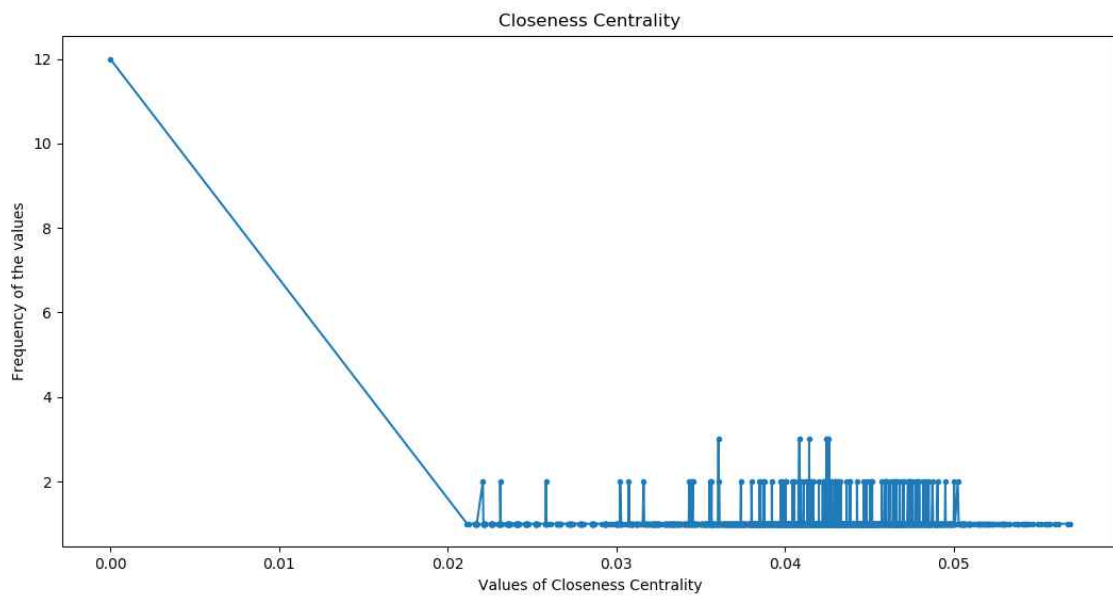
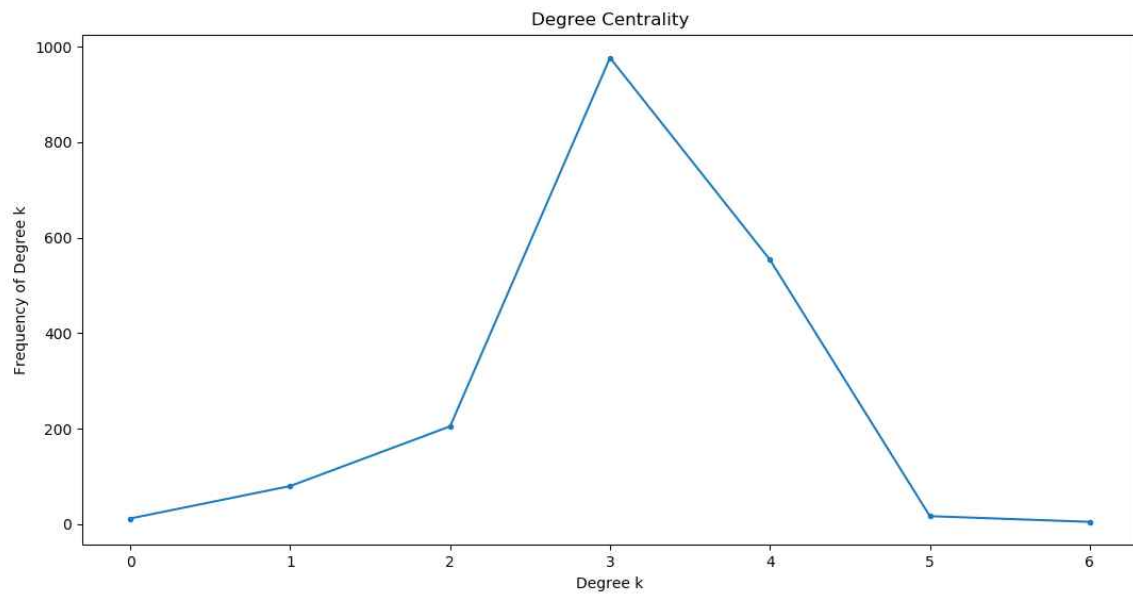


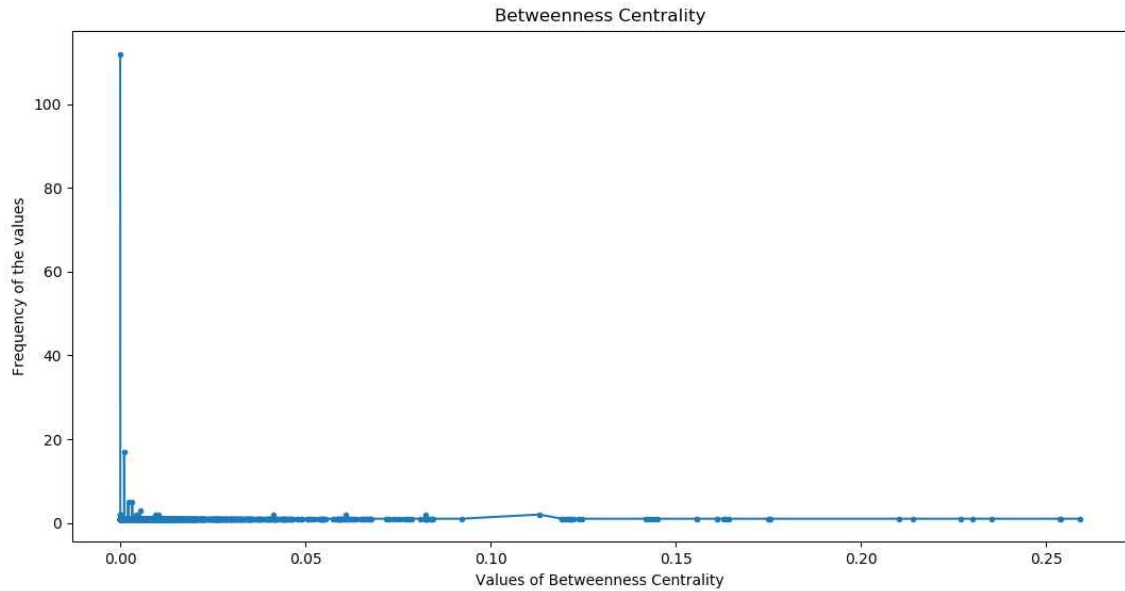
수리적 빅데이터 개론 중간고사 보고서

2014024007

김민기

1. 네트워크 종류





위 그림은 각각 Degree, Betweenness, Closeness Centrality의 분포를 나타낸 그림입니다.

1. Degree Centrality Distribution

Degree 분포를 봤을 때, Degree의 값이 3인 곳에서 가장 큰 빈도를 나타내는 것을 알 수 있고, 이것은 평균이 3인 포아송분포의 모양과 흡사 하다고 할 수 있습니다. 예를 들어, Scale Free Network의 경우와 비교해봤을 때, Scale Free Network 같은 경우, Degree 값이 크게 나타나는 Hub의 역할을 하는 곳의 빈도가 0으로 수렴할 만큼 작고, Degree의 값이 작게 나타나는 곳의 빈도가 기하급수적으로 올라가는 형태를 보이기 때문에, 위 네트워크는 Scale Free Network는 아니라고 할 수 있습니다.

2. Closeness Centrality Distribution

Closeness Centrality Distribution의 분포를 보았을 때, 0.05보다 큰 값을 가지는 노드들의 비율이 0.047을 나타내는데, 이것은 소수의 노드들이 Hub의 역할을 수행하는 것이 아니라, 많은 노드들이 다른 노드로 갈 때 거쳐야하는 위치에 존재한다는 것, 근접성이 높은 역할을 수행하고 있다는 것입니다. 즉, 각각의 노드들이 서로 많이 연결되어 있다는 것을 나타냅니다.

3. Betweenness Centrality Distribution

Betweenness Centrality의 분포를 보았을 때, 0.25보다 큰 값을 가지는 노드의 경우, 다른 노드로 가기 위해, 많은 노드들이 이를 거친다는 것을 의미하는데, 이것의 빈도가 3으로 매우 작아서, Scale Free의 Hub에 해당한다고 생각할 수 있습니다. 하지만, 0~0.1사이에 존재하는 값들이 0.97의 비중을 차지하므로, 대부분의 노드들이 다른 노드로 가기위해 거쳐가야 하는 노드가 존재함을 의미합니다. 따라서, 대부분의 노드들이 서로 연결되어 있음을 의미합니다.

위 세 개 중심성 척도의 결과를 종합해봤을 때, 대구시 도로 네트워크는 Random Graph Network라는 결론을 내릴 수 있습니다.

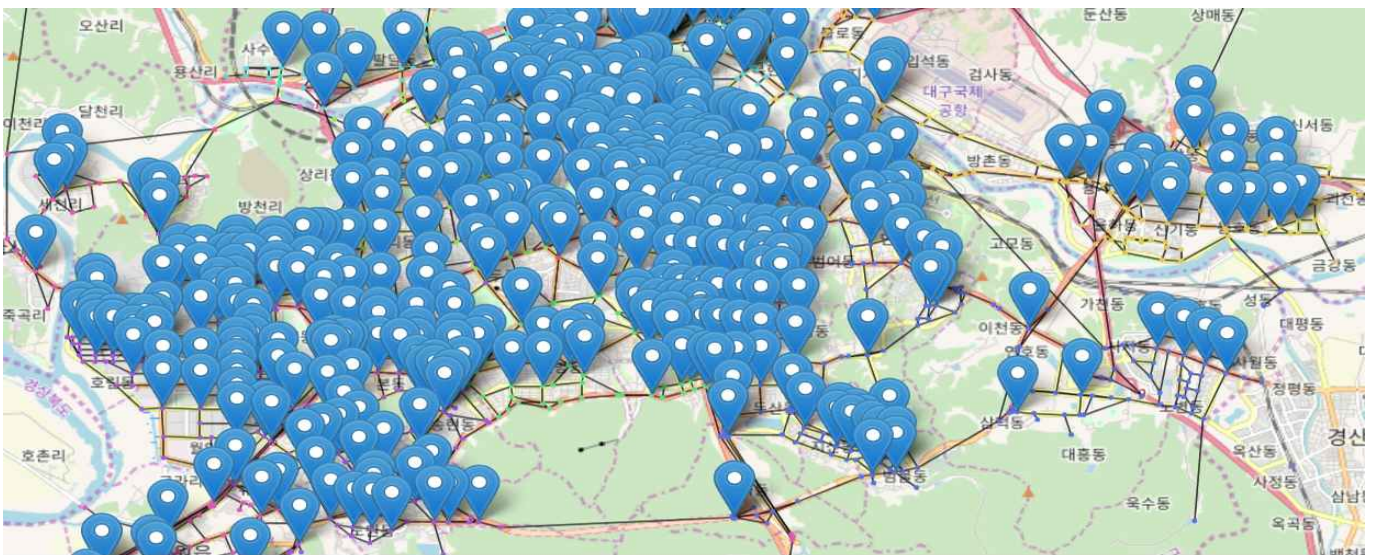
2. Folium을 이용한 시각화와 정체구간 결정 및 활용방안 모색

1) 정체구간 분석

1)경기도 교통정보센터에서 발간한 2013년 교통분석 보고서에 따르면, 소통상태의 분류를 고속도로의 경우 속도가 40km/h 미만이면 '정체', 40km/h~70km/h 사이면 '서행' 그 이상의 속도이면 소통이 원활한 상태라고 분류했고, 일반국도, 국지도, 지방도, 시군도, 기타도로는 20km/h 미만의 속도이면 '정체', 20km/h~40km/h 사이면 '부분정체' 그 이상의 속도이면 소통이 원활한 상태로 분류하고 있고, 정체유형별 특성을 4가지 유형으로 분류하고 있습니다. 유형1은 적은 교통량에도 정체가 발생하는 곳, 유형2는 평균 교통량에 정체가 발생하는 곳, 유형3은 많은 교통량 집중으로 인하여 정체가 발생하는 곳, 유형4는 강설, 강우, 대형사고등으로 발생하는 비반복적 곳입니다.

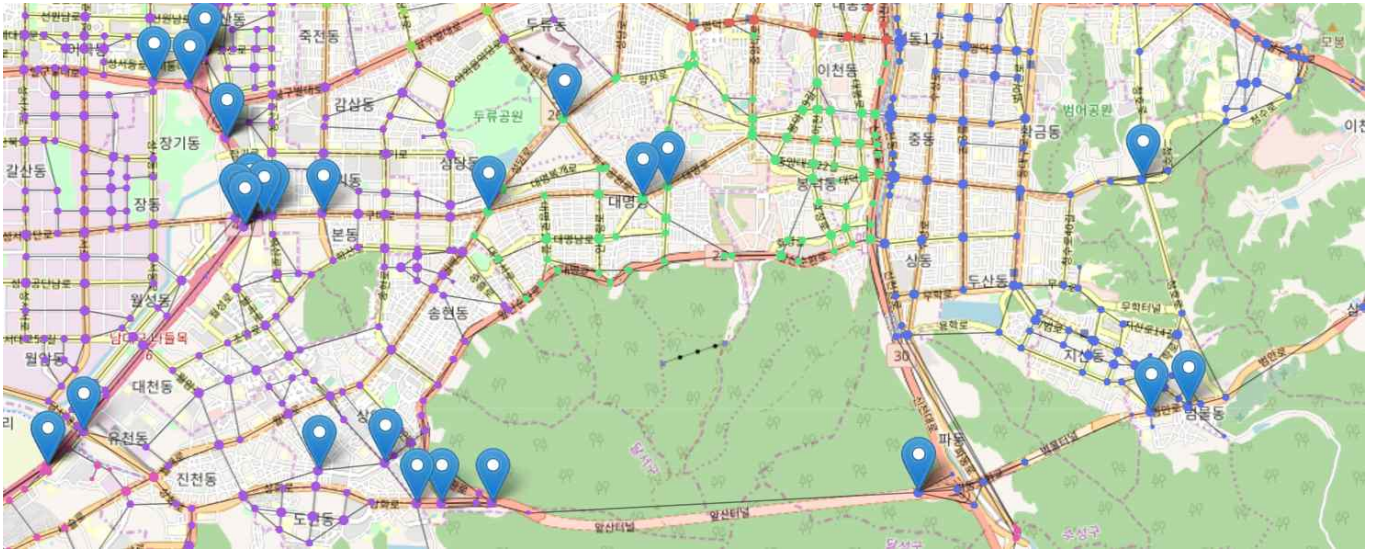
주어진 데이터를 통해 판단할 수 있는 정체유형은 유형3이라고 판단했고, 따라서, 정체구간을 정의할 때, 많은 교통량이 집중되는 곳을 위주로 정체구간이라 판단할 것입니다.

4개의 도로가 교차하는 4거리부터 교통량이 집중되기 때문에, 4거리 이상의 Degree를 가지는 노드를 정체구간이라고 정의할 수 있다고 생각했습니다.



위 그림은 Degree가 4이상의 값을 가지는 곳을 지도에 Marking 한 것입니다. 위 그림에서 볼 수 있듯이, Degree만으로 정체구간을 판단하는 것은 무의미함을 알 수 있습니다. 왜냐하면, Degree의 수가 높은 곳, 즉 많은 도로가 연결된 곳이라고 해서 꼭 정체구간이라고 볼 수 없기 때문입니다.(예를들어, 고속도로 IC 같은 곳) 따라서, 위 분석에서 썼었던 Betweenness Centrality와 Closeness Centrality도 함께 이용하여 정체구간을 판단하기로 결정하였습니다.

첨부된 코드에서 알 수 있듯이, Betweenness Centrality, Closeness Centrality를 고려하기 위해, Betweenness Centrality를 0.05이상, Closeness Centrality를 0.04이상, Degree가 4이상 인 값을 지도에 표시하면 다음과 같습니다.



위 그림에서 볼 수 있듯이, BC와 CC가 고려되어, 다른 곳으로 가는 중간적인 위치를 수행할 수 있는 곳이 지도에 표시된 것을 알 수 있습니다. (사거리이상이면서 다른 곳으로의 연결이 많이 되어 있는 곳) 그리고 대구의 주요 간선도로인 앞산순환로, 달구벌대로, 팔공로, 상화로, 신천동로가 지도에 포함된 것을 알 수 있습니다.

위 결과를 통해 얻어진 대표적인 정체구간은 평리네거리, 이곡역네거리, 본리네거리, 앞산네거리, 두류공원 네거리 등의 네거리입니다. 네거리는 도로 4개가 겹쳐있고, 다른 곳으로 가는 교통의 요지가 되는 경우가 많기 때문에, 필연적으로 교통량이 많아, 정체구간이 될 가능성이 높습니다. 그리고, 서대구IC, 동대구IC, 성서IC 등의 고속도로와 국도를 연결하는 도로의 경우도 교통량이 많이 몰리는 지역이므로, 교통이 정체될 가능성이 높은 구간임을 알 수 있습니다.

2) 정체구간 분석의 활용방안 모색

앞서 언급한 보고서에서는 유형3의 경우 정체관리 전략은 교통량 분산 또는 특송대책에 상응하는 대응체계 구축, 사전홍보 및 실시간 홍보를 통한 유도 등의 우회유도를 언급하고 있습니다.

3) 'Network 도로분석'의 특이점

관습적으로 상습정체구간을 정의할 때, 도로 교통량의 속도로 정의하고 있었습니다. 하지만, 위 분석의 경우 Nodes, Links 즉, Network를 이용하여 도로 정체구간을 분석할 수 있다는 것을 보여주었다는 점에서 다른 정체구간 분석과는 다른 특이점을 찾을 수 있을 것입니다. 도로 교통량의 속도 데이터와 Network 데이터를 결합한다면, 더욱 정확한 대구시 도로 정체구간을 설정, 정체구간의 해소 방안 마련을 할 수 있을 것이라 생각합니다.