

1. Độ đo cơ bản của mạng (Basic Network Metrics):

- Average Degree

Ý nghĩa: Average Degree của một đồ thị vô hướng được sử dụng để đo lường số lượng cạnh so với số lượng đỉnh.

Công thức:

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

- E: Số cạnh (edges)
- N: Số đỉnh (nodes)

Phạm vi: Từ 0 đến N-1

Chỉ số được cho là tốt: Vừa phải không quá cao cũng không quá thấp. Chắc là tại vì khi kết nối quá thấp thì cộng đồng hơi thiếu gắn kết nếu kết nối quá cao thì cảm thấy không phù hợp, trông không giống một xã hội mà giống một gia đình hơn.

- Network Diameter

Ý nghĩa: Là độ dài ngắn nhất của đường đi giữa cặp đỉnh xa nhất trong mạng.

Công thức: Đo khoảng cách ngắn nhất giữa tất cả các cặp đỉnh và lấy giá trị lớn nhất.

Phạm vi: từ 1 => vô cùng / luôn dương và khác không.

Chỉ số được cho là tốt: Mạng có đường kính nhỏ thể hiện các đỉnh có thể giao tiếp nhanh chóng với nhau (tốt cho mạng cần hiệu quả cao).

- Graph Density

Ý nghĩa: **Graph Density (Mật độ đồ thị)** là một chỉ số quan trọng để đánh giá mức độ kết nối của một mạng. Mật độ đồ thị tính toán tỷ lệ giữa số lượng cạnh thực tế có trong đồ thị và số lượng cạnh tối đa có thể có nếu tất cả các nút trong mạng đều được kết nối với nhau.

Công thức:

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

- E: Số cạnh (edges)
- N: Số đỉnh (nodes)

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

Giải thích:

- **Mật độ cao:** Nếu mật độ gần 1, có nghĩa là đồ thị có rất nhiều kết nối, gần như đầy đủ.
- **Mật độ thấp:** Nếu mật độ gần 0, có nghĩa là đồ thị có ít kết nối, có thể chỉ có một vài nút kết nối với nhau.

Chỉ số được cho là tốt: Mạng có mật độ cao cho thấy các nút kết nối chặt chẽ với nhau, nhưng vẫn sẽ tùy vào mục đích sử dụng mạng.

- Connected Components

Ý nghĩa: phân ra từng nhóm, trong mỗi nhóm đó thì tất cả các nút đều có kết nối với nhau, một group. Hình thức đại diện có thể hiểu như kiểu một hội nhóm, câu lạc bộ. Mỗi nhóm như vậy gọi là một thành phần kết nối.

Công thức: Đếm số thành phần kết nối trong mạng.

- E: Số cạnh (edges)
- N: Số đỉnh (nodes)

Phạm vi: Từ 1 (mạng liên thông) đến nhiều hơn nếu mạng bị phân mảnh.

Chỉ số được cho là tốt: Một mạng có ít thành phần kết nối (càng gần 1 càng tốt).

- Average Path Length

Ý nghĩa: trong một đồ thị đo lường khoảng cách trung bình giữa tất cả các cặp nút trong mạng. Nó là số bước (hoặc cạnh) trung bình cần thiết để đi từ một nút này đến một nút khác trong đồ thị.

Công thức:

$$\text{Average Path Length} = \frac{\sum_{\text{Pair of nodes}} \text{Shortest Path Length}}{\text{Total number of pairs of nodes}}$$

- **Shortest Path Length:** Là số bước hoặc cạnh ngắn nhất giữa hai nút trong đồ thị.
- **Total number of pairs of nodes:** Là tổng số cặp nút có thể kết nối với nhau.

Phạm vi: Từ 1 đến vô cùng (tùy thuộc vào cấu trúc mạng).

- **Đồ thị liên thông:** Nếu đồ thị có một thành phần kết nối duy nhất, tức là mọi nút đều có thể liên kết với nhau, thì Average Path Length sẽ tính toán khoảng cách trung bình giữa tất cả các nút trong mạng.
- **Đồ thị không liên thông:** Nếu đồ thị có nhiều thành phần kết nối, một số cặp nút không thể kết nối được, do đó không tính vào chiều dài đường đi.

Chỉ số được cho là tốt: Mạng có chiều dài đường đi trung bình thấp thường hiệu quả hơn, cho thấy kết nối giữa các nút nhanh và dễ dàng.

- Average Clustering Coefficient

Ý nghĩa: **Average Clustering Coefficient** (Hệ số cụm trung bình) là một chỉ số dùng để đo lường mức độ các nút trong mạng kết nối chặt chẽ với nhau. Nó cho thấy mức độ các nút trong mạng có xu hướng tạo thành các nhóm nhỏ, nơi mà các nút trong nhóm đó đều kết nối với nhau.

Công thức:

$$C = \frac{1}{N} \sum_i c_i$$

- với c_i là hệ số cụm của nút i .
Hệ số cụm của mỗi nút được tính bằng tỷ lệ các liên kết thực tế có giữa các láng giềng của nó so với số lượng liên kết tối đa có thể có.
- **Phạm vi:** Từ 0 đến 1.
- **Chỉ số tốt:** Hệ số cụm cao cho thấy các nút trong mạng có sự kết nối với nhau rất chặt chẽ.

2. Độ đo tính trung tâm (Centrality Metrics)

Degree Centrality (Tính trung tâm theo bậc)

Ý nghĩa: Đo lường mức độ quan trọng của một nút dựa trên số liên kết mà nó có. Trong đồ thị có hướng, có hai loại degree centrality:

- **In-degree:** Số lượng liên kết hướng vào nút.
- **Out-degree:** Số lượng liên kết hướng ra từ nút.
- **Công thức:**

In-degree:

$$C_D^{in}(v) = deg^{in}(v)$$

(Số liên kết đến nút v).

Out-degree:

$$C_D^{out}(v) = deg^{out}(v)$$

(Số liên kết từ nút v).

- **Phạm vi:** Từ 0 đến $N-1$ (tùy vào số lượng đỉnh trong mạng).
- **Chỉ số tốt:** Các nút có số liên kết cao (degree centrality cao) được xem là quan trọng trong mạng, vì chúng dễ dàng tiếp cận hoặc có ảnh hưởng lớn.

- Betweenness Centrality

Ý nghĩa: Betweenness centrality đo lường mức độ mà một nút xuất hiện trên các con đường ngắn nhất giữa các nút khác trong mạng. Nút có betweenness cao có thể đóng vai trò quan trọng trong việc kết nối các phần khác nhau của mạng.

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

Phạm vi: Từ 0 đến $\frac{(N-1)(N-2)}{2}$, nơi N là số lượng đỉnh trong mạng

Chỉ số tốt: Một nút có betweenness centrality cao thường có vai trò quan trọng trong việc điều phối thông tin và duy trì sự kết nối trong mạng. Đây có thể là một "cầu nối" quan trọng giữa các cụm trong mạng.

Closeness Centrality (Tính trung tâm theo gần gũi)

- **Ý nghĩa:** Closeness centrality đo lường mức độ gần gũi của một nút với tất cả các nút còn lại trong mạng. Một nút có closeness cao có thể tiếp cận các nút khác một cách nhanh chóng, vì nó gần với nhiều nút.
- **Công thức:** $CC(v) = \frac{1}{\sum_{u \neq v} d(v, u)}$

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

$d(v, u)$ là khoảng cách ngắn nhất từ nút v đến nút u.

- **Phạm vi:** Từ 0 đến ∞
- **Chỉ số tốt:** Nút có closeness cao thường có khả năng truyền thông tin nhanh chóng và hiệu quả trong mạng. Nó là một nút "trung tâm" trong mạng với khả năng tiếp cận gần như tất cả các nút khác.

Eigenvector Centrality (Tính trung tâm theo vector riêng)

- **Ý nghĩa:** Eigenvector centrality đo lường mức độ quan trọng của một nút không chỉ dựa trên số lượng liên kết của nó mà còn dựa trên mức độ quan trọng của các nút mà nó kết nối. Nếu một nút kết nối với các nút quan trọng, nó cũng sẽ có eigenvector centrality cao.
- **Công thức:** Được tính thông qua việc giải hệ phương trình: $Ax = \lambda x$
 - A là ma trận kề của đồ thị.
 - x là vector riêng (eigenvector).
 - λ là giá trị riêng (eigenvalue).
- **Phạm vi:** Không giới hạn.
- **Chỉ số tốt:** Nút có eigenvector centrality cao là nút có ảnh hưởng lớn trong mạng, đặc biệt nếu nó kết nối với nhiều nút quan trọng khác.

PageRank

- **Ý nghĩa:** PageRank là một thuật toán do Google phát triển để đánh giá tầm quan trọng của các trang web, trong đó tầm quan trọng của một trang được xác định bởi số lượng và chất lượng của các liên kết trỏ đến trang đó.
- **Công thức:**

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

- d: Hệ số giảm (thường là 0.85).
 - N: Tổng số nút trong mạng.
 - $\text{In}(v)$: Tập hợp các nút có liên kết vào v.
 - $\text{Deg out}(u)$: Số lượng liên kết ra của nút u.
- **Phạm vi:** Từ 0 đến 1.
 - **Chỉ số tốt:** Nút có PageRank cao là một nút quan trọng và có ảnh hưởng trong mạng, bởi vì nó được kết nối với các nút quan trọng khác.

HITS (Hub and Authority)

- **Ý nghĩa:** Thuật toán HITS phân biệt các nút thành hai loại:
 - **Hub:** Nút có nhiều liên kết hướng ra.
 - **Authority:** Nút có nhiều liên kết hướng vào. HITS đo lường tầm quan trọng của nút dựa trên vai trò "Hub" và "Authority" của nó trong mạng.
- **Công thức:**

Hub score: $h(v) = \sum_{u \in In(v)} a(u)$

Authority score: $a(v) = \sum_{u \in out(v)} h(u)$

-
- **Phạm vi:** Không giới hạn.
- **Chỉ số tốt:** Nút "Hub" có thể dẫn dắt lưu lượng thông tin đến các nút "Authority", trong khi các nút "Authority" có vai trò cung cấp thông tin quan trọng.

Eccentricity (Tính ngoại lệ)

- **Ý nghĩa:** Eccentricity đo lường độ xa nhất từ một nút đến bất kỳ nút nào trong mạng. Một nút có eccentricity thấp có thể tiếp cận nhanh chóng các nút khác, trong khi nút có eccentricity cao thường là nút ngoại vi.
- **Công thức:**
- $Eccentricity(v) = \max d(v, u)$

U thuộc V

- $d(v, u)$ là khoảng cách ngắn nhất từ nút v đến u.
- **Phạm vi:** Từ 0 đến đường kính của mạng (network diameter).
- **Chỉ số tốt:** Nút có eccentricity thấp là trung tâm của mạng, có khả năng tiếp cận hầu hết các nút nhanh chóng.