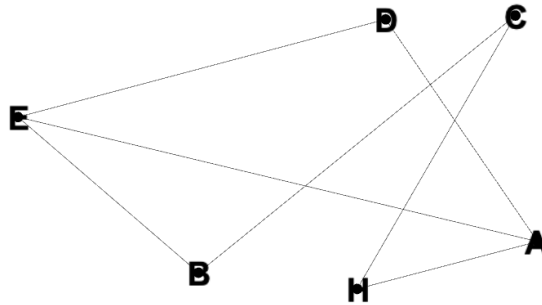


Môn thi: *Mạng xã hội*
Thời gian làm bài: *90 phút*
Không được sử dụng tài liệu

Câu 1: (4đ) Tính các độ đo *Closeness Centrality*, *Harmonic Closeness Centrality* và *Betweenness Centrality* cho tất cả các đỉnh của đồ thị:



Hướng dẫn trình bày dưới dạng:

– *Closeness Centrality*

$$\circ c_C(A) = \frac{1}{1+2+3+\dots} = \frac{1}{\dots}$$

– *Harmonic Closeness Centrality*

$$\circ c_H(A) = \frac{1}{1} + \frac{1}{2} + \frac{1}{3} + \dots = \dots$$

– *Betweenness Centrality*

Đỉnh bắt đầu	Đỉnh đích	Đỉnh trung gian	Đường đi
A	B	(...) or (...nếu các đường đi khác nhau)	

$$\circ c_B(A) = \frac{1}{(\text{Tổng số đường đi})}$$

Câu 2: (3đ) Trình bày chi tiết thuật toán PageRank xếp hạng đỉnh của đồ thị (hệ số xác suất =0.85) cho đồ thị có ma trận kề như sau, cho biết kết quả xếp hạng các đỉnh?

	A	B	C	D	E	H
A	0	0	1	0	0	1
B	0	0	0	1	0	0
C	1	0	0	1	1	1
D	0	1	1	0	1	0
E	0	0	1	1	0	0
H	1	0	1	0	0	0

Hướng dẫn trình bày:

Lần lặp	A	B	C	D	E	H
0	1	1	1	1	1	1
1
2
...

Trong đó sinh viên chỉ cần diễn giải công thức ở lần lặp 1:

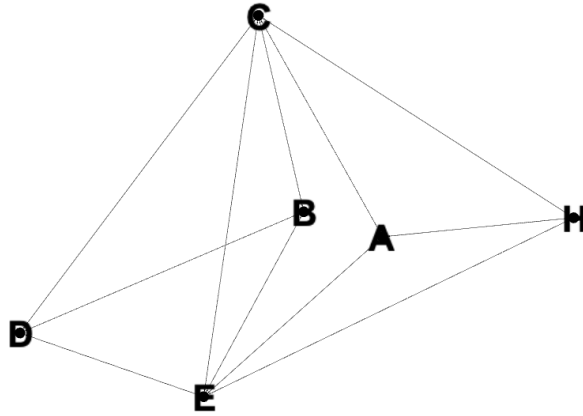
$PR(A_1)$ = công thức tính PageRank cho đỉnh A ở lần lặp 1

(Tương tự cho các đỉnh còn lại)

Câu 3: (3đ) Sinh viên chỉ chọn một trong hai phần (Phần A hoặc Phần B):

Phần A:

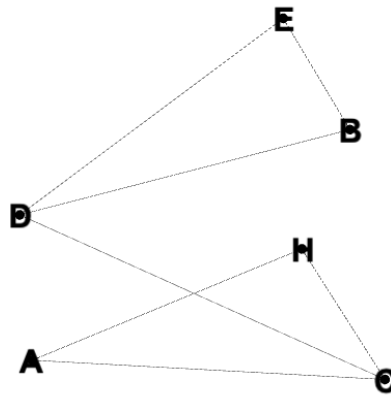
Tính hệ số cụm (Clustering Coefficient) của mỗi đỉnh đồ thị sau, sử dụng thuật toán K-Means gom cụm các đỉnh dựa trên giá trị các hệ số cụm này.



Phần B

Sử dụng thuật toán Girvan Newman cho bài toán phân cụm đồ thị sau. Kết quả phân cụm?

Lưu ý: sinh viên tính độ đo Edge Betweenness dựa trên duyệt cây BFS hoặc công thức gốc (kẻ bảng) đều được.



TRƯỞNG KHOA

GIẢNG VIÊN RA ĐỀ

Ghi chú: cán bộ coi thi không giải thích gì thêm.

Câu 1 đáp ứng chuẩn đầu ra G1, G2 của đề cương môn học.

Câu 2, 3 đáp ứng chuẩn đầu ra G2 của đề cương môn học.