

# べき則の性質を有するトポロジにおける 輻輳伝播に関する一考察

平山 孝弘<sup>†</sup> 荒川 伸一<sup>†</sup> 村田 正幸<sup>†</sup>

<sup>†</sup> 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: †{t-hirayama,arakawa,murata}@ist.osaka-u.ac.jp

あらまし インターネットのトポロジを観測した結果、出線数分布がべき則に従うことが明らかにされており、ノードの出線数分布がべき則に従うトポロジを生成する様々なモデルが提案されている。しかし、出線数分布が同じでも、トポロジによって構造上の特徴が異なり、ネットワークの性能も大きく異なることが指摘されている。本稿ではインターネットにおけるパケットの混雑に着目し、既存のトポロジ生成手法によって作られたトポロジと、実際のISPトポロジにおいて、パケットの混雑状況がどのように伝播するのかを明らかにする。計算機シミュレーションの結果、BAモデルで生成されたトポロジは輻輳がネットワークに伝播しやすい傾向を持つ一方で、ISPのルータレベルトポロジは輻輳の伝播を抑制する構造を有することがわかった。

キーワード べき則、ルータレベルトポロジー、輻輳、フロー制御、BAモデル

## Congestion Propagation in ISP Topologies Having Power-law Degree Distribution

Takahiro HIRAYAMA<sup>†</sup>, Shin'ichi ARAKAWA<sup>†</sup>, and Masayuki MURATA<sup>†</sup>

<sup>†</sup> Graduate School of Information Science and Technology, Osaka University

1-5 Yamadaoka, Suita, Osaka 565-0871, Japan

E-mail: †{t-hirayama,arakawa,murata}@ist.osaka-u.ac.jp

**Abstract** Measurement studies of the Internet topology have revealed that the degree distribution follows a power-law distribution. In the power-law distribution, the probability  $P(k)$  that a node has  $k$  out-going links follows  $P(k) \sim k^{-\gamma}$ . Large networks such as social network and metabolic network also follow a power-law. However, it is known that the Internet topology structure is different from social network structure. So, a research focused only on the degree distribution is not enough to evaluate the network performance. In this paper, we investigate how the congestion propagates on the Internet topology. We examine the traffic on the BA topology obtained by topology generating method and the real Internet topology with flow-control between routers to understand how the traffic jam spreads over a network. As a result, we find that congestion propagates widely in the BA topology, while the Internet topology prevents the congestion propagation. With the comparison of their structure, congestion does not propagate widely in the topology that has highly modulated structures.

**Key words** power-law network, router-level topology, congestion propagation, flow control, BA model

### 1. はじめに

インターネットトポロジを観測した結果、ノードの出線数分布がべき則に従うことが明らかとなっている。出線数分布がべき則に従うトポロジでは、 $k$ 本のリンクを持つノードの出現確率  $P(k)$  が  $k^{-\gamma}$  ( $\gamma$ は定数) に比例する。すなわち、インターネットのトポロジは、一部のノード (以降、ハブノードと呼ぶ)

に多数のリンクが接続され、多数のノードはごく少数のリンクが接続されるという特徴を有する。

出線数分布がべき則に従うトポロジの生成手法の一つに、BAモデルがある [1]。BAモデルは、段階的なノードの追加と、新たなリンクを得る確率が既存のリンク数に比例するという2つの規則に基づくトポロジ生成手法であり、BAモデルで生成

したトポロジの特徴について、様々な研究がなされている [2]。BA モデルで生成したトポロジは、ハブノードとハブノードが隣接しており、その結果、ノード間の平均ホップ数が小さくなる特徴がある。文献 [3] では、上述の規則に基づくトポロジ生成モデルを用いて生成したトポロジにおいて、最短経路制御を用いた場合に、ノードを経由するノードペア数の分布がべき則に従うことが示されている。これは、多くのリンクを持つハブノードは、最短経路上のノードになりやすいためである。文献 [4] では、ハブノードには多くのパケットが集中するためハブノードで輻輳を起しやすいたことが述べられている。ここでいう輻輳とは、バッファに空きがなくなり新たにパケットを受け取ることができなくなった状態である。さらに、輻輳を起したノードと隣接するノード間でのフロー制御により輻輳を起したノードへのパケット転送が停止されることから、ハブノードとハブノードが近接するトポロジでは輻輳が伝播しやすいことが指摘されている。すなわち、あるノードで生じた輻輳がネットワーク全体に広がりやすく、トラフィック量の一時的な増加に対して脆弱となる。文献 [5] においても、リンク数の多いハブノードで輻輳が生じると、リンク数の少ないノードへと輻輳が伝播することが示されている。輻輳の伝播を回避する手法として、経路制御アルゴリズムにランダム性を取り入れること、および、一定時間輻輳が続いたノードのバッファからパケットを廃棄することが述べられている。

これらの文献においては、べき則に従うトポロジとして BA モデルで生成したトポロジが広く用いられている。しかし、BA モデルで生成したトポロジと現実のインターネットのトポロジとは構造が大きく異なり、トポロジ上でのネットワーク性能も大きく異なるという指摘がなされている [6]~[8]。結局のところ、出線数分布がべき則というトポロジ特性のひとつのみに着目してネットワーク性能の良し悪しを議論することは不十分であり、文献 [3] や [4] で得られているネットワークの輻輳に関する知見が、現実の ISP のルータレベルトポロジにも適用できるとは限らない。特に ISP のルータレベルトポロジにおいて輻輳がどのようなリンクで発生するのか、また、どのように輻輳が伝播するのかを知ることは、ルータの接続関係や回線容量などの効率のよい設計手法へと応用することが期待される。

本稿では、出線数分布がべき則に従うルータレベルトポロジにおいて、どのようなリンクで輻輳が発生しやすいかを明らかにする。次に、あるリンクで輻輳が生じた場合に、その輻輳がネットワークにどのように波及するのかを計算機シミュレーションによって明らかにする。トポロジ構造の観点から評価を行った結果、BA モデルで生成されたトポロジは輻輳がネットワーク全体に伝播しやすい性質を持つ一方で、ISP のルータレベルトポロジは輻輳の伝播を抑制する構造を有することを示す。

本稿の構成は以下の通りである。2. 章で本稿で行ったシミュレーションのモデル、3. 章でその結果について述べる。最後に 4. 章でまとめと今後の課題を述べる。

## 2. ネットワークモデル

ここでは、本稿で使用するシミュレーションのモデルについ

て説明する。

### 2.1 ネットワークトポロジ

ネットワークトポロジは、現実の ISP トポロジとして米国の通信会社である AT&T 社 (523 ノード、2608 リンク) と Sprint 社 (467 ノード、2560 リンク) のルータレベルトポロジを用いる。これらのトポロジは Rocketfuel ツール [9] を用いて計測されたものであり、ノードの出線数分布はべき則に従う [8]。また、比較のために、2 社のトポロジとノード数およびリンク数が等しいトポロジを BA モデルに生成した。以降では、AT&T と同一のノード数・リンク数を有するトポロジを BA AT&T トポロジと呼ぶ。

### 2.2 トラフィックモデル

各ノードでは、ポアソン過程に従ってパケットが生成されるものとし、その宛先ノードはランダムに選ぶものとする。 $R$  を 1 単位時間あたりにネットワークに到着するパケット数、 $N$  をノード数とし、指数分布の到着率  $\lambda$  は、 $R/N$  とする。次章において、 $R$  を変化させることでネットワークで生じるパケット数を変化させることでネットワークの輻輳がどのように変化するのかを明らかにする。

### 2.3 ノード処理モデル

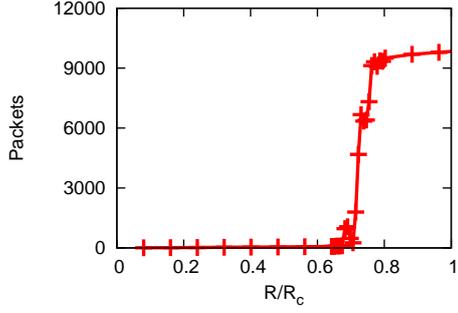
ネットワーク内の各ノードは出力リンク毎に有限長のバッファを持ち、バッファサイズはすべて等しく 10 パケットであるものとする。パケットを受け取ったノードは、自身が宛先ノードである場合、そのパケットをネットワークから破棄する。自身が宛先ノードでない場合、経路制御アルゴリズムによって定まる次ホップのノード向けの出力リンクのバッファに格納される。ただし、バッファに空きが無ければパケットの受信を拒否する。この時、パケットは前ホップのノードでの出力リンクのバッファに格納する。各ノードでは、出力リンクの空き状況を常に観測しているものとし、出力リンクに空きが生じた場合には FIFO (First-In First-Out) に基づいてバッファ中のパケットを次ホップのノードへ転送する。各リンクの伝搬遅延時間は 1 単位時間とする。

## 3. シミュレーション評価

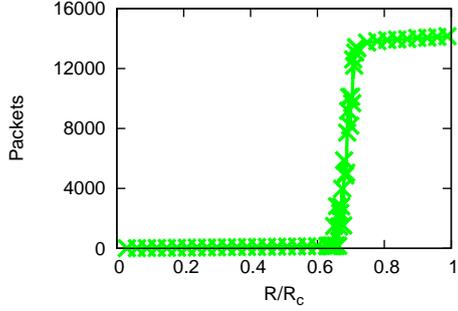
本章では、現実の ISP トポロジおよび従来のトポロジ生成手法に基づくトポロジを対象とした計算機シミュレーションを行い、ルータ間のフロー制御がある場合において輻輳がどのように伝播するのかを明らかにするとともに、輻輳が伝播しにくいトポロジ構造の特徴を明らかにする。計算機シミュレーションでは、以下の 3 つの経路制御アルゴリズムを適用し、経路制御の違いにより輻輳伝播がどのように変化するのかについても明らかにする。

- 最短経路制御
- 最小遅延時間経路制御
- next-nearest-neighbor (nnn) サーチ経路制御 [4]

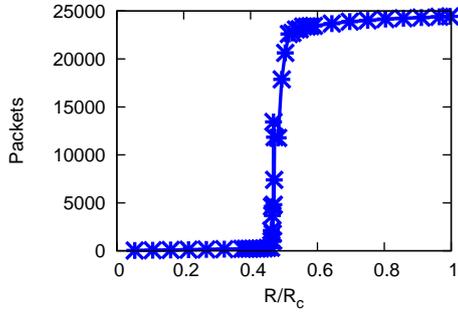
nnn サーチは、文献 [4] で用いられている経路制御手法であり、自身が宛先ノードでなければ自身の周辺数ホップの範囲のノードを探索し、宛先ノードが見つければそこまでの最短経路を、見つからなければランダムに選んだ隣接ノードへパケットを転



(a) AT&amp;T トポロジ

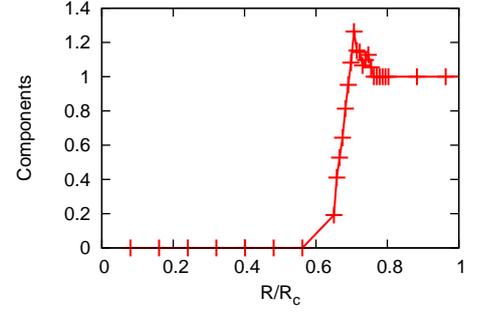


(b) Sprint トポロジ

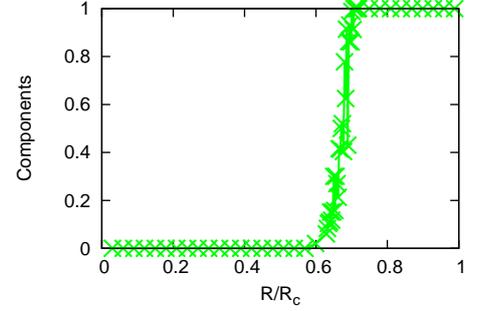


(c) BA AT&amp;T トポロジ

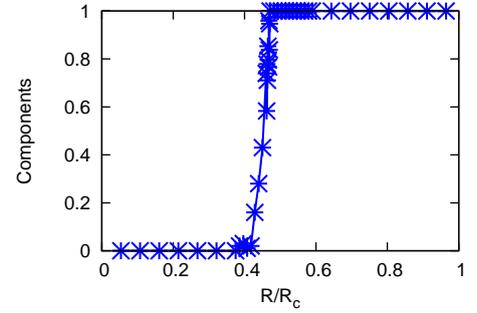
図1 最短経路制御での平均パケット数の推移



(a) AT&amp;T トポロジ



(b) Sprint トポロジ



(c) BA AT&amp;T トポロジ

図2 最短経路制御での平均コンポーネント数の推移

送する。

計算機シミュレーションによる評価では、ネットワークのパケット発生量  $R$  を変化させ、定常状態での平均系内パケット数の推移によりネットワークの輻輳状況を明らかにする。なお、パケット発生量  $R$  がある値  $R_{max}$  を超えると、トポロジを構成する大部分のリンクが輻輳し、パケットが転送できない状態に陥る。 $R_{max}$  は適用する経路制御に依存するため、最短経路を使用した場合のパケット発生量の閾値  $R_c$  によって正規化する。 $R_c$  は、リンクを経由するノードペア数 (Betweenness centrality) の最大値  $B_{max}$  をもとに、以下の式で算出する [10]。

$$R_c = \frac{CN(N-1)}{B_{max}}, \quad (1)$$

ここで、 $C$  はリンクの伝播遅延であり、 $N$  はノード数である。各トポロジの  $B_{max}$  および  $R_c$  を表1に示す。

また、トポロジの輻輳状況を評価するもう一つの指標として、定常状態での、輻輳が生じているリンク、および、そのリンクに連結したノードを抽出することで得られるコンポーネント

表1 各トポロジにおける  $B_{max}$  と  $R_c$  の値

	$B_{max}$	$R_c$
AT&T	21916	12.457
Sprint	6519	33.383
BA AT&T	2926	93.303

(部分グラフ) の数を用いる。異なるコンポーネントに含まれるリンクの間では輻輳が伝播していないと考えることができる。従って、多くのコンポーネントが形成されていれば、離れた場所で輻輳が発生しているものの、輻輳が伝播していない状況であると考えられる。また、多くのリンクが輻輳し系内パケット数が増加している際にコンポーネントの数が少ない場合には、あるリンクで起こった輻輳が広域に伝播していると判断できる。

### 3.1 最短経路制御

最短経路制御では、全てのパケットは常に宛先ノードまでのホップ数が最小となる経路を辿る。経路上のリンクが輻輳している場合は前段のリンクのバッファに格納される。

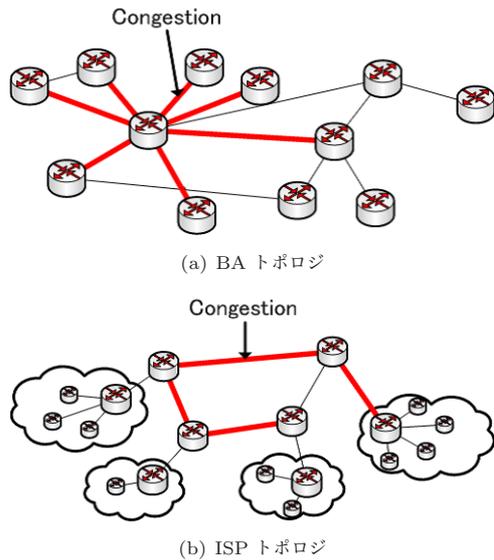
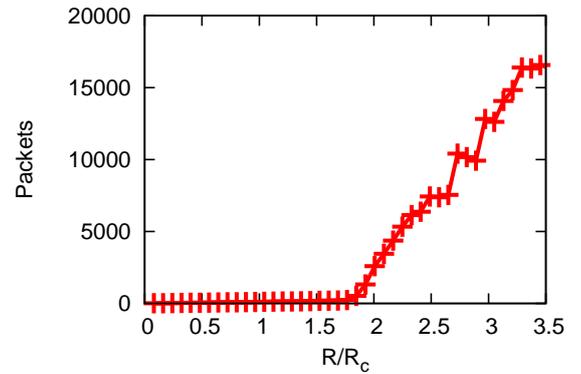


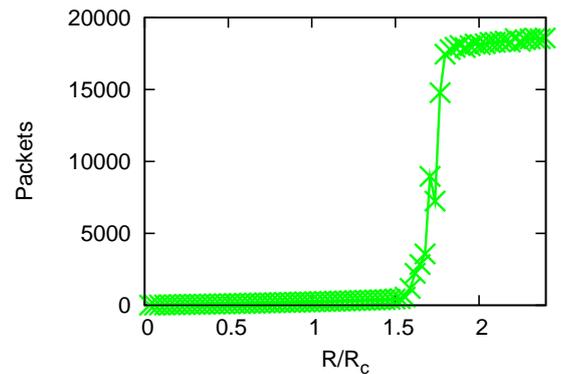
図3 トポロジのモデル図

図1に、正規化されたパケット発生量  $R/R_c$  を変化させた際の、平均系内パケット数を示す。図中の数値はシミュレーション時間 9000 から 10000 の間の系内パケット数の平均値である。また、同様に平均コンポーネント数を算出した結果が図2である。 $R$  が十分に小さい場合、輻輳が発生せず、パケットは滞りなく最短時間で宛先ノードまで届けられるため、系内パケット数も小さくなる。 $R$  が大きくなるとともに、いくつかのリンクで輻輳が発生し、輻輳によりパケットの転送が滞るため宛先ノードまでの転送時間も長くなる。パケットの系内滞在時間が長くなるため、結果として  $R$  が上昇すると系内パケット数も上昇する。さらに  $R$  が上昇すると、パケットが集中しやすいリンクでは1パケット処理するとすぐに次のパケットが到着し、輻輳が解消できない状態に陥る。輻輳が解消できなければ、輻輳が前ホップのリンクへと伝播してパケットの転送が滞り、系内パケット数が多くなる。なお、図1および2を見ると、すべてのトポロジにおいて系内パケット数が急激に増加する  $R/R_c$  の値は、1よりも低い。これは、 $R_c$  の算出の際に無限長のバッファを仮定している一方で、シミュレーション時には有限バッファとしているためである。

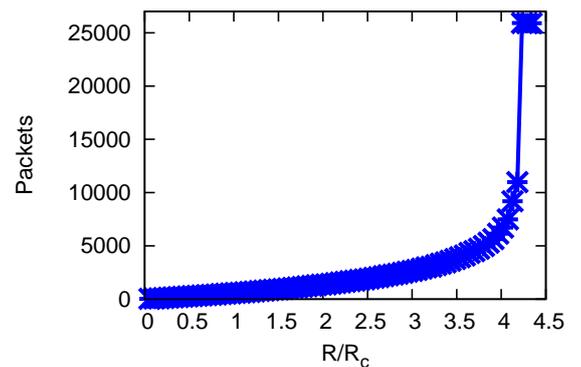
図2を見ると、 $R$  が大きくなり系内パケット数が増加するとともに、平均コンポーネント数も急激に増加していることがわかる。しかし、図2(a)と、図2(b)・図2(c)の結果を比較すると、AT&Tでは平均コンポーネントが1を超えることがあるのに対し、図2(b)・図2(c)では常に1以下になっていることがわかる。これは、BAトポロジでは、 $B_{max}$  を与えるリンクに隣接しているノードはハブノードであり、さらにハブノードに多くのパケットが集中するためである。すなわち、BAトポロジではハブノードと隣接するリンクで輻輳が生じやすく、ハブノードを中心とするコンポーネントが生成される(図3(a))。一方AT&Tトポロジでは、 $B_{max}$  を与えるリンクは大都市間を連結するリンクであり、地域間でパケットを転送する際にその経路上のリンクで一時的に輻輳が生じることはあるものの、そ



(a) AT&T トポロジ



(b) Sprint トポロジ



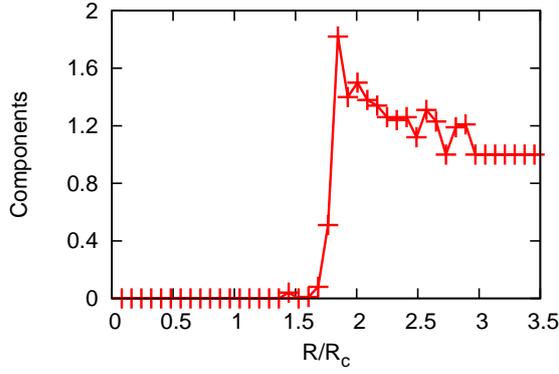
(c) BA AT&T トポロジ

図4 遅延時間を最小にする経路制御でのパケット数の推移

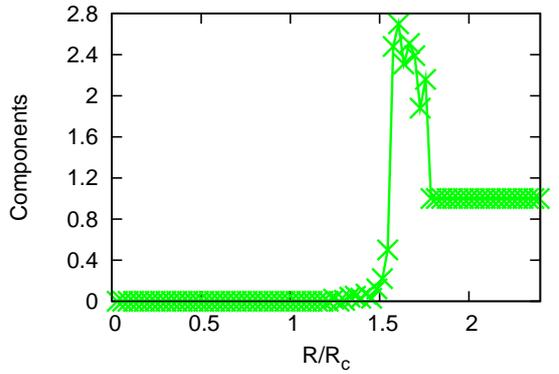
の輻輳は必ずしも他の地域に伝播するものではない(図3(b))。従ってAT&Tトポロジでは複数箇所輻輳が生じる可能性があり、結果として平均コンポーネント数が1を超える状況が発生している。

### 3.2 最小遅延時間経路制御

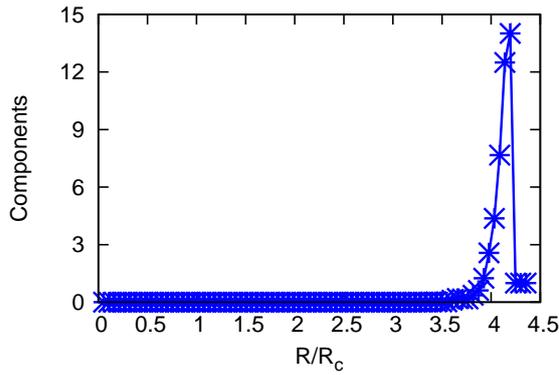
次に、宛先ノードまでの遅延時間が最小となる経路制御(最小遅延時間経路制御)を適用した際のネットワークの輻輳状況について評価を行った。最小遅延時間経路制御では、ノードは全リンクのバッファ内のパケット数を常に把握できるものと仮定し、バッファ内に待機しているパケットの数と残存ホップ数から最も短い時間で宛先ノードに到達可能な経路を選択する。



(a) AT&T トポロジ



(b) Sprint トポロジ

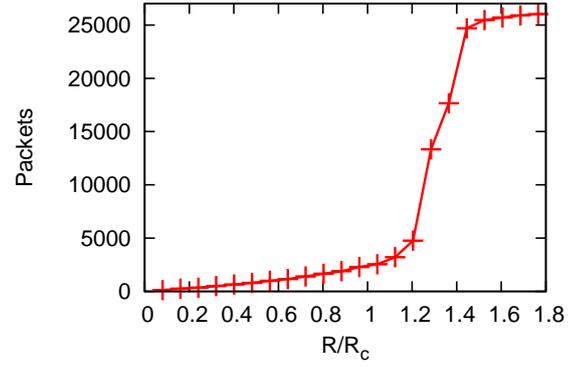


(c) BA AT&T トポロジ

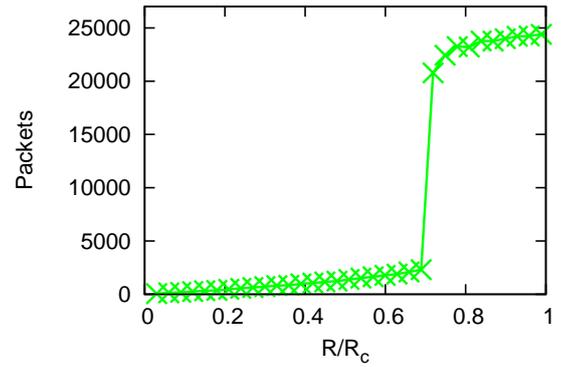
図5 遅延時間を最小にする経路制御でのコンポーネント数の推移

図4を見ると、どのトポロジでも、 $R/R_c$ が1を超えていても、系内パケット数は低く、最短経路制御と比較して多くのトラフィックを収容できることがわかる。これは、より短い時間でパケットを転送できる経路へ迂回され、輻輳が生じているリンクが回避されるためである。

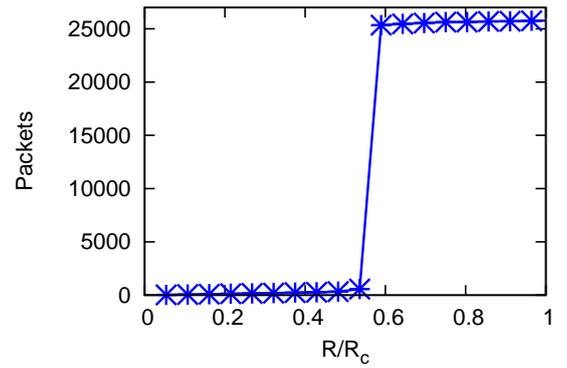
ISPトポロジではパケット発生率 $R$ が $R_c$ の1.5倍を超えると系内パケット数が飽和しパケットの転送ができなくなるのに対し、BAトポロジでは $R$ が $R_c$ の4倍以下であればパケットを転送することが可能であり、BAトポロジの方がより多くのパケットを転送することができる。この要因として、ルータ、回線を配置するISPトポロジでは、ルータ



(a) AT&T トポロジ



(b) Sprint トポロジ



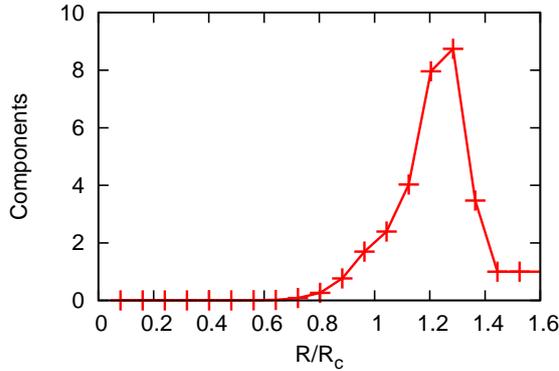
(c) BA AT&T トポロジ

図6 nnn サーチを用いた場合のパケット数の推移

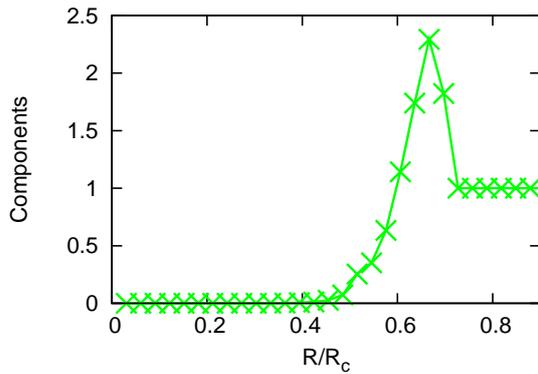
のポート数の制限もしくは回線費用などの技術的・経済的な制約があるため、地域間のリンク数が少なく、地域内リンク数が多いことが挙げられる。ISPトポロジでは地域間のリンクで輻輳が生じると、トラフィックを迂回することができなくなるため収容可能なトラフィック量が少なくなる。

### 3.3 nnn サーチ

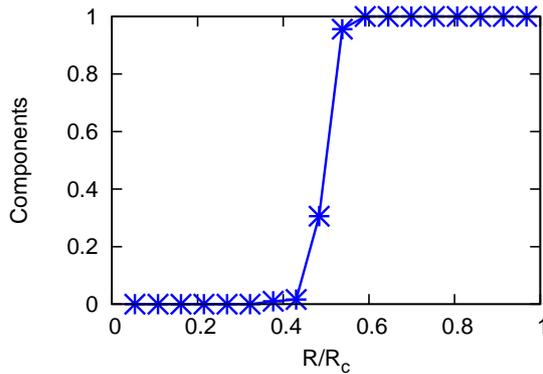
文献[4]では、自身の周辺数ホップの範囲のみを検索し、パケットの宛先となるノードが見つからなければランダムに選んだ隣接ノードへパケットを送る経路制御アルゴリズム nnn が提案されており、ハブノードのホップ間距離が短いトポロジにおける有効性が示されている。図6および図7は、nnn を経路制



(a) AT&T トポロジ



(b) Sprint トポロジ



(c) BA AT&T トポロジ

図7 nnn サーチを用いた場合のコンポーネント数の推移

御アルゴリズムとして使用した場合の、各トポロジにおけるパケット発生量  $R$  と平均系内パケット数および平均コンポーネント数を示したものである。図7を見ると、いずれのトポロジにおいても系内トポロジ数が増加するとともに、平均コンポーネント数も増加することがわかる。特にAT&Tトポロジでは、他のトポロジと比較して平均コンポーネント数が多くなっている。nnnを適用した場合、AT&Tトポロジでは、輻輳が生じるリンクは地域内のリンクとなっており、コンポーネントの位置が地理的に分散しているためである。

#### 4. まとめと今後の課題

本稿では、既存のトポロジ生成手法であるBAモデルによって生成したトポロジと、ISP ルータレベルトポロジにおいて、どのようなリンクで輻輳が起こりやすく、その輻輳がネットワークにどのように波及するのかを計算機シミュレーションによる比較を行った。そして、BAモデルで生成されたトポロジは輻輳がネットワーク全体に伝播しやすい性質を持つ一方で、ISPのルータレベルトポロジは輻輳の伝播を抑制する構造を有することがわかった。

今後は、エンドホスト間フロー制御およびリンク毎に異なる回線容量を定めたモデルでの評価を行う必要がある。

#### 謝 辞

本研究の一部は、文部科学省科学研究費基盤研究(A)18200004によっている。ここに記して謝意を表す。また、本研究の遂行にあたり、有意義な議論をしていただいたNTTコミュニケーション科学基礎研究所の新井賢一氏に感謝いたします。

#### 文 献

- [1] A.-L. Barabási and R. Albert, "Emergence of scaling in random networks," *Science*, vol. 286, pp. 509–512, Oct. 1999.
- [2] R. Cohen, S. Havlin, and D. Avraham, *Handbook of Graphs and Networks – From the Genome to the Internet*, ch. 4. WILEY-VCH GmbH & Co., 2003. Structural Properties of scale-free networks.
- [3] K.-I. Goh, B. Kahng, and D. Kim, "Universal behavior of load distribution in scale-free networks," *Physical Review Letters*, vol. 87, Dec. 2001.
- [4] B. Tadić, S. Thurner, and G. Rodgers, "Traffic on complex networks: Towards understanding global statistical properties from microscopic density fluctuations," *Physical Review E*, vol. 69, Mar. 2004.
- [5] Z. X. Wu, W. X. Wang, and K. H. Yeung, "Congestion diffusion and decongestion strategy in networked traffic," arXiv:0711.1655v1, Nov. 2007.
- [6] L. Li, D. Alderson, W. Willinger, and J. Doyle, "A first-principles approach to understanding the Internet's router-level topology," *ACM SIGCOMM Computer Communication Review*, vol. 34, pp. 3–14, Oct. 2004.
- [7] T. Bu and D. Towsley, "On distinguishing between Internet power law topology generators," in *Proceedings of Twenty-First Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM 2002)*, vol. 2, pp. 1587–1596, June 2002.
- [8] R. Fukumoto, S. Arakawa, T. Takine, and M. Murata, "Analyzing and modeling router-level Internet topology," in *Proceedings of The International Conference on Information Networking (ICOIN)*, Jan. 2007.
- [9] N. Spring, R. Mahajan, D. Wetherall, and T. Anderson, "Measuring ISP topologies with rocketfuel," *IEEE/ACM Transactions on Networking*, vol. 12, pp. 2–16, Feb. 2004.
- [10] L. Zhao, Y.-C. Lai, K. Park, and N. Ye, "Onset of traffic congestion in complex networks," *Physical Review E*, vol. 71, Feb. 2005.