Network Science Chapter5

坂田研 M2
友清雄太
2024/07/11

Barabashi-Albert Model

Barabashi-Albert Modelの特徴

- (A)成長
  - 各時間のステップで、毎回、新たなノードを一つ加え、すでに存在するm個のノードとの間にリンクを張る
- (B)優先的選択
  - 新たなノードは次数が大きいノードに接続しやすい

The probability $\Pi(k)$ that a link of the new node connects to node i depends on the degree k_i as

$$\Pi(k_i) = \frac{k_i}{\sum_j k_j} \qquad (5.1)$$

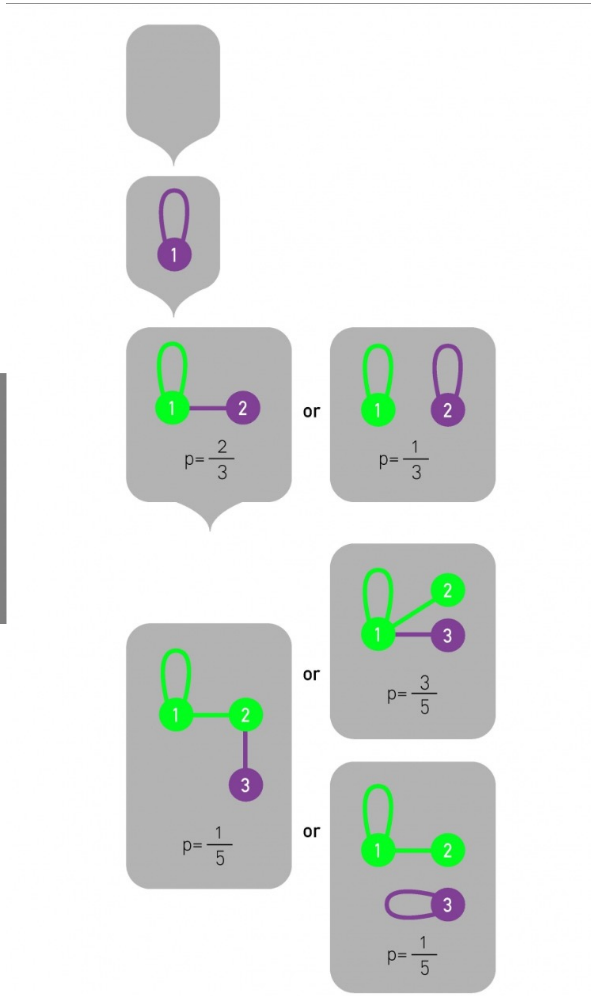
Ki: ノードiの次数

↑新しいノードのリンクがノードiに接続する確率

→Rich get richerを再現している

Barabashi-Albert Modelでのノード接続の具体例

$$p = \begin{cases} \frac{k_i}{2t-1}, & \text{if } 1 \leq i \leq t-1 \\ \frac{1}{2t-1}, & \text{if } i = t \end{cases}$$



$G_1^{(0)}$: We start with an empty network.

$G_1^{(1)}$: The first node can only link to itself, forming a self-loop. Self-loops are allowed, and so are multi-links for $m>1$.

$G_1^{(2)}$: Node 2 can either connect to node 1 with probability 2/3, or to itself with probability 1/3. According to (5.2), half of the links that the new node 2 brings along is already counted as present. Consequently node 1 has degree $k_1=2$ at node 2 has degree $k_2=1$, the normalization constant being 3.

$G_1^{(3)}$: Let us assume that the first of the two $G_1^{(t)}$ network possibilities have materialized. When node 3 comes along, it again has three choices: It can connect to node 2 with probability 1/5, to node 1 with probability 3/5 and to itself with probability 1/5.

次数のダイナミクス

$$\frac{dk_i}{dt} = m\Pi(k_i) = m\frac{k_i}{\sum_{j=1}^{N-1} k_j} \qquad (5.3)$$

m: 新たに追加されるノードが張るリンクの数→2など
←ノードiはm回リンク先として選ばれる可能性がある

The coefficient *m* describes that each new node arrives with *m* links. Hence, node *i* has *m* chances to be chosen. The sum in the denominator of (5.3) goes over all nodes in the network except the newly added node, thus

$$\sum_{j=1}^{N-1} k_j = 2mt - m \qquad (5.4)$$

←あるタイミングでのネットワーク中の既存ノードの次数の和

]

Therefore (5.4) becomes

$$\frac{dk_i}{dt} = \frac{k_i}{2t-1} \qquad (5.5)$$

(5,3)と(5,4)より

For large t the (−1) term can be neglected in the denominator, obtaining

$$\frac{dk_i}{k_i} = \frac{1}{2}\frac{dt}{t} \qquad (5.6)$$

(5,3)の-1は無視できるので

By integrating (5.6) and using the fact that $k_i(t_i)=m$, meaning that node *i* joins the network at time t_i with *m* links, we obtain

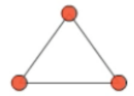
(5,6)を積分してあげる

$$k_i(t) = m\left(\frac{t}{t_i}\right)^{\beta} \qquad (5.7)$$

K_i(t)なのでt_i 以降のことを考えればいいので、このような形に

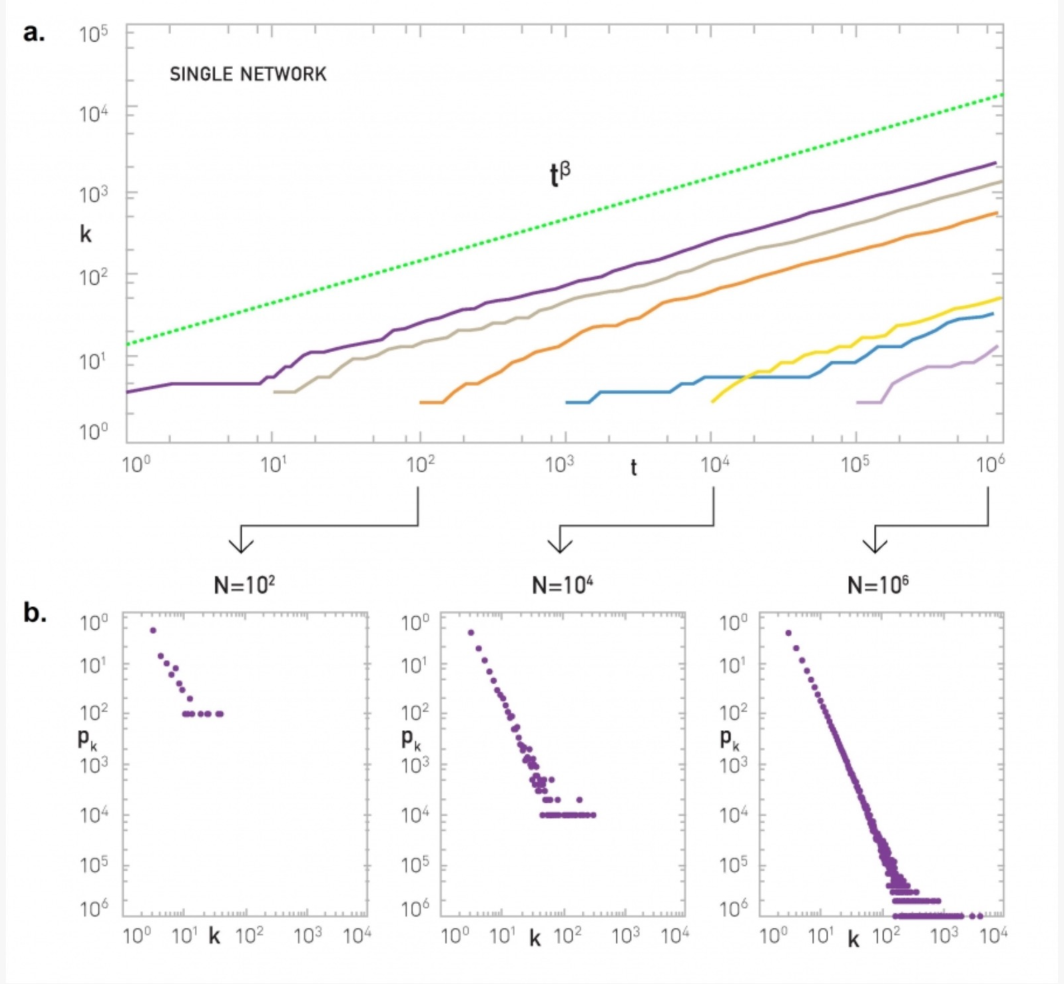
We call β the *dynamical exponent* and has the value

$$\beta = \frac{1}{2}$$



↑具体的な動画

次数ダイナミクスの可視化



さっきの次数ダイナミクスに従うと、このように次数が増えてくる
→スケールフリー性が徐々に出現する

- 次数の成長は劣線型（β<1）。新しいノードは、前の追加ノードよりも、多くのノードをリンク先候補とする。なので、既存のノードは時間と共に増えていく自分以外のノードとの間でリンクを奪い合う

- ノードiが加えられる時間が早いほど、その次数K_i(t)は大きくなる。ハブの次数が大きいのは、早くから存在していたから。マーケティングやビジネスでの先行者利益現象が起きている

$$\frac{dk_i(t)}{dt} = \frac{m}{2}\frac{1}{\sqrt{t_i t}} \qquad (5.8)$$

←ノードiが単位時間あたりに獲得するリンク数 (5,7)を微分

→ノードが獲得できるリンク数は時間と共に減少

次数分布

連続体理論

Box 5.3

Continuum Theory

To calculate the degree distribution of the Barabási-Albert model in the continuum approximation we first calculate the number of nodes with degree smaller than k, i.e. $k_i(t) < k$. Usin

$$k_i(t) = m\left(\frac{t}{t_i}\right)^{\beta} \qquad (5.7)$$

$$t_i < t\left(\frac{m}{k}\right)^{1/\beta} \qquad (5.12)$$

kよりも小さい次数を持つノード数k_i(t)<kより

In the model we add a node at equal time step (BOX 5.2). Therefore the number of nodes with degree smaller than k is

$$t\left(\frac{m}{k}\right)^{1/\beta} \qquad (5.13)$$

Altogether there are $N=m_0+t$ nodes, which becomes $N \approx t$ in the large t limit. Therefore the probability that a randomly chosen node has degree k or smaller, which is the cumulative degree distribution, follows

$$P(k) = 1 - \left(\frac{m}{k}\right)^{1/\beta} \qquad (5.14)$$

←無作為に選ばれたノードがkまたはそれより小さい次数を持つ確率

By taking the derivative of (5.14) we obtai

$$p_k = \frac{\partial P(k)}{\partial k} = \frac{1}{\beta}\frac{m^{1/\beta}}{k^{1/\beta+1}} = 2m^2k^{-3} \qquad (5.15)$$

←(5,14)を微分すると次数分布を得る

which is (5.9).

次数分布はγ=3の冪乗則に従う

$$p(k) \approx 2m^{1/\beta}k^{-\gamma} \qquad (5.9)$$

with

$$\gamma = \frac{1}{\beta} + 1 = 3 \qquad (5.10)$$

より正確には↓

$$p_k = \frac{2m(m+1)}{k(k+1)(k+2)} \qquad (5.11)$$

→Pkは定常的なスケールフリー状態の出現を予測する(時間に依存しない)
→歴史、大きさ、年齢が異なるネットワークにおいて類似の次数分布が見られることを説明

成長性と優先的選択、どちらか一つでもよくない？という指摘に対しての反論

モデルA: 成長性のみ
モデルB: 優先的選択のみ

モデルA: 成長性のみ

- **Preferential Attachment**

The probability that a new node links to a node with degree k_i is

$$\Pi(k_i) = \frac{1}{(m_0+t-1)} \quad (5.16)$$

$$k_i(t) = m \ln\left(e\frac{m_0+t-1}{m_0+t_i-1}\right) \quad (5.17)$$

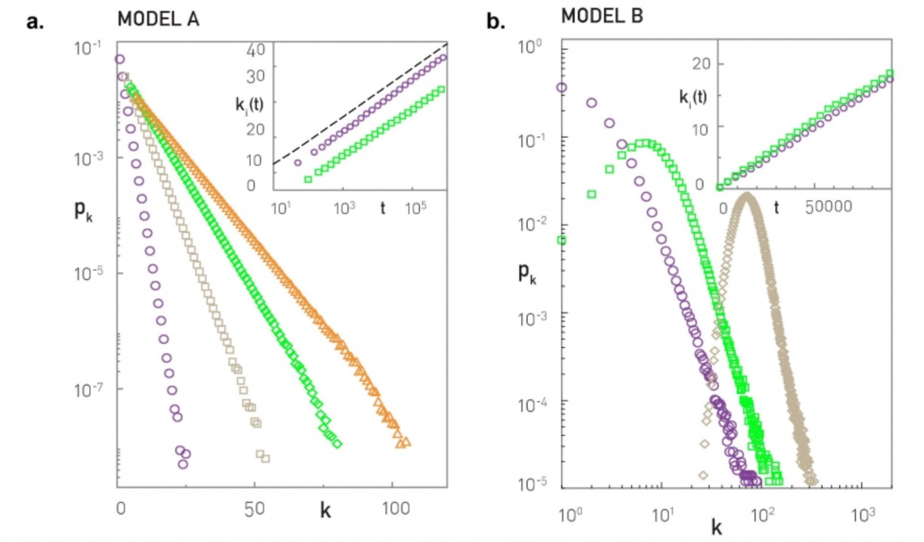
$$p(k) = \frac{e}{m}\exp\left(-\frac{k}{m}\right) \quad (5.18)$$

優先的選択がないと、定常的ではあるが、次数分布が指数関数であるネットワークが誕生

モデルB: 優先的選択のみ

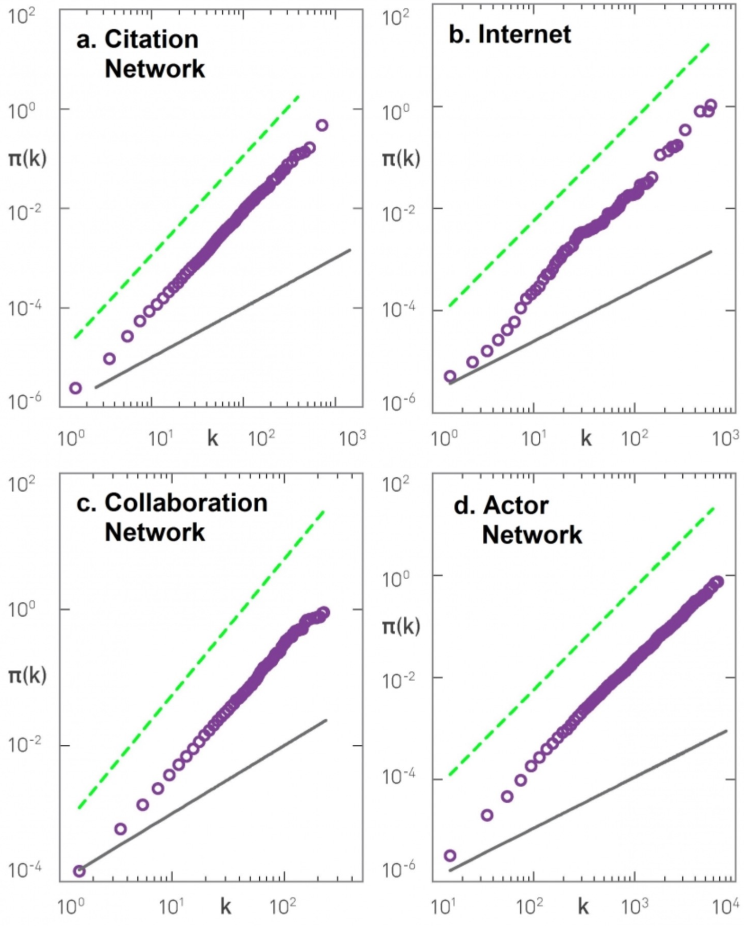
tがN(N-1)/2に近づくにつれ、ネットワークは完全グラフになる

成長が存在しないと、定常性は失われ、ネットワークが完全グラフに収束



優先的選択の度合いについて

実際のネットワークを見ると、~K から ~k^2の間に収まっている



累積優先的選択関数

$$\pi(k) = \sum_{k_i=0}^{k} \Pi(k_i) \qquad (5.21)$$

$$\Pi(k) \sim k^{\alpha} \qquad (5.22)$$

優先的選択の度合いによってネットワークの構造は変わってくる

Sublinear Preferential Attachment (0 < α < 1)

For any $\alpha > 0$ new nodes favor the more connected nodes over the less connected nodes. Yet, for $\alpha < 1$ the bias is weak, not sufficient to generate a scale-free degree distribution. Instead, in this regime the degrees follow the stretched exponential distribution (SECTION 4.10)

$$p_k \sim k^{-\alpha} \exp\left(\frac{-2\mu(\alpha)}{\langle k\rangle(1-\alpha)} k^{1-\alpha}\right) \qquad (5.23s)$$

where $\mu(\alpha)$ depends only weakly on α. The exponential cutoff in (5.23) implies that sublinear preferential attachment limits the size and the number of the hubs.

Sublinear preferential attachment also alters the size of the largest degree, k_{max}. For a scale-free network k_{max} scales polynomially with time, following (4.18). For sublinear preferential attachment we have

$$k_{max} \sim (\ln t)^{1/(1-\alpha)} \qquad (5.24)$$

a logarithmic dependence that predicts a much slower growth of the maximum degee than the polynomial. This slower growth is the reason why the hubs are smaller for $\alpha < 1$ (Image 5.11).

Superlinear Preferential Attachment (α > 1)

For $\alpha > 1$ the tendency to link to highly connected nodes is enhanced, accelerating the *rich-gets-richer process*. The consequence of this is most obvious for $\alpha > 2$, when the model predicts a *winner-takes-all* phenomenon: almost all nodes connect to a few super-hubs. Hence we observe the emergence of a hub-and-spoke network, in which most nodes link directly to a few central nodes. The situation for $1 < \alpha < 2$ is less extreme, but similar.

This winner-takes-all process alters the size of the largest hub as well, finding that (Image 5.11).

$$k_{max} \sim t \qquad (5.25)$$

In summary, nonlinear preferential attachment changes the degree distribution, either limiting the size of the hubs ($\alpha < 1$), or leading to super- hubs ($\alpha > 1$, Image 5.12). Consequently, $\Pi(k)$ needs to depend strictly linearly on the degrees for the resulting network to have a pure power law p_k. While in many systems we do observe such a linear dependence, in others, like the scientific collaboration network and the actor network, preferential attachment is sublinear. This nonlinear $\Pi(k)$ is one reason the degree distribution of real networks deviates from a pure power-law. Hence for systems with sublinear $\Pi(k)$ the stretched exponential (5.23) should offer a better fit to the degree distribution.