

衛星 IP ネットワークのための輻輳制御 TCP Hybla の解析モデルの提案

Analytical Model of TCP Hybla for Satellite IP Networks

遠藤 崇江[†] 金 帝演[†] 内海 哲史[†] 加藤 靖[†] Salim Zahir[‡]
 Takanobu Endo Jeyeon Kim Satoshi Utsumi Yasushi Kato Salim Zahir

1. まえがき

衛星 IP ネットワークは通信をカバーできる地理的範囲が広く、災害時でも接続性が優れているため非常時の通信手段として有効である。衛星 IP ネットワークは、カバーできる範囲が広い、接続性が優れている、RTT(Round Trip Time) が大きい、リンクエラー率が高い、と言った特徴を持つ [1]。

有線ネットワークのための代表的な輻輳制御の手法としては TCP Reno [2] [3] が事実上の標準として用いられている。しかし、TCP Reno は RTT が小さい有線ネットワークを想定しているため、時間当たりのウィンドウサイズの増加量が小さく、RTT が大きい衛星 IP ネットワークでは十分な性能を維持できないことが知られている [4]。

一方、衛星 IP ネットワークのための輻輳制御の手法として TCP Hybla が提案されている [4]。TCP Hybla は Linux カーネル 2.6.13 以上で既に実装されている。また、TCP Hybla は時間当たりのウィンドウサイズの増加量が大きいため、RTT が大きい場合でもスループットが優れている。現在、TCP Reno についてはいくつもの解析モデルが存在している [5] [6] [7] [8] [9]。TCP Reno の解析モデルについては付録で述べる。しかし、TCP Hybla については解析がまだ十分でない。解析モデルを利用すれば、TCP 通信を行い続けることなく、Ping 等を使ってパケットロス率と RTT を測定することで、衛星 IP ネットワークにおける TCP Hybla のスループットを見積もることが可能となる。

本稿では、End-to-End で衛星 IP ネットワークにおける TCP Hybla のスループットを簡単に見積もるために、TCP Hybla の解析モデルを提案し、検証を行う。

2. TCP Hybla

TCP Hybla は C.Caini と R.Firriencieli によって提案された衛星 IP ネットワークのための輻輳制御の手法の一つである。TCP Reno の輻輳回避フェーズでは 1RTT 毎に 1 パケットずつウィンドウサイズが増加するため、RTT が大きい場合にスループットが低くなっていた。TCP Hybla の輻輳回避フェーズではウィンドウサイズの増加量は 1RTT 毎に ρ^2 ($\rho = \text{RTT}/\text{RTT}_0$) パケットずつ増加する。ここでの RTT_0 は性能を均一化するための基準となる値であり、たとえば文献 [4] では $\text{RTT}_0 = 25\text{ms}$ と設定している。RTT が大きくなれば ρ の値も大きくなるため、RTT が大きい場合でも高いスループットが維持できる。ACK を受信する毎に TCP

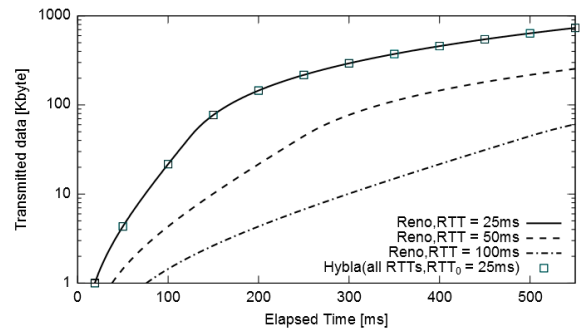


図1 TCP Hybla と TCP Reno のデータ送信量 [4]

Hybla のウィンドウサイズは次式に従って変化する。

$$W_{i+1} = \begin{cases} W_i + 2^\rho - 1 & \text{SS} \\ W_i + \rho^2/W_i & \text{CA} \end{cases} \quad (1)$$

ここで、SS (Slow Start) は輻輳制御におけるスロースタートフェーズ、CA (Congestion Avoidance) は輻輳回避フェーズを表している。 i は、受け取った ACK の個数を示す。

文献 [4] より時刻 t までの TCP Hybla のデータ送信量 $T(t)$ は次式で計算できる。

$$T(t) = \begin{cases} \frac{2^{t/\text{RTT}_0} - 1}{\ln(2)} & 0 \leq t < t_{\gamma,0} \quad \text{SS} \\ \frac{\gamma - 1}{\ln 2} + \frac{(t - t_{\gamma,0})^2}{2\text{RTT}_0^2} + \frac{(t - t_{\gamma,0})}{\text{RTT}_0} & t \geq t_{\gamma,0} \quad \text{CA} \end{cases} \quad (2)$$

ここで、 γ はスロースタートスレッシュホールド、 $t_{\gamma,0}$ はスロースタートフェーズから輻輳回避フェーズに移行するまでの時間である。また、 $t_{\gamma,0} = \text{RTT} \log_2 \gamma$ と計算される。式 (2) を利用して、図 1 に TCP Hybla と TCP Reno のデータ送信量を示す。ここで $\gamma = 32$ としている。式 (2) より、TCP Hybla の送信量は現在の RTT にかかわらず RTT_0 によって送信量が定まる。すなわち、TCP Hybla はパケットロスが発生しないとき、どの RTT でも TCP Reno の $\text{RTT} = \text{RTT}_0$ の場合とデータ送信量が同じになるようにウィンドウサイズが増加する。

3. TCP Hybla の解析モデルの提案

TCP Hybla の SS は式 (1) より $2^\rho - 1$ パケットずつウィンドウサイズが増加するため、RTT が大きくなると 1RTT でスロースタートスレッシュホールドを超える。よって、CA のウィンドウサイズの変化がスループットに大きく影響すると考えられる。本章では、TCP Hybla の輻輳回避フェーズにおけるスループットについて、解析モデルを提案する。提案する解析モデルは ρ とパケッ

[†] 国立高等専門学校機構鶴岡工業高等専門学校, National Institute of Technology, Tsuruoka College

[‡] オレンジジャパン株式会社, Orange Labs, Tokyo

トロス率, 及び RTT の値よりスループットの計算が可能となる. パラメータと関数を以下のように定義する.

- p : パケットロス率.
- R : RTT の平均値.
- Y : 輻輳回避 (CA: Congestion Avoidance)/早期再送 (FR: Fast Retransmit) フェーズにおいて送信されるパケット数の期待値.
- X : CA フェーズにおける RTT の回数の期待値.
- A : CA/FR フェーズにおける時間の期待値.
- $A(W, m)$: 輻輳ウィンドウサイズ W において m パケットがロスする確率.
- ρ : RTT/RTT_0 . (ここで, RTT_0 は, 参照するコネクシオンの RTT. たとえば, $RTT_0 = 25ms$.)

3.1. ピーク時の輻輳ウィンドウサイズの平均値が最大値に達しない場合

追加のパラメータを以下のように定義する.

- W : ピーク時の輻輳ウィンドウサイズの平均値.
- L : CA/FR フェーズにおいてロスするパケット数の期待値.
- D : 輻輳ウィンドウサイズの縮小率の期待値.

定義より, ピーク時の輻輳ウィンドウの最初のパケットは常にロスするので, ピーク時の輻輳ウィンドウサイズ W について, m パケットロスする確率は, 次のようになる.

$$A(W, m) = C_{m-1}^{W-1} (1-p)^{W-m} p^{m-1}. \quad (3)$$

最初のパケットロスは, 3 つの重複確認応答によって検出されるため, $m \leq W-3$ となる. 従って,

$$\begin{aligned} L &= \sum_{m=1}^{W-3} mA(W, m) \\ &= \sum_{m=0}^{W-4} (1+m) C_m^{W-1} (1-p)^{W-1-m} p^m \\ &\approx \sum_{m=0}^{W-1} C_m^{W-1} (1-p)^{W-1-m} p^m \\ &\quad + \sum_{m=0}^{W-1} m C_m^{W-1} (1-p)^{W-1-m} p^m \\ &= 1 + p(W-1) \\ &\approx 1 + pW. \end{aligned} \quad (4)$$

また, TCP Hybla の場合, 輻輳ウィンドウあたり複数回パケットロスが発生したとき, その回数だけ, 輻

輳ウィンドウを半減する. 従って,

$$\begin{aligned} D &= \sum_{m=1}^{W-3} \left(\frac{1}{2}\right)^m A(W, m) \\ &= \frac{1}{2} \times \sum_{m=1}^{W-3} \left(\frac{1}{2}\right)^{m-1} C_{m-1}^{W-1} (1-p)^{W-m} p^{m-1} \\ &= \frac{1}{2} \times \sum_{m=0}^{W-4} \left(\frac{1}{2}\right)^m C_m^{W-1} (1-p)^{W-1-m} p^m \\ &\approx \frac{1}{2} \times \sum_{m=0}^{W-1} C_m^{W-1} \left(\frac{p}{2}\right)^m (1-p)^{W-1-m} \\ &= \frac{1}{2} \times \left(\frac{p}{2} + (1-p)\right)^{W-1} \\ &= \frac{(2-p)^{W-1}}{2^W}. \end{aligned} \quad (5)$$

最初のパケットロスが発生するまでにその CA/FR フェーズに送信されるパケット数は, $1/p$ である [5] [8]. 損失パケットを送信してから, 輻輳ウィンドウ W だけパケットを送信し続け, L パケットがロスするので,

$$Y = \frac{1}{p} + W - L. \quad (6)$$

平均輻輳ウィンドウを D 倍 ($D \leq 1/2$) だけ縮小してから, $X \times RTT$ で輻輳ウィンドウが平均値に戻るので,

$$W = DW + \rho^2 X. \quad (7)$$

変形すると,

$$X = \frac{1-D}{\rho^2} W. \quad (8)$$

また, CA フェーズに送信されるパケット数の期待値は $\sum_{k=1}^X (DW + \rho^2 k)$, FR フェーズに送信されるパケット数の期待値は $\frac{W}{2} + \frac{W}{4}(1 - \frac{1}{L})$ なので,

$$\begin{aligned} Y &= \sum_{k=1}^X (DW + \rho^2 k) + \frac{W}{2} + \frac{W}{4}(1 - \frac{1}{L}) \\ &= X(DW + \frac{\rho^2}{2}(X+1)) + \frac{3W}{4} - \frac{W}{4L}. \end{aligned} \quad (9)$$

式 (4) ~ (9) より, W についての方程式が導ける. また, CA フェーズの時間の期待値は $X \times RTT$, FR フェーズの時間の期待値は $(1 + (1 - \frac{1}{L})) \times RTT$ なので,

$$A = (X + 1 + (1 - \frac{1}{L})) \times R. \quad (10)$$

スループット B は,

$$B = \frac{Y}{A} = \frac{\frac{1}{p} + W - L}{A}. \quad (11)$$

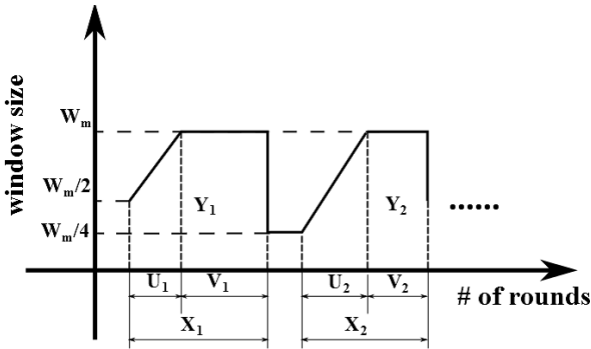


図 2 ウィンドウサイズに上限がある場合のウィンドウサイズのモデル

3.2. ピーク時の輻輳ウィンドウサイズの平均値が最大値に達する場合

TCP の受信側は受信できる最大のウィンドウサイズ W_m を送信側へ通知する機能を持つ。ここで、ピーク時の輻輳ウィンドウサイズの平均値が最大値に達する場合を考える。このモデルは、最大ウィンドウサイズ W_m よりも、前節の方程式で導かれる W の値が大きい場合に適用する。図 2 に輻輳ウィンドウサイズが最大値に達する場合のウィンドウサイズのモデルを示す。追加のパラメータを以下のように定義する。

- W_m : 輻輳ウィンドウサイズの最大値。
- L_m : CA/FR フェーズにおいてロスするパケット数の期待値。
- D_m : 輻輳ウィンドウサイズの縮小率の期待値。
- U : CA フェーズにおいて、輻輳ウィンドウサイズが増加する RTT の回数の期待値。
- V : CA フェーズにおいて、輻輳ウィンドウサイズが最大値となる RTT の回数の期待値。

ピーク時の平均輻輳ウィンドウサイズが最大値に達しない場合と同様、 $m \leq W_m - 3$ となる。従って、

$$L_m = \sum_{m=1}^{W_m-3} mA(W, m) \approx 1 + pW_m, \quad (12)$$

$$D_m = \sum_{m=1}^{W_m-3} \left(\frac{1}{2}\right)^m A(W, m) \approx \frac{(2-p)W_m^{-1}}{2W_m}. \quad (13)$$

最初のパケットロスが発生するまでにその CA/FR フェーズに送信されるパケット数は、 $1/p$ であり、損失パケットを送信してから、最大輻輳ウィンドウ W_m 分だけパケットを送信し続け、 L_m パケットがロスするので、

$$Y = \frac{1}{p} + W_m - L_m. \quad (14)$$

最大輻輳ウィンドウを D_m 倍 ($D_m \leq 1/2$) だけ縮小してから、 $U \times RTT$ で輻輳ウィンドウが最大値に戻るので、

$$W_m = D_m W_m + \rho^2 U. \quad (15)$$

変形すると、

$$U = \frac{1 - D_m}{\rho^2} W_m. \quad (16)$$

ここで、CA フェーズにおける RTT の回数の期待値について、以下の式が成り立つ。

$$X = U + V. \quad (17)$$

また、CA フェーズに送信されるパケット数の期待値は $\frac{1}{2}(D_m + 1)W_m U + W_m V$ 、FR フェーズに送信されるパケット数の期待値は $\frac{W_m}{2} + \frac{W_m}{4}(1 - \frac{1}{L_m})$ なので、

$$Y = \frac{1}{2}(D_m + 1)W_m U + W_m V + \frac{W_m}{2} + \frac{W_m}{4}(1 - \frac{1}{L_m}). \quad (18)$$

CA フェーズの時間の期待値は $X \times RTT$ 、FR フェーズの時間の期待値は $(1 + (1 - \frac{1}{L_m})) \times RTT$ なので、

$$A = (X + 1 + (1 - \frac{1}{L_m})) \times R. \quad (19)$$

スループット B は、

$$B = \frac{Y}{A} = \frac{\frac{1}{p} + W_m - L_m}{A}. \quad (20)$$

4. 解析モデルの検証

ns-2 [10] を使用して TCP Hybla のシミュレーションを行い、シミュレーションと提案する解析モデルのスループットの比較を行う。解析モデルについては、Newton 法を用いて、近似解を求める。シミュレーションはリンクエラー率(パケットロス率)を変更して、それぞれ 10 回測定し、そのスループットの平均値を取る。また、提案した解析モデルと TCP Reno の解析モデルについても比較を行う。シミュレーションにおけるパラメータを表 1 に、シミュレーショントポロジーを図 3 に示す。バッファオーバーフローによるパケットロスの影響を小さくするため、バッファ長を十分大きく取る(ここでは、1,000packets。)図 3 のリンク容量(Link Capacity)は W_m/RTT とする。

表 1 シミュレーションパラメータ

Packet Size	1,000 bytes
W_m	550 packets
Buffer Size	1,000 packets
Simulation Time	500 sec

RTT=100ms の時の評価結果を図 4 に、RTT=250ms の時の評価結果を図 5 に、RTT=550ms の時の評価結果を図 6 に示す。

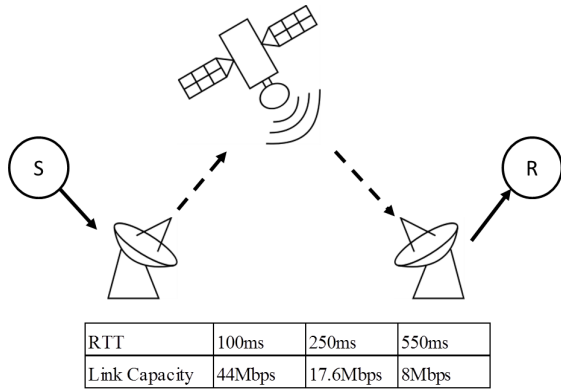


図3 シミュレーショントポロジー

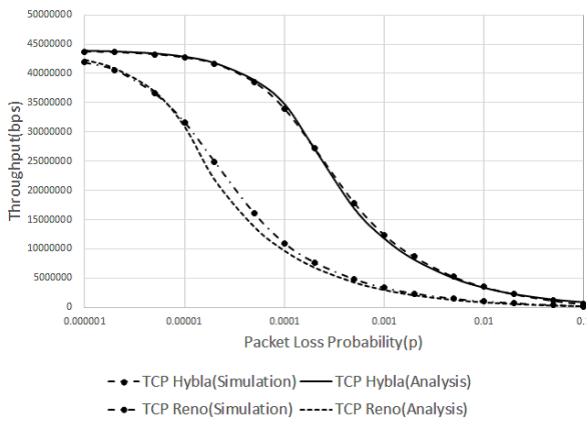


図4 RTT=100msの時のシミュレーションと解析モデルのスループット

TCP Hybla について, RTT=100ms ではパケットロス率が 0.01 まで, 差異が 5% 以内に収まった. また, RTT=250ms でもパケットロス率が 0.01 までについて, 差異が 5% 以内に収まった. RTT=550ms ではパケットロス率が 0.01 までについて, 最大の差異が 13% 程度になった.

パケットロス率が 0.01 以上については, 今回は解析モデルにおいてタイムアウトを考慮していないため, シミュレーションとの差異の割合が大きくなった.

5. まとめ

本稿では End-to-End で衛星 IP ネットワークにおける TCP Hybla のスループットを簡単に見積もるために, TCP Hybla の解析モデルの提案と検証を行った. 評価の結果, パケットロス率が 0.01 以下について, RTT=100ms, 250ms ではシミュレーションとの差異が 5% 以内に収まった. RTT=550ms の場合は最大で 13% 程度となった. これらの結果から, End-to-End で衛星 IP ネットワークのスループットを簡単に見積もるための基本的な知見を得ることができた.

今後の課題として, タイムアウトを考慮した TCP

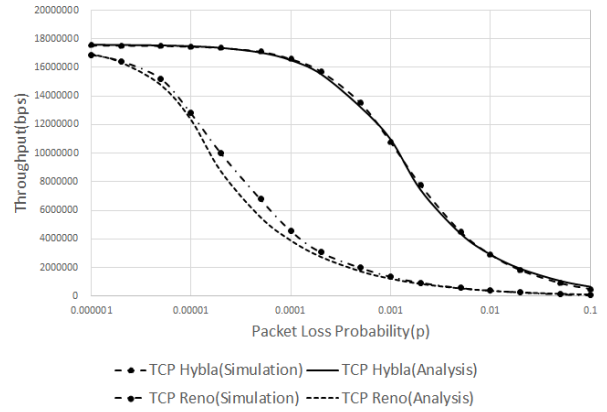


図5 RTT=250msの時のシミュレーションと解析モデルのスループット

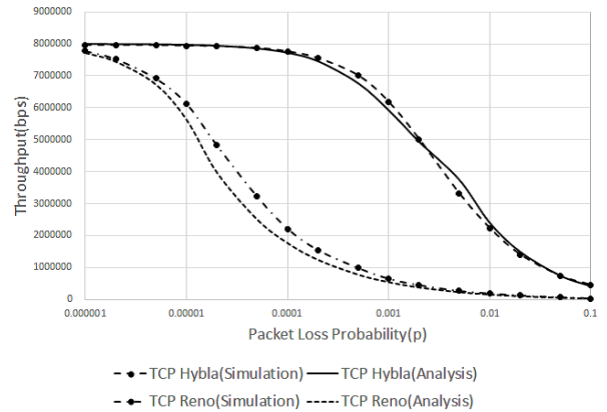


図6 RTT=550msの時のシミュレーションと解析モデルのスループット

Hybla の性能解析を行い, より大きなパケットロス率についても高い精度の解析モデルを検討する必要がある.

謝辞

詳細なコメントをくださった匿名の査読者に感謝する.

参考文献

- [1] M. Hassan and R. Jain, High Performance TCP/IP Networking: Concepts, Issues, and Solutions, Pearson Prentice Hall, 2003.
- [2] V. Jacobson, "Modified TCP congestion avoidance algorithm," end2end-interest mailing list, ftp://ftp.isi.edu/end2end/end2end-interest-1990.mail. April 1990.
- [3] W. Stevens, "TCP Slow Start, Congestion Avoidance, Fast Retransmit, and Fast Re-

covery Algorithms,” IETF, <http://www.rfc-editor.org/rfc/rfc2001.txt>, Jan.1997.

- [4] C. Caini, R. Firrincieli, ”TCP Hybla: a TCP enhancement for heterogeneous networks,” International Journal of Satellite Communications and Networking, vol.22, no.5, pp547-566, Aug.2004.
- [5] J. Padhye, V. Firoiu, D.F. Towsely, and J. Kurose, ”Modeling TCP Reno Performance: A Simple Model and Its Empirical Validation,” IEEE/ACM Transactions on Networking, vol.8, no.2, pp133-145, April 2000.
- [6] N. Cardwell, S. Savage, and T. Anderson, ”Modeling TCP Latency,” Proc. of IEEE INFOCOM 2000, vol.3, pp1742-1751, March 2000.
- [7] D. Zheng, G. Y. Lazarou, and R. Hu, ”A Stochastic Model for Short-lived TCP Flows”, IEEE International Conference on Communications 2003 (ICC 2003), May 2003, pp.11-15.
- [8] Z. Chen, T. Bu, M. Ammar, and D. Towsely, ”Comments on Modeling TCP Reno Performance: A Simple Model and Its Empirical Validation,” IEEE/ACM Transactions on Networking, vol.14, no.2, pp451-453, April 2006.
- [9] N. Parvez, A. Mahanti, and C. Williamson, ”An analytic throughput model for TCP NewReno,” IEEE/ACM Transaction on Networking, vol. 18, Issue 2, April 2010.
- [10] UCB/LBNL/VINT : Network Simulator - ns (version2), <http://www.isi.edu/nsnam/ns/>.

付録

A. TCP Reno の解析モデル [5] [8]

ここでは TCP Reno の解析モデルについて示す。この解析モデルでは、ウィンドウサイズが増加するのに必要な RTT の回数 b とパケットロス率 p 、及び RTT からスループットの計算が可能となる。

A.1. ピーク時の輻輳ウィンドウサイズの平均値が最大値に達しない場合

図 A-1 に、パケットロスが検出されてから、次のパケットロスを検出するまでの TCP Reno のパケット送信モデルを示す。パケット送信数の期待値を Y 、ピーク時の平均輻輳ウィンドウサイズを W とすると、

$$Y = \frac{1-p}{p} + W \tag{A-1}$$

また、TCP Reno は b 回の RTT 毎にウィンドウサイズが 1 パケットずつ増加するため、 Y は次式で計算

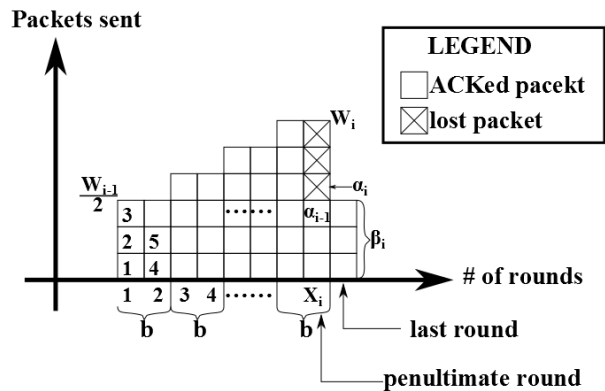


図 A-1 TCP Reno のパケット送信モデル [5]

できる。

$$\begin{aligned} Y &= \sum_{k=0}^{X/b-1} \left(\frac{W}{2} + k \right) b + \frac{W}{2} \\ &= \frac{X}{2} \left(\frac{W}{2} + W \right) + \frac{W}{2} \end{aligned} \tag{A-2}$$

また、 W は次式で計算できる。

$$W = \frac{W}{2} + \frac{X}{b} - 1 \tag{A-3}$$

式 (A-3) を変形すると、

$$X = b \left(\frac{W}{2} + 1 \right) \tag{A-4}$$

となり、式 (A-1)、式 (A-2)、式 (A-4) より、 W は次式で計算できる。

$$W = \sqrt{\frac{8(1-p)}{3bp} + \left(\frac{3b-2}{3b} \right)^2} - \frac{3b-2}{3b} \tag{A-5}$$

CA/FR フェーズの時間の期待値 A は、

$$A = (X + 1) \cdot R \tag{A-6}$$

スループット B は、

$$B = \frac{Y}{A} = \frac{\frac{1-p}{p} + W}{A} \tag{A-7}$$

A.2. ピーク時の輻輳ウィンドウサイズの平均値が最大値に達する場合

ウィンドウサイズに上限があるモデルを図 A-2 に示す。最大ウィンドウサイズ W_m は次式で計算できる。

$$W_m = \frac{W_m}{2} + \frac{U}{b} - 1 \tag{A-8}$$

変形すると、

$$U = \left(\frac{W_m}{2} + 1 \right) b \tag{A-9}$$

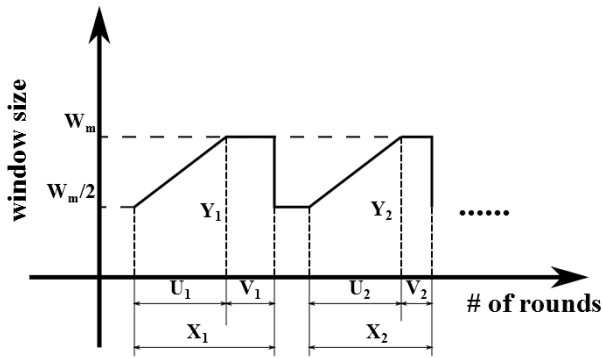


図 A-2 ウィンドウサイズに上限がある場合の TCP Reno のウィンドウサイズのモデル [5]

パケットロスを検出してから，次のパケットロスを検出するまでに送信したパケット送信数の期待値 Y は，

$$Y = \frac{1-p}{p} + W_m, \quad (\text{A-10})$$

$$Y = \frac{3}{4}W_m U + W_m V + \frac{W_m}{2}. \quad (\text{A-11})$$

式 (A-9) ~ 式 (A-11) より，

$$V = \frac{1-p}{W_m p} - \frac{3b}{8}W_m + \frac{3b+2}{4}. \quad (\text{A-12})$$

CA フェーズの RTT の平均回数 X は，次式で計算できる．

$$\begin{aligned} X &= U + V \\ &= \frac{1-p}{W_m p} + \frac{b}{8}W_m + \frac{b+2}{4}. \end{aligned} \quad (\text{A-13})$$

CA/FR フェーズの時間の期待値 A は，

$$A = (X + 1) \cdot R. \quad (\text{A-14})$$

スループット B は，

$$B = \frac{Y}{A} = \frac{\frac{1-p}{p} + W_m}{A}. \quad (\text{A-15})$$