

特別研究報告

題目

局所情報に基づく高信頼トポロジ生成手法の提案

指導教員

村田 正幸 教授

報告者

倉田 園子

平成 19 年 2 月 20 日

大阪大学 基礎工学部 情報科学科

内容梗概

インターネット利用者数の増大によるネットワークの大規模化にともない、分散ネットワーク制御の重要性が高まっている。その一方で、大規模なネットワークにおいて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。べき則に従うネットワークとは、隣接ノード数が k であるノードの出現確率 $P(k)$ が $k^{-\gamma}$ (γ は定数) に比例するネットワークである。この場合、一部のノードに多くのリンクが接続される一方で、他の多くのノードには少数のリンクが接続される。

出線数分布がべき則に従うトポロジ生成手法のひとつに BA モデルがある。BA モデルとは、ノードを一つずつ追加する段階的ノード追加、既存のトポロジの出線数情報に基づいた優先的リンク追加の二つの規則に基づくトポロジ生成手法である。BA モデルで生成されたトポロジのネットワーク性能として、ノード間のホップ間距離が小さくなるスモールワールドの性質を有することが明らかにされている。また、BA モデルで生成されるトポロジはノードにランダムに発生する障害には耐性があるものの、ノードの出線数の降順に発生する障害には脆弱であることが示されている。

BA モデルでは、ノードをトポロジに追加する際に、既存のトポロジの全てのノードの出線数情報 (大域情報) が必要となるが、ネットワークのノード数が大きくなるとともに大域情報の取得が困難となる。それにも関わらず、大規模ネットワークにあらわれるべき則に関する多くの研究では、大域情報の取得を前提としたトポロジ生成手法を対象としてネットワーク性能の評価が行われてきた。一部の研究では、局所的な範囲の情報もしくは制限された情報 (局所情報) に基づいてべき則に従うトポロジを生成する手法の提案がなされている。しかし、これらの研究では、出線数分布がべき則に従うトポロジ生成が目的とされており、局

所情報を用いて生成したトポロジが有するトポロジ構造ならびにネットワーク性能はほとんど評価されていない。また、同一の出線数分布を有するトポロジに対するネットワークの性能評価が行われており、ネットワークの性能はトポロジ構造に依存していることが明らかにされている。そのため、局所情報を用いたトポロジ生成手法の提案のみでは不十分であり、得られるトポロジのネットワーク性能を高めるために必要となる局所情報ならびに局所情報を用いたトポロジ生成手法を明らかにすることが重要である。

そこで本報告では、分散制御の適用が必須であるような大規模ネットワークを対象とし、局所情報を用いたトポロジ生成手法を提案する。その際に、BA モデルで生成されるトポロジのホップ間距離に関する利点およびランダム障害への耐性を生かすために出線数分布に関するべき則を保持しつつ、アタック障害に関する脆弱性を補うことでネットワークの信頼性が向上するトポロジを生成する。提案手法では、BA モデルの利点を生かすために BA モデルの生成モデルを基に、局所情報としてノード負荷を用いたトポロジ生成手法と、クラスタリング手法を適用して定めたノードのクラスタ番号を局所情報として用いたトポロジ生成を行った。アタック障害に対する耐性を BA モデルで生成したトポロジと比較した結果、ノード負荷を局所情報として用いたトポロジ生成手法ではおよそ 100%、クラスタ番号を局所情報として用いたトポロジ生成手法ではおよそ 70%耐性が向上することが明らかとなった。その一方で、ランダム障害に対する耐性は、ノード負荷を局所情報として用いたトポロジ生成手法ではおよそ 22%、クラスタ番号を局所情報として用いたトポロジ生成手法ではおよそ 4%劣化することも明らかとなった。

主な用語

べき則、大域情報、局所情報、クラスタリング手法、アタック障害、ランダム障害

目次

1	はじめに	5
2	関連研究	8
2.1	大域情報に基づくトポロジ生成	8
2.2	局所情報に基づくトポロジ生成	9
3	局所情報に基づくトポロジ生成手法の提案と評価	10
3.1	局所情報	10
3.2	局所情報を用いたトポロジ生成手法	10
3.3	生成されたトポロジの特性評価	11
4	クラスタリング手法に基づくトポロジ生成手法の提案と評価	23
4.1	クラスタリング	23
4.2	クラスタリング手法に基づくトポロジ生成手法	23
4.3	生成されたトポロジの特性評価	25
5	まとめ	33
	謝辞	34
	参考文献	35

目 次

1	局所情報に基づくトポロジの出線数分布： $j = 2$ の場合	13
2	局所情報に基づくトポロジの出線数分布： $j = 3$ の場合	14
3	局所情報に基づくトポロジの出線数分布： $j = 4$ の場合	15
4	局所情報に基づくトポロジの出線数分布： $j = 5$ の場合	16
5	局所情報に基づくトポロジの出線数分布： $j = 10$ の場合	17
6	局所情報に基づくトポロジの出線数分布： $j = 200$ の場合	18
7	局所情報に基づくトポロジのアタック障害への耐性評価	19
8	局所情報に基づくトポロジのランダム障害への耐性評価	20
9	局所情報に基づくトポロジ：ノード数が 1000 個の場合	21
10	局所情報に基づくトポロジ：アタック障害の耐性が悪化する例	22
11	クラスタリング手法に基づくトポロジの出線数分布： $j = 2$ の場合	27
12	クラスタリング手法に基づくトポロジの出線数分布： $j = 5$ の場合	28
13	クラスタリング手法に基づくトポロジの出線数分布： $j = 10$ の場合	29
14	クラスタリング手法に基づくトポロジの出線数分布： $j = 100$ の場合	30
15	クラスタリング手法に基づくトポロジの出線数分布： $j = 500$ の場合	31
16	クラスタリング手法に基づくトポロジのアタック障害への耐性評価	32
17	クラスタリング手法に基づくトポロジのランダム障害への耐性評価	32

1 はじめに

インターネット利用の普及にともないネットワークが社会に浸透し、SNS サービス・P2P サービスなどの様々なアプリケーションがインターネット上で展開されている。これらのアプリケーションではルータの連結によって構築された物理的なトポロジ上にユーザによる論理的なトポロジを構築することから、アプリケーションにとっての効果的かつ効率的な論理トポロジを構築することが課題となっている。また、インターネット利用者数の増大によるネットワークの大規模化にともない、分散型のトポロジ構築の重要性が高まっている。その一方で、大規模なネットワークにおいてノードの出線数分布がべき則に従うことが観測されている [1]。出線数分布に関するべき則とは、ノードの出線数が k となる確率 $P(k)$ が、 $\alpha \times k^{-\gamma}$ となることである (α, γ は定数)。

出線数分布がべき則に従うトポロジ生成手法のひとつに BA モデルがある [2]。BA モデルは、(1) ノードを一つずつ追加する段階的ノード追加 (Incremental Growth)、(2) 既存のトポロジの出線数情報に基づいた優先的リンク追加 (Preferential Attachment)、の二つの規則に基づいてノードおよびリンクを追加するトポロジ生成手法である。BA モデルで生成されたトポロジを対象として、ネットワーク性能の評価がなされており [3]、ノード間のホップ間距離が小さくなるスモールワールドの性質を有することが明らかにされている。BA モデルで生成されるトポロジの障害に対する信頼性を評価した結果、ノードにランダムに障害が発生する場合 (ランダム障害) には耐性があるものの、ノードの出線数の降順に障害が発生する場合 (アタック障害) には脆弱であることが示されている [4-7] BA モデルの適用例として、P2P ネットワークのトポロジ生成がある。P2P ネットワークにおけるピアをノード、コネクションをリンクに見立てることで、P2P ネットワークのトポロジを BA モデルの成長モデルで説明できることが示されている [8]。

BA モデルでは、ノードをトポロジに追加する際に、既に生成されたトポロジの全てのノードの出線数情報 (大域情報) を用いて接続先ノードを決定する。このことは、例えば BA モデルによるトポロジ生成を P2P ネットワークに適用した場合、新規参加ピアの接続先を決定する際に P2P ネットワークに参加中の全てのピアの出線数情報を用いることに相当する。しかし、P2P ネットワークにおけるピアは他のすべてのピアの出線数情報を取得する

ことはない。また、ネットワークのノード数が大きくなるとともに大域情報の取得を前提としたトポロジ生成は困難となる。それにもかかわらず、大規模ネットワークにあらわれるべき則に関する多くの研究では、大域情報の取得を前提としたトポロジ生成手法を対象としており、大域情報を用いて生成されたトポロジ上でのネットワーク性能の評価が行われてきた [2, 9–12]。一部の研究では、局所的な範囲の情報もしくは制限された情報に基づいてべき則に従うトポロジを生成する手法に関する研究がなされている [13, 14]。しかし、これらの研究では、BA モデルと同等の出線数分布を有するトポロジ生成が目的であり、局所的な範囲の情報を用いて生成したトポロジが有するトポロジ構造ならびにネットワーク性能はほとんど評価されていない。

文献 [15] では同一の出線数分布を有するトポロジをいくつか列挙し、トポロジの構造に依存してネットワーク性能が大きく異なることを示している。すなわち、単に局所情報を用いて BA モデルと同等の出線数分布を有するトポロジを生成するのみではアプリケーションにとって効率的・効果的な論理トポロジを構築するには不十分であり、得られるトポロジのネットワーク性能を高めるために必要となる局所情報ならびに局所情報を用いたトポロジ生成手法を明らかにすることが重要である。

そこで本報告では分散制御の適用が必須であるような大規模ネットワークを対象とし、局所情報を用いた高信頼トポロジ生成手法を提案する。ここでいう高信頼とは、BA モデルで生成されるトポロジの性質として明らかとなっているアタック障害に対する耐性が高いことである。提案手法では単に局所情報を用いて BA モデルで生成されるトポロジと同等の出線数分布を有するトポロジを生成するのではなく、BA トポロジのホップ間距離に関する利点活かすために出線数分布に関するべき則を保持しつつ、アタック障害に関する脆弱性を補うことでネットワークの信頼性が向上するトポロジを生成する。本報告では、局所情報としてノード負荷を用いたトポロジ生成手法と、クラスタリング手法を適用して定めたノードのクラスタ番号を局所情報に用いたトポロジ生成手法を比較評価する。アタック障害に対する耐性を BA モデルで生成したトポロジと比較した結果、ノード負荷を局所情報として用いたトポロジ生成手法ではおよそ 100%、クラスタ番号を局所情報として用いたトポロジ生成手法ではおよそ 70%向上することが明らかとなった。その一方で、ランダム障害に対する耐性は

ノード負荷を局所情報として用いたトポロジ生成手法ではおよそ 22%、クラスタ番号を局所情報として用いたトポロジ生成手法ではおよそ 4%劣化することも明らかとなった。

本報告の内容は以下の通りである。2章において大域情報に基づくトポロジ生成、局所情報に基づくトポロジ生成に関する関連研究の内容を述べる。3章、4章ではそれぞれ局所情報に基づくトポロジ生成手法、クラスタリング手法に基づくトポロジ生成手法を提案し、それぞれのトポロジの特性および障害への耐性を評価する。最後に本報告のまとめと今後の課題を5章で述べる。

2 関連研究

2.1 大域情報に基づくトポロジ生成

大域情報に基づくトポロジ生成手法のひとつに、BA (Barabási-Albert) モデルがある [2]。BA モデルは、(1) ノードを一つずつ追加する段階的ノード追加、(2) 既存のトポロジの出線数情報に基づいた優先的リンク追加、の二つの規則に基づいてノードおよびリンクを追加するトポロジ生成手法である。

BA モデルでは、ノード i の出線数 k_i が大きい程、ノード i へのリンクが接続される確率 $\Pi(k_i)$ が高くなるように接続確率を与えている。文献 [2] では以下の方法でトポロジを生成する。

Step 1 : 初期ノードとして m_0 個のノードを配置する。

Step 2 : ネットワークのノード数が N 未満である場合、Step 3 へ。ノード数が N であればトポロジ生成を終了する。

Step 3 : ノードを 1 個追加する。

Step 4 : 追加されたノードから式 (1) の確率に従い、他の m ($\leq m_0$) 個の相異なるノードへリンクを接続する。Step 2 へ。

$$\Pi(k_i) = k_i / \sum_j k_j \quad (1)$$

BA モデルで生成されるトポロジには、多くのリンクが集中するハブと呼ばれるノードが存在し、出線数分布がべき則に従うというスケールフリー (Scale-free) の性質を有することが示されている。さらに、ノード間の最小ホップ数が小さくなるスモールワールド (Small-world) の性質を有することも明らかとなっている [16]。また、ランダムに発生する障害に対しては耐性があるものの、多くのリンクが集中するハブに対する意図的な障害 (アタック障害) には脆弱であることが示されている [7]。

BA モデルに代表される大域情報を用いたトポロジ生成手法を用いてトポロジを構築することは、ネットワークのノード数が多くなるとともに、そのモデルを用いたトポロジ生成は非現実的となる。したがって大域情報の取得を必要としない局所情報に基づいたトポロジ生成手法が必要である。

2.2 局所情報に基づくトポロジ生成

大域情報に基づいてトポロジを生成する BA モデルに対し、局所的な範囲の情報に基づいてべき則に従うトポロジを生成する手法に関する研究がなされている [13,14]。いずれの文献においても、BA モデルと同様の、ノードを一つずつ追加する段階的追加の規則を保持するものの、リンク接続の規則が BA モデルと異なる。文献 [13] では、各ノードに類似度をあらわす指標を導入し、類似度が一致しているノード集合に対してのみ BA モデルと同様の優先的リンク接続を適用するトポロジ生成手法を提案している。類似度を一様分布で与えて評価を行った結果、優先リンク接続の対象ノード集合を、全体のノード数に対して 70%程度確保すれば、BA モデルで生成したトポロジと出線数分布および平均ホップ長に関して同等の性質を有することが示されている。優先リンク接続の対象ノード集合が少ない場合においては、出線数分布にべき則が現れるものの平均ホップ長が大きくなることが明らかとなっている。

文献 [14] では、ノード追加時に新規リンクの接続先を決定する際に、既存のトポロジの構造に関する情報を用いることなく BA モデルに似たトポロジを生成する手法を示している。具体的には、既存のトポロジからランダムにノード (v とする) を選択し、確率 β で v にリンクを接続する。確率 $1 - \beta$ でノード v の隣接ノード集合からランダムにノードを 1 つ選び、そのノードにリンクを接続する。 β が 1.0 の場合は出線数分布は指数分布となる一方で、 β が 0.25 未満の場合に出線数分布に関してべき則の性質が現れることを解析的に示している。

しかし、いずれの文献においても、出線数分布がべき則に従うトポロジを生成することが目的となっており、生成されたトポロジに対する局所情報を用いて生成したトポロジのネットワーク特性、特にノード障害への耐性については評価されていない。

3 局所情報に基づくトポロジ生成手法の提案と評価

本章では、局所情報になりうる値について述べ、局所情報として、ノードの負荷情報に相当する Betweenness Centrality を用いたトポロジ生成手法の提案と評価について述べる。

3.1 局所情報

既存のネットワークに新たに追加したノードのリンクの接続先を決定する際に、ネットワーク全体の情報 (大域情報) を用いて最適な接続先を探すのではなく、分散制御化のために接続先候補ノードを選出する必要があるが、どのようにその範囲を定めるかが問題となる。また、障害への耐性を高めるために、ノードでどのような局所情報を保持すればよいのか、および、その情報を使ってどのようにリンクの接続先を決定すれば良いのか、ということも考慮する必要がある。

前者に関しては、ユークリッド距離が近いノードに限定するアプローチと、文献 [14] のようにノードをランダムに選択し、その近傍の情報を取得してその中から選択するアプローチが考えられる。後者に関しては、保持すべき情報として、既存の研究で広く用いられている出線数、ノードの負荷情報に相当する Betweenness Centrality [17]、分散クラスタリング手法 [18] により得られるクラスタ番号が考えられる。

ノードが保持する情報のうち、ノード負荷を用い、負荷の大きいノードに優先的にリンクを接続することにより、ノード間のホップ数が小さくなることが考えられる。そのことに加え、ユークリッド距離が近いノードをリンク接続先候補とするアプローチを用いることにより、ノード負荷の分散が可能となり、負荷の大きいノードに障害が発生した際のネットワーク接続性の急激な悪化を阻止できると考える。そこで、ノード負荷に相当する Betweenness Centrality を用いたトポロジ生成を行うこととする。

3.2 局所情報を用いたトポロジ生成手法

べき則に従うトポロジ生成モデルの一つである、BA (Barabási-Albert) モデル [2] では、トポロジを生成する際に大域情報が必要となるが、このモデルの特性のひとつである、ス

モールワールドの性質は、アプリケーションの性能の効率化を考える際に重要となる。そこで、BAモデルのトポロジ生成手法を基に、上記の、ユークリッド距離が近いノードをリンク接続先候補とするアプローチとノードの負荷情報に相当する Betweenness Centrality を局所情報として用いるアプローチを組み合わせ、局所情報を用いたトポロジ生成手法を提案する。ただし、新たに追加したノードがリンクの接続先を決定する前に、既存のネットワークの探索を行い、既存のノード中のあるノードに情報提供を求め、そのノードの保持する情報(局所情報)を得るということを前提とする。

以上のことを考慮し、提案手法では文献 [2] の生成手法に加え、各ノードに座標を与え、新たなノードを追加した際に既存のノードとのユークリッド距離を求め、リンク接続先候補を近隣 $j (\geq m)$ 個のノードとする。ただし、実ネットワークにおいては、ユークリッド距離を求める必要はなく、ノードが保持している局所情報によってこの情報が得られることを前提とする。

さらに、BAモデルのトポロジ生成では指標が各ノードの出線数となっているが、提案手法では指標を Betweenness Centrality とし、ノード i の Betweenness Centrality の値 l_i が大きい程、ノード i へのリンクが接続される確率 $\Pi(l_i)$ が高くなるように式 (2) の確率を与え、その確率に従って、他の $m (\leq m_0)$ 個の相異なるノードへリンクを接続することとした。確率 $\Pi(l_i)$ が高いノード i に優先的にリンクを接続することによって、ハブノードが現れやすくなり、各ノード間の平均ホップ数が小さくなる。

$$\Pi(l_i) = l_i / \sum_j l_j \quad (2)$$

3.3 生成されたトポロジの特性評価

生成されたトポロジにおける特性評価として、BAモデルのトポロジとの比較、および、アタック障害、ランダム障害への耐性の評価を行った結果を述べる。評価にはノード数1000個で生成したトポロジを用い、提案手法においてはリンク接続候補数 j の値を2から5、10、200と変化させて生成したトポロジを用いた。

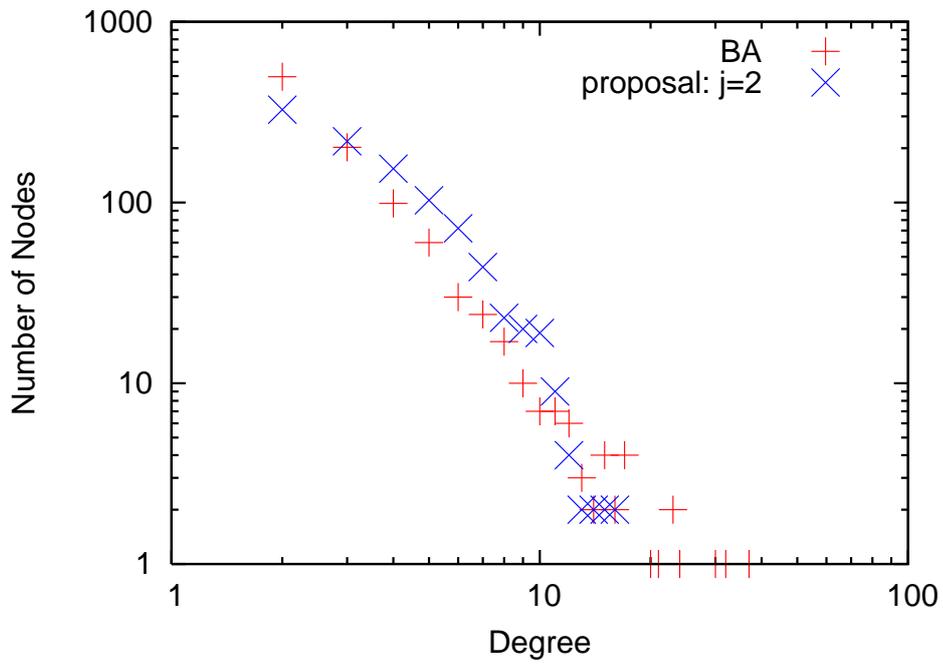
BA モデル、提案手法それぞれから生成されたトポロジの出線数分布、出線数の累積分布の比較を図 1 から図 6 に示す。これらの図 (a) の横軸がノードの出線数、縦軸がその出線数を持つノード数を示す。図 (b) の横軸がノードの出線数、縦軸が出線数を降順に並べた際の順位を示す。どちらの図においても、べき則に従う場合は右肩下がりの直線で近似できる。

図 1 から図 5 を比較すると、リンク接続先候補数 j の値を 10 に近づけるほど、BA モデルの出線数分布に近づき、べき則に従うことがわかる。しかし、図 6 のように j の値を大きくすると、BA に比べ、出線数に大きな偏りが生じてしまい、傾斜が急なべき則に従うこともわかる。これは、接続確率を与える際に使用している指標である Betweenness Centrality の値が、出線数に比べ偏りが生じやすく、リンク接続先候補ノードを増やす程 Betweenness Centrality の値が大きい特定のノードが候補ノードに含まれ、リンク接続先ノードとして選出されやすくなるのが原因と考える。

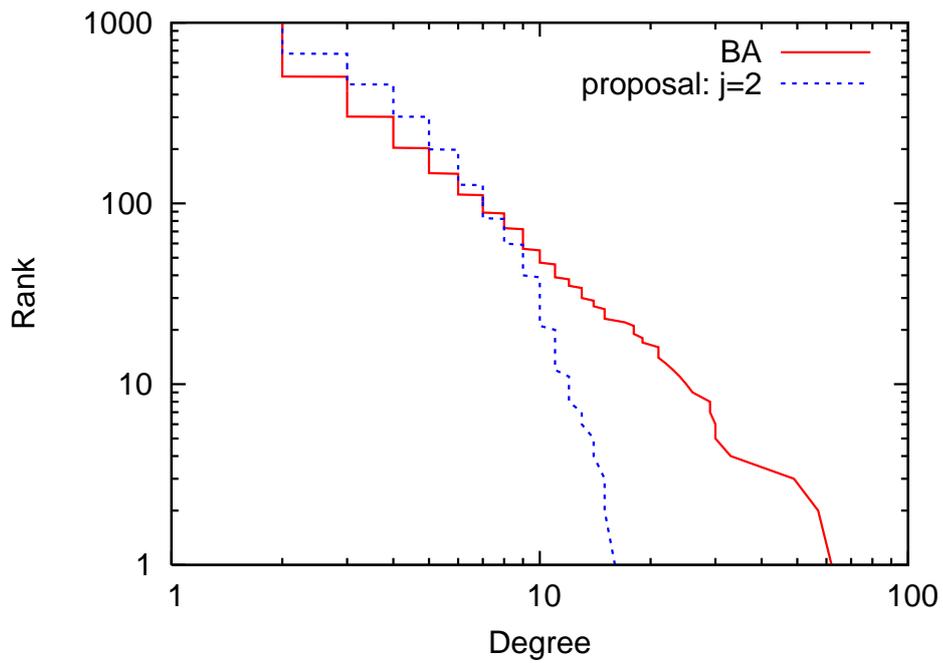
また、 j の値を変化させ生成したトポロジにおいてアタック障害、ランダム障害への耐性を評価した。その結果を図 7、図 8 に示す。これらの図の横軸が障害により故障したノード数、縦軸がカバー率を示している。カバー率とはトポロジ全体のうち、通信可能なノードの割合をあらわす。

図 7(b) より、 j の値が 5 の場合は故障ノード数が 50 程度に到達するまでは BA モデルの評価結果に比べ、耐性が向上していることがわかるが、 j の値を 10、200 とした場合、BA モデルに比べカバー率が悪化していることがわかる。しかし、BA モデルの欠点であるアタック障害への耐性の脆弱さを改善することを目的としているため、 j の値を 5 までとして図 7(a)、8 に示すように障害への耐性評価を行った。

図 7(a) より、アタック障害への耐性が BA モデルの耐性と比べ、故障ノード数が 70 程度に達するまではカバー率が BA モデルの結果を上回ることがわかる。しかし、70 程度に達した際に、耐性が急激に悪化することもわかる。この原因として、故障するノード数がある値に達すると、生成されていたトポロジがツリー状に変化し、その中心部となるノードに故障が発生した際にこのような状態になることがあげられる (図 9、図 10 参照)。また、図 8 より、BA モデルに比べ、ランダム障害への耐性が低いこともわかる。

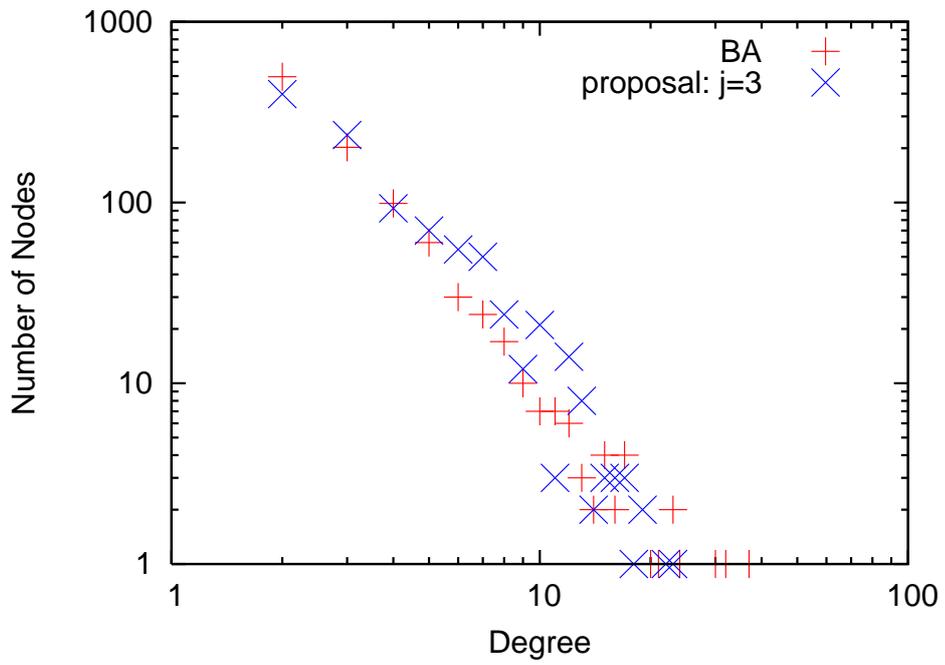


(a) 出線数分布

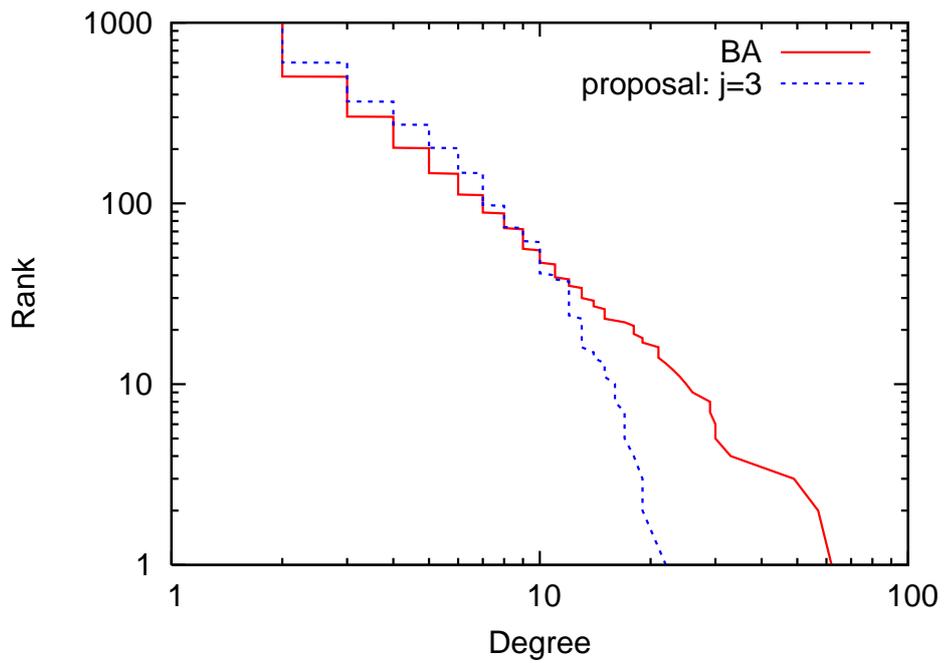


(b) 出線数の累積分布

図 1: 局所情報に基づくトポロジの出線数分布 : $j = 2$ の場合

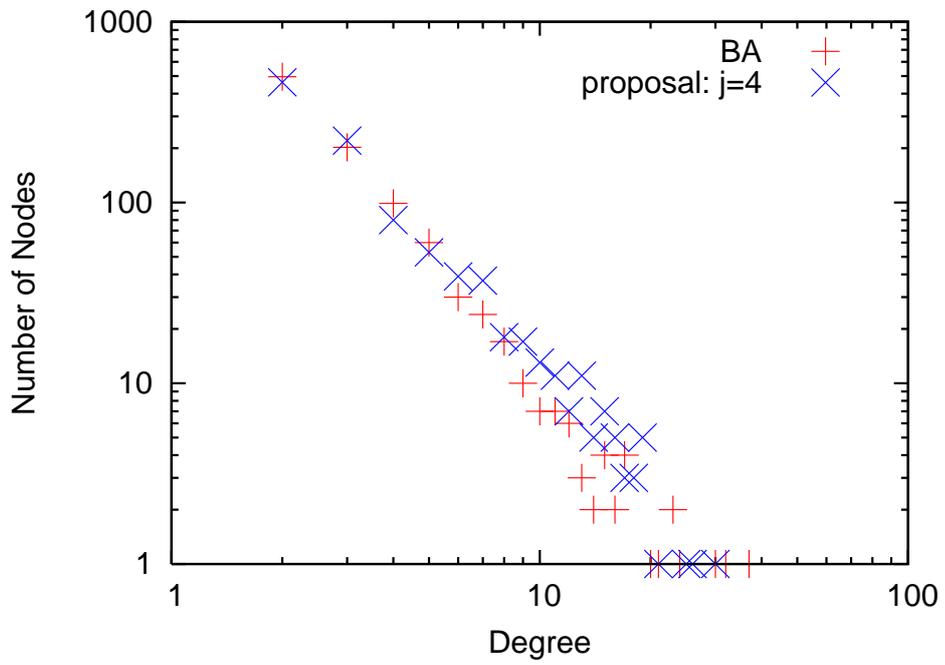


(a) 出線数分布

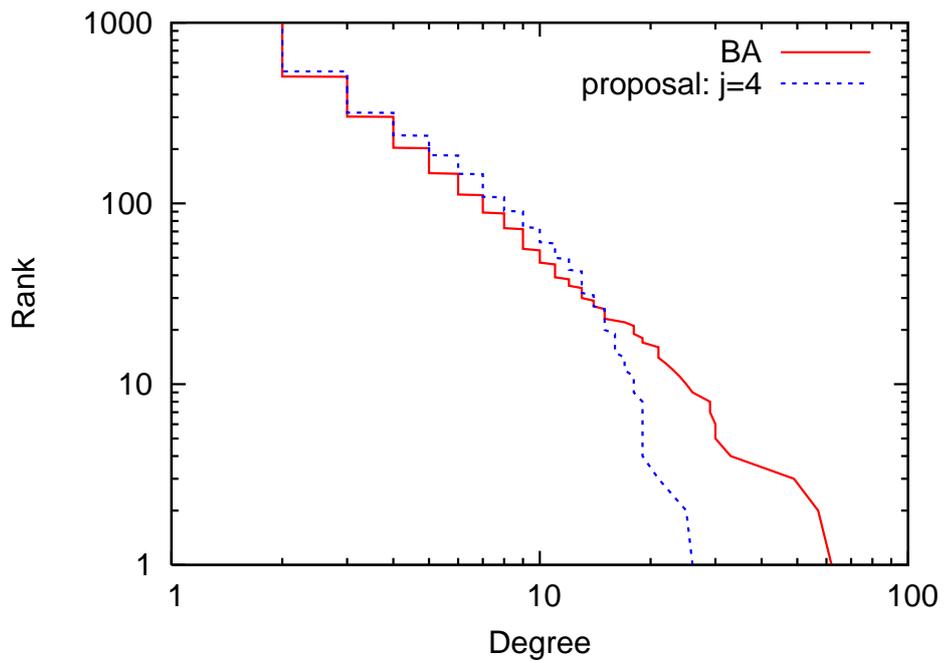


(b) 出線数の累積分布

図 2: 局所情報に基づくトポロジの出線数分布 : $j = 3$ の場合

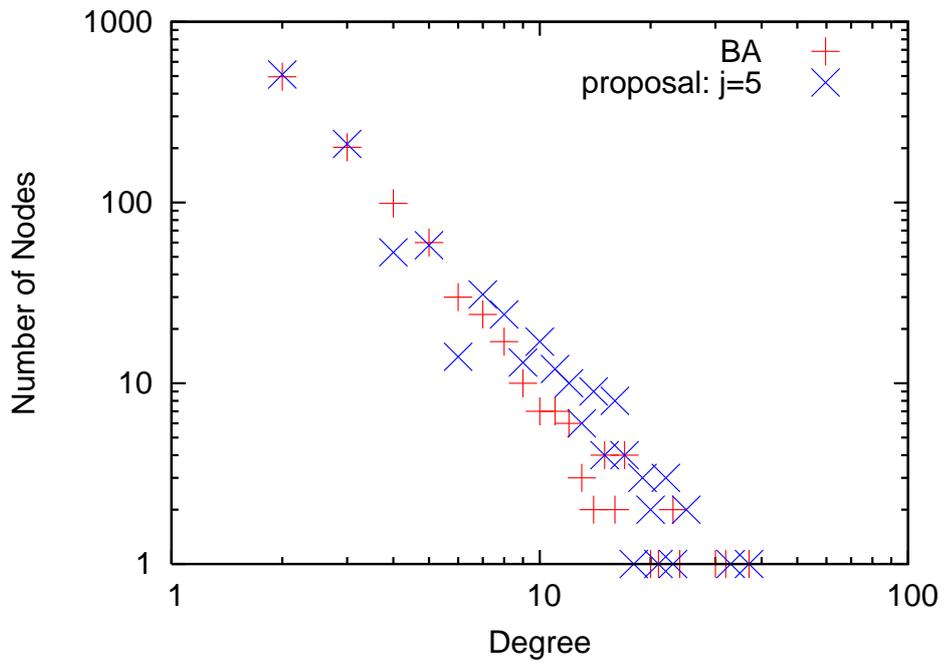


(a) 出線数分布

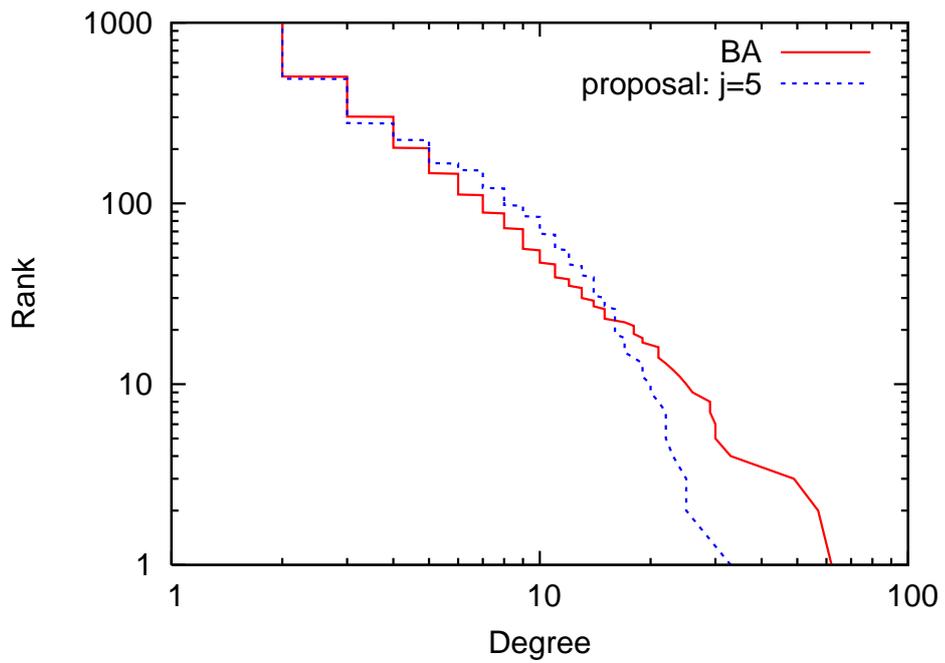


(b) 出線数の累積分布

図 3: 局所情報に基づくトポロジの出線数分布 : $j = 4$ の場合

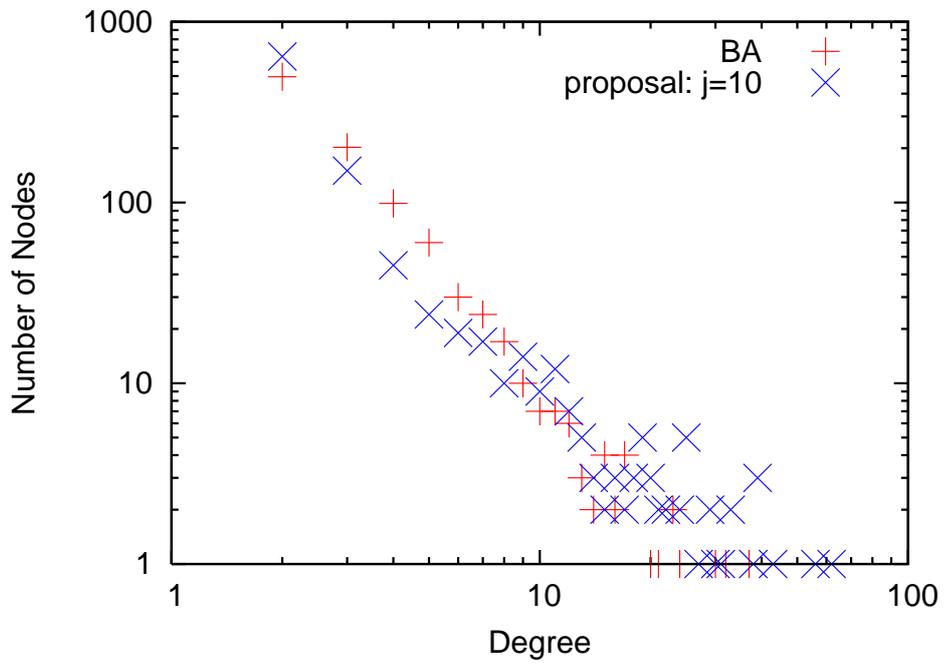


(a) 出線数分布

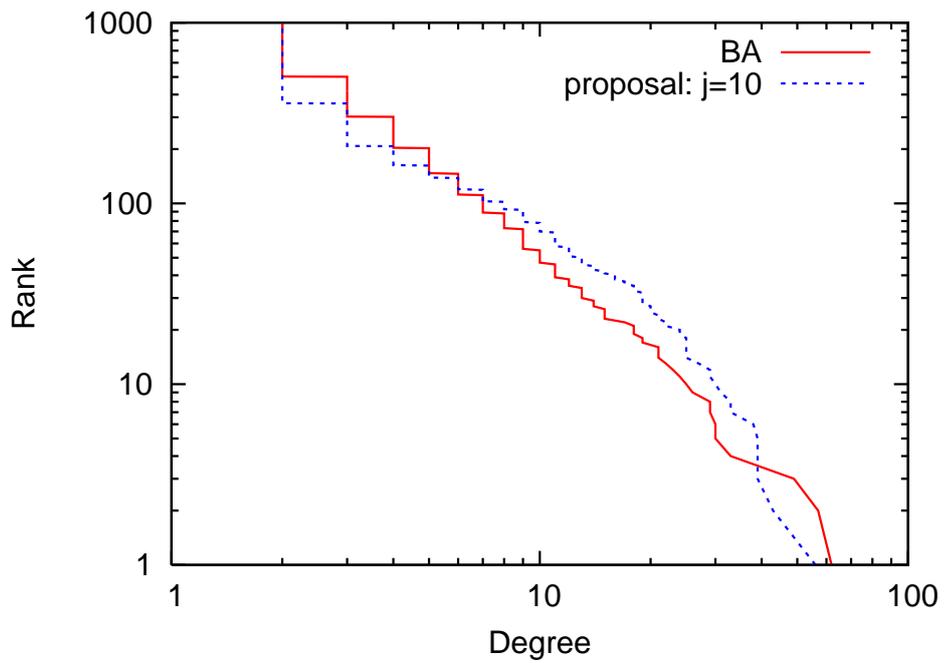


(b) 出線数の累積分布

図 4: 局所情報に基づくトポロジの出線数分布 : $j = 5$ の場合

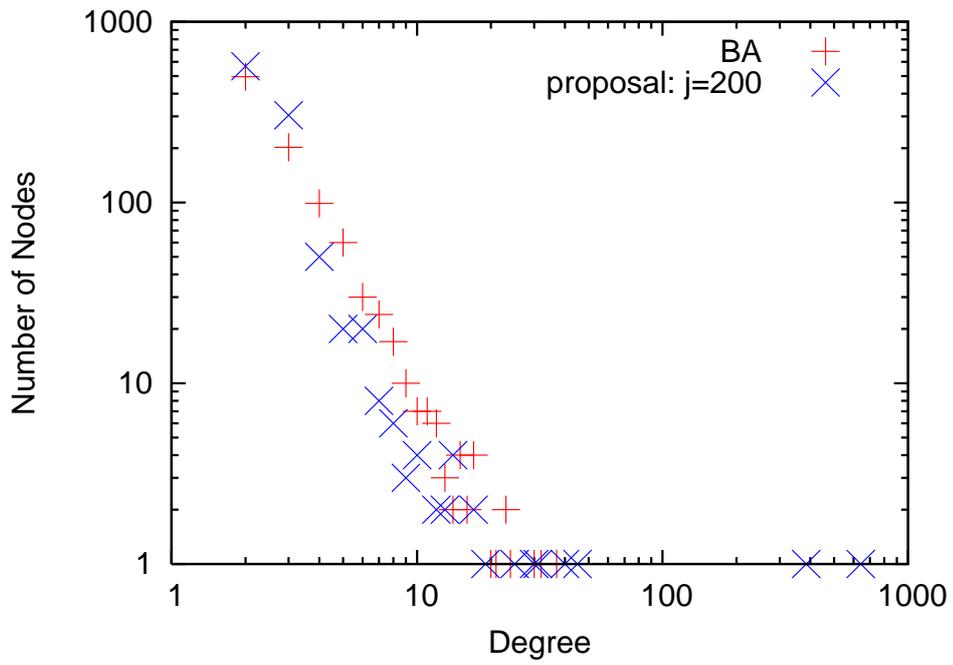


(a) 出線数分布

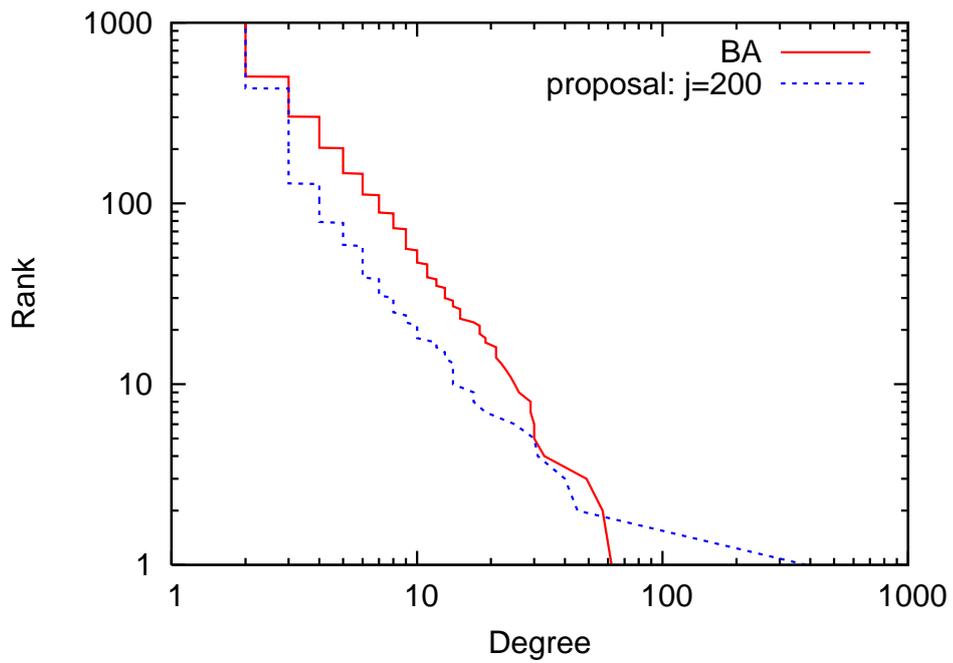


(b) 出線数の累積分布

図 5: 局所情報に基づくトポロジの出線数分布 : $j = 10$ の場合

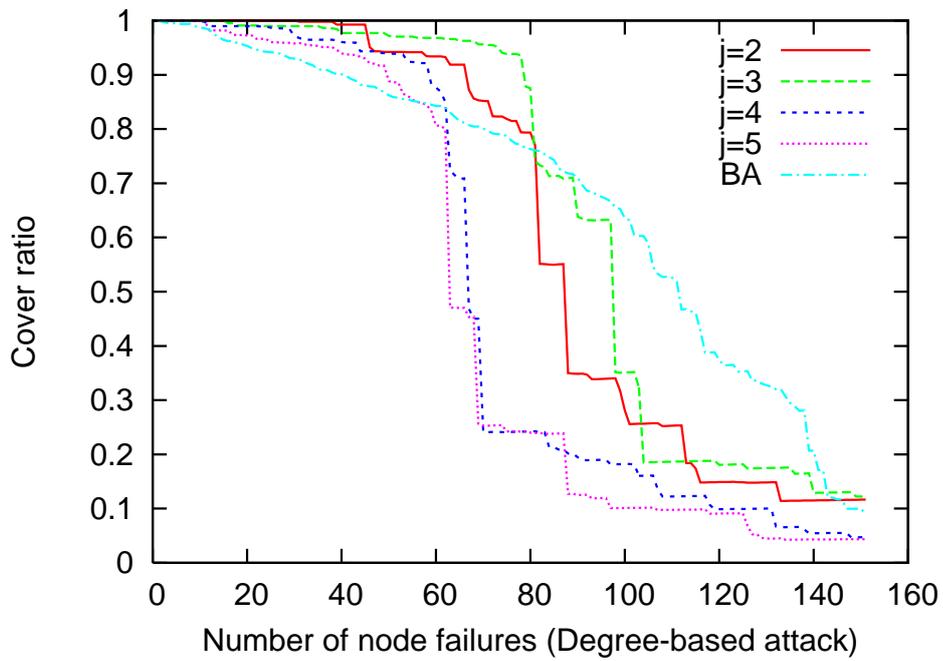


(a) 出線数分布

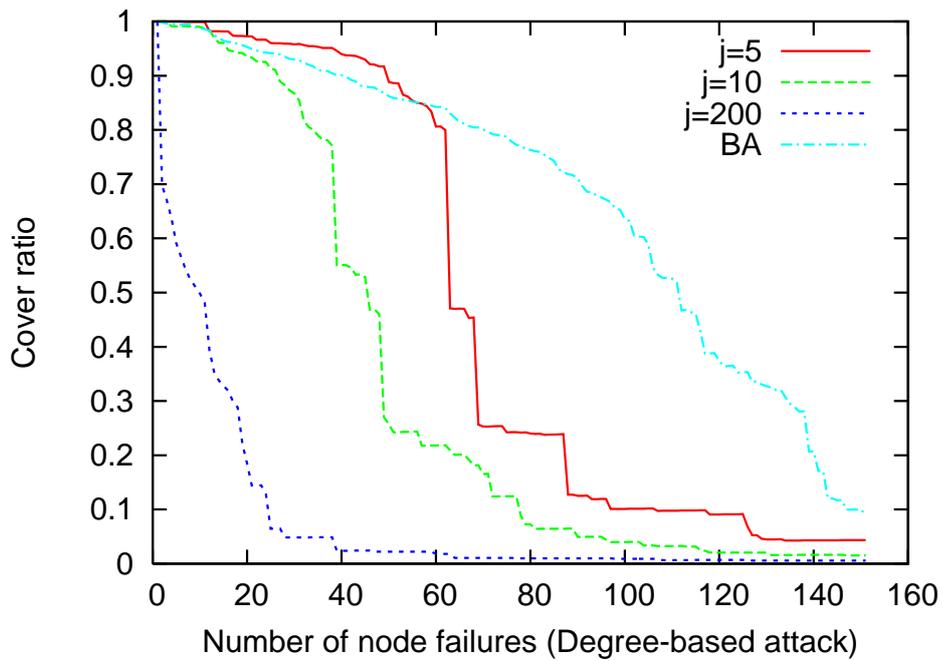


(b) 出線数の累積分布

図 6: 局所情報に基づくトポロジの出線数分布: $j = 200$ の場合



(a) $2 \leq j \leq 5$ の場合



(b) $5 \leq j \leq 200$ の場合

図 7: 局所情報に基づくトポロジの攻撃障害への耐性評価

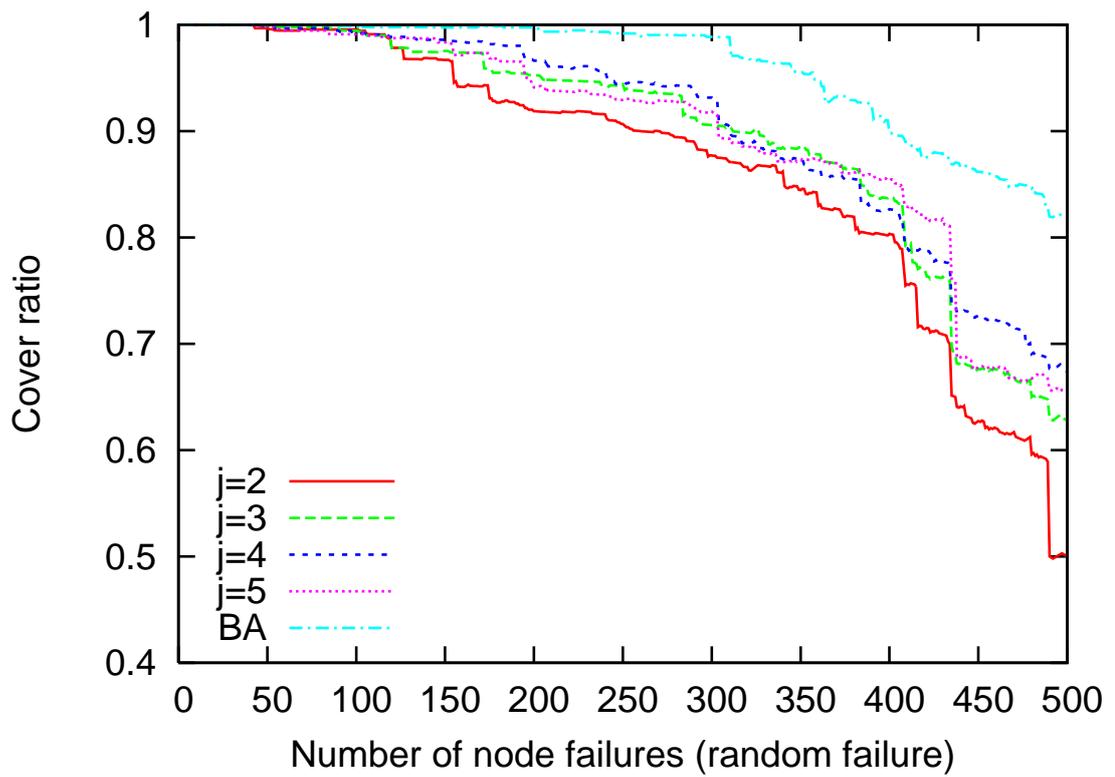


図 8: 局所情報に基づくトポロジのランダム障害への耐性評価

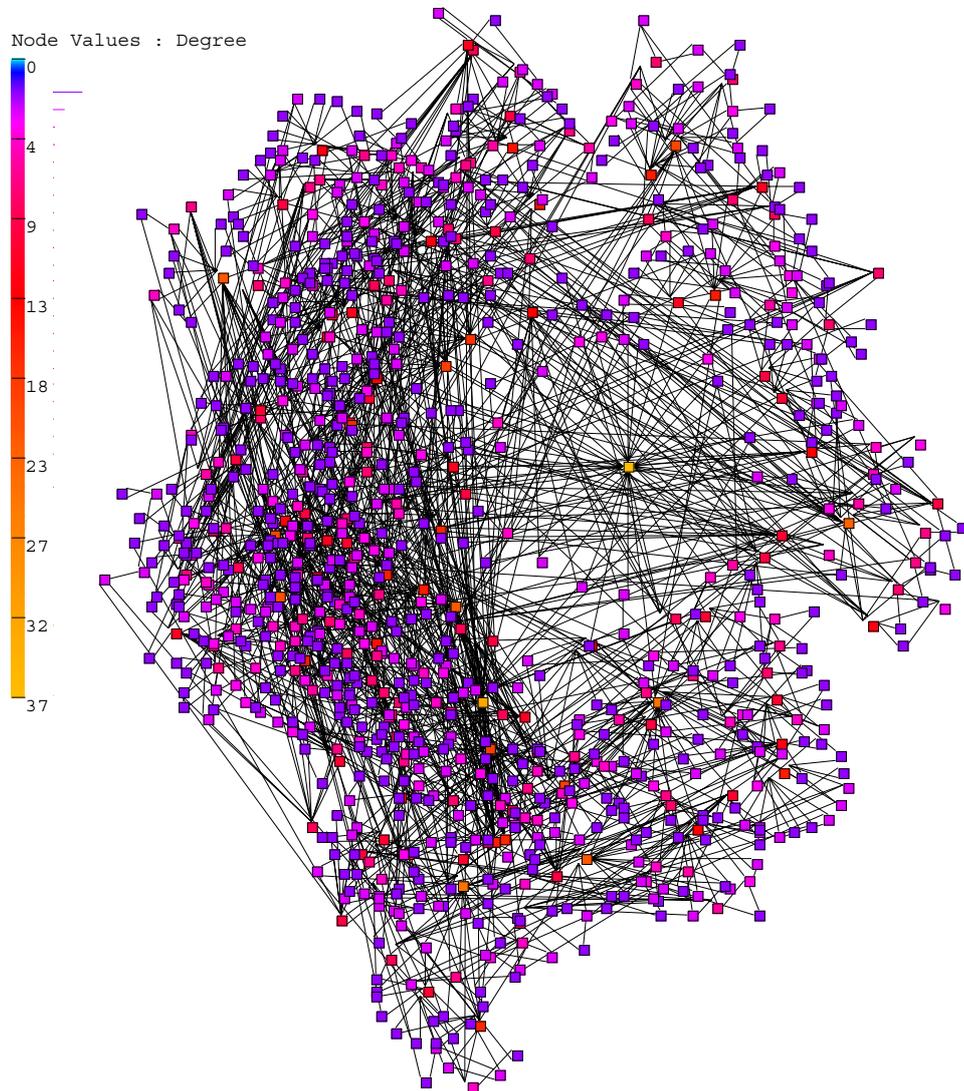
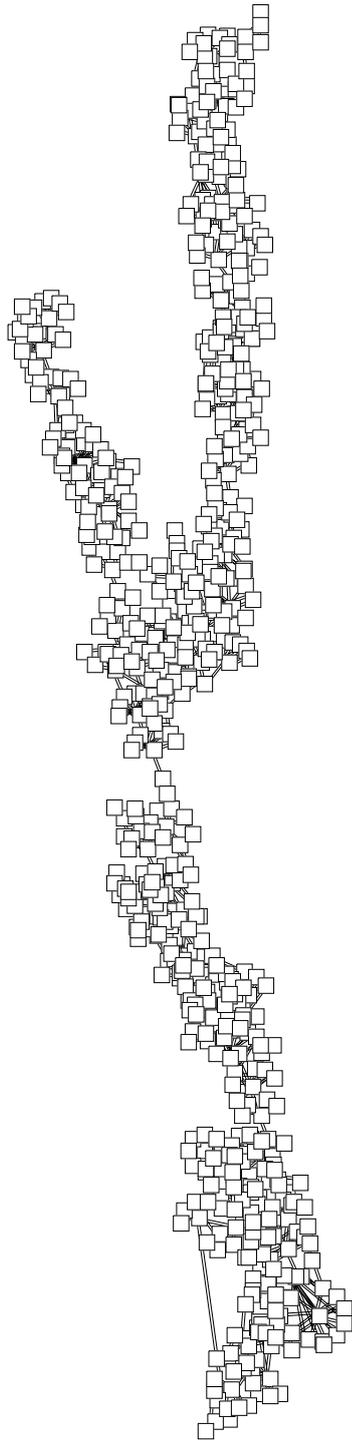
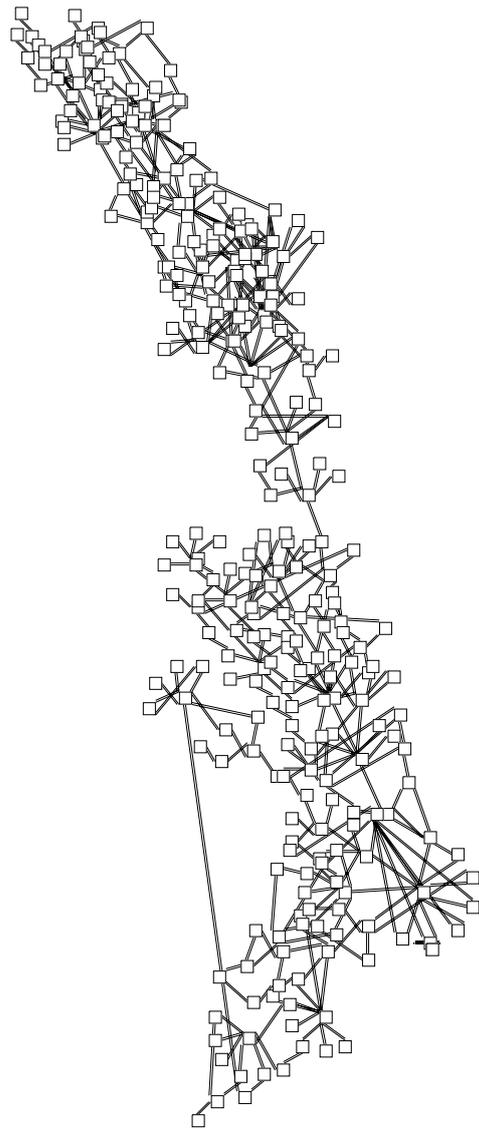


図 9: 局所情報に基づくトポロジ：ノード数が1000個の場合



(a) 61 個のノードが故障した場合



(b) 62 個のノードが故障した場合

図 10: 局所情報に基づくトポロジ : アタック障害の耐性が悪化する例

4 クラスタリング手法に基づくトポロジ生成手法の提案と評価

本章では、クラスタリング手法により既存のネットワークを複数のクラスタに分割し、そのクラスタ番号が異なるものにリンクを接続するトポロジ生成手法の提案および評価を述べる。

4.1 クラスタリング

本節では 3.1 節で述べた局所情報のうち、クラスタ番号を用いる手法について述べる。

局所情報を基にネットワークを分割する分散クラスタリング手法 [18] によって、クラスタリングが行われた際に、ノードは属するクラスタを識別するためのクラスタ番号という情報を保持していることを前提とする。各ノードの保持するクラスタ番号を用いて異なるクラスタにリンクを接続することにより、あるクラスタ内で障害が発生した際に各ノードは他のクラスタにリンクを接続しているため、ネットワークの接続性が保たれ、アタック障害に対して耐性を高めることが可能と考える。そのため、これらのクラスタ番号を局所情報として用いたトポロジ生成を行う。

ただし、評価においては、クラスタ番号を局所情報として用いたトポロジ生成手法においてアタック障害への耐性を高めることを目的としているため、集中型のクラスタリング手法 [19] を用いた。

4.2 クラスタリング手法に基づくトポロジ生成手法

べき則に従うネットワークトポロジ生成モデルの一つである、BA (Barabási-Albert) モデル [2] では、トポロジを生成する際に大域情報が必要となるが、このモデルの特性のひとつである、スモールワールドの性質は、アプリケーションの性能の効率化を考える際に重要となる。そこで、BA モデルのトポロジ生成手法を基に、クラスタリング手法に基づくトポロジ生成手法を提案する。提案手法では、文献 [18] のような分散クラスタリング手法を用いて、既存のネットワークをクラスタに分割し、各ノードがクラスタ番号を保持していることを前提とする。

新たなノードを追加した際に、新規追加ノードが既存のトポロジに対して探索を行い、近隣 j 個のノードをリンク接続先の候補とする。そして、ノード i の出線数 c_i の値が大きい程、ノード i へのリンクが接続される確率 $\Pi(c_i)$ が高くなるように式 (3) の確率を与え、その確率に従って、リンクの接続先候補となるノードを選出する。確率 $\Pi(c_i)$ が高いノード i に優先的にリンクを接続することによって、ハブノードが現れやすくなり、各ノード間の平均ホップ数が小さくなる。選出されたノードのクラスタ番号と、追加ノードがリンクを接続しているノードのクラスタ番号を比較し、異なる場合はそのノードにリンクを接続し、一致する場合は再び確率 $\Pi(c_i)$ に従って接続先候補となるノードを選出する。

相異なるクラスタ番号を保持するノードにリンクを接続することによって、ノードのリンク接続先のクラスタが分散し、アタック障害への耐性が高まることが期待される。ただし、リンクを接続する候補のノード全てのクラスタ番号が一致する場合は、接続確率に従ってリンクを接続する。

提案するトポロジ生成手法の手順は以下の通りである。

Step 1 : 初期ノードとして m_0 個のノードを配置する。

Step 2 : ネットワークのノード数が N 未満である場合、Step 3 へ。ノード数が N であればトポロジ生成を終了する。

Step 3 : ノードを 1 個追加する。

Step 4 : 追加されたノードとユークリッド距離が近い $j (\geq m)$ 個のノードをリンクを接続する候補とする。

Step 5 : 追加されたノードから式 (3) の確率に従い、他の $m (\leq m_0)$ 個の相異なるノードへリンクを接続する。ただし、選ばれたノードのクラスタ番号がすでに接続済みのノードのクラスタ番号と一致する場合、もう一度式 (3) の確率に従い、ノードを選ぶ。 m 個のノードへリンクが接続された後、Step 2 へ。

$$\Pi(c_i) = c_i / \sum_j c_j \quad (3)$$

4.3 生成されたトポロジの特性評価

生成されたトポロジにおける特性評価として、BA モデルのトポロジとの比較、および、アタック障害への耐性の評価を行った結果について述べる。評価にはノード数 1000 個で生成したトポロジを用い、提案手法においてはリンク接続候補数 j の値を 2、5、10、500 と変化させて生成したトポロジを用いた。

4.1 節で述べたように、評価では集中型のクラスタリング手法 [19] を用いている。そこで、用いたクラスタリング手法について述べる。

文献 [19] では、トポロジを 2 つのクラスタに分割する手法を再帰的に用いていくつかのクラスタに分割する。クラスタを 2 つに分割する際には、行列 B を式 (4) のように定義し、この行列を用いて式 (5) のように定義されるベクトル s と行列 B を用いて式 (6) のように定義される指標 Q (≤ 1) の値が大きくなる場合にクラスタリングを行う。

Step 1 : 対象トポロジの行列 B を式 (4) と定義する。ただし、 $A_{(i,j)}$ はノード i 、 j 間にリンクが接続されている場合 1 となり、そうでない場合は 0 となる。また、 k_i はノード i の出線数、 m はトポロジ内の総リンク数を表す。

$$B_{(i,j)} = A_{(i,j)} - k_i k_j / 2m \quad (4)$$

Step 2 : 行列 B の、正でありかつ最も値の大きい固有ベクトル u_i と固有値 β_i を求める。また、一次ベクトル s を式 (5) と定義する。ただし、 n はそのトポロジ内のノード数とする。

$$s = \sum_{i=1}^n a_i u_i \quad a_i = u_i^T s \quad (5)$$

Step 3 : 式 (6) で示される指標に従ってネットワークを 2 つのクラスタに分割する。その際、クラスタの境界線上に存在するリンク数が最少となるようにする。

$$Q = s^T B s \quad (6)$$

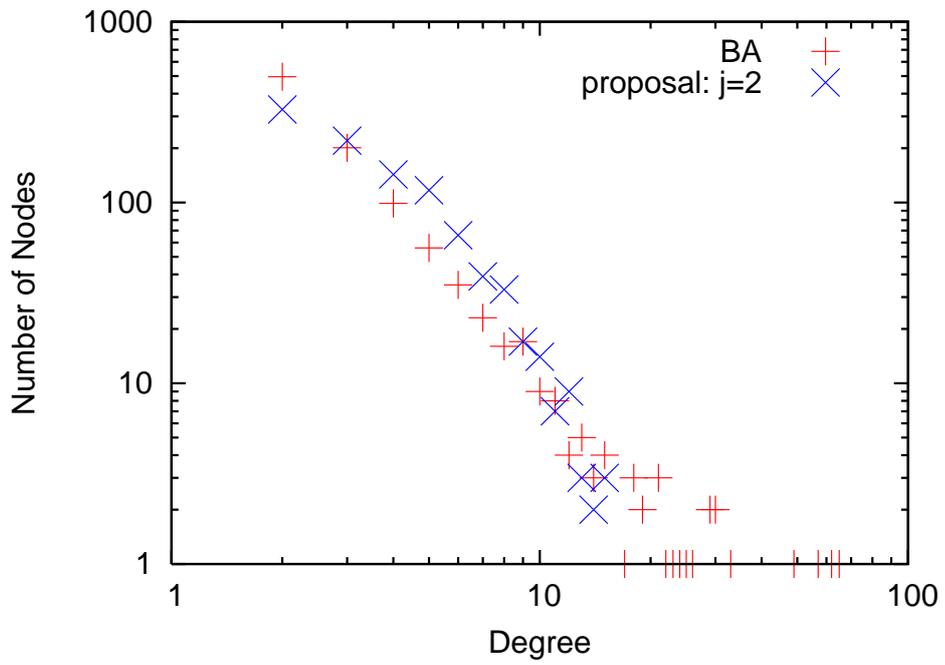
Step 4 : $Q \leq 0$ の場合、それ以上クラスタに分割できないため、クラスタ化を終了する。

そうでない場合は Step 1 へ。

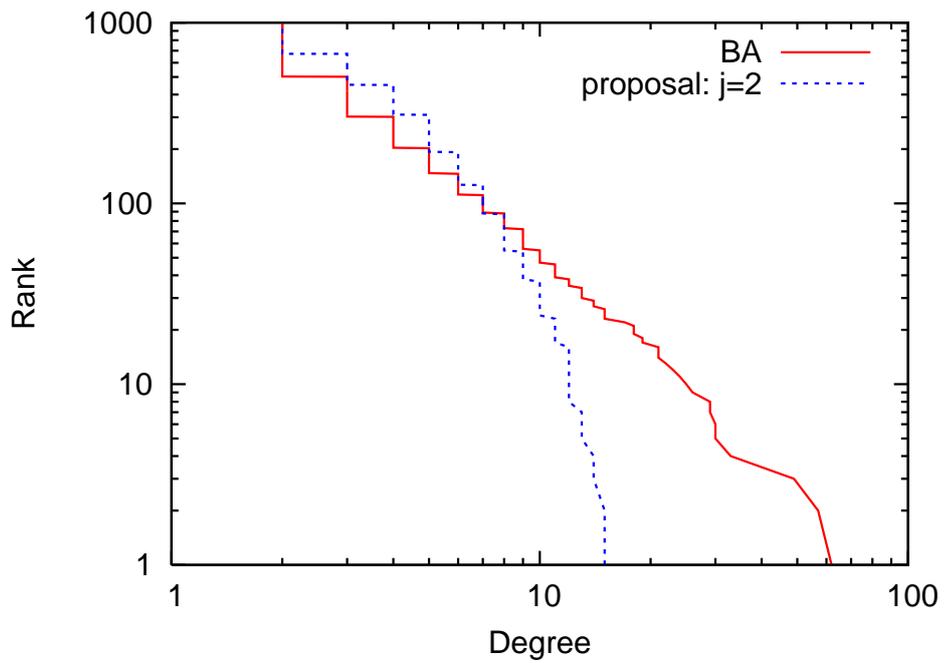
このクラスタリング手法を用いてクラスタ番号を各ノードに保持させ、クラスタ番号を基にトポロジの生成を行った。

ここでは、BA モデル、提案手法それぞれから生成されたトポロジの出線数分布、出線数の累積分布の比較を図 11 から図 15 に示す。これらの図 (a) の横軸がノードの出線数、縦軸がその出線数を持つノード数を示す。図 (b) の横軸がノードの出線数、縦軸が出線数を降順に並べた際の順位を示す。図 11 から図 15 を比較すると、リンク接続先候補数 j の値を増やし、トポロジ内のノード数に近づけるほど、べき則に従い、BA モデルの出線数分布と近くなるということがわかる。

また、 j の値を変化させ生成したトポロジにおいてアタック障害、ランダム障害への耐性を評価した。その結果を図 16、17 に示す。これらの図の横軸が障害により故障したノード数、縦軸がカバー率を示している。これらの図より、 j の値を大きくすると、BA モデルの結果とほぼ一致することがわかる。この原因として、追加したノードのリンクの接続候補数を増やすほど、クラスタ番号の不一致の確率が高くなり、クラスタ番号を用いることなくリンクの接続が行われていることがあげられる。そこで、 j の値を小さいものに注目すると、 $j = 5$ の場合はアタック障害への耐性が高いものの、ランダム障害への耐性が BA モデルに比べ劣化していることがわかる。 $j = 10$ の場合はアタック障害への耐性が高くなり、ランダム障害への耐性は障害の起きるノード数が増えるに従い、BA モデルのものに近づくことがわかる。

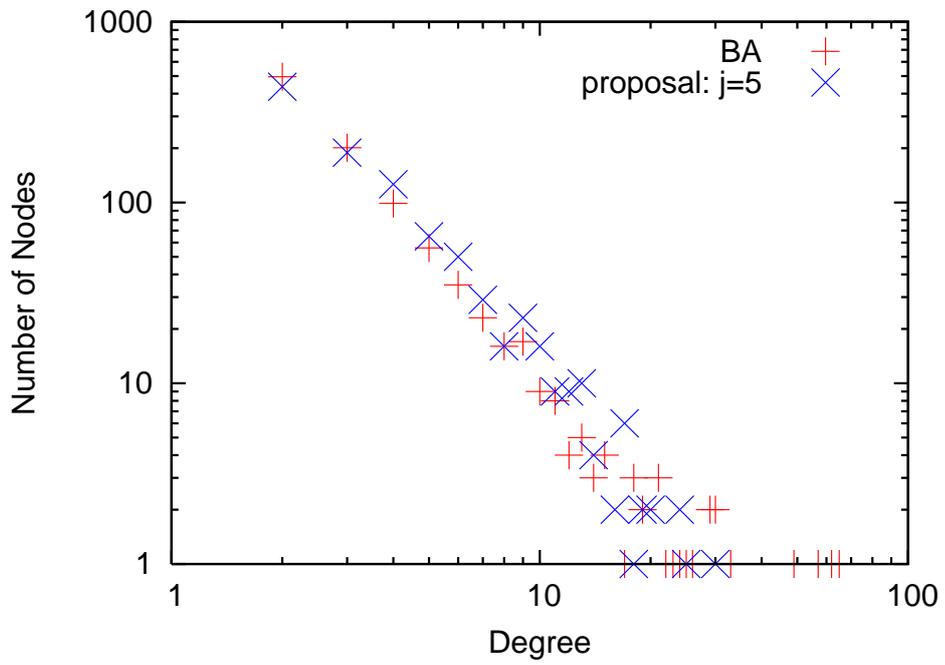


(a) 出線数分布

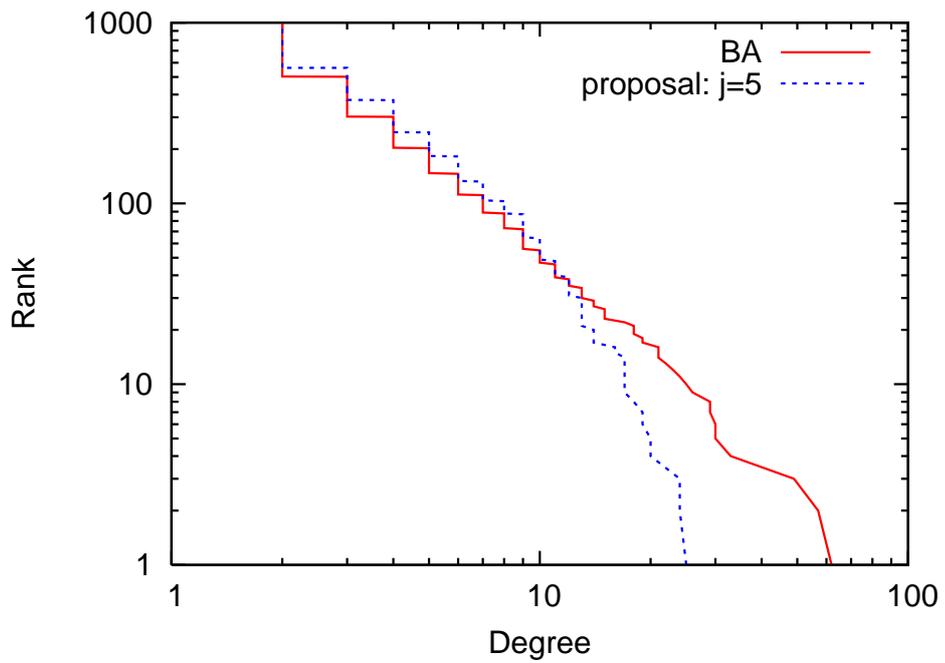


(b) 出線数の累積分布

図 11: クラスタリング手法に基づくトポロジの出線数分布 : $j = 2$ の場合

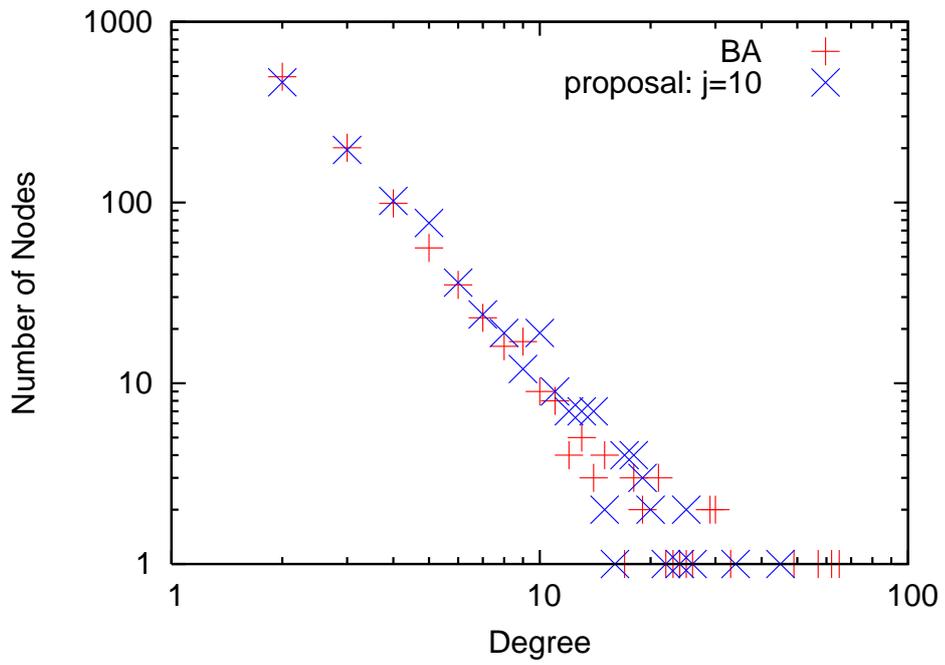


(a) 出線数分布

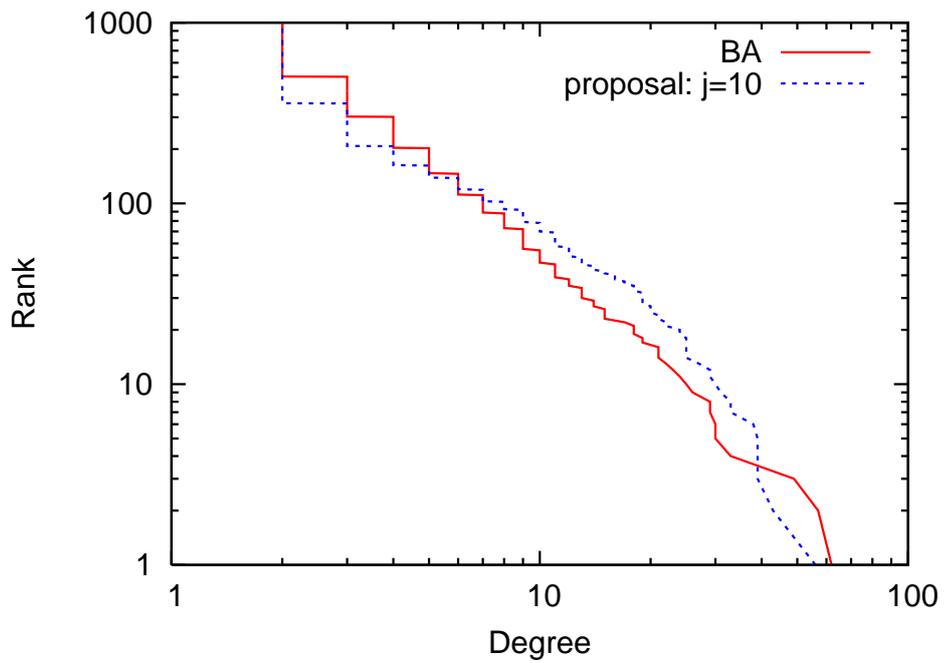


(b) 出線数の累積分布

図 12: クラスタリング手法に基づくトポロジの出線数分布 : $j = 5$ の場合

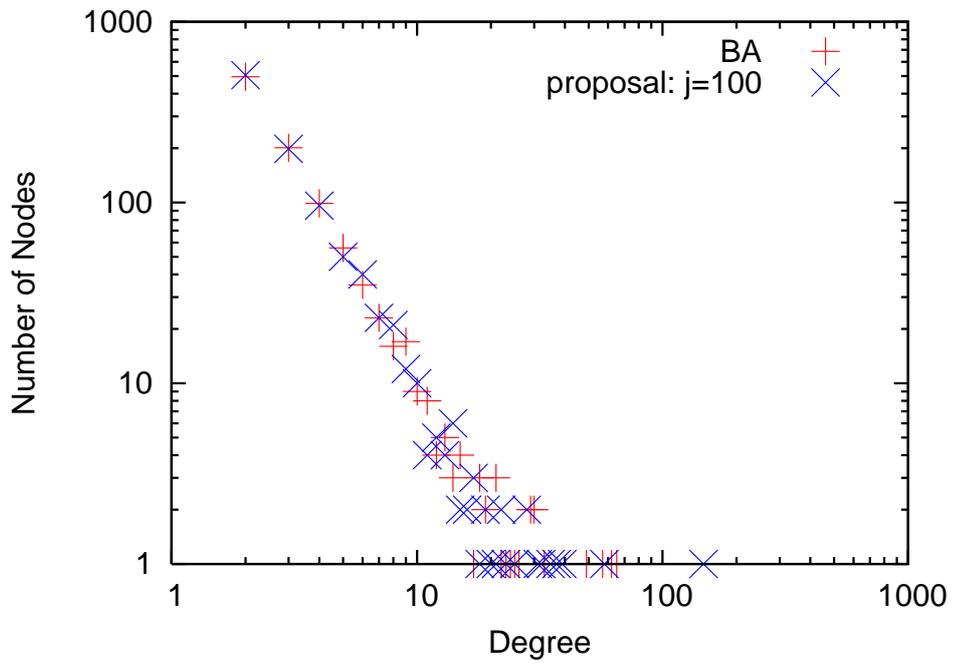


(a) 出線数分布

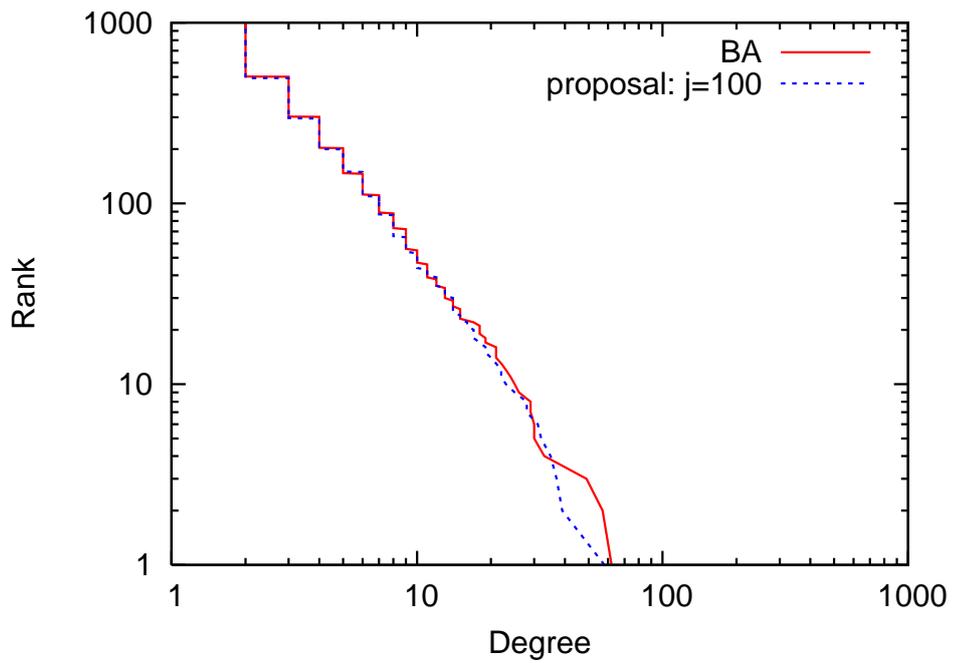


(b) 出線数の累積分布

図 13: クラスタリング手法に基づくトポロジの出線数分布 : $j = 10$ の場合

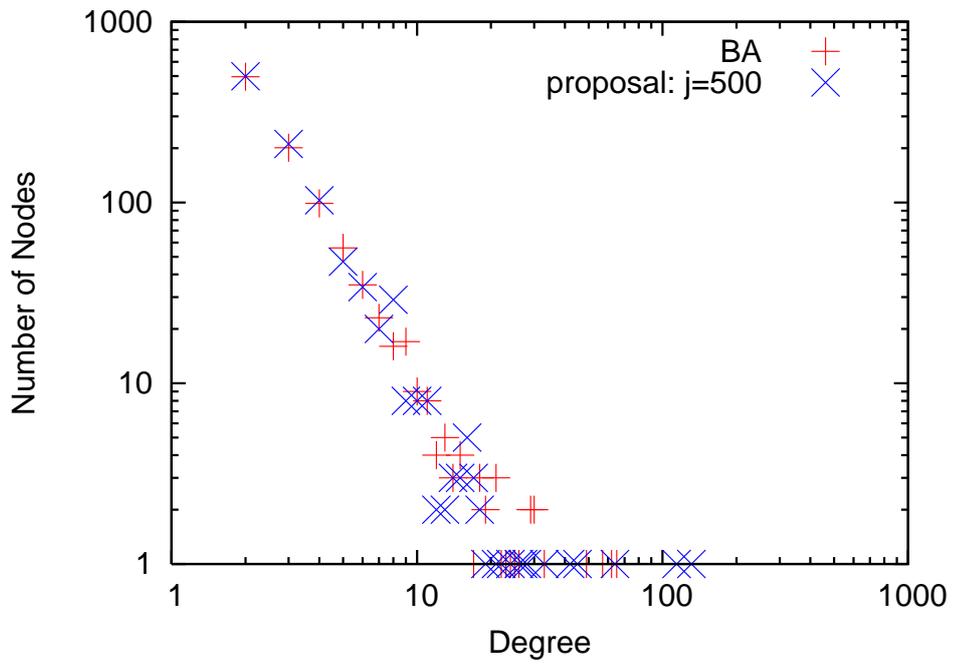


(a) 出線数分布

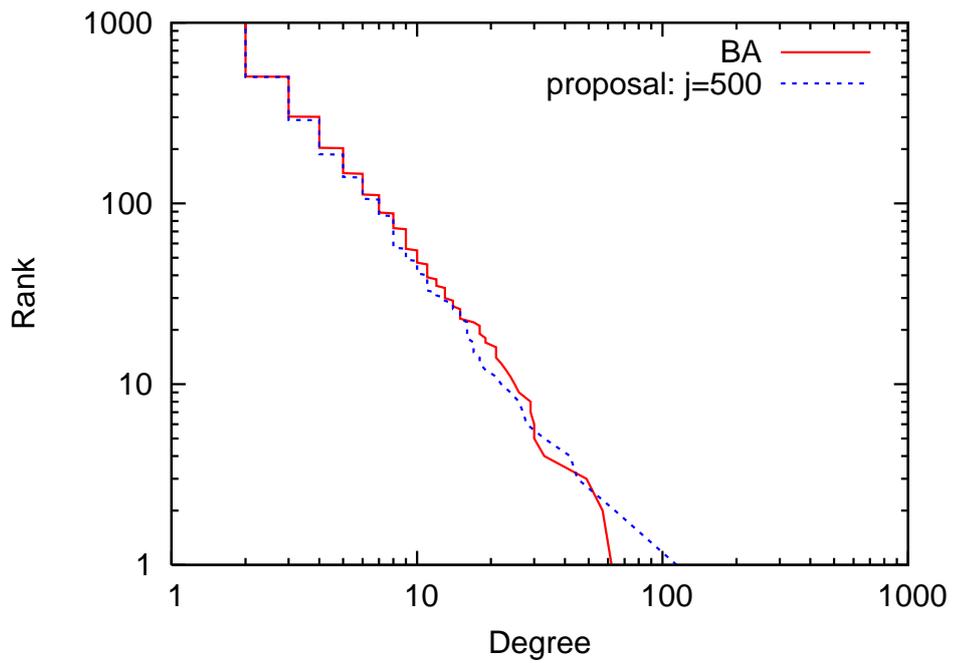


(b) 出線数の累積分布

図 14: クラスタリング手法に基づくトポロジの出線数分布 : $j = 100$ の場合



(a) 出線数分布



(b) 出線数の累積分布

図 15: クラスタリング手法に基づくトポロジの出線数分布 : $j = 500$ の場合

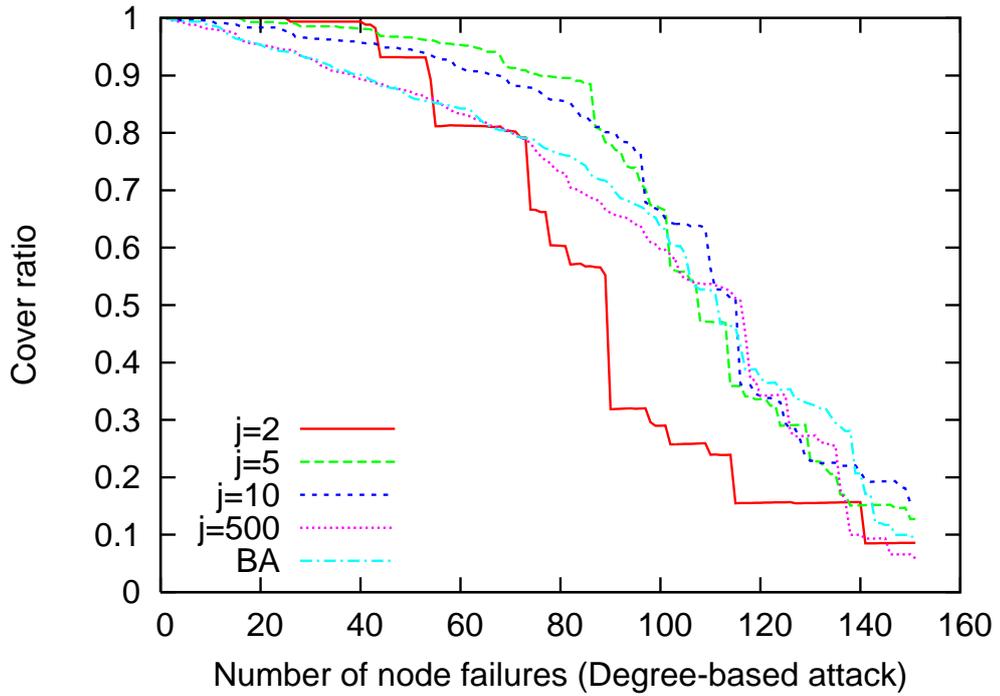


図 16: クラスタリング手法に基づくトポロジの攻撃障害への耐性評価

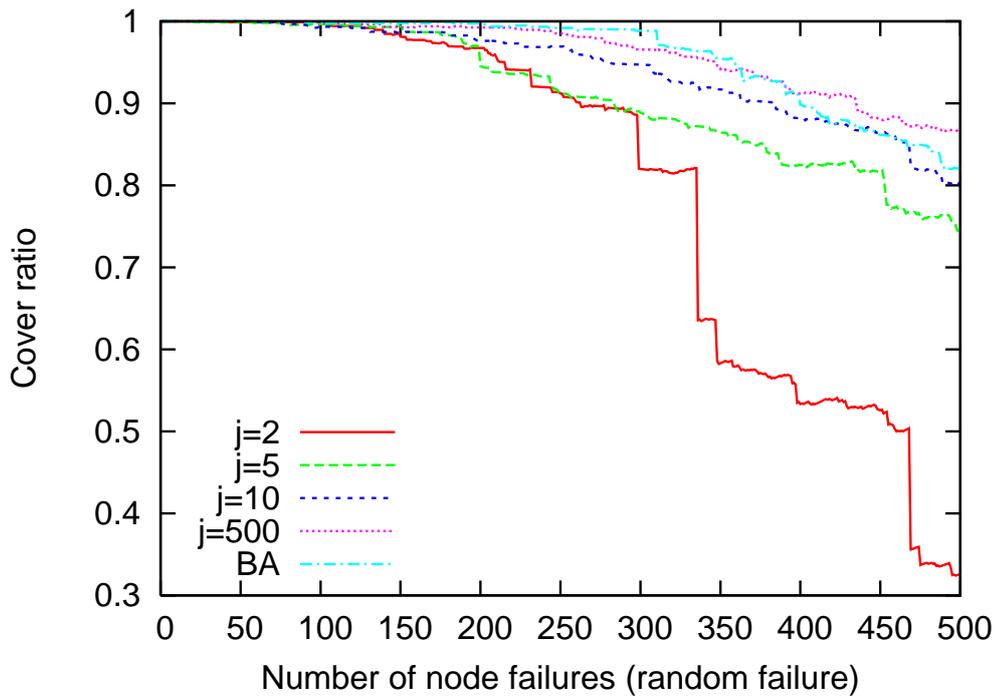


図 17: クラスタリング手法に基づくトポロジのランダム障害への耐性評価

5 まとめ

本報告では、局所情報に基づいたトポロジ生成手法で生成されたトポロジにおいてネットワークの性能評価が行われていないということに着目し、分散制御の適用が必須であるような大規模ネットワークを対象とし、局所情報を用いたトポロジ生成手法を提案し、生成されたトポロジのネットワークの性能評価を行った。まず大域情報に基づくトポロジ生成手法の一つである BA モデルの特性であるスモールワールドの性質、ランダム障害への耐性の高さを考慮し、BA モデルの生成手法を基に二つのトポロジ生成手法を提案した。いずれの手法においても、追加されたノードが近隣のノードの情報を取得し、リンクを接続する候補を選出することによって分散制御を可能とした。また、BA モデルにおいて脆弱とされるアタック障害への耐性を高めることを目的とし、トポロジの生成を行った。シミュレーションによって障害への耐性を評価した結果、BA モデルに比べ、ノード負荷を用いて生成したトポロジにおいてはアタック障害への耐性がおよそ 100%向上し、クラスタ番号を用いて生成したトポロジにおいてはアタック障害への耐性がおよそ 70%向上することが明らかとなった。しかし、ランダム障害への耐性は、ノード負荷を用いて生成したトポロジにおいてはおよそ 22%、クラスタ番号を用いて生成したトポロジにおいてはおよそ 4%劣化することも明らかとなった。

本報告で提案した手法では、大規模なアタック障害が発生した際に耐性が急激に悪化することも明らかとなっている。このことから、提案手法をネットワークの分散制御に応用するためには、リンクの再接続を考慮したトポロジ生成が必要不可欠であり、そのようなトポロジ生成手法を考案することが今後の課題としてあげられる。

謝辞

本報告の作成にあたり、ご指導ならびにご教授いただきました大阪大学情報科学研究科の村田正幸教授に心より感謝を申し上げます。また、本報告において終始ご指導いただきました大阪大学情報科学研究科の荒川伸一助手には、日頃から様々な相談に乗っていただき、助言をしていただきました。深く感謝致します。ならびに、平素から適切な助言をいただきました大阪大学情報科学研究科の若宮直紀助教授、大阪大学サイバーメディアセンターの長谷川剛助教授、大阪大学サイバーメディアセンターの笹部昌弘助手、大阪大学経済学研究科の天下裕一助手に感謝を申し上げます。

最後に、常日頃から様々な相談に応じていただき、支えてくださいました小泉佑揮氏、福元良太氏、堀内英斗氏、大橋正稔氏、木利友一氏をはじめとする村田研究室の皆様心より御礼申し上げます。

参考文献

- [1] M. Faloutsos, P. Faloutsos, and C. Faloutsos, “On power-law relationships of the Internet topology,” in *Proceedings of ACM SIG-COMM’ 99*, pp. 251–262, Aug. 1999.
- [2] A. L. Barabasi and R. Albert, “Emergence of scaling in random networks,” *Science*, vol. 286, pp. 509–512, Oct. 1999.
- [3] K. L. Gohn, B. Kahng, and D. Kim, “Universal behavior of load distribution in scale-free networks,” *Physical Review Letters*, vol. 87, no. 27.
- [4] L. Zhao, K. Park, and Y.-C. Lai, “Attack vulnerability of scale-free networks due to cascading breakdown,” *Physical Review E*, vol. 70, Sept. 2004.
- [5] L. Gallos, P. Argyrakis, A. Bunde, R. Cohen, and S. Havlin, “Tolerance of scale-free networks: from friendly to intentional attack strategies,” *Physica A: Statistical Mechanics and its Applications*, vol. 344, pp. 504–509, Dec. 2004.
- [6] R. Cohen, K. Erez, D. Avraham, and S. Havlin, “Breakdown of the Internet under intentional attack,” *Physical Review Letters*, vol. 86, Apr. 2001.
- [7] R. Albert, H. Jeong, and A. Barabasi, “Error and attack tolerance of complex networks,” *Nature*, vol. 406, pp. 378–382, Aug. 2000.
- [8] S. Saroiu, P. Gummadi, and S. Gribble, “A measurement study of peer-to-peer file sharing systems,” in *Proceedings of Multimedia Computing and Networking 2002 (MMCNf02)*, pp. 156–170, Jan. 2002.
- [9] G. Siganos, M. Faloutsos, P. Faloutsos, and C. Faloutsos, “Power laws and the AS-level Internet topology,” *IEEE/ACM Transactions on Networking*, vol. 11, pp. 514–524, Aug. 2003.

- [10] M. Kim and M. Médard, “Robustness in large-scale random networks,” in *Proceedings of INFOCOM 2004*, vol. 4, pp. 2364–2373, Mar. 2004.
- [11] R. Albert and A. Barabási, “Statistical mechanics of complex networks,” *Reviews of modern physics*, vol. 74, pp. 47–97, Jan. 2002.
- [12] M. E. J. Newman, *Random graphs as models of networks*, ch. 2, pp. 35–68. WILEY–VCH, 2002, Nov. 2002.
- [13] J. Gómez-Gardeñes and Y. Moreno, “Local versus global knowledge in the Barabási–Albert scale-free network model,” *Physical Review E*, vol. E69, pp. 1–4, Mar. 2004.
- [14] D. Smith, C. F. Lee, and N. F. Johnson, “Realistic network growth using only local information: From random to scale-free and beyond,” arXiv:cond-mat/0608733v2, Sept. 2006.
- [15] 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.
- [16] X. Wang and G. Chen, “Complex networks: Small-world, scale-free and beyond,” *Circuits and Systems Magazine, IEEE*, vol. 3, pp. 6–20, Sept. 2003.
- [17] V. Latora and M. Marchiori, “A measure of centrality based on the network efficiency,” arXiv:cond-mat/0402050v1, Feb. 2004.
- [18] S. Basagni, “Distributed clustering for ad hoc networks,” in *Proceedings of the 1999 International Symposium on Parallel Architectures, Algorithms and Networks (ISPAN ’99)*, pp. 310–315, June 1999.
- [19] M. Newman, “Modularity and community structure in networks,” in *Proceedings of the National Academy of Sciences*, vol. 103, pp. 8577–8582, June 2006.