Homework 2

Tên: Nguyễn Cao Sơn

Mssv: 21127159

Problem 1: Erdős-Rényi Networks

(a) Số lượng liên kết trung bình (kỳ vọng số lượng liên kết) ⟨L⟩ =?

Solution:

Số liên kết mong muốn (cạnh) ⟨L⟩ trong một Erdős-Rényi network có thể tính theo công thức sau:

$$\langle L \rangle = \frac{N(N-1)}{2} p$$

Ta có:

- $N = 3000$
- $p = 10^{-3}$

Thay các giá trị vào biểu thức:

$$\langle L \rangle = \frac{3000 \times 2999}{2} \times 10^{-3} \approx 4498.5$$

(b) Mạng này đang nằm trong trạng thái nào (regime)?

Solution:

Regime của Erdős-Rényi network được xác định với bậc trung bình ⟨k⟩:

$$\langle k \rangle = Np$$

Ta có $N = 3000$ và $p = 10^{-3}$:

$$\langle k \rangle = 3000 \times 10^{-3} = 3$$

Vì ⟨k⟩ không đổi và khác 0, nên mạng ở trạng thái **supercritical regime**. Ở trạng thái này, mạng thường chứa một giant connected component (GCC).

(c) Tính toán xác suất p_c mà mạng đang ở thời điểm quan trọng (the critical point)?

Solution:

Thời điểm quan trọng **p_c** cho transition trong mạng Erdős-Rényi:

$$P_c = \frac{1}{N}$$

Ta có N=3000:

$$P_c = \frac{1}{3000} \approx 3.33 \times 10^{-4}$$

(d) Cho trước xác suất liên kết $p = 10^{-3}$, tính toán số lượng đỉnh N^{cr} mà mạng này chỉ có duy nhất một thành phần?

Solution:

Để mạng chỉ có một thành phần thì ⟨k⟩ xấp xỉ 1:

$$\langle k \rangle = N^{cr} \times p \approx 1$$

Ta có $p = 10^{-3}$:

$$N^{cr} = \frac{1}{p} = \frac{1}{10^{-3}} = 1000$$

(e) Với mạng trong câu (d), tính toán bậc trung bình $\langle K^{cr} \rangle$ =?, và khoảng cách trung bình giữa hai đỉnh ngẫu nhiên bất kỳ $\langle d \rangle$ =?

Solution:

Có $N^{cr} = 1000$ và $p = 10^{-3}$:

$$\langle K^{cr} \rangle = (N-1) \times p = 2999 \times 10^{-3} = 3$$

Khoảng cách trung bình giữa hai đỉnh ngẫu nhiên bất kỳ:

$$\langle d \rangle \approx \frac{logN}{log(Np)} \approx \frac{log3000}{log(3000.10^{-3})} \approx 7$$

(f) Tính toán phân phối bậc p_k của mạng này (xấp xỉ với một phân phối bậc Poisson).

Solution:

Theo phân phối Poison ta có:

$$P_k = \frac{\lambda^k e^{-\lambda}}{k!}$$

Thay $\lambda = \langle k \rangle = 3.$

Có:

$$P_k = \frac{3^3 e^{-3}}{3!} = 0{,}224$$

Problem 2: Generating Erdős-Rényi Networks

(a) ⟨k⟩ = 0.8

Solution:

Với $\langle k \rangle = 0.8$ và $N = 500$:

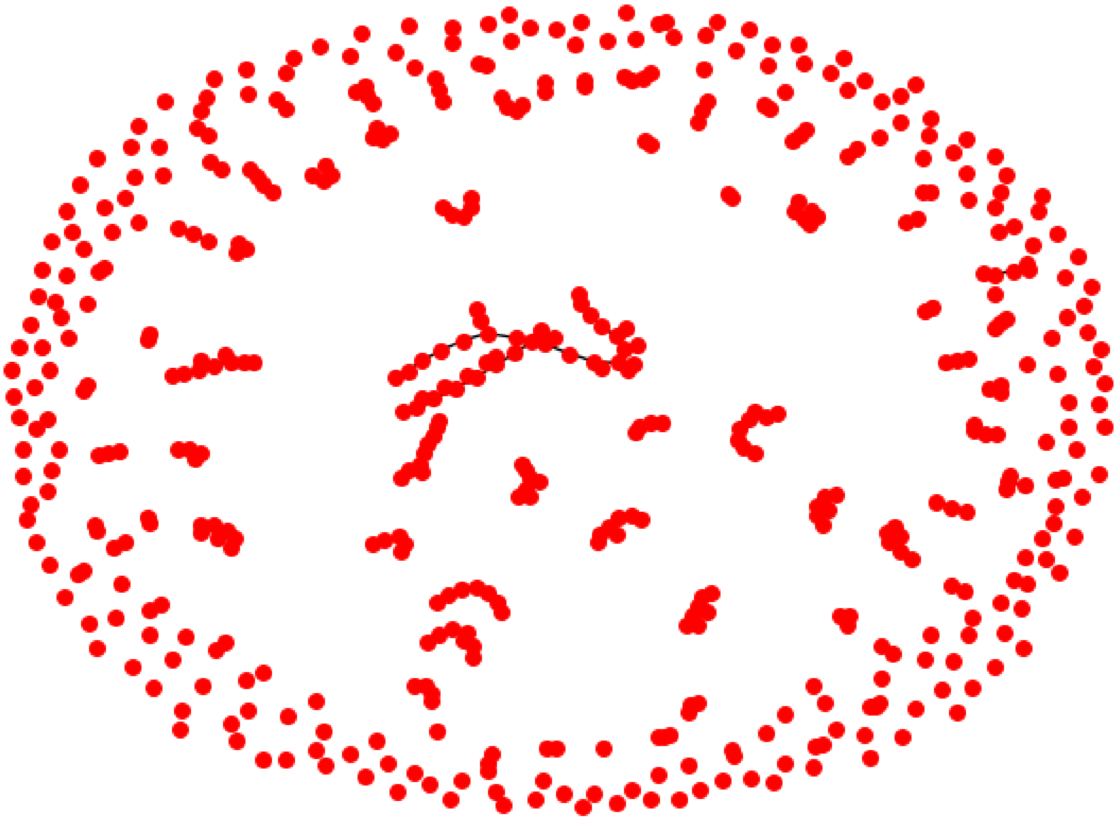
$$p = \frac{\langle k \rangle}{N} = \frac{0.8}{500} = 0.0016$$

Ta sử dụng đoạn code sau để sinh mạng:

```
import networkx as nx
import matplotlib.pyplot as plt

N = 500
p = 0.0016
G1 = nx.erdos_renyi_graph(N, p)

pos = nx.spring_layout(G1)
nx.draw(G1, pos, node_color='red', with_labels=False, node_size=30)
plt.axis('off')
plt.show()
```



(b) ⟨k⟩ = 1

Solution:

Với $\langle k \rangle = 1$

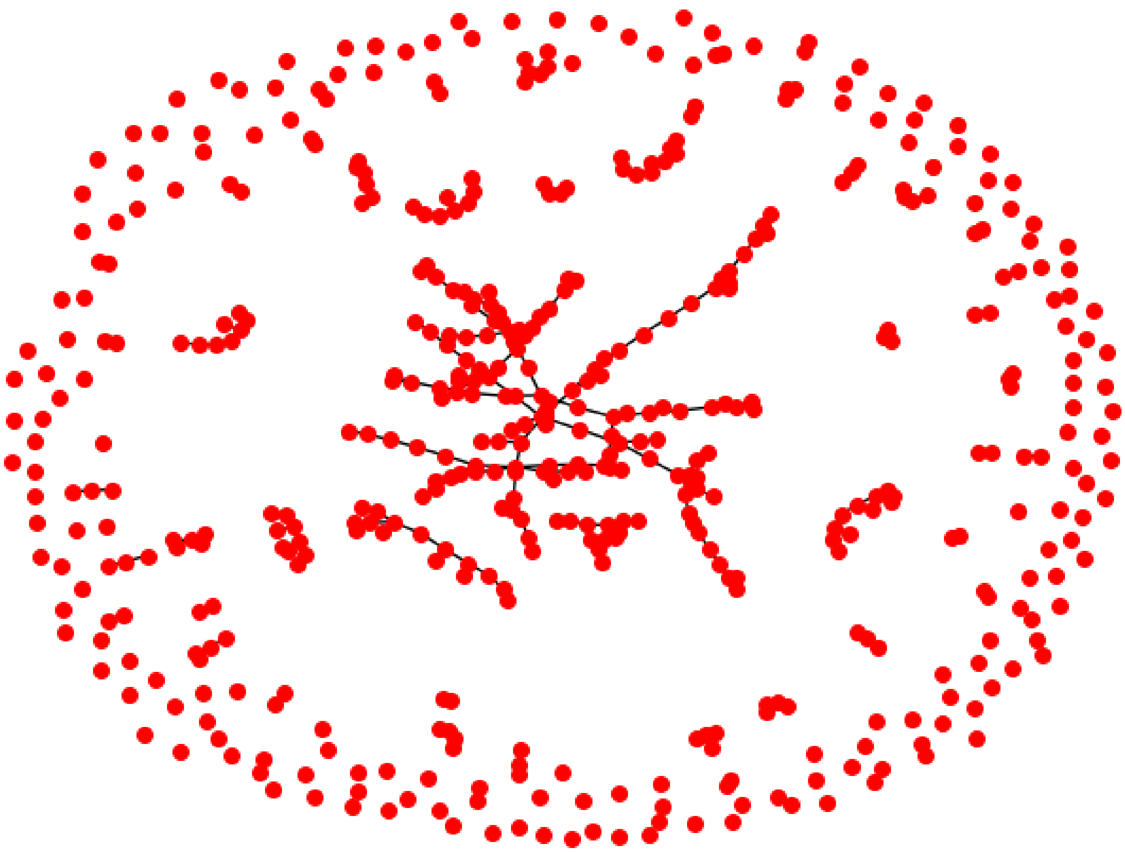
$$p = \frac{\langle k \rangle}{N} = \frac{1}{500} = 0.002$$

Ta sử dụng đoạn code sau để sinh mạng:

```
import networkx as nx
import matplotlib.pyplot as plt

N = 500
p = 0.002
G1 = nx.erdos_renyi_graph(N, p)
```

```
pos = nx.spring_layout(G1)
nx.draw(G1, pos, node_color='red', with_labels=False, node_size=30)
plt.axis('off')
plt.show()
```



(c) ⟨k⟩ = 8

Solution:

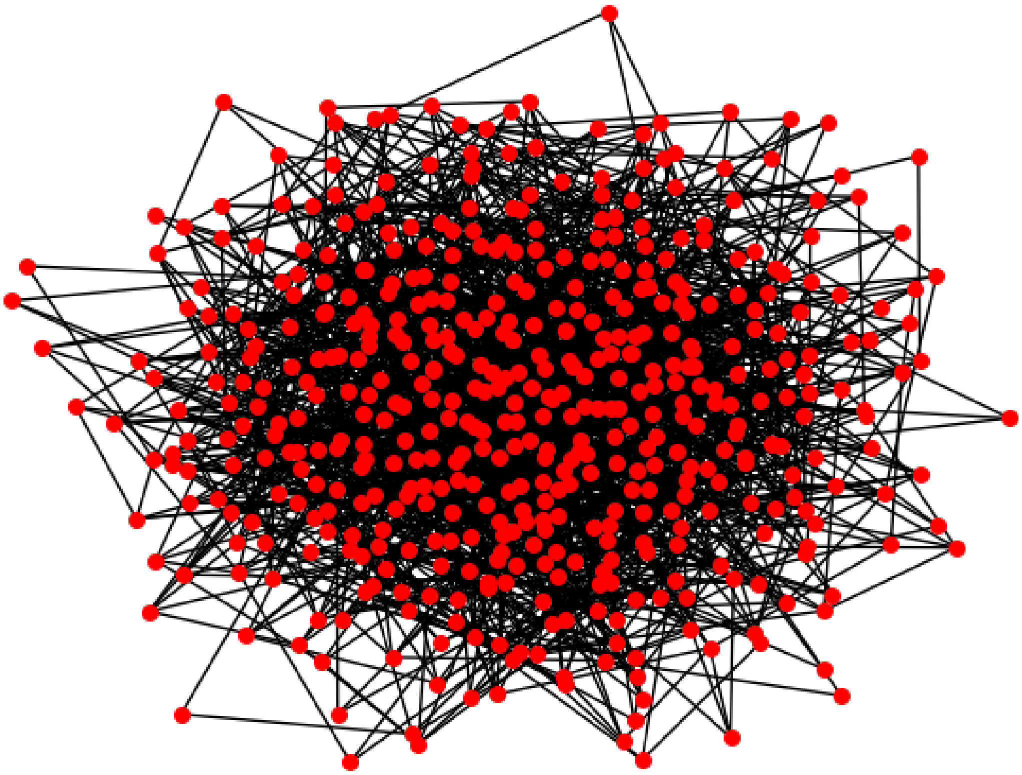
$$\boldsymbol{P = \frac{8}{500} = 0.016}$$

Ta sử dụng đoạn code sau để sinh mạng:

```
import networkx as nx
import matplotlib.pyplot as plt

N = 500
p = 0.016
G1 = nx.erdos_renyi_graph(N, p)
```

```
pos = nx.spring_layout(G1)
nx.draw(G1, pos, node_color='red', with_labels=False, node_size=30)
plt.axis('off')
plt.show()
```



Problem 3: Nghịch lý Tình bạn

(a) Tìm nhân tử chuẩn hóa A.

Solution:

Ta có: $q_k = A * k * p_k$

Giả sử mạng có phân phối bậc luật lũy thừa với $2 < \gamma < 3$, bậc nhỏ nhất là k_{min} và bậc lớn nhất là k_{max}, ta có:

$$P_k \sim k^{(-\gamma)} \quad \text{(với } kmin \leq k \leq kmax\text{)}$$

Ta có điều kiện: $\Sigma(k = kmin \rightarrow kmax)q_k = 1$ thay q_k vào điều kiện trên ta có:

$$A * \sum_{k=k_{\min}}^{k_{\max}} k * p_k = 1$$

⇔

$$A * \sum_{k=k_{\min}}^{k_{\max}} k * (1-\gamma) = 1$$

Suy ra:

$$A = \frac{1}{\sum_{k=k_{\min}}^{k_{\max}} k * (1-\gamma)}$$

$$= \frac{2-\gamma}{k_{\max}^{(2-\gamma)} - k_{\min}^{(2-\gamma)}}$$

(b) Bậc trung bình của các hàng xóm của một nút được chọn ngẫu nhiên là:

$$\langle k_{\mathrm{nn}} \rangle = \sum_k k * q_k$$

Thay $q_k = A * k * p_k$ vào biểu thức trên ta có:

$$\langle k_{\mathrm{nn}} \rangle = A * \sum_k k^{(2-\gamma)}$$

Với A đã tính từ câu (a) và dung xấp xỉ tích phân để tính tổng ta có:

$$\langle k_{nn} \rangle = A \frac{k_{\max}^{3-\gamma} - k_{\min}^{3-\gamma}}{3-\gamma} = \frac{(2-\gamma)}{k_{\max}^{2-\gamma} - k_{\min}^{2-\gamma}} \frac{k_{\max}^{3-\gamma} - k_{\min}^{3-\gamma}}{3-\gamma}$$

Với các giá trị γ , $k_{\min}$, $k_{\max}$ cho trước thì ta có thể tính được bậc trung bình.

(c) Tính toán bậc trung bình của một nút được chọn ngẫu nhiên trong một mạng

Ta có :

$$\langle k \rangle = \sum_k k \cdot p_k \approx \int_{k_{\min}}^{k_{\max}} x^{1-\gamma} dx = \frac{x^{2-\gamma}}{2-\gamma} \Big|_{k_{\min}}^{k_{\max}} = \frac{k_{\max}^{2-\gamma} - k_{\min}^{2-\gamma}}{2-\gamma} = \frac{1000^{2-2.3} - 1^{2-2.3}}{2-2.3} \approx 27.59$$

Tương tự ta tính $\langle k_{nn} \rangle$ với công thức đã tìm ra ở câu (b) và thay các giá trị $\gamma = 2.3, k_{min} = 1, k_{max} = 1000$ ta được:

$$\langle k_{nn} \rangle = \frac{(2-2.3)}{1000^{2-2.3} - 1^{2-2.3}} \frac{1000^{3-2.3} - 1^{3-2.3}}{3-2.3} \approx 52.98$$

Ta thấy $\langle k_{nn} \rangle$ > $\langle k \rangle$, điều này phù hợp với "nghịch lý tính bạn" trong mạng lưới. Các nút bậc cao thường có nhiều hang xóm hơn, làm tang bậc trung bình của các hang xóm so với bậc trung bình của mạng.

<u>(d) Và bây giờ, bạn giải thích "nghịch lý" trong câu (c), rằng là bạn bè của một nút có nhiều bạn bè hơn chính nút đó?</u>

"Nghịch lý tình bạn" (friendship paradox) trong mạng lưới xã hội là một hiện tượng khá phổ biến và có thể giải thích như sau:

Trong một mạng lưới, các nút (đại diện cho người dùng) có số lượng liên kết (bạn bè) khác nhau, phân bố theo một phân phối bậc nhất định (thường là phân phối luật lũy thừa). Các nút có bậc cao (nhiều bạn bè) thường có nhiều khả năng được chọn làm "hàng xóm" (bạn bè) của một nút khác hơn so với các nút có bậc thấp.

Khi chúng ta chọn một nút ngẫu nhiên trong mạng, bậc trung bình của các hàng xóm (bạn bè) của nút đó (⟨k_nn⟩) thường sẽ lớn hơn bậc trung bình của toàn mạng (⟨k⟩). Điều này xảy ra vì các nút có bậc cao có nhiều khả năng được chọn làm hàng xóm hơn, do đó làm tăng bậc trung bình của các hàng xóm so với bậc trung bình của toàn mạng.