

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

Average Degree (Độ trung bình của mạng)

Average Degree được dùng để biểu thị mức độ kết nối trung bình của các đỉnh trong mạng.

Công thức của Average Degree:

$$\text{Average Degree} = \frac{2 \times |E|}{|V|}$$

Trong đó:

- $|E|$ : Số cạnh trong mạng.
- $|V|$ : Số đỉnh trong mạng.

Average Degree có khoảng giá trị từ  $[0, |V|-1]$ . Giá trị cao cho thấy các đỉnh được kết nối tốt. Tuy nhiên, mức kết nối quá cao có thể không mong muốn trong một số trường hợp (như trong mạng máy tính).

Network Diameter (Đường kính mạng)

Network Diameter là khoảng cách dài nhất (số cạnh) giữa hai đỉnh bất kỳ trong mạng, cho thấy khoảng cách xa nhất mà thông tin phải truyền qua.

Công thức của Network Diameter:

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

Trong đó:

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

Network Diameter có khoảng giá trị giao động từ  $[1, \infty)$ . Đường kính nhỏ là tốt trong các mạng giao tiếp vì đảm bảo thông tin được truyền nhanh hơn.

Graph Density (Độ đặc của đồ thị)

Graph Density có tác dụng biểu thị tỷ lệ giữa số cạnh hiện có và số cạnh tối đa có thể có, cho biết mức độ kết nối tổng thể của mạng.

Công thức:

$$Density = \frac{2 \times |E|}{|V| \times (|V| - 1)}$$

Trong đó:

- $|E|$ : Số cạnh hiện có.
- $|V|$ : Số đỉnh trong mạng.

Graph Density có phạm vi từ  $[0,1]$ , với:

- 0: Mạng rời rạc (không có cạnh).
- 1: Mạng đầy đủ (mọi đỉnh đều được kết nối với nhau).

Mạng có giá trị gần 1 thể hiện sự kết nối mạnh. Tuy nhiên, giá trị lý tưởng phụ thuộc vào ứng dụng (ví dụ, mạng xã hội cần độ đặc cao, nhưng mạng giao thông thì không).

Connected Components (Thành phần liên thông)

Connected Components là số nhóm các đỉnh trong đó mỗi cặp đỉnh trong cùng nhóm có ít nhất một đường đi nối với nhau. Có phạm vi giao động từ  $[1,|V|]$  với:

- 1: Mạng hoàn toàn liên thông.
- $|V|$ : Từng đỉnh tách biệt (không có cạnh)

Mạng tốt nhất là khi chỉ có một thành phần liên thông, đặc biệt trong mạng giao tiếp hoặc mạng internet.

Average Path Length (Độ dài đường đi trung bình)

Average Path Length là giá trị trung bình của khoảng cách ngắn nhất giữa mọi cặp đỉnh, thể hiện hiệu quả trong việc truyền thông tin.

Công thức:

$$\text{Average Path Length} = \frac{\sum_{u \neq v} d(u, v)}{|V| \times (|V| - 1)}$$

Trong đó:

- $d(u,v)$ : Khoảng cách ngắn nhất giữa hai đỉnh  $u$  và  $v$

Average Path Length có khoảng giá trị giao động từ  $[1,\infty)$ . Giá trị càng thấp càng tốt, vì thông tin sẽ được truyền nhanh hơn.

### Average Clustering Coefficient (Hệ số cụm trung bình)

Average Clustering Coefficient dùng để đo lường khả năng các đỉnh trong mạng có xu hướng tạo thành cụm (các đỉnh láng giềng của một đỉnh có xu hướng được kết nối với nhau).

Công thức:

$$C = \frac{1}{|V|} \sum_{v \in V} \frac{2 \sum_{u \in N(v)} |E(u,v)|}{k_v(k_v - 1)}$$

Trong đó

- $|E(v)|$ : Số cạnh giữa các láng giềng của đỉnh  $v$ .
- $k_v$ : Số láng giềng của đỉnh  $v$ .

Average Clustering Coefficient có phạm vi vào khoảng  $[0,1]$ , với:

- 0: Không có cụm.
- 1: Mỗi đỉnh và láng giềng của nó tạo thành một mạng đầy đủ.

Giá trị cao thể hiện mạng có tính "xã hội" cao (các cụm chặt chẽ), phù hợp trong các mạng xã hội.

## 2. Độ đo tính trung tâm (Centrality Metrics)

Degree Centrality (Trung tâm theo số liên kết) được dùng để đo lường số lượng liên kết trực tiếp mà một đỉnh (node) có trong mạng.

- Với đồ thị có hướng:
  - In-degree: Số liên kết đi vào đỉnh.
  - Out-degree: Số liên kết đi ra từ đỉnh.

Công thức:

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

Trong đó:

- $\deg(v)$ : Tổng số liên kết của đỉnh  $v$ .
- $|V|$ : Tổng số đỉnh trong mạng.

Degree Centrality có phạm vi rơi vào khoảng từ  $[0,1]$ , với giá trị 1 khi đỉnh kết nối với tất cả các đỉnh khác. Giá trị cao cho thấy đỉnh đóng vai trò trung tâm với nhiều liên kết trực tiếp. Phù hợp để xác định các nút "nổi bật" trong mạng.

Betweenness Centrality (Trung tâm trung gian) được dùng để đo lường mức độ mà một đỉnh nằm trên các đường đi ngắn nhất giữa các cặp đỉnh khác trong mạng. Phản ánh vai trò "cầu nối" của đỉnh.

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  $v$ .

Betweenness Centrality có phạm vi nằm trong khoảng từ  $[0,1]$ , giá trị cao biểu thị vai trò quan trọng của đỉnh trong việc kết nối các phần của mạng. Các đỉnh có betweenness centrality cao là "nút cổ chai" hoặc "điểm trung gian" quan trọng.

Closeness Centrality (Trung tâm gần gũi) là chỉ số dùng để đo lường mức độ gần gũi của một đỉnh với tất cả các đỉnh khác trong mạng. Phản ánh tốc độ mà thông tin có thể lan truyền từ đỉnh đó đến các đỉnh khác.

Công thức:

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

Trong đó:

- $d(v, u)$ : Khoảng cách ngắn nhất từ đỉnh  $v$  đến  $u$ .

Closeness Centrality có phạm vi rơi vào khoảng từ  $[0,1]$ , giá trị cao cho thấy đỉnh gần với các đỉnh khác. Đỉnh có closeness centrality cao thường đóng vai trò quan trọng trong việc truyền thông tin nhanh chóng.

Eigenvector Centrality (Trung tâm theo giá trị riêng) là chỉ số dùng để đo lường tầm ảnh hưởng của một đỉnh dựa trên tầm ảnh hưởng của các đỉnh kết nối với nó.

Công thức:

$$x_v = \frac{1}{\lambda} \sum_{u \in V} A_{uv} x_u$$

Trong đó:

- $A_{uv}$ : Phần tử trong ma trận kề.
- $x_v$ : Giá trị eigenvector centrality của đỉnh vvv.
- $\lambda$ : Giá trị riêng lớn nhất.

Eigenvector Centrality có phạm vi từ khoảng  $[0,1]$ , giá trị cao biểu thị đỉnh kết nối với các đỉnh "quan trọng". Eigenvector centrality cao thể hiện đỉnh không chỉ có nhiều liên kết mà còn liên kết với các đỉnh có tầm ảnh hưởng cao.

PageRank (Xếp hạng theo Google) là chỉ số dùng để đo lường tầm quan trọng của một đỉnh dựa trên số lượng và chất lượng các liên kết đến nó, được sử dụng rộng rãi trong xếp hạng trang web.

Công thức:

$$PR(v) = \frac{1-d}{|V|} + d \sum_{u \in In(v)} \frac{PR(u)}{deg_{out}(u)}$$

Trong đó:

- $d$ : Hệ số suy giảm (thường là 0.85).
- $In(v)$ : Tập các đỉnh liên kết đến v.
- $deg_{out}(u)$ : Số liên kết đi ra từ u.

PageRank có phạm vi rơi vào khoảng  $[0,1]$ , giá trị cao cho thấy đỉnh có tầm quan trọng lớn. Phù hợp để đánh giá các nút trong mạng web hoặc các hệ thống có cấu trúc tương tự.

HITS (Hub and Authority)

Hub: Đỉnh kết nối tới nhiều đỉnh có vai trò "authority".

Authority: Đỉnh được kết nối từ nhiều "hub".

Công thức của :

Hub:

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

Authority:

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

Hub có phạm vi nằm ở khoảng  $[0,1]$ , giá trị cao thể hiện đỉnh đóng vai trò quan trọng như hub hoặc authority.

Phù hợp trong mạng web để đánh giá tầm quan trọng của trang dựa trên liên kết.

Eccentricity (Độ lệch tâm) là khoảng cách lớn nhất từ một đỉnh đến bất kỳ đỉnh nào khác trong mạng, phản ánh mức độ "xa" của đỉnh đó trong mạng.

Công thức của Eccentricity:

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

Trong đó:

$d(v, u)$  Khoảng cách ngắn nhất từ  $v$  đến  $u$

Eccentricity có phạm vi rơi vào  $[1, \infty)$ . Giá trị thấp cho thấy đỉnh ở gần trung tâm mạng hơn, trong khi giá trị cao biểu thị đỉnh ở rìa mạng.