

コンフィグレーションモデルの巨大連結成分について

情報科学研究科 システム情報科学専攻

尾畑伸明研究室 五十嵐 有美

1 複雑ネットワークのモデル

ネットワークとは数学のグラフ理論でいう「グラフ」のことであり、ここでは単純無向グラフを考える。複雑ネットワークというときは、そのノード数が巨大であることを前提とする。各ノードに接続するエッジの個数を次数と呼び、ここでは k と書く。複雑さの観点は様々あるが、特に次数分布 $\{P(k)\}$ に注目する。ランダムグラフは複雑さを実現する最も基本的な確率的ネットワークモデルを与える。本論文ではコンフィグレーションモデルを扱う。

2 コンフィグレーションモデル

Erdős-Rényi ランダムグラフの拡張版であり、本研究では次数分布をべき則で与える。

定義 1 (べき分布). 最小次数を k_{min} とする。次数分布がべき則に従うとは、

$$P(k) = ck^{-\gamma}, \quad k \geq k_{min} \quad (1)$$

が成り立つことである。ここで、 c は規格化定数である。

本研究では連結成分のでき方を研究対象とするので、 $k_{min} = 1$ とする。最大次数 k_{max} を ∞ と見なせるほどノード数 N が大きく、その全次数範囲 ($1 \leq k < \infty$) で $P(k)$ が $k^{-\gamma}$ に比例しているとする、

$$P(k) = \frac{1}{\zeta(\gamma)} k^{-\gamma}, \quad k = 1, 2, \dots \quad (2)$$

が成り立つ。ここで、 $\zeta(x) = \sum_{n=1}^{\infty} n^{-x}$ はツェータ関数である。平均値 $\langle k \rangle$ と 2 次のモーメント $\langle k^2 \rangle$ は、

$$\langle k \rangle = \sum_{k=1}^{\infty} kP(k) = \frac{\zeta(\gamma-1)}{\zeta(\gamma)} \quad (3)$$

$$\langle k^2 \rangle = \sum_{k=1}^{\infty} k^2 P(k) = \frac{\zeta(\gamma-2)}{\zeta(\gamma)} \quad (4)$$

で与えられる。ここで、 $\gamma > 3$ ならば、 $\langle k \rangle, \langle k^2 \rangle$ 共に有限値である。

定理 2 (Molloy-Reed). ノード数 $n \rightarrow \infty$ で次数分布 $\{P(k)\}$ を実現するある条件を満たす (詳細は修士論文参照) ランダムグラフの列 Ω_n において、

1. $\langle k^2 \rangle > 2\langle k \rangle$ のとき、巨大連結成分が存在する。さらに、 $\langle k^2 \rangle < \infty$ ならば、巨大連結成分は 1 つである。

2. $\langle k^2 \rangle < 2\langle k \rangle$ のとき、すべての連結成分が小さい。

巨大連結成分とは、ランダムグラフの列 Ω_n において、ある定数 C が存在して、サイズが $Cn + o(n)$ となる連結成分のことである。Molloy-Reed の結果より、 $\frac{\langle k^2 \rangle}{\langle k \rangle} = 2$ が臨界点を与える。Molloy-Reed の定理を準用すると、コンフィグレーションモデルについては、 $\frac{\zeta(\gamma-2)}{\zeta(\gamma-1)} = 2$ を満たすような γ が臨界点 γ_c となることが示唆される。この超越方程式の解は、 $\gamma_c \approx 3.47875$ が知られている。

コンフィグレーションモデルは Molloy-Reed の定理の前提を満たしていないので臨界点 γ_c がわかっていない。3.47 を目安にして γ_c を見出していく。

3 シミュレーション結果と臨界点における挙動

連結成分の個数や最大連結成分のサイズなどは、 γ や N に依存して変化する。それらのパラメータを変化させ、ネットワーク構造を詳しく見ていくことで臨界点 γ_c の特定を試みる。臨界点の特定のため、次に示す 4 つの観点からシミュレーションを行い考察した。シミュレーション回数は 20000 回とする。

3.1 最大連結成分のサイズ S_1 の平均値の N 依存性

シミュレーション結果より、

$$S_1 \propto N \quad (\gamma < \gamma_c) \quad (5)$$

$$S_1 \propto N^\alpha, \quad \alpha < 1 \quad (\gamma > \gamma_c) \quad (6)$$

を得た。べき指数 α は、

$$\gamma = 3.5 \text{ のとき, } \alpha \sim 0.56 \quad (7)$$

$$\gamma = 3.6 \text{ のとき, } \alpha \sim 0.46 \quad (8)$$

となることがわかった。

3.2 第 2 連結成分のサイズ S_2 の平均値の N 依存性

シミュレーション結果より、

$$S_2 \propto \log N \quad (\gamma < \gamma_c) \quad (9)$$

$$S_2 \propto N^\beta, \quad \beta < 1 \quad (\gamma > \gamma_c) \quad (10)$$

を得た。べき指数 β は、

$$\gamma = 3.5 \text{ のとき, } \beta \sim 0.56 \quad (11)$$

$$\gamma = 3.6 \text{ のとき, } \beta \sim 0.48 \quad (12)$$

となることがわかった。 $\gamma = 3.5$ のとき、 $\alpha \simeq \beta \simeq 0.56$ が成り立ち、 $\gamma = 3.6$ のとき、 $\alpha \simeq \beta \simeq 0.48$ が成り立つことが数値的に確認された。このことから、 $\gamma > \gamma_c$ の範囲について、 S_1, S_2 は共に N のべきに比例して成長するが、そのべき指数の値は等しくなることが見出された。

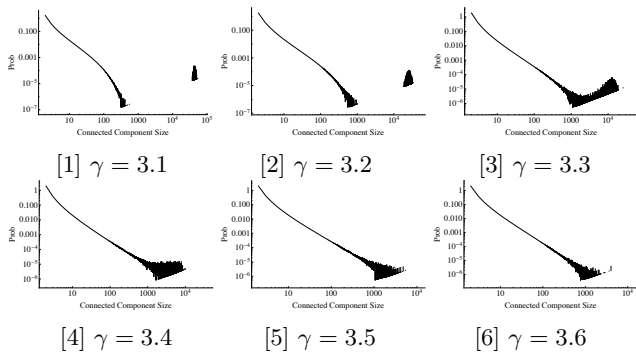
3.3 ネットワーク全体における連結成分のサイズ分布

定義 3 (サイズ分布). ランダムにノードを 1 点を選んだときに、そのノードがサイズ s 以上の連結成分に属する確率を $F(s)$ とおく。このとき、

$$F(s) = \int_s^\infty w(t) dt \quad (13)$$

を満たす (連続近似) ような $w(t)$ をサイズ分布の密度関数という。

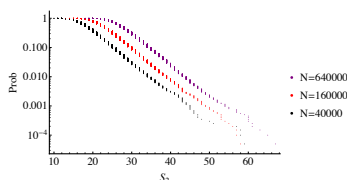
全ノード数 N を固定して γ を変化させていったときにその連結成分のサイズ分布がどのように変化するかを見た。 $N = 320000$ の結果を以下に示す。



[1], [2] では、山が 2 つ見える。つまり巨大連結成分が 1 つ存在していることがいえる。 [4], [5] では、最大連結成分の揺らぎが大きく、グラフは単調減少している。これより、ネットワークの形状が臨界点 γ_c 前後で異なることがわかった。

3.4 第 2 連結成分のサイズ S_2 の平均値の分布

20000 回のシミュレーションの S_2 の揺らぎから、 γ_c の特定を試みる。横軸に S_2 、縦軸にその累積出現頻度を対数でとり、 $N = 40000$ (黒)、 $N = 160000$ (赤)、 $N = 640000$ (紫) の場合の結果を同時にプロットした。

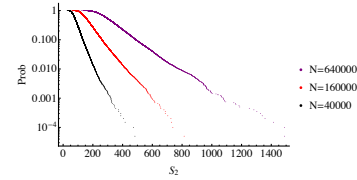


[7] $\gamma = 2.6$ のシミュレーション結果

[7] より、 $\gamma < \gamma_c$ では、密度関数は指数減衰する。密度関数を $S_2(x) = Ce^{-\mu x}$ (C は任意の定数) とかくとき、指数の値 μ はサイズ N に依らずに一定である。 $\gamma = 2.6$ の場合、

$$\mu \sim 0.25 \quad (14)$$

という結果を得た。 $\gamma > \gamma_c$ についても S_2 分布の N 依存性を観察した。



[8] $\gamma = 3.6$ のシミュレーション結果

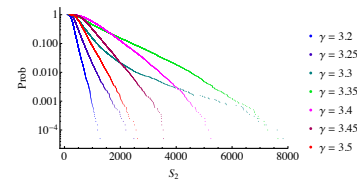
[8] より、 $\gamma > \gamma_c$ では、密度関数は指数減衰し、指数の値 μ はサイズ N によって異なる。 $\gamma = 3.6$ の場合、

$$N = 40000, \quad \mu \sim 0.03 \quad (15)$$

$$N = 160000, \quad \mu \sim 0.015 \quad (16)$$

$$N = 640000, \quad \mu \sim 0.0084 \quad (17)$$

という結果を得た。臨界点付近の γ については、興味深い複雑な挙動が観察された。



[9] $N = 640000, \gamma$ 別の S_2 分布の挙動

[9] より、 S_2 分布が指数減衰しないような γ の存在が観察された。そのような γ の範囲は、 N 毎に決まるが、 $N \rightarrow \infty$ としたとき、極限として臨界点が得られると予想される。

4 まとめ

前節の考察よりコンフィグレーションモデルの臨界点は、Molloy-Reed の結果である $\gamma_c \approx 3.47875$ とは異なり、

$$3.3 < \gamma_c < 3.4 \quad (18)$$

を予想する。第 2 連結成分のサイズ S_2 に関しては、先行研究がほとんどない。他の統計量も検討した結果、やはり (18) 式を支持することが見えるが、さらに詳しい解析は今後の課題とした。(修士論文)

参考文献

- [1] 矢久保孝介：複雑ネットワークとその構造，共立出版，2013.
- [2] M. Molloy and B. Reed: The size of the giant component of a random graph with a given degree sequence, Combin. Probab. Comput. 7 (1998), 295–305.