MẠNG XÃ HỘI

Bài 2. ĐỒ THỊ MẠNG XÃ HỘI

ThS. Lê Nhật Tùng

Mục lục

Nội dung

2.1.1 Đồ thị

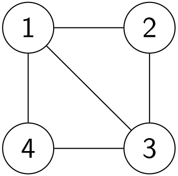
- **Đồ thị** $G = (V, E)$ gồm một tập V gọi là tập đỉnh và một tập E gọi là tập cạnh hay cung.
- Tập $E \subseteq V^2$ gồm các cặp phần tử của V.
- Giả sử u và v là hai đỉnh của đồ thị G $(u, v \in V)$, nếu cặp đỉnh (u, v) không được sắp thứ tự thì (u, v) gọi là cạnh nối hai đỉnh u và v. Ngược lại, nếu cặp đỉnh (u, v) được sắp thứ tự thì (u, v) gọi là cạnh có hướng (hay cung), trong đó u được gọi là đỉnh đầu và v được gọi là đỉnh cuối.

Các loại đồ thị

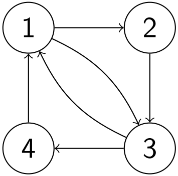
- **Đồ thị vô hướng** là đồ thị chỉ chứa các cạnh trong đồ thị vô hướng, cạnh (u, v) tương đương với cạnh (v, u).
- **Đồ thị có hướng** là đồ thị chỉ chứa các cạnh có hướng (cung). Trong đồ thị có hướng, cung (u, v) khác với cung (v, u).

Các loại đồ thị

Đồ thị vô hướng



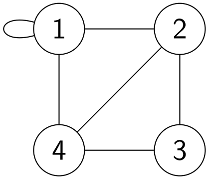
Đồ thị có hướng



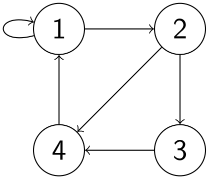
- **Đồ thị vô hướng**: cạnh (u, v) tương đương với cạnh (v, u)
- **Đồ thị có hướng**: cung (u, v) khác với cung (v, u)

Một số khái niệm trên đồ thị

Khuyên trong đồ thị vô hướng



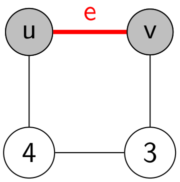
Khuyên trong đồ thị có hướng



- **Khuyên:** Cạnh nối một đỉnh với chính nó được gọi là một khuyên.
- Ví dụ: Trong cả hai đồ thị trên, đỉnh 1 có một khuyên.

Một số khái niệm trên đồ thị

Đỉnh kề và cảnh liên thuộc

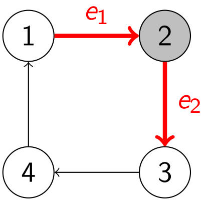


- **Đỉnh kề và cạnh liên thuộc:** Trong đồ thị, hai đỉnh $u, v \in V$ $(u \neq v)$ được gọi là kề nhau nếu tồn tại cạnh $e = (u, v) \in E$.
- Khi đó, cạnh e được gọi là liên thuộc với đỉnh u và v.

- Trong hình minh họa:
  - Đỉnh u và v là các đỉnh kề nhau (tô xám)
  - Cạnh e là cạnh liên thuộc với u và v (tô đỏ)

Cạnh (cung) kề nhau

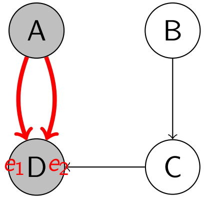
Cung kề nhau qua đỉnh 2



- **Cạnh (cung) kề nhau:** Hai cung e_1 và e_2 được gọi là kề nhau nếu chúng có đỉnh chung
- Với cung, không phụ thuộc vào:
  - Đỉnh chung là đỉnh đầu hay đỉnh cuối của e_1
  - Đỉnh chung là đỉnh đầu hay đỉnh cuối của e_2

- Trong hình minh họa:
  - e_1 và e_2 là hai cung kề nhau (tô đỏ)
  - Đỉnh 2 là đỉnh chung (tô xám)
  - e_1 kề e_2 qua đỉnh 2 (đỉnh cuối của e_1 và đỉnh đầu của e_2)

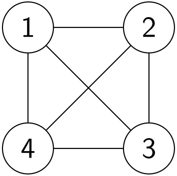
Cung song song giữa A và D



- **Cạnh (cung) song song:** Hai cạnh (cung) được gọi là song song nếu nó nối hai cặp đỉnh giống nhau.
- Trong hình minh họa:
  - e_1 và e_2 là hai cung song song
  - Cùng nối từ đỉnh A đến đỉnh D
  - Các đỉnh A và D được tô xám

2.1.2 Một số dạng đồ thị đặc biệt

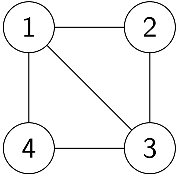
Ví dụ về đơn đồ thị



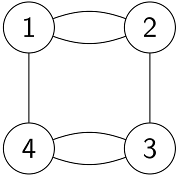
- **Đơn đồ thị** là đồ thị:
  - Không chứa khuyên
  - Mỗi cặp đỉnh chỉ được nối bởi một cạnh duy nhất
- Trong ví dụ:
  - Không có đỉnh nào nối với chính nó
  - Giữa hai đỉnh bất kỳ chỉ có tối đa một cạnh

2.1.2 Một số dạng đồ thị đặc biệt

Đơn đồ thị vô hướng



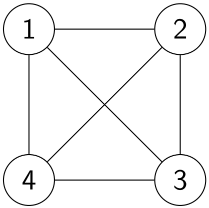
Đa đồ thị vô hướng



- **Đa đồ thị:** là đồ thị mà mỗi cặp đỉnh có thể được nối bởi nhiều hơn một cạnh.
- Trong ví dụ đa đồ thị:
  - Có hai cạnh song song giữa đỉnh 1 và 2
  - Có hai cạnh song song giữa đỉnh 3 và 4

2.1.2 Một số dạng đồ thị đặc biệt

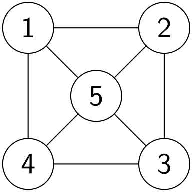
Đồ thị đầy đủ K_4



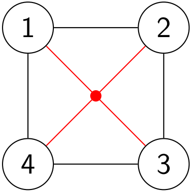
- **Đồ thị đầy đủ** $G = (V, E)$ là đồ thị mà mỗi cặp đỉnh được nối với nhau bằng đúng một cạnh
- Trong ví dụ K_4:
    - Có 4 đỉnh
    - Mỗi đỉnh được nối với tất cả các đỉnh còn lại
    - Tổng số cạnh: 6 cạnh
    - Mỗi đỉnh có bậc 3

2.1.2 Một số dạng đồ thị đặc biệt

Đồ thị phẳng

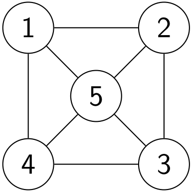


Biểu diễn không phẳng

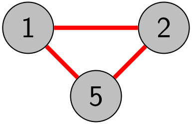


- **Đồ thị phẳng** là đồ thị có thể biểu diễn hình học trên một mặt phẳng mà các cạnh chỉ cắt nhau ở đỉnh
- Ví dụ:
  - Bên trái: Đồ thị phẳng (các cạnh không cắt nhau)
  - Bên phải: Biểu diễn không phẳng (có điểm cắt nhau được đánh dấu đỏ)

2.1.2 Một số dạng đồ thị đặc biệt



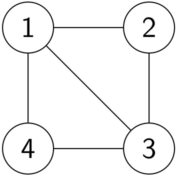
Đồ thị G $G = (V_G, E_G)$



Đồ thị con G_A $G_A = (V_A, E_A)$

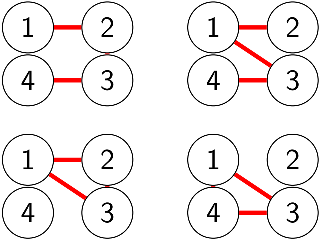
- **Đồ thị con** $G_A = (V_A, E_A)$ của đồ thị $G = (V_G, E_G)$ khi:
  - V_A là tập con của V_G $(V_A \subseteq V_G)$
  - E_A chỉ gồm các cạnh/cung của G mà hai đỉnh nó liên thuộc thuộc tập V_A

2.1.2 Một số dạng đồ thị đặc biệt



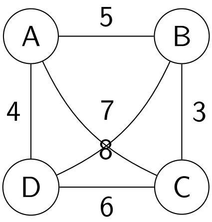
Đồ thị G $\quad G = (V_G, E_G)$

Các đồ thị bộ phận $G_1 = (V_G, E_1)$



- **Đồ thị bộ phận** $G_1 = (V_G, E_1)$ của đồ thị $G = (V_G, E_G)$ khi:
  - Giữ nguyên tất cả các đỉnh của G
  - E_1 là tập con của E_G $(E_1 \subseteq E_G)$

Đồ thị có trọng số

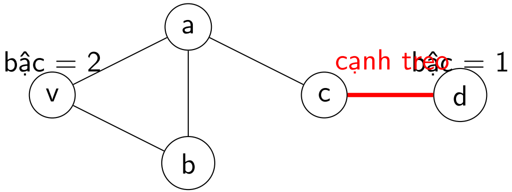


Cạnh	Trọng số
(A,B)	5
(B,C)	3
(C,D)	6
(D,A)	4
(A,C)	8
(B,D)	7

- **Đồ thị có trọng số:** là đồ thị mà mỗi cạnh (u, v) có một giá trị $c(u, v)$ gọi là trọng số của cạnh
- Trọng số có thể biểu diễn:
  - Khoảng cách giữa các đỉnh
  - Chi phí di chuyển
  - Dung lượng đường truyền
  - ...

2.1.3 Bậc của đỉnh đồ thị

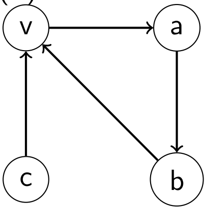
- **Bậc của đỉnh:** Trong đồ thị vô hướng (hoặc có hướng), bậc của đỉnh v là số cạnh liên thuộc với đỉnh v.
- **Đỉnh treo và đỉnh cô lập:**
  - Nếu bậc của đỉnh bằng 1, đỉnh được gọi là đỉnh treo
  - Nếu bậc của đỉnh bằng 0, đỉnh được gọi là đỉnh cô lập
- **Cạnh (cung) treo:** là cạnh (cung) có ít nhất một đầu là đỉnh treo



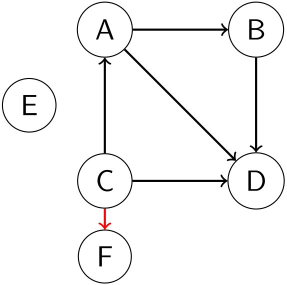
2.1.3 Bậc của đỉnh đồ thị (tiếp)

- **Nửa bậc trong:** của đỉnh v, ký hiệu $d^-(v)$ là số cung có v là đỉnh cuối
- **Nửa bậc ngoài:** của đỉnh v, ký hiệu $d^+(v)$ là số cung có v là đỉnh đầu
- Trong đồ thị có hướng, bậc của đỉnh v, ký hiệu d(v), bằng $d^-(v) + d^+(v)$

$$d(v) = 3$$
$$d^+(v) = 1$$
$$d^-(v) = 2$$



2.1.3 Bậc của đỉnh đồ thị (ví dụ)



Đỉnh	Nửa bậc trong	Nửa bậc ngoài	Bậc
A	1	2	3
B	1	1	2
C	0	3	3
D	3	0	3
E	0	0	0
F	1	0	1

- Đỉnh E là đỉnh cô lập
- Đỉnh F là đỉnh treo
- Cung (C,F) là cung treo (tô đỏ)

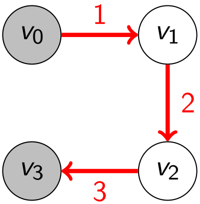
Một số tính chất về bậc của đồ thị

- 1. Tổng số bậc của tất cả các đỉnh gấp đôi số cạnh.
- 2. Số đỉnh có bậc lẻ luôn là một số chẵn.
- 3. Nếu đồ thị có nhiều hơn hai đỉnh thì có ít nhất hai đỉnh cùng bậc.
- 4. Nếu một đồ thị với n đỉnh (n>2) có đúng hai đỉnh cùng bậc thì hai đỉnh này không thể có bậc 0 hoặc n-1.
- 5. Luôn tồn tại đồ thị n đỉnh (n>2) mà 3 đỉnh bất kỳ của đồ thị đều không cùng bậc.
- 6. Cho đồ thị G=(V,E) với ít nhất kn+1 đỉnh, mỗi đỉnh có bậc không bé hơn (k-1)n+1, luôn tồn tại đồ thị con đầy đủ của G gồm k+1 đỉnh.

2.1.4 Đường đi và chu trình

Đường đi: dãy $v_0, v_1, ..., v_n$ $(v_i \in V, i = 0, 1, ..., n)$ được gọi là đường đi từ v_0 đến v_n nếu:

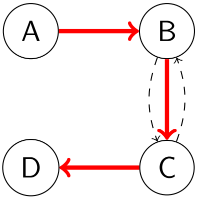
- $\forall i (1 \leq i \geq n-1)$ cặp đỉnh v_i và v_{i+1} kề nhau
- $(v_i, v_{i+1}) \in E$
- v_0 là đỉnh bắt đầu
- v_n là đỉnh kết thúc
- Độ dài đường đi là số cạnh trong dãy



- Đường đi từ v_0 đến v_3 có độ dài 3
- Các cạnh và thứ tự được đánh số từ 1 đến 3

2.1.4 Đường đi và chu trình (tiếp)

Đường đi sơ cấp: là đường đi mà các đỉnh trong đường đi không bị lặp lại.



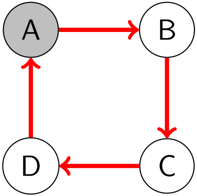
Ví dụ:

- Đường đi sơ cấp: A → B → C → D
- Không phải đường đi sơ cấp: A → B → C → B → D
  - Đường màu đỏ: đường đi sơ cấp
  - Đường nét đứt: ví dụ về đường đi không sơ cấp (có lặp đỉnh B)

2.1.4 Đường đi và chu trình (tiếp)

Chu trình: là đường đi có đỉnh bắt đầu và đỉnh kết thúc trùng nhau.

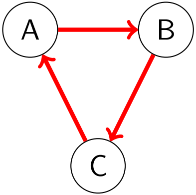
Ví dụ: A $\rightarrow$ B $\rightarrow$ C $\rightarrow$ D $\rightarrow$ A là một chu trình



- Đỉnh A vừa là đỉnh bắt đầu vừa là đỉnh kết thúc
- Độ dài chu trình là 4 (số cạnh trong chu trình)

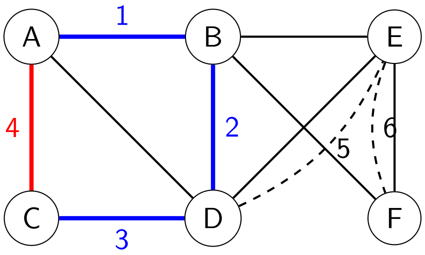
2.1.4 Đường đi và chu trình (tiếp)

Chu trình sơ cấp: là chu trình có các đỉnh không bị lặp lại, trừ đỉnh đầu và đỉnh cuối.



- Chu trình sơ cấp: A $\rightarrow$ B $\rightarrow$ C $\rightarrow$ A
- Không phải chu trình sơ cấp: A $\rightarrow$ B $\rightarrow$ C $\rightarrow$ B $\rightarrow$ A
  - Tất cả các đỉnh chỉ xuất hiện một lần (trừ A)
  - Chu trình có độ dài tối thiểu 3

2.1.4 Đường đi và chu trình (Ví dụ tổng hợp)



- **Đường đi:**
  A — B — D — C
  (cạnh 1,2,3)
- **Đường đi sơ cấp:**
  A — B — D — C
- **Chu trình:**
  A — B — D — C — A (thêm cạnh 4)
- **Đường đi không sơ cấp:**
  D — E — F — E
  (qua E hai lần)

- Màu xanh: đường đi cơ bản độ dài 3 (A — B — D — C)
- Màu đỏ: cạnh tạo chu trình độ dài 4 (A — B — D — C — A)
- Nét đứt: đường đi không sơ cấp qua cạnh 5,6

Một số tính chất

1. Giả sử G là đồ thị vô hướng với n đỉnh (n>2) và các đỉnh đều có bậc không nhỏ hơn 2. Khi đó, G chứa ít nhất một chu trình sơ cấp.
2. Giả sử G là đồ thị vô hướng với n đỉnh (n>3) và các đỉnh đều có bậc không nhỏ hơn 3. Khi đó, G chứa ít nhất một chu trình sơ cấp có độ dài chẵn.

Nội dung

2.2 BIỂU DIỄN MẠNG XÃ HỘI BẰNG ĐỒ THỊ

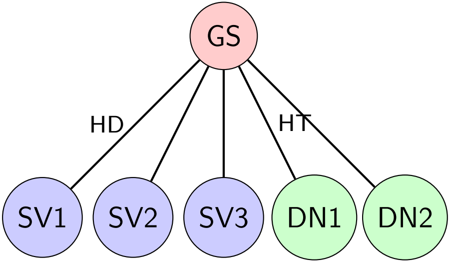
- Mạng xã hội xác định một tập hữu hạn các người dùng (actor) và các liên kết giữa chúng.
- **Actor** có thể là:
  - Cá nhân, tổ chức, công ty
  - Thành viên trong một cộng đồng
  - Cán bộ trong công ty
  - Nhà nghiên cứu trong tổ chức
- **Liên kết** giữa các actor thể hiện mối quan hệ:
  - Quan hệ kinh doanh: mua bán, hợp đồng, đối tác
  - Quan hệ tổ chức: cấp trên - cấp dưới, đồng nghiệp
  - Quan hệ xã hội: bạn bè, gia đình, cộng đồng

2.2 BIỂU DIỄN MẠNG XÃ HỘI BẰNG ĐỒ THỊ (tiếp)

- **Ví dụ về mối quan hệ trong mạng xã hội:**
  - Trong công ty:
    - Giữa nhân viên và cấp trên
    - Giữa các phòng ban
    - Giữa các chi nhánh
  - Trong giáo dục và nghiên cứu:
    - Giữa giảng viên và sinh viên
    - Giữa các nhà nghiên cứu
    - Giữa các tổ chức nghiên cứu
  - Trong kinh doanh:
    - Giữa công ty và đối tác
    - Giữa nhà cung cấp và khách hàng
    - Giữa các doanh nghiệp cùng ngành

2.2 BIỂU DIỄN MẠNG XÃ HỘI BẰNG ĐỒ THỊ (tiếp)

- Đồ thị G = (V, E) biểu diễn mạng xã hội:
  - V: tập các đỉnh (actor)
  - E: tập các cung (liên kết)

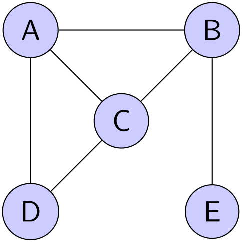


HD: Hướng dẫn HT: Hợp tác

Biểu diễn đồ thị bằng ma trận kề

- Ma trận kề của đồ thị vô hướng là ma trận đối xứng. Mỗi phần tử của ma trận kề phản ánh một cung giữa hai actor và được ký hiệu như sau:

$$x_{ij} = \begin{cases} 1 & \text{Khi có liên kết giữa đỉnh } x_i \text{ và đỉnh } x_j \\ 0 & \text{khi không có liên kết giữa đỉnh } x_i \text{ và đỉnh } x_j \end{cases}$$



Hình 2.14. Mạng gồm 5 đỉnh

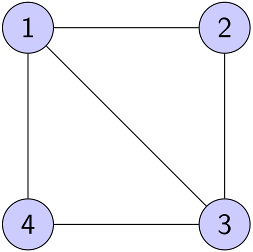
Ma trận kề A:

$$A = \begin{pmatrix} 0 & 1 & 1 & 1 & 0 \\ 1 & 0 & 1 & 0 & 1 \\ 1 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \end{pmatrix}$$

Bài tập: Xác định ma trận kề

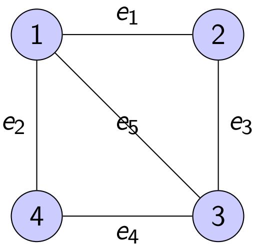
Bài tập: Cho đồ thị vô hướng G như hình vẽ. Hãy:

1. Xác định ma trận kề
2. Cho biết bậc của mỗi đỉnh



Bài tập: Xác định ma trận kề (Lời giải)

Đồ thị G:



Lời giải:

1. Ma trận kề A:

$$A = \begin{pmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \end{pmatrix}$$

2. Bậc của các đỉnh:
   - $d(1) = 3$ (nối với 2,3,4)
   - $d(2) = 2$ (nối với 1,3)
   - $d(3) = 3$ (nối với 1,2,4)
   - $d(4) = 2$ (nối với 1,3)

Nội dung

2.3.1 Mật độ của mạng

- **Mật độ mạng (density)** là một trong những số đo quan trọng của mạng xã hội.
- Khi hệ số gắn kết của mạng càng lớn:
  - Mức độ gắn kết, sự chặt chẽ giữa các actor càng lớn
  - Sự tương trợ, hỗ trợ giữa các actor càng nhiều
  - Ảnh hưởng lên hành vi của actor càng mạnh mẽ
- Đánh giá tổng quát:
  - Mật độ cao $\rightarrow$ tính gắn kết mạnh, thông tin truyền đi tốt
  - Mật độ thấp $\rightarrow$ tính gắn kết yếu, thông tin truyền đi kém

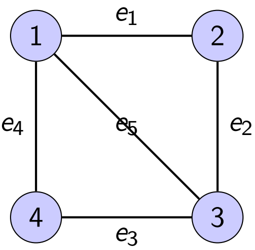
2.3.1 Mật độ của mạng (tiếp)

- Mật độ mạng được tính bằng tỷ lệ giữa tổng các mối liên hệ thực tế và tổng các mối quan hệ có thể có.
- Công thức tính mật độ:

$$\text{Mật độ} = \frac{k}{n(n-1)/2}$$

  Trong đó:
  - k: tổng các đường liên kết thực tế của toàn mạng
  - n: tổng các tác nhân (actor) trong mạng xã hội
  - $\frac{n(n-1)}{2}$: tổng các mối liên kết khả dĩ có trong mạng xã hội
- Giá trị của số đo mật độ mạng nằm trong đoạn [0,1]

Ví dụ: Tính mật độ của mạng xã hội trong hình vẽ.



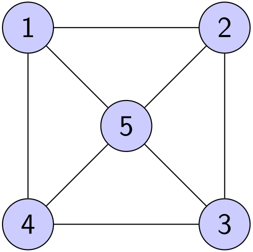
Giải:

- Số đỉnh: $n = 4$
- Số cạnh thực tế: $k =$
- Số cạnh tối đa có thể
  $\frac{n(n-1)}{2} = \frac{4(4-1)}{2} = 6$
- Mật độ mạng:
  Mật độ $= \frac{5}{6} \approx 0.833$

=> Mật độ khá cao (0.83
thấy mạng có tính kết nối

Bài tập: Cho mạng xã hội G như hình vẽ. Hãy:

1. Xác định ma trận kề
2. Tính mật độ của mạng
3. Nhận xét về tính kết nối của mạng



1. Ma trận kề A:

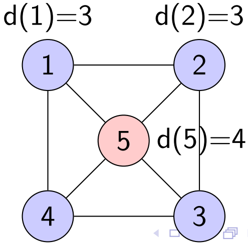
$$A = \begin{pmatrix} 0 & 1 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 & 1 \\ 1 & 1 & 1 & 1 & 0 \end{pmatrix}$$

2. Tính mật độ:

- $n = 5$ đỉnh
- $k = 8$ cạnh
- Số cạnh tối đa: $\frac{n(n-1)}{2} = \frac{5(4)}{2} = 10$
- Mật độ $= \frac{k}{n(n-1)/2} = \frac{8}{10} = 0.8$

3. Nhận xét:

- Mạng có mật độ cao (0.8)
- Mỗi đỉnh đều kết nối với ít nhất 3 đỉnh khác
- Đỉnh 5 kết nối với tất cả các đỉnh còn lại
- $\Rightarrow$ Mạng có tính kết nối mạnh, thông tin truyền đi dễ dàng giữa các actor



2.3.2 Số đo bậc trung tâm (Degree centrality)

- 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 tính:

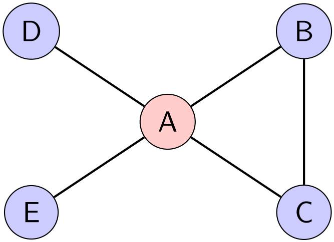
$$C_D(v) = \frac{deg(v)}{n-1}$$

  Trong đó:
  - n: là số đỉnh của đồ thị
  - $deg(v)$: tổng số liên kết trực tiếp đến đỉnh v (bậc của đỉnh)
- Giá trị nằm trong đoạn [0,1]:
  - Giá trị càng gần 1: tính trung tâm càng lớn
  - Dùng để xác định actor quan trọng (key players)

Độ phức tạp:

- $O(V^2)$ nếu đồ thị là dày
- $O(E)$ nếu đồ thị thưa

2.3.2 Số đo bậc trung tâm (Ví dụ)



Tính bậc trung tâm cho các đỉnh:

- $C_D(A) = \frac{4}{4} = 1$
- $C_D(B) = \frac{2}{4} = 0.5$
- $C_D(C) = \frac{2}{4} = 0.5$
- $C_D(D) = \frac{1}{4} = 0.25$
- $C_D(E) = \frac{1}{4} = 0.25$

Nhận xét:

- Đỉnh A có bậc trung tâm cao nhất (1.0)
- $\Rightarrow$ A là actor quan trọng nhất trong mạng

2.3.2 Số đo bậc trung tâm (Lời giải)

1. Ma trận kề:

$$A = \begin{pmatrix} 0 & 1 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 & 1 \\ 1 & 1 & 1 & 1 & 0 \end{pmatrix}$$

2. Số đo bậc trung tâm:

- $C_D(1) = \frac{3}{4} = 0.75$
- $C_D(2) = \frac{3}{4} = 0.75$
- $C_D(3) = \frac{3}{4} = 0.75$
- $C_D(4) = \frac{3}{4} = 0.75$
- $C_D(5) = \frac{4}{4} = 1.00$

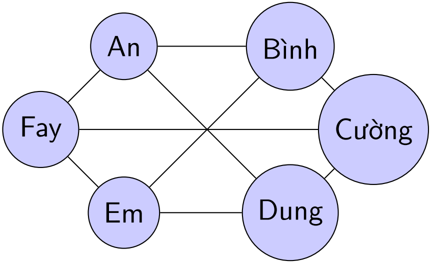
3. Nhận xét:

- Actor 5 có số đo bậc trung tâm cao nhất (1.00)
- Các actor còn lại có số đo bằng nhau (0.75)
- $\Rightarrow$ Actor 5 đóng vai trò quan trọng nhất vì:
  - Kết nối trực tiếp với tất cả các actor khác
  - Có khả năng truyền thông tin tốt nhất
  - Có ảnh hưởng trực tiếp đến toàn bộ mạng

5 **Key player**

2.3.2 Số đo bậc trung tâm (Bài tập thực hành)

Bài tập: Một nhà nghiên cứu thu thập dữ liệu về mối quan hệ giữa 6 sinh viên trong lớp học như hình vẽ bên. Các cạnh thể hiện mối quan hệ "thường xuyên trao đổi bài tập". Hãy:



1. Tính số đo bậc trung tâm cho từng sinh viên
2. Xác định sinh viên nào có vai trò quan trọng nhất trong việc trao đổi bài tập? Tại sao?

Hướng dẫn:

- Xác định bậc của mỗi đỉnh
- Áp dụng công thức $C_D(v) = \frac{deg(v)}{n-1}$ với $n = 6$

2.3.2 Số đo bậc trung tâm (Lời giải)

Bậc của các đỉnh:

- deg(An) = 3
- deg(Bình) = 3
- deg(Cường) = 3
- deg(Dung) = 3
- deg(Em) = 3
- deg(Fay) = 3

Số đo bậc trung tâm:

- $C_D(An) = \frac{3}{5} = 0.6$
- $C_D(Bnh) = \frac{3}{5} = 0.6$
- $C_D(Cng) = \frac{3}{5} = 0.6$
- $C_D(Dung) = \frac{3}{5} = 0.6$
- $C_D(Em) = \frac{3}{5} = 0.6$
- $C_D(Fay) = \frac{3}{5} = 0.6$

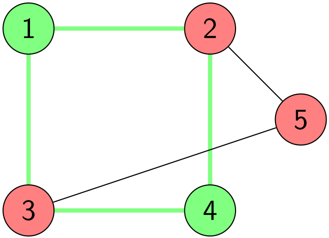
Nhận xét:

- Tất cả sinh viên đều có số đo bậc trung tâm bằng nhau (0.6)
- Mỗi sinh viên đều trao đổi bài tập với 3 sinh viên khác
- Mạng lưới trao đổi có tính đồng đều cao
- $\Rightarrow$ Không có sinh viên nào đóng vai trò trung tâm, việc trao đổi bài tập diễn ra đồng đều giữa các thành viên

Đây là một mạng lưới cân bằng!

2.3.3 Đường đi ngắn nhất

- **Định nghĩa:** Đường đi ngắn nhất giữa hai đỉnh là con đường kết nối hai đỉnh và có chiều dài từ đỉnh này đến đỉnh kia là ngắn nhất.
- **Ý nghĩa:**
  - Tốc độ liên lạc, trao đổi thông tin diễn ra nhanh chóng
  - Xác định các node quan trọng trên đường đi ngắn nhất
  - Giúp xác định điểm gắn kết mạng
- **Ví dụ:** Trong hình vẽ, giữa node 1 và 4 có hai đường đi ngắn nhất độ dài 2:
  - $1 \to 2 \to 4$ hoặc $1 \to 3 \to 4$



Đường đi ngắn nhất

2.3.4 Số đo trung tâm gần gũi (Closeness centrality)

- **Đặc điểm:**
  - Khắc phục điểm yếu của số đo bậc (chỉ xét quan hệ trực tiếp)
  - Xét khả năng tiếp cận với toàn bộ mạng của một actor
  - Actor có thể ít liên kết trực tiếp nhưng vẫn "gần gũi"với mạng
- **Công thức tính:**

$$C_C(v) = \frac{1}{\sum_{t \in V/v} d_G(v,t)}$$

Trong đó: $d_G(v,t)$ là chiều dài đường đi ngắn nhất từ đỉnh v tới đỉnh t

- **Công thức chuẩn hóa:**

$$CC(v) = (n-1)C_C(v)$$

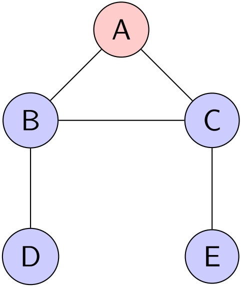
Trong đó n là số đỉnh của đồ thị

2.3.4 Số đo trung tâm gần gũi - Ví dụ tính toán

Ý nghĩa

Số đo này tương ứng với thời gian cần thiết để thông tin truyền từ một actor tới các actor khác. Khoảng cách càng nhỏ, khả năng truyền tin càng lớn.

2.3.4 Số đo trung tâm gần gũi - Ví dụ tính toán



Khoảng cách từ A đến:

- B: $d_G(A, B) = 1$
- C: $d_G(A, C) = 1$
- D: $d_G(A, D) = 2$
- E: $d_G(A, E) = 2$

Tính $C_C(A)$:

1. Tính tổng khoảng cách:

$$\sum_{t \in V/A} d_G(A, t) = 1 + 1 + 2 + 2 = 6$$

2. Tính giá trị chưa chuẩn hóa:

$$C_C(A) = \frac{1}{6} \approx 0.167$$

3. Chuẩn hóa (với n = 5):

$$CC(A) = (n-1)C_C(A) = 4 \times \frac{1}{6} \approx 0.667$$

2.3.4 Số đo trung tâm gần gũi - Ví dụ tính toán

So sánh với các đỉnh khác

- B: $\sum d_G(B, t) = 1 + 1 + 1 + 2 = 5,\ CC(B) = 4 \times \frac{1}{5} = 0.800$
- D: $\sum d_G(D, t) = 2 + 1 + 2 + 3 = 8,\ CC(D) = 4 \times \frac{1}{8} = 0.500$

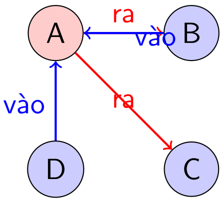
Nhận xét

- Đỉnh B có số đo trung tâm gần gũi cao nhất (0.800) do có vị trí thuận lợi để tiếp cận các đỉnh khác
- A dù có bậc cao (2 liên kết) nhưng có số đo thấp hơn B do cách xa D, E
- D có số đo thấp nhất do phải qua nhiều bước để đến các đỉnh khác

2.3.4 Số đo trung tâm gần gũi (tiếp)

Áp dụng cho đồ thị có hướng:

- **Các cung ra:**
  - Đo khả năng tiếp cận từ đỉnh được chọn đến các đỉnh khác
  - Đo khoảng cách từ đỉnh nguồn đến tất cả đỉnh đích
- **Các cung vào:**
  - Đo khả năng tiếp cận từ các đỉnh khác đến đỉnh được chọn
  - Đo khoảng cách từ tất cả đỉnh đến đỉnh đích



2.3.4 Số đo trung tâm gần gũi (Closeness centrality)

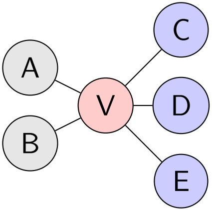
Ví dụ minh họa

- Độ gần gũi cung ra của A: xét các đường đi A→B, A→C
- Độ gần gũi cung vào của A: xét các đường đi B→A, D→A

2.3.5 Số đo trung tâm trung gian (betweenness centrality)

- **Định nghĩa:** Xác định actor đóng vai trò "cầu nối"trong mạng, dù có thể:
  - Không có nhiều kết nối trực tiếp
  - Không "gần gũi"với nhiều thành viên
- **Công thức:**

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

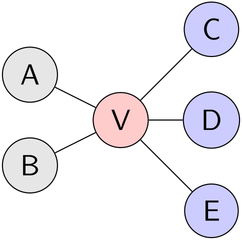


V: node trung gian

Trong đó:

  - σ_{st}: số đường đi ngắn nhất từ s đến t
  - $\sigma_{st}(v)$: số đường đi ngắn nhất từ s đến t qua v
- **Công thức chuẩn hóa:**
  - Đồ thị vô hướng: $C'_B(v) = \frac{C_B(v)}{(n-1)(n-2)/2}$
  - Đồ thị có hướng: $C'_B(v) = \frac{C_B(v)}{(n-1)(n-2)}$

2.3.5 Số đo trung tâm trung gian - Ví dụ tính toán



Thông tin đồ thị:

- Tổng số đỉnh: $n = 6$
- Đồ thị vô hướng

Tính $C_B(V)$:

1. Liệt kê các cặp đỉnh qua V:
   - A $\rightarrow$ C: 1 đường qua V / 1 tổng = 1
   - A $\rightarrow$ D: $1/1 = 1$
   - A $\rightarrow$ E: $1/1 = 1$
   - B $\rightarrow$ C: $1/1 = 1$
   - B $\rightarrow$ D: $1/1 = 1$
   - B $\rightarrow$ E: $1/1 = 1$
2. $C_B(V) = \sum \frac{\sigma_{st}(v)}{\sigma_{st}} = 6$

2.3.5 Số đo trung tâm trung gian (betweenness centrality)

Chuẩn hóa kết quả

- Hệ số chuẩn hóa (vô hướng): $(n-1)(n-2)/2 = (6-1)(6-2)/2 = 10$
- $C'_B(V) = \frac{C_B(V)}{(n-1)(n-2)/2} = \frac{6}{10} = 0.6$

Nhận xét

- V có hệ số trung tâm trung gian cao (0.6) do nằm trên tất cả các đường đi ngắn nhất giữa nhóm trái (A,B) và nhóm phải (C,D,E)
- V đóng vai trò cầu nối quan trọng trong mạng

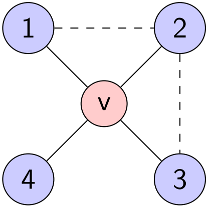
Ý nghĩa

Số đo này càng lớn thì actor càng quan trọng trong việc kiểm soát thông tin và giao dịch trong mạng.

- Trong mạng xã hội, số đo gom cụm được Watts và Strogatz đề xuất làm tiêu chuẩn đo mức độ gắn kết giữa các actor trong mạng.
- Số đo gom cụm của một actor được xác định bởi các actor láng giềng có mối liên kết trực tiếp với nhau.

Giải thích công thức:

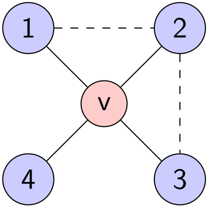
- C_i: Hệ số gom cụm của đỉnh i
- $|[e_{jk}]|$: Số cạnh thực tế giữa các láng giềng của đỉnh i
- k_i: Bậc của đỉnh i (số láng giềng)
- $k_i(k_i - 1)$: Số cạnh tối đa có thể có giữa các láng giềng



- Trong mạng xã hội, số đo gom cụm được Watts và Strogatz đề xuất làm tiêu chuẩn đo mức độ gắn kết giữa các actor trong mạng.
- Số đo gom cụm của một actor được xác định bởi các actor láng giềng có mối liên kết trực tiếp với nhau.

Giải thích công thức:

- C_i: Hệ số gom cụm của đỉnh i
- $|[e_{jk}]|$: Số cạnh thực tế giữa các láng giềng của đỉnh i
- k_i: Bậc của đỉnh i (số láng giềng)
- $k_i(k_i - 1)$: Số cạnh tối đa có thể có giữa các láng giềng



2.3.6 Số đo gom cụm (clustering centrality) - Phần 2

Công thức tính:

1. Đồ thị có hướng:

$$C_i = \frac{|[e_{jk}]|}{k_i(k_i - 1)}$$

   - Mẫu số là số cạnh tối đa có thể có trong đồ thị có hướng
   - Mỗi cặp đỉnh chỉ tính một chiều

2. Đồ thị vô hướng:

$$C_i = \frac{2|[e_{jk}]|}{k_i(k_i - 1)}$$

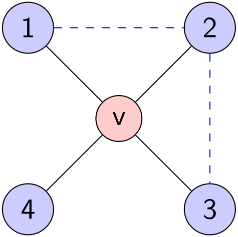
   - Nhân 2 ở tử số vì mỗi cạnh được tính hai lần
   - Mẫu số vẫn giữ nguyên do tính tổng số cặp đỉnh có thể

3. Trung bình toàn mạng:

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

   - n: Số đỉnh trong mạng

2.3.6 Số đo gom cụm (clustering centrality) - Ví dụ



Đồ thị vô hướng

Phân tích:

- Đỉnh v có 4 láng giềng: $k_v = 4$
- Số cạnh tối đa có thể có giữa các láng giềng:

$$k_v(k_v - 1) = 4(4 - 1) = 12$$

- Số cạnh thực tế giữa các láng giềng:

$$|[e_{jk}]| = 2 \text{ (cạnh 1-2 và 2-3)}$$

Tính toán:

$$C_v = \frac{2|[e_{jk}]|}{k_v(k_v - 1)} = \frac{2 \times 2}{12} = \frac{1}{3} \approx 0.333$$

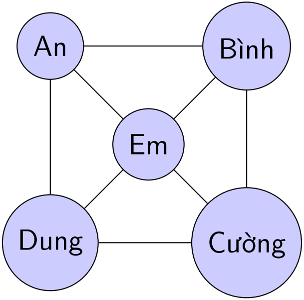
2.3.6 Số đo gom cụm (clustering centrality) - Ví dụ

Nhận xét

- Hệ số gom cụm $\frac{1}{3}$ cho thấy mức độ kết nối trung bình giữa các láng giềng
- Có thể tăng hệ số này bằng cách thêm các cạnh giữa các láng giềng (ví dụ: 1-4, 3-4)

Bài tập 1: Phân tích mạng học tập

Tình huống: Một nghiên cứu về mối quan hệ học tập giữa 5 sinh viên trong một nhóm thực hành. Mỗi cạnh thể hiện việc "thường xuyên trao đổi bài tập".

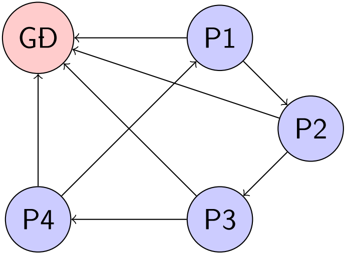


Yêu cầu:

1. Tính mật độ mạng
2. Xác định:
   - Số đo bậc trung tâm
   - Số đo trung tâm gần gũi
   - Số đo trung tâm trung gian
3. Tính số đo gom cụm cho mỗi sinh viên
4. Nhận xét vai trò của Em trong nhóm

Bài tập 2: Phân tích luồng thông tin trong tổ chức

Tình huống: Sơ đồ luồng thông tin giữa các phòng ban trong một công ty. Mũi tên chỉ hướng báo cáo/trao đổi thông tin.

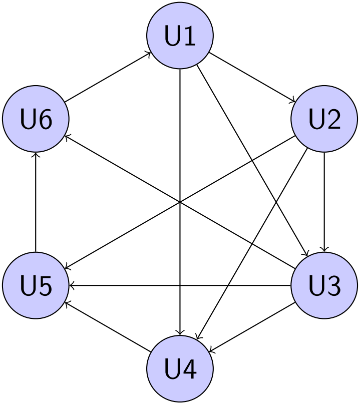


Yêu cầu:

1. Tính mật độ mạng
2. Xác định:
   - Bậc vào và bậc ra của mỗi phòng ban
   - Số đo trung tâm gần gũi (cung vào/ra)
3. Tính hiệu quả truyền thông tin trong tổ chức
4. Đề xuất cải thiện luồng thông tin

Bài tập 3: Phân tích mạng xã hội trực tuyến

Tình huống: Một nhóm 6 người tham gia diễn đàn trực tuyến. Mũi tên thể hiện người A theo dõi/tương tác với người B.



Yêu cầu:

1. Tính mật độ mạng
2. Xác định:
   - Người có ảnh hưởng nhất (bậc ra cao nhất)
   - Người được quan tâm nhất (bậc vào cao nhất)
3. Tính các số đo trung tâm
4. Phân tích vai trò "người kết nối"
5. Đề xuất cách tăng tương tác trong nhóm

Chúc các bạn học thật tốt!