

where $\mu$ is the average error rate, $E(\cdot)$ represents the expected value and $\sigma^2 = p(1-p)$ is the variance of $e_i$. It can be observed that $\mu = 0$ and the value of $\mu$, $i > 1$ can be computed from (3). The second measure is derived using the information theoretic approach. The dependence of error events is considered and a measure in terms of mutual information is described.
The computation of $a_{(1)}$ for different modulation formats is illustrated by an example.

**Example:**

In this example, attention is restricted to symmetrical modulation formats namely FSK and PSK systems. We also assume coherent reception. Generalization to incoherent reception is straightforward. The memoryless detector for coherent FSK decides that $i$ was transmitted if $X_k^i > X_k$. Thus, an error occurs if $X_k^i > X_k$. In this example, we compute $a_{(1)}$ to illustrate the statistical correlation of error events for binary coherent FSK and PSK. To compute $a_{(1)}$ and $b_{(1)}$, we need to compute the probability of two consecutive errors, i.e., $P(e_k = 1, e_{k-1} = 1)$. For coherent FSK, this probability is given by

$$P(e_k = 1, e_{k-1} = 1) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi(2N_0)}} \exp \left( -\frac{x^2 - 2p_2}{2N_0} \right) dx dy$$

where $p$ is the correlation coefficient for correlated Gaussian noise. In a similar manner, for binary PSK

$$P(e_k = 1, e_{k-1} = 1) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi(1-p^2)}} \exp \left( -\frac{x^2 - 2p_2}{2(1-p^2)} \right) dx dy$$

Table 1. Values of $a_{2(1)}$ for FSK and PSK systems.

<table>
<thead>
<tr>
<th>SNR</th>
<th>FSK</th>
<th>PSK</th>
<th>FSK</th>
<th>PSK</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>0.27975</td>
<td>0.23469</td>
<td>0.27975</td>
<td>0.23469</td>
</tr>
<tr>
<td>4</td>
<td>0.23469</td>
<td>0.1902</td>
<td>0.23469</td>
<td>0.1902</td>
</tr>
<tr>
<td>6</td>
<td>0.20365</td>
<td>0.11830</td>
<td>0.20365</td>
<td>0.11830</td>
</tr>
<tr>
<td>8</td>
<td>0.15902</td>
<td>0.07394</td>
<td>0.15902</td>
<td>0.07394</td>
</tr>
<tr>
<td>10</td>
<td>0.11286</td>
<td>0.05756</td>
<td>0.11286</td>
<td>0.05756</td>
</tr>
</tbody>
</table>

Table 2. Values of $b_{2(1)}$ for FSK and PSK systems.

<table>
<thead>
<tr>
<th>SNR</th>
<th>FSK</th>
<th>PSK</th>
<th>FSK</th>
<th>PSK</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>0.27975</td>
<td>0.23469</td>
<td>0.27975</td>
<td>0.23469</td>
</tr>
<tr>
<td>4</td>
<td>0.23469</td>
<td>0.1902</td>
<td>0.23469</td>
<td>0.1902</td>
</tr>
<tr>
<td>6</td>
<td>0.20365</td>
<td>0.11830</td>
<td>0.20365</td>
<td>0.11830</td>
</tr>
<tr>
<td>8</td>
<td>0.15902</td>
<td>0.07394</td>
<td>0.15902</td>
<td>0.07394</td>
</tr>
<tr>
<td>10</td>
<td>0.11286</td>
<td>0.05756</td>
<td>0.11286</td>
<td>0.05756</td>
</tr>
</tbody>
</table>
Receivers with memory for channels with memory have been proposed recently. These receivers try to utilize the channel memory to yield a better performance. These receivers are expected to alter the correlation of errors and it is the objective here to examine this relationship. A receiver with memory for fading channels was considered in [5]. The system considered in this section and its error clustering properties are considered in the next section. Fading is assumed to be slow so that it can be assumed to be constant during a signalling interval and its effect may be represented as a sequence of dependent random variables \( \{v_k, b_k\} \) where \( v_k \) and \( b_k \) denote the contributions of amplitude and phase fading respectively. The additive noise \( n(t) \) is assumed to be white Gaussian with power spectrum \( N_0 \).

Let \( H^i \) denote the hypothesis that \( i \) is transmitted during \( k \)th signalling interval. The received signal during the \( k \)th signalling interval \( [(k-1)T, kT) \) under the hypothesis \( H^i \) is given by

\[
r_k(t) = r(t + (k - 1)T) = v_k^i(t, \theta_k^i) + n_k(t) \quad 0 \leq t < T; \quad i = 0, 1
\]  

The relationship between \( (x_k^i, y_k^i) \) and \( (v_k^i, \theta_k^i) \) if \( H^i \) is true, for \( i = 0, 1 \), is given by

\[
\begin{align*}
\gamma_k^i &= \delta_{i,j} v_k^i \cos \theta_k^i + n_k^i \quad \delta_{i,j} = \delta_{j,k} = \delta_{k,j} = 0, 1 \\
\gamma_k^i &= \delta_{i,j} v_k^i \sin \theta_k^i + n_k^i \quad \delta_{i,j} = \delta_{j,k} = \delta_{k,j} = 0, 1
\end{align*}
\]  

where \( \delta_{i,j} \) is the Kronecker delta function.

The receiver with memory has an adaptive structure and consists of an estimator and a detector. A limited-memory filter with decision feedback computes an estimate of the fading parameters \( \hat{a}_k \) and \( \hat{b}_k \) on the basis of the previous \( M \) observations and decisions. The estimates \( \hat{a}_k \) and \( \hat{b}_k \) are assumed to be linear functions of the past \( M \) observations with the following structure

\[
\begin{align*}
\hat{a}_k &= \sum_{i=1}^{M} \alpha_i (x_{k-i}^i)^{1/2} + (1 - z_{k-i}^i) y_{k-i}^i) \\
\hat{b}_k &= \sum_{i=1}^{M} \beta_i (x_{k-i}^i)^{1/2} + (1 - z_{k-i}^i) y_{k-i}^i)
\end{align*}
\]  

where \( z_{k,i} \) denotes the detector decision on the \( k \)th bit. The estimator coefficients \( \{\alpha_i, \beta_i\} \) are computed so as to minimize the MSE. Central limit theorem for dependent random variables is employed to obtain asymptotic approximations under the strong mixing condition. The detector is optimized under the minimization of probability of error criterion. The following optimum decision rule is obtained by minimizing the Bayes risk function.

\[
\begin{align*}
&f(x_{k-1}^0, y_{k-1}^0, x_0, y_0 | H_k^i, z_k) = \frac{f(x_{k-1}^0, y_{k-1}^0, x_0, y_0 | z_k) \delta_{i,k}}{\sum_{i=0}^{1} f(x_{k-1}^0, y_{k-1}^0, x_0, y_0 | z_k) \delta_{i,k}} \quad i = 0, 1
\end{align*}
\]  

where \( f(., ., ., ., z_k) \) represents the conditional density under a given hypotheses \( H_k^i \) and the past decision sequence \( z_k \). The conditional density based on the estimates of the fading parameters can be used to obtain the above density functions, i.e.,

\[
\begin{align*}
&f(x_{k-1}^0, y_{k-1}^0, x_0, y_0 | z_k) = e^{-\frac{1}{2} \left( x_{k-1}^0 - \hat{a}_k \right)^2 + \left( y_{k-1}^0 - \hat{b}_k \right) \beta_k} \delta_{i,k} \quad i = 0, 1
\end{align*}
\]  

The conditional probability of bit error is given by

\[
\begin{align*}
P(e_k | z_k) &= \frac{1}{2} \left[ \text{erf}(T_k - \xi_k) + N_0 \left( 1 + \xi_k \right) \right] + \text{erfc}(T_k / N_0) \\
&= \frac{1}{2} \left[ \text{erf}(T_k - \xi_k) + N_0 \left( 1 + \xi_k \right) \right]
\]  

Similarly
The average probability of bit error, $p$, and the average probability of two consecutive errors is computed from

$$P(e_k = 1) \cap (e_{k-1} = 1) \mid Z_k = \text{erfc}(T_k / N_0^{1/2}) \text{erfc}(T_k / N_0^{1/2}) + 2 \text{erf}(T_k - Z_k)/(2N_0 + Z_k^2))$$

$$= \frac{1}{2} \pi \left[ \frac{\pi (N_0 + Z_k^2(1 - p^2))}{2(1 - p^2)(N_0 + Z_k^2)} \right]^{-1} \exp \left[ - \frac{x^2 - 2\omega x + \omega^2}{2(1 - p^2)(N_0 + Z_k^2)} \right] dx dy$$

(29)

The average probability of bit error, $p$, and the average probability of two consecutive errors is computed from

$$P = 2^M \sum_{Z_k} P(e_k = 1) \mid Z_k$$

(30)

$$P(e_k = 1) \cap (e_{k-1} = 1) = 2^M \sum_{Z_k} P(e_k = 1) \mid Z_k$$

(31)

These expressions can then be used in (19) and (20) to obtain $\alpha(1)$ and $\beta(1)$ are presented in Figs. 1 and 2. For these results, the numerical values chosen are $m = 5, p = 0.9$ and $M = 8$. The discrepancy in the behavior of $\alpha(1)$ and $\beta(1)$ for low SNR can be attributed to the approximation made in the design of the receiver with memory. The approximate design implied small error probability which certainly is not valid for the low SNR case.

VII. SUMMARY AND CONCLUSIONS

In this paper, we have examined error clustering in digital communication systems. The relationship between the actual communication systems and the errors generated by them has been investigated and some preliminary results have been presented. Two measures which represent the relationship quantitatively have been defined. Two classes of digital communication systems have been considered. First we considered the relationship of various modulation formats and error clustering in the presence of correlated noise. The other class considered a communication system with a receiver with memory over fading channels. The examples considered the measures related to double errors. This study is expected to find applications in the design of more efficient communication systems since more information about the occurrence of error clusters is available.

REFERENCES

Figure 1. $a_2(t)$ for the receivers with and without memory.

Figure 2. $b_2(t)$ for the receivers with and without memory.