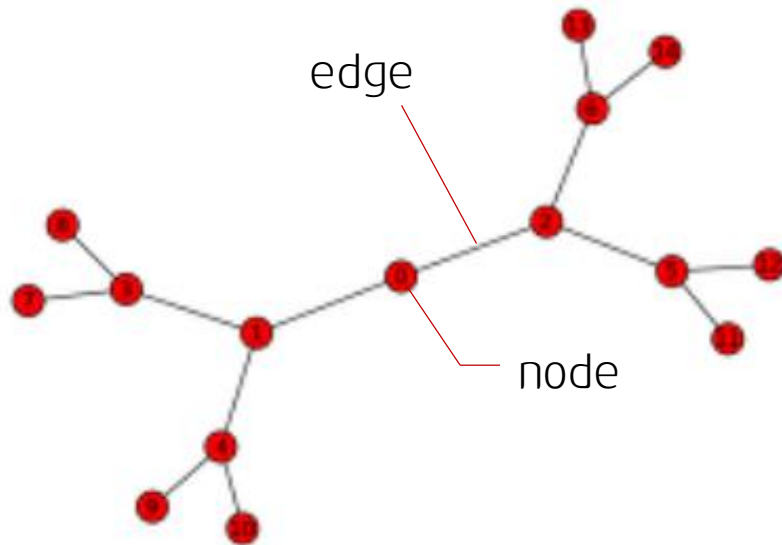


Network Analysis +) Network X

디자인팀 김수정

Network

- 점(node)과 연결선(edge)들로 이루어진 집합

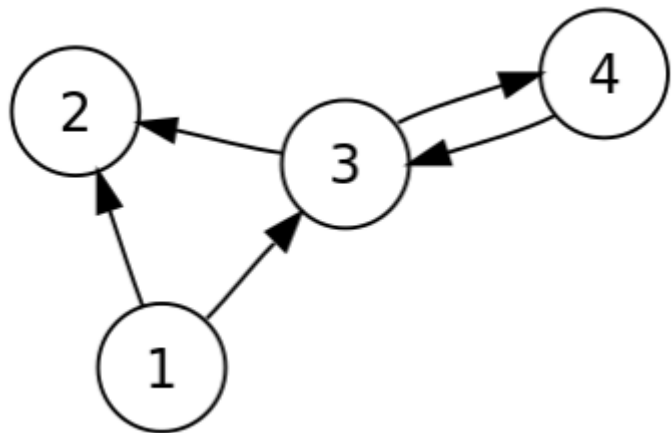


node : 개체

edge : 개체들 간의 연결관계

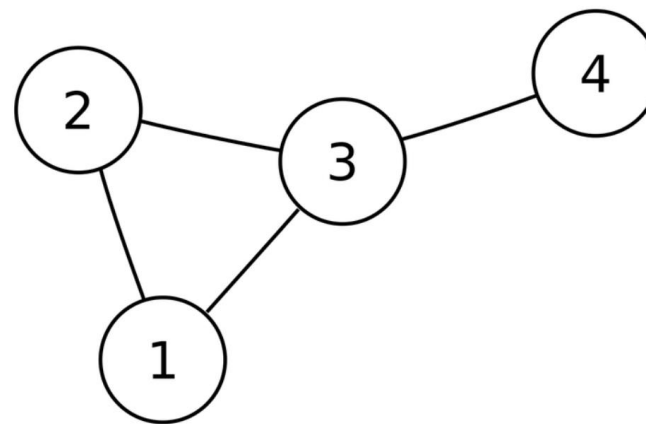
Graph의 종류

- Directed Graph



방향성이 있는 그래프
 $(3,4) \neq (4,3)$

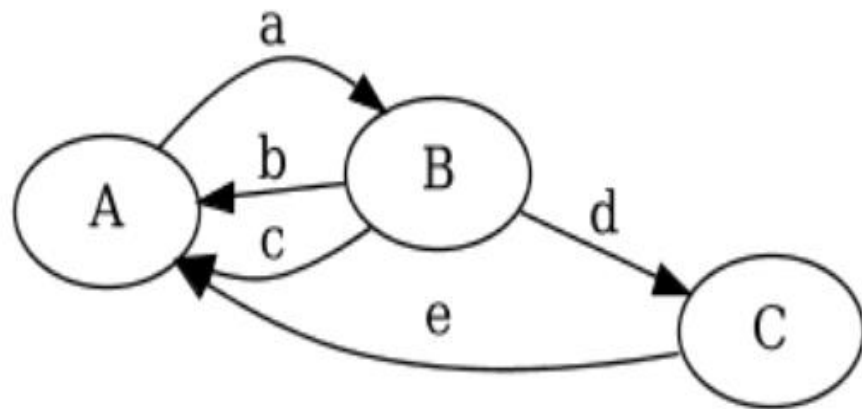
- Undirected Graph



방향성이 없는 그래프
 $(3,4) = (4,3)$

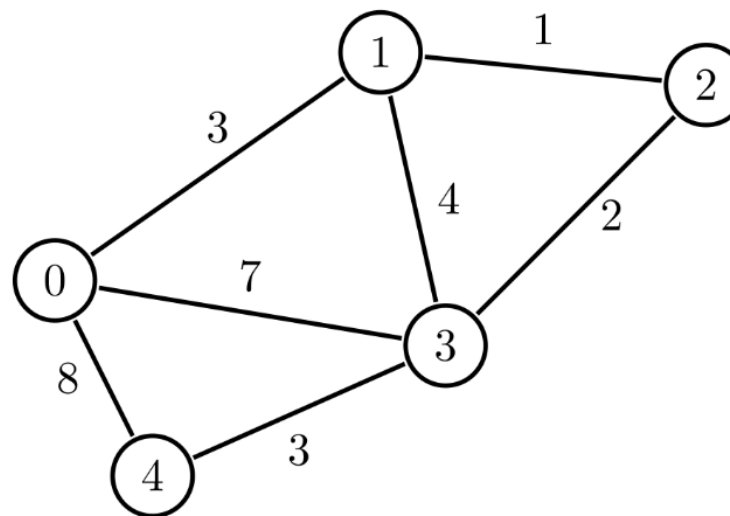
Graph의 종류

- Multigraphs



Node들 사이에 복수의 edge를 허용
Multigraph / MultiDigraph

- Weighted graph



각 edge에 weight를 추가

주요 네트워크 분석 기법

- 노드 중요도 측정 (Node Centrality)
- 네트워크 구조 추정 (Network Structure)
- 커뮤니티 탐지 (Community Detection)

Node Centrality

- 네트워크 구조를 고려했을 때, **각 노드들이 얼마나 중요한 위치에 있는지를** 정량화하는 것
- 중요한 node가 없어질 경우, 전체 네트워크의 흐름 자체에 영향을 받게 된다. -> 중요한 node를 찾아 전체 네트워크 흐름을 제어할 수 있다.

Node Centrality

Degree Centrality

- 각 node가 다른 node와 얼마나 많이 연결되어 있는지를 계산
- 연결된 node의 **개수**를 count
- node의 연결 방향성에 따라 In-degree centrality와 Out-degree centrality로 나눌 수 있음

In-degree centrality: 해당 node로 들어오는 관계만을 고려(인기도)

Out-degree centrality: 해당 node에서 나가는 관계만을 고려(영향력)



Node Centrality

Eigenvector Centrality

- 해당 node의 centrality를 계산할 때 해당 node와 연결된 **다른 node의 centrality**를 반영하여 계산
- 단순히 node의 개수를 계산하는 Degree Centrality의 단점을 개선



Node Centrality

Betweenness Centrality

- 노드 간의 흐름을 고려하여 중요도를 측정
- 다른 node들 사이의 **shortest path**에 해당 node가 많이 포함될수록 중요한 node
라고 판단
- 많이 거쳐가야 하는 node가 중요한 node



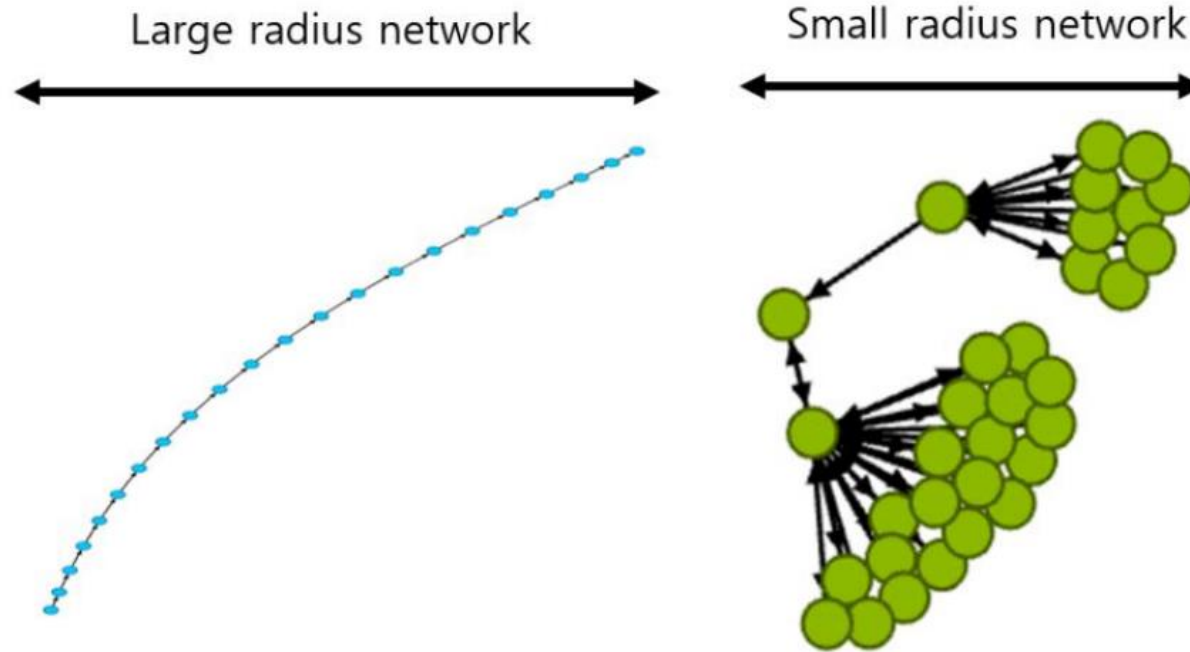
Node Centrality

Closeness Centrality

- 거리를 기준으로 측정
- 모든 다른 node까지의 거리의 평균이 작을수록 중요한 node
- 접근성 측면을 고려한 것

Network Structure

Radius

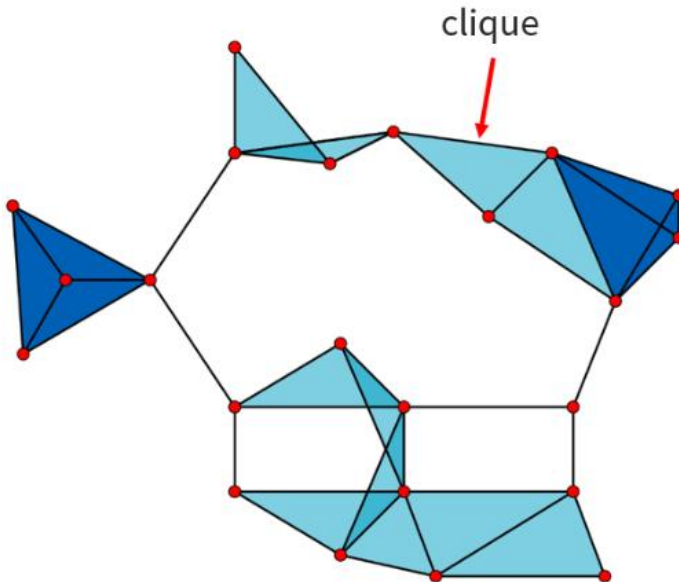
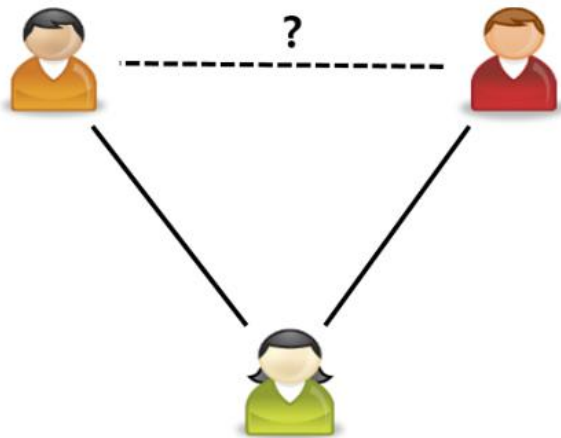


- Closeness Centrality가 가장 높은 node와 가장 낮은 node사이의 거리

Network Structure

Clustering Coefficient

- 어떤 노드와 연결된 두 개의 임의의 노드가 연결되어 있을 확률
- 네트워크가 얼마나 긴밀하게 뭉쳐있는지를 알려주는 지표

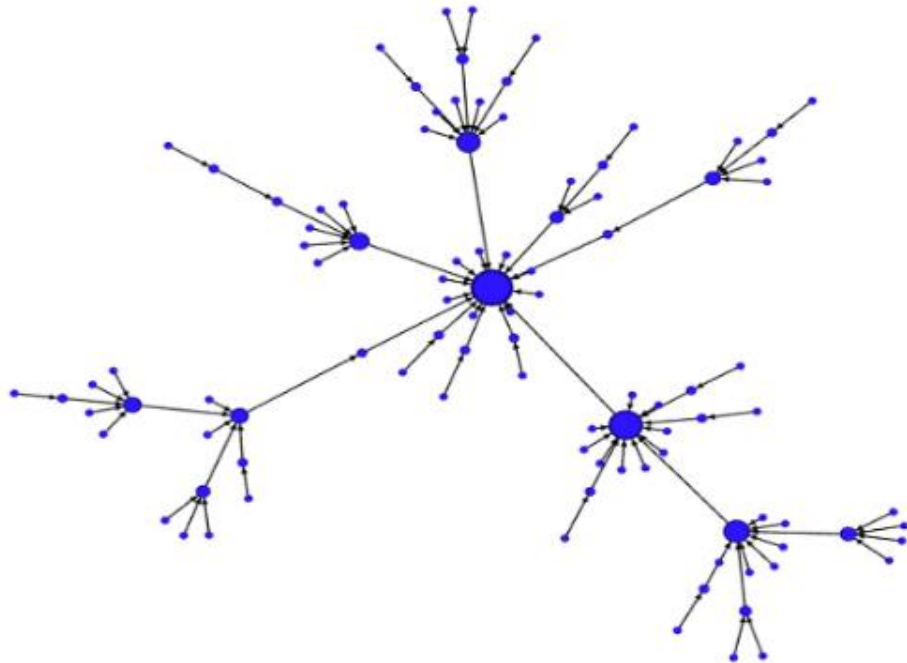


* Clique : 3개의 node가 연결되어
있는 구조

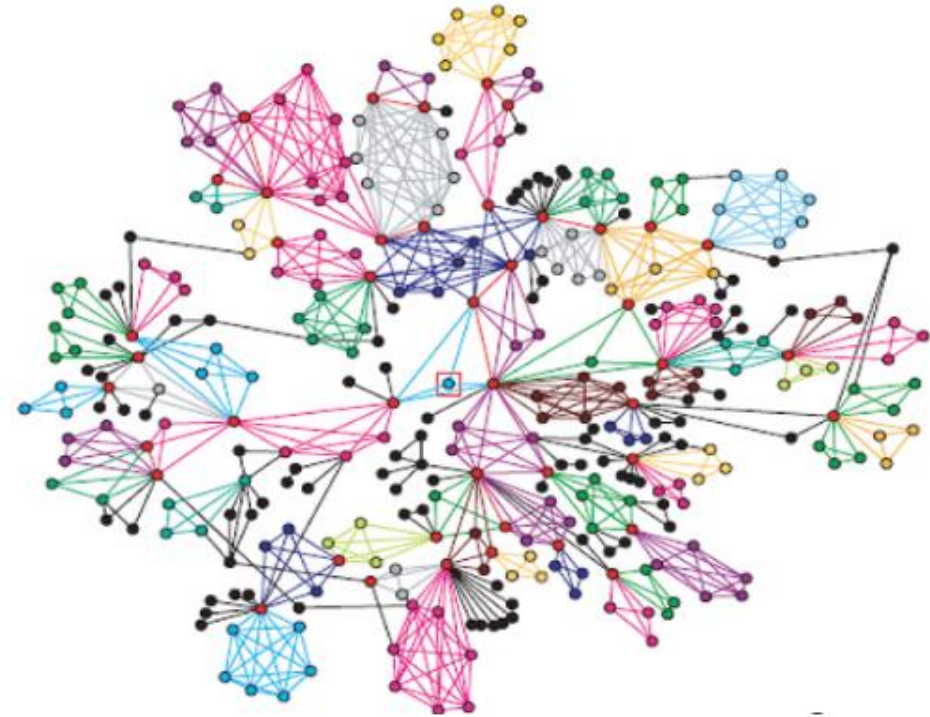
Network Structure

Clustering Coefficient

Low clustering coefficient



High clustering coefficient



Network Structure

Degree Assortativity

- 서로 연결된 node 쌍에 대해 각 node의 **degree Centrality에 대한 상관관계**를 측정하는 값
- -1에서 1사이의 값을 가짐

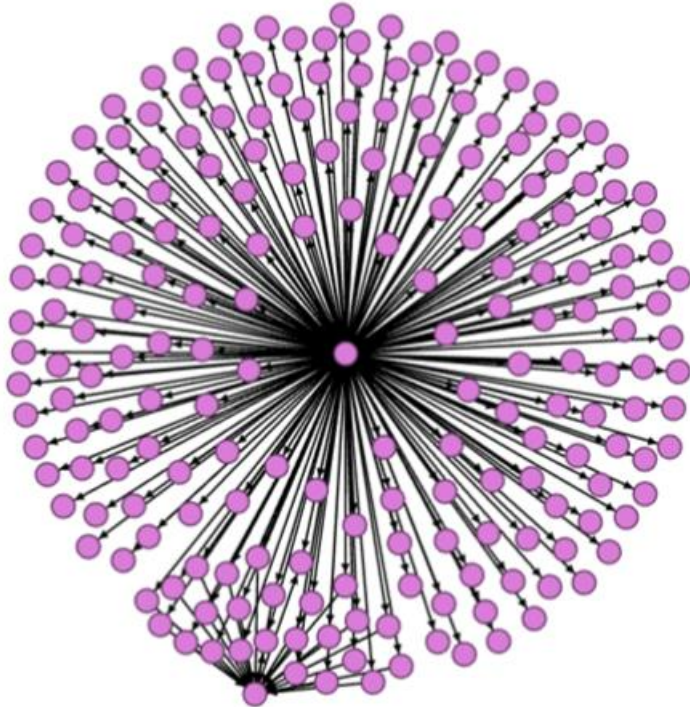
Degree Centrality가 비슷한 node들끼리 연결된 경우: Degree Assortativity가 1에 가까움

Degree Centrality가 낮은 node와 높은 node가 연결된 경우: Degree Assortativity가 -1에 가까움

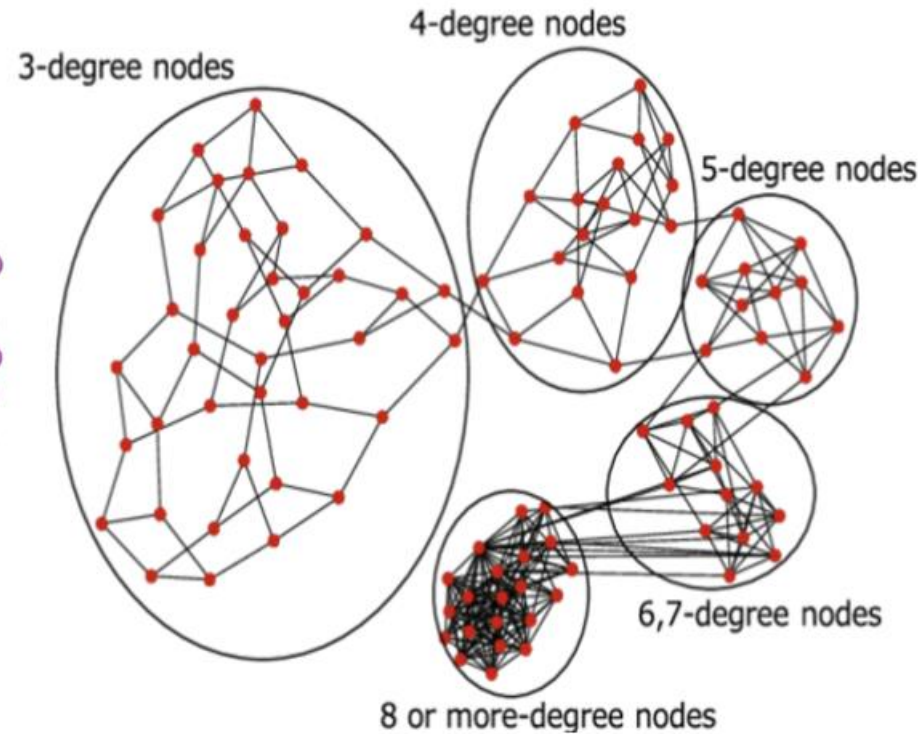
Network Structure

Degree Assortativity

Low degree assortativity network



High degree assortativity network

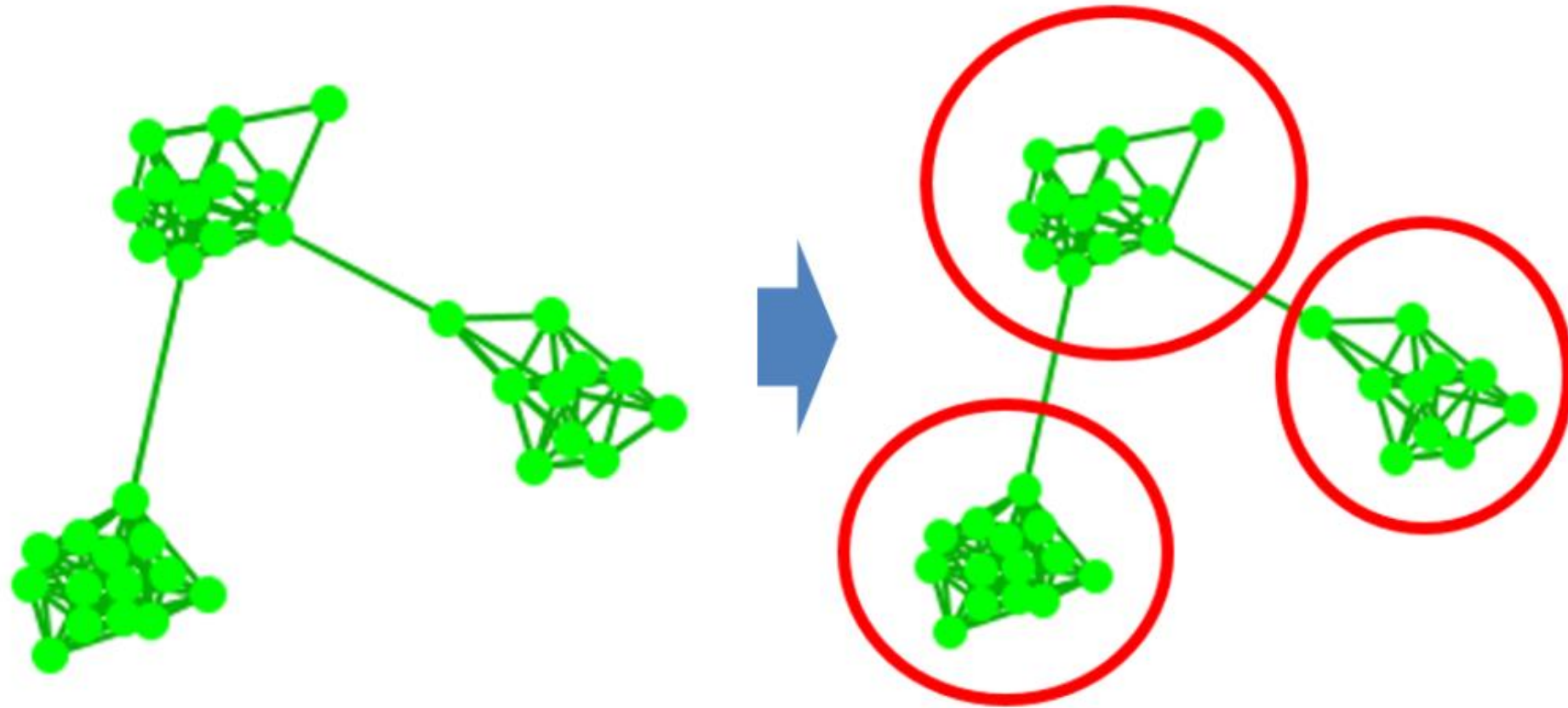


Community Detection

- 전체 네트워크를 연결밀도가 상대적으로 높은 소집단끼리 묶어서 분석하는 기법
- 분석하고자 하는 네트워크 규모가 커서 분석해야 할 node 개수가 많고 구조가 복잡할 때 사용하는 기법
- 전체 구조에서는 발견하기 힘든 여러 가지 정보를 얻을 수 있음

community: 다른 집단에 비해 상대적으로 연결 밀도가 높은 node 집단

Community Detection



분석 사례

- 드라마/소설 속 인물 간의 관계 분석
- 소셜네트워크 분석 (SNS Network Analysis)
- 악성코드 탐지
- 온라인 게임 유저 분석

참고: <https://brunch.co.kr/@gimmesilver/46>

<http://www.dator.co.kr/encore/textyle/2571789>

Network X

- Python Library
- 무방향성, 방향성, 다중그래프 등의 다양한 데이터 구조 지원
- Node, Edge의 자유도 높음 (거의 모든 데이터 형식 지원)
- matplotlib과 연계된 시각화 지원
- Node Centrality 척도들, Network Structure 등을 쉽게 계산할 수 있음

실습 / 과제