

Câu 1: Ý nghĩa của các chỉ số độ đo cơ bản của mạng (Basic Network Metrics)

1. Average Degree

- Là số lượng cạnh trung bình nối với mỗi đỉnh trong đồ thị. Nó cho biết mức độ kết nối trung bình của các đỉnh.
- Công thức:

$$\text{Average Degree} = \frac{2 \times k}{n}$$

Trong đó:

- k là tổng liên kết thực (cạnh) của toàn mạng.
- n là tổng các tác nhân (actor - đỉnh) trong mạng.
- Ứng dụng: Đánh giá mức độ kết nối tổng thể của mạng; nếu chỉ số này cao, mạng thường có mật độ kết nối cao hơn.

2. Network Diameter

- Là khoảng cách dài nhất (số cạnh) giữa hai đỉnh bất kỳ trong đồ thị, tính trên đường đi ngắn nhất giữa chúng.
- Công thức:

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

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

- Ứng dụng: Đánh giá kích thước "toàn cầu" của mạng và thời gian lan truyền trong hệ thống.

3. Graph Density

- Là tỷ lệ giữa số lượng cạnh thực tế so với số lượng cạnh tối đa có thể có trong đồ thị.
- Công thức:

$$\text{Density} = \frac{k}{\frac{n(n-1)}{2}}$$

Trong đó:

- k là tổng số cạnh có trong mạng.

- n là số đỉnh có trong mạng.
 - $\frac{n(n-1)}{2}$ là tổng số cạnh khả dĩ có trong mạng.
- Ứng dụng: Cho biết mức độ lấp đầy kết nối trong mạng. Đồ thị có mật độ cao thường mô tả hệ thống rất kết nối.

4. Connected Components

- Là các tập hợp các đỉnh mà trong đó mọi cặp đỉnh đều được kết nối trực tiếp hoặc gián tiếp thông qua các cạnh.
- Ứng dụng: Giúp phân tích cấu trúc mạng, xác định các phần rời rạc, hay đo lường sự cô lập của các nhóm trong mạng.

5. Average Path Length

- Là độ dài trung bình của các đường đi ngắn nhất giữa mọi cặp đỉnh trong mạng.
- Công thức:

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

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

- Ứng dụng: Đánh giá hiệu quả của mạng trong việc truyền thông tin hoặc lan truyền tín hiệu.

6. Average Clustering Coefficient

- Đo lường mức độ mà các đỉnh trong mạng có xu hướng tạo thành các nhóm cụm (cliques).
- Công thức:

$$C = \frac{1}{n} \times \sum_{v \in V} \frac{\text{Số cạnh giữa các láng giềng của } v}{\text{Số kết hợp có thể có giữa các láng giềng}}$$

- Ứng dụng: Đánh giá tính cục bộ của mạng, đặc biệt trong mạng xã hội hoặc mạng sinh học.

Câu 2: Ý nghĩa các chỉ số đo tính trung tâm

1. Degree Centrality

- Đo lường tầm quan trọng của một đỉnh dựa trên số lượng cạnh kết nối trực tiếp với nó.
- Công thức:

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

Trong đó:

- n là số đỉnh của đồ thị.
 - $\deg(v)$ là tổng số liên kết trực tiếp đỉnh v (bậc của đỉnh).
- Phân loại:
 - In-degree: Số cạnh đi vào một đỉnh (trong đồ thị có hướng). Thể hiện mức độ được tham chiếu hoặc nhận thông tin.
 - Out-degree: Số cạnh đi ra từ một đỉnh (trong đồ thị có hướng). Thể hiện mức độ tác động hoặc phát tán thông tin.
 - Ứng dụng: Phân tích các nút quan trọng trong mạng xã hội, mạng lưới giao thông, hoặc mạng máy tính.

2. Betweenness Centrality

- Đo lường mức độ mà một đỉnh nằm trên đường đi ngắn nhất giữa các cặp đỉnh khác.
- Công thức:

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

Trong đó:

- σ_{st} là số đường đi ngắn nhất giữa s và t .
 - $\sigma_{st}(v)$ là số đường qua v .
- Ứng dụng: Xác định các điểm quan trọng trong mạng để truyền thông tin, kiểm soát dòng chảy, hoặc phát hiện các nút "nút cổ chai".

3. Closeness Centrality

- Đo lường mức độ một đỉnh có thể tiếp cận các đỉnh khác một cách nhanh chóng.
- Công thức:

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

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

- Ứng dụng: Tìm các nút có khả năng lan truyền thông tin hoặc ảnh hưởng mạnh nhất trong mạng.

4. Eigenvector Centrality

- Đo lường tầm quan trọng của một đỉnh dựa trên tầm quan trọng của các đỉnh mà nó kết nối.
- Công thức:

$$C_E(v) \propto \sum_{u \in N(v)} C_E(u)$$

Trong đó $N(v)$ là tập láng giềng của v .

- Ứng dụng: Tìm các nút có ảnh hưởng toàn cục mạnh mẽ trong mạng, chẳng hạn như trong mạng xã hội hoặc hệ thống tài chính.

5. PageRank

- Là một biến thể của Eigenvector Centrality, nhưng thêm trọng số cho các cạnh dựa trên mức độ quan trọng của đỉnh nguồn.
- Công thức:

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

Trong đó: d là hệ số giảm dần, thường là 0.85.

- Ứng dụng: Được sử dụng trong các công cụ tìm kiếm như Google để xếp hạng các trang web.

6. HITS

- Ý nghĩa:
 - Hub: Đỉnh có khả năng liên kết đến nhiều đỉnh quan trọng (authority).
 - Authority: Đỉnh được liên kết bởi nhiều hub quan trọng.
- Công thức: Dựa trên hai điểm số lặp lại:

$$h(v) = \sum_{u \in N^{out}(v)} a(u), \quad a(v) = \sum_{u \in N^{in}(v)} h(u)$$

Trong đó: $N^{out}(v)$ và $N^{in}(v)$ là tập các đỉnh láng giềng có hướng đi ra và đi vào.

- Ứng dụng: Phân tích mạng liên kết như hệ thống web hoặc hệ thống học thuật (trích dẫn bài báo).

7. Eccentricity

- Là khoảng cách lớn nhất từ một đỉnh đến bất kỳ đỉnh nào khác trong đồ thị.
- Công thức:

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

Trong đó: $d(v, u)$ là khoảng cách từ v đến u .

- Ứng dụng: Đánh giá mức độ "xa cách" của một đỉnh trong mạng. Đỉnh có eccentricity nhỏ thường nằm gần trung tâm.