"The BLER depends largely on the mean block length and is not very sensitive to the type of distribution involved."

Computation of Block Error Rates with Randomly Varying Block Sizes

by

Efren F. Abaya, Ph.D.*

ABSTRACT

The block error rate (BLER) is a basic parameter of data communication channels which is usually computed on the assumption of independent and identically distributed bit errors and fixed block sizes. This paper presents a method of computing block error rates when the block sizes are allowed to vary according to a known probability distribution. It is shown that under some simple conditions the block error rate is relatively insensitive to the actual distribution or to the variance of the block length, and that it depends largely on the average block length. Some analytic and numerical computations are presented to illustrate the method.

INTRODUCTION

The block error rate is a basic parameter that influences the utilization and time delay of data communication channels utilizing an ARQ procedure. In analytic studies, it is usually computed on the assumption of independent and identically distributed bit errors and fixed block sizes. This paper presents a method of computing block error rates when the block sizes are allowed to vary according to a known probability distribution. It is shown, under some simple conditions which are realized in most real situations, that the block error rate is relatively insensitive to the actual distribution of block sizes, and that it depends largely on the average block length. Thus, the assumption of fixed block sizes is basically sound and reasonable in practice. Some analytic bounds and numerical calculations are presented to explore the limits of this assumption.

These results are useful in building analytic models of data communication channels where an expression for the block error rate in terms of more basic quantities such as bit error rate and block length is needed. Although the assumption of fixed block size is common, in practice the block size may be allowed to vary quite a bit. One study on computers with time-sharing terminals showed that a geometric distribution can be used to model the number of characters transmitted in one burst [2]. In other applications, transmissions from heterogeneous sources may be time interleaved (multiplexed) resulting in variable block sizes.

---

*Professor of Communication Systems, Department of Electrical Engineering, U.P. College of Engineering, Diliman, Quezon City, Philippines.

STATEMENT OF THE PROBLEM

The bit error rate (BER), symbolized by P_B, and the block error rate (BLER), symbolized by P_F, are the probabilities of encountering errors in one bit, and in one block of contiguous bits, respectively. In this paper, it is assumed that error bits are independent of one another, and can affect all bit positions with the same probability (i.e., errors are independent and identically distributed). With this assumption, the error rates for a block of n bits are related by the equation [4].

$$P_F = 1 - (1-P_B)^n \quad . \tag{1}$$

Usually P_B is much less than unity and the product nP_B is much smaller than one. For example, P_B is on the order of 10^{-3} or smaller, while nP_B is less than 0.1. Then an easily computed approximation to the BLER is [4].

$$P_F \approx nP_B \quad . \tag{2}$$

The error in this approximation is $0(n^2 P_B{}^2)$.

In this paper, the block length N is allowed to take a known probability distribution P(n) where

$$P(n) = Pr[N=n] \qquad n=1,2,3,...$$

The notation $Pr[\cdot]$ is used for probabilities. The distribution may have a finite or infinite number of atoms.

The problem discussed in this paper is the computation of BLER when the block length N has distribution P(n) and the BER is specified.

In practice, the average block length $\bar{N}$ is used in (1) or (2) giving the approximations

$$P_F \approx 1 - (1-P_B)^{\bar{N}} \tag{3}$$

or

$$P_F \approx \bar{N} P_B \tag{4}$$

The closeness of these approximations to the exact value (1) will also be examined in this paper.

NOTATION

The following symbols are used throughout the paper. Other symbols are defined as needed.

P_B	=	bit error rate (BER)
Q_B	=	$1 - P_B$
P_F	=	block error rate (BLER)
N	=	random variable representing block length (bits); $N \geq 1$
P(n)	=	probability distribution of N; n = 1,2,3,...
$\bar{N}$	=	mean block length (bits)
$\sigma_N{}^2$	=	variance of N
$Pr[\cdot]$	=	probability of the event in brackets
$Pr[\cdot/\cdot]$	=	conditional probability
$E[\cdot]$	=	expectation
G(s)	=	generating function of P(n)

ANALYSIS

An expression for BLER is simply obtained by using conditional probabilities. Let Ω symbolize the event that a block contains errors. Then

$$P_F = Pr[\Omega]$$
$$= E[\, Pr[\Omega/N]\,].$$

Since errors are independent and identically distributed

$$Pr[\Omega/N] = 1 - Q_B{}^N$$

Therefore

$$P_F = 1 - E[Q_B{}^N] \quad . \tag{5}$$

The generating function G(s) of a distribution on the non-negative integers is defined as [1], [3, p. 11-12].

$$G(s) = E[s^N]$$
$$= \sum_{n=0}^{\infty} s^n P(n). \tag{6}$$

This is related to the moment generating function E[exp (tN)] by an exponential transformation of the variable. Although the moment generating function is more common, the formulation G(s) is more convenient for our purpose and will be used here. Tables of generating functions may be found in [3, p. 16]. Some examples are shown in Table 1 of this paper.

Using the generating function, the previous expression (5) for BLER becomes

$$P_F = 1 - G(Q_B) \quad . \tag{7}$$

Equation (7) is the basic formula used in this paper to calculate and investigate the BLER.

Since $Q_B{}^N$ is a convex (upward) function of N, Jensen's Inequality [3, p. 249] may be applied to (5). Then a binomial expansion and approximation may be used on (3). There results an order relationship among the BLER and its approximations

$$P_F \leqq 1 - Q_B{}^{\overline{N}} \leqq \overline{N} P_B \tag{8}$$

Hence both approximations are on the conservative side. Of course, (3) is always valid, whereas (4) is meaningless if $\overline{N}P_B$ exceeds unity. It is interesting to note that the BLER for the case of variable length blocks is less than the BLER for fixed length blocks of $\overline{N}$ bits. How much less this is will be examined next in this section.

Taking a Taylor's Series expansion of (7) about s = 1 and using the fact that G(1) = 1 gives

$$P_F = \sum_{k=1}^{\infty} (-1)^{k+1} \frac{G^{(k)}(1) P_B{}^k}{k!} \tag{9}$$

where $G^{(k)}(1)$ is the k-th derivative evaluated at s = 1.

The derivatives of the generating function are easy to calculate in terms of the moments of N[1, p. 265]. By direct differentiation of (6)

$$G^{(1)}(1) = E[N] = \bar{N} \qquad (10a)$$

$$G^{(2)}(1) = E[N^2] - E[N]$$

$$= \sigma_N^{\,2} + \bar{N}^2 - \bar{N} \qquad (10b)$$

$$G^{(3)}(1) = E[N^3] - 3E[N^2] + 2E[N] \qquad (10c)$$

$$G^{(k)}(1) = E[N(N-1)\ldots(N-k+1)] \qquad (10d)$$

Substituting Equations (10) in (9) and neglecting third and higher order terms yields

$$P_F \approx \bar{N}P_B - \frac{(E[N^2] - \bar{N})P_B^{\,2}}{2}$$

$$= \bar{N}P_B - \frac{(\sigma_N^{\,2} + \bar{N}^2 - \bar{N})P_B^{\,2}}{2} \qquad (11)$$

The error in the approximation may be bounded since all of the derivatives in Equations (10) are positive, so that (9) is an alternating series. The result is

$$\left| P_F - \bar{N}P_B \right| \leq \frac{(E[N^2] - \bar{N})P_B^{\,2}}{2} \qquad (12a)$$

or

$$\left| \frac{P_F - \bar{N}P_B}{\bar{N}P_B} \right| \leq \frac{1}{2} \left| \frac{E[N^2]}{\bar{N}} - 1 \right| P_B \qquad (12b)$$

The right hand side of Equations (12) may be used to estimate the accuracy of the approximation $\bar{N}P_B$. Since P_B itself and $\bar{N}P_B$ should be very small, the right hand side of (12) is dominated by $(\sigma_N^{\,2}/\bar{N})P_B$. Hence this approximation is useful if this last expression is small and also $\bar{N}P_B$ is small.

Next, the approximation (3) is expanded by using the Binomial Theorem into

$$1 - Q_B^{\,\bar{N}} = \bar{N}P_B - \frac{\bar{N}(\bar{N}-1)P_B^{\,2}}{2} + \frac{\bar{N}(\bar{N}-1)(\bar{N}-2)P_B^{\,3}}{6} - \ldots$$

Subtracting this from (9) gives, up to the second order term

$$P_F - (1-Q_B^{\,\bar{N}}) \approx - \frac{\sigma_N^{\,2}P_B^{\,2}}{2} \qquad (13)$$

A rough estimate of the relative error in the approximation is

$$\left| \frac{P_F - (1-Q_B^{\,\bar{N}})}{1-Q_B^{\,\bar{N}}} \right| \approx \left| \frac{P_F - (1-Q_B^{\,\bar{N}})}{\bar{N}P_B} \right|$$

$$\approx \frac{1}{2}\,\frac{\sigma_N^{\,2}P_B}{\bar{N}} \qquad (14)$$

Hence this approximation is useful under the same condition as the previous one, namely, that $(\sigma_N{}^2/\bar{N})P_B$ is small.

Equation (13) provides an interesting comparison of the BLER for variable length blocks and for fixed length blocks of $\bar{N}$ bits. It was noted in (8) that variable length blocks with mean length $\bar{N}$, regardless of distribution, have smaller BLER than fixed length blocks of $\bar{N}$ bits. It may be seen in (13) that the difference in BLER is proportional to the variance of N. Hence, the greater the variability of block lengths, the smaller the BLER is. This conclusion is tempered by the fact that as $\sigma_N{}^2$ increases, the higher order moments also increase so that the terms which were neglected in (13) become significant. The numerical examples presented in the next section, however, support this observation.

This phenomenon can be explained in the following way. If the mean length $\bar{N}$ is fixed, the only way to increase the variance is to have very long blocks with small probabilities and very short blocks with large probabilities. Then the overall BLER is dominated by those of the small blocks. As the variance increases, the influence of the small blocks increases, and therefore the BLER decreases. Such behavior may be seen clearly in the two-point test distribution presented in the next few sections.

EXAMPLES

The following examples serve to illustrate the analytic results of the previous section. The distributions used are the following:

(a) The uniform distribution on the integers L, L + 1, . . . , U designated Unif (L, U).

(b) The "binomial" distribution on the integers 1, 2, . . . , U designated Bin (U, p). The conventional binomial distribution starts from 0. Since the minimum block length is 1 bit, the distribution is translated by 1.

(c) The 'geometric' distribution on the positive integers, designated Geom(p). This is the conventional geometric distribution translated to begin from N=1.

(d) A two-point test distribution whose mean and variance can be arbitrarily chosen to match those of the first three distributions. This models a system with only two block sizes.

The uniform and binomial distributions both have a maximum block length, while the geometric theoretically permits unlimited block lengths. The geometric distribution can be used to model the number of characters transmitted in time-sharing terminal computers [2].

Some properties of these distributions are given in Table 1. In the next few sections expressions for BLER will be presented.

UNIFORM DISTRIBUTION

From (7) and Table 1, the BLER with a Unif (L, U) distribution of block lengths is

$$P_F = 1 - \frac{Q_B{}^L(1 - Q_B{}^{U-L+1})}{P_B(U-L+1)} \tag{15}$$

From the inequality

$$(U-L+1)P_B \geq 1 - Q_B{}^{U-L+1}$$

it follows that

$$1 - Q_B{}^L \leq P_F \quad .$$

Combined with (8), this implies that

$$\left|P_F - \bar{N}P_B\right| \leq \left|1 - Q_B{}^L - \bar{N}P_B\right|$$
$$\approx \left|LP_B - \frac{(L+U)P_B}{2}\right|$$
$$= \frac{(U-L)P_B}{2} \qquad (16)$$

$$\left|\frac{P_F - \bar{N}P_B}{\bar{N}P_B}\right| \approx \frac{(U-L)}{U+L} < 1 \qquad (17)$$

Hence the relative error should not exceed 100 percent if P_B is even moderately small.

"BINOMIAL" DISTRIBUTION

For the Bin (U, p_D) distribution of block lengths, a little algebraic manipulation of (7) with the expression for G from Table 1 gives

$$P_F = 1 - (1 - p_D P_B)^{U-1} Q_B \quad . \qquad (18)$$

Recall that

$$(1-x)^m \leq 1 - mx + \frac{m(m-1)x^2}{2} \quad .$$

Letting $x = p_D P_B$ and $m = U-1$ in (18) gives

$$P_F \geq 1 - \left[1 - (U-1)p_D P_B + \frac{(U-1)(U-2)p_D{}^2 P_B{}^2}{2}\right] Q_B$$

$$= P_B\left[1 + (U-1)p_D Q_B\right] - \frac{(U-1)(U-2)p_D{}^2 P_B{}^2 Q_B}{2}$$

Combining this inequality with (8) and using the Triangle Inequality gives a bound on the error of the approximation

$$\left|P_F - \bar{N}P_B\right| \leq P_B\left|\left[1 + (U-1)p_D Q_B\right] - (Up_D + q_D)\right|$$
$$+ \frac{(U-1)(U-2)p_D{}^2 P_B{}^2 Q_B}{2}$$

$$= (U-1)p_D P_B{}^2 + \frac{(U-1)(U-2)p_D{}^2 P_B{}^2}{2}$$

$$\leq (U-1)p_D P_B{}^2 + \frac{(U-1)(U-2)p_D P_B{}^2}{2}$$

$$= \frac{U(U-1)p_D P_B{}^2}{2} \qquad (19)$$

$$\left| \frac{P_F - \bar{N}P_B}{\bar{N}P_B} \right| = \frac{U}{2} \frac{(U-1)p_D}{(U-1)p_D + 1} P_B$$

$$\leq \frac{UP_B}{2} \tag{20}$$

Since UP_B is usually less than 1, the relative error is less than 50 percent.

"GEOMETRIC" DISTRIBUTION

For the Geom (p_D) distribution the resulting expression for BLER is

$$P_F = \frac{P_B}{p_D + q_D P_B} \tag{21}$$

Then

$$P_F - \bar{N}P_B = \frac{P_B}{p_D + q_D P_B} - \frac{P_B}{p_D}$$

$$= \frac{-q_D P_B^2}{p_D(p_D + q_D P_B)} \tag{22}$$

$$|P_F - \bar{N}P_B| = \frac{P_B^2}{p_D(\frac{p_D}{q_D} + P_B)}$$

$$\leq P_B/p_D$$

$$= \bar{N}P_B$$

$$\left| \frac{P_F - \bar{N}P_B}{\bar{N}P_B} \right| \leq 1$$

Since $\bar{N}P_B$ should be less than unity, the relative error should not exceed 100 percent. In this case (22) provides the best expression for the error in the approximation $\bar{N}P_B$.

NUMERICAL CALCULATIONS

Exact and approximate block error rates are presented in Table 2 for the three distributions discussed above. The parameters were chosen so that the mean block length $\bar{N}$ = 50.5 for all of them. However, the variances of the distributions differ widely.

In addition, four cases of the two-point test distribution are included. The mean block lengths of these are also $\bar{N}$ = 50.5. In the first three test cases, the variances are matched as closely as possible to the variances of the binomial, uniform and geometric distributions, respectively. The last test case has a variance much larger than any of the others.

Thus, this set of distributions permits a comparison of BLER among different distributions (both bounded and unbounded), among distributions with different variances, and among the same family of distributions with different variances. The last rows in Table 2 are the approximations based on mean block length, for comparison with the actual values of BLER.

The following observations may be made:

(1) The BLER is not very sensitive to the type of distribution, especially if the BLER is small. Even when the BER is not small (e.g., 10^{-2} and up), the BLER does not vary much if the variance is not too large. This observation is consistent with (13) which says that the BLER stays within approximately $\sigma_N{}^2 P_B{}^2/2$ of $(1\text{-}Q_B{}^{\bar{N}})$ regardless of the distribution.

(2) The BLER is also not very sensitive to the variance of the distribution. It is only when the BER is large (e.g., 0.1) and the variance is very large that the BLER's begin to be markedly different.

(3) The BLER depends mostly on the mean block length, with type of distribution or variance providing only a small perturbation.

(4) The approximations (3) and (4) are very good for small BER. Even for large BER, the approximations, particularly (3), are reasonable if the variance is not too large (e.g., in these examples, when $\sigma_N{}^2/\bar{N} < 1$). This is predicted by (11)-(14).

(5) As noted before, the BLER goes down as the variance increases. This pattern is evident even if the distributions are of different types (e.g., the binomial, uniform and geometric). It is most apparent in the case of the test distribution where, as the variance goes up, Pr[N=1] approches unity. A reason for this phenomenon was already suggested in an earlier section.

CONCLUSION

The effect of a randomly varying block size on the block error rate has been studied. An analytic expression for BLER which involves the generating function of the distribution of block lengths was derived. It was shown analytically and through numerical examples that both $1\text{-}Q_B{}^{\bar{N}}$ and $\bar{N}P_B$ are reasonable approximations when either the BER or the variance of the distribution of block lengths is not large. The BLER thus depends largely on the mean block length and is not very sensitive to the type of distribution involved. Thus the assumption of fixed block size may be regarded as reliable and reasonable in practice.

REFERENCES

[1] WILLIAM FELLER, *An Introduction to Probability Theory and Its Applications,* vol. I, 3rd ed., N.Y.: John Wiley & Sons, 1968, p. 265.

[2] E. FUCHS and P.E. JACKSON, "Estimates of Distributions of Random Variables for Certain Computer Communications Traffic Models," *Communications of the ACM,* vol.13 (Dec. 1970), pp. 752-757; reprinted in P.E. Green and R.W. Lucky, *Computer Communications,* N.Y., IEEE Press, 1975.

[3] SAMUEL KARLIN and HOWARD M. TAYLOR, *A First Course in Stochastic Processes,* 2nd ed., N.Y.: Academic Press, 1975.

[4] ANDREW S. TANNENBAUM, *Computer Networks,* N.J.: Prentice-Hall, Inc., 1981, p. 126.

Table 1
Properties of Discrete Distributions

NAME	DISTRIBUTION P(n)	MEAN	VARIANCE	G (s)
Unif(L,U)	$(U-L+1)^{-1}$ n=L, . . . , U	$\frac{L+U}{2}$	$\frac{(U-L+1)^2-1}{12}$	$\frac{s^L(1-s^{U-L+1})}{(1-s)(U-L+1)}$
Bin(U,p)	$\binom{U-1}{n-1} p^{n-1}q^{U-n}$ q=1−p n=1, . . . , U	Up+q	(U−1)pq	$s(q+ps)^{U-1}$
Geom(p)	pq^{n-1} n=1,2, . . . q=1−p	1/p	q/p^2	ps/(1−qs)
Two point	P(1)=p $P(\frac{N-p}{q})=q$ q=1−p	$\bar{N}$	$(\bar{N}-1)^2\frac{p}{q}$	$sp+s^{(\bar{N}-p)/q}q$

Table 2
Comparison of BLER for $\bar{N}$ = 50.5

DISTRIBUTION	BIT ERROR RATE			
	10^{-5}	10^{-3}	10^{-2}	10^{-1}
Bin(100,0.5) var=24.75	5.04874E−4	4.92583E−2	3.97274E−1	9.94391E−1
Unif(1,100) var=833.25	5.04795E−4	4.88735E−2	3.72372E−1	9.10002E−1
Geom(2/101) var=2,499.8	5.04750E−4	4.81182E−2	3.37793E−1	8.48739E−1
Two point p_D=.01 var.=24.75	5.04874E−4	4.92581E−2	3.97134E−1	9.86408E−1
Two point p_D=.25 var=816.75	5.04834E−4	4.88771E−2	3.70011E−1	7.74355E−2
Two point p_D=.505 var=2,499.8	5.04750E−4	4.80808E−2	3.20676E−1	5.45488E−1
Two point p_D=.9 var=22,052.	5.03774E−4	4.00189E−2	1.08316E−1	1.90000E−1
Fixed length $1-Q_B^{\bar{N}}$	5.04875E−4	4.92701E−2	3.98027E−1	9.95111E−1
Approximation $\bar{N}P_B$	5.05000E−4	5.05000E−2	5.05000E−1	-----