

Phần 1:

Id	Label	Degree ▾	Betweenness Centrality	Closeness Centrality
1	1	3553	7850454.57368	0.448688
2	2	1696	2824223.129625	0.447388
3	3	1283	1467088.627146	0.424776
5	5	902	1094386.128834	0.420905
4	4	876	1066551.739696	0.424809
10	10	755	973872.641197	0.435596
8	8	679	634973.095482	0.423677
7	7	673	616516.556674	0.425713
11	11	577	540574.100325	0.447351
177	177	543	281136.303354	0.423976
6	6	510	355618.17523	0.417523
15	15	482	918759.260014	0.432631
9	9	435	348405.453455	0.422848
26	26	390	154791.824412	0.422617
22	22	369	284741.302119	0.435
33	33	365	176176.364564	0.430118
12	12	360	141686.902076	0.404219
95	95	332	266097.827925	0.414863
13	13	328	333687.021037	0.396013
16	16	327	206075.617894	0.409208
30	30	316	119447.821311	0.414227

Degree Centrality:

Top 3 nút theo Degree Centrality:

1. **Id 123:** Degree = 3553
2. **Id 1696:** Degree = 2824223
3. **Id 1283:** Degree = 1467088.627146

Ý nghĩa:

- **Degree Centrality** đo lường số lượng kết nối trực tiếp của mỗi nút trong mạng. Những nút có Degree Centrality cao nhất là những nút có nhiều kết nối nhất.

- **Giải thích:**

- **Id 123:** Là nút có nhiều kết nối nhất, đại diện cho cá nhân hoặc tổ chức có vai trò trung tâm giao tiếp, kết nối nhiều thành phần trong mạng xã hội.
- **Id 1696 và Id 1283:** Là các nút quan trọng khác, cũng có nhiều kết nối nhưng ít hơn Id 123, cho thấy vai trò của họ có thể là trung gian trong một số nhóm hoặc cộng đồng nhỏ hơn.

Betweenness Centrality:

Top 3 nút theo Betweenness Centrality:

1. **Id 123:** Betweenness Centrality = 7850454.57368
2. **Id 1696:** Betweenness Centrality = 2824223.129625
3. **Id 1283:** Betweenness Centrality = 1467088.627146

Ý nghĩa:

- Betweenness Centrality đo lường mức độ mà một nút nằm trên nhiều đường đi ngắn nhất giữa các cặp nút khác.
- **Giải thích:**
 - **Id 123:** Đóng vai trò như một "cầu nối" quan trọng nhất trong mạng xã hội, giúp kết nối các cộng đồng hoặc phần khác nhau của mạng. Nếu nút này bị loại bỏ, mạng có thể bị phân mảnh.
 - **Id 1696 và Id 1283:** Cũng đóng vai trò quan trọng trong việc kết nối các phần của mạng, nhưng mức độ ảnh hưởng thấp hơn so với Id 123.

Closeness Centrality:

Top 3 nút theo Closeness Centrality:

1. **Id 123:** Closeness Centrality = 0.448688
2. **Id 1696:** Closeness Centrality = 0.447388
3. **Id 1283:** Closeness Centrality = 0.424776

Ý nghĩa:

- **Closeness Centrality** đo lường mức độ gần gũi của một nút với tất cả các nút khác trong mạng. Nút có giá trị cao có thể truyền thông tin nhanh chóng đến các nút khác.
- **Giải thích:**

- **Id 123:** Là nút gần gũi nhất với tất cả các nút khác, cho thấy khả năng lan truyền thông tin nhanh chóng trong toàn mạng.
- **Id 1696 và Id 1283:** Cũng có khả năng lan truyền thông tin hiệu quả, nhưng ít hơn Id 123.

Phần 2:

Thuật toán Louvain



Modularity: 0.248
Modularity with resolution: 0.248
Number of Communities: 20

🔍 **Modularity = 0.248:**

- Mạng có cấu trúc cộng đồng trung bình. Mặc dù có sự chia tách các cộng đồng, các mối quan hệ giữa các nút không hoàn toàn bị giới hạn bên trong cộng đồng, mà vẫn có các kết nối giữa các cộng đồng khác nhau.

📌 **Số lượng cộng đồng = 20:**

- Mạng được phân tách thành nhiều nhóm nhỏ. Điều này có thể phản ánh tính chất phức tạp của mạng, nơi nhiều nhóm hoặc cộng đồng tồn tại song song nhưng không có sự tách biệt hoàn toàn

Thuật toán Girvan-newman

LPA

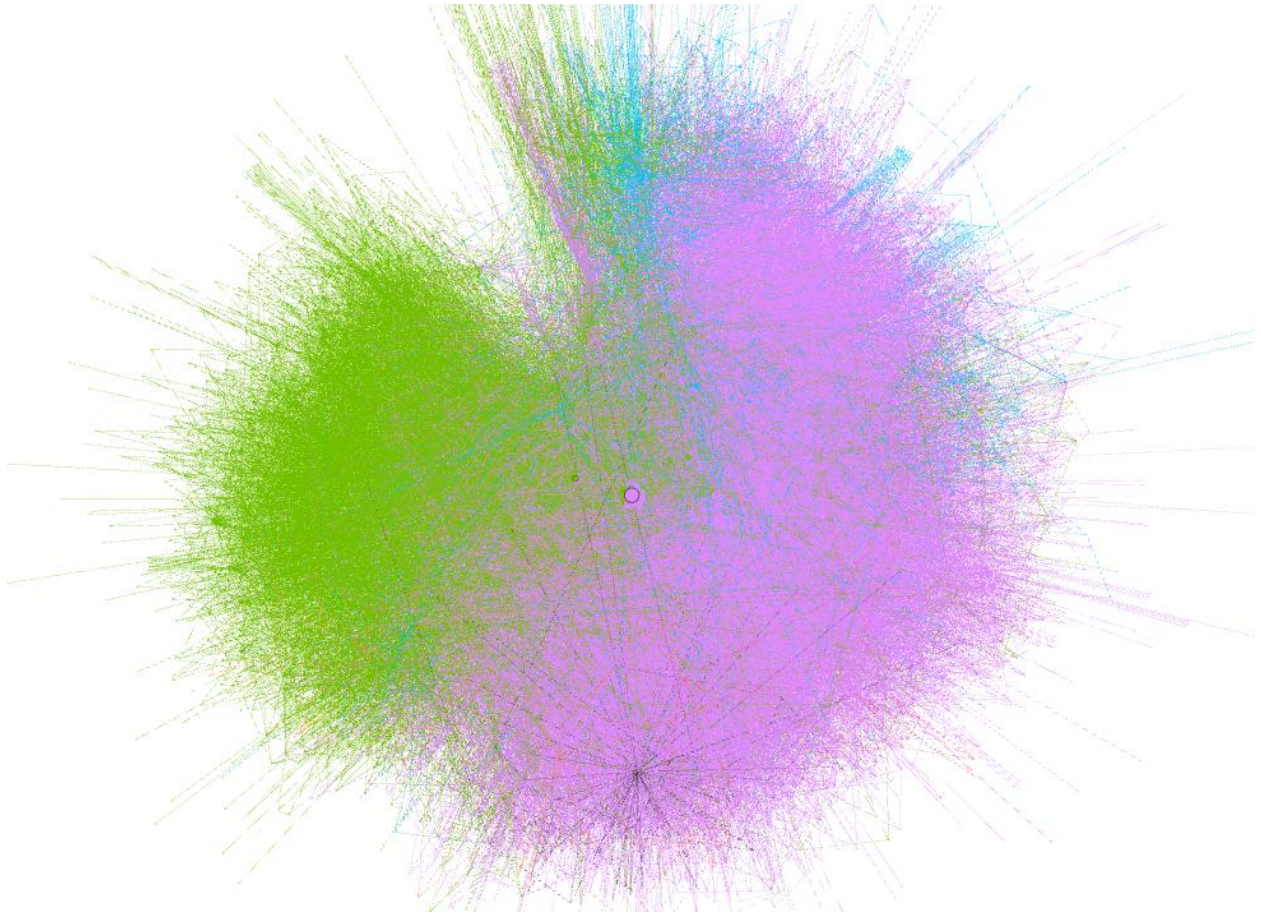


▼ Threads	
Threads number	15
▼ Performance	
Tolerance (speed)	1.0
Approximate Repulsion	<input checked="" type="checkbox"/>
Approximation	1.2
▼ Tuning	
Scaling	2.0
Stronger Gravity	<input type="checkbox"/>
Gravity	1.0

Điều chỉnh tham số ForceAtlas2:

- Scaling: Tăng giá trị này nếu đồ thị quá rối để tăng khoảng cách giữa các node.
- Gravity: Tăng để giữ các node gần trung tâm hơn.
- Prevent Overlap: Chọn mục này để ngăn chồng lấn các node.
- Edge Weight Influence: Điều chỉnh để các cạnh có trọng số lớn ảnh hưởng mạnh hơn.

Theo độ đo trung tâm



Phần 4:

So sánh dựa trên kết quả:

1. Số lượng cộng đồng:

- **Girvan-Newman** phát hiện nhiều cộng đồng nhất (27), phù hợp nếu bạn muốn phân tích sâu hơn và chi tiết hơn.
- **LPA** phát hiện ít cộng đồng hơn (15), cho thấy nó có xu hướng gom nhiều nút vào cùng một cộng đồng.
- **Louvain** cân bằng nhất với 20 cộng đồng, phù hợp nếu bạn cần một cái nhìn tổng quan không quá chi tiết.

2. Modularity:

- **LPA** có Modularity cao nhất (0.276), cho thấy cộng đồng được phân tách rõ ràng nhất.

- **Louvain** đứng thứ hai (0.248), vẫn đảm bảo cấu trúc cộng đồng nhưng không bằng LPA.
- **Girvan-Newman** có Modularity thấp nhất (0.238), điều này có thể do mạng được phân chia quá chi tiết, dẫn đến một số cộng đồng nhỏ không có kết nối chặt chẽ.

Giải thích ý nghĩa của các cộng đồng được phát hiện trong ngữ cảnh của mạng xã hội

Cộng đồng trong mạng xã hội đại diện cho các nhóm người dùng, tổ chức hoặc các phần tử có mối quan hệ chặt chẽ hơn với nhau so với phần còn lại của mạng. Ý nghĩa cụ thể:

1. Louvain (20 cộng đồng):

- Phù hợp để nhận diện các nhóm trung bình, có kết nối rõ ràng nhưng không quá chi tiết.
- Trong mạng xã hội, các cộng đồng này có thể là các nhóm bạn bè hoặc nhóm đồng nghiệp trong cùng một môi trường.

2. Girvan-Newman (27 cộng đồng):

- Chia mạng thành nhiều nhóm nhỏ hơn, phản ánh sự phân chia rõ ràng giữa các nhóm khác biệt.
- Các cộng đồng này có thể đại diện cho các nhóm chuyên biệt (ví dụ: các bộ phận trong công ty hoặc các cộng đồng theo sở thích nhỏ hơn).

3. LPA (15 cộng đồng):

- Gồm nhiều nút vào các cộng đồng lớn, phù hợp với các mạng có cấu trúc tổng quát.
- Các cộng đồng này có thể là các nhóm lớn như mạng bạn bè của các cá nhân có ảnh hưởng (influencers) hoặc các nhóm xã hội lớn.

Đề xuất: Louvain

- **Lý do chọn Louvain:**

1. Cân bằng giữa số lượng cộng đồng và Modularity:

- Louvain cung cấp số lượng cộng đồng vừa phải (20), không quá ít (như LPA) và không quá nhiều (như Girvan-Newman), phù hợp để phân tích mạng ở mức trung bình.
- Modularity đạt 0.248, cho thấy cấu trúc cộng đồng rõ ràng nhưng không quá chi tiết.

2. Hiệu suất cao:

- Louvain hoạt động hiệu quả trên các mạng lớn, phù hợp với mạng xã hội thực tế.

3. Tính ổn định:

- Louvain thường cho kết quả ổn định hơn so với LPA, và tốc độ xử lý nhanh hơn Girvan-Newman.