

研究速報

マルコフ決定過程の動作時間と受信バッファ容量が有限の選択再送 ARQ への適用

前田 康成^{†a)} (正員) 雨宮 康二^{††}
 小林 直人^{†††} (正員) 吉田 秀樹[†]
 鈴木 正清[†] (正員) 松嶋 敏泰^{†††} (正員)

Applying Markov Decision Processes to Selective-Repeat ARQ with Finite Operation Time and Finite Receiver Buffer

Yasunari MAEDA^{†a)}, *Member*,
 Koji AMEMIYA^{††}, *Nonmember*,
 Naoto KOBAYASHI^{†††}, *Member*,
 Hideki YOSHIDA[†], *Nonmember*, Masakiyo SUZUKI[†],
 and Toshiyasu MATSUSHIMA^{†††}, *Members*

[†] 北見工業大学情報システム工学科, 北見市

Dept. of Computer Sciences, Kitami Institute of Technology, 165 Koen-cho, Kitami-shi, 090-8507 Japan.

^{††} 早稲田大学大学院理工学研究科, 東京都

Waseda University, 3-4-1 Okubo, Shinjuku-ku, Tokyo, 169-8555 Japan.

^{†††} 東京工科大学メディア学部, 八王子市

School of Media Science, Tokyo University of Technology, 1404-1, Katakura-cho, Hachioji-shi, 192-0914 Japan.

^{††††} 早稲田大学応用数学科, 東京都

Department of Applied Mathematics, Waseda University, Tokyo, 169-8555 Japan.

a) E-mail: maeda@cs.kitami-it.ac.jp

あらまし 本研究では, メッセージブロックに誤りを生じるブロック誤り率が既知で, 動作時間と受信バッファ容量が有限の選択再送 ARQ を研究対象とする. マルコフ決定過程を用いて定式化し, スループットの期待値を最大化する方法を導出し, その理論的境界を計算する.

キーワード 選択再送 ARQ, マルコフ決定過程, スループット, 動的計画法

1. まえがき

ブロック誤り率が既知で, 受信バッファ容量が有限の選択再送 ARQ (自動再送要求方式) に関する研究は数多くある [1]~[7]. しかし, 今日まで受信バッファ容量が有限の選択再送 ARQ のスループットを最大化する選択再送 ARQ 方式は導出されていない.

そこで, 本研究ではスループットの最大化をマルコフ決定過程 (Markov decision processes, 以下 MDP と略す) [8] の期待総利得の最大化問題として定式化する. 次にその最大化問題を動的計画法 (Dynamic Programming, 以下 DP と略す) で解く. これにより, ブロック誤り率が定常かつ既知, 動作時間と受信バッファ容量が有限の選択再送 ARQ において, スループットの期待値を最大化する点で最適な選択再送

ARQ 方式を MDP の最適な決定関数として導出する.

しかし, DP で最大化問題を解く際の計算量が膨大なため, 最適な選択再送 ARQ 方式をそのまま実用化することは考えていない. この方式の用途として, 例えば以下の 2 点がある. 1 点目は, 求められた最適な選択再送 ARQ 方式の特徴的なメッセージブロックの選択例をルール化して, 既存の選択再送 ARQ 方式と融合することによる既存方式の改善である. 2 点目は, 最大化問題を DP で解く過程の計算結果を利用してスループットの理論的境界を計算し, その理論的境界を既存選択再送 ARQ 方式の性能評価に応用することである. 本研究では後者の具体例を示し考察する.

前述のとおり受信バッファ容量が有限のスループットを最大化する方式は今日まで導出されていないが, そのスループットの理論的境界もまたまだ計算されていない. そのため, 従来研究では提案方式のスループットを, 受信バッファ容量が無限の理想的な選択再送 (Ideal Selective-Repeat) ARQ 方式 (以下 ISR と略す) のスループットと比較している. しかし, 受信バッファ容量が有限と無限では大きく条件が異なり, 厳密な評価はできない. そこで, 本研究で計算する理論的境界を既存方式の評価に応用し, より厳密な評価を可能にする.

以下, 2. で選択再送 ARQ の概要, ISR のスループット, MDP の概要について述べ, 3. で MDP と選択再送 ARQ の対応の説明, スループットを最大化する選択再送 ARQ 方式の導出, スループットの理論的境界の計算を行い, 4. でスループットの理論的境界の計算結果を既存方式の評価に応用する例を示し, 5. で全体のまとめについて述べる.

2. 準備

2.1 選択再送 ARQ の概要

本研究では動作時間が有限時間 T の選択再送 ARQ (再送による重複も含めて延べ T 個のメッセージブロックを送信する) を対象とする. 受信者に伝えたい順番を示す番号が添字として付与されたメッセージブロックを m_i , $m_i \in M$, $i \geq 1$ と表記する. 送信者は各時点においてメッセージブロックを 1 個送信する. M はメッセージブロックの集合である. 再送によって同じ番号のメッセージブロックが複数回送信されることがあるので, 動作時間 T 内に実際に送信されるメッセージブロック番号の最大値は T 以下である.

受信者は, 受信したメッセージブロックにブロック誤りがなければ ACK 信号を, ブロック誤りがあれば

NAK 信号を確認信号として送信者に送る。受信者は正しく受信できたメッセージブロックを上位層に送る。上位層とは、選択再送 ARQ 方式を介して受け取った情報（メッセージブロック）の受渡し先を示す。具体的にはユーザや情報を利用するシステムなどである。ただし、受信者には番号順に上位層にメッセージブロックを送るという制約がある。受信者は、正しく受信してもまだ上位層に送ることができないメッセージブロックを、一時的に受信バッファに格納する。C は受信バッファ容量で、本研究では既知で有限かつ $C \geq 1$ である。受信バッファ容量が有限なので、受信バッファが満杯で格納できない場合（オーバフロー）がある。なお、オーバフローが起きた場合には、正しく受信していても特別に NAK を送信者に送る。

D は往復伝搬遅延を示し、送信者があるメッセージブロックを送信してから、そのメッセージブロックに対する確認信号を受信するまでに要する時間である。つまり、送信者は、時点 t に送信したメッセージブロックに対する確認信号を時点 t + D に受け取る。 $T > D \geq 2$ である。

図 1 に選択再送 ARQ の動作例を示す。時点 1 に送信者がメッセージブロック m_1 を送信し、 m_1 は受信者に正しく届き上位層に送られている。受信者は送信者に ACK の確認信号を送り、送信者はこの確認信号を時点 4 に受け取っている。時点 2 に送信者が m_2 を送信し、 m_2 は受信者に間違って届いている。受信者は送信者に NAK の確認信号を送り、送信者はこの確認信号を時点 5 に受け取って m_2 を再送している。以下、時点 3 及び時点 4 に送られた m_3 及び m_4 は正しく届くが、 m_2 がまだ正しく届いていないので、 m_3 及び m_4 は受信バッファに格納される。時点 7 に送られた m_6 は正しく届くがオーバフローになり、受信者は

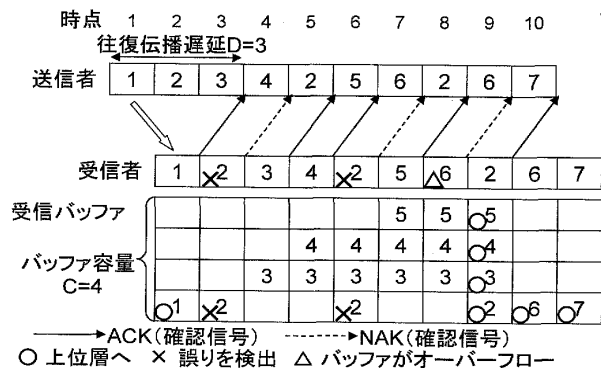


図 1 選択再送 ARQ の動作例

Fig. 1 Transmitting and receiving example of SR-ARQ.

NAK を送っている。

2.2 ISR のスループット

ISR のスループットは、受信バッファ容量が無限という条件下における理論的境界である。選択再送 ARQ のスループットは式 (1) で定義され、無限の動作時間を仮定した場合の ISR のスループットは式 (2) で計算できることが知られている。

$$TH = \frac{\text{上位層へ送られたメッセージブロック数}}{\text{送信したメッセージブロック数}} \quad (1)$$

$$TH_{ISR} = 1 - (\text{ブロック誤り率}) \quad (2)$$

式 (2) が、受信バッファ容量が無限という条件下におけるスループットの理論的境界である。

2.3 MDP の概要

MDP は、状態 s_i , $s_i \in S$, $S = \{s_1, s_2, \dots, s_{|S|}\}$ ($|S|$ は有限), 各状態で選べる行動 a_i , $a_i \in A$, $A = \{a_1, a_2, \dots, a_{|A|}\}$ ($|A|$ は有限), 状態 s_i で行動 a_j を選んだもとで状態 s_k へ遷移する遷移確率 $p(s_k | s_i, a_j, \xi^*)$ (ξ^* は遷移確率分布を支配する真のパラメータで既知), 遷移に伴い発生する利得 $r(s_i, a_j, s_k)$ で構成される。図 2 に状態数と行動数が 2 の MDP の例を示す。MDP の目的は、行動を選び、状態が遷移し、利得を得るという一連のプロセスを繰り返しながら総利得を最大化することである。プロセスの繰り返し回数が有限の場合には、総利得の期待値（期待総利得）を最大化する最適な決定関数 $d^*(\cdot, \cdot)$ を DP によって求めることができる。具体的には式 (3) を用いて、時点 t の状態が s_i という条件下における時点 t + 1 以降の期待総利得の最大値 $V(s_i, t)$ を逐次的に計算する。決定関数は状態と時点を受け取って、その時点で選ぶべき行動を返す関数である。

$$V(s_i, t) = \max_{a_j \in A} \sum_{s_k \in S} p(s_k | s_i, a_j, \xi^*) (r(s_i, a_j, s_k) + V(s_k, t + 1)) \quad (3)$$

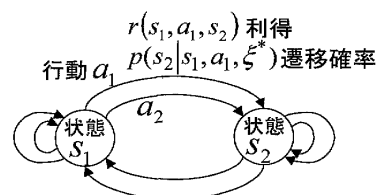


図 2 MDP の例

Fig. 2 An example of MDP.

3. MDP の選択再送 ARQ への適用

3.1 MDP と選択再送 ARQ の対応

最初にいくつかの定義を行う。\$w_t\$ は時点 \$t\$ の最初に送信者が受け取った確認信号 (ACK または NAK) を示す。\$y_t\$ は時点 \$t\$ で送信者が送信したメッセージブロック番号を示す。\$z_t\$ は、上位層に送られたことを時点 \$t\$ でまだ送信者が確認していない (時点 \$t\$ で送信者がまだ上位層に送られていないと認識している) メッセージブロック番号のうちで、最小の番号を示し、\$y_t \ge z_t\$ である。本研究では、受信ブロック中の誤りは間違いなく検出でき、かつ確認信号を送信する帰還通信路には誤りは生じないと仮定する。よって、ブロック誤り率と NAK 信号が返送される確率は同じである。そこで、ブロック誤り率を \$p(\text{NAK} | \theta^*)\$ と表記する。本来は確率関数の表記法に従って \$Pr\{w_t = \text{NAK} | \theta^*\}\$ と表記すべきだが、本研究では確率関数を略記する。\$\theta^*\$, \$\theta^* \in \Theta\$ はブロック誤りが生じる確率分布の真のパラメータで既知である。例えば、ビット誤り率が 0.001 の 2 元対称通信路と \$n\$ ビットのブロックの場合、ブロック誤りを生じる確率分布は \$\theta^* = 1 - (1 - 0.001)^n\$ の二項分布である。

MDP における時点 \$t\$ の状態 \$x_t\$ は、時点 \$t\$ に送信者が把握している受信バッファの状況と、時点 \$t-D+1\$ から時点 \$t-1\$ に送信したメッセージブロック番号に関する情報から構成され、次式のようになる。

$$x_t = (x_{t,0}, x_{t,1}, \dots, x_{t,C-1}, y'_{t,t-D+1}, \dots, y'_{t,t-1}), \quad (4)$$

ただし、\$x_{t,i}\$ は受信バッファの状況を示し、時点 \$t\$ においてメッセージブロック \$m_{z_t+i}\$ が受信バッファに格納されていることを送信者が確認しているかどうかによって、以下のような値をとる。

$$x_{t,i} = \begin{cases} 1, & \text{時点 } t \text{ において格納を確認済;} \\ 0, & \text{時点 } t \text{ において格納を未確認.} \end{cases} \quad (5)$$

また、\$y'_{t,i}\$ は \$z_t\$ を基準値としてメッセージブロック番号 \$y_i\$ を次式によって変換した数値である。

$$y'_{t,i} = \begin{cases} -2, & i < 1 \text{ または } i > T; \\ -1, & \text{当該メッセージブロック} \\ & \text{が既に上位層に送られて} \\ & \text{いる, またはバッファに} \\ & \text{格納されていることを時} \\ & \text{点 } t \text{ に送信者が確認済;} \\ y_i - z_t, & \text{それ以外,} \end{cases} \quad (6)$$

ただし、最大限効率的に送信できた場合 (ACK のみ発生した場合) を想定すると、バッファ容量 \$C\$ と往復伝搬遅延 \$D\$ の関係から送信可能な最大のメッセージブロック番号が決まり、\$y'_{t,i} \in \{-2, -1, 0, 1, \dots, C+D-2\}\$ である。よって、選択再送 ARQ に適用した場合の MDP の状態集合 \$S\$ について \$x_t \in S\$, \$|S| = 2^C (C+D+1)^{D-1}\$ である。

MDP における時点 \$t\$ の行動 \$y'_{t,t}\$ は、時点 \$t\$ に送信者が送信するメッセージブロック番号 \$y_t\$ を式 (6) で変換した数値である。\$y'_{t,t} \in \{-2, -1, 0, 1, \dots, C+D-2\}\$ であり、集合 \$\{-2, -1, 0, 1, \dots, C+D-2\}\$ が MDP の行動集合 \$A\$ に相当する。

時点 \$t\$ で送信者が、送信するメッセージブロック番号 \$y_t\$ を決定 (MDP で行動 \$y'_{t,t}\$ を選ぶことに相当) してメッセージブロック \$m_{y_t}\$ を送信すると、時点が \$t+1\$ に進む。次に送信者は時点 \$t+1\$ の確認信号 \$w_{t+1}\$ を受け取る。\$w_{t+1}\$ は時点 \$t-D+1\$ に送信したメッセージブロック \$m_{y_{t-D+1}}\$ に対する確認信号である。確認信号 \$w_{t+1}\$ を受け取ることによって、上位層に送られたことを時点 \$t+1\$ においてまだ送信者が確認していないメッセージブロック番号のうちで最小の番号 \$z_{t+1}\$ と、時点 \$t+1\$ の状態 \$x_{t+1}\$ が決まる。

本研究で扱う MDP ではブロック誤り率が状態遷移確率に相当する。状態 \$x_t\$ から状態 \$x_{t+1}\$ へ遷移する際の MDP の遷移確率は \$p(x_{t+1} | x_t, y'_{t,t}, \xi^*) = p(w_{t+1} | \theta^*)\$ である。なお、オーバフローの場合には特別に確率 1 で確定的に NAK が生じるが、オーバフロー後の状態は ACK でも NAK でも同一の状態なので、正しく届いたのに NAK が送信者に送られても MDP における行動の決定には影響を与えない。

本研究では動作時間 \$T\$ は固定なので、式 (1) のスループットの値は分子の上位層に送られたメッセージブロック数で決まる。そこで、時点 \$t\$ の状態 \$x_t\$ においてメッセージブロック \$m_{y_t}\$ を送信 (行動 \$y'_{t,t}\$ を選ぶ) し、時点 \$t+1\$ に確認信号 \$w_{t+1}\$ を受け取って状態 \$x_{t+1}\$ に遷移した際に、送信者が新たに上位層に送られたことを確認するメッセージブロック数を MDP の利得 \$r(x_t, y'_{t,t}, x_{t+1})\$ として次式で定義する。

$$r(x_t, y'_{t,t}, x_{t+1}) = \begin{cases} \sum_{i=0}^{C-1} x_{t,i} - \sum_{i=0}^{C-1} x_{t+1,i+1}, & y'_{t,t-D+1} = 0 \\ & \text{かつ } w_{t+1} = \text{ACK}; \\ 0, & \text{その他の場合,} \end{cases} \quad (7)$$

ただし、右辺の $y'_{t,t-D+1} = 0$ かつ $w_{t+1} = ACK$ という条件は、まだ上位層に送られたことを確認していない最小番号のメッセージブロック m_{z_t} ($m_{y_{t-D+1}}$ と同じ) が上位層に送られた場合を示す。利得が発生するのは、この条件が成立する場合のみである。式 (7) では受信バッファ状況の差分を用いて、新たに上位層に送られたメッセージブロック数を計算する。

時点 $t+1$ の最初に確認信号 w_{t+1} を受け取ると、 w_{t+1} と時点 t の状態 x_t と z_t に基づいて、時点 $t+1$ の x_{t+1} と z_{t+1} が決まる (各変数が更新される)。以下で各変数の更新方法を説明する。本研究では、まだ何も送信していない状態を初期状態 x_1 として想定している。初期値は z_1 が $z_1 = 1$ 、 $x_{1,i}$ が $x_{1,i} = 0$ 、 $0 \leq i \leq C-1$ 、 $y'_{1,i}$ が $y'_{1,i} = -2$ 、 $t-D+1 \leq i \leq t-1$ である。最初に $y'_{t,t-D+1} = 0$ ($y_{t-D+1} = z_t$ と同じ) かつ $w_{t+1} = ACK$ の場合を説明する。この場合は、上位層に送られたことをまだ確認していない最小番号のメッセージブロック m_{z_t} ($m_{y_{t-D+1}}$ と同じ) が上位層に送られた場合である。 z_{t+1} は次式で決まる。

$$z_{t+1} = z_t + b, \quad (8)$$

ただし、 $x_{t,i}$ 、 $1 \leq i \leq C-1$ の中に $x_{t,i} = 0$ が成立する i が存在する場合には、成立する最小の i を b とし、成立する i が存在しない場合には、 $b = C$ である。 $x_{t+1,i}$ 、 $0 \leq i \leq C-1$ は次式で決まる。

$$x_{t+1,i} = \begin{cases} x_{t,z_{t+1}-z_t+i}, & z_{t+1} - z_t + i < C; \\ 0, & \text{その他の場合.} \end{cases} \quad (9)$$

$y'_{t+1,i}$ 、 $t-D+2 \leq i \leq t$ は次式で決まる。

$$y'_{t+1,i} = \begin{cases} -2, & y'_{t,i} = -2; \\ -1, & y'_{t,i} = -1 \text{ または} \\ & 0 \leq y'_{t,i} < z_{t+1} - z_t; \\ y'_{t,i} - (z_{t+1} - z_t), & \text{その他の場合.} \end{cases} \quad (10)$$

次に $0 < y'_{t,t-D+1} < C$ かつ $w_{t+1} = ACK$ の場合を説明する。この場合は、受信バッファ内に新たなメッセージブロックが格納されたか、または格納済のメッセージブロックの ACK 信号を再度受け取った場合である。 $z_{t+1} = z_t$ であり、 $x_{t+1,i}$ 、 $0 \leq i \leq C-1$ は次式で決まる。

$$x_{t+1,i} = \begin{cases} 1, & i = y'_{t,t-D+1}; \\ x_{t,i}, & \text{その他の場合.} \end{cases} \quad (11)$$

$y'_{t+1,i}$ 、 $t-D+2 \leq i \leq t$ は次式で決まる。

$$y'_{t+1,i} = \begin{cases} -1, & y'_{t,i} = y'_{t,t-D+1}; \\ y'_{t,i}, & \text{その他の場合.} \end{cases} \quad (12)$$

上記以外の場合には、 $z_{t+1} = z_t$ であり、 $x_{t+1,i}$ 、 $0 \leq i \leq C-1$ は $x_{t+1,i} = x_{t,i}$ 、 $y'_{t+1,i}$ 、 $t-D+2 \leq i \leq t$ は $y'_{t+1,i} = y'_{t,i}$ である。

ここで図 1 の時点 1 から時点 4 を例に、各変数の具体例を示す。時点 1 ではまだ何も送信されていない初期状態 $x_1 = (0, 0, 0, 0, -2, -2)$ から始まり、上位層にもまだ何も送られていないので $z_1 = 1$ で、メッセージブロック m_{y_1} 、 $y_1 = 1$ を送信している。往復伝搬遅延 $D = 3$ なので、確認信号や利得 (上位層に送られたメッセージブロック数の送信者による確認) は時点 4 から発生する。時点 2 では $z_2 = 1$ で状態 $x_2 = (0, 0, 0, 0, -2, 0)$ に遷移し、 m_{y_2} 、 $y_2 = 2$ を送信している。時点 3 では $z_3 = 1$ で状態 $x_3 = (0, 0, 0, 0, 0, 1)$ に遷移し、 m_{y_3} 、 $y_3 = 3$ を送信している。時点 3 までは確認信号を受け取っていないので、状態遷移は送信したメッセージブロック番号に依存した確定的な遷移である。時点 4 では確認信号 $w_4 = ACK$ を受け取り、 m_{y_1} 、 $y_1 = 1$ が上位層に送られたことを送信者が確認するので $z_4 = 2$ となり、状態 $x_4 = (0, 0, 0, 0, 0, 1)$ に遷移し、利得 $r(x_3, 2, x_4) = 1$ が発生し、 m_{y_4} 、 $y_4 = 4$ を送信している。状態 x_4 への遷移は確認信号に依存しているので状態遷移確率による遷移である。

3.2 スループット最大化の MDP による定式化

本研究で対象とする動作時間 T の選択再送 ARQ を MDP の問題として定式化すると、往復伝搬遅延との関係で $T + D$ 期間の MDP の問題になる。時点 1 から時点 D の間には送信者は、メッセージブロックを送信するが確認信号は受け取らないので、メッセージブロックが上位層に送られたことを確認することはない。時点 $D + 1$ から時点 T の間には送信者は、確認信号を受け取って上位層に送られたメッセージブロック数を確認し、メッセージブロックを送信する。時点 $T + 1$ から時点 $T + D$ の間には送信者は、確認信号を受け取って上位層に送られたメッセージブロック数を確認するが、メッセージブロックは送信しない。以下、統計的決定理論 [9] に基づいて定式化を進める。

ブロック誤り率の真のパラメータが θ^* 、行動 (送信するメッセージブロック) を選ぶ決定関数が $d(\cdot, \cdot)$ という条件のもとで、状態が状態系列 $x^{T+D} = x_1 x_2 \cdots x_{T+D}$ のように遷移し、確認信号を確認信号系列 $w^T =$

$w_{D+1}w_{D+2}\cdots w_{T+D}$ のように受け取った場合の MDP における $T+D$ 期間の総利得を示す効用関数を次式のように定義する.

$$U(x^{T+D}, d(\cdot, \cdot), w^T, \theta^*) \quad (13)$$

$$= \sum_{i=D}^T r(x_i, y'_{i,i}, x_{i+1}) + \sum_{i=T+1}^{T+D-1} r'(x_i, w_{i+1}, x_{i+1}),$$

ただし, $d(\cdot, \cdot)$ は $d(x_t, t) = y'_{t,t}$ のように状態と時点を受け取って, 行動を選ぶ決定関数を示す. また, 時点 $T+1$ から時点 $T+D$ の間には行動を選ばないので, 利得関数の引数の一部を確認信号に代えた表記 $r'(x_i, w_{i+1}, x_{i+1})$ にしている. 効用は T 個のメッセージブロックを送信した際に上位層に送られたメッセージブロック数に相当し, $U(x^{T+D}, d(\cdot, \cdot), w^T, \theta^*) \leq T$ である.

期待効用関数は次式のようにになる.

$$EU(d(\cdot, \cdot), \theta^*) = \sum_{x^{T+D-1}, w^T} p(x^{T+D-1}, w^T | x_1, d(\cdot, \cdot), \theta^*)$$

$$U(x^{T+D}, d(\cdot, \cdot), w^T, \theta^*), \quad (14)$$

ただし, x_1 は初期状態 (時点 1 の状態) で既知とする. 式 (14) では, すべての状態系列及び確認信号系列について効用関数の期待値を計算している. これはある決定関数 $d(\cdot, \cdot)$ を用いた場合の MDP における $T+D$ 期間の期待総利得である. また, 期待総利得は, メッセージブロックのある選択方法 $d(\cdot, \cdot)$ を用いて T 個のメッセージブロックを送信した際に上位層へ送られるメッセージブロック数の期待値に相当する.

式 (14) の最大値が上位層へ送られるメッセージブロック数の期待値の最大値で, その最大値を T で割った値がスループットの期待値の最大値である. スループットの期待値を最大化する点で最適な選択再送 ARQ 方式が, 最適な決定関数として次式で定義される.

$$d^*(\cdot, \cdot) = \arg \max_{d(\cdot, \cdot)} EU(d(\cdot, \cdot), \theta^*). \quad (15)$$

式 (15) で定義された最適な選択再送 ARQ 方式である最適な決定関数を, 次節で DP を用いて求める.

3.3 最適な選択再送 ARQ 方式の導出

有限期間の MDP の期待総利得を最大化する最適な決定関数は DP で求められる. 式 (15) で定義された最適な選択再送 ARQ 方式である最適な決定関数も同様に DP で求められる. しかし, その際の DP の計算量が膨大なので, 本研究で求める最適な選択再送 ARQ

方式をそのまま実用化することは考えていない. また, DP で $T+D$ 期間の MDP を解くのは選択再送 ARQ の動作前である. つまり, 実際に送信を行う前に, 想定されるすべての状態について送信すべきメッセージブロックを求めておくことになる. よって, 選択再送 ARQ の動作時間中に随時メッセージブロックの選択を行う従来方式とは多少概念が異なる. DP では時点 $T+D-1$ から時点 1 までさかのぼりながら解く. 次に DP における各時点ごとの処理について説明する.

時点 $T+D-1$ では各状態 x_{T+D-1} について, 式 (16) で状態が x_{T+D-1} という条件のもとで, 時点 $T+D$ に得る利得の期待値 $V'(x_{T+D-1}, T+D-1)$ を計算する. なお, 時点 $T+D-1$ では行動を選ばない (メッセージブロックを送信しない) ので, 利得の期待値計算のみ行う.

$$V'(x_{T+D-1}, T+D-1)$$

$$= \sum_{w_{T+D}} p(w_{T+D} | \theta^*) r'(x_{T+D-1}, w_{T+D}, x_{T+D}). \quad (16)$$

時点 t_1 , $T+1 \leq t_1 \leq T+D-2$ では各状態 x_{t_1} について, 式 (17) で状態が x_{t_1} という条件のもとで, 時点 t_1+1 以降に得る期待総利得 $V'(x_{t_1}, t_1)$ を計算する.

$$V'(x_{t_1}, t_1) = \sum_{w_{t_1+1}} p(w_{t_1+1} | \theta^*)$$

$$\{r'(x_{t_1}, w_{t_1+1}, x_{t_1+1}) + V'(x_{t_1+1}, t_1+1)\}. \quad (17)$$

時点 T では各状態 x_T について, 式 (18) で状態が x_T という条件のもとで, 時点 $T+1$ 以降に得る期待総利得の最大値 $V(x_T, T)$ を計算する. そして, 式 (18) の右辺を最大にする行動を最適な決定 $d^*(x_T, T)$ として選ぶ.

$$V(x_T, T) = \max_{y'_{T,T}} \sum_{w_{T+1}} p(w_{T+1} | \theta^*)$$

$$\{r(x_T, y'_{T,T}, x_{T+1}) + V'(x_{T+1}, T+1)\}. \quad (18)$$

時点 t_2 , $D \leq t_2 < T$ では各状態 x_{t_2} について, 式 (19) で状態が x_{t_2} という条件のもとで, 時点 t_2+1 以降に得る期待総利得の最大値 $V(x_{t_2}, t_2)$ を計算する. そして, 式 (19) の右辺を最大にする行動を最適な決定 $d^*(x_{t_2}, t_2)$ として選ぶ.

$$V(x_{t_2}, t_2) = \max_{y'_{t_2, t_2}} \sum_{w_{t_2+1}} p(w_{t_2+1} | \theta^*)$$

$$\{r(x_{t_2}, y'_{t_2, t_2}, x_{t_2+1}) + V(x_{t_2+1}, t_2 + 1)\}. \quad (19)$$

時点 $D - 1$ では各状態 x_{D-1} について、式 (20) で状態が x_{D-1} という条件のもとで、時点 $D + 1$ 以降に得る期待総利得の最大値 $V''(x_{D-1}, D - 1)$ を計算する。そして、式 (20) の右辺を最大にする行動を最適な決定 $d^*(x_{D-1}, D - 1)$ として選ぶ。

$$V''(x_{D-1}, D - 1) = \max_{y'_{D-1, D-1}} V(x_D, D). \quad (20)$$

なお、時点 1 から時点 D の間には利得が発生しないので、時点 $D + 1$ 以降の総利得を計算している。

時点 t_3 , $1 \leq t_3 \leq D - 2$ では各状態 x_{t_3} について、式 (21) で状態が x_{t_3} という条件のもとで、時点 $D + 1$ 以降に得る期待総利得の最大値 $V''(x_{t_3}, t_3)$ を計算する。そして、式 (21) の右辺を最大にする行動を最適な決定 $d^*(x_{t_3}, t_3)$ として選ぶ。

$$V''(x_{t_3}, t_3) = \max_{y'_{t_3, t_3}} V''(x_{t_3+1}, t_3 + 1). \quad (21)$$

このように、DP を用いて時点 $T + D - 1$ から時点 1 までさかのぼって処理することによって、式 (15) で定義された期待総利得を最大化する決定関数が求まる。この決定関数が、ブロック誤り率が既知、動作時間と受信バッファ容量が有限の選択再送 ARQ において、スループットの期待値を最大化する選択再送 ARQ 方式である。

3.4 スループットの理論的境界の計算

ここでは、3.3 の DP を用いて最適な選択再送 ARQ 方式を求める過程の計算結果を利用して、スループットの理論的境界を計算する。

上位層へ送られるメッセージブロック数の期待値の最大値が式 (21) によって、 $V''(x_1, 1)$, $V''(x_1, 1) \leq T$ として計算される。 $V''(x_1, 1)$ を利用すると、次式によってスループットの理論的境界を計算できる。

$$TH_{Proposed} = \frac{V''(x_1, 1)}{T}. \quad (22)$$

なお、往復伝搬遅延 D との関係で $T + D$ 期間の MDP を解くが、選択再送 ARQ の動作時間や実際に送信するメッセージブロックの延べ数は T なので、式 (22) の分母は T である。

4. スループットの理論的境界の既存選択再送 ARQ 方式の評価への応用

ここでは、受信バッファ容量が有限の既存選択再送 ARQ 方式を評価する際の、3.4 の式 (22) で計算され

る受信バッファ容量が有限という条件下におけるスループットの理論的境界の有効性をシミュレーションによって検証する。具体的には従来研究で提案されている連続多重送信 (Continuous Multiple Transmissions) 選択再送 ARQ 方式 (以下 CMT と略す) [3] のスループットの評価に応用する。

往復伝搬遅延 $D = 4$, 動作時間 $T = 1000$ の場合について、ブロック誤り率を 0.001, 0.01, 0.1 から 0.9 まで変化させて、式 (22) で計算した理論的境界 (図中では Derived と表示) と、CMT のシミュレーションを 1000 回行った平均のスループット (図中では CMT と表示) とを比較した。参考のため、受信バッファ容量が無限の ISR のスループットも図中に ISR として表示している。CMT については事前に予備シミュレーションを行って、CMT で利用する各種パラメータの設定を行った。受信バッファ容量が $C = 4$ の結果を図 3 に示し、 $C = 8$ の結果を図 4 に示す。

図 3 及び図 4 から、CMT はブロック誤り率が低い場合には Derived を達成し、ブロック誤り率が高い

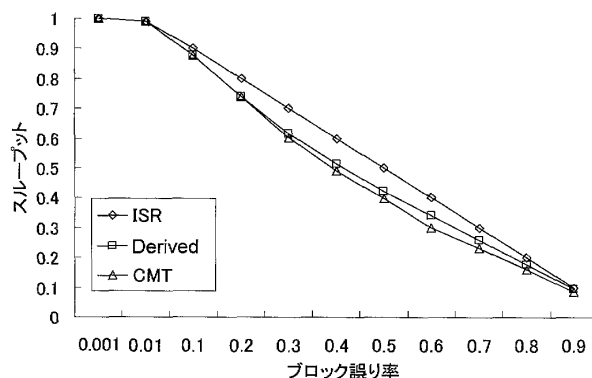


図 3 従来研究と理論的境界の比較 ($C = 4$)
Fig. 3 Comparison of theoretical limits with previous method ($C = 4$).

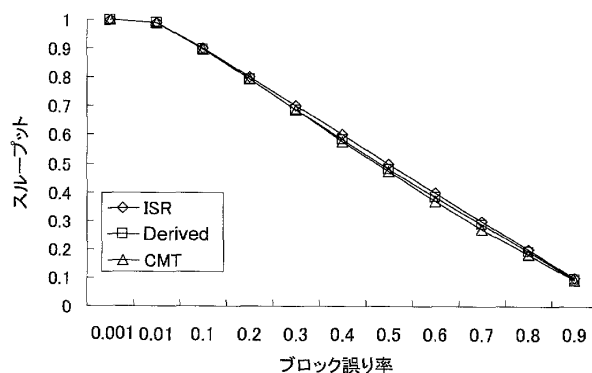


図 4 従来研究と理論的境界の比較 ($C = 8$)
Fig. 4 Comparison of theoretical limits with previous method ($C = 8$).

場合には Derived よりも少し低くなっている。また、ISR が一番高いが、ISR と CMT の間に Derived がある。よって、受信バッファ容量が有限の CMT の評価では、受信バッファ容量が無限の理論的境界である ISR と比較するのではなく、受信バッファ容量が有限の理論的境界である Derived と比較することによって、どれだけ理論的境界に近いかという点でより厳密な評価が可能になる。

なお、本研究では有限の動作時間を仮定したが、CMT は無限の動作時間を仮定しているため、本シミュレーション結果から単純に性能の判断はできない。しかし、本論文には掲載していない $T = 10000$ の場合でも同様の結果が得られているので、動作時間をこれ以上長くしても結果は大きく変わらないと思われる。他方、動作時間を短くした場合については、 $T = 100$ で確認したところ Derived と CMT の差が $T = 1000$ の場合よりも大きくなった。このことから、 $T = 100$ という動作時間は、無限の動作時間を仮定した CMT の性能を十分に引き出すには短すぎる事が分かる。

5. 考察と今後の課題

5.1 MDP の適用と最適な選択再送 ARQ 方式について

本研究ではブロック誤り率が定常かつ既知、動作時間と受信バッファ容量が有限の選択再送 ARQ を対象とした。受信バッファ容量が有限の選択再送 ARQ は従来から数多く研究されているが、今日まで受信バッファ容量が有限の選択再送 ARQ のスループットの最大化する選択再送 ARQ 方式は導出されていない。

そこで、本研究では最初に選択再送 ARQ のスループットの最大化が MDP の期待総利得の最大化問題として定式化できることを示した。次にその最大化問題を DP で解き、スループットの期待値を最大化するという意味で最適な選択再送 ARQ 方式を導出した。

しかし、最適な選択再送 ARQ 方式を求める際の DP の計算量は膨大である。MDP の状態数が $2^C(C+D+1)^{D-1}$ なので、式 (19) などの処理が最大で $2^C(C+D+1)^{D-1}(T+D-1)$ 回必要である。また、実用に際しては事前に求めた最適な決定関数(具体的には各条件下で選択すべきメッセージブロックを書き出したリスト)の記憶用のメモリ容量も計算量と同様に膨大になる。そのため、本研究で導出した最適な選択再送 ARQ 方式をそのまま実用化することは考えていない。今後の課題には、1. で本方式の用途例の 1 点目として述べた既存方式の改善が挙げられ

る。上記のリストから特徴的なメッセージブロックの選択例をルール化して、CMT などの既存方式と融合することによって、既存方式の改善が可能である。また、より現実に近い通信路であるブロック誤り率が非定常な通信路の検討も今後の課題としたい。

5.2 スループットの理論的境界について

5.1 で述べたように、本研究で導出した最適な選択再送 ARQ 方式をそのまま実用化することは考えていない。しかし、3.4 で述べたように、最適な選択再送 ARQ 方式を DP を用いて求める過程の計算結果を利用すると、受信バッファ容量が有限のスループットの理論的境界を計算できる。式 (22) で計算できる理論的境界の有効性をシミュレーションによって確認した。CMT のような受信バッファ容量が有限の選択再送 ARQ 方式の評価を行う際には、式 (22) で計算できるスループットの理論的境界と比較することによって、どれだけ境界に近いかという点でより厳密な評価が可能になった。

謝辞 本研究の一部は日本学術振興会科学研究費新学術領域 (No.20200044) 及び早稲田大学特定課題研究 (No.2009A-058) の援助による。

文 献

- [1] G. Benelli, "A selective ARQ protocol with a finite-length buffer," *IEEE Trans. Commun.*, vol.41, no.7, pp.1102-1111, July 1993.
- [2] Y. Chang and C. Leung, "On Weldon's ARQ strategy," *IEEE Trans. Commun.*, vol.COM-32, no.3, pp.297-300, March 1984.
- [3] 楠堂忠夫, 岡 育生, 藤原値賀人, "連続多重送信選択再送方式に関する考察," *信学論 (B-I)*, vol.J81-B-I, no.8, pp.531-539, Aug. 1998.
- [4] S. Lin and D.J. Costello, Jr., *Error Control Coding*, Prentice Hall, New Jersey, 1983.
- [5] M.J. Miller and S. Lin, "The analysis of some selective-repeat ARQ schemes with finite receiver buffer," *IEEE Trans. Commun.*, vol.COM-29, no.9, pp.1307-1315, Sept. 1981.
- [6] E.J. Weldon, Jr., "An improved selective-repeat ARQ strategy," *IEEE Trans. Commun.*, vol.COM-30, no.3, pp.480-486, March 1982.
- [7] P.S. Yu and S. Lin, "An efficient selective-repeat ARQ scheme for satellite channels and its throughput analysis," *IEEE Trans. Commun.*, vol.COM-29, no.3, pp.353-363, March 1981.
- [8] 森村英典, 高橋幸雄, *マルコフ解析*, 日科技連, 東京, 1979.
- [9] J.O. Berger, *Statistical Decision Theory and Bayesian Analysis*, Springer-Verlag, New York, 1980. (平成 21 年 12 月 24 日受付, 22 年 3 月 19 日再受付)