On Adaptive Go-Back-N ARQ Protocol for Variable-Error Rate Channels

A. Annamalai, *Student Member, IEEE,* Vijay K. Bhargava, *Fellow, IEEE,* and W.-S. Lu, *Senior Member, IEEE*

***Abstract*— This letter presents a simple method to simultaneously optimize a multiplicity of design parameters for the adaptive automatic repeat request strategy previously reported, and subsequently provides a quantitative measurement that reflects the appropriateness of the selected parameters. An exact analytical expression that allows us to compute the throughput crossover probability between the two different protocols is derived. The results provide fundamental insights into how these key parameters interact and determine the system performance.**

***Index Terms*— Automatic repeat request, optimization methods, time-varying channels, adaptive systems, protocol, feedback communication.**

I. INTRODUCTION

IN A RECENT study [1], a novel channel state estimation (CSE) technique was introduced by which the count of the previously received ACK (positive acknowledgment) and NACK (negative acknowledgment) status signals were used to switch between different Go-Back-N ARQ protocols, and subsequently adapted to the time varying nature of the wireless channels. The adaptive automatic repeat request (ARQ) strategy for a Gilbert–Elliott channel is depicted in Fig. 1. While in the "good" channel state, the transmitter follows the basic Go-Back-N procedure. In this operation mode, the transmitter goes back N blocks upon reception of an NACK. Whereas, in the "bad" channel state, the transmitter operates in an n-copy transmission mode, which is similar to the basic Go-Back-N except for sending n copies of a data block in each transmission and, if necessary, in each retransmission.

Notice that in [1], the system design parameters were selected by the trial-and-error method. In contrast, here we adopt Broyden–Fletcher–Goldfarb–Shanno (BFGS) algorithm [2], a well-known quasi-Newton optimization method, to obtain the suboptimal values for these parameters in a systematic and efficient manner. This approach is particularly attractive in cases where a large number of variables need to be optimized simultaneously.

We model this problem as the minimization of mean square error to the desired performance, with design parameters α and β as the optimization variables. Consequently, this method lends itself to a quantitative measure of the suitability of the selected parameters. The design variables correspond to the observation interval (in terms of number of packets) associated

Paper approved by K.-C. Chen, Editor for Wireless Data Communications of the IEEE Communications Society. Manuscript received February 3, 1997; revised July 3, 1997 and November 3, 1997. This work was supported by a Strategic Project Grant from the Natural Sciences and Engineering Research Council (NSERC) of Canada.

The authors are with the Department of Electrical and Computer Engineering, University of Victoria, Victoria, B.C., Canada V8W 3P6 (e-mail: bhargava@ece.uvic.ca).

Publisher Item Identifier S 0090-6778(98)08130-6.

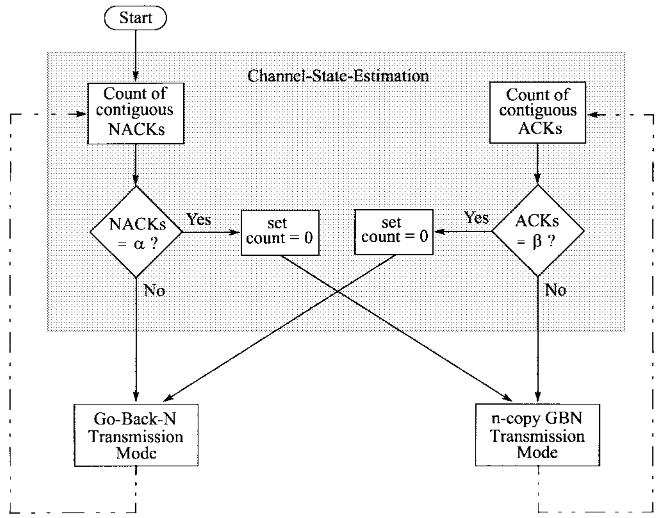


Fig. 1. An adaptive Go-Back-N ARQ scheme with transition between two operation modes.

with each distinct channel state. Next, we have derived an exact analytical expression that enables us to determine the throughput crossover probability between the basic Go-Back-N $(n = 1)$ and n-copy transmission strategy. The knowledge of this parameter is essential in the proposed adaptive scheme because the switching occurs at the vicinity of this point. The results presented in this letter provide fundamental insights into how these key parameters interact and affect the system performance.

II. PROBLEM FORMULATION

Let $T(P_e)$ and $\hat{T}(P_e)$ denote the throughput performance of the desired (ideal) and adaptive ARQ protocol, respectively, where P_e corresponds to the block error probability. Our task is to find the optimum design parameters such that $\hat{T}(P_e)$ best approximates $T(P_e)$ in the sense that the mean square error (MSE) function is minimized

$$\begin{aligned}\underset{\{\alpha,\beta\}\in Z}{\text{minimize}}\; E(\alpha,\beta) &= \int_0^1 [T(P_e) - \hat{T}(P_e)]^2\, dP_e\\ &\cong \Delta \sum_{k=1}^{K} [T(P_{ek}) - \hat{T}(P_{ek})]^2\\ \text{subject to:}\quad & \alpha_{\min} < \alpha < \alpha_{\max}\\ & \beta_{\min} < \beta < \beta_{\max}\end{aligned} \tag{1}$$

where K corresponds to the sample size, P_{ek} denotes the block error probability of the kth sample, and the optimization variables can assume any value from the set Z, which consists of positive integers. In other words, the error function (MSE) is our objective function, and its local minimum point contains the information of the optimum design parameters, $\boldsymbol{x}^* = [\alpha^* \quad \beta^*]^T$. Discrete determination of $E(\alpha, \beta)$ is valid if the step size Δ between the consecutive data points is selected to be relatively small. In this letter, we decide to choose the samples to be equally spaced in the region $0 \leq P_e \leq 1$, with a step-size of $\Delta = 0.0015$. The throughput formulas are described by [1]

$$T(P_{ek}) = \begin{cases} S_{1k} = \dfrac{1 - P_{ek}}{1 + (N-1)P_{ek}}, & \text{if } P_{ek} \leq P_{co} \\ S_{2k} = \dfrac{1 - P_{ek}^n}{n + (N-1)P_{ek}^n}, & \text{if } P_{ek} > P_{co} \end{cases} \tag{2}$$

$$\begin{aligned} \hat{T}(P_{ek}) &= \Lambda_{1k} S_{1k} + \Lambda_{2k} S_{2k} \\ &= [p_{2k}/(p_{1k} + p_{2k})] S_{1k} + [p_{1k}/(p_{1k} + p_{2k})] S_{2k} \end{aligned} \tag{3}$$

where P_{co} is the crossover probability between the Go-Back-N and n-copy transmission modes, whereas Λ_{1k} and Λ_{2k} correspond to the steady-state probabilities of the channel at low- and high-error rate states, respectively. Identical to the definitions in [1], $p_{1k} = P_{ek}^{\alpha}$ and $p_{2k} = (1 - P_{ek})^{\beta}$ are the state transition probabilities, and the system throughput for low- and high error state are denoted as S_{1k} and S_{2k}, respectively.

As we will describe shortly, the optimal solution to our problem (i.e., when the performance curve of the proposed CSE scheme coincides with the desired performance envelope over the entire $0 \leq P_e \leq 1$ range) exists in the infinite $\alpha - \beta$ space. In this case, the optimal transition probabilities will be equal at $P_{ek} = P_{\text{co}}$ since $\Lambda_{1k} = \Lambda_{2k}$. Therefore, it can be readily shown that the optimal design parameters are related by

$$\frac{\beta^*}{\alpha^*} = \frac{\ln(P_{\text{co}})}{\ln(1 - P_{\text{co}})}. \tag{4}$$

Since the optimal solution (i.e., local minimum point) does not lie in a reasonable value range, one can resort to the suboptimal solutions with some sacrifice in performance. If we select α and β values to be very large, then this scheme will lose its ability to adapt to moderately fast channel variations. On the other hand, extremely small values of α and β will result in premature (unnecessary) switching, and poor fit to the desired performance curve. Therefore, we have introduced additional boundary constraints to the design parameters, which will be specified by the channel behavior and/or the intended application. In our minimization problem, these boundary constraints can be eliminated via transformation $y = (e^z - e^{-z}/e^z + e^{-z}) \equiv \tanh(z)$. The hyperbolic tangent is a monotonically increasing function with respect to z that maps the entire 1-D space $-\infty < z < \infty$ to $-1 < \tanh(z) < 1$. Subsequently, it is easy to show that the linear relationship described in (5) gives a map from $(-\infty, \infty)$ to $(x_{\min}, \ x_{\max})$

$$x = \left(\frac{x_{\max} - x_{\min}}{2}\right) \tanh(z) + \left(\frac{x_{\max} + x_{\min}}{2}\right). \tag{5}$$

Finally, the objective function for an unconstrained optimization is obtained by substituting

$$\begin{aligned} &[\alpha \quad \beta]^T \\ &\quad = \begin{bmatrix} 0.5(\alpha_{\max} - \alpha_{\min}) \tanh(z_\alpha) + 0.5(\alpha_{\max} + \alpha_{\min}) \\ 0.5(\beta_{\max} - \beta_{\min}) \tanh(z_\beta) + 0.5(\beta_{\max} + \beta_{\min}) \end{bmatrix} \end{aligned}$$

into (1), and minimizing the MSE function with respect to $\boldsymbol{z} = [z_\alpha \quad z_\beta]^T$, i.e.,

$$\underset{\{z_\alpha, z_\beta\} \in \mathcal{R}}{\text{minimize}} \; E(z_\alpha, z_\beta) = \int_0^1 [T(P_e) - \hat{T}(P_e)]^2 \, dP_e \tag{6}$$

where $\mathcal{R}$ consists of the set of real numbers. The corresponding gradient function (partial derivatives) can be obtained without much difficulty.

It is worth mentioning that the BFGS algorithm has been adopted in our optimization process for the following reasons: 1) it is an efficient nonlinear optimization routine because it only requires the computation of the gradient vector, and it is unnecessary to manipulate or invert the Hessian matrix (which could be difficult or time-consuming, especially when the exact expression for the gradient vector is not available); 2) it is reliable and possesses a unique property that guarantees the subsequent updating formulas to be positive definite if the initial matrix is forced to be positive definite; and 3) the algorithm is quite tolerant (robust) to line-search imprecisions. Readers are encouraged to refer to [2] for further details of this algorithm.

We now derive the exact throughput crossover probability between the basic Go-Back-N and n-copy transmission schemes. After some algebraic manipulations, the throughput difference between the two protocols can be restated as

$$S_2 - S_1 = \frac{\left\{\dfrac{P_e[1 - P_e^{n-1}]}{1 - P_e} - \dfrac{n-1}{N}\right\}(1 - P_e)N}{[n + (N-1)P_e^n][1 + (N-1)P_e]}. \tag{7}$$

It is evident that the denominator is always greater than 0. Thus, with the assumption of noiseless feedback communication, the n-copy transmission outperforms the basic Go-back-N when the first numerator term of (7) is positive, namely,

$$\begin{aligned} &\frac{P_e}{1 - P_e}[1 - P_e^{n-1}] - \frac{n-1}{N} \\ &\qquad = P_e^{n-1} + P_e^{n-2} + \cdots + P_e - \frac{n-1}{N} > 0. \end{aligned} \tag{8}$$

Therefore, the crossover probability P_{co} occurs at $P_e = 1.0$ and when the inequality in (8) is replaced with an equality. The knowledge of this exact probability (which can be computed numerically) is essential in the design of our adaptive protocol because the switching will occur at this transition point. For the special cases of $n = 2$ and $n = 3, P_{co}$ reduces to $1/N$ and $(-1 + \sqrt{1 + (8/N)})/2$, respectively. However, for noisy feedback channels, the n-copy transmission performs better than basic Go-Back-N transmission mode only if expression

TABLE I
SUBOPTIMAL DESIGN PARAMETERS AND THEIR CORRESPONDING ERROR FUNCTION (MSE) FOR DIFFERENT VALUES OF $n, N,$ AND $\alpha_{\max}$

n	N	$\alpha_{max} = 3$		$\alpha_{max} = 5$		$\alpha_{max} = 10$	
		$\hat{\mathbf{x}}^* = \begin{bmatrix}\alpha \\ \beta\end{bmatrix}$	$E(\hat{\mathbf{x}}^*)$	$\hat{\mathbf{x}}^* = \begin{bmatrix}\alpha \\ \beta\end{bmatrix}$	$E(\hat{\mathbf{x}}^*)$	$\hat{\mathbf{x}}^* = \begin{bmatrix}\alpha \\ \beta\end{bmatrix}$	$E(\hat{\mathbf{x}}^*)$
2	5	$[3\ 22]^T$	4.545×10^{-6}	$[5\ 36]^T$	9.735×10^{-7}	$[10\ 72]^T$	1.217×10^{-7}
	10	$[3\ 66]^T$	1.491×10^{-6}	$[5\ 109]^T$	3.225×10^{-7}	$[10\ 219]^T$	4.035×10^{-8}
3	5	$[3\ 10]^T$	1.545×10^{-5}	$[5\ 16]^T$	3.300×10^{-6}	$[10\ 32]^T$	4.200×10^{-7}
	10	$[3\ 28]^T$	4.785×10^{-6}	$[5\ 47]^T$	1.028×10^{-6}	$[10\ 94]^T$	1.289×10^{-7}
4	5	$[3\ 6]^T$	2.745×10^{-5}	$[5\ 9]^T$	6.165×10^{-6}	$[10\ 19]^T$	7.260×10^{-7}
	10	$[3\ 16]^T$	8.025×10^{-6}	$[5\ 27]^T$	1.725×10^{-6}	$[10\ 55]^T$	2.115×10^{-7}

TABLE II
COMPARISON OF THE RATIO β/α BETWEEN THE OPTIMAL AND SUBOPTIMAL DESIGN VARIABLES

n	N	$c = \hat{\beta}^*/\hat{\alpha}^*$			P_{co}	$\frac{\beta^*}{\alpha^*} = \frac{\ln(P_{co})}{\ln(1-P_{co})}$
		$\alpha_{max} = 3$	$\alpha_{max} = 5$	$\alpha_{max} = 10$		
2	5	7.33	7.20	7.20	0.20000	7.2126
	10	22.00	21.80	21.90	0.10000	21.8543
3	5	3.33	3.20	3.20	0.30623	3.2368
	10	9.33	9.40	9.40	0.17082	9.4339
4	5	2.00	1.80	1.90	0.38937	1.9122
	10	5.33	5.40	5.50	0.23304	5.4898

(9) is satisfied:

$$\frac{\lambda}{1-\lambda}[1-\lambda^{n-1}] - \frac{n-1}{N} = \lambda^{n-1} + \lambda^{n-2} + \cdots + \lambda - \frac{n-1}{N} > 0 \qquad (9)$$

where $\lambda = P_e(1-P_f) + P_f$, and P_f denotes the feedback channel error probability (i.e., probability that an acknowledgment message is corrupted). Similar to [1], we assume that the feedback channel error can only make ACK and NACK messages indistinguishable and the transmitter will handle this erred ACK/NACK message as a NACK. It is important to note that (9) is a general expression which is valid for both noiseless and noisy feedback channels. For instance, the 2-copy transmission mode yields higher throughput than a single copy transmission when $P_e > P_{co} = [1/N - P_f]/[1 - P_f]$. It is apparent that (9) reduces to (8) when $P_f = 0$.

III. NUMERICAL RESULTS

The optimized α and β values for a given $\alpha_{\max}, N,$ and n are depicted in Table I. It is apparent from this table that α always assumes the value of $\alpha_{\max}$ (i.e., suboptimal solution exists on the boundary of the specified region), and the objective function approaches its absolute minimum point as the upper limit for α increases. This trend challenges the conclusion drawn in [1]. Also for each α, there exists an optimum value for β that minimizes the error function.

It is also worthwhile to investigate as to how the suboptimal design parameters will be related to each other, so that we can interpolate the results in broader ranges. Surprisingly, we observed that the relationship between $\hat{\alpha}^*$ and $\hat{\beta}^*$ can be approximated by a linear function, $\hat{\beta}^* = c\hat{\alpha}^*$, where c is a scalar. As an example, Fig. 2 illustrates the linear relationship between α and β for the adaptive protocol with $n = 2$, and varying buffer sizes N. We would like to point out that this insight is not obvious from (3) because we were unable to reduce this analytical expression containing both the variables in terms of their ratio only. As well, while the relationship described in (4) is strictly true only for infinite (extremely large) α and β, our optimization results (see Table II) reveal that this expression is a very good approximation even for finite (small) α and β. In other words, the scalar c is dependent on the throughput crossover probability and can be closely estimated by the ratio $\ln(P_{co})/\ln(1-P_{co})$. This function provides a rule of thumb allowing a handy calculation of β for a given α, or vice versa. Another interesting point to note here is that the ratio β^*/α^* becomes larger as the buffer size increases (correspond to systems with large roundtrip delay), but declines for higher values of n (refer to Table II and Fig. 2).

Comparison of the throughput performance between the $\{\alpha, \beta\}$ values suggested in [1] and our suboptimal solutions are illustrated graphically in Fig. 3(a). To make a fair comparison, we have selected α to be the same for both cases. It is evident from this figure that our optimized $\{\alpha, \beta\}$ pair yields a very close match with the desired performance curve

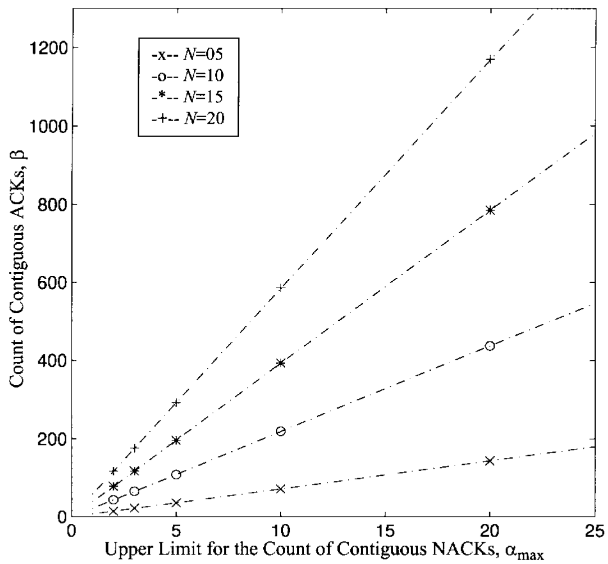
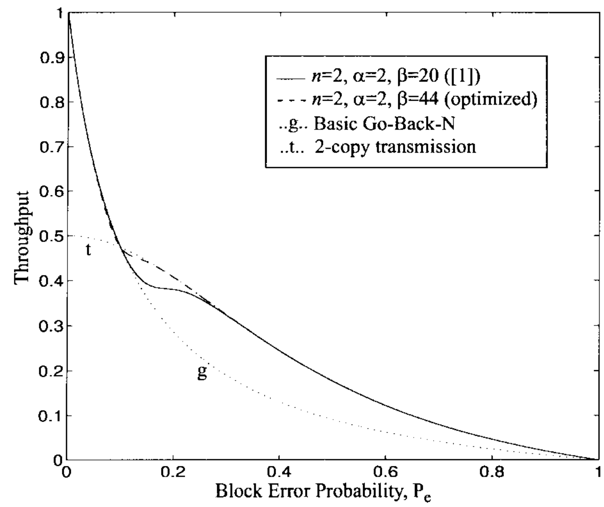


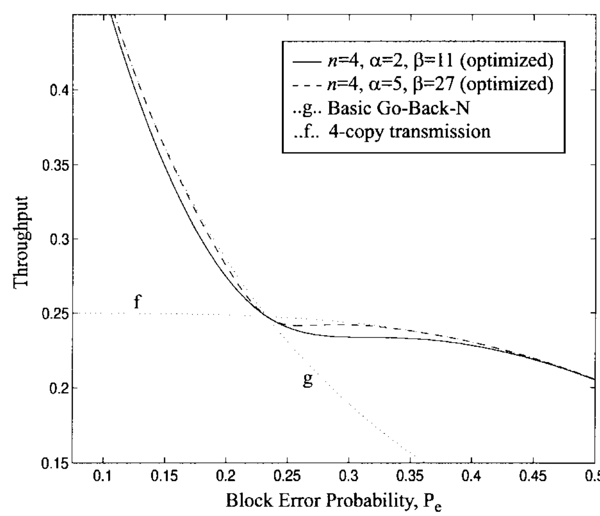
Fig. 2. Relationship between design parameters α and β of the adaptive ARQ protocol with $n = 2$, as a function of buffer size N and $\alpha_{\max}$.

even with a small α, and obviously much better fit than the performance curve with parameters suggested in [1] over a wide range of block error rates. Using the quantitative criteria developed here, the mean square errors for both sets of parameters are given by $E(2, 20) = 1.779 \times 10^{-4}$ and $E(2, 44) = 4.995 \times 10^{-6}$, respectively; whereas in Fig. 3(b), it is shown that a substantial improvement in terms of matching the performance curve of the proposed CSE scheme with the desired performance envelope can be attained by selecting a slightly larger value for α. This observation becomes more pronounced for large n but small N values, as illustrated in Table I. The MSE for each set of the design parameters are $E(2, 11) = 2.655 \times 10^{-5}$ and $E(5, 27) = 1.725 \times 10^{-6}$.

It should be noted that the selection of α and β is mainly determined by the channel fluctuation rate, and the switching reliability criterion (i.e., the MSE value) will be used as a secondary metric. In a rapidly changing environment, both the design parameters should be selected as small as possible (however, one should also remember that for a given α, there always exists a value for β that minimizes the error function). In this situation, the numerical value of $E(\alpha, \beta)$ may be used only to quantify the CSE algorithm switching reliability. On the other hand, in a modest or in an extremely slowly varying channel, both α and β can be selected to be quite large. In this case, the MSE function may allow us to choose the design parameters based on a specified quality criterion. For instance, in Fig. 3(a) we have shown that by using a moderate value for α and its corresponding β, the switching reliability can be quite good. Increasing the value of α beyond a certain value (say, $\alpha = 3$) will result only in a slight improvement of the quality criterion. Consequently, choosing a very large value for these variables is not very desirable, specifically when the channel variation rate is not too slow.



(a)



(b)

Fig. 3. Comparison of the throughput performance between different sets of design parameters. Buffer size N is assumed to be 10.

ACKNOWLEDGMENT

The authors would like to thank the anonymous reviewers for their helpful comments and constructive suggestions.

REFERENCES

[1] Y.-D. Yao, "An effective Go-Back-N ARQ scheme for variable-error-rate channels," *IEEE Trans. Commun.*, vol. 43, pp. 20–23, Jan. 1995.
[2] R. Fletcher, *Practical Methods of Optimization, Volume 1—Unconstrained Optimization*. New York: Wiley, 1980.