

데이터문제해결1 과제1

컴퓨터공학과 17011599 안정연

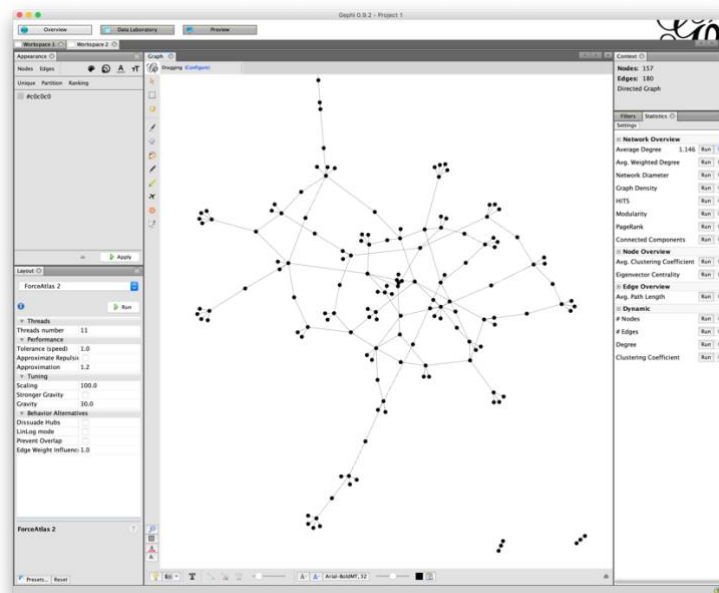
한국 영화계의 감독과 배우들의 관계를 파악해 보기 위해 네트워크를 구현해보았습니다.

1. 데이터 수집

데이터는 최근 흥행한 영화들의 감독과 출연한 배우들의 데이터를 수집하여 스프레드시트에 저장하였습니다.

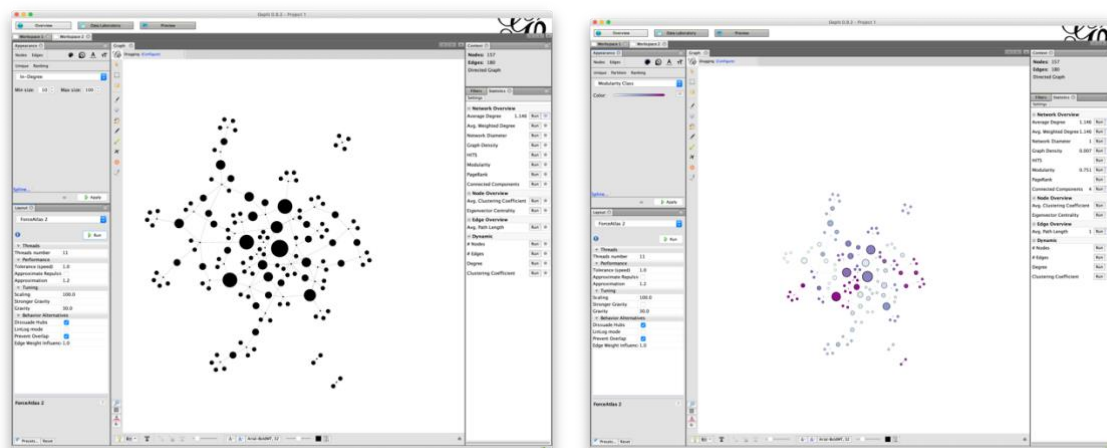
2. Gephi에서 불러오기

Gephi에서 감독에서 배우로 향하는 방향그래프로 다음과 같이 불러왔습니다. Layout을 ForceAtlas2로 설정하고 Scaling을 100으로 하고 Gravity를 30으로 하여 다음과 같이 분포하게 하였습니다.

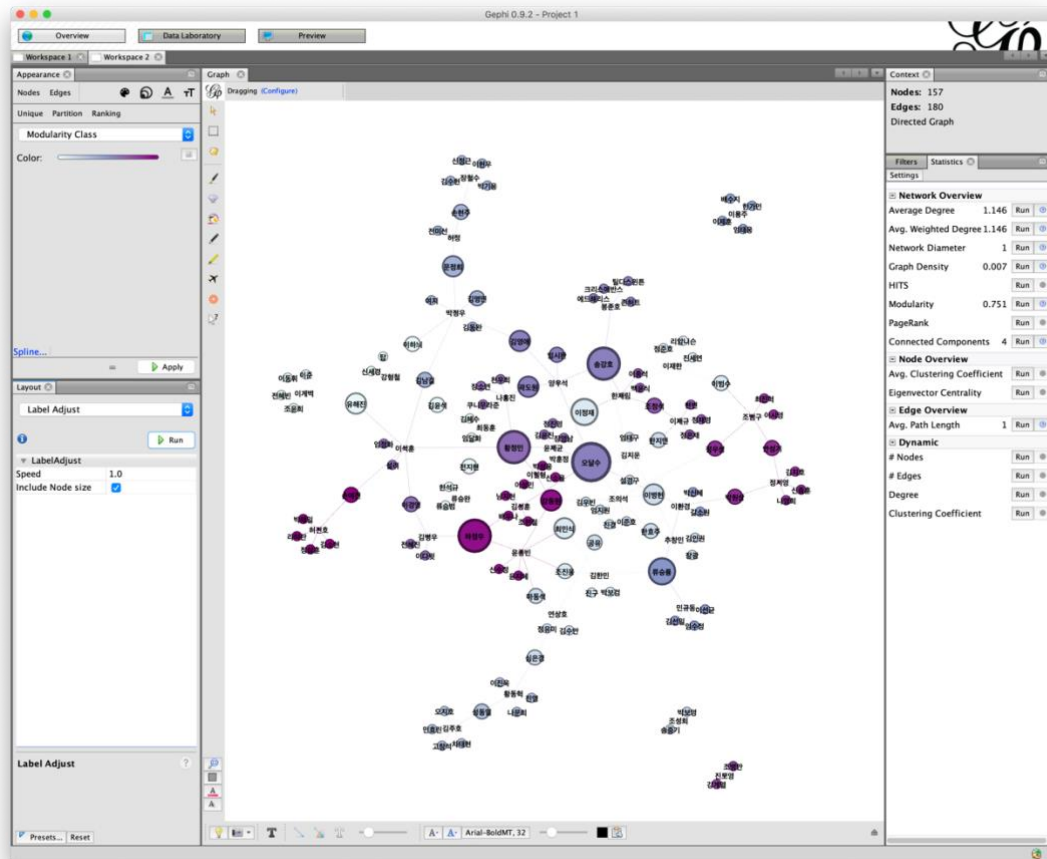


3. 분석 및 시각화

영화에 많이 출연한 배우들을 알기 위해 Node의 in-degree가 클 수록 노드의 사이즈가 크게 만들었습니다. 또한, 소집단 파악을 위해 Modularity Class 기준으로 색상을 적용하였습니다.



마지막으로 노드의 이름을 보이게 해 다음과 같이 시각화 하였습니다.



시각화를 통해 각각의 배우들이 얼마나 많이 영화에 출연하였고 어떤 감독의 작품에 출연했는지 한눈에 알 수 있었습니다. 예를 들어, 송강호 배우의 경우 봉준호 감독, 양우석 감독, 한재림 감독, 김지운 감독의 영화에 출연했음을 알 수 있습니다. 또한, 감독 노드를 누르면 그 감독에 출연한 배우들을 한눈에 알 수 있습니다. 예시로, 이환경 감독의 영화에 오달수, 박원상, 박신혜, 갈소원 배우가 출연했음을 알 수 있습니다.

