

# Độ đo cơ bản của mạng (Basic Network Metrics)

## 1. Average Degree

**Ý nghĩa:** Trung bình số kết nối (cạnh) mà một nút trong mạng có.

**Công thức:**

$$\text{Average Degree} = \frac{2E}{N}$$

Trong đó:

- $E$ : Số cạnh
- $N$ : Số nút

**Phạm vi:**

- $0 \leq \text{Average Degree} \leq N - 1$

**Chỉ số tốt:** Giá trị lớn hơn thể hiện mạng dày đặc hơn, tuy nhiên cần xét trong ngữ cảnh cụ thể.

## 2. Network Diameter

**Ý nghĩa:** Khoảng cách lớn nhất giữa hai nút bất kỳ trong mạng (tính bằng số cạnh).

**Công thức:**

$$\text{Diameter} = \max(d(u, v))$$

Trong đó  $d(u, v)$  là độ dài đường đi ngắn nhất giữa hai nút  $u$  và  $v$ .

**Phạm vi:**

- $1 \leq \text{Diameter} \leq N - 1$

**Chỉ số tốt:** Giá trị nhỏ hơn thể hiện mạng liên kết mạnh và hiệu quả.

## 3. Graph Density

**Ý nghĩa:** Tỷ lệ giữa số cạnh hiện có và số cạnh tối đa có thể.

**Công thức:**

$$\text{Density} = \frac{2E}{N(N-1)}$$

Trong đó:

- $E$ : Số cạnh
- $N$ : Số nút

**Phạm vi:**

- $0 \leq \text{Density} \leq 1$

**Chỉ số tốt:** Giá trị cao thể hiện mạng dày đặc hơn, nhưng cần cân nhắc độ phức tạp của mạng.

## 4. Connected Components

**Ý nghĩa:** Số lượng các thành phần liên thông (cụm nút liên kết với nhau).

**Công thức:** Đếm số cụm liên thông trong đồ thị.

**Phạm vi:**

- $1 \leq \text{Connected Components} \leq N$

**Chỉ số tốt:** Số lượng thấp (thường là 1) thể hiện mạng liên kết tốt.

## 5. Average Path Length (APL)

**Ý nghĩa:** Độ dài trung bình của đường đi ngắn nhất giữa hai nút bất kỳ trong mạng.

**Công thức:**

$$\text{APL} = \frac{1}{\binom{N}{2}} \sum_{i < j} d(i, j)$$

Trong đó  $d(i, j)$  là khoảng cách giữa hai nút  $i$  và  $j$ .

**Phạm vi:**

- $1 \leq \text{APL}$

**Chỉ số tốt:** Giá trị nhỏ thể hiện mạng hiệu quả hơn.

## 6. Average Clustering Coefficient (ACC)

**Ý nghĩa:** Đo mức độ các nút trong mạng có xu hướng tạo thành cụm.

**Công thức:**

$$ACC = \frac{1}{N} \sum_{i=1}^N C_i$$

Với:

$$C_i = \frac{2e_i}{k_i(k_i - 1)}$$

Trong đó:

- $e_i$ : Số cạnh giữa các nút láng giềng của nút  $i$
- $k_i$ : Số láng giềng của nút  $i$

**Phạm vi:**

- $0 \leq ACC \leq 1$

**Chỉ số tốt:** Giá trị cao thể hiện mạng có tính cụm chặt chẽ, thường tốt cho mạng xã hội.

## Độ đo tính trung tâm (Centrality Metrics)

### 1. Degree Centrality

**Ý nghĩa:** Số đo này giúp đo số lượng các mối quan hệ trực tiếp của một tác nhân với các thành viên khác trong mạng xã hội.

**Công thức:**

$$C_D(v) = \frac{\deg(v)}{n - 1}$$

Trong đó:

- $n$ : Số đỉnh của đồ thị
- $\deg(v)$ : Tổng số liên kết trực tiếp đến đỉnh (bậc của đỉnh)

**Phạm vi:**

- Giá trị nằm trong  $[0,1]$

**Chỉ số tốt:**

- Giá trị càng gần 1: tính trung tâm càng lớn.
- Dùng để xác định actor quan trọng (key players).

## 2. Betweenness Centrality

**Ý nghĩa:** Đo lường mức độ mà một nút nằm trên các đường đi ngắn nhất giữa các cặp nút khác.

**Công thức:**

$$C_B(v) = \sum_{s \neq v \neq t} \frac{\sigma_{st}(v)}{\sigma_{st}}$$

Trong đó:

- $\sigma_{st}$ : Tổng số đường đi ngắn nhất giữa  $s$  và  $t$ .
- $\sigma_{st}(v)$ : Số đường đi ngắn nhất giữa  $s$  và  $t$  đi qua nút  $v$ .

**Phạm vi:**

- $0 \leq C_B(v) \leq 1$

**Chỉ số tốt:** Giá trị cao cho thấy nút đóng vai trò trung gian quan trọng trong việc kết nối các phần khác nhau của mạng. Tuy nhiên, nếu chỉ một nút có giá trị quá cao, mạng có thể dễ bị phân mảnh nếu nút này bị xóa.

## 3. Closeness Centrality

**Ý nghĩa:** Đo lường mức độ gần gũi của một nút với tất cả các nút khác trong mạng.

**Công thức:**

$$C_C(v) = \frac{1}{\sum_{u \neq v} d(u, v)}$$

Trong đó:

- $d(u, v)$ : Khoảng cách ngắn nhất giữa nút  $u$  và  $v$ .

**Phạm vi:**

- Giá trị lớn hơn thể hiện nút gần gũi với các nút khác hơn.

**Chỉ số tốt:** Giá trị cao cho thấy nút có thể giao tiếp hiệu quả với các nút khác trong mạng, giảm thời gian truyền thông tin. Đặc biệt quan trọng trong mạng xã hội và mạng giao thông.

## 4. Eigenvector Centrality

**Ý nghĩa:** Đo lường tầm ảnh hưởng của một nút dựa trên tầm ảnh hưởng của các nút mà nó liên kết.

**Công thức:**

$$C_E(v) = \frac{1}{\lambda} \sum_{u \in \text{Adj}(v)} C_E(u)$$

Trong đó:

- $\lambda$ : Hằng số tỷ lệ (eigenvalue).
- $\text{Adj}(v)$ : Tập các nút láng giềng của nút  $v$ .

**Chỉ số tốt:** Giá trị cao thể hiện nút không chỉ kết nối với nhiều nút khác mà các nút liên kết cũng quan trọng. Điều này rất hữu ích trong việc xác định các tác nhân có sức ảnh hưởng lớn trong mạng.

## 5. PageRank

**Ý nghĩa:** Đo lường mức độ quan trọng của một nút trong mạng, được sử dụng phổ biến trong các thuật toán tìm kiếm.

**Công thức:**

$$\text{PR}(v) = \frac{1-d}{N} + d \sum_{u \in \text{Adj}(v)} \frac{\text{PR}(u)}{\deg(u)}$$

Trong đó:

- $d$ : Hệ số giảm dần (thường  $d = 0.85$ ).
- $N$ : Số nút trong mạng.
- $\deg(u)$ : Bậc của nút  $u$ .

**Chỉ số tốt:** Giá trị cao cho thấy nút được liên kết bởi nhiều nút quan trọng khác. Điều này đặc biệt hữu ích trong việc xác định các trang web hoặc nút nổi bật trong hệ thống.

## 6. HITS (Hub and Authority)

**Ý nghĩa:** Đo lường vai trò của một nút như là trung tâm (hub) hoặc nguồn thông tin (authority).

**Công thức:**

$$H_v = \sum_{u \in \text{Adj}(v)} A_u$$
$$A_v = \sum_{u \in \text{Adj}(v)} H_u$$

Trong đó:

- $H_v$ : Điểm số hub của nút  $v$ .
- $A_v$ : Điểm số authority của nút  $v$ .

**Chỉ số tốt:**

- Điểm hub cao: Nút là một nguồn cung cấp thông tin quan trọng.
- Điểm authority cao: Nút là một nguồn thông tin được tham chiếu nhiều.

## 7. Eccentricity

**Ý nghĩa:** Đo khoảng cách xa nhất từ một nút đến bất kỳ nút nào khác trong mạng.

**Công thức:**

$$\text{Ecc}(v) = \max_{u \in V} d(v, u)$$

Trong đó:

- $d(v, u)$ : Khoảng cách ngắn nhất giữa nút  $v$  và  $u$ .
- $V$ : Tập tất cả các nút trong mạng.

**Chỉ số tốt:** Giá trị nhỏ thể hiện nút ở vị trí trung tâm của mạng, dễ tiếp cận với các nút khác. Điều này đặc biệt hữu ích trong các mạng giao thông hoặc mạng thông tin.