

## I. ĐỘ ĐO CƠ BẢN CỦA MẠNG

Độ đo cơ bản của mạng là các chỉ số định lượng được sử dụng để mô tả và phân tích cấu trúc, tính chất và hành vi của một mạng. Những độ đo này giúp các nhà nghiên cứu và chuyên gia hiểu rõ hơn về cách các phần tử trong mạng kết nối với nhau và ảnh hưởng lẫn nhau.

### 1. Average Degree (Độ trung bình của các nút)

**Ý nghĩa:** Trung bình số liên kết (degree) của các nút trong mạng, phản ánh mức độ kết nối trung bình.

**Công thức:**

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

Trong đó:

- E: Số cạnh (edges) trong mạng.
- N: Số nút (nodes) trong mạng.

**Phạm vi:** Giá trị dao động từ 0 (mạng không có kết nối) đến  $N-1$  (mạng đầy đủ, nơi mỗi nút kết nối với tất cả các nút khác).

**Tốt:** Giá trị phụ thuộc vào mục đích của mạng. Mạng giao thông cần kết nối cao, trong khi mạng bảo mật cần kết nối thấp.

### 2. Network Diameter (Đường kính của mạng)

**Ý nghĩa:** Khoảng cách dài nhất giữa hai nút bất kỳ trong mạng, tính theo số cạnh trong đường đi ngắn nhất.

**Công thức:**

$$\text{Diameter} = \max_{i,j} \{\text{Shortest Path Length}(i, j)\}$$

Trong đó:

- $i, j$ : Hai nút bất kỳ trong mạng.

**Phạm vi:** Từ 1 (mạng đầy đủ) đến vô hạn (nếu mạng không liên thông).

**Tốt:** Đường kính nhỏ thường tốt cho hiệu quả giao tiếp, ví dụ: mạng xã hội hoặc mạng viễn thông.

### 3. Graph Density (Mật độ đồ thị)

**Ý nghĩa:** Tỷ lệ số cạnh hiện có so với số cạnh tối đa có thể có.

**Công thức:**

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

Trong đó:

- $E$ : Số cạnh.
- $N$ : Số nút.

**Phạm vi:** Từ 0 (mạng không có kết nối) đến 1 (mạng đầy đủ).

**Tốt:** Mật độ cao thường tốt cho mạng cộng tác, nhưng có thể không hiệu quả trong mạng quy mô lớn.

#### 4. Connected Components (Thành phần liên thông)

**Ý nghĩa:** Số nhóm con của mạng mà trong đó tất cả các nút được kết nối trực tiếp hoặc gián tiếp.

**Công thức:** Đếm số lượng thành phần liên thông trong mạng bằng thuật toán duyệt đồ thị (DFS/BFS).

**Phạm vi:** Từ 1 (mạng hoàn toàn liên thông) đến  $N$  (tất cả nút đều rời rạc).

**Tốt:** Mạng ít thành phần liên thông hoặc chỉ có một thành phần liên thông thường được xem là tốt cho tính kết nối.

#### 5. Average Path Length (Độ dài đường đi trung bình)

**Ý nghĩa:** Trung bình khoảng cách ngắn nhất giữa tất cả các cặp nút trong mạng.

**Công thức:**

$$\text{Average Path Length} = \frac{\sum_{i \neq j} \text{Shortest Path Length}(i, j)}{N(N-1)}$$

**Phạm vi:** Giá trị thấp cho thấy mạng nhỏ gọn và hiệu quả.

**Tốt:** Giá trị thấp thường tốt cho tốc độ truyền tải thông tin (ví dụ: mạng xã hội, mạng máy tính).

#### 6. Average Clustering Coefficient (Hệ số phân cụm trung bình)

**Ý nghĩa:** Đo lường mức độ mà các nút trong mạng có xu hướng tạo thành cụm (cluster).

**Công thức:**

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

**Phạm vi:** Từ 0 (không có cụm) đến 1 (mọi nút tạo thành cụm hoàn chỉnh).

**Tốt:** Giá trị cao phù hợp với mạng xã hội, mạng sinh học, hoặc các hệ thống có tính cụm tự nhiên.

## II. ĐỘ ĐO TRUNG TÂM

### 1. Degree Centrality (In-degree và Out-degree với đồ thị có hướng):

- **Ý nghĩa:**

- Đo lường mức độ kết nối trực tiếp của một node trong mạng.
  - **In-degree:** Số cạnh đi vào một node (được trở đến).
  - **Out-degree:** Số cạnh đi ra từ một node (trở đến node khác).
- Phản ánh mức độ tương tác hoặc phổ biến của một node trong mạng.

- **Công thức:**

$$C_D(v) = \text{Degree of node } v$$

- Với đồ thị không có hướng:  $C_D(v) = \text{In-degree} + \text{Out-degree}$ .

- **Phạm vi:**

- $[0, N - 1]$ , với  $N$  là số lượng node trong mạng.

- **Chỉ số tốt:**

- Giá trị cao: Node có nhiều kết nối, thường là trung tâm hoặc phổ biến.
- Ví dụ: Trong mạng xã hội, các node có nhiều In-degree có thể đại diện cho người nổi tiếng (được nhiều người theo dõi).

### 2. Betweenness Centrality

- **Ý nghĩa:**

- Đo lường mức độ một node đóng vai trò cầu nối (bridge) trong mạng, tức là số lần một node nằm trên đường đi ngắn nhất giữa hai node khác.
- Phản ánh tầm quan trọng của node trong việc kiểm soát dòng thông tin qua mạng.

- Công thức:

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

Trong đó:

- $\sigma_{st}$ : Số đường đi ngắn nhất giữa  $s$  và  $t$ .
- $\sigma_{st}(v)$ : Số đường đi ngắn nhất giữa  $s$  và  $t$  đi qua  $v$ .
- Phạm vi:
  - $[0, 1]$  (được chuẩn hóa).
- Chỉ số tốt:
  - Giá trị cao: Node đóng vai trò quan trọng trong việc kết nối các cộng đồng khác nhau.

### 3. Closeness Centrality

- Ý nghĩa:
  - Đo lường mức độ gần gũi của một node với tất cả các node khác trong mạng (ngược lại với khoảng cách trung bình từ một node đến các node khác).
  - Phản ánh khả năng tiếp cận của một node với các node khác.
- Công thức:

$$C_C(v) = \frac{1}{\sum_t d(v, t)}$$

Trong đó  $d(v, t)$  là đường đi ngắn nhất từ  $v$  đến  $t$ .

- Phạm vi:
  - $[0, 1]$ .
- Chỉ số tốt:
  - Giá trị cao: Node có thể tiếp cận nhanh chóng tất cả các node khác trong mạng.

### 4. Eigenvector Centrality

- Ý nghĩa:
  - Đo lường tầm quan trọng của một node dựa trên mức độ kết nối với các node quan trọng khác.
  - Phản ánh ảnh hưởng gián tiếp và trực tiếp của một node trong mạng.

- Công thức:

$$C_E(v) = \frac{1}{\lambda} \sum_{u \in N(v)} A_{vu} C_E(u)$$

Trong đó:

- $A_{vu}$ : Ma trận kề của đồ thị.
- $\lambda$ : Giá trị riêng (eigenvalue).
- Phạm vi:
  - $[0, 1]$ .
- Chỉ số tốt:
  - Giá trị cao: Node được kết nối với các node có ảnh hưởng lớn trong mạng.

## 5. PageRank

- Ý nghĩa:
  - Đo lường tầm quan trọng của một node dựa trên sự chuyển hướng ngẫu nhiên trong mạng (random walk). Thường dùng để xếp hạng các trang web trong Google Search.
  - Phản ánh sự phổ biến dựa trên sự chú ý hoặc truy cập.
- Công thức:

$$PR(v) = \frac{1-d}{N} + d \sum_{u \in M(v)} \frac{PR(u)}{\text{Out-degree}(u)}$$

Trong đó:

- $d$ : Hệ số giảm (thường là 0.85).
- $M(v)$ : Các node trỏ đến  $v$ .
- Phạm vi:
  - $[0, 1]$ .
- Chỉ số tốt:
  - Giá trị cao: Node nhận được nhiều sự chú ý hoặc được truy cập thường xuyên.

## 6. HITS (Hub and Authority)

- **Ý nghĩa:**

- Phân tích vai trò của node thành hai loại:
  - **Hub:** Node trở đến nhiều **Authority**.
  - **Authority:** Node được trở đến bởi nhiều **Hub**.
- Phản ánh vai trò nguồn cung cấp thông tin (Hub) và sự ảnh hưởng (Authority) của node.

- **Công thức:**

$$\text{Authority}(v) = \sum_{u \in \text{In-neighbors}(v)} \text{Hub}(u)$$

$$\text{Hub}(v) = \sum_{u \in \text{Out-neighbors}(v)} \text{Authority}(u)$$

- Thuật toán lặp để hội tụ giá trị.

- **Phạm vi:**

- $[0, 1]$ .

- **Chỉ số tốt:**

- **Hub cao:** Node là nguồn thông tin quan trọng.
- **Authority cao:** Node có ảnh hưởng lớn trong mạng.

## 7. Eccentricity

- **Ý nghĩa:**

- Khoảng cách lớn nhất từ một node đến bất kỳ node nào khác trong mạng.
- Phản ánh "độ xa" của một node trong mạng

- **Công thức:**

$$E(v) = \max_u d(v, u)$$

Trong đó  $d(v, u)$  là khoảng cách ngắn nhất giữa  $v$  và  $u$ .

- **Phạm vi:**

- $[1, \text{Diameter of the Network}]$ .

- **Chỉ số tốt:**

- Giá trị nhỏ: Node nằm ở vị trí trung tâm của mạng.

