김종윤
2023-28318
협동과정 인공지능 전공

소셜컴퓨팅

과제 3 - 4주차

1. Network Book Exercise 3.7 - Q2 ~ 5

1.1.Q2

The connection between B and C must be weak tie.

Let's assume that the connection between B and C is strong tie.

For node B, it is strongly tied with node E and C this must result at least weak tie between node E and C. Similarly for node C, as it is strongly tied with node B and F, node B and F must have at least weak tie. However, the tie between E and C, B and F is not constructed, the assumption violates the triadic closure property.

Therefore, the connection between B and C is weak tie.

1.2.Q3

There are 5 nodes on the Figure 3.22. Let's go around each node to see whether it satisfies the triadic closure property.

i) node A (O)

The node A is having strong tie with node B and D. There is a weak tie between node B and D.

The triadic closure assumption is satisfied.

ii) node B (O)

There is strong tie with node A and C. The weak tie is constructed between node A and C.

The triadic closure property is satisfied.

iii) node C

There is strong tie with node B and E. However, there is no connection between node B and E directly.

This violates the triadic closure assumption.

iv) node D (O)

There is strong tie with node A and E. There is weak tie between node A and E.

This satisfies triadic closure property.

v) node E

Similar to node C, although there is strong tie with node C and D, no connection is constructed between node C and D.

The triadic closure is violated.

Therefore, node A, B and D satisfies the property.

1.3.Q4

The network described on Figure 3.23 is identical to Figure 3.22.

Therefore the node C and E violates the property.

1.4.Q5

Similar question is given, similar solution will be given.

i) node A

The strong tie with node B and C. There is a strong tie between node B and C.

This satisfies the property.

ii) node B

Same triangle as node A.

Satisfies the property.

iii) node C

The node C has strong tie with node A, B, and E. However, there is no connection between node A and E.

This violates the triadic closure assumption.

iv) node D

The node D only has weak tie.

Nothing to consider further.

v) node E

The node E has only one strong tie.

No more further investigation required.

In a nut shell, node A, B, D and E satisfies the property.

2. 레미제라블 데이터셋 분석

2.1.Centrality가 높은 상위 5명 인물

2.1.1.Degree Centrality

The degree centrality gets high value when there is more connected nodes. As expected the centrality of the characters are a sequence of 'Valjean', the son of Thenardier, 'Gavroche', the son-in-law 'Marius', the police 'Javert' and 'Thenardier' the innkeeper, Fantine left Cosette at the inn.

The Valjean gets the highest value, about 0.47 where others get lower than 0.3.

Character	Centrality
Valjean	0.47368421052631600
Gavroche	0.2894736842105260
Marius	0.25
Javert	0.22368421052631600
Thenardier	0.21052631578947400

2.1.2.Betweenness Centrality

The betweenness centrality measures how other nodes establishes connection via the node. Easily can be found from the term 'betweenness', it shows how much the node is between others.

The betweenness centrality of 'Valjean' is the highest value even compared with others. This describes that Valjean is the main character that is located between a number of nodes.

Character	Centrality
Valjean	0.569989052783619
Myriel	0.176842105263158
Gavroche	0.165112502425848
Marius	0.132032488621946
Fantine	0.129644540988194

2.1.3.Closeness Centrality

The closeness centrality is how the node is close to other nodes.

'Valjean' shows highest closeness and 'Marius' (the son-in-law), 'Thenardier' (the innkeeper) follow.

Because the other major characters also have connectivity between other characters, their value is quite large than expected. Still, 'Valjean' gets conspicuously high value.

Character	Centrality
Valjean	0.644067796610169
Marius	0.531468531468532
Thenardier	0.517006802721089
Javert	0.517006802721089
Gavroche	0.513513513513514

2.1.4.Eigenvector Centrality

The eigenvector centrality shows high value if the node is connected to large network. Unlike other centrality measures, Valjean does not get the highest centrality measure, while Gavorche (the son of Thenardier, the innkeeper) who has quite a lot of connectivity between others and have possibility to be connected to other large group as his parents run an inn.

Character	Centrality
Gavroche	0.317838939774977
Valjean	0.267618175988539
Enjolras	0.267178632823567
Marius	0.259111145341788
Bossuet	0.242130786374741

2.2.인물들의 커뮤니티 그룹

2.2.1.적당한 커뮤니티 그룹 수

With `greedy_modularity_communities` method from `networkx` returns 5 community group. The `girvan_newman` method shows 76 possible community groups which can be divided by 2~77 groups. As `greedy_modularity_communities` method returns 5 groups, the number around 5 will show best description. The graph of "modularity - number of community" will help which is the best number of community group.

Therefore, best number of community group is 11.

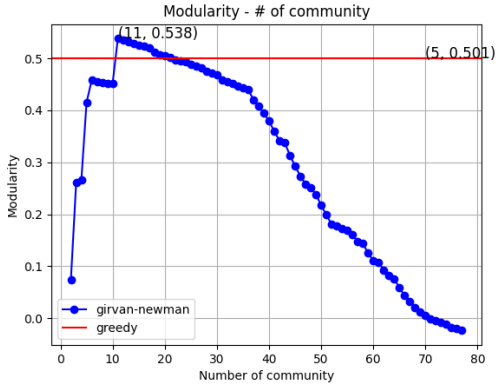

Figure 1: Modularity - Number of community.
(Greedy method shows approximately 0.5, while girvan-newman shows 0.538 when there are 11 groups.)

2.2.2.커뮤니티 그룹 시각화

The visualization of the community group is described on the following figures.

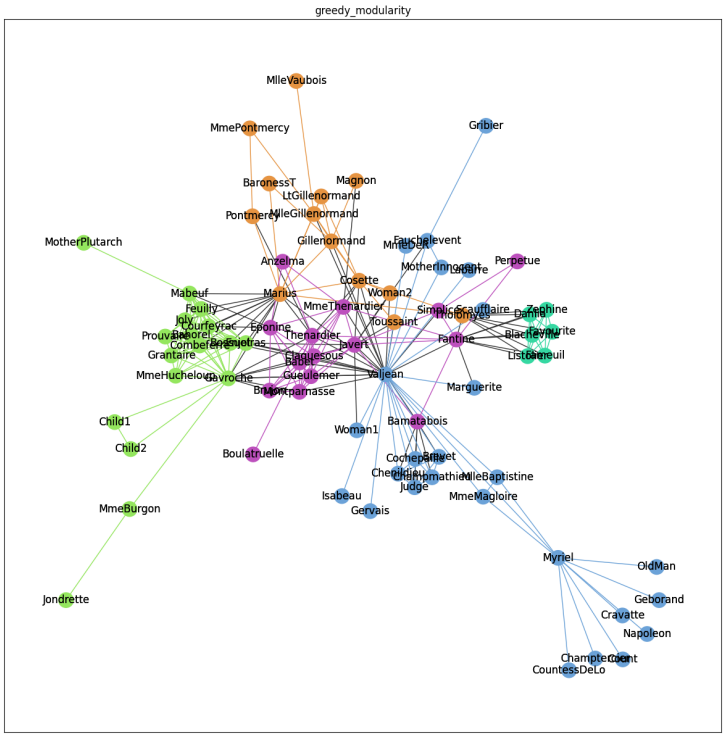

Figure 2: Visualization of greedy modularity

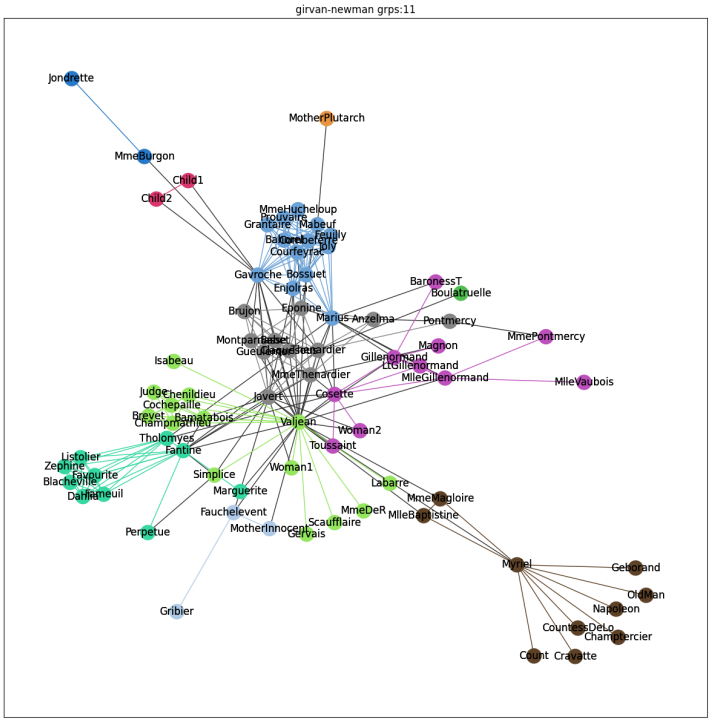

Figure 3: Visualization of Girvan-Newman modularity with group size 11.