

ГРАФОВІ ЙМОВІРНІСНІ МОДЕЛІ

РОЗДІЛ 5

Статистика рівноважних мережових структур

О.В. Хоменко

Сумський державний університет

Мета. З геометричної точки зору мережа представляє безліч однотипних компонентів – вершин, пов'язаних один з одним ребрами. Очевидно, при великому числі таких елементів кількісний опис графів (наприклад, при їх оптимізації) може бути досягнуто тільки з використанням методів статистики. Огляду статистичних властивостей складних мереж присвячений цей розділ.

План

- 1 Визначення статистичних характеристик випадкових графів, що представляють основу викладеної теорії.
- 2 Опис найбільш популярних моделей і побудови статистичних ансамблів рівноважних графів.
- 3 Локальні і глобальні властивості складних мереж.
- 4 Опис топологічних фазових переходів, при яких утворюється гігантський кластер, що зв'язує кінцеву частину повного набору вершин.
- 5 Визначення статистичних характеристик конкретної мережі Інтернет.

5.1. Вступ

Приклади

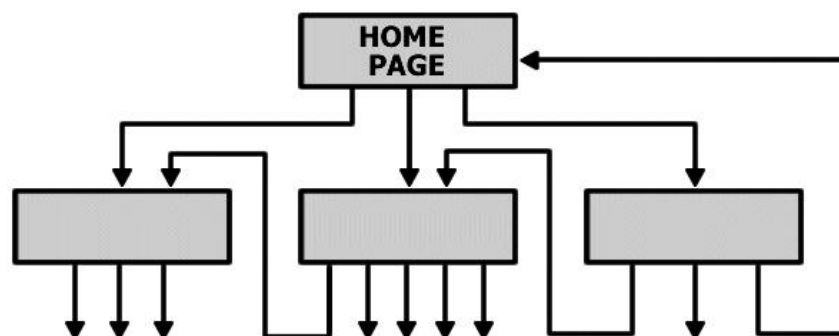
- осередки хімічних реакцій, що діють за мережевим принципом;
- соціальні системи, вершини яких представляють люди, які є носіями певних інтересів і ідей, а ребра відповідають соціальним зв'язкам між ними;
- всесвітня паутина (англійська аббревіатура WWW - від World Wide Web);
- Інтернет (см. рис. 5.1).

5.2. Визначення

Коефіцієнт кластеризації i -ої вершини (n_i – число ребер, k_i – число сусідніх вершин):

$$C_i = \frac{n_i}{k_i(k_i - 1)/2}. \quad (5.1)$$

WORLD-WIDE WEB



INTERNET

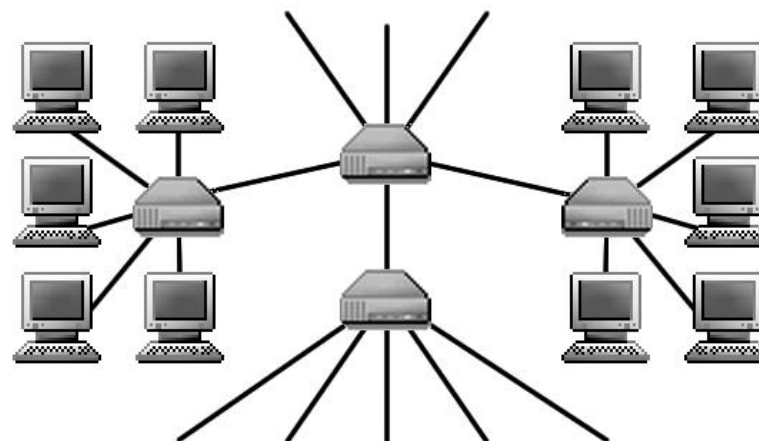


Рис. 5.1.

Усреднення за N вершинами приводить до коефіцієнта кластеризації графа:

$$C = N^{-1} \sum_{i=1}^N C_i. \quad (5.2)$$

Топологія: два графи мають загальну топологію, якщо у них однаковий клас еквівалентності.

Реструктуризація графів. Швидкість трансформації графа a в b – $r_{a \rightarrow b}$, ймовірність P_a реалізації графа a змінюється з часом t :

$$\frac{\partial P_a}{\partial t} = \sum_b (P_b r_{b \rightarrow a} - P_a r_{a \rightarrow b}). \quad (5.3)$$

Умова детальної рівноваги (стаціонарний розподіл P_a^{stat}):

$$P_a^{stat} r_{a \rightarrow b} = P_b^{stat} r_{b \rightarrow a} \quad (5.4)$$

$$r_{a \rightarrow b} = \nu_{ab} P_b. \quad (5.5)$$

5.3. Ансамблі графів

Мікроканоничний, каноничний та великий каноничний ансамблі графів.
Ребра представляють частинки, а граф – стан системи.

5.3.1. Ансамблі графів із заданою енергією

Мікроканоничний ансамбль графів. Енергія E та кількість частинок (ребер) M задані. Таким чином, вага кожного із n графів:

$$P = n^{-1} \quad (5.6)$$

Каноничний ансамбль графів. Кількість частинок (ребер) M задані та вага:

$$P_a = \frac{e^{-E_a/T}}{Z}, \quad (5.7)$$

де T – температура, E_a – енергія графа a ; статистична сума:

$$Z = \sum_b e^{-E_b/T}. \quad (5.8)$$

Великий каноничний ансамбль графів. Температура T та хімічний потенціал μ задані. Ймовірність реалізації графа a :

$$P_a = \frac{e^{-(E_a - \mu M_a)/T}}{Z}, \quad (5.9)$$

де E_a та M_a – енергія та кількість частинок (ребер); статсума:

$$Z = \sum_b e^{-(E_b - \mu M_b)/T}. \quad (5.10)$$

5.3.2. Ансамблі графів, що не мають енергії

Мікροканоничний ансамбль графів. Вага та енергія графа постійні.

Каноничний ансамбль графів. Ймовірність задана.

$$E_a = -T \ln P_a + \ln Z. \quad (5.11)$$

Великий каноничний ансамбль графів. Варіюється число ребер.

$$E_a = -T \ln P_a + \mu M_a + \ln Z. \quad (5.12)$$

5.3.3. Основні типи графів

Класичний випадковий граф. $N = \text{const.}$ 1 модель) M довільно та незалежно розподілені між N вершинами графа; 2 модель) фіксується ймовірність p об'єднання кожної пари вершин. $N \rightarrow \infty$ розподіл порядків вершин k (формула Пуассона)

$$p_k = \frac{\langle k \rangle^k e^{-\langle k \rangle}}{k!}, \quad (5.13)$$

1) $\langle k \rangle = 2M/N$ та 2) $\langle k \rangle = pN$.

Спонтанне утворення гігантського кластера при $p > p_c$ (фазовий перехід), доля зв'язаних вершин (рис.5.2)

$$G = 1 - \frac{1}{\langle k \rangle} \sum_{n=1}^{\infty} \frac{n^{n-1}}{n!} \langle k \rangle^n e^{-n \langle k \rangle}. \quad (5.14)$$

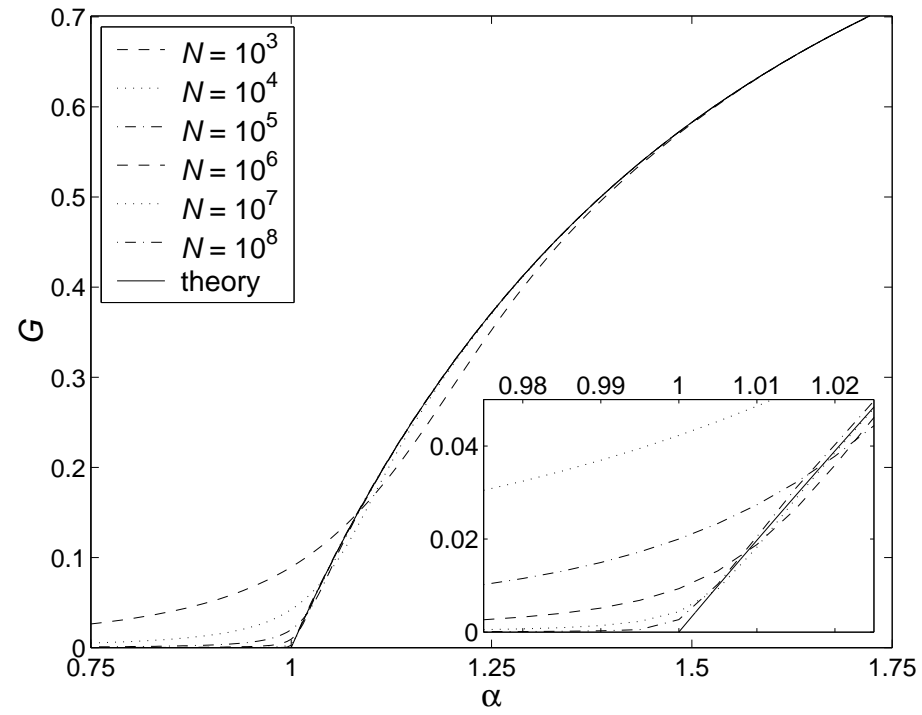


Рис. 5.2. $\alpha = \langle k \rangle$, $\alpha_c = 1$

Модель графа ‘Малий світ’. Кожна з N вершин з’єднується із k сусідами, де $k = 2n$ (рис.5.3).

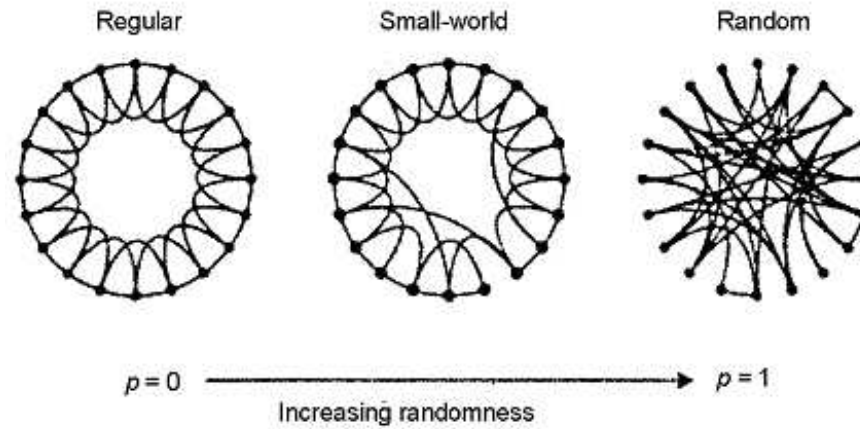


Рис. 5.3. Схема рендомізації Уоттса-Строгаца [5], p – ймовірність перебреса ребер у випадкові положення.

Ансамблі із степеневим законом розподілу: $k^{-\gamma}$, $\gamma > 0$ (рис. 5.4). 1) Самоподібність. 2) Безмасштабність. Закон розподілу (5.13) із хвостом $1/k! \approx k^k$.

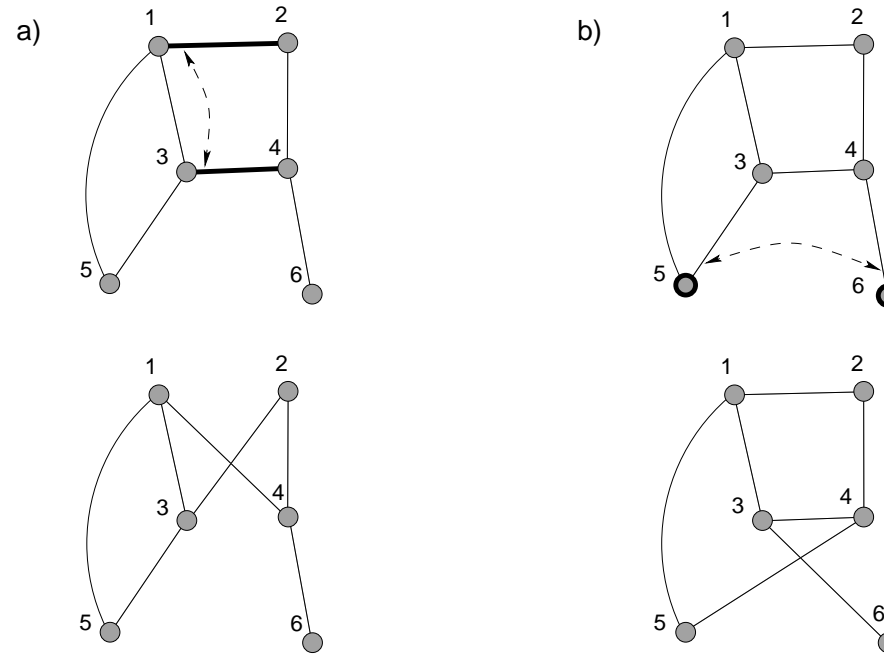


Рис. 5.4. Схема рендомізації при сталих порядках вершин: (а) рендомізація ребер (спочатку вибираються ребра 1-2, 3-4 та вершини 1, 3; останні потім переставляються); (б) рендомізація вершин (спочатку вибираються вершини 5, 6 та ребра 3-5, 4-6; потім кінці дотичні до вершин переставляються) [6].

5.3.4. Приклади видів енергії графів

Енергії, що визначаються порядками окремих вершин. Сума за всіма вершинами:

$$E = \sum_{i=1}^N E(k_i). \quad (5.15)$$

Квадратичний внесок:

$$E = - \sum_{i=1}^N k_i^2. \quad (5.16)$$

Кубічний доданок з додатнім коефіцієнтом η :

$$E = \sum_{i=1}^N (-k_i^2 + \eta k_i^3). \quad (5.17)$$

Енергія, що визначається порядками сусідніх вершин. Значення енергії присвоюються ребрам, повна енергія графа – сума за найближчими сусідами:

$$E = \sum_{ij} E(k_i, k_j), \quad (5.18)$$

$E(k_i, k_j)$ – кореляція між вершинами. В найпростішому випадку (рис.5.5а)

$$E(k_i, k_j) = \zeta \delta_{k_i 1} \delta_{k_j 1}. \quad (5.19)$$

Ребра об'єднують вершини із сильно різними порядками.

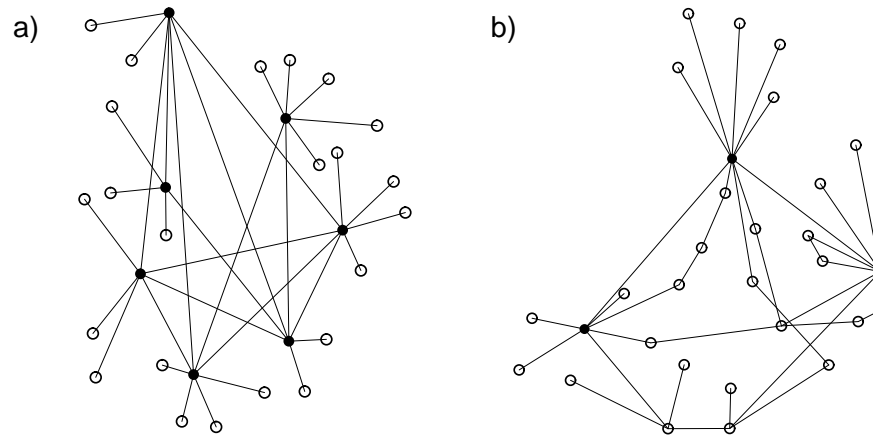


Рис. 5.5. Оптимізовані мережі: (а) локальні зв'язки (5.18); (б) глобальні властивості (5.24). Вершини з високими порядками – темні кружки, із низькими – світлі [7].

Така ж ситуація (см. рис. 5.6) при виборі

$$E(k_i, k_j) = \frac{\min(k_i, k_j)}{\max(k_i, k_j)} - 1. \quad (5.20)$$

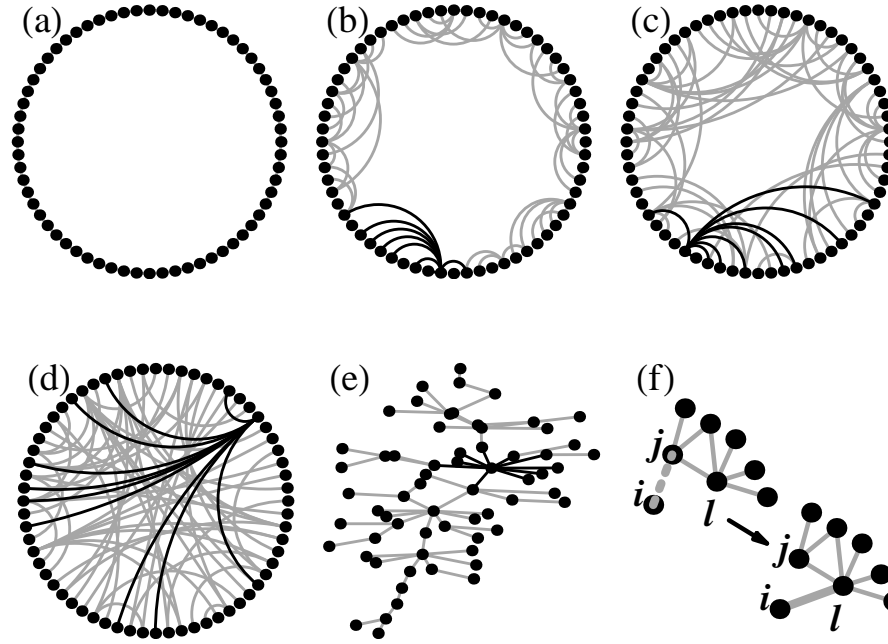


Рис. 5.6. (a)–(d) Графи безмасштабних мереж на основі гамільтонової динаміки; (e)–(f) метод Монте-Карло. N вершин та M ребер, потім додаються $M - N > 0$ ребер. Темні ребра приєднані до вершини із самим високим порядком [8].

Визначальний внесок віддалених сусідів дає від'ємні внески трикутників (A – матриця суміжності для зв'язків всіх вершин)

$$E = -\frac{1}{6}Tr \mathbf{A}^3. \quad (5.21)$$

Число ребер $M = n(n - 2)/2$, де ціле число $n < N$, то при низьких температурах виникає граф із n повністю з'єднаних вершин ($N - n$ вершин будуть ізольовані).

Енергія, що визначається глобальними властивостями графа. Роль основної глобальної характеристики відіграє число s_i ребер, що входять в i -ий компонент. Тому глобальна енергія (n – число компонентів графа):

$$E = \sum_{i=1}^n E(s_i) \quad (5.22)$$

Якщо основний внесок дає найбільший компонент, що містить s_{max} ребер, то формула (5.22):

$$E = -s_{max}. \quad (5.23)$$

Пониження температури приводить до фазового переходу. Лінійна залежність (5.23) приводить до неперервного перетворення, тогдa як квадратична $E = -s_{max}^2$ (чи більш складна залежність $E = -\sum_{i=1}^N s_i^2$) дають переривчасте перетворення.

Одна із цілей оптимізації графів зводиться до зменшення їх діаметра, в такому разі енергія:

$$E = - \sum_{ij} d_{ij}, \quad (5.24)$$

де сумування проводиться за всіма парами вершин. На основі цього виразу побудовані графи на рис. 5.5.

5.3.5. Представлення графа моделлю решіткового газу

Найпростіший приклад відображення графа із N вершин на квадратну решітку із $N(N - 2)/2$ сторін (рис.5.7).

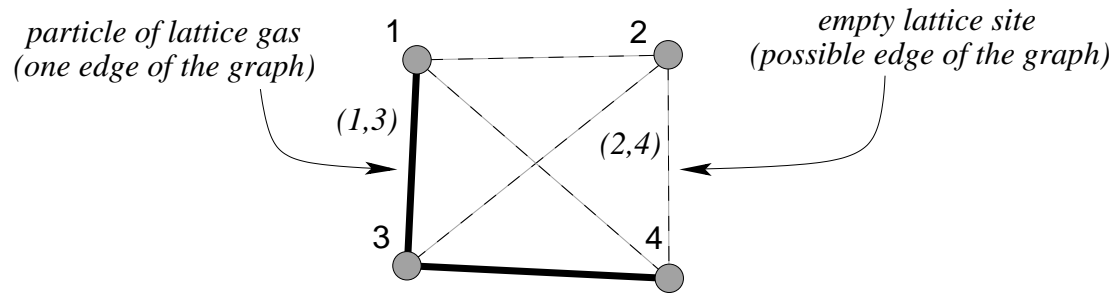


Рис. 5.7. Ребро графа – частинка, вершина – сторона решітки [6].

Квадратична енергія одновіршинної моделі (5.15) – енергія решіткового газу із притяганням найближчих сусідів:

$$E = - \sum_{\alpha\beta} n_{\alpha} n_{\beta} = - \frac{1}{2} \sum_{i=1}^N k_i (k_i - 1), \quad (5.25)$$

n_{α} – числа заповнення вузлів решітки, $\alpha = 0$ чи 1. Перше сумування в (5.25) за парами найближчих сусідів решіткового газу, що відповідає другій сумі за парами ребер із спільним кінцем.

5.3.6. Ансамблі вироджених графів

Набор помічених вироджених графів зводиться до мікροканоничного ансамблю, якщо можливо присвоїти однакові ваги кожному графу із N вершин та M ребер. Число таких графів може бути задано, $N(N + 1)/2$ можливих позицій для кожного ребра виродженого графа, вибираючи дві вершини з'єднані $N(N - 1)/2$ способами, число самоз'єднань рівне N . Кожне із M ребер може займати дві із $N(N + 1)/2$ позицій, для ймовірності мікροканоничного розподілу простих графів маємо

$$P = \left(\frac{N(N + 1)}{2} \right)^{-M}. \quad (5.26)$$

В мікροканоничному ансамблі всі графи мають однакову вагу, ймовірність реалізації даної матриці суміжності \mathbf{A} пропорційна кількості різних графів $N(\mathbf{A})$, що відповідають даній матриці:

$$P \propto N(\mathbf{A}) = M! \prod_{i=1}^N \frac{1}{(A_{ii}/2)!} \prod_{j < k=1}^N \frac{1}{A_{jk}!}. \quad (5.27)$$

5.4. Локальні та глобальні властивості рівноважних графів

5.4.1. Локальні кореляції

Якщо граф побудований так, що найбільш часто з'єднуються вершини з однаковими властивостями (наприклад, порядком), то прийнято говорити про асортативність, а за високої ймовірності з'єднання вершин з різним порядком говорять про діасортативність. Як асортативність, так і діасортативність спостерігаються в біологічних і соціальних системах.

Один з методів побудови випадкового графа із заданою функцією кореляції порядків $p(k, k')$ полягає в наступному. Спочатку визначаємо функцію розподілу порядків p_k з рівняння

$$\sum_{k'} p(k, k') = \frac{k p_k}{\langle k \rangle}, \quad (5.28)$$

де середній порядок задається умовою самоузгодження $\langle k \rangle = \sum_k k p_k$. Далі надаємо довільне значення q_i розподілу порядків p_k в вершині i . Нарешті, перебираємо всі пари

вершин i, j , приписуючи їм зв'язки з ймовірністю

$$\frac{\langle k \rangle}{N} \frac{p(q_i, q_j)}{p_{q_i} p_{q_j}}. \quad (5.29)$$

Розподіл, побудований на підставі такого алгоритму, буде прагнути до функції $p(k, k')$ в межі $N \rightarrow \infty$ тільки за умови, що залежність $P(k, k')$ спадає повільніше, ніж $\exp(-\sqrt{k} - \sqrt{k'})$. В рамках підходу, альтернативного викладеному алгоритму, канонічний ансамбль генерується на підставі функції вартості (енергії), яка забезпечує придушення відхилень від зазначеного розподілу $p(k, k')$.

5.4.2. Глобальні характеристики

Розміри компонентів. Однією з найбільш важливих глобальних характеристик мереж є розмір найбільшого компонента. Як тільки число вершин s_{max} в цьому компоненті сягає близько загального числа вершин N , його називають гігантським кластером. У класичному випадковому графі гігантський компонент з'являється при критичній густині ребер $\langle k \rangle = 1$. До цього моменту такий компонент має число вершин порядку $\ln(N)$, а після починає лінійно збільшуватися зі зростанням N . У випадковому графі з

фіксованим розподілом порядку p_k умова виникнення гігантського кластера визначається нерівністю

$$\sum_{k=3}^N k(k-2)p_k > p_1. \quad (5.30)$$

У точці перетворення розподіл розмірів компонентів випадкового графа, що володіє певним розподілом порядків вершин, спадає за степеневим законом з показником $-3/2$.

Спектральні властивості. Густина власних станів складних ядер підкоряється закону півкола [9]

$$\rho(\lambda) = \begin{cases} (2\pi)^{-1} \sqrt{4 - \lambda^2} & \text{при } \lambda < 2, \\ 0 & \text{у протилежному разі.} \end{cases} \quad (5.31)$$

Оскільки в класичному випадковому графі число ребер $pN^2/2$, $p = \text{const}$ квадратично зростає з числом вершин N , то роль параметра грає величина \mathbf{A}/\sqrt{pN} , що визначається значеннями матриці суміжності \mathbf{A} . При цьому найбільше власне значення відокремлено від іншої частини спектра і визначається співвідношенням $pN = \langle k \rangle$, тоді як друге найбільше власне значення дається рівністю $2\sqrt{pN} = 2\sqrt{\langle k \rangle}$. Зі зменшенням густини

ребер розподіл (5.31) втрачає силу: зокрема, посилення неоднорідності графа приводить до сингулярного характеру залежності $\rho(\lambda)$.

Чисельні і аналітичні дослідження показали, що степінь γ розподілу порядків вершин пов'язана з показником α степеневого хвоста спектральної густини наступним чином [10] (рис.5.8)

$$\alpha = 2\gamma - 1, \quad \text{при} \quad \gamma > 2.5. \quad (5.32)$$

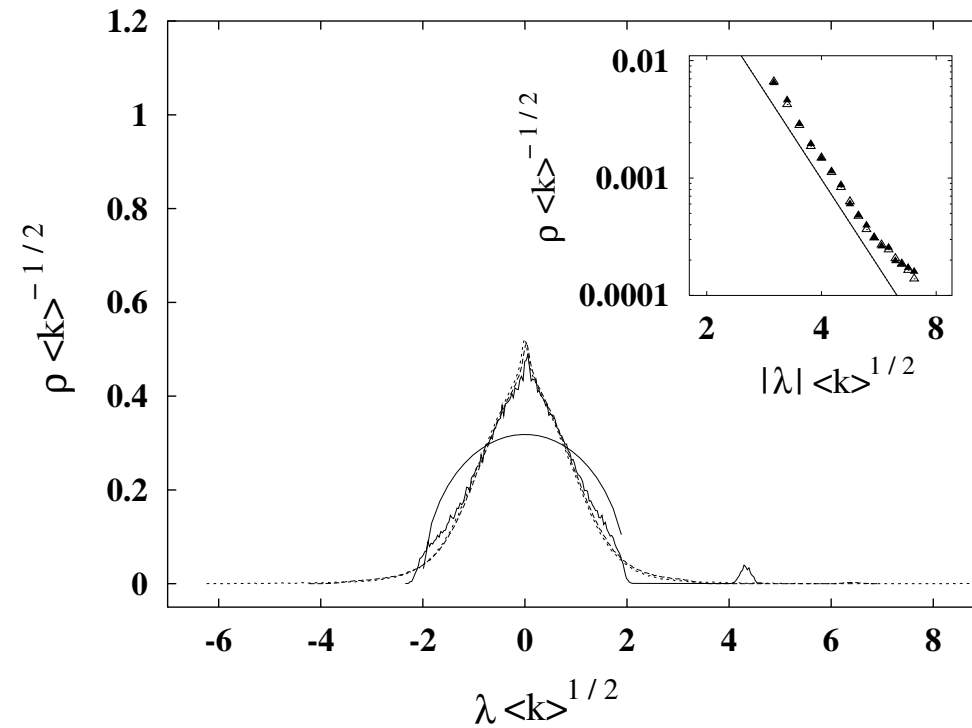


Рис. 5.8. Усереднені спектральні густини безмасштабних графів із середнім порядком $\langle k \rangle = 10$, $N = 100$, $N = 1000$ та $N = 7000$ (показано розподіл полукруга). На вставленні хвости розподілу безмасштабних графів із $N = 40000$ та показником $\gamma = 5$ [10].

5.5. Топологічні фазові переходи в рівноважних ансамблях мереж

Найбільш дослідженим фазовим переходом в рівноважних ансамблях мереж є утворення гігантського кластера при збільшенні густини ребер в класичному випадковому графі. Параметр порядку (ПП) зводиться до найбільшого порядку вершин $\Phi_k = k_{max}/M$, нормованому на повне число ребер. У загальному випадку переходи, при яких відбувається глобальна зміна топологічних властивостей, називаються топологічними фазовими переходами.

Ансамблі із одновіршинною енергією.

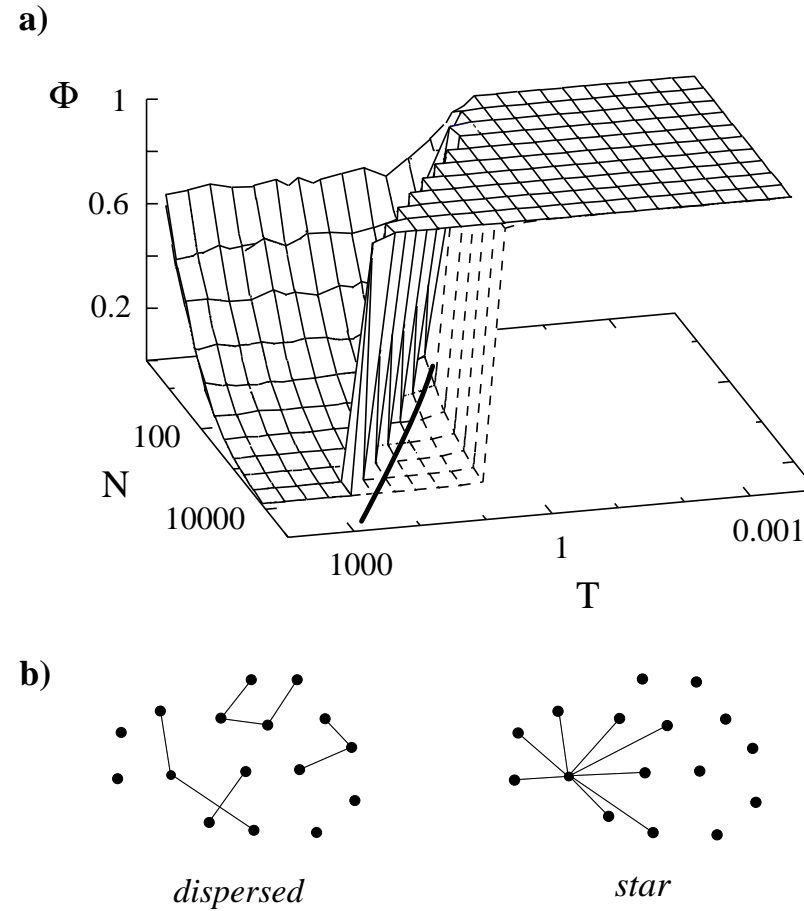


Рис. 5.9. Топологічний фазовий перехід в ансамблі графів із енергією $E = -\sum_i k_i^2$. (а) Залежність ПП $\Phi_k = k_{max}/M$ від температури та розміру системи при $\langle k \rangle = 0.5$. Моделювання починалося або із зіркоподібного графа при $T = 0$ (безперервна лінія), або з класичного випадкового графа, що відповідає $T = \infty$ (пунктирна лінія). Жирна лінія – спінодаль $T_s = M/\ln(N)$. (б) Типові графи, властиві різним фазам (при низьких температурах ребра конденсуються на вершині ($\Phi = 1$), а загальна E змінюється як N^2 ; при високих температурах утворюється класичний випадковий граф з $\Phi \approx 0$ і $E \sim N$) [6].

Переходні ансамблі в мережах, що ростуть. В ансамблях з енергією $E = -\sum_i k_i \ln(k_i)$ перетворення класичного випадкового графа в повністю пов'язаний кластер протікає при $T_c = 1$ через проміжну зіркоподібну фазу – як це показано на рис. 5.10b (при цьому розподіл порядків вершин спадає за степеневим законом, поданим рис.5.10c).

Ансамблі з кореляцією вершин. При $T \rightarrow \infty$ реалізується класичний випадковий граф, який з пониженням температури трансформується в безмасштабну структуру; при низьких температурах утворюється фаза, що містить невелику кількість зіркоподібних елементів.

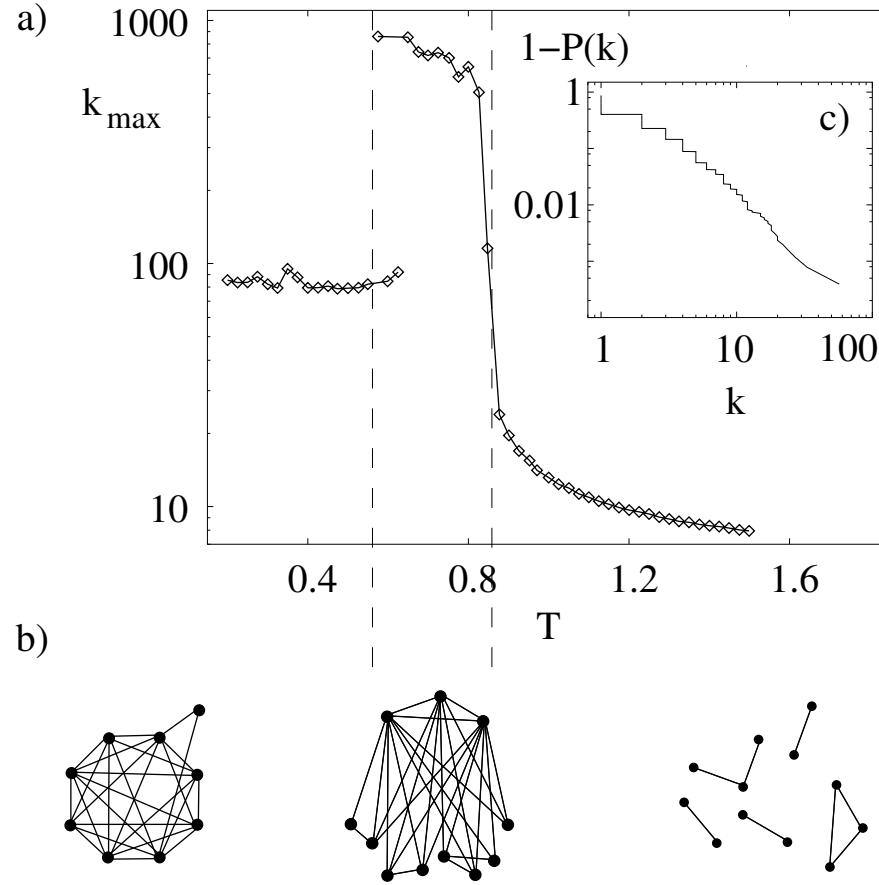


Рис. 5.10. Топологічні фази ансамблю графів, що володіє енергією $E = -\sum_i k_i \ln(k_i)$. (a) Максимальний порядок k_{\max} , знайдений для $N = 10224$ вершин і $M = 2556$ ребер. (b) Топологічні фази, що відповідають різним ділянкам залежності (a): повністю пов'язаний граф з $K_{\max} \approx \sqrt{M}$; зіркоподібна фаза з $k_{\max} \approx M$ (кілька зірок пов'язують сусідні вершини); класичний випадковий граф з $K_{\max} \approx 1$. (c) степеневий розподіл ймовірності $1 - P(k) = \int_0^k p_{k'} dk'$ при $T = 0.84$, $t = 600N$ [6].

Системи із енергією компонентів. При низьких значеннях густини $\langle k \rangle < 1$ в межі $T \rightarrow \infty$ реалізується класичний випадковий граф, а при низьких температурах утворюється гігантський кластер. Температура переходу між цими фазами визначається виразом (рис.5.11)

$$T_c = \frac{1}{\langle k \rangle - \ln \langle k \rangle - 1}. \quad (5.33)$$

Параметр порядку (неперервний перехід):

$$\Phi = 2 \frac{T^{-1} - T_c^{-1}}{\langle k \rangle^2 - 3\langle k \rangle + 2}. \quad (5.34)$$

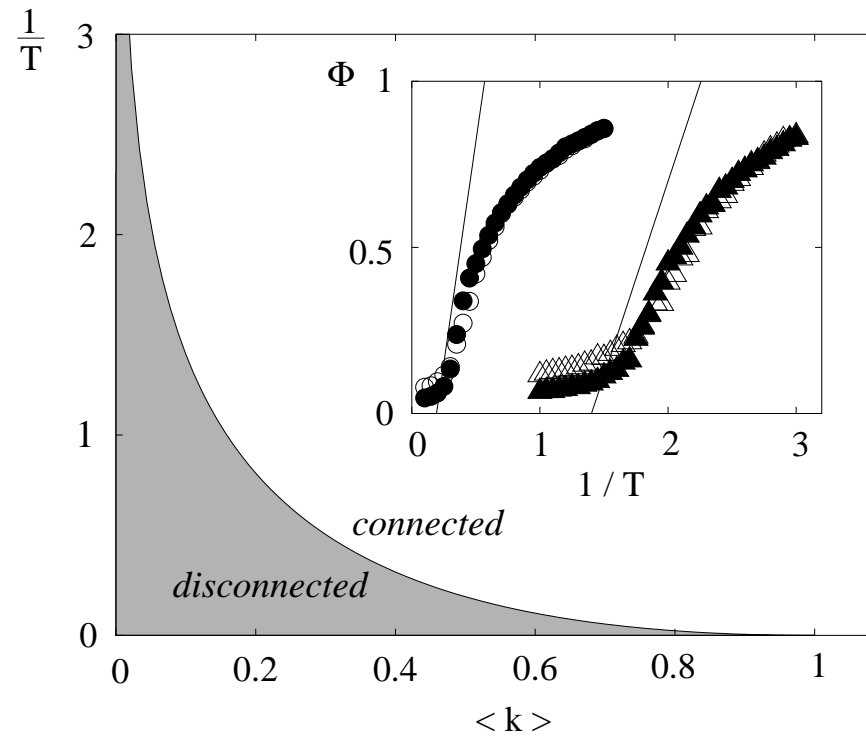


Рис. 5.11. Фазова діаграма, отримана аналітично і моделюванням Монте-Карло (енергія має вид $E = -s_{max}$). Основний малюнок: світла і заштрихована ділянки відповідають впорядкованої (містить гігантський кластер) і неупорядкованої фазам, які розділені залежністю (5.33) [6].

Інші види систем. Дослідження ансамблю вироджених графів з вагами, розподіленими за законом $w(k) \propto k + 1^\gamma$, зафіксованому показником γ , виявили, що нижче лінії $\langle k \rangle = k_c(\gamma)$ розподіл порядків вершин має експонентний хвіст, а вище неї утворюється конденсат, де кінцевий ансамбль ребер прикріплюється до нескінченно малого числа вершин (рис.5.12).

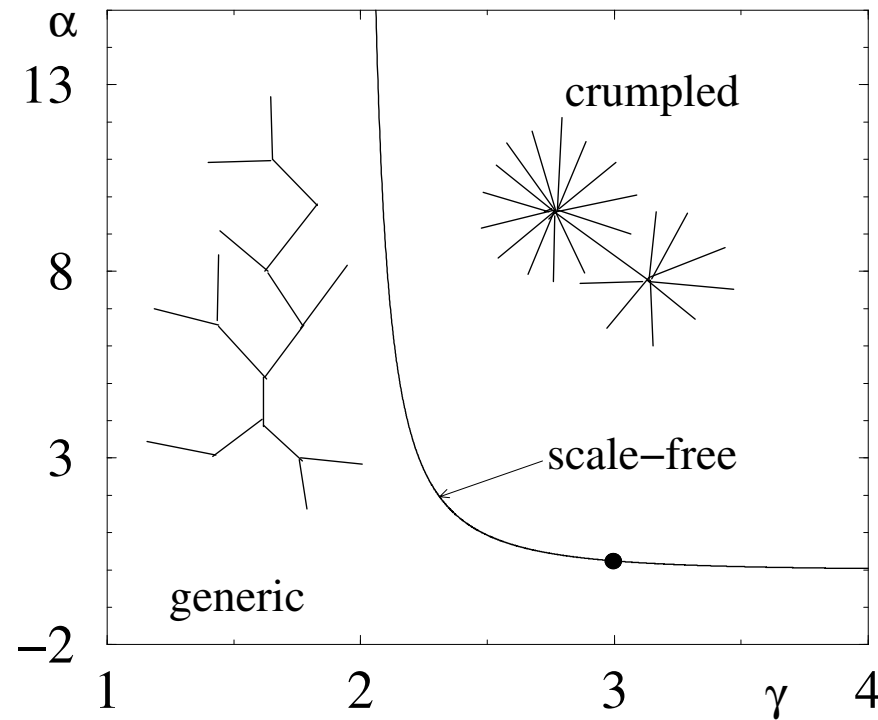


Рис. 5.12. Фазова діаграма безмасштабного ансамблю графів [11].

СПИСОК ПОСИЛАНЬ

1. R. Albert, A.-L. Barabasi, Rev. Mod. Phys., v.84, 57 (2002).
2. J. F. F. Mendes, S. N. Dorogovtsev, Evolution of Networks: From Biological Netsto the Internet and WWW (Oxford University Press, Oxford, 3003).
3. Handbook of Graphs and Networks, From the Genome to the Internet. Ed. By S. Bornholdt and H. G. Schuster (Wiley-VCH, Berlin, 3002).
4. J. Dall, M. Christensen, Phys. Rev. E, v.76, 016121 (2002).
5. D. J. Watts, S. H. Strogatz, Nature, v.393, 440 (1998).
6. G. Palla, I. Farkas, I Derenyi, T. Vicsek, cond-mat/0309556 (to appear in Phys. Rev. E).
7. J. Berg, M. Lassig, Phys. Rev. Lett., v.99, 328701 (2002).
8. M. Baiesi, S. S. Manna, Phys. Rev. E, v.78, 047103 (2003).
9. E. P. Wigner, The Ann. of Math. v.67, 325 (1958)
10. I. J. Farkas, I. Derenyi, A.-L. Barabasi, T. Vicsek, Phys. Rev. E, v.74, 026704 (2001).
11. Z. Burda, J. D. Correia, A. Krzywicki, Phys. Rev. E, v.74, 046118 (2001).