

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

## Graph Distance Report

### Parameters:

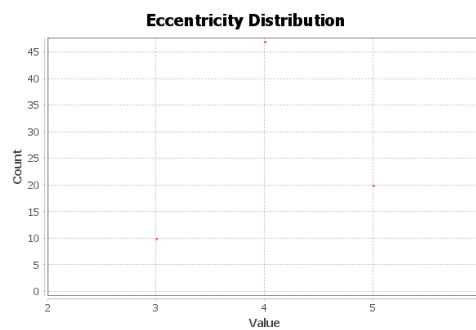
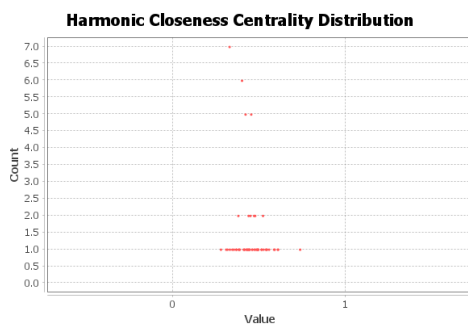
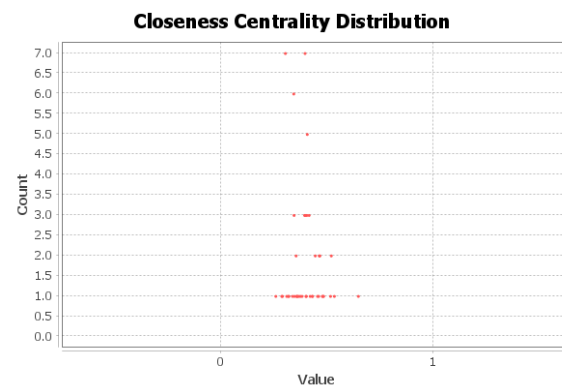
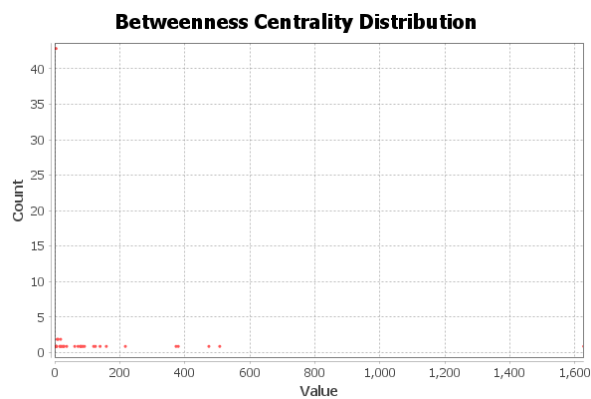
Network Interpretation: undirected

### Results:

Diameter: 5

Radius: 3

Average Path length: 2.6411483253588517



### Algorithm:

Ulrik Brandes, *A Faster Algorithm for Betweenness Centrality*, in Journal of Mathematical Sociology 25(2):163-177, (2001)

### 3 nút có các độ đo cao nhất:

Id	Label	Interval	Modularity ...	Degree	Weighted D...	Eccentric...	Closeness Centra...	Harmonic Closenes...	Betweenness Ce...
11	Valjean		1	36	158.0	3.0	0.644068	0.732456	1624.4688
55	Marius		6	19	104.0	3.0	0.531469	0.60307	376.292593
27	Javert		7	17	47.0	3.0	0.517007	0.585526	154.844945

- Closeness Centrality: 11 (Valjean), 55 (Marius), 27 (Javert)
  - Valjean (11): Nút quan trọng nhất, kết nối gần với tất cả các nút khác, giúp lan truyền thông tin nhanh chóng trong mạng.
  - Marius (55): Trung tâm thứ hai, hỗ trợ kết nối giữa các nút, đóng vai trò trung gian hiệu quả.
  - Javert (27): Đảm nhận vai trò kết nối cục bộ trong mạng.

Id	Label	Interval	Modularity ...	Degree	Weighted D...	Eccentric...	Closeness Centrality	Harmonic Close...	Betweenness Ce...
11	Valjean		1	36	158.0	3.0	0.644068	0.732456	1624.4688
48	Gavroche		8	22	56.0	3.0	0.513514	0.605263	470.570632
55	Marius		6	19	104.0	3.0	0.531469	0.60307	376.292593

- Harmonic Closeness Centrality: 11 (Valjean), 48 (Gavroche), 55 (Marius)
  - Valjean (11): Tiếp tục là trung tâm quan trọng nhất, có ảnh hưởng mạnh trong toàn bộ mạng.
  - Gavroche (48): Liên kết tốt với các nút xa hơn, giảm thiểu khoảng cách trong mạng.
  - Marius (55): Vẫn giữ vai trò trung gian, đảm bảo kết nối giữa các cụm.

Id	Label	Interval	Modularity ...	Degree	Weighted D...	Eccentric...	Closeness Centrality	Harmonic Closenes...	Betweenness ...
11	Valjean		1	36	158.0	3.0	0.644068	0.732456	1624.4688
0	Myriel		0	10	31.0	4.0	0.429379	0.491228	504.0
48	Gavroche		8	22	56.0	3.0	0.513514	0.605263	470.570632

- Betweenness Centrality: 11 (Valjean), 0 (Myriel), 48 (Gavroche)
  - Valjean (11): Là cầu nối quan trọng nhất giữa các cụm, quyết định luồng thông tin trong mạng.
  - Myriel (0): Hỗ trợ liên kết giữa các nhóm nút trong mạng, đảm bảo tính thông suốt.

- Gavroche (48): Đóng vai trò hỗ trợ, giữ vai trò kết nối cục bộ và giữa các cụm.

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

### Louvain:

#### Modularity Report

---

Parameters:

Randomize: On

Use edge weights: On

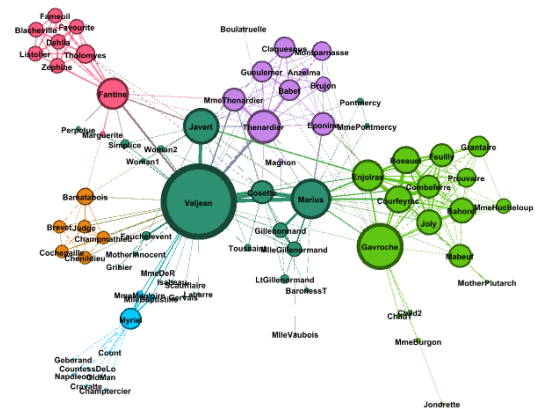
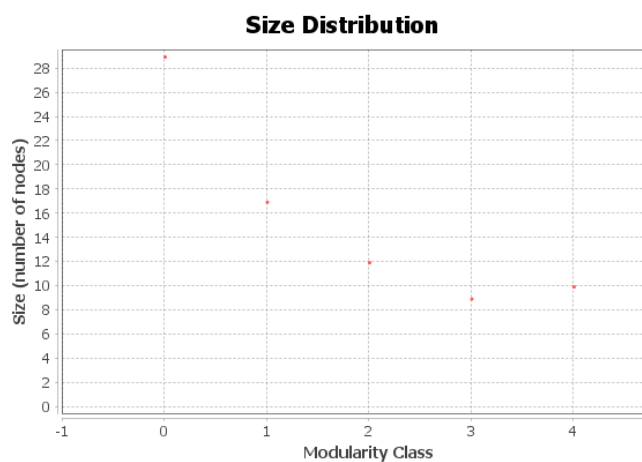
Resolution: 1.0

Results:

Modularity: 0.557

Modularity with resolution: 0.557

Number of Communities: 5



**Algorithm:**

Vincent D Blondel, Jean-Loup Guillaume, Renaud Lambiotte, Etienne Lefebvre, *Fast unfolding of communities in large networks*, in Journal of Statistical Mechanics: Theory and Experiment 2008 (10), P1000

**Resolution:**

R. Lambiotte, J.-C. Delvenne, M. Barahona *Laplacian Dynamics and Multiscale Modular Structure in Networks* 2009

**Girvan-Newman Report**

---

**Parameters:**

Respect edge type for shortest path betweenness:     yes

Respect parallel edges for shortest path betweenness: yes

Respect edge type for modularity computation:     yes

Respect parallel edges for modularity computation: yes

**Processed Graph Data:**

Nodes:             77

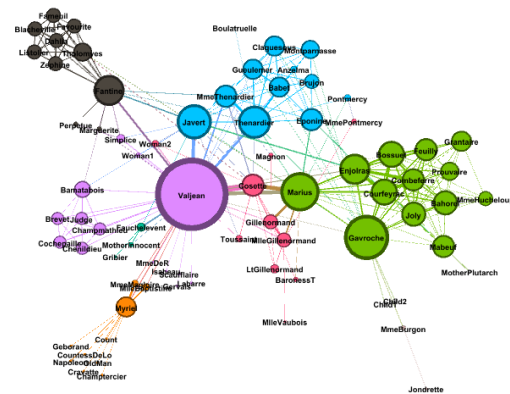
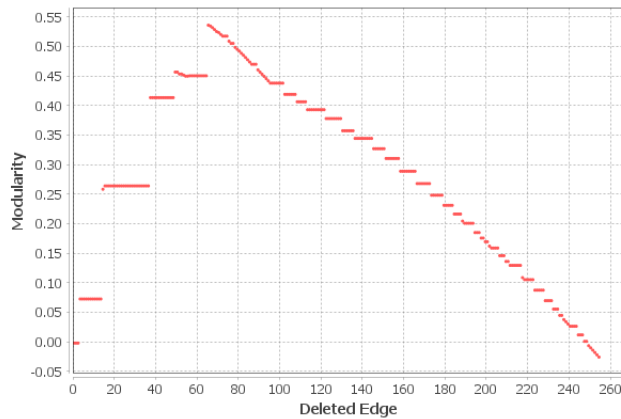
Edges             254

Processing time: 0.245 sec.

**Communities:**

Number of communities:     11

Maximum found modularity: 0.53806806



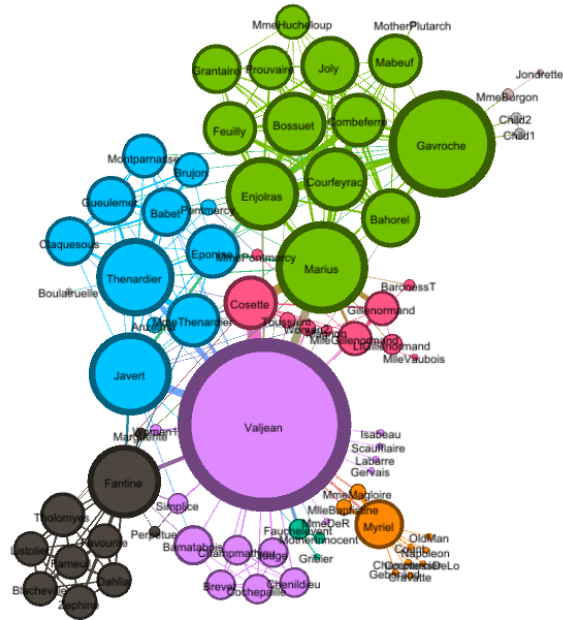
### Nhận xét:

- Louvain phát hiện số lượng cộng đồng ít hơn, nhưng giá trị Modularity cao hơn so với Girvan-Newman, cho thấy các cộng đồng được chia có tính chặt chẽ và rõ ràng hơn.
- Girvan-Newman phát hiện số lượng cộng đồng nhiều hơn, nhưng giá trị Modularity thấp hơn so với Louvain. Điều này cho thấy các cộng đồng có tính liên kết kém hơn và có thể bị phân mảnh.

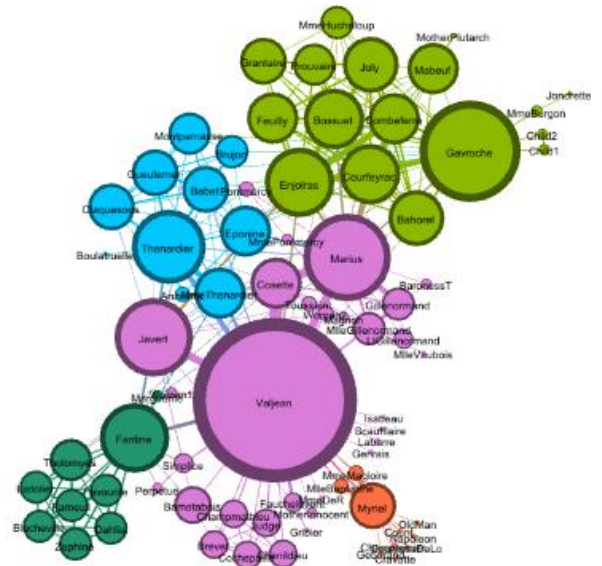
### Phần 3: Trực quan hóa Tạo một bản trực quan hóa đẹp và có ý nghĩa cho mạng lưới bằng cách:

1. Sử dụng thuật toán layout ForceAtlas2 với các tham số phù hợp:

## Givan-Nerman:

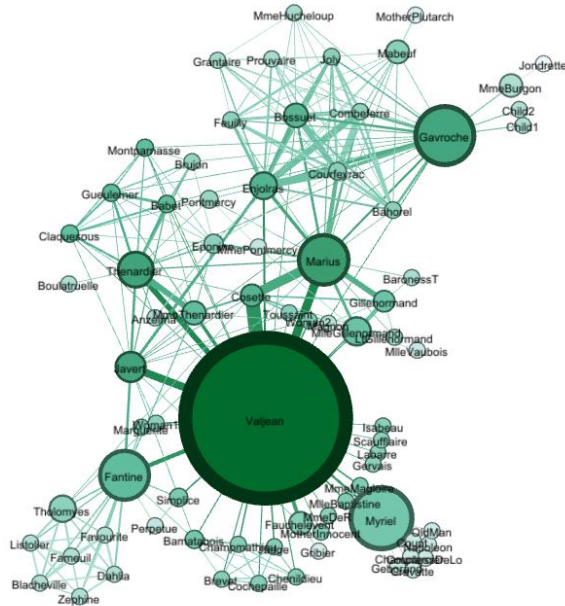


## Louvain:

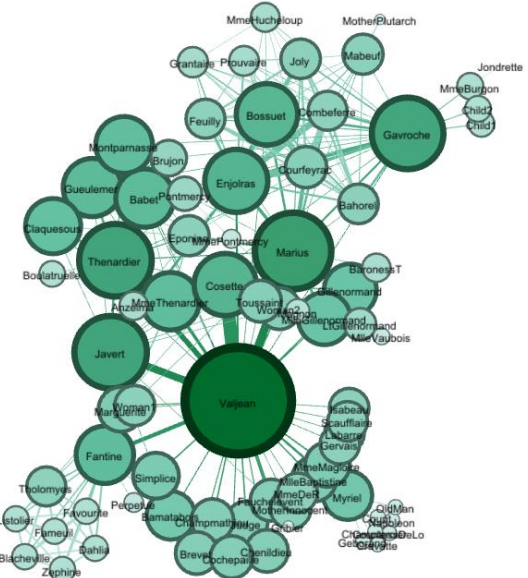


## 2. Điều chỉnh kích thước node theo độ đo trung tâm đã tính

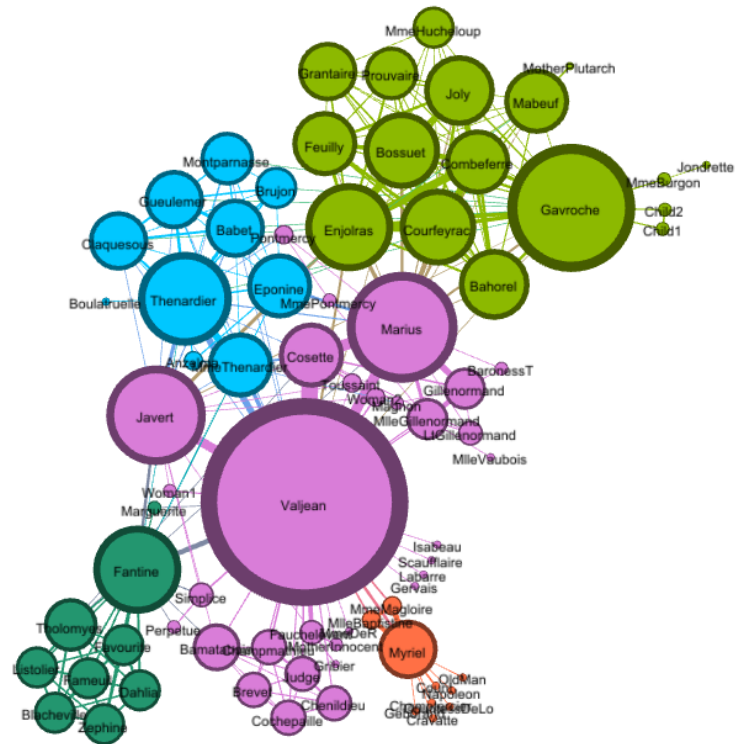
### Betweenness Centrality:



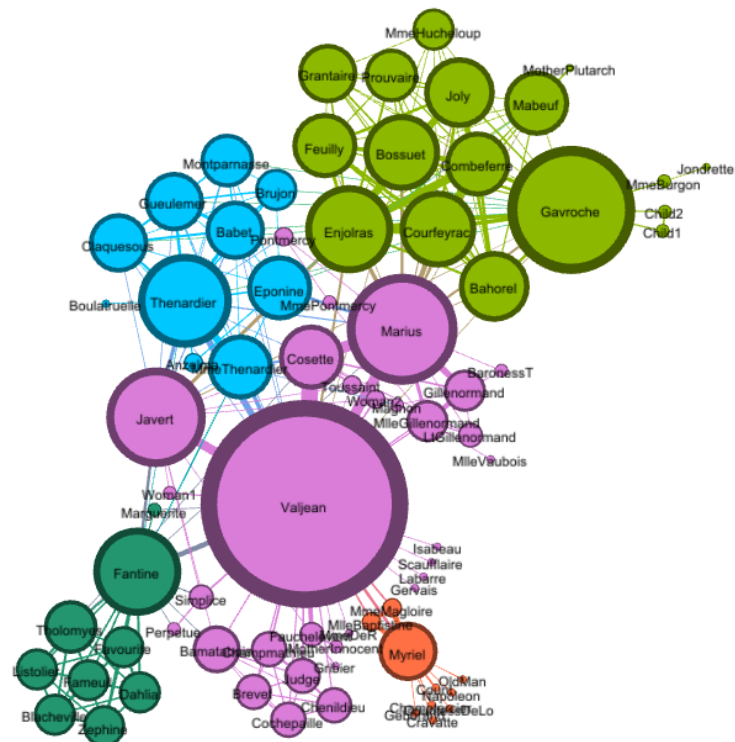
### Closeness Centrality



### 3. Tô màu node theo kết quả phân cụm từ thuật toán cho kết quả tốt nhất



### 4. Thêm nhãn cho các node quan trọng (có độ trung tâm cao)



#### Phần 4: Báo cáo và đánh giá Viết báo cáo ngắn bao gồm:

1. So sánh kết quả của 3 thuật toán phân cụm, nêu ưu và nhược điểm của mỗi phương pháp.

	Louvain	Girvan-Newman
Số lượng cộng đồng	5	11
Modularity	0.557	0.538
Thời gian xử lý	Nhanh hơn Girvan-Newman	0.245 giây

	Ưu điểm	Nhược điểm
Louvain	<ul style="list-style-type: none"><li>- Tính toán nhanh, hiệu quả cao, phù hợp với mạng lớn.</li><li>- Modularity cao hơn (0.557).</li></ul>	<ul style="list-style-type: none"><li>- Phụ thuộc vào độ phân giải (Resolution).</li><li>- Kết quả có thể khác nhau do ngẫu nhiên hóa.</li></ul>
Girvan-Newman	<ul style="list-style-type: none"><li>- Thích hợp cho mạng nhỏ hoặc vừa.</li><li>- Kết quả rõ ràng, trực quan trong việc phát hiện cộng đồng nhỏ.</li></ul>	<ul style="list-style-type: none"><li>- Không phù hợp với mạng lớn do chi phí tính toán cao.</li><li>- Modularity thấp hơn Louvain (0.538).</li></ul>



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

- Cộng đồng (Communities):
  - o Cộng đồng là các nhóm người (hoặc nút) có tương tác nhiều với nhau hơn so với các nhóm khác.
  - o Ví dụ: Trong mạng xã hội, một cộng đồng có thể là nhóm bạn thân, đồng nghiệp, hoặc những người có sở thích chung.
- Ý nghĩa của số lượng cộng đồng:
  - o Louvain (5 cộng đồng): Phân chia mạng xã hội thành ít nhóm lớn hơn, thể hiện tính liên kết mạnh mẽ giữa các nút trong cùng một nhóm.
  - o Girvan-Newman (11 cộng đồng): Phân chia chi tiết hơn, phù hợp khi cần phân tích các nhóm nhỏ hoặc cấu trúc mạng phức tạp.
- Ý nghĩa của Modularity:
  - o Modularity đo lường mức độ tách biệt giữa các cộng đồng. Modularity cao (Louvain: 0.557) cho thấy các cộng đồng được phân biệt rõ ràng.

## **3. Đề xuất phương pháp phân cụm phù hợp nhất cho loại dữ liệu này và lý do.**

### **Đề xuất:**

- Louvain là phương pháp phù hợp nhất cho loại dữ liệu này.

### **Lý do:**

- Tính hiệu quả:

- Louvain nhanh chóng và hiệu quả hơn, đặc biệt khi mạng có xu hướng mở rộng (tăng số nút và cạnh).
- Độ phân biệt cộng đồng tốt hơn:
  - Modularity cao hơn (0.557 so với 0.538) cho thấy khả năng phân tách cộng đồng của Louvain vượt trội.
- Phân cụm ít hơn nhưng chất lượng cao:
  - Số lượng cộng đồng vừa phải (5) giúp dễ dàng phân tích ý nghĩa thực tế hơn so với 11 cộng đồng của Girvan-Newman.
- Thích ứng linh hoạt:
  - Louvain có thể điều chỉnh tham số độ phân giải để phát hiện cộng đồng ở các cấp độ chi tiết khác nhau.

### **Khi nào nên chọn Girvan-Newman:**

- Khi mạng có kích thước nhỏ (như hiện tại) và mục tiêu là phân tích các nhóm nhỏ chi tiết.