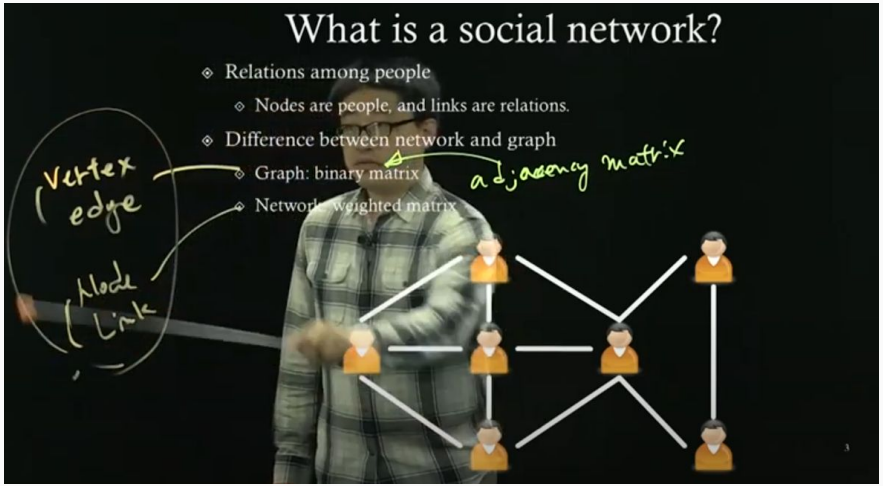
6. Non Linear Structure, Optimization, and Algorithm

- Social Network
- Centrality
- K Clique and Newman Clustering
- Visualization of Network Structure

Social Network



social network 란?
사람들 사이의 관계

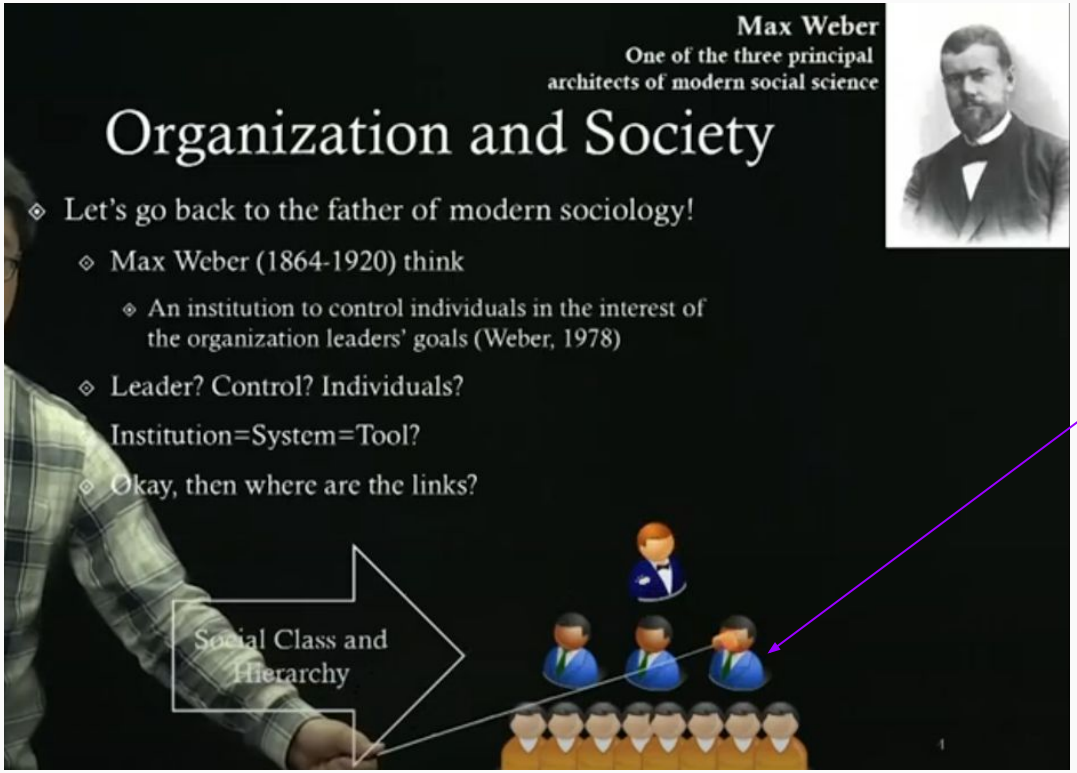
Nodes = 사람
Links = 관계

0과1로 이루어져 있으면
Graph : binary matrix
vertex
edge

0보다 큰 실수로 이루어져있으면
Network : weighted matrix
node
link

둘 다 adjacency matrix인접행렬

Social Network

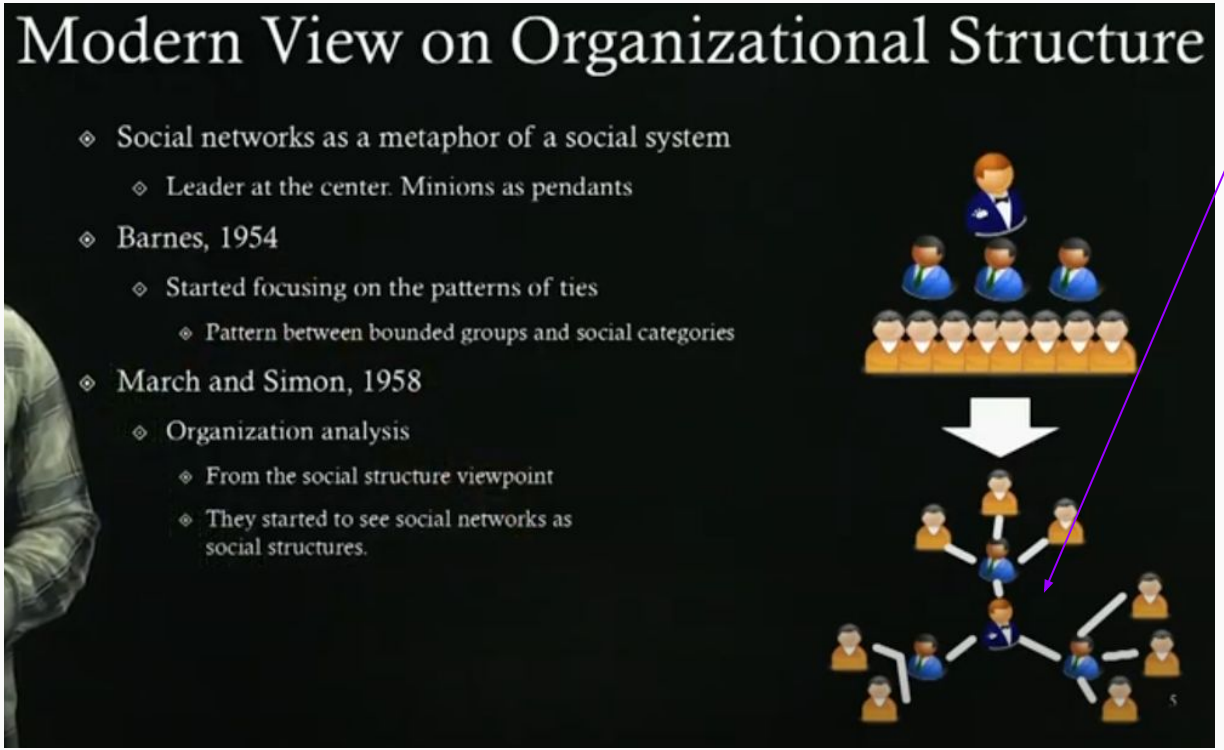


소셜네트워크의 시작은 분야마다 보는 시각이 다르다

sociology관점으로 본다면 Max Weber가 시작

리더의 이익에 맞춰 개인들을 컨트롤하는것이 사회 시스템이다 라고 max weber는 생각했다

Social Network

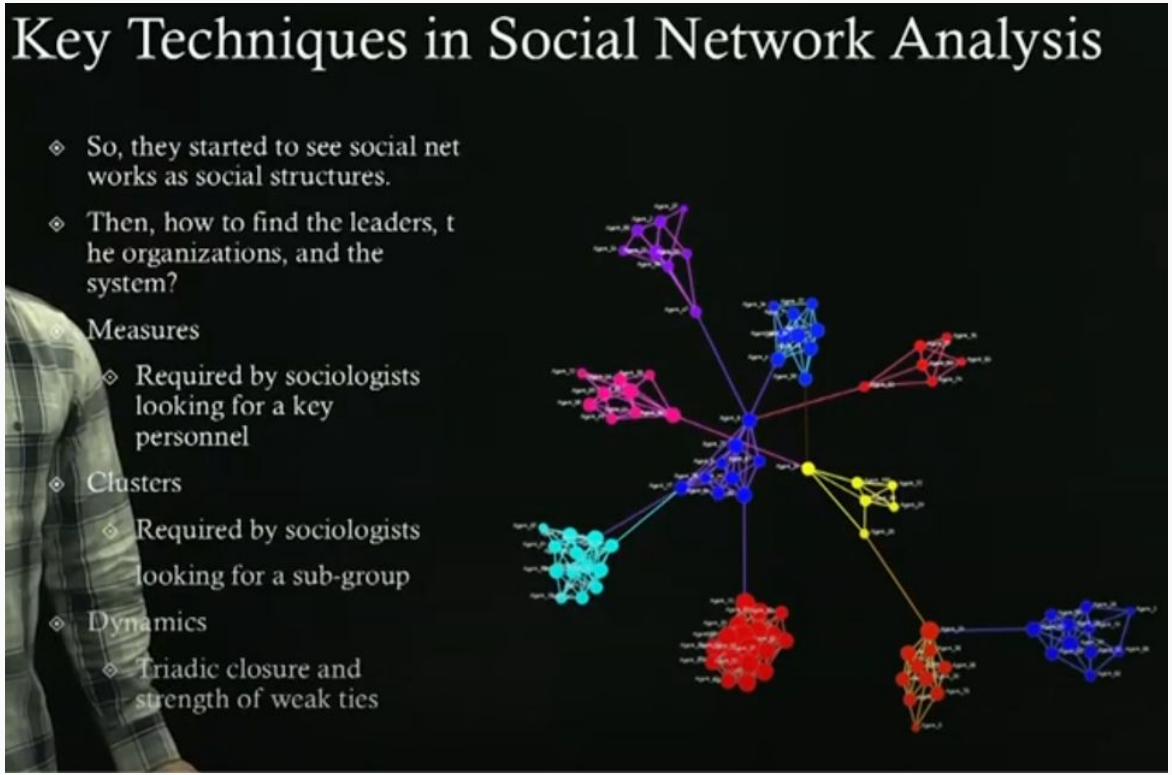


근대적인 관점에서 본 조직
리더가 정보의 흐름을 컨트롤해서
시스템을 만드는게 근대 조직

barnes
tie의 패턴에 대해 얘기해야한다
카테고리또는 묶여있는 그룹가
넘어가는 과정에서
어떤식으로 패턴이 생기는지
공부해야한다고 함

march 와 simon
조직에 대한 분석
소셜네트워크를 소셜스트럭쳐로
표현할 수 있는 방법으로 봄

Centrality

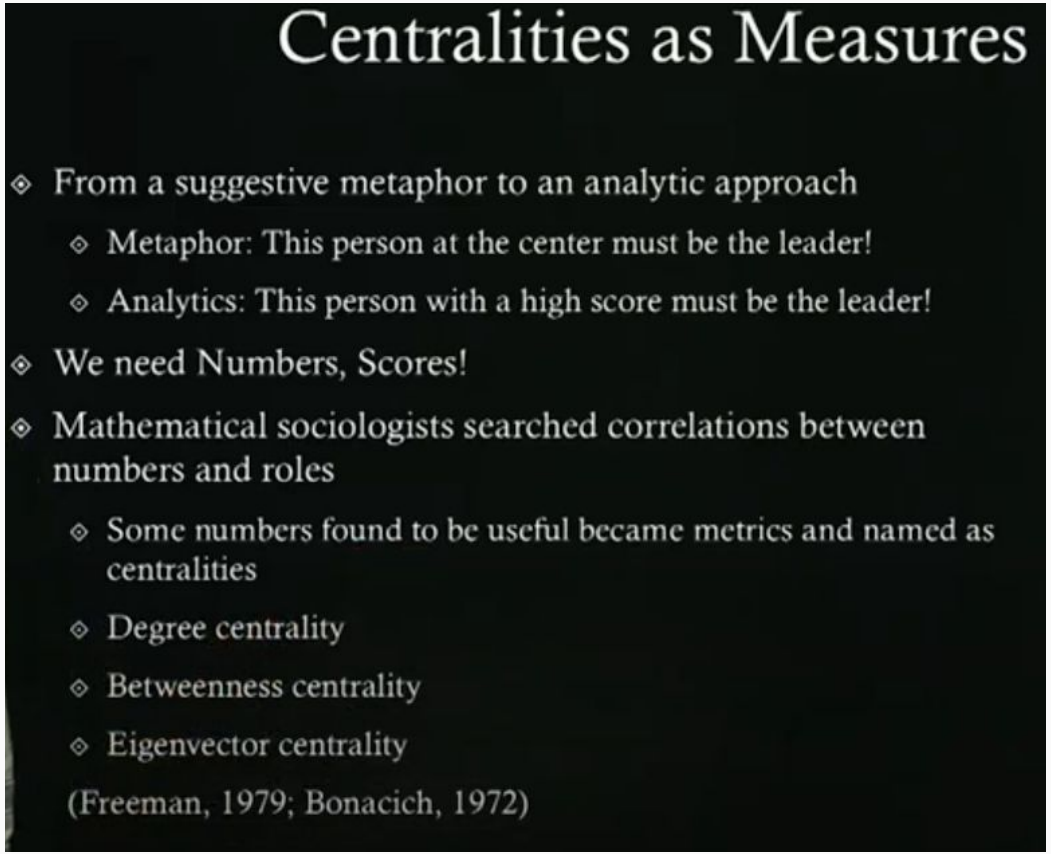


그림은 cellular network라고 부른다

어떻게 리더를 찾고 분석을 하는가?

- 점 하나하나에 중요도 점수를 매겨서 중심도라는걸 구한다(Measures)
- 어떻게 하면 내부조직을 구하는가?(Clusters)
- 서로서로가 잘 숨겨질 수 있도록 만들어진 구조이다(Dynamics)

Centrality



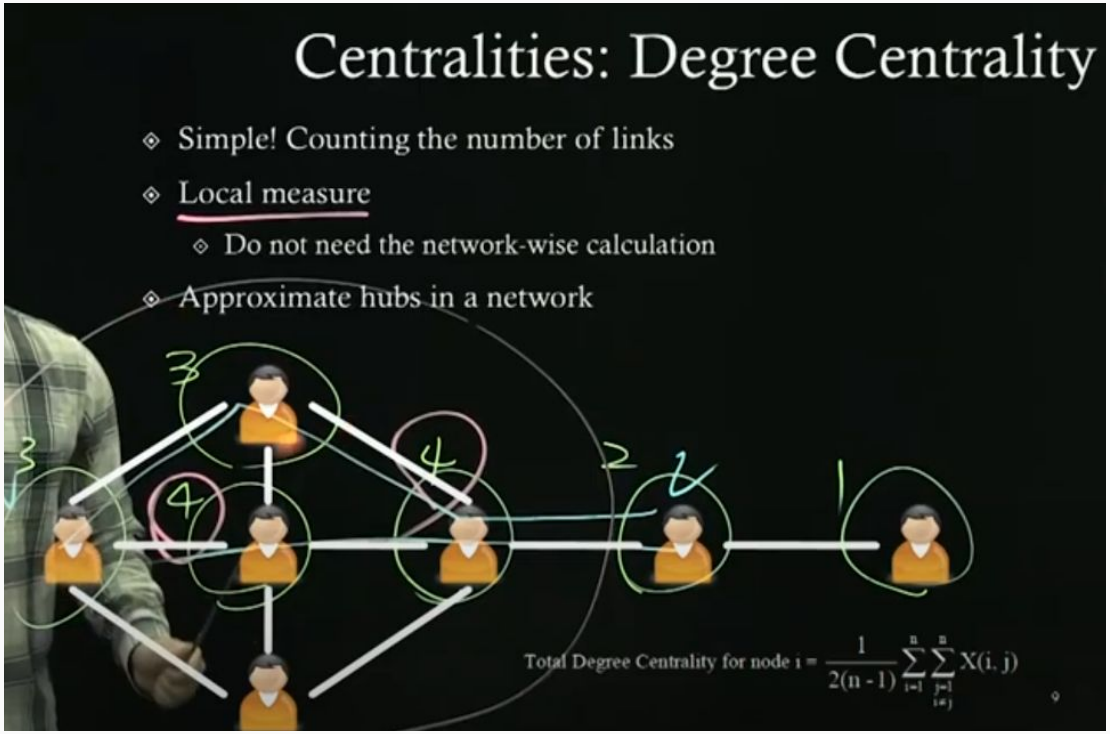
Metaphor = 이 사람은 리더이다(선언적)
Analytcs = 가장 높은 점수를 가지고 있으니 이 사람은 리더이다(구분, 정략적인 분석)

수리적인 방법을 활용해서 사회적인 현상을 분석해야한다고 한다
매우 유용한 점수체계를 가지고 있는 중심도 3가지

- Degree centrality
- Betweenness centrality
- Eigenvector centrality

중에 Degree와 Betweenness를 알아본다

Centrality



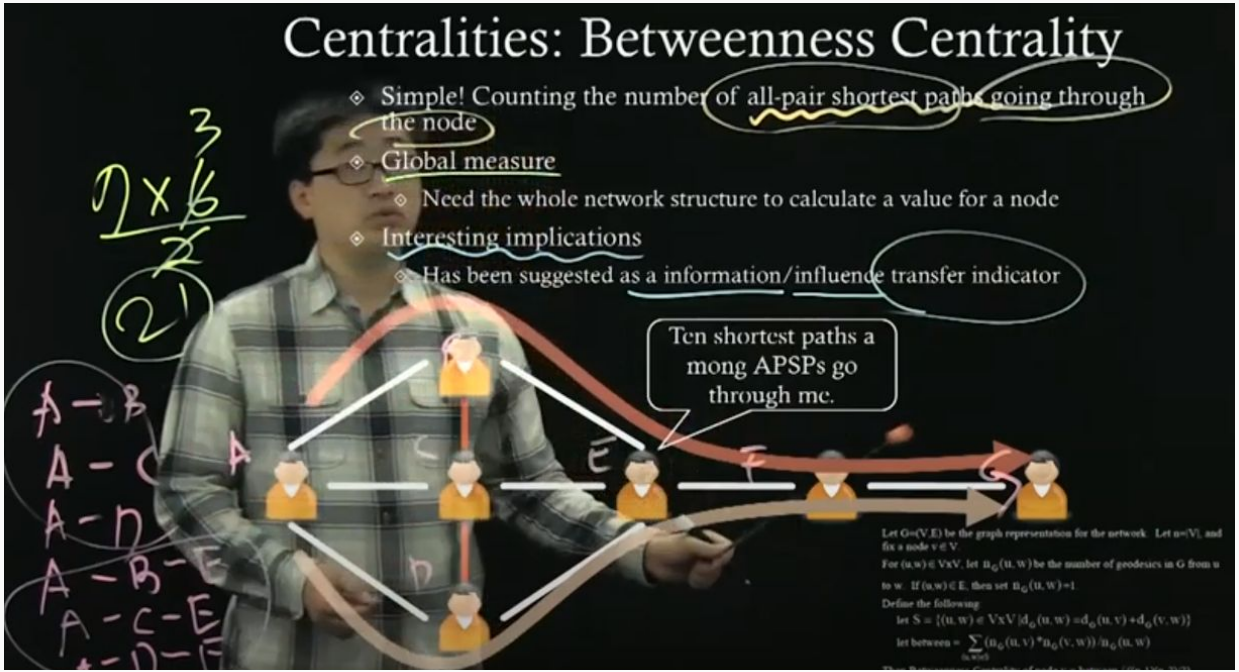
Degree Centrality는
노드는 몇개의 링크를 가지고있는가를 계산한다

Local measure

장점 = 빠른 계산

문제점 = 전체를 보지 못하고 일부분만 보고 리더가 결정되는것이 문제이다

Centrality



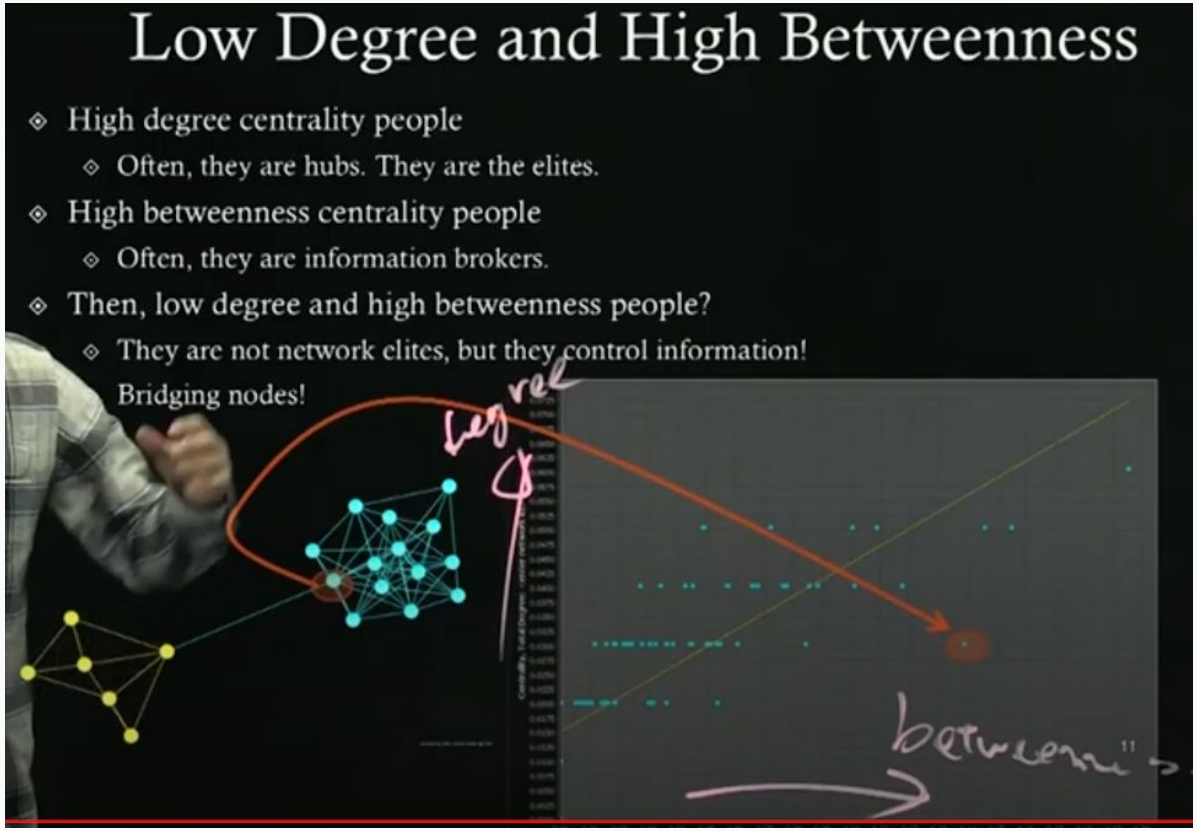
shortest path에 얼마나 등장하는가

특정 노드를 지나가는
all-pair-shortest를 전부 계산하는것
all-pair-shortest는 ex) a-b ,a-c ,a-d ,a-b-e ,a-c-e....

Global measure

shortest path를 구해서 중요도를 통해 정보나 영향력이 전파된다고 생각해서 만든 중심도

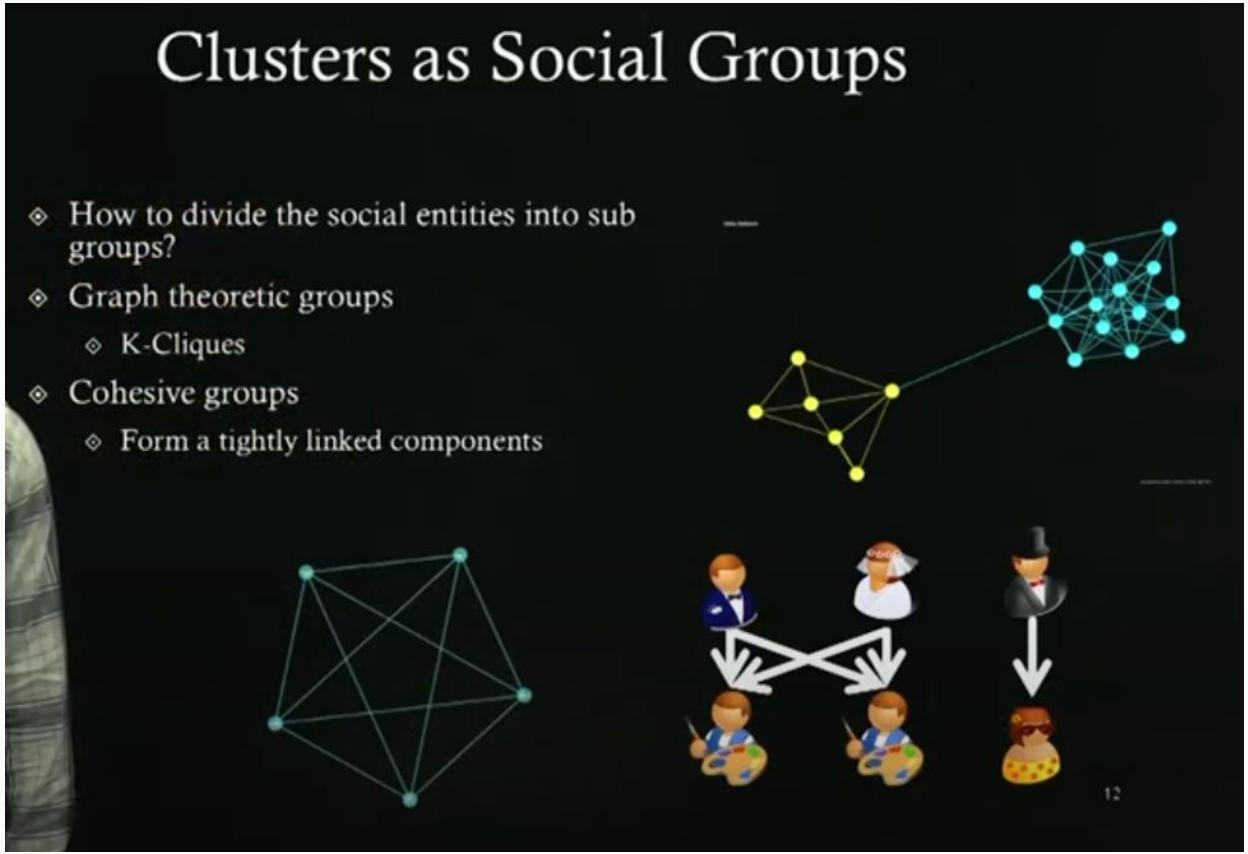
Centrality



높은 degree centrality 사람은 허브의 역활을 한다 또는 엘리트역활

높은 Betweenness centrality 사람은 브로커의 역활을 한다

K Clique and Newman Clustering



서브그룹에 대한 분석
군집화 방법

군집화 방법이 여러가지 있지만
그 중에

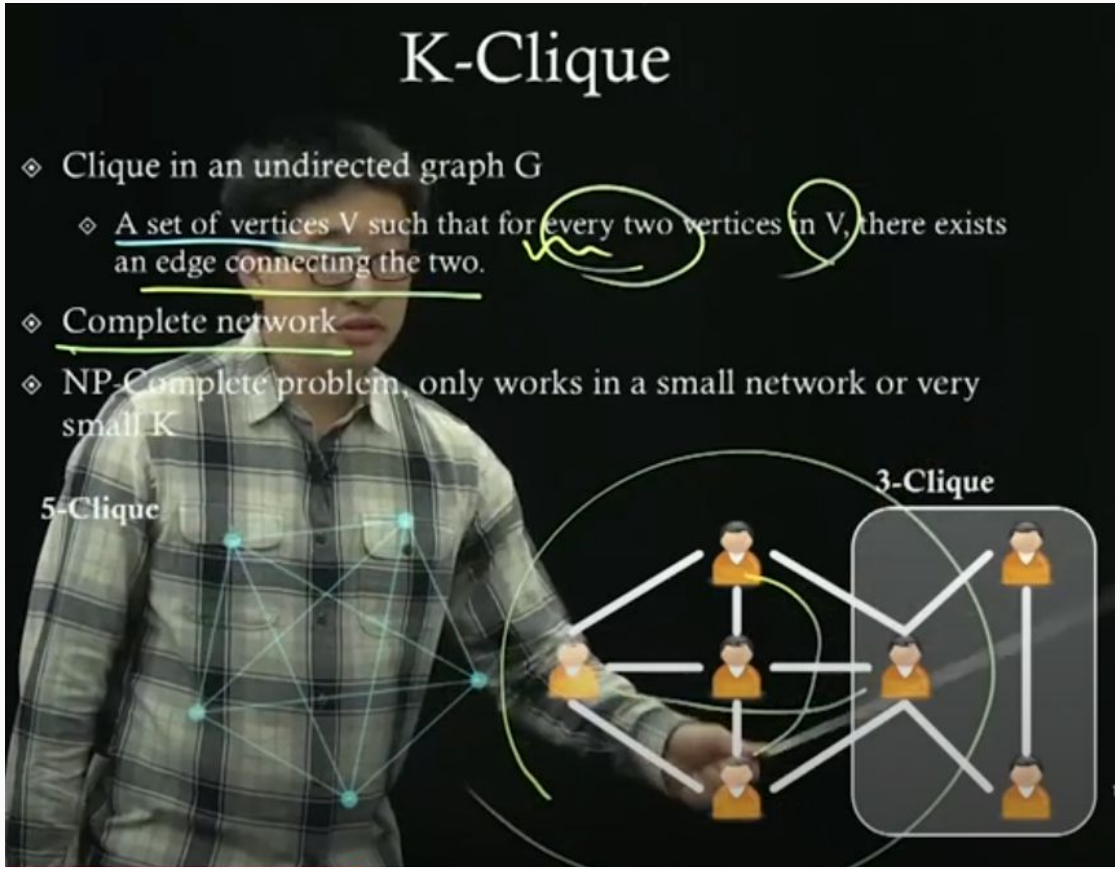
그래프 수열에 많이 쓰이는
군집화 방법
Graph theoretic groups

- K-Cliques

그리고
Cohesive groups(잘 엉겨붙는 집합)

- Newman Clustering

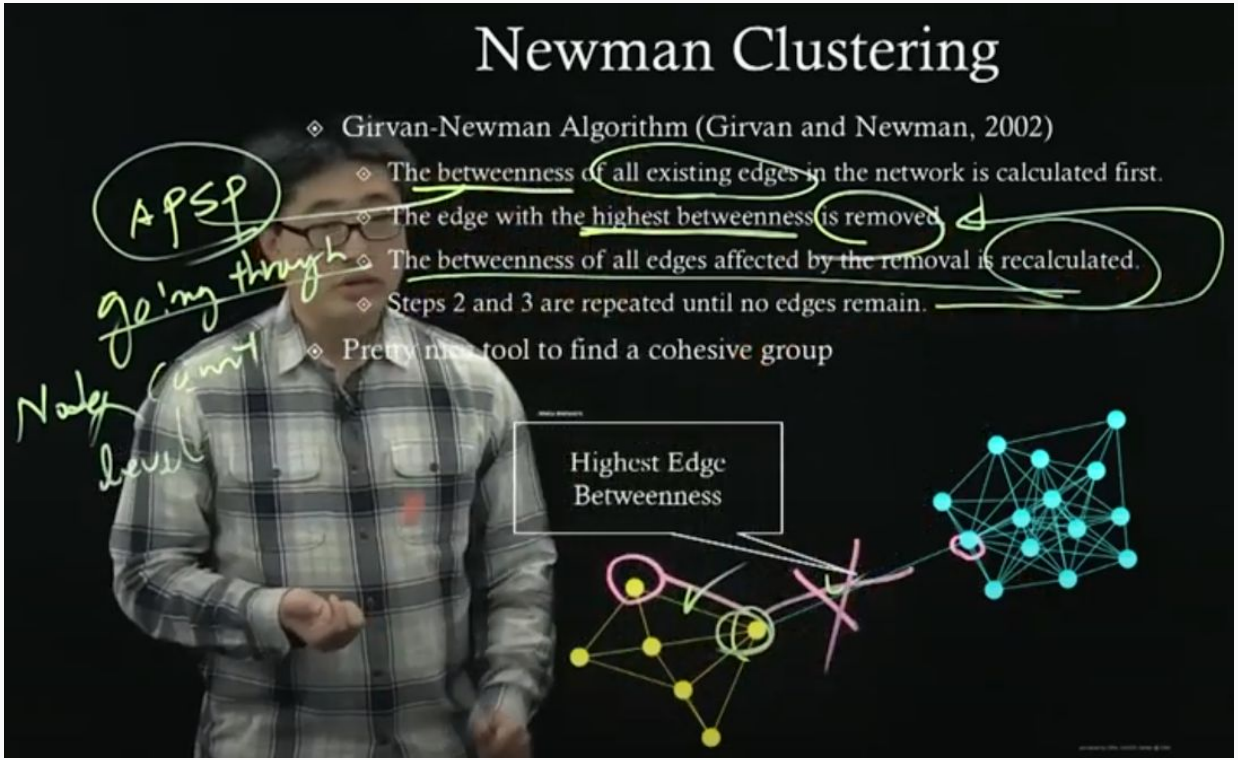
K Clique and Newman Clustering



K-Clique
Clique들을 찾는 군집화 방법

서브셋이 있고 모든 꼭지점들은 서로 edge에 의해 연결되어 있을때 이것을 Clique라 한다

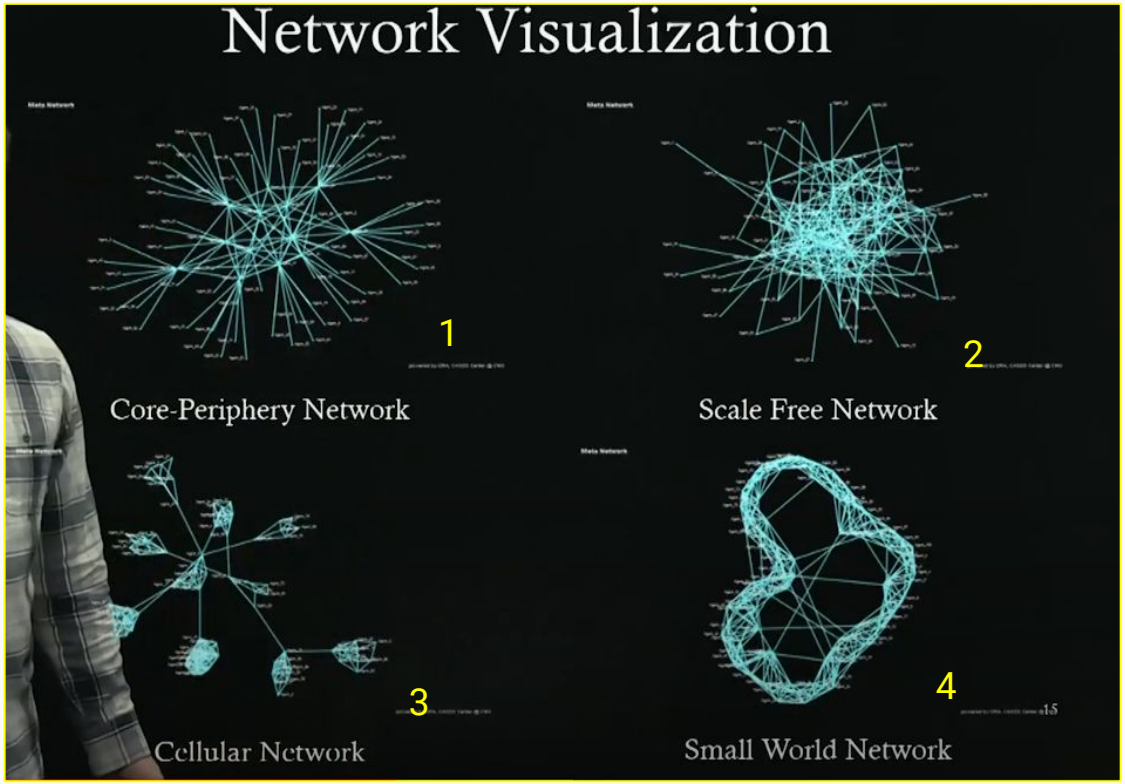
K Clique and Newman Clustering



Newman Clustering Girvan-Newman Algorithm이라고도 부르고 두 개의 다른 컴포넌트로 분리시켜서 서브그룹을 나누는게 Newman Clustering입니다

Edge도 betweenness를 구할 수 있는데 가장 높은 점수의 엣지를 지우고 컴퍼넌트로 나누어진다면 그만두지만 나누어지지 않았다면 다시 betweenness계산하고 다시 엣지 지우고 반복해서 나누는게 newman Clustering 알고리즘

Visualization of Network Structure



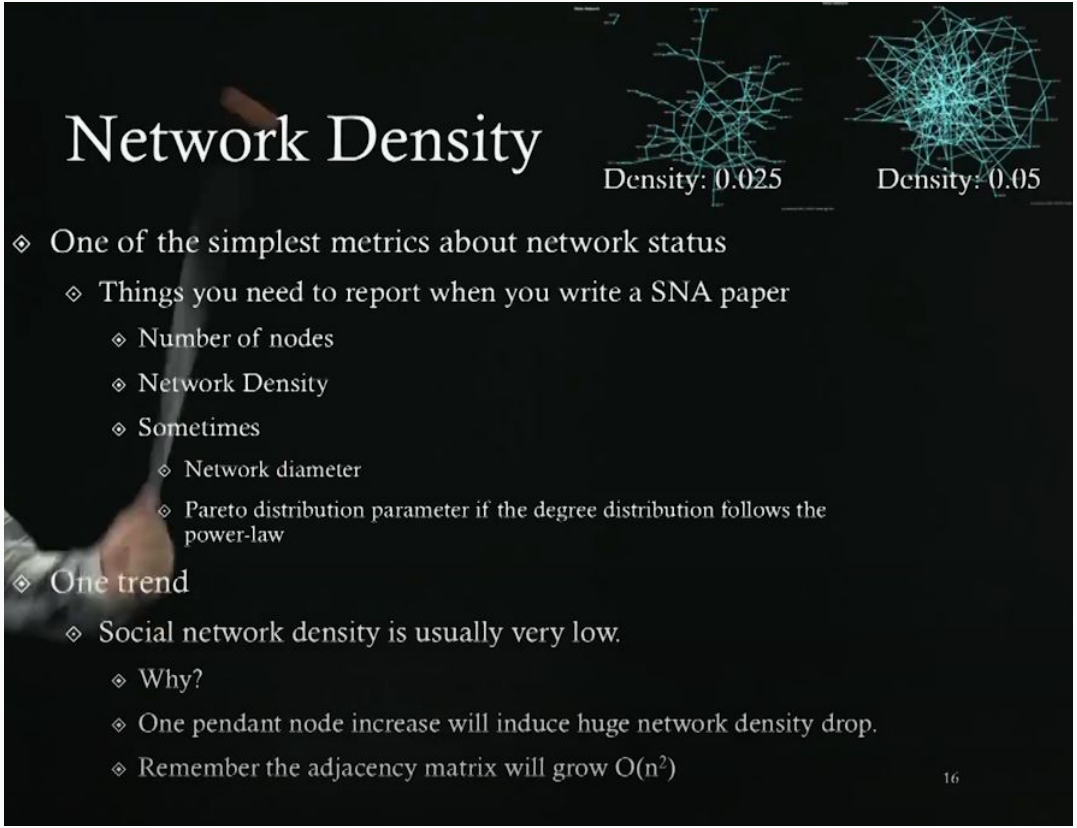
1.주변에 있는 periphery라는 존재들은 코어에만 달려있는 형태 코어들은 서로 잘 엮어있어서 정보를 주고받지만 periphery는 코어에게서만 정보를 얻을 수 있다

2.극단적인 허브가 존재 개별 노드들의 링크가 적다

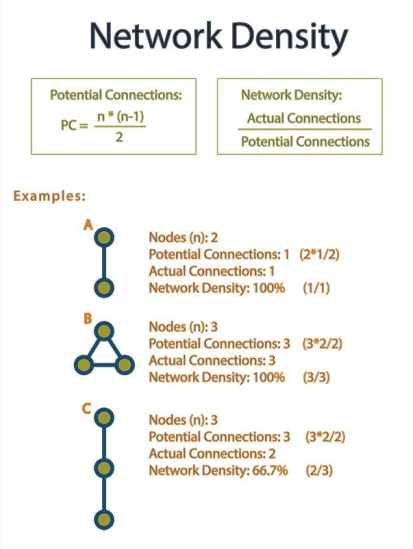
3.세포처럼 결합되있고 점조직으로 이루어져있다

4.주변끼리는 잘 결합 되있지만 멀리있으면 잘 안되있다
하지만 한 두개씩 서로 크로싱 연락하는 존재들이 있는데 네트워크 연결에 큰 역활을 한다

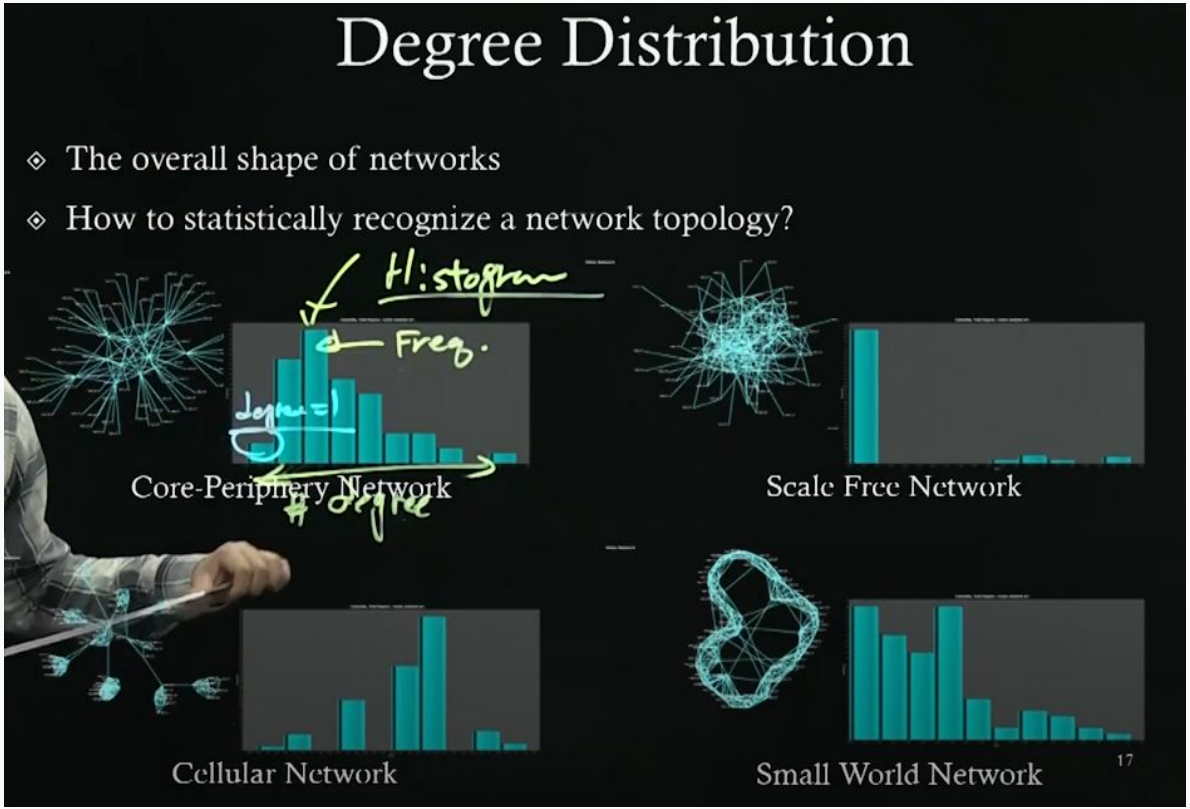
Visualization of Network Structure



Network Density
모양만 보고 정량적인 수치를 알기는 힘드니 Density가 필요하다



Visualization of Network Structure



시각화 , density로 표현 할 수 있지만 둘 다 명확하진 않음
네트워크 구조를 명확하게 표현하고 싶으면 Degree Distribution으로 표현
(히스토그램)

그래프 y축은count x축은 node가 가진 link count