

Centrality를 이용한 서울시 교통 밀집도 분석*

이재용^{○1} 이희주¹ 김기환¹ 이상훈¹ 한치근²

경희대학교 컴퓨터공학과¹, 경희대학교 컴퓨터공학부 교수(교신저자)²

ljyh0117@naver.com leehj1621@naver.com pilot1129@khu.ac.kr

a01b01c01@nate.com cghan@khu.ac.kr

Analysis of Traffic Density in Seoul Using Centrality

Jaeyong Lee[○] Heeju Lee Kihwan Kim Sanghoon Lee Chigeun Han

School of Computer Science and Engineering, KyungHee University

요 약

새로운 교통수단을 도입하거나 새로운 노선을 계획하기 위하여 Centrality를 이용한 네트워크 분석들이 많이 연구되었다. 대부분 단순 정류장과 노선만을 이용한 가중 매개 중심성(Weighted Betweenness Centrality)을 이용한 분석들이 많이 연구되었다. 본 논문에서는 정류장과 노선, 그리고 탑승객을 이용한 가중 매개 중심성(Weighted Betweenness Centrality) 지표를 이용하여 정류장간의 경로 중심성을 분석한다. 분석 결과, 정류장을 거치는 노선이 많고 해당 노선에 승객이 많은 경우 중심성값에 많은 영향을 주며, 현재 교통 시스템은 중심성을 효율적으로 분산하지 못한다는 것을 확인하였다. 그리고 높은 중심성을 가진 노선을 기준으로 새로운 노선을 제안하였다. 제안된 결과는 서울시 도로교통에서 교통수단을 도입하거나 새로운 노선을 계획할 시 효율적으로 중심성을 분산하여 승객의 편의성을 향상시킬 수 있다.

1 서 론

새로운 교통수단을 도입하거나 새로운 노선을 계획하기 위하여 Centrality를 이용한 네트워크 분석들이 많이 연구되었다^{[1][2]}. 대부분의 연구에서는 중심성을 계산하기 위하여 서울시의 전체 버스노선과 해당 노선이 지나가는 정류장만을 이용하였다^[3]. 하지만 해당 중심성은 현재 설치되어있는 정보만을 확인하고 현재의 문제점을 찾을 수 없다. 특히 유동인구를 고려하지 않은 중심성 분석은 현재 네트워크의 문제점을 찾지 못하고 방치할 수 밖에 없다^[4]. 기존 연구에서는 버스 노선만을 고려하여 중심성을 분석하였다면, 본 논문에서는 버스 노선과 유동인구인 승객을 이용하여 중심성이 높은 곳을 찾고, 해당 높은 중심성을 주위에 분산할 수 있는 새로운 노선을 제안한다.

중심성에 영향을 주는 요소는 사용한 중심성지표의 성격에 따라 다르다. Degree Centrality의 경우 정류장과 노선수에만 영향을 받지만, Betweenness Centrality의 경우 노선을 이용한 최단거리 확률 및 유동인구에 영향을 받는 등, 중심성 지표의 성격에 따라 다양하다.

본 논문의 구성은 다음과 같다. 2장에서는 네트워크 중심지표를 방식으로 분류하여 설명하고, 3장에서는 그 중 가중 매개 중심성 지표를 이용한 상위 정류장 경로를 분석한다. 4장에서는 상위 정류장 경로의 중심성을 분산하기 위하여 새로운 노선을 제안하고 마지막으로 5장에서는 결론을 맺는다.

2 네트워크 중심성 지표

네트워크의 개별 노드의 특성과 개별 엣지의 특성을 정량화

하고, 엣지의 가중치를 이용한 분석을 하기 위하여 중심성 지표를 사용한다. 네트워크 중심성 지표로는 연결 중심성(Degree Centrality), 고유벡터중심성(Eigenvector Centrality), 매개중심성(Betweenness Centrality), 페이지랭크(PageRank), 근접 중심성(Closeness Centrality), 등이 있다. 이중 자주 사용되는 것이 페이지랭크와 매개중심성이다.

먼저 매개 중심성(Betweenness Centrality)지표는 해당 노드가 중계자 역할을 얼마나 잘하는지를 측정하는 지표이다. 대중교통 네트워크에서 한 정류장에서 다른 정류장으로 가기 위하여 지나는 최단 경로를 구할 때, 해당 경로 위의 정류장이 전체 경로의 수에서 얼마나 많은 확률을 차지하고 있는가를 통하여 도출한다. 다음으로 페이지랭크(PageRank)지표는 노드간 랭크 값을 주고 받게 되는데, 이를 반복하다 보면 전체 노드가 특정한 페이지 랭크 값을 수렴한다. 이러한 사실을 통해 계산하는 지표가 페이지 랭크이다.

본 논문에서는 노드에 가중치를 두어 계산하는 매개 중심성인 가중 매개 중심성(Weighted Betweenness Centrality) 지표를 이용하여 중심성을 계산하고, 효율적으로 중심성을 분산할 수 있는 노선을 제안한다.

3 서울시 버스 네트워크 중심성 지표 계산

본 장에서는 실 데이터를 이용하여 서울시 교통 네트워크의 가중 매개 중심성을 구하였다.

3.1 데이터 수집과 계산 과정

본 연구를 위한 계산 과정으로는 크게 세 단계가 존재하였다. 첫째, 일자, 버스, 정류장 별 승차, 하차 데이터

와 버스, 정류장, 시간 별 1월 승차, 하차 데이터^[5]를 이용하여 일자, 버스, 정류장간 경로, 시간 별 버스의 포화도 계산하였으며 이는 서울시에서 가장 많이 사용되는 디젤 도시저상형 차량을 기준으로 하였다. 해당 차량은 좌석 29석, 입석 26석으로 총 55명의 정원을 가지고 있으며, 55명이 탑승했을 시에 100%의 포화도를 부여하였다. 둘째, 해당 정류장의 경로들을 이용하여 노드인 정류장의 매개 중심성을 구하였다. 셋째, 이렇게 구한 경로의 포화도와 노드의 중심성을 토대로 경로의 중심성을 구하였다.

서울시 교통 데이터는 해당 달이 끝나고 15일 이후에 업로드를 하기 때문에, 본 연구에서는 실질적인 데이터가 존재하는 가장 최근 자료인 2019년 1월 자료를 사용하였다. 모두 611개의 버스노선과 12610개의 정류장, 16157개의 경로를 분석 대상으로 하였으며 중심성 계산을 위한 조합의 수는 203,739,770(=12610*16157)개다.

3.2 승객 수를 이용한 경로 별 포화도

3.1의 데이터를 이용하여 특정 날짜의 특정 시간의 특정 경로에서의 포화도를 계산하였다. 포화도가 높다는 것은 버스가 해당 노선을 지날 때에 버스에 타고 있는 승객 수가 버스를 얼마나 채우고 있는지를 의미한다. 이중 상위 9개를 그림 1에 나타냈다.

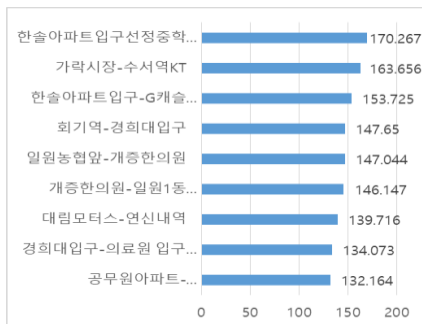


그림 1 Top9 Saturation of the path

한솔아파트입구-갈현동미미아파트가 170.267%로 가장 높았으며, 이 포화도 수치는 승차 및 하차량이 많은 역이 아닌 승차 한 상태로 이동하는 경로인 곳이 포화도가 높음을 표현한다. 그 뒤로 가락시장-수서역KT가 163.656%, 한솔아파트입구-G캐슬아파트가 153.725% 순으로 표현되었다.

3.3 연결관계를 이용한 정류장 매개 중심성(Betweenness)

정류장 중심성은 노드인 정류장과 엣지인 정류장의 연결 관계만을 가지고 매개 중심성 지표를 사용하여 계산하였다. 두 노드를 연결하는 여러 최단경로($\sigma_a(v)$) 속 해당 노드를 통과하는 경로($\sigma_{all}(v)$)의 비율을 임의의 두 노드의 경로에 대해 비율을 합한 수치로 표현된다.

$$g(v) = \sum_{a=1}^{79,499,745} \frac{\sigma_a(v)}{\sigma_{all}} * 100 \quad (1)$$

식 1을 이용하여 계산한 결과가 높다는 것은 해당 정류장을 다른 정류장으로 가기 위해 최단거리 루트에 속한다는 것을

의미하며, 교통상 요지라는 것을 의미한다. 이중 중심성이 높은 상위 9개를 그림 2에 나타냈다.

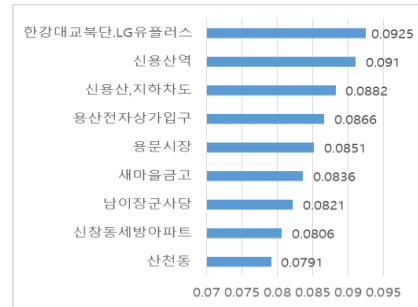


그림 2 Top9 Centrality of the station

한강대교북단-LG유플러스가 0.092로 가장 높았으며, 이는 한강대교북단이 교통의 요지임을 나타낸다. 그 뒤로 한강대교북단과 연결된 신용산역이 0.091, 신용산 지하차도가 0.088 순으로 나타났다.

3.4 포화도와 노드중심성을 이용한 경로 별 가중 매개 중심성

정류장 경로별 가중 매개 중심성은 3.2에서 구한 경로별 포화도 $S(v)$ 와, 3.3에서 구한 정류장 중심성 지표 $g(v)$ 를 이용하여 계산하였다. 정류장은 노드로 구성하고, 엣지는 노선인 정류장의 경로, 가중치는 해당 경로의 포화도로 구성하였다. 앞에서 구한 식 1을 정규화를 위하여 노드중심성의 최대비율과 최소비율을 이용한 이원 연립일차 방정식을 세워 상수값을 구하고, 25를 곱하였다.

$$wbc(v) = \left(\frac{15,000}{14,493} * g(v) + \frac{1,499,237}{1,499,300} \right) * S(v) * 25 \quad (2)$$

경로중심성은 정류장중심성보다는 포화도에 큰 영향을 받는 경향이 있으므로, 본 논문에서는 정류장 중심성의 영향을 적게 조정하고 포화도에 의한 영향을 크게 받게 조정하였다.

식 2를 이용하여 이용한 계산 결과는 본 논문에서 구하고자 한 경로의 중심성을 나타내며, 지역적으로도 중요하고 포화도가 높아 승객의 혼잡성이 예상되는 경로라는 것을 의미한다. 이중 중심성이 높은 상위 9개를 그림 3에 나타냈다.

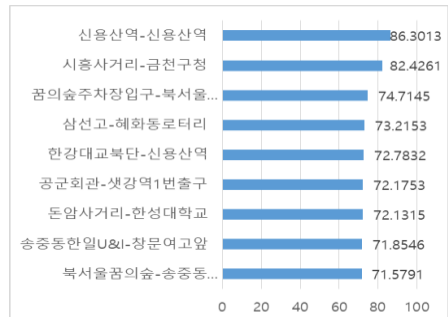


그림 3 Top9 Centrality of the path

신용산역에서 회차하는 경로가 86.30으로 가장 높았으며, 이는 중요하면서도 포화도가 높은 구간임을 나타낸다. 그 뒤로 시흥사거리-금천구청이 82.42, 꿈의숲주차장입구-북서울꿈의숲이 74.71 순으로 나타났다.

4 경로 매개 중심성 지표를 이용한 노선 제안

본 장에서는 도출된 경로중심성을 이용한 새로운 노선을 구성하고, 해당 노선에 의한 중심성 변화를 구하였다.

4.1 매개 중심성 지표 노선 분석

위 3장에서 도출한 경로의 매개 중심성을 이용하여 그래프를 표현하면 다음과 같다. 해당 그래프는 12,610개의 노드와 16,157개의 엣지로 구성되었다.

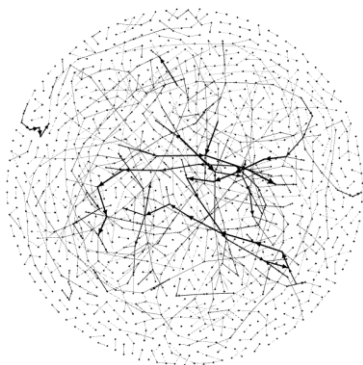


그림 4 Path Centrality Graph

그림 4는 노선의 매개 중심성 지표를 나타내는 그래프이다. 그래프의 노드는 정류장이고, 실제 버스와 같은 방향성을 가지며, 굵기는 경로중심성을 weight로 삽입하였을 시의 그래프를 나타낸다. 중심성이 높은 엣지가 분리되어있지 않고 연결되어있는 것을 확인할 수 있는데, 이는 특정 구간이 포화도가 높은 상태로 이동하면서, 해당 구간이 중요하기 때문에 연결되었음을 알 수 있다.

4.2 노선 제안

4.1에서 도출된 그래프에서, 상위 1,000개의 정류장 중 연결이 되어있으며 높은 중심성을 가진 경로를 연결하여 새로운 노선을 제작하였다.



그림 5 New BusRoute

새로 제안한 노선은 평균 중심성이 45.92이며, 최고 중심성이 74.71인 정류장을 포함하는 노선이다. 해당 노선은 30개의 정류장을 가지고 있으며, 중계 목화 아파트 4단지로부터 시작하여 광화문까지 운행한다. 해당 노선은 운행시간이 왕복 2시간이며, 총 29.34km를 운행한다.

4.3 제안한 노선으로 인한 중심성 변동 분석

4.2절에서 제안한 노선으로 변동된 중심성을 표로 나타내면 다음과 같다.

전 정류장	후 정류장	전 중심성	후 중심성	감소폭
꿈의숲주치장입구	북서울꿈의숲	74.7145	37.3572	-37.3573
상선교,한성대학교	해화동로터리	73.2153	68.9086	-4.3067
돈암사거리,성신여대입구	상선교,한성대학교	72.1315	67.8885	-4.2430
송중동한일유엔아이	창문여고앞	71.8546	35.9274	-35.9272
북서울꿈의숲	송중동한일유엔아이	71.5791	35.7896	-35.7895
오현초등학교	꿈의숲주치장입구	71.2279	35.6139	-35.6140
해화동로터리	명동3가,성대입구	69.3502	65.2708	-4.0794
길음뉴타운	미아리고개,미아리예술극장	65.7160	62.2572	-3.4588
미아리고개,미아리예술극장	돈암사거리,성신여대입구	65.7152	62.5859	-3.1293
송곡초등학교입구	월곡뉴타운	65.0528	48.7896	-16.2632

그림 6 Saturation Decrease

새로운 노선에 속하는 정류장 중, 중심성이 높은 10개 정류장의 감소폭을 확인하였다. 많이 감소한 것은 37.3573이 감소하였고, 가장 적게 감소한 것은 3.1293이 감소하였다. 많이 감소한 것은 해당 경로를 지나는 버스의 수가 적은 경우이고, 적게 감소한 것은 해당 경로를 지나는 버스의 수가 많은 경우이다. 이를 통하여 가상 노선을 신설하여 중심성이 분산 됨을 확인 하였다.

5 향후 연구 및 결론

본 논문에서는 가중 매개 중심성 지표를 사용하여 서울 버스가 정류하는 각 정류장과 경로를 분석하고 이로부터 서울 버스 네트워크와 승객의 흐름을 분석하였다. 그리고 새로운 노선을 제안해 중심성을 분산하여 승객의 혼잡성을 감소시키는 방안도 제안하였다.

향후 연구로는 버스만이 아닌 지하철의 중심성 또한 적용하여 대중교통 전체를 이용한 중심성 지표를 연구할 계획이다.

참고문헌

- [1] S.U.Choi, C.M.Jun, Multi-Layer Network Analysis on the Importance of Bus Lines in Seoul, The Korean Society for Geo-Spatial Information System, Vol.2015 No.9, pp135-140, 2015
- [2] J.W.LEE, K.W.Lee, Analysis of Seoul Metropolitan Subway Network Characteristics Using Network Centrality Measures, Journal of the Korean Society for Railway VOL.20, NO.3, pp413-422, 2017
- [3] G.D.Park, Network Analysis of the Public Transportation System in Seoul, Inha University, 2018
- [4] Freeman, Linton C, Centrality in Social networks Conceptual Clarification, "Social Network, Vol1, NO.3, pp215-239, 1978
- [5] Seoul Open Data homepage : <https://data.seoul.go.kr/>