

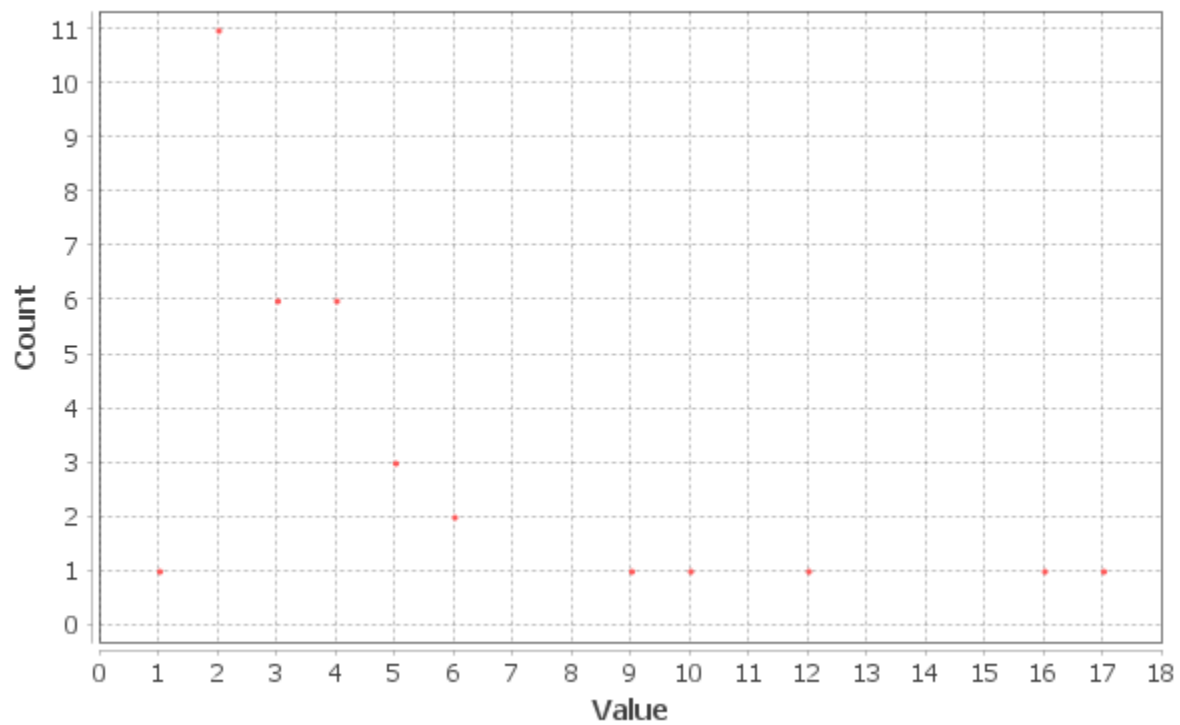
Thực Hành Lab05
Bài số 2


Yêu cầu thực hiện

Phần 1: Phân tích cấu trúc mạng Tính toán và so sánh các độ đo tính trung tâm (Centrality Measures) sau:

- Degree Centrality
- Betweenness Centrality
- Closeness Centrality Hãy xác định 3 nút có độ trung tâm cao nhất theo mỗi độ đo và giải thích ý nghĩa của chúng trong mạng lưới.

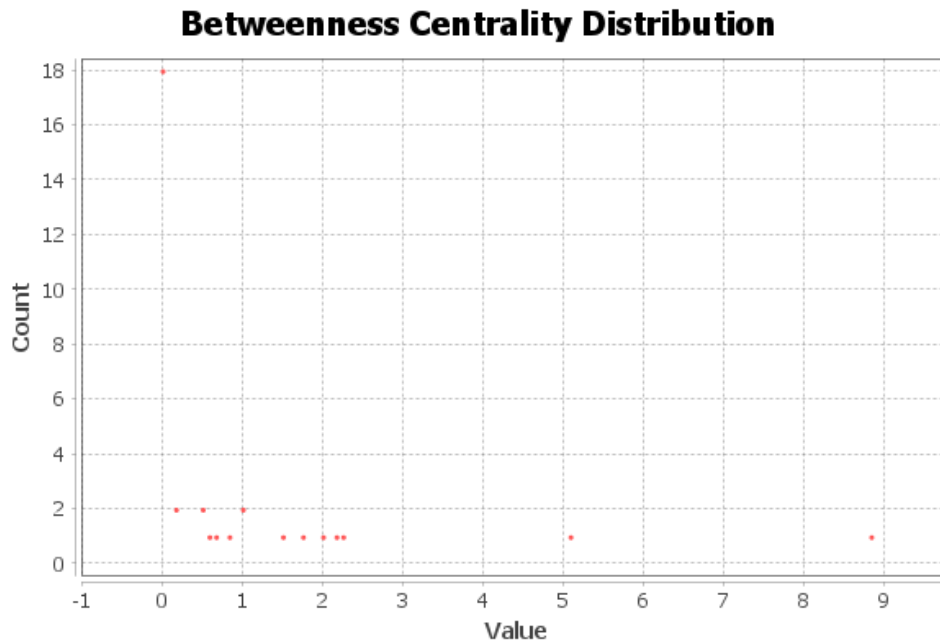
Degree Distribution



Id	Label	Interval	In-Degree	Out-Degree	Degree 
34	34		0	17	17
1	1		16	0	16
33	33		1	11	12
3	3		8	2	10
2	2		8	1	9
4	4		3	3	6
32	32		2	4	6
24	24		5	0	5
9	9		3	2	5
14	14		1	4	5
6	6		3	1	4
30	30		2	2	4
31	31		2	2	4
7	7		1	3	4
28	28		1	3	4
8	8		0	4	4
25	25		3	0	3
5	5		2	1	3
29	29		2	1	3
20	20		1	2	3
26	26		1	2	3
11	11		0	3	3
15	15		2	0	2
16	16		2	0	2

Nút có id là 34 có bậc cao nhất vì nó có chỉ số DC cao nhất, bậc của nó là 17

Betweenness Centrality



Id	Label	Interval	In-Degree	Out-Degree	Degree	Weighted In-Degree	Weighted Out-Degree	Weighted Degree	Eccentricity	Closeness Centrality	Harmonic Closeness Centrality	Betweenness Centrality
3	3		8	2	10	8.0	2.0	10.0	1.0	1.0	1.0	8.833333
32	32		2	4	6	2.0	4.0	6.0	3.0	0.636364	0.761905	5.083333
9	9		3	2	5	3.0	2.0	5.0	2.0	0.75	0.833333	2.25
29	29		2	1	3	2.0	1.0	3.0	2.0	0.6	0.666667	2.166667
4	4		3	3	6	3.0	3.0	6.0	1.0	1.0	1.0	2.0
14	14		1	4	5	1.0	4.0	5.0	1.0	1.0	1.0	1.75
7	7		1	3	4	1.0	3.0	4.0	1.0	1.0	1.0	1.5
26	26		1	2	3	1.0	2.0	3.0	1.0	1.0	1.0	1.0
30	30		2	2	4	2.0	2.0	4.0	1.0	1.0	1.0	1.0
31	31		2	2	4	2.0	2.0	4.0	2.0	0.666667	0.75	0.833333
28	28		1	3	4	1.0	3.0	4.0	2.0	0.714286	0.8	0.666667
20	20		1	2	3	1.0	2.0	3.0	1.0	1.0	1.0	0.583333
2	2		8	1	9	8.0	1.0	9.0	1.0	1.0	1.0	0.5
6	6		3	1	4	3.0	1.0	4.0	1.0	1.0	1.0	0.5
10	10		1	1	2	1.0	1.0	2.0	2.0	0.6	0.666667	0.166667
33	33		1	11	12	1.0	11.0	12.0	2.0	0.73913	0.823529	0.166667
1	1		16	0	16	16.0	0.0	16.0	0.0	0.0	0.0	0.0
5	5		2	1	3	2.0	1.0	3.0	1.0	1.0	1.0	0.0
8	8		0	4	4	0.0	4.0	4.0	1.0	1.0	1.0	0.0
11	11		0	3	3	0.0	3.0	3.0	1.0	1.0	1.0	0.0
12	12		0	1	1	0.0	1.0	1.0	1.0	1.0	1.0	0.0
13	13		0	2	2	0.0	2.0	2.0	2.0	0.666667	0.75	0.0
15	15		2	0	2	2.0	0.0	2.0	0.0	0.0	0.0	0.0
16	16		2	0	2	2.0	0.0	2.0	0.0	0.0	0.0	0.0

1. Đa số các nút có Betweenness Centrality bằng 0:

- Nhiều nút trong mạng có giá trị Betweenness Centrality bằng 0. Điều này cho thấy các nút này không đóng vai trò trung gian trong bất kỳ đường đi ngắn nhất nào giữa các cặp nút khác. Chúng có thể là các nút nằm ở rìa mạng hoặc không phải là điểm kết nối quan trọng.

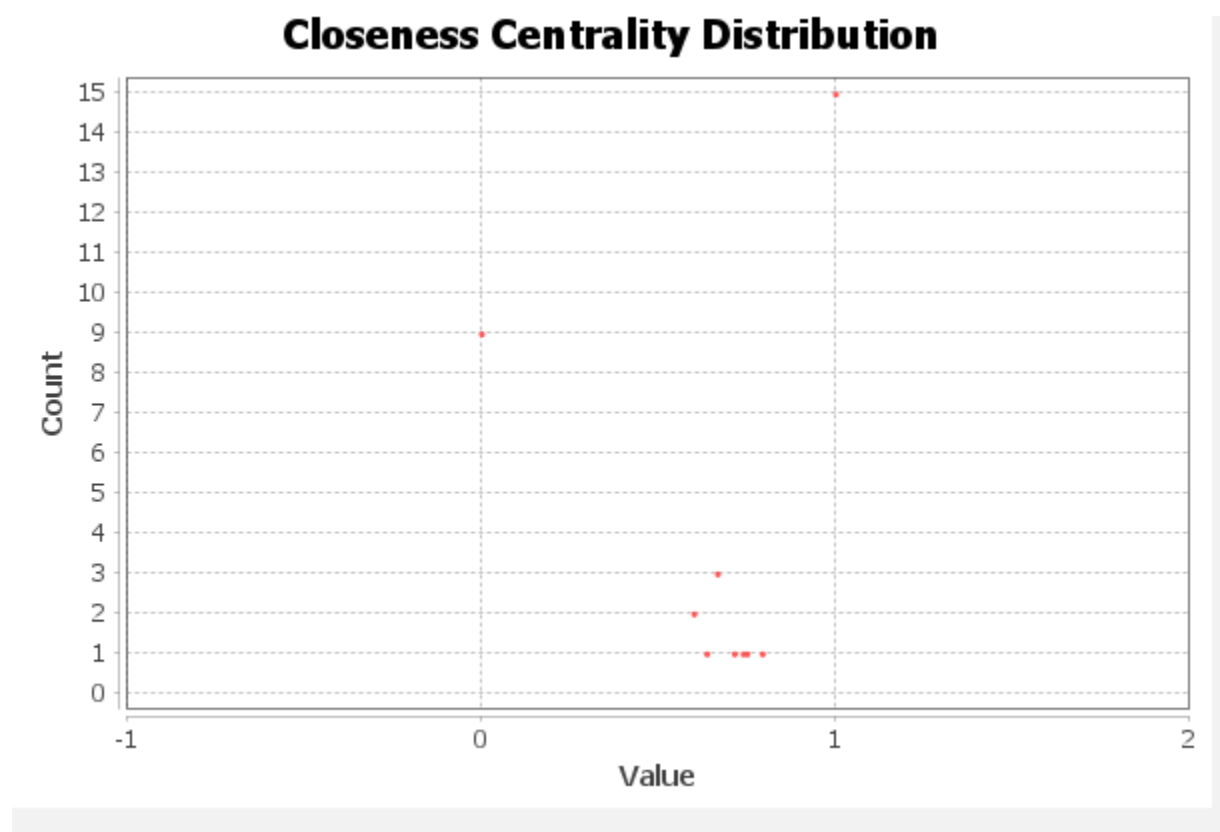
2. Một số ít nút có giá trị Betweenness Centrality cao hơn đáng kể:

- Các nút như **Id = 3**, **Id = 32**, và **Id = 9** có giá trị Betweenness Centrality cao nhất (lần lượt là 8.8333, 5.0833, và 2.25). Điều này chỉ ra rằng chúng đóng vai trò trung gian quan trọng trong mạng, ảnh hưởng đến luồng thông tin hoặc kết nối giữa các phần khác nhau của mạng.

3. Phân phối không đồng đều:

- Phân phối Betweenness Centrality có dạng bất đối xứng, với một số ít nút tập trung giá trị cao trong khi phần lớn các nút có giá trị thấp hoặc bằng 0. Đây là đặc điểm thường thấy trong các mạng có tính chất phân cấp hoặc mạng có một vài điểm nút quan trọng.

Closeness Centrality



Id	Label	Interval	In-Degree	Out-Degree	Degree	Weighted In-Deg...	Weighted Out-Deg...	Weighted Deg...	Eccentricity	Closeness Cen...
2	2		8	1	9	8.0	1.0	9.0	1.0	1.0
3	3		8	2	10	8.0	2.0	10.0	1.0	1.0
4	4		3	3	6	3.0	3.0	6.0	1.0	1.0

Closeness Centrality trong bảng đều có giá trị bằng 1. Điều này cho thấy rằng tất cả các nút (nodes) trong mạng được tính toán đều nằm trong khoảng cách gần nhau, hoặc mạng lưới này rất nhỏ và liên thông mạnh mẽ. Một số nhận xét cụ thể:

1. **Giá trị Closeness Centrality = 1:** Nếu tất cả các nút đều có giá trị bằng 1, có thể điều này xảy ra do đồ thị được chuẩn hóa hoặc tính toán trên mạng nhỏ, nơi mà mỗi nút có thể tiếp cận trực tiếp hoặc trong khoảng cách tối thiểu đến các nút khác.
2. **Tầm quan trọng của Closeness Centrality:** Trong các mạng lớn hơn hoặc phức tạp hơn, chỉ số này thường giúp xác định mức độ "gần gũi" của một nút với toàn bộ mạng. Tuy nhiên, ở đây các giá trị giống nhau không cung cấp sự khác biệt về tầm quan trọng giữa các nút.

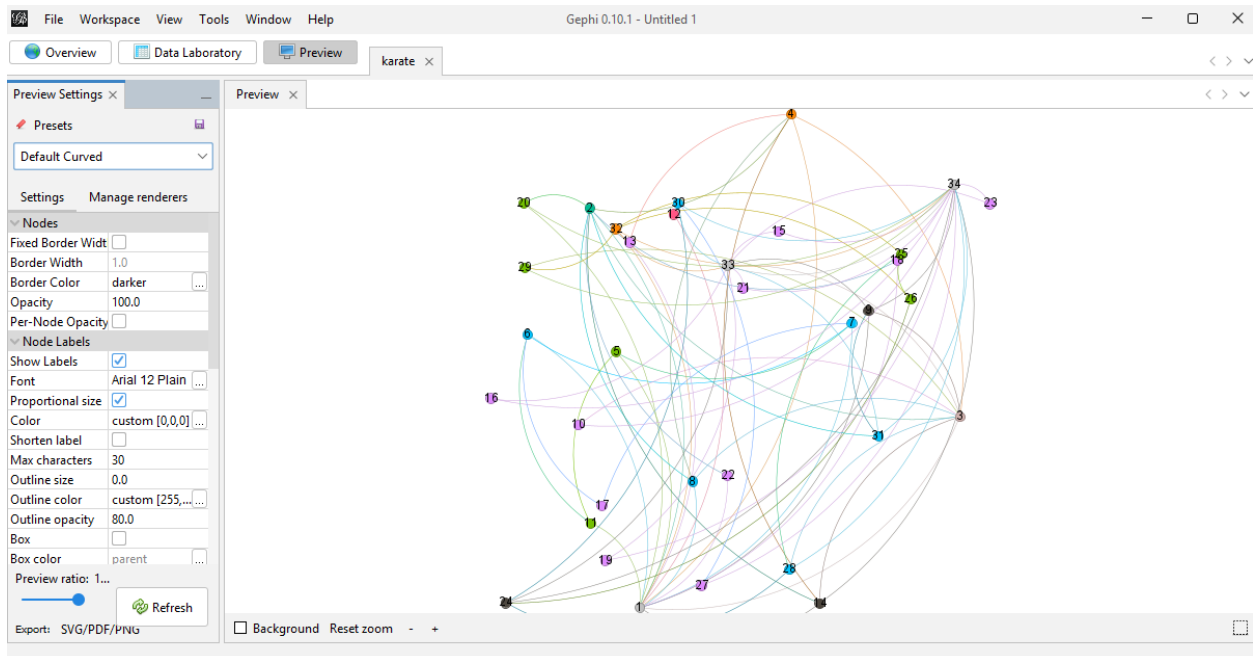
Phần 2: Phát hiện cộng đồng Thực hiện phân cụm mạng lưới sử dụng 3 thuật toán sau:

1. Thuật toán Louvain
2. Thuật toán Girvan-newman (gephi.org/plugins/#/plugin/girvan-newman-clustering hoặc gephi.org/plugins/#/plugin/newman-girvan-plugin)
3. Thuật toán LPA (gephi.org/plugins/#/plugin/label-propagation-clustering)

Với mỗi thuật toán, hãy:

- Ghi lại số lượng cộng đồng được phát hiện
- Tính toán độ đo Modularity của kết quả phân cụm
- Lưu ảnh kết quả phân cụm với các node được tô màu theo cộng đồng

1. Louvain



1. Cấu trúc của Mạng:

- Biểu đồ hiển thị một mạng xã hội với các nút (nodes) và các cạnh (edges).
- Các nút đại diện cho các cá nhân hoặc thực thể, trong khi các cạnh đại diện cho các mối quan hệ hoặc kết nối giữa các thực thể đó.

2. Phân Cụm Mạng:

- Mạng được chia thành nhiều cụm khác nhau, có thể thấy qua các màu sắc khác nhau của các nút và cạnh.
- Các cụm này có thể đại diện cho các nhóm hoặc cộng đồng trong mạng xã hội, nơi mà các nút bên trong cụm có nhiều liên kết hơn so với các nút bên ngoài cụm.

3. Độ Kết Nối (Connectivity):

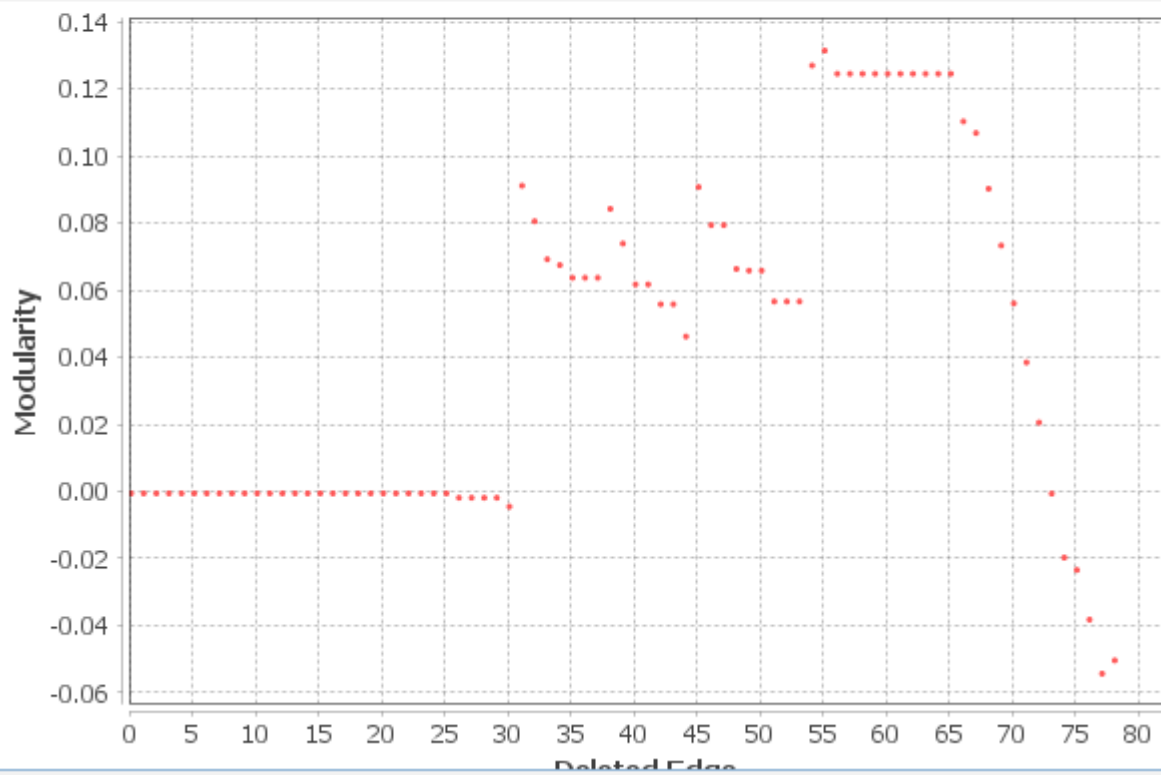
- Một số nút có nhiều cạnh kết nối hơn, cho thấy họ có vai trò quan trọng hoặc là các nút trung tâm trong mạng.

Girvan-Newman

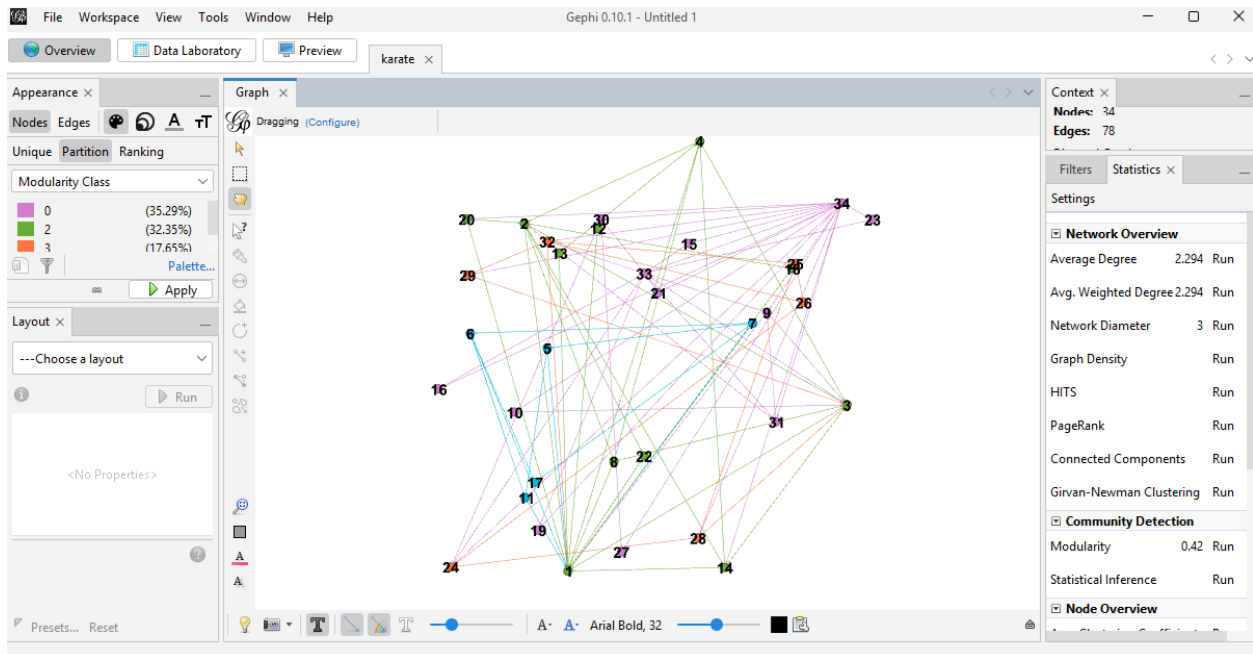
Communities

Number of communities: 20

Maximum found modularity: 0.13198553



LPA



• Cấu trúc Biểu đồ:

- Biểu đồ mạng hiển thị 34 nút (nodes) và 78 cạnh (edges).
- Các nút được tô màu theo các lớp modularity (Modularity Class), thể hiện các cộng đồng phát hiện được trong mạng.

• Phân Cụm Cộng Đồng:

- Mạng được phân chia thành các cộng đồng khác nhau, thể hiện qua các màu sắc khác nhau của các nút.
- Mỗi cộng đồng được gán một nhãn (label) từ 0 đến 3, với tỷ lệ phần trăm nút trong mỗi cộng đồng:
 - Lớp 0: 35.29%
 - Lớp 2: 32.35%
 - Lớp 3: 17.65%
- Lớp modularity 0 có số lượng nút lớn nhất, tiếp theo là lớp 2 và lớp 3.

• Thống Kê Mạng:

- **Average Degree:** Trung bình có 2.294 cạnh kết nối mỗi nút.
- **Network Diameter:** Đường kính mạng là 3, nghĩa là khoảng cách dài nhất giữa hai nút bất kỳ trong mạng là 3 cạnh.

- **Modularity:** Giá trị modularity là 0.42, cho thấy mạng có phân cụm cộng đồng khá rõ ràng. Modularity là thước đo để đánh giá chất lượng phân chia mạng thành các cộng đồng, giá trị càng cao thì phân cụm càng tốt.

Phần 3:

