2007年夏の課題
氏名:奥村　浩徳

～3本の論文の要約～

①. Vulnerability of weighted network

1.はじめに

　複雑ネットワークの構造はインターネットや道路網、遺伝子のつながりなど様々な分野に見られるが、それらが外部からの攻撃に対してどの程度の耐性を持つかは非常に興味深いテーマである。ここではWAN（Worldwide Air transportation Network) を例にとり、色々な”測度”からネットワークに攻撃を加えその変化を観察してみる。

2. 準備

　用意したデータは2002年の世界の主要な空港のつながりと距離、その路線で運ぶことのできる人数をあらわしたものである（出典:International Air Transportation Association) 。このデータでは、

N(空港の数) = 3880, E(辺の数)=18810, ⟨k⟩(平均次数)=9.7, ⟨l⟩(平均頂点間距離) = 4.4

となっている。また次数 k のノードの数はベキ乗則に従うというスケールフリーネットワーク構造になっている。

3. 様々な”測度”

　最初に、空港 i と j の距離を $d_{ij}(=d_{ji})$ 、運ぶことのできる人の数を w_{ij} （$=w_{ji}$）とする。あるノード i に与えられた重み S_i を次のように定義する。

$$S_i = \Sigma_{j \in V(i)} w_{ij}$$

ここで V(i)はノード i に隣接するノードの集合である。また、ノード i に与えられた距離の総和 D_i 、並びに距離と重みの積の総和 O_i を以下のように定義する。

$$D_i = \Sigma_{j \in V(i)} d_{ij} \quad , \quad O_i = \Sigma_{j \in V(i)} w_{ij} \times d_{ij}$$

　ノード h からノード j に行くときの最短距離を σ_{hj} として、そのうちノード i を通る

ものを $\sigma_{hj}(i)$ とする。またグラフに重みがついた時にはそれぞれ σ^w_{hj} , $\sigma^w_{hj}(i)$ と表す。このときノード i の中心性 b_i 、重み付き中心性 b^w_i を以下のように定義する。

$$b_i = \sum_{h,j} \frac{\sigma_{hj}(i)}{\sigma_{hj}} \quad , \quad b^w_i = \sum_{h,j} \frac{\sigma^w_{hj}(i)}{\sigma^w_{hj}}$$

ここで h と j は $h \neq j$ となる全てのノードである。

4. ”測度”による違いについて

3. で定義した測度はWANの分析にどのような違いをもたらすのか？Table1は色々な測度によってノードに順位をつけて、それが異なった測度の順位とどの程度一致しているかを示す表である。

この表における τ は以下の式で定義される。

$$\tau = \frac{|n_c - n_d|}{n(n-1)/2}$$

この式の意味は次のようである。ある1つの測度によるノードの順位のリストと、別の1つの測度によるノードの順位のリストがあった時に、一致しているノードのペアの数を n_c 、異なっているノードのペアの数を n_d として、それをペアの総数で割れば測度による違いが数値として出てくるというものである。

表では 0.2~0.3 などの小さな数値もでてくるが、3880 個のノードのランダムなリストを2つ作ってその τ を求めると大体 0.01 となるので 0.2 や 0.3 といった数値でも実は強い一致を示している事になる。

5. ネットワークへの攻撃

様々な測度からWANに攻撃を加えたときに、その構造がどのように変化していくかを見ていく。まずはネットワークの全ノードの何割に攻撃を与えたかを表す記号 g を用意する。攻撃を加えるときは各測度に基づくノードの順位のリストの高いものから攻撃する。このとき1攻撃ごとにノードの順位を計算しなおし、その新しくなったリストの上位から攻撃を加える。

また、WANの変化を見るために $I_O(g) = O_g / O_0$ という測度を用意する。ここで $O_0 = \Sigma_i O_i$ であり、 $O_g = max_G \Sigma_{i \in G} O_i$ （Gは攻撃後に残っているネットワークを表す）である。

Figure2の上図におけるWANの振る舞いを見てみると、各測度ごとにネットワークの生き残りの様子はかなり違ってみえる(N_g / N_0 は攻撃後につながっている最

大のネットワークのノードの数を、初期のネットワークのノード数で割ったものである)。BC や WBC に基づく攻撃がネットワークのつながりを早い段階で断ち切ってしまう様子が観察できる。

しかしFigure2の下図を見てみると、$I_O(g)$ はどれも同じような割合で低下している事が分かる。つまり攻撃後に残る最大のネットワーク内の O_g はどのような測度から攻撃を受けても変わらないという事を示している。さらに g = 0.04 付近で k,s,O,D の N_g/N_0 の値は約 0.8 となっているにも関わらず、$I_O(g)$ の値はかなり 0 に近くなっていることが分かる。つまり見かけ上はネットワークが生き残っていても、人を運ぶ距離とその人数という観点から見ればWANは多大な被害を受けていることが分かる。なお、$I_O(g)$ は距離と重みの積の和という測度からつくられているので、単純に距離だけや重みだけで受けるダメージを調べてみても大体同じような結果になる。

5. ノードの順位の再計算について

Figure2では攻撃毎に各測度に基づく順位の再計算を行ったが、最初に与えられたノードの順位のリストだけで攻撃を行ったらどうなるのであろうか？

Figure3の上図ではその違いが見て取れる。O でも BC でも再計算を行ってネットワークに攻撃を加えたほうが、つながりをより早く断ち切ることができる。O に対して BC の再計算後の受ける影響が大きいのは、各ノードの中心性が他のノードの存在に大きく左右されるためと考えられるがここではあまり深入りしない。

むしろ重要なのはFigure3の下図において再計算したものもしないものも、ほぼ同じように減少をしているということである。Figure2の下図と同じように g = 0.04 で $I_O(g)$ はほとんど 0 に近づく。これはノードの順位の再計算をしようがしまいが、主要なノードをいくつか消すだけでネットワークに多大な被害を与えることができるという事を意味している。

6. まとめ

以上のことをまとめて考えると、WANに大きな損害をもたらすためにはある1つの”測度”を用いた最初のノードの順位のリストがあればそれで充分ということになる。そのリストの上位のノードから順に攻撃を加えていけば、WANの実質的な機能はすぐに停止する。逆に守る側としては全ての”測度”に応じた攻撃に対処しなければならず、またあるノードが潰されたときの順位の再計算も行わなければならないので、極めて防御しにくいという結論が導かれる。現実にこのような事態になるかどうかはわからないが、少なくとも空港関係者は空港ネットワークが思った以上にダメージを受けやすいものであるということを認識しておく必要がありそうである。

②. Congestion-gradient driven transport on complex networks

1．はじめに

ネットワークのあるノードに目的地を持たせた一つの粒子を発生させて時間毎に動かしていった場合、直感的には一番短い距離を辿っていけば一番早く目的地に着きそうである。しかしそれはネットワークに一つの粒子しかないような場合で、ネットワーク上に多くの粒子が存在する場合ハブに粒が集中して混雑がおこる。この論文では各粒子にどのような経路をとらせればネットワークの混雑を回避できるかという事を探っていく。

2．設定

ネットワークのモデルとしてはランダムネットワークとBAネットワークの２つを用いる。ランダムグラフの各ノードが繋がっているかどうかはノードの総数をＮとして、p = 6/(N-1)で決まるものとする。一方BAモデルでは初期のノード数を7とし、時間毎に3つのリンクをもつ新しいノードが追加される。そのリンクがノードｉに結びつく確率はそのノードの次数を k_i として、

$$p(k_i)=\frac{k_i}{\sum_{j=1}^{n} k_j}$$

で決まる。

各ノードは待ち行列を持ち、FIFOルール（先に来たものを先に処理する）を適用する。時間毎にランダムな1つの粒子が処理される（別のノードに移動する）か、ある確率でどこかのノードに発生するかどちらかである。粒子は目的地に到達するとネットワークから取り除かれる。

この論文でネットワークが混雑しているというのは、粒子の発生に対して粒子の消滅が低い状態を意味している。つまり時間のステップを限りなく大きくしていった時にネットワーク上の粒子の数が一定とはならずにどんどん増大していくような状態である。

ある粒子がノードｉからｊに移動する確率は、

$$P_{ji}=\frac{A_{ji}(q_j+1)^{-\beta}}{\sum_{k=1}^{N} A_{ki}(q_k+1)^{-\beta}}$$

で表される。ここで、 q_j, q_k はそれぞれノードｊとｋの待ち行列、 A_{ji} はノードｊとｉがつながっていれば1、それ以外は0となる隣接行列成分である。またβは任意に

設定できる数値で、β の値を大きくするほど待ち行列の大きいノードに粒子が移動する確率を低くできる。これを”迂回ルール”と呼ぶことにする。

３．シミュレーション

ランダムネットワークとBAモデルネットワークを用意して、R（粒子の発生確率）とβを調節しながら100回のシミュレーションを行った結果がFIG．１である。縦軸は100回の試行のうち何割が混雑した状態になったかを表している。

この結果を見ると β が 0 の時にネットワークが最も混雑しやすく、β が 1 のときに最も混雑しにくいことが分かる。また、傾向としては β が 0 から 1 まで増加すればネットワークは混雑しにくくなるが 1 を超えると徐々にキャパシティが下がるようである。

さらにFIG．２ではノードの総数 N と β の値によって R*（ネットワークの混雑回数が5割になるような R）がどのように変わるかが見て取れる。まずランダムモデルの方がBAモデルよりも、高い R に対して耐性をもっている事がわかる。そして（a）、（b）共にノードの数の増加に伴って R*が減少していく。またFIG．１でも観察されたが β が 1 の時にネットワークは一番混雑しにくい状態になるようである。

４．中心性を用いた分析

粒子がノード s から t へ向かうときの、ノード i の中心性を定める。ある粒子がノード j から i に移動する確率は $P_{ij}=A_{ij}/d_j$ で与えられる。 d_j はノード j の次数である。 P_{ij} を個々の成分としたN×N行列 Ｐ が、ある粒子が次の移動でどのノードに移るかを示す確率行列となる。さらにその粒子が目的地 t にたどり着いたときに Ｐ の t 列成分を全て 0 にすれば、その粒子はノード t から移動しないことになる。こうして得られた確率行列を P^t とする。ノード s から出発してノード t に向かう1つの粒子が n ステップの中でノード i を通過する確率は $(P^t)^n_{it}$ （ P^t を n 回かけてその s、t 成分を抜き出したもの）であらわされる。これを用いて出発ノード s、目的地ノード t に関するノード i の中心性は

$$b_i^{st}=\sum_{n=0}^{\infty}(P^t)^n_{is}$$

で定義される。ただし $(P^t)^0$ は単位行列 I とする 。さらに、一般的な中心性はノード s と t をそれぞれ 1 から N まで変化させてその平均をとればいいので、

$$b_i=\frac{1}{N^2}\sum_{s,t=1}^{N}\sum_{n=0}^{\infty}(P^t)^n_{is}$$

で定義される。

FIG．3とFIG．4はノードの中心性によって流れ込んでくる粒子の数の平均〈w〉

とその時の待ち行列の平均$\langle q \rangle$を表したものである。$\langle w \rangle$の最大値が 1 なのは、少なくとも一つのノードが平均して 1 個以上の粒子を受け取っていればネットワークが混雑している状態だと見なせるからである。ノードの数による変化はあまり無いの N=30 で固定して、100 個のネットワークに対して実験を行った。タイムステップは 10^5 とした。FIG．3はランダムネットワーク、FIG．4はBAネットワークに対する結果である。またFIG．3と4の β の値は(a),(b)で 0、(c),(d)で 0.5、(e),(f)で 10 となっている。ここで注意しておかなければならないのは中心性の計算のときには単純なランダムウォークによって粒子を移動させていたが、実際には2．の P_{ij} の式に基づいて粒子は移動するという点である（β =0 の場合を除く）。これはコンピューターによる計算が煩雑になるのを避けるためと、中心性をランダムウォークで計算したものと P_{ij} で計算したものとの間に強い相関関係があるためである。

さて、FIG．3と4の(a)と(b)を見てみると誤差を含めてある一定の線上に点が乗っていることが分かる。FIG．3の上の点の集まりは R=0.4、FIG．4の上の点の集まりは R=0.3 とした時のものであり、下の点の集まりは両方とも R=0.1 の時のものである。そして実はFIG．3と4の(a)の線の傾きは R の値と一致している。よって中心性が 1/R となるノードがあるとそのノードに平均 1 以上の粒子が訪れることになり、ネットワークは混雑する。またFIG．3と4の(b)の上下の線は $\langle q \rangle = \frac{\langle w \rangle}{1-\langle w \rangle}$ で求められるが、ここではあまり重要なことではない。むしろ大切なのは(a)に対して(c)と(e)（FIG．3とFIG．4共に）の方がバランスよくノードに粒子を振り分けてネットワークの耐性を上げている点である。(a)で上の線に集まっていた点は、(b),(c)では下の点の集まりとなっているのだが基本的には$\langle w \rangle$が 1 以上になる点は現れていない。これは中心性の低いノードにもある程度の粒子を移動させることによって、中心性の高いノードに対する負荷を和らげているからである。そしてもう一点、β を 0.5 から 10 に上げた時ランダムネットワークとBAネットワーク共に低い R に対してはネットワークのキャパシティが上がっているようだが、高い R に対しては逆にキャパシティが下がっているというのも重要な点である。これはつまりネットワークに多くの粒子を発生させた場合、中心性の高いノードを避けるようなモデルをつくると中心性の低いノードへの負荷が大きくなりすぎて逆にネットワークが混雑してしまうという事を意味している。この傾向はランダムネットワークよりもBAネットワークの方がより顕著である。

これはFIG．5によっても確認することができる。FIG．5は N=30 のBAネットワークにおいて R=0.55 で粒子を発生させた場合、それぞれの粒子がどのくらいの時間で目的地にたどり着けるかを確率で表したものである。タイムステップは 30000 として、(a)では β=1、(b)では β=10 とする。見て分かるようにほとんどの粒子が 2000 ステップの内に目的のノードの到達しているが(a)にくらべて(b)はより長いステップで目的地に到達する粒子も存在するようである。これからわかることは混雑を回避しようとして、待ち行列の長いノードに行かないようにすると返って目的地までの道のりが長

くなってしまい余計に時間がかかるという事である。しかしこれだけで混雑の仕組みの全てが説明できるわけでなく、次の章で見るようにネットワークの”トラップ”もまた混雑の原因となっているのである。

5．ネットワークのトラップについて

ネットワークの平均次数が低い時にはある種のトラップが発生して、全体の粒子の流れを滞らせてしまう場合がある。いま、N=30、R=0.65、〈d〉=6として50個のランダムネットワーク上で粒子の動きを観察した。その結果46個のネットワークで”ピンポントラップ”が、全体の粒子の流れを阻害しているのが分かった。

FIG．6は2個のノードによるネットワークのピンポントラップを表したものである。次数の高いあるノードにリンクが1つしかないノードがくっ付いていてるもの（1と5）がトラップである。この状態で1から5に粒子が移動した場合、5が目的地でない粒子は再び1へと返ってくる。βの値が低ければその粒子はその後1から5以外のノードに移る可能性が高いが、βの値が低いと待ち行列がある他のノードには行かずに再び粒子は5に移動するという状態が起こる（ピンポン現象）。こうして1と5の間だけで粒子のやり取りが発生して、目的地にたどり着けない粒子が1にどんどんたまっていくという状態が続いてしまうのである。FIG．5は全粒子のうち何割がどのノードに存在するかを時間ごとに表したものであるが、予想通りに1が最も高く5が最も低くなっている。注目したいのは1の値が高くなっている時には5は低くなっていて、5の値が高いときには1の値は低くなっている点である。さらに他のノードの数値の変化は時間が経つにつれてあまり見られなくなる。これは1と5の間でしか粒子のやり取りが行われていない状態の裏づけデータとなっている。そして5の待ち行列が増えると”迂回ルール”により、1や5を目的地とした粒子は5のノードを避けるようになる。その結果FIG．8に見られるように、ネットワーク上に1と5を目的地とした処理されない粒子がたまっていくのである。これがランダムネットワークのピンポントラップの仕組みである。

BAネットワークで同じ条件で実験を行うと”トライアングルトラップ”が多く観察される。BAモデルでは基本的に次数の高いノードに新しいノードがくっ付いてネットワークが成長していくのであるが次数の低いノード同士の結合もたまに発生する。そうするとFIG．9に見られるように、次数の高いノードに次数の低い2個のノードがくっ付いてトライアングルを形成することになる。これもβの値を上げていくと”迂回ルール”が働き、2個のノード間（11と20）の間でしか粒子のやりとりが行われなくなる。その結果11と20には目的地にたどり着けない粒子がたまっていくことになるが、そこそこ11と20の待ち行列が長くなると次数の高いノード（FIG．9の上の部分のノード）からは11と20に粒子を送らなくなり、待ち行列の小さいノード（FIG．9の下の部分のノード）に粒子を送るようになる。そうすると11と20を目的地とした粒子は処理され

ずにネットワークにどんどんたまっていきネットワークは混雑状態になる。これがトライアングルトラップの仕組みである。

　FIG. 12から分かるようにランダムネットワーク、BAネットワーク共にβの値を大きくしていくと”迂回ルール”が強く働き、トラップを形成しているノードを目的地とした粒子の単位時間当たりの増加率が大きくなっていくことが分かる。またβが0.5~1.0の時にどの目的地へ行く粒子の増加率も低く抑えられてことがわかる。

6. まとめ

　これまでの観察においてβ=1辺りで”迂回ルール”を適応させればネットワークが最もうまく機能することが分かった。完全にランダムな移動のさせ方はネットワークの混雑を引き起こすが、待ち行列のあるノードを避けすぎるとそれも混雑状態をよぶということである。要はバランスの問題である。またネットワークのトラップについても分析したが、これらのトラップは平均次数が高ければ基本的には形成されない。よってノードの平均次数を高くするというのも混雑緩和策としては有効である。

　しかしβ=1付近がなぜ最適な値なのかについての理論的な分析はこの論文ではきちんとなされていない。これは迂回ルールの式のつくり方が大きく関係していると思われるが、そのあたりの事は別の文献や論文を参照するのが望ましい。

③. Traffic dynamics based on local routing protocol on a scale-free network

1. はじめに

この論文の主旨は②と大体似ているが、若干違う設定でシミュレーションを行っている。そしてなぜ②において β=1 で（この論文では α=-1 で）ネットワークのキャパシティが最大となるのかについてある回答が与えられている。

2. 設定

ネットワークのモデルはBAモデルとし、新たなノードは5つリンクを持ってネットワークに繋がるものとする。パケットは1ステップに R 個、ネットワーク上に目的地をもって発生し、そこに到着知ればネットワークかた取り除かれる。またノードは1度通ったリンクを2度通過することは無いものとする（PIAルール）。各ノードは待ち行列を持ち、それぞれのノードが一度に C 個のパケットを処理できるものとする。各ノードのパケットはFIFOルールに従って処理される。ノード h にあるパケットが繋がっているある1つのノード i に移動する確率は

$$\Pi_i = \frac{k_i^{\alpha}}{\sum_j k_j^{\alpha}} \quad \text{（j は h につながっている全てのノード）}$$

で決まる。ただし繋がっているノードの中に目的地があれば、パケットは直接そこに移動するものとする。

3. シミュレーションと理論的な解析

FIG. 1はノードが 0 から 1000 になるまでシミュレーションを行った時の R に対する η の変化を表したものである。各ノードの処理能力は C=10 で固定した。ここで η は、

$$\eta = \lim_{t \to 1000} \frac{C}{R} \frac{\langle N(t+\Delta t) - N(t) \rangle}{\Delta t}$$

で定義される。N(t)はその時間におけるパケットの総数で、⟨⟩は各 t についての平均をとったものである。この式の正確な意味は分からないが、ある時間からネットワークを少しだけ進めたときに総パケット数が増加していれば N(t+Δt)-N(t)の値も増加するので、そのまま時間を進めていくといづれネットワーク上にパケットが溢れかえってしまう事態になる。それなので η の値が 0 より大きくならない最大の R（ R_c とする）が大きいほどネットワークのキャパシティは大きいと言える。図から明らかなように α=-1 の時の R_c は 46,47 で最も大きい。次いで α=-1.6 の時の値が 36,37 で2番目

に大きい。それからα=-0.5,0,0.5の順に R_c の値は小さくなっていき、α=1で最も小さくなる。FIG．2はαの値をいろいろ変えてFIG．1と同様の実験を行った時の、αに対する R_c の値を表したものである。Aが1.0から小さくなるにつれての R_c 値は上昇していき、α=-1.0で最大値をとる。その後αの減少に伴って R_c も減少していく。

このことに対して理論的な分析をする。時間tにおけるノードiのパケット処理数 $n_i(t)$ が周りから流れ込んでくるパケット数に等しいとすると、

$$0=-n_i(t)+\sum_{j=1}^{N} A_{ij}n_j(t)\frac{k_i^{\alpha}}{\sum_{l=1}^{N} A_{jl}k_l^{\alpha}} \leftarrow (1)$$

という等式が成り立つ。ここで A_{ij} は②の2．で用いた隣接行列成分である。BAネットワークでのノードの次数分布はベキ乗則にしたがっているので、大半のノードの次数は同じだとみなすことができる。するとWを定数として、

$$\sum_{l=1}^{N} A_{jl}k_l^{\alpha}=\sum_{l=1}^{N} A_{jl}W=k_jW \leftarrow (2)$$

と書き変える事ができる。(1)と(2)より、

$$n_i=\sum_{j=1}^{N} A_{ij}n_j\frac{k_i^{\alpha}}{k_jW} \leftarrow (3)$$

が導かれる。この式は簡単に解けないので、

$$n_i=Ck_i^{\theta} \leftarrow (4)$$

という仮定を設ける。そして(4)を(3)に代入すると、

$$Ck^{\theta}=\frac{Ck_i^{\alpha}}{W}\sum_{j=1}^{N} A_{ij}k_j^{\theta-1} \leftarrow (5)$$

という式が出てくる。ここでまた(2)の式を使うと、

$$Ck_i^{\theta}=\frac{Ck_i^{\alpha}}{W}k_iW=Ck_i^{1+\alpha} \leftarrow (6)$$

となって、最終的に $n(k) \sim k^{\theta}(\theta=1+\alpha) \leftarrow (7)$ が導かれる。(7)が正しいかどうか確かめるためにN=1000のBAネットワークにおけるn(k)の分布を見てみると、確かにkのベキ乗に比例していて、その傾きはθ(=1+α)となっている事が分かる。よって(7)と仮定(4)の正しさが示された。もしPIAルールが無ければ違う結果になるかも知れないという恐れがあるが、FIG．3の小さなウインドからはPIAルールがあっても無くてもθとαの関係はほとんど変わらないことが分かる。

さて(7)においてα=-1とおいてみるとθ=0となるので $k^{\theta}=1$ となる。よってn(k)はkに関係なく一定の値をとることになり、その値は(4)よりCとなる。つまりα=-1というのは各ノードに処理できるだけのパケットを送るという戦術を表していたのである。

これでようやく”迂回ルール”をα=-1 で適用したときにネットワークが最大の耐性を持つという理由が分かった。

この論文では他にαの値に応じた待ち行列の分布の様子や目的地までの平均時間などについての解説があるが②で同じような事例を取り上げているのでここでは割愛することにする。

ただ1つ面白い結果としてはFIG. 7で見られるようにR が R_c を超えるとネットワークのパケット増加率はαの値に関係なく上昇していくというものがある。そして上昇率はほぼ直線となっているのであるが、その傾きは大体 0.7 となっているのが非常に興味深い。これは 30%のパケットはネットワークが混雑状態になっていようがいまいが関係なく目的地までたどり着けるという事を意味している。これらの事が起こる原因については論文では触れられていないが、混雑状態のネットワークにおいてもある程度のパケットを処理できるというのは有用な事実である。

●用いた論文

①. Vulnerability of weighted networks
(Luca Dall'Asta, Alain Barrat , et al.;2006,DOI:10.1088/1742-5468/2006/04/P04006)

②. Congestion-gradient driven transport on complex networks
(Bogdan Danila, Yong Yu, et al. ;2006,DOI:10.1103/PhysRevE.74.046114)

③. Traffic dynamics based on local routing protocol on a scale-free network
(Wen-Xu Wang, Bing-Hong Wang, et al.;2006,DOI:10.1103/PhysRevE.73.026111)