社団法人　電子情報通信学会
THE INSTITUTE OF ELECTRONICS,
INFORMATION AND COMMUNICATION ENGINEERS

信学技報
TECHNICAL REPORT OF IEICE.

インターネットにおけるルータ間の輻輳伝播現象

杉山　浩平† 　奥村　治† 　大崎　博之† 　今瀬　眞†

† 大阪大学 大学院情報科学研究科
〒 565-0871 大阪府吹田市山田丘 1-5
E-mail: †{k-sugi,o-okumr,oosaki,imase}@ist.osaka-u.ac.jp

あらまし　近年、インターネットにおいて様々な非線型現象が発生することが報告されている。例えば、輻輳が発生しているルータからその周辺のルータへ、波のように輻輳が伝わるという、ルータ間の輻輳伝播が報告されている。これまでルータ間の輻輳伝播に着目した研究がいくつか行われている。しかしこれらの研究では、ルータ間の輻輳伝播がなぜ発生するのか、また、ルータ間の輻輳伝播がどのような条件下で発生するのかが十分明らかにされていない。本稿では、シミュレーション実験により、ルータ間の輻輳伝播が発生する原因の一つを明らかにするとともに、どのような条件下でルータ間の輻輳伝播が発生するかを明らかにする。その結果、リンクの帯域や伝搬遅延の大きさ、トラヒックのランダム性によって、ルータ間の輻輳伝播の周期が変化すること、ネットワークのランダム性が強くなるに従い、ルータ間の輻輳伝播の周期性が弱まること、などが明らかになった。
キーワード　輻輳伝播、TCP (Transmission Control Protocol)、ルータ、非線型現象

Congestion Propagation among Routers in the Internet

Kouhei SUGIYAMA†, Osamu OKUMURA†, Hiroyuki OHSAKI†, and Makoto IMASE†

† Graduate School of Information Science and Technology, Osaka University
1-5 Yamadaoka, Suita, Osaka, 565-0871 Japan
E-mail: †{k-sugi,o-okumr,oosaki,imase}@ist.osaka-u.ac.jp

Abstract　In recent years, various non-linear phenomena of the Internet have been discovered. For instance, it is reported that the congestion of a router propagates to neighboring routers like a wave. Several researches on the congestion propagation between routers have been performed. However, in these researches, cause of the congestion propagation between routers and condition that the congestion propagation between routers occurs have not been sufficiently investigated. In this paper, we reveal a cause of the congestion propagation between routers, and also investigate under what conditions the congestion propagation between routers is observed. Consequently, we show that the speed of the congestion propagation between routers is affected by the bandwidth and the propagation delay of links, and that the periodicity of the congestion propagation between routers becomes less obvious as the randomness of network traffic increases.
Key words　Congestion Propagation, TCP (Transmission Control Protocol), Router, Nonlinear Phenomena

1　は じ め に

インターネットは巨大な非線型システムであり、インターネットの非線型なダイナミクスに注目が集まっている。近年、インターネットにおいて、様々な非線型現象が観測されることが報告されている。例えば、イーサネットやインターネットのトラヒックが自己相似性を持つこと [1,2]、TCP コネクションのウィンドウサイズの時間的な変動がカオス的な挙動を示すこと [3]、ルータの輻輳が近隣のルータへ伝播すること [4] などが報告されている。

ルータ間の輻輳伝播とは、輻輳が発生しているルータからその周辺のルータへ、波のように輻輳が伝わる現象である。輻輳伝播の様子を図 1 に示す。まず、ルータ 1 で輻輳が発生すると、それに遅れて、ルータ 2、ルータ 3、ルータ 4、ルータ 5 で輻輳が発生するとともに、ルータ 1 の輻輳が解消される。同様に、ルータ 6、ルータ 7、ルータ 8、ルータ 9 で輻輳が発生するとともに、ルータ 2、ルータ 3、ルータ 4、ルータ 5 の輻輳が解消される。

これまで、インターネットにおけるルータ間の輻輳伝播に関する研究がいくつか行われている [5,6]。文献 [5] では、現実のネットワークにおいて、ルータ間の輻輳伝播が観測されることを報告している。また、文献 [6] では、シミュレーション実験

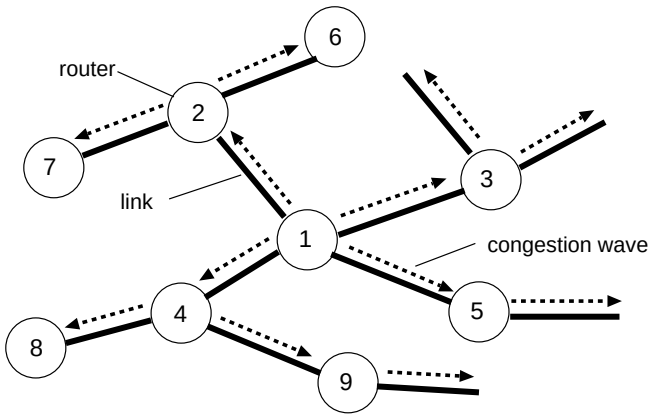

図 1: インターネットにおけるルータ間の輻輳伝播の例

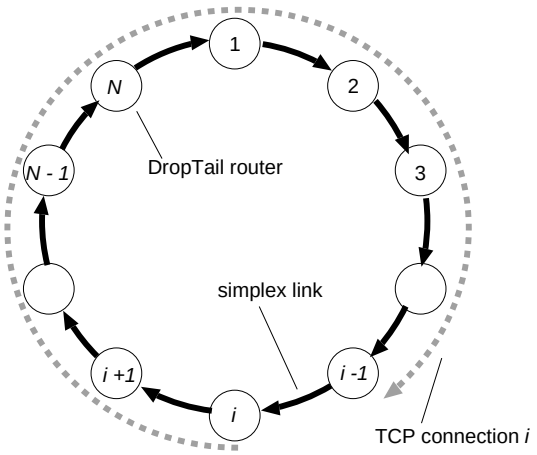

図 2: シミュレーションモデル

表 1　シミュレーションで用いたパラメータ設定

ノード数	N	10
リンクの帯域	B	10 [Mbit/s]
リンクの伝搬遅延	τ	31 [ms]
ルータのバッファサイズ	L	300 [packet]
ルータのキュー管理機構		DropTail
パケット長	S	552 [byte]
TCP バージョン		TCP Tahoe

により、ルータ間の輻輳伝播を調査している。図 2 に示すような、単方向リンクにより接続されたリング型ネットワークにおいて、i 番目のルータから $i-1$ 番目のルータへ向けて、連続的な TCP トラヒックを生成している。その結果、ルータ間の輻輳伝播が観測されるとともに、各 TCP コネクションの転送レートが周期的に変動することを示している。しかしこれらの研究では、ルータ間の輻輳伝播がなぜ発生するのか、また、ルータ間の輻輳伝播がどのような条件下で発生するのかが十分明らかにされていない。

そこで本稿では、ルータ間の輻輳伝播が発生する原因の一つを明らかにするとともに、どのような条件下でルータ間の輻輳伝播が発生するかを明らかにする。本稿では、文献 [6] と同じリング型のネットワークモデルを用い、輻輳伝播が発生する条件を明らかにするため、さまざまなネットワークパラメータやシステムパラメータを変化させてシミュレーションを行う。特に、リンクの帯域や伝搬遅延、ルータのバッファサイズやキュー管理機構 (DropTail や RED)、TCP のオプションやバージョンの違いが、ルータ間の輻輳伝播に与える影響を明らかにする。

本稿の構成は、以下の通りである。まず 2 章では、シミュレーション実験によりルータ間の輻輳伝播の一例を示す。3 章では、ルータ間の輻輳伝播が発生する原因について議論する。4 章では、さまざまなネットワークパラメータやシステムパラメータを変化させてシミュレーションを行い、ルータ間の輻輳伝播が発生する条件を明らかにする。最後に 5 章において、本稿のまとめと今後の課題を述べる。

2　ルータ間の輻輳伝播

シミュレーションに用いたネットワークモデルを図 2 に示す。文献 [6] と同様に、N 台のルータが単方向リンクで接続された、リング型ネットワークを用いる。以下では、$i\,(1 \leqq i \leqq N)$ 番目のルータをルータ i と表記する。図 2 に示すように、ルータ i $(1 \leqq i \leqq N)$ からルータ $i-1$ に向けて (ただし、ルータ 1 からはルータ N に向けて)、TCP コネクション i が連続的にデータを転送する。図 2 には、ルータ i からルータ $i-1$ への TCP コネクションのみを示しているが、実際には N 本の TCP コネクションが存在することに注意されたい。

シミュレーションで用いたパラメータ設定を表 1 に示す。特に断りの無い限り、以下のシミュレーションでは 表 1 に示すパラメータを用いている。シミュレーションには ns-2 (version 2.28) [7] を用いた。

この時の、各ルータのキュー長 (バッファ内パケット数) の時間的変動を図 3 に示す。図 3 では、X 軸は時間、Y 軸はルータの番号 i を示している。図中には、10 [s] ごとに計測したルータのキュー長を色の明暗により示している。この図より、あるルータで発生した輻輳 (キュー長の増加) が、他のルータへ順番に伝播している様子が確認できる。また、キュー長の増加・減少が、下流のルータから上流のルータへ (ルータ i からルータ $i-1$ へ) と伝播していることが分かる。この結果より、ルータの輻輳が、下流のルータから上流のルータへと規則的に伝播していることがわかる。

さらに、この時の各 TCP コネクションの転送レートの時間的変動を図 4 に示す。図 4 では、X 軸は時間、Y 軸は TCP コネクションの番号 i を示している。図中には、10 [s] ごとに計測した TCP コネクションの転送レートを色の明暗により示している。この図より、ある TCP コネクションの転送レートの増加・減少が、他の TCP コネクションへ順番に伝播している様子が確認できる。また、ルータのキュー長の変動と同様に、TCP コネクションの転送レートの変動が、下流のコネクションから上流のコネクションへと規則的に伝播していることがわかる。

図 3 および図 4 を比較することにより、ルータ間の輻輳伝播と TCP コネクションの転送レートの変動が同じ周期を持つことが分かる。このことから、ルータ間の輻輳伝播の影響により、

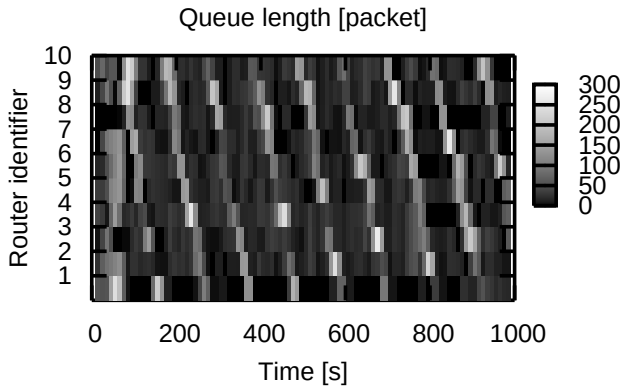

図 3: ルータのキュー長の時間的変動 ($B = 10$ [Mbit/s])

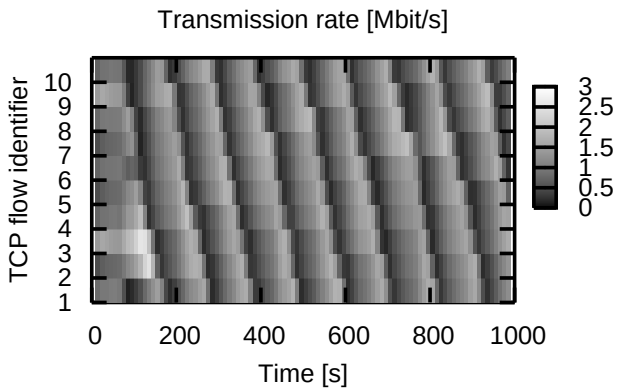

図 4: TCP コネクションの転送レートの時間的変動 ($B = 10$ [Mbit/s])

TCP コネクションの転送レートが周期的に変動している、もしくは逆に、TCP コネクションの転送レートの周期的な変動の影響により、ルータ間で輻輳が伝播していると予想される。

そこで次章では、図 2 に示すリング型のネットワークにおいて、なぜルータ間の輻輳伝播が発生するのか、その理由について議論する。

3 なぜルータ間の輻輳伝播が発生するか?

2 章のシミュレーションにおける、ルータ 1 のキュー長および TCP コネクション 1 の転送レートの時間的変動を図 5 に示す。2 章で議論したように、ルータのキュー長の変動と TCP コネクションの転送レートの変動が同じ周期を持つことがこの図からも分かる。

ルータのキュー長と、TCP コネクションの転送レートはほぼ同じタイミングで変動しているが、TCP コネクションの転送レートの変動が、ルータのキュー長の変動よりもわずかに (10 [s] 程度) 早いことがわかる。このことから、TCP コネクションの転送レートの周期的な変動により、ルータ間の輻輳伝播

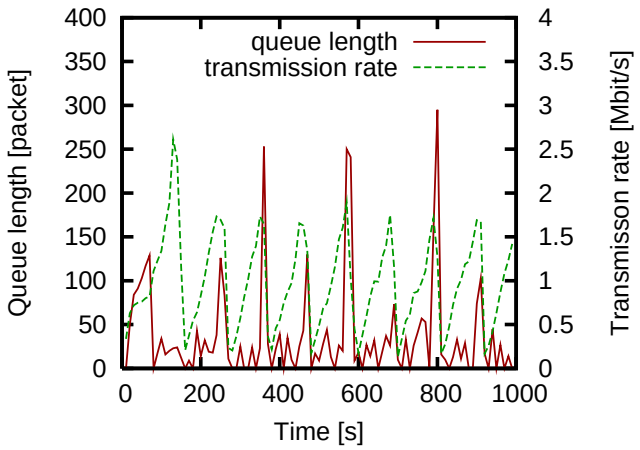

図 5: ルータ 1 のキュー長と TCP コネクション 1 の転送レートの時間的変動

(キュー長の周期的変動) が引き起こされていると考えられる。

そこで、なぜ TCP コネクションの転送レートが周期的に変動するのかを考える。図 3 から図 5 より、TCP コネクション 1 の転送レートの増減と、ルータ 1 のキュー長の増減に、強い正の相関が見られる。このことから、TCP コネクション 1 の転送レートが高い時には、TCP コネクション 1 から送出された多くのパケットが、ルータ 1 のキューに格納されている (バッファリングされている) と考えられる。

ここで、ある時点において、すべての TCP コネクションの中で、最も転送レートの大きい TCP コネクション i に着目する。この時、TCP コネクション i から送出された多くのパケットが、ルータ i のキューに格納されている。このため、ルータ i のキューがオーバーフローした時には、**TCP** コネクション i から送出されたパケットが廃棄される確率が高くなる。

TCP コネクション i のパケットが廃棄されれば、TCP コネクション i はウィンドウフロー制御により転送レートを減少する。パケット廃棄直前に、TCP コネクション i の転送レートが最大であったため、TCP コネクション i が転送レートを減少すると、ルータ i のキュー長が大きく減少させる。この直後に、ルータ i のキューにパケットを格納できる確率が高いのは、ルータ i と送信側ホストとの距離が最も短かい **TCP** コネクション $i-1$ である。つまり、すべての TCP コネクションの中で、TCP コネクション i の転送レート減少の直後に、転送レートを増加できる可能性が高いのは TCP コネクション $i-1$ である。

上記のような処理が繰り返されることにより、結果として、TCP コネクション i の転送レートの増加・減少が、TCP コネクション $i-1$ へ順番に伝播する、つまり、TCP コネクションの転送レートの変動が、下流のコネクションから上流のコネクションへと伝播すると考えられる。

4 ルータ間の輻輳伝播が発生する条件

本章では、さまざまなネットワークパラメータやシステムパラメータを変化させてシミュレーションを行う。特に、リンク

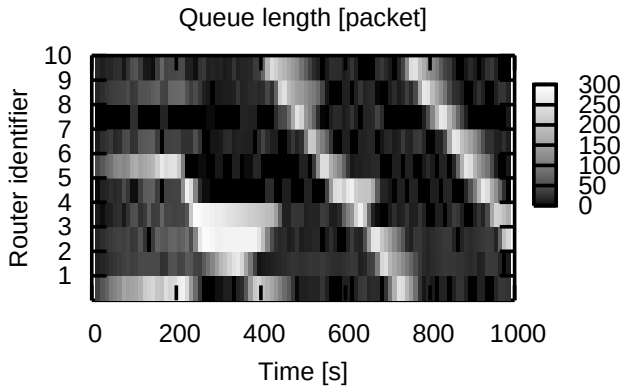

図 6: ルータのキュー長の時間的変動 ($B = 1$ [Mbit/s])

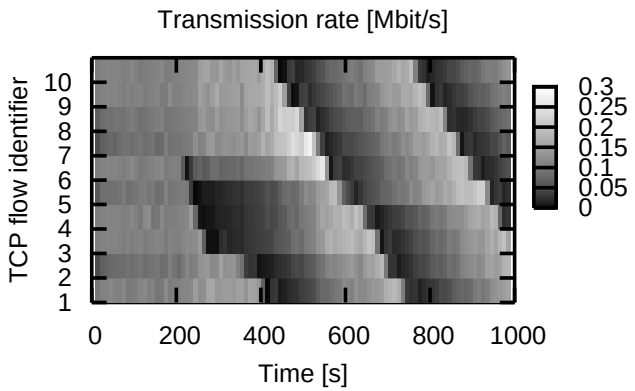

図 7: TCP コネクションの転送レートの時間的変動 ($B = 1$ [Mbit/s])

の帯域や伝搬遅延、ルータのバッファサイズやキュー管理機構 (DropTail や RED)、TCP のオプションやバージョンの違いが、ルータ間の輻輳伝播に与える影響を明らかにする。

4.1 システムパラメータの影響

まず、システムパラメータとして、リンクの帯域、リンクの伝搬遅延、ルータのバッファサイズが、ルータ間の輻輳伝播に与える影響を調査する。

すべてのリンクの帯域を $B = 1$ [Mbit/s] に設定した時の、各ルータのキュー長の時間的変動を図 6 に、各 TCP コネクションの転送レートの時間的変動を図 7 に示す。

図 3 ($B = 10$ [Mbit/s]) および図 6 ($B = 1$ [Mbit/s]) を比較すると、リンクの帯域にかかわらず、各ルータのキュー長が周期的に変動している (ルータ間で輻輳が伝播している) ことがわかる。ただし、図 3 ($B = 10$ [Mbit/s]) および図 6 ($B = 1$ [Mbit/s]) を比較すると、ルータ間の輻輳伝播の周期が約 3 倍に大きくなっていることがわかる。同様に、TCP コネクションの転送レートの変動の周期も約 3 倍に大きくなっている (図 7)。これは、各 TCP コネクションのラウンドトリップ時間が増加したために、その結果として、ルータ間の輻輳伝播の周期が大きくなっていると考えられる。

なお、紙面の制約のため結果は省略するが、リンクの伝搬遅延を $\tau = 55$ [ms] と設定した場合や、ルータのバッファサイズを $L = 600$ [packet] と設定した場合にも、同様にルータ間の輻輳伝播は発生していた。ただし、ルータ間の輻輳伝播の周期はそれぞれ異なっていた。これらはすべて、各 TCP コネクションのラウンドトリップ時間が変化したために、ルータ間の輻輳伝播の周期が変化したと考えられる。

以上の結果から、システムパラメータである、リンクの帯域、リンクの伝搬遅延、ルータのバッファサイズは、ルータ間の輻輳伝播の有無には影響を与えないこと、ルータ間の輻輳伝播の周期は TCP コネクションのラウンドトリップ時間によって変化することが分かった。

4.2 トラヒックのランダム性の影響

DropTail ルータに複数の TCP コネクションが収容された場合、それらの TCP コネクションの動作が同期してしまうという現象 (フェーズ効果) が知られている [8]。文献 [8] では、TCP トラヒックにランダム性を持たせることにより、フェーズ効果が発生しなくなることが示されている。具体的には、TCP 送信側ホストからのパケット送信タイミングをランダムに遅らせることにより、フェーズ効果が発生しなくなることが示されている。

3 章で述べたように、ルータ間の輻輳伝播は、各 TCP コネクションの周期的な転送レートの変動が原因である。そこで、TCP トラヒックにランダム性を持たせることにより、TCP コネクションの周期的な転送レートの変動が消失し、その結果、ルータ間の輻輳伝播も発生しなくなるという可能性が考えられる。

そこで、TCP 送信側ホストからのパケット送信タイミングに、ランダムな遅延を加えた場合のシミュレーションを実行した。具体的には、TCP 送信側ホストからのパケット送信時に、0–0.1 [s] のランダムな遅延を加えた。この時の、各ルータのキュー長の時間的変動を図 8 に、各 TCP コネクションの転送レートの時間的変動を図 9 に示す。図 3 および図 8 を比較すると、TCP トラヒックにランダム性を加えた場合でも、ルータ間の輻輳伝播は観測されるが、その規則性が弱まっていることがわかる。

以上の結果から、TCP トラヒックにランダム性を加えることにより、ルータ間の輻輳伝播の規則性は弱まるが、輻輳伝播自体は消失しないことが分かった。

4.3 ルータのキュー管理機構の影響

DropTail ルータにおけるフェーズ効果を解消する別の方法として、RED (Random Early Detection) [9] に代表されるアクティブキュー管理機構が提案されている [9–12]。そこで、ルータのキュー管理機構を DropTail から RED に変更してシミュレーションを行った。

DropTail ルータの代わりに、RED ルータを用いた時の、ルータのキュー長の時間的変動を図 10 に、各 TCP コネクションの

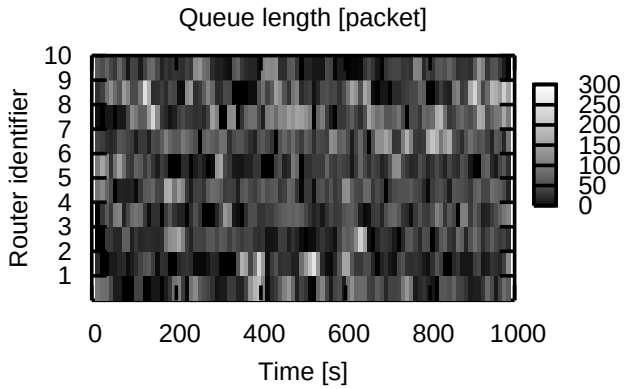

図 8: ルータのキュー長の時間的変動 (TCP トラヒックにランダム性を加えた場合)($B = 10$ [Mbit/s])

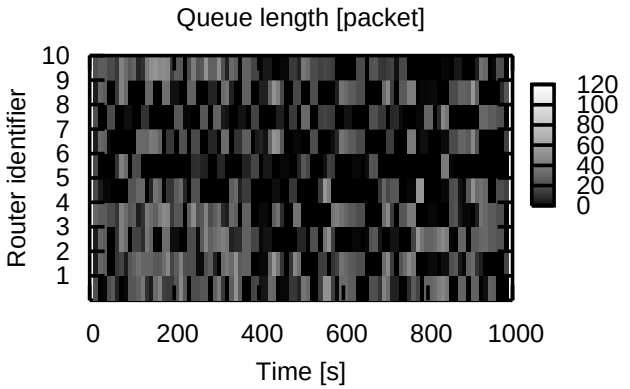

図 10: ルータのキュー長の時間的変動 (RED ルータの場合)($B = 10$ [Mbit/s])

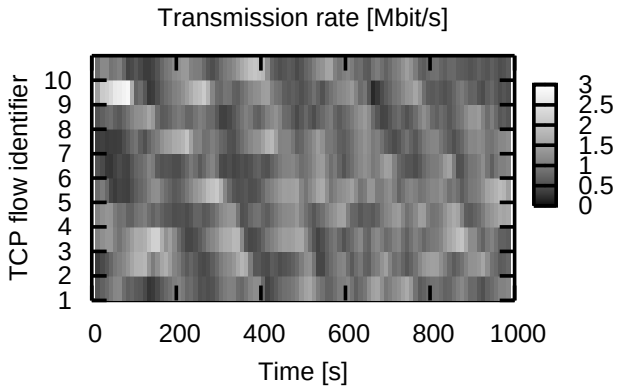

図 9: TCP コネクションの転送レートの時間的変動 (TCP トラヒックにランダム性を加えた場合)($B = 10$ [Mbit/s])

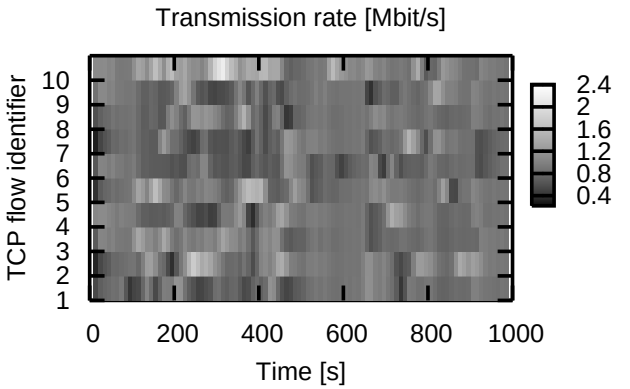

図 11: TCP コネクションの転送レートの時間的変動 (RED ルータの場合)($B = 10$ [Mbit/s])

転送レートの時間的変動を図 11 に示す。図 3 および図 10 を比較すると、RED ルータを用いた場合でも、ルータ間の輻輳伝播は観測されるが、その規則性が弱まっていることがわかる。

以上の結果から、ルータのキュー管理機構を RED にすることにより、ルータ間の輻輳伝播の規則性は弱まるが、輻輳伝播自体は消失しないことが分かった。

4.4 TCP バージョンの影響

最後に、TCP のバージョンが、ルータ間の輻輳伝播に影響を与えるかどうかを調査する。

3 章で議論したように、ルータ間の輻輳伝播が発生する原因の一つは、TCP コネクション i が転送レートを減少させた時 (つまり、帯域を解放した時) に、TCP コネクション $i-1$ が解放された帯域を獲得する確率が高いことにある。TCP にはさまざまな派生バージョンが存在し、TCP の輻輳制御アルゴリズムは TCP のバージョンによって異なる。このため、TCP の輻輳制御アルゴリズムが変化すれば、それによってルータ間の輻輳伝播も変化すると予想される。

そこで、TCP Tahoe ではなく、TCP Vegas [13] を用いた時の、各ルータのキュー長の時間的変動を図 12 に、各 TCP コネクションの転送レートの時間的変動を図 13 に示す。

図 4 (TCP Tahoe) および図 12 (TCP Vegas) を比較すると、TCP Vegas を用いた場合でも、ルータ間で輻輳が伝播していることがわかる。ただし、輻輳伝播の周期が大きく異なっている。また、図 4 (TCP Tahoe) および図 13 (TCP Vegas) を比較すると、TCP Vegas を用いた場合には、TCP コネクションの転送レートに周期性が見られないことがわかる。つまり、TCP Vegas を用いた場合には、TCP コネクションの転送レートが周期的に変動していないにもかかわらず、ルータ間の輻輳伝播が発生している。つまり、3 章で議論した、ルータ間の輻輳伝播の発生原因とは別の原因により、ルータ間の輻輳伝播が発生していると考えられる。これについては、今後さらなる検討が必要である。

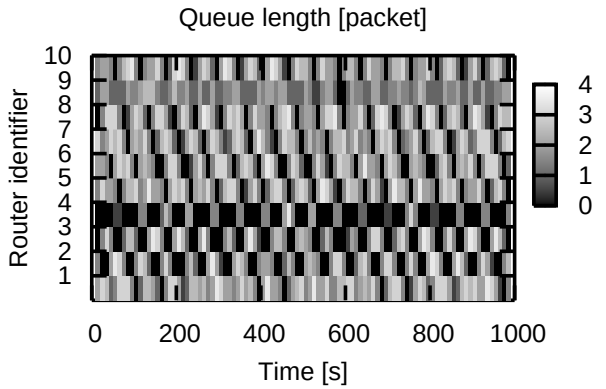

図 12: ルータのキュー長の時間的変動 (TCP Vegas の場合)

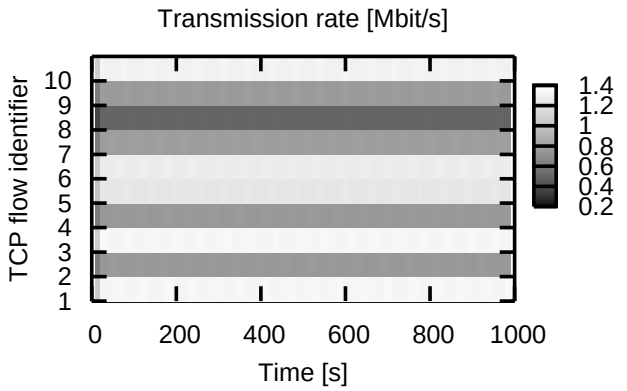

図 13: TCP コネクションの転送レートの時間的変動 (TCP Vegas の場合)

5 まとめと今後の課題

本稿では、リング型ネットワークにおいて、ルータ間の輻輳伝播が発生する原因の一つを明らかにした。また、リング型ネットワークを用い、さまざまなネットワークパラメータやシステムパラメータを変化させてシミュレーションを行った。その結果、リンクの帯域や伝搬遅延の大きさによって、ルータ間の輻輳伝播の周期が変化すること、ネットワークのランダム性が強くなるに従い、ルータ間の輻輳伝播の周期性が弱まること、などが明らかになった。

今後の課題として、ルータ間の輻輳伝播の他の発生原因、例えば、TCP Vegas を用いた場合に輻輳伝播が発生する原因を明らかにする必要がある。また、ルータ間の輻輳伝播により、TCP コネクションのエンド-エンド性能がどの程度影響を受けるのかを定量的に評価する予定である。さらに、リング型ネットワークだけでなく、より一般的なネットワークトポロジを対象とした評価を行う予定である。

文　献

[1] W. Leland, M. Taqqu, W. Willinger, and D. Wilson, “On the self-similar nature of ethernet traffic (extended version),” *IEEE/ACM Transactions on Networking*, vol. 2, pp. 1–15, Feb. 1994.

[2] M. Crovella and A. Bestavros, “Self-similarity in world wide web traffic: Evidence and possible causes,” *IEEE/ACM Transactions on Networking*, vol. 5, pp. 160–169, May 1997.

[3] A. Veras and M. Boda, “The chaotic nature of tcp congestion control,” *inProc. IEEE INFOCOM*, pp. 1715–1723, Mar. 2000.

[4] D. Newman, N. Sizemore, V. Lynch, and B. Carreras, “Growth and propagation of disturbances in a communication network model,” *in-Proceedings of th 35th Hawaii International Conference on System Sciences*, Jan. 2002.

[5] K. Fukuda, H. Takayasu, and M. Takayasu, “Spatial and temporal behavior of congestion in Internet traffic,” *Fractals*, vol. 7, pp. 23–31, 1999.

[6] J. Steger, P. Vaderna, and G. Vattay, “On the propagation of congestion waves in the Internet,” *Physica A*, vol. 359, pp. 784–792, Jan. 2006.

[7] S. McCanne and S. Floyd, “Ns (network simulator),” 1995. available at `http://www-nrg.ee.lbl.gov/ns`.

[8] S. Floyd and V. Jacobson, “On traffic phase effects in packet-switched gateways,” *Internetworking:Research and Experience*, vol. 3, pp. 115–156, Sept. 1992.

[9] S. Floyd and V. Jacobson, “Random early detection gateways for congestion avoidance,” *IEEE/ACM Transactions on Networking*, vol. 1, pp. 397–413, Aug. 1993.

[10] S. Floyd, “Recommendations on using the gentle variant of RED,” May 2000. available at `http://www.aciri.org/floyd/red/gentle.html`.

[11] J. Aweya, M. Ouellette, and D. Y. Montuno, “A control theoretic approach to active queue management,” *Computer Networks*, vol. 36, pp. 203–235, July 2001.

[12] S. Athuraliya, S. H. Low, V. H. Li, and Q. Yin, “REM: Active queue management,” *IEEE Network*, pp. 48–53, May/June 2001.

[13] L. S. Brakmo and L. L. Peterson, “TCP Vegas: End to end congestion avoidance on a global Internet,” *IEEE journal on Selected Areas in Communications*, vol. 13, pp. 1465–1480, Oct. 1995.