自律的ネットワークにおける選択的攻撃の影響と媒介中心性を考慮した修復手法

有澤　俊裕* 大沢　英一†
(公立はこだて未来大学大学院)‡ (公立はこだて未来大学)§

1　はじめに

特定の目的で作られたセンサネットワークや，各種レーダーシステムなどに代表される大規模分散システムは，様々な事象の観測など多くの分野で利用が広がっており，今日の我々の生活において重要な役割を担っている．このような大規模分散システムはその用途の性質から，優れた耐障害性など“頑健”であることが求められる．そのため，ネットワークの構造を工夫するなどして耐障害性を向上させる研究がなされてきた [1][2][3][4].

ところで，2011 年に発生した東日本大震災以降，レジリエンスという言葉が注目されている．レジリエンス (Resilience) とは，Resistance と Recovery の 2 つの意味を併せ持つ単語である．Resistance は機能の維持を，Recovery は障害からの回復力をそれぞれ意味している．すなわち，レジリエンスを持つということは,「想定外」に対応するための備えを持つと解釈することができる．

本論文では，このレジリエンスという言葉を大規模分散システムに当てはめて考える．自然災害と同様に，システムの設計者が想像もしなかった障害が発生する可能性を否定することはできない．先に述べたように，システムの耐障害性を向上させる研究は多くなされてきた．しかし，攻撃などによってシステムの性能が低下してしまった場合，そこからいかにシステムを自律的に修復し，低下した性能の回復を図るかという問題に対する研究は十分なされているとは言いがたい．

本研究は，大規模分散システムのレジリエンス性の向上に貢献するため，選択的攻撃を受けたシステムの自律的回復手法を提案し，その手法の効果を検証することを目的とする．そこで，選択的攻撃が大規模分散システムにどの程度被害を与えるのか，自律的な性能回復としてどのような方法が有効なのかを考察した．また，本論文では自律的回復手法を,「末端にあるノードが，局所情報のみを用いて自らシステムの性能を回復させる手法」と定義する．

2　先行研究

本章では，本論文の内容に関係する先行研究について，その概要を述べる．

2.1　カスケード故障

ネットワーク内のあるノードの故障から過負荷故障が連鎖し，最終的に大規模な故障に至る現象をカスケード故障と呼ぶ．あるノードが担っていた負荷を周辺に肩代わりさせることによって，肩代わりしたノードも過負荷によって故障してしまい，さらにそれが連鎖することを表している．

林らはこの問題に対し，結合相関の正負がカスケード故障に与える影響について調査した上で，隣接ノード間のリンク張り替えによる防御戦略を提案した [1]．結果として，ネットワークの結合相関はカスケード故障に対し影響を与えるが，それだけでは被害規模は単純に決まらないことを示唆した．また，防御戦略についてはノードの負荷容量に関する耐久性に応じて，提案手法と従来手法を使い分けることが効果的であると論じている．

2.2　ネットワークの機能不全と構造最適化

谷澤は故障と攻撃の双方に強いネットワークについて考察している [2]．まず，故障 (ここではネットワーク内のノードがランダムで故障するとしている) に対しては，ハブの存在が有効であると述べている．ハブが低次数のノードを支配下に置き，更にハブ同士も連結しあうことによってランダムな故障に耐えうるネットワーク構造を実現する．一方で，ハブを狙った選択的な攻撃に対して脆弱になってしまうため，ハブを多く導入することはできない．

結論として，故障と攻撃の双方に対して最も頑強であるネットワーク構造は，二極次数分布構造を持つネットワークであるとしている．二極次数分布とは，2 つの次数 k_1，k_2 のみの次数分布である．しかし，このような極端な次数分布を持つネットワーク構造を持つネットワークは実世界に多く存在しているとは言いがたく，応用面での難点を持つ．

3　前提知識

本章では，本論文の内容で用いる技術や知識について，その概要を述べる．

3.1　複雑ネットワークの概要

実世界におけるネットワークの大多数は単純ではないネットワーク，すなわち複雑ネットワークである [5]．例えば，人間の社会的ネットワークやインターネット，神経細胞，交通網なども複雑ネットワークに属する．複雑ネットワークはレギュラーネットワーク (完全に規則的なネットワーク) とランダムネットワーク (規則性の見られ

* g2114002@fun.ac.jp
† osawa@fun.ac.jp
‡ 函館市亀田中野町 116-2 公立はこだて未来大学大学院 システム情報科学研究科
§ 函館市亀田中野町 116-2 公立はこだて未来大学大学 システム情報科学部複雑系知能学科

ないネットワーク) の中間のような構造をしている．その形は多種多様であるが，スケールフリー性やスモールワールド性など，共通する特性を見出すことができる．

また，複雑ネットワークには，ハブと呼ばれるノードが存在することが多い．ハブは他のノードと多く繋がっており，複雑ネットワークにおいて重要な役割を果たしている．このようなハブは複雑ネットワーク全体の平均経路長の短縮などに貢献しているため，選択的攻撃の対象になりやすい．次数の集中した上位 5 パーセントのノードが故障したとすると，系全体の平均経路長は約 2 倍にまで増大してしまうことがわかっている．

3.2 複雑ネットワークの特性

以下に，複雑ネットワークの特性について示す.

3.2.1 スケールフリー性

実世界に存在するネットワークのノードが持つ次数 (ノードに接合するリンクの数) の分布はべき乗則に従っていることが知られている．つまり，一部のノードは大きな次数を持っている一方で，大多数のノードは次数が小さいということを表している．べき乗則に従ったネットワークはスケールフリー性を有している．また，スケールフリー性を有するネットワークは，ランダムに発生する障害 (選択的に発生しない障害) に対して高い頑強性を有することが知られている．

3.2.2 スモールワールド性

スモールワールド性は，ネットワークの平均経路長がネットワークの規模に対して小さいという性質である．具体的な内容については後述するが，平均経路長とは任意のノード v_i から v_j に行くまでの最短距離の平均である．つまり，スモールワールド性を満たすということは，ネットワーク内の任意の 2 つのノードが，その中間にわずかな数のノードを介するだけで接続されているということである．

3.3 複雑ネットワークの統計的指標と中心性

本節では，複雑ネットワークの統計的な指標とネットワークの中心性について説明する．特に，本研究と関連するものに絞って述べる．

3.3.1 平均経路長

平均経路長とは，ネットワーク内の任意の 2 ノード間の最短距離の平均を表す指標であり，ネットワークのおおよその大きさを表現することが可能である．平均経路長 L は以下の式で求めることができる．

$$L = \frac{\sum_{i=1}^{n}\sum_{j=1}^{n} l(i,j)}{{}_nC_2} \tag{1}$$

L がネットワーク内の全ノード数 n に対して，$log(n)$ に比例する程度であれば，スモールワールド性を満たすと定義されている．

3.3.2 次数分布

次数分布は，ある次数 k を持つノードの分布を表したものであり，ネットワークの持つ性質によってその分布が異なる．先述の通り，この分布がべき乗則に従えばスケールフリー性を満たす．また，ランダムネットワークの次数分布は正規分布であり，特定のピークを持つことが知られている．

3.3.3 媒介中心性

媒介中心性は，情報を橋渡しすることに関与している度合いを示すものである．ネットワーク内の任意の 2 ノードを選び，それぞれの組み合わせの最短経路内にあるノードが存在する確率によって決定される．ノード v_i の媒介中心性 B_i は，始点 v_{i_s} から終点 v_{i_t} の最短経路の中で v_i を通るものを $g_i^{(i_s i_t)}$ とし，最短経路の総数を $N_{i_s i_t}$ とした時，以下の式で算出する．

$$B_i = \frac{\sum_{i_s=1;i_s\neq i}^{n}\sum_{i_t=1;i_t\neq i}^{i_s-1}\frac{g_i^{(i_s i_t)}}{N_{i_s i_t}}}{(N-1)(N-2)/2} \tag{2}$$

また，媒介中心性の高いノードが破壊された場合には平均経路長の増大など，ネットワーク内の情報伝達に大きな影響が生じる可能性が高い．

3.4 BA モデル

BA(Barabasi-Albert) モデルとは，ネットワーク生成モデルのひとつであり，複雑ネットワークの特性を有する [6]．以下に，生成の手順を示す.

1. ノード数 m_0 の完全グラフを用意する．
2. m 本のリンクを持ったノードを 1 個追加する．ただし, $m \leqq m_0$ とする．
3. 追加したノードと既存のノードを結びつける．ただし，同じノードから 2 本以上のリンクをつなぐことは禁止とする．また，新規ノードと結合するノード v_i の選択は，v_i の次数を k_i，既存ノードの数を n として以下の式によって確率的に決定する．

$$\prod(k_i) = \frac{k_i}{\sum_{j=1}^{n} k_j} \tag{3}$$

4. 目的のノード数になるまで 2.，3. を繰り返す.

この BA モデルでは，(2) 式により，次数の高いノードに新しいリンクが加わりやすいとしている．これを優先的選択という．優先的選択をしている理由は，等確率で新たに結びつくノードを選択すると，スケールフリー性を満たさないネットワークが生成されるためである．この優先的選択によって，一度次数が高くなったノードは

その後も新しいリンクを獲得しやすく，ハブになりやすい．このモデルでは，ノードの次数分布がべき乗則に従い，スケールフリー性を満たす．また，平均経路長は短くなり，スモールワールド性も満たす．しかし，BA モデルはクラスター性は満たさない．

4 実験モデル

本章では，シミュレーションのための実験モデルや，修復手法の詳細について説明する．本実験モデルは，実世界のセンサネットワークを想定したものである．ただし，シミュレーションの単純化のために，センサネットワークの設置における地理的な制約や物理的な距離などについては考慮せず，あくまで数理的なモデル上でのシミュレーションを行うものとする．

4.1 各エージェントと環境のパラメータ

ノードをエージェント a_iとみなし，その集合を $A = \{a_1, a_2, ..., a_n\}$ とする．ここで各エージェントはセンサネットワークにおける，末端のセンサに相当する．また各エージェントは，タスクを処理実行する上での能力に対応する R_iを持つものとする．

処理能力 R_i は，エージェント a_i が単位時間あたりに処理できるタスクの量を表す．この処理能力は，攻撃前のネットワークにおける媒介中心性を元に決定する．エージェント a_i の媒介中心性を B_i としたとき，以下の式に従って R_i を決定する．

$$R_i = (1+\alpha)B_i, i = 1, 2, ..., n \quad (\alpha\text{は定数}) \quad (4)$$

媒介中心性の値が高いほど，処理能力も高くなる．これは，実世界のセンサネットワークにおいて媒介中心性が高いノードは処理能力を高く設定しているという前提での設定である．また，ネットワークを修復する際，能力の高いものに張り替えが集中しないようにするための許容量 $CA(a_i)$ を設定する．詳細については後述する．

タスク $T = \{t_1, t_2, ..., t_m\}$($m$ はタスクの総数) は各エージェントに対する環境情報取得要求に相当する．本実験モデルではタスクの処理内容を，タスタを割り当てられたエージェントから距離 2 以内のエージェントを協力者と定め，協調してセンサから環境情報を取得し，出力することと想定した．タスクの処理を完了させるためにはコスト $Cost(t_i)$ を要するものとする．エージェント a_iから距離 2 位内のエージェントの数を $NE(a_i)$ とし，$Cost(t_i)$ を以下の式で定義する．

$$Cost(t_i) = \beta NE(a_i), i = 1, 2, ..., n \quad (\beta\text{は定数}) \quad (5)$$

センサネットワークは，各ノードの接続関係を表す無向グラフ $G = (V, E)$ で表す．エージェント a_iは自身から距離 2 位内のエージェントと通信が可能であるとする．それ以外の場合は通信が行えず，各種情報伝達は不可能であるとする．ただし，攻撃によってエージェントが孤立してしまった場合には，故障したエージェントの近傍からランダムで接続できるものとする．また，本実験ではシミュレーションの単純化のため，通信遅延や，情報伝達に必要なコストなどについては考慮しない．

4.2 シミュレーターと各エージェントの動作

本研究では，センサネットワークの動作によって提案手法の性能を評価する．センサネットワークは Fig. 1 のような Step により動作する．以下に，シミュレーターと各エージェントの動作を示す．

Step1 タスクの割り当て
: 各エージェントが処理するタスクは，ネットワーク内のエージェントにランダムで割り当てられる．タスクにかかるコストは上記の式に従って決定する．

Step2 環境情報の取得
: 割り当てられたエージェント a_i は，自身から距離 2 以内に存在するエージェント (協力者と定義する) に対し，環境情報 (各エージェントがセンサから取得できる情報) を取得するように指示を送信する．指示を受けたエージェントは，センサから環境情報を取得する．

Step3 環境情報の統合
: 協力者となったエージェントは環境情報を取得した後，その環境情報をエージェント a_i に送信する．

上記の 3 つの Step を完了するまでにかかる時間を処理時間とする．

4.3 効率の指標

この実験モデルの効率は，(1)1 タスクの平均処理時間，(2) 平均協力者数，(3) 平均処理能力で評価する．ここでの処理能力とは，協力者数を処理時間で割った値とする．この指標は，協力者数によってコストが変化することを考慮して設定した．さらに，ネットワークの統計的指標についても比較する．

4.4 ネットワークの修復手法

ネットワーク修復の手順を以下に示す．

Step1 攻撃の検知
: 攻撃を受けたエージェント (仮に a_i とする) の近傍に位置するエージェント $V_j(a_i)$ は，次数の変化によって攻撃を検知する．

Step2 張り替え先の決定
: エージェント $V_j(a_i)$ は，自身から距離 2 に位置するエージェントの中から許容量 CA の最も高いエージェント (仮に a_j とする) を選択し，接続する．その後，a_j の許容量 $CA(a_j)$ を γ だけ減少させる．ただし，γ は定数とする．

ここで，Step2 の 許容量 $CA(a_j)$ について説明する．常に許容量の最も高いエージェントを張り替え先として

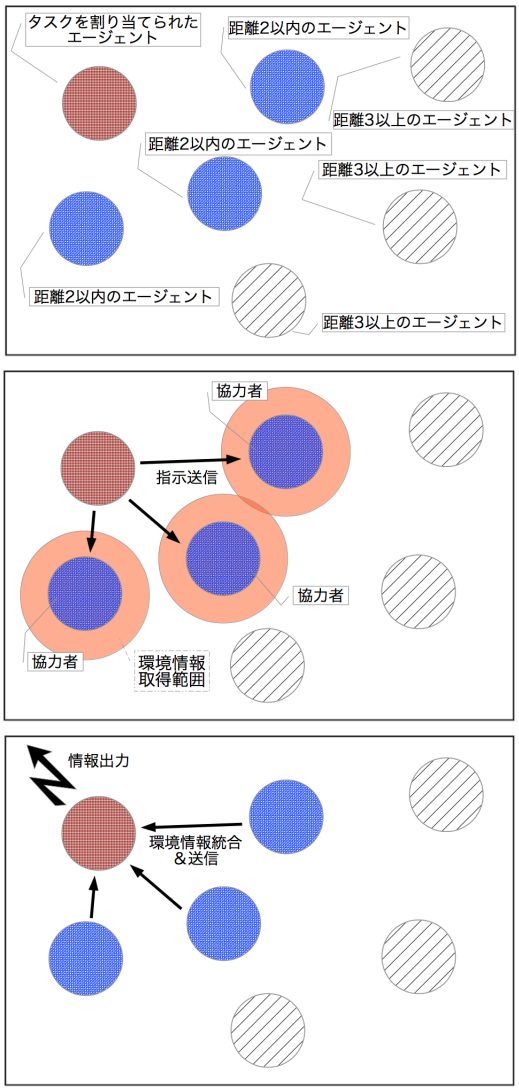


Fig. 1 シミュレーターの各ステップ

選んだ場合，同じエージェントに張り替えが集中してしまう可能性がある．このような状況では，カスケード故障など他の障害を引き起こしてしまうことが考えられる．そこでこのような許容量を設定し，張り替えの際にその許容量を下げることで次回修復時の張り替え先候補として選択されにくくすることにより，リンクの集中を防ぐものである．

5 実験

攻撃によるネットワークの性能低下，および提案する修復手法の性能を評価するための実験を行った．以下にその詳細を示す．

5.1 実験のパラメータ

ネットワークはノード数 N=1000 とし，BA モデルを用いて生成した．ネットワークを生成するパラメータは，m_0=7，m=4 とした．タスクの総数は 1000 とし，タスクのコストは前章の式に従い決定する．また，単位時間毎のタスク生起数は 10 と設定した．攻撃を行う個数 BR は 1,2,10,20,30,40,50,100 とし，次数上位 5%の中からランダムで $BR/2$ 個，同じく処理能力上位 5%の中から $BR/2$ 個に対して攻撃を行った．ただし，BR=1 の時は次数最上位のエージェントに対し攻撃を行った．

5.2 実験結果および考察

各実験結果と考察を以下に示す．

5.2.1 平均処理時間の比較

平均処理時間を Fig. 2 に示す．Fig. 2 の平均処理時間の結果を見ると，攻撃直後と修復後には大きな違いが見られなかった．また，攻撃個数が 100 個を超えた場合，処理時間が大きく増大することが観測された．しかし，先述の通りタスクの処理にかかる時間は，協力者数に影響されることに注意されたい．後述の平均協力者数の結果を見ると，修復後のほうが協力者が多い．つまり，修復後の方が，より大きなコストのタスクを処理しているということになる．

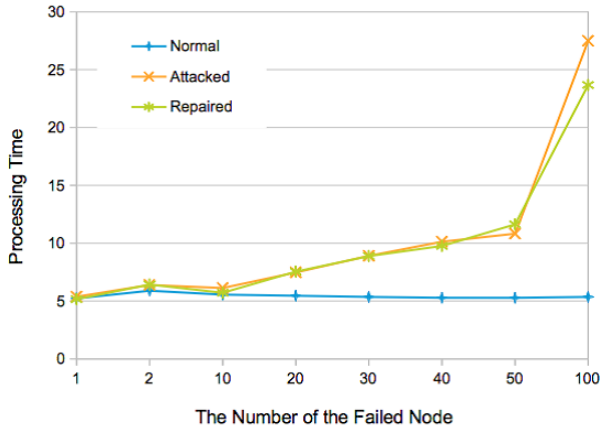


Fig. 2 平均処理時間

5.2.2 平均協力者数の比較

平均協力者数を Fig. 3 に示す．Fig. 3 から，修復を行わなかった場合，協力者数の減少が著しいという結果が得られた．修復を行った場合，協力者数は攻撃前と同じか，あるいは改善が見られた．これは，修復を行ったエージェントが攻撃前よりも次数の高いエージェントにリンクを張り替えることにより，協力者数を増加させることが可能となったことが理由と思われる．しかし，攻撃個数が 100 個の場合に協力者数が減少していることから，回復には限界があると考えられる．前提知識で述べたとおり，スケールフリーネットワークは次数の上位 5%が破壊された場合に大きく性能が低下することがわかっている．本実験の結果でも，その影響が現れたと考えられる．

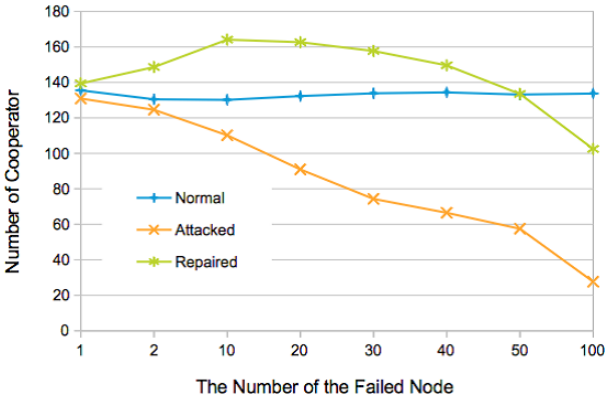


Fig. 3 平均協力者数

5.2.3 平均処理能力の比較

平均処理能力を Fig. 4 に示す．Fig. 4 より，修復手法を適用することにより，ある程度処理能力の回復が見られる．しかし，攻撃前の能力と同等までに回復しているのは攻撃個数が全体の 1%の場合までで，それ以降は攻撃直後よりは高い数値を示しているものの，処理能力が低下してしまうという結果が得られた．処理能力は，協力者数を処理時間で割ることにより算出している．本実験結果では協力者数の上昇が見られているため，処理時間を改善することができればこの処理能力も改善されると思われる．よって，処理時間を改善するための対策が必要となる．

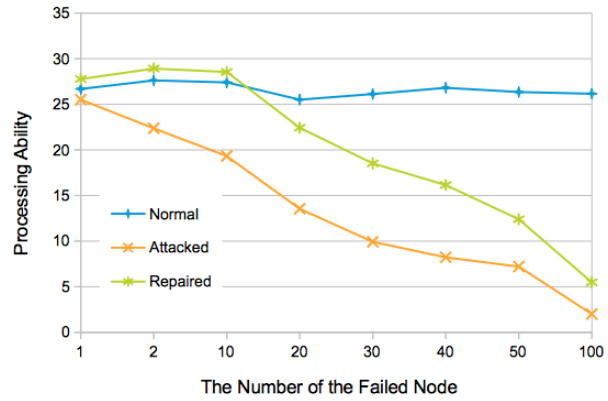


Fig. 4 平均処理能力

5.2.4 次数分布の比較

Fig. 5 の (a)，(b)，(c)，(d) にノードの破壊個数を 1 個，30 個，50 個，100 個と設定した時の次数分布を示す．ノードの破壊個数の増加に伴って，攻撃後の次数分布が変化していることがわかる．ハブなどの高次数のノードが破壊されることにより，低次数のノード (例えば次数が 10 以下のノードなど) が増加するという結果はどの場合においても見られるが，破壊個数が多い場合はその傾向が強くなることが見て取れる．しかし，攻撃前と修復後の次数分布は，破壊個数 100 個の場合を除き，ほとんど違いが見られない．この結果は，修復手法は破壊されたハブの代わりとなるノードを作り出し，その構造を保つことに成功したからであると考えられる．ただし，破壊個数 100 個の場合に次数分布が攻撃前とは異なる形を見せている．よって，平均協力者数の比較で述べたように修復手法で対応できる破壊個数には限界があると考えられる．

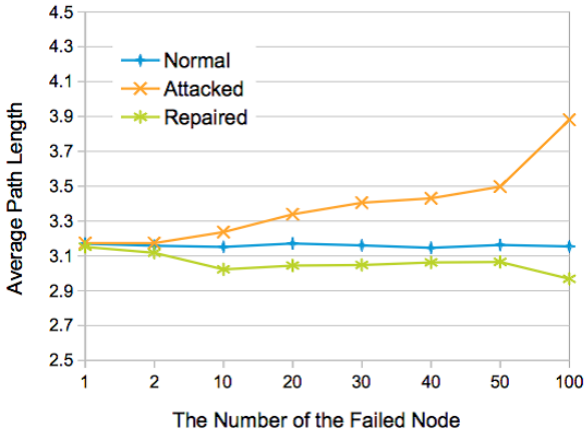


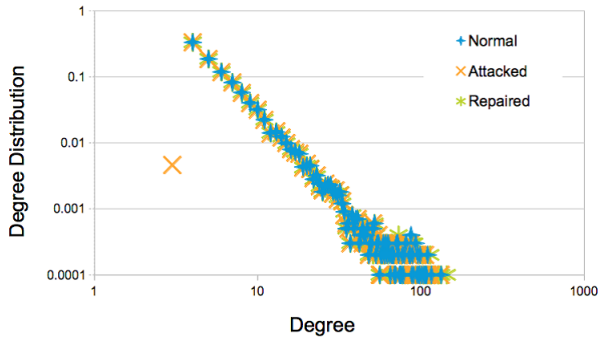
Fig. 6 平均経路長

5.2.5 平均経路長の比較

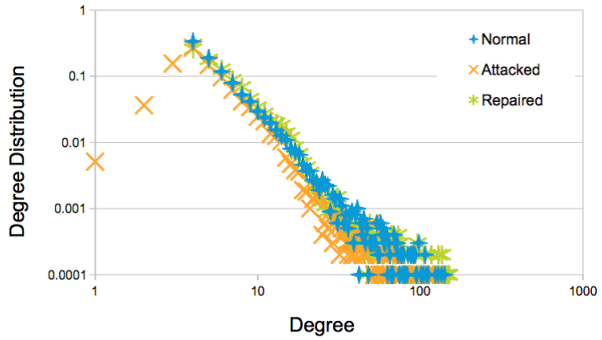
Fig. 6 に平均経路長を示す．平均経路長を比較すると，全てにおいてスモールワールド性を満たしているものの，攻撃後では平均経路長が長くなっていることが確認できる．逆に修復後では攻撃前よりも平均経路長が短くなっていることがわかる．これは平均協力者数の比較でも述べた通り，リンクの張り替えによって元々繋がっていたノードより平均経路長の短縮に貢献するノード (例えば高次数のノードや媒介中心性の高いノードなど) に繋がることが可能となったためであると考えられる．しかし，攻撃によって平均経路長が増加しているとは言え，予測よりもその増加量は少なかった．これはネットワークの持つ次数分布に影響されるものと考えられる．次数分布がより極端な分布であれば，選択的攻撃によって平均経路長は今回の実験結果よりも増加すると考えられる．よって，様々な次数分布のネットワークを用意し，実験を行う必要がある．

5.2.6 全体の考察

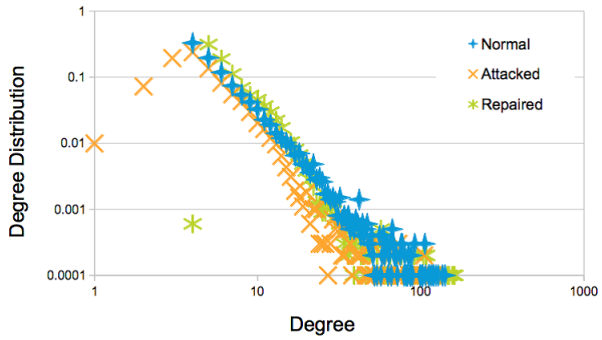
以上の結果より，攻撃個数がネットワーク全体の 1%程度であれば，攻撃前の性能を維持することが可能であることが示された．また，1%以上破壊された場合には，破壊個数の増加とともに性能も低下することがわかった．統計的指標の観点からの比較では，破壊個数 50 個までは攻撃前と修復後でほぼ同じ次数分布しているのに対し，



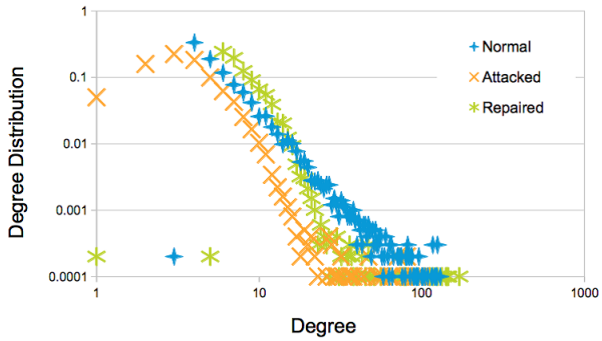
(a) 破壊個数：1



(b) 破壊個数：30



(c) 破壊個数：50



(d) 破壊個数：100

Fig. 5　次数分布

破壊個数 100 個では分布に変化が見られた．このことから，修復手法で対応できる破壊個数には限界があると考えられる．今後の課題としては，修復手法の更なる改善による平均処理時間の短縮と，処理能力の向上が挙げられる．

6　まとめ

本論文では，選択的攻撃が大規模分散システムに与える被害の規模の調査と，どのような方法が自律的な性能回復として有効なのかを考察した．実験では，提案手法によって協力者数を維持，あるいは増加させることが可能であるという結果が得られた．

本研究では，BA モデルのみ用いてネットワークを構成したが，提案手法の有用性を確認するためには，様々な次数分布を持つネットワークや，生成モデルの異なるネットワークを用いて実験を行う必要がある．さらに，各パラメータの設定によっては，結果に違いが見られる可能性もある．今後は，本実験と異なるネットワーク構造を実験環境に設定し，各種パラメータの設定を変更して実験と考察を行う予定である．さらに，その結果を踏まえてより良い自律的性能回復手法の模索を行いたい．

参考文献

[1] 林　幸雄，宮崎敏幸，“結合相関を持つ Scale-Free ネットワーク上のカスケード故障に対する防御戦略，” 情報処理学会論文誌，vol.47, No.3, pp.802–812, March 2006.

[2] 谷澤俊弘，“故障と攻撃の両方に強い繋がり方とは?–ネットワークの機能不全と構造最適化–，” 情報処理学会論文誌，vol.49, No.3, pp.282–289, March 2008.

[3] 山本聡彦，生天目　彰，“進化手法による最適同期ネットワークの設計，” 情報処理学会研究報告，vol.2009-MPS-75, No.10, September 2009.

[4] 小林直樹，白山　晋，“ネットワークトポロジーの最適化とその過程の分析，” JAWS2010 論文，2010.

[5] 増田直紀，今野紀雄 “複雑ネットワーク–基礎から応用まで–，” 近代科学社，2010.

[6] A.L. Barabasi, R. Albert ,“Emergence of scaling in random networks,” Science, vol.286, pp.509–512, 1999.