

다층 네트워크를 활용한 서울시 버스노선 중요도 분석†

Multi-Layer Network Analysis on the Importance of Bus Lines in Seoul

최승우*, 전철민**

Seungu Choi, Chulmin Jun

서울시립대학교 공간정보공학과

ellun7@gmail.com, cmjun@uos.ac.kr

요약

최근 도시 대중교통 정책은 단순 공급위주 정책에서 다수단 연계 통행을 고려한 최적의 비용대비편익을 충족하는 정책 방향으로 전환되고 있다. 대중교통은 네트워크 구조로 표현되는 대표적인 사례로 네트워크 구조를 활용한 많은 분석이 이뤄져왔다. 그러나 두 노드를 연결하는 다양한 노선들로 존재로 단일한 링크로 표현하기 어렵기 때문에, 다양한 교통수단 및 노선 성격을 반영한 분석을 어렵게 한다. 다층 네트워크는 서로 다른 성질을 갖는 링크들을 하나의 층으로 구성하여 기존 단일 네트워크 개념을 확장한 구조로, 다층 네트워크를 이용하여 노선별로 성질이 다른 대중교통 네트워크를 단일 네트워크보다 정확하게 표현하고 분석할 수 있다. 본 연구에서는 서울시 버스노선 및 교통카드 통행 데이터를 이용하여 노드 매개 중심성(Betweenness Centrality)을 통해 각 버스노선이 승객들이 최단거리로 이동하는 데 미치는 영향력을 기준으로 노선 중요도를 분석하였다.

1. 서론

최근 글로벌 경제위기와 기존 대중교통 체계의 운영적자 문제 등으로 인하여 대중교통 정책이 도시구조와 통행 패턴 특성을 검토하여 최적의 비용대비 편익(B/C)을 만족하는 대중교통 시스템을 설계하는 방향으로 패러다임이 전환되고 있다. 따라서 다양한 교통수요에 부응하기 위해 기존의 버스·지하철에 더하여 BRT, 트램, 경전철, 모노레일 등 신 교통수단을 도입한 세분화된 교통 시스템이 고려되고 있다. 따라서 대중교통 분석의 중요도가 높아지고 있다. 특히 2004년 수도권 대중교통 통합요금제가 시행된 이래 전국의 주요 도시에서 환승 통행이 보편화 되었으

며, 선진국의 많은 도시에서도 신 교통수단을 결합하여 다수단(multimodal) 대중교통 이용을 고려한 교통시스템을 구성하고 있어 교통수단, 노선특성 및 환승통행을 고려한 대중교통 분석의 필요성이 높아지고 있다.

대중교통은 정해진 노선을 따라 이동하며 특정 정류소에서만 승객들이 승하차할 수 있는 특성이 있기 때문에 노드는 정류장으로, 링크는 노선의 이동경로로 구성할 수 있는 대표적인 네트워크 구조 중 하나로[1], 대중교통 분석에 네트워크 구조를 많이 활용하고 있다.

그러나 대중교통은 같은 구간을 연결하는 다양한 성질의 교통수단 및 노선이 존

† 본 연구는 2014년 정부(국토교통부)의 제원으로 공간정보 융복합 핵심인재 양성사업의 지원을 받아 수행된 연구임(2014-04-10)

본 연구는 2012년 정부(교육과학기술부)의 제원으로 한국연구재단의 지원을 받아 수행된 연구임(NRF-2012S1A5B8A03045234)

* 서울시립대학교 공간정보공학과 석사과정

** 서울시립대학교 공간정보공학과 교수 (교신저자)

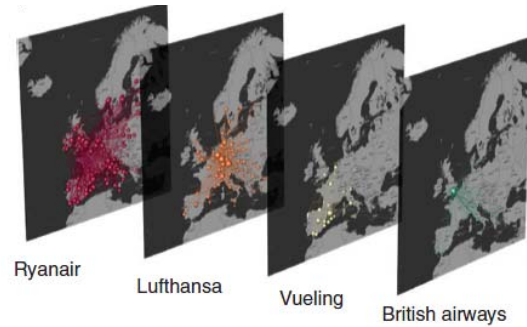
재하기 때문에, 노선 숫자만큼 링크를 구성해야 하며 동일한 노드에 연결된 링크의 경우라도 환승을 통해 연결되는 경우를 고려해야한다. 그러므로 기존의 단일 네트워크로는 다양한 특성을 가진 대중교통 수단 및 노선 특성을 반영하기 어렵고 [2], 분석결과가 올바르게 못한 결과를 가져올 수 있다[3].

이러한 단일 네트워크의 단점을 극복하기 위하여 기존 네트워크 구조를 확장하여 서로 다른 성질을 가지는 네트워크를 하나의 층으로 하고 여러 개 층에서 공유되는 노드를 통해 다른 층의 네트워크와 연결되는 다층 네트워크(Multi-Layer Network, Multiplex Network)의 개념이 제시되었으며, 기존 네트워크 구조를 분석할 때 사용되었던 다양한 지표 및 네트워크 성질을 다층 네트워크에서 계산할 수 있게 하는 방법론들이 연구되었다. 성질이 다른 링크로 연결된 네트워크는 기존의 단일 네트워크로 표현 및 분석이 어렵기 때문에 연구하는 사례에 따라 연구자가 개별적으로 네트워크 모델과 분석방법을 개발해야 하는 어려움이 있었으나, 다층 네트워크를 활용함으로써 다양한 성질을 갖는 네트워크에 대하여 새로운 개념의 추가나 수정이 필요없이 보편적으로 활용되는 네트워크 분석 프레임워크를 그대로 이용할 수 있는 장점이 있다.

이에 본 연구에서는 다층 네트워크 프레임워크를 이용하여 서울시 버스노선을 다층 네트워크로 표현하고, 각각의 버스노선에 대하여 전체 버스 네트워크에서 제외하였을 때 노드 중심성 지표 중 매개 중심성(Betweenness Centrality)이 감소하는 정도를 계산하여, 기존 승객의 최적통행에 해당 노선이 얼마나 기여하는 지 확인하고 버스 노선의 중요도를 분석한다.

2. 다층 네트워크 구조

다층 네트워크는 동일한 성질의 연결로 구성된 네트워크를 하나의 레이어로 하고, 레이어 간 노드들의 연결관계를 표현하여 복합적인 구성된 복잡계(Complex Systems)를 기술한다.



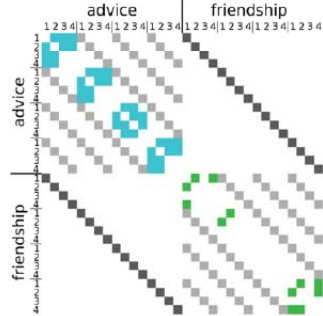
(그림 1) 유럽 항공사별 항공노선 다층 네트워크(Source : figure 1 of [3])

네트워크에서 노드의 연결 상태는 일반적으로 인접성 행렬(Adjacency Matrix)로 표현하며, 네트워크 내의 노드의 수가 N 개 일 때 인접성 행렬의 크기는 $(N \times N)$ 이며, 가중치가 없는 네트워크에서 노드 i, j 가 연결되어 있을 때 인접성 행렬의 성분 $a_{ij} = 1$ 로, 연결되지 않을 때 $a_{ij} = 0$ 으로 정의한다.

일반적인 단일 네트워크에서는 2차원 인접성 행렬로 노드 간 연결상태가 표현되나, 다층 네트워크에서는 노드가 레이어 별로 구분되므로 다층 네트워크에서 인접성은 α -레이어의 노드 i 에서 β -레이어의 노드 j 로 연결되는 것으로 정의된다. 따라서 시·종점 노드의 Index 2개의 정보에 레이어 정보가 추가되어 다층 네트워크에서 인접성은 Rank-4 인접성 텐서 $M_{ij}^{\alpha\beta}$ 또는 $M_{ij}^{\alpha\beta}$ 로 표기한다[3,4].

인접성 텐서는 4차원으로 표현되기 때문에 직관적으로 이해하기 어렵다는 문제가 있어, 종종 2차원 형태로 평활화(flattening)하여 표현하는데 이것을 초인접성 행렬(Supra-adjacency Matrix)이라고 하며(수식 1)[3-5], 레이어 간 인접성 행렬 A_{nm} 을 순서에 따라 2차원 행렬 모양으로 펼쳐진 모습을 한다(수식 1). 초인접성 행렬에서 $A_{11}, A_{22} \dots$ 와 같이 $n=m$ 인 대각성분은 레이어 내의 연결상태(intra-layer connections)를 나타내는 행렬이며, $n \neq m$ 인 그 외의 성분은 레이어 간 연결상태(inter-layer connections)를 나타내는 행렬이 된다.

$$M = \begin{pmatrix} A_{11} & A_{12} & \cdots & A_{1m} \\ A_{21} & A_{22} & \cdots & A_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ A_{m1} & A_{m2} & \cdots & A_{mm} \end{pmatrix} \quad (1)$$



(그림 2) 소셜 다층 네트워크의 초 인접성행렬(Source : fig 4 of [5])

(그림 2)에서 소셜 네트워크는 친분관계(Friendship)와 조언관계(Advice) 2개 레이어로 구성된 네트워크이며, 레이어가 2개이므로 $2 \times 2 = 4$ 개의 인접성 행렬이 포함되어 있다. 왼쪽상단과 오른쪽하단의 인접성 행렬은 각각 조언관계와 친분관계 내의 노드 간 연결상태를 보여주며, 오른쪽상단과 왼쪽하단의 인접성 행렬은 조언관계 ↔ 친분관계 레이어 사이의 노드의 연결성을 보여준다.

3. 분석 방법

네트워크에서 각 노드가 얼마나 중요한 위치를 가지는 지를 나타내는 지표로 중심성(Centrality)이 있다. 중심성은 다양하게 정의되며, 연결 중심성(Degree Centrality), 근접 중심성(Closeness Centrality), 매개 중심성(Betweenness Centrality), 고유벡터 중심성(Eigenvector Centrality), 페이지랭크(PageRank) 등이 있다[6,7].

이 중 매개 중심성은 두 노드를 연결하는 여러 최단경로(σ_{st}) 중 해당 노드 v 를 통과하는 경로($\sigma_{st}(v)$)의 비율을 자신을 제외한 나머지 모든 임의의 두 노드의 경로에 대하여 해당 비율을 합한 수치이다[8].

$$g(v) = \sum_{s \neq v \neq t} \frac{\sigma_{st}(v)}{\sigma_{st}} \quad (2)$$

매개 중심성이 높은 노드는 네트워크 상에서 이동하는 객체가 해당 노드를 많이 경유한다는 의미로 다른 노드들을 연결해주는 중간 다리(Bridge)로서 역할을 수행한다고 할 수 있다. 매개 중심성은 정보나 전염병의 전파력, 교통에서의 통행량, 교통체증과 밀접한 연관이 있다[9,10].

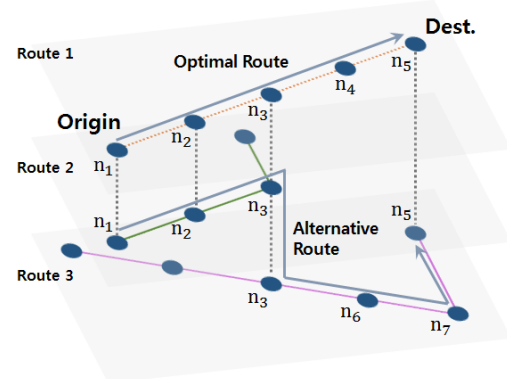
대중교통을 이용하는 승객은 대체로 최단시간 내에 목적지에 도착하기 위하여 최적경로를 경유하는 노선을 선택하여 이동한다고 할 수 있다. 이 때 전체 버스 네트워크에서 버스 노선 하나를 제외할 경우 해당 노선을 이용하여 최적경로로 이동하는 승객들은 다른 경로로 우회해야 한다. 승객들이 우회하여 이동하면 제외한 버스 노선이 통과하던 노드는 그만큼 매개 중심성이 감소하므로, 하나의 버스노선을 제외함으로써 해당 노선을 통과하는 노드 전체의 매개 중심성 감소량 Δ_r 을 구할 수 있다.(수식 3)

$$\Delta_r = \sum_{n \in N_r} (C_n - C'_n) \quad (3)$$

(N_r : 노선 r 을 통과하는 노드 집합)

C_n : 노드 n 의 매개 중심성

C'_n : 노선 r 을 네트워크에서 제외하였을 때 노드 n 의 매개 중심성)



(그림 3) 한 노선을 제외하였을 때 최단경로 변화 예시

(그림 3)의 예에서 노드 n_1 에서 n_5 를 이동하는 경로가 있다고 하면 원래 네트워크에서 최적 경로는 Route 1을 따라 이동하는 $n_1 \rightarrow n_2 \rightarrow n_3 \rightarrow n_4 \rightarrow n_5$ 이며, 통과 노드인 n_2, n_3, n_4 는 매개 중심성이 높아진다.

그러나 Route 1을 네트워크에서 제외한 후 노드 n_1 에서 n_5 를 이동하는 최적 경로는 Route 2, Route 3을 따라 이동하는 $n_1 \rightarrow n_2 \rightarrow n_3 \rightarrow n_6 \rightarrow n_7 \rightarrow n_5$ 가 되므로 노드 n_4 는 해당 경로의 최적 경로를 통과하는 노드에서 제외됨으로써 매개 중심성이 감소하게 된다.

매개 중심성은 일반적으로 많은 노선이 통과하는 노드일수록 커지므로 주요 간선 도로를 통과하는 노선들이 이에 비례하여 전반적으로 높은 매개 중심성 감소량을 나타내게 된다. 따라서 노드마다 매개 중심성 감소량을 원래 네트워크에서 구한 매개 중심성으로 나누어, 해당 버스노선이 노드마다 승객들의 최단거리 이동에 얼마나 기여하는 지의 비율을 이용하여 노선의 중요도를 계산하는 방법도 고려할 수 있다.

$$I_r = \sum_{n \in N_r} \left(\frac{C_n - C_n^r}{C_n} \right) \quad (4)$$

위의 두 방법으로 구한 매개 중심성 감소량과 감소비율은 순수하게 네트워크의 구조적 특성만을 반영하며 실제 해당 노선을 이용하는 수요를 반영하지 않는다. 따라서 해당 노드 통과 수요를 반영하기 위해, 교통카드 통행 데이터를 이용하여 서울시 버스를 이용하는 승객의 이동경로를 통해 경유 노드를 탐색하여 각 노드마다 승객들의 통과량을 계산하였다. 이 통과량을 가중치로 하여 매개 중심성 감소 비율에 곱함으로써 수요를 고려한 노선 중요도를 구할 수 있다.

$$I'_r = \sum_{n \in N_r} \left(P_n \times \frac{C_n - C_n^r}{C_n} \right) \quad (5)$$

(P_n : 노드 n 의 승객 통과량)

매개 중심성을 구하는 방법으로 시간 복잡도가 노드 수와 링크 수를 곱한 $O(V \times E)$ 인 Brandes 알고리즘[11]을 다층 네트워크로 확장한 알고리즘이 제안되어[12], 해당 방법으로 매개 중심성을 구하였다.



(그림 4) 서울시 버스정류장별 승객 통과량



(그림 5) 서울시 버스정류장별 매개 중심성

분석 대상은 서울시 버스(마을버스 포함) 599개 노선을 대상으로 하였으며, 버스 정류장을 하나의 노드로, 하나의 노선이 이동하는 경로를 링크로 다층 네트워크를 구축하였다. 버스 네트워크에서 노드의 수는 15,882개, 링크 수는 35,051개이나, 버스 정류장은 방향에 따라 또는 노선성격에 따라 비슷한 지점에 버스정류장이 다수로 분리되어 분포하므로 네트워크의 정합성과 연산 속도의 증대를 위하여 인근에 위치한 버스정류장들을 통합하여 노드 6,041개, 링크 33,054개의 보정된 버스 네트워크를 분석에 활용하였다. 또한 버스 정류장 및 노선 정보, 승객 통행 데이터는 2011년 10월 19일 수요일 서울시 교통카드 데이터를 활용하였다.



(그림 6) 서울시 버스 네트워크

4. 분석 결과

분석 결과는 노드 통과량 가중치를 고려하였을 때 매개 중심성의 절대 감소량/매개 중심성 감소 비율의 두 가지 경우로 결과를 도출하였다.

(표 1) (노드 통과량)×(중심성 감소량) 기준 상위 20개 노선

버스 노선	감소량
150번(도봉산-석수역)	2.7396E+9
262번(중랑차고지-여의도)	2.3402E+9
152번(화계사-삼막사사거리)	2.2504E+9
163번(월계동-목동)	2.1523E+9
108번(양주 덕정-종로5가)	1.9985E+9
720번(기차촌-답십리)	1.8422E+9
149번(하계동-서빙고)	1.8292E+9
461번(장지공영차고지-여의도)	1.7905E+9
9403번(성남분당-을지로5가)	1.7202E+9
143번(정릉-개포동)	1.6554E+9
462번(송파차고지-영등포)	1.5570E+9
160번(도봉산-운수동)	1.5076E+9
260번(중랑차고지-양천차고지)	1.4668E+9
370번(강동공영차고지-은평공영차고지)	1.3107E+9
703번(문산향양리-서울역 환승센터)	1.2941E+9
6411번(구로동-개포동)	1.2889E+9
271번(면목동-상암동)	1.2852E+9
421번(염곡동-옥수동)	1.2602E+9
3412번(강동공영차고지-우면동)	1.2587E+9
140번(도봉차고지-송파차고지)	1.2485E+9

(표 2) (노드 통과량)×(중심성 감소비율) 기준 상위 20개 노선

버스 노선	감소량
461번(장지공영차고지-여의도)	263,253.60
152번(화계사-삼막사사거리)	262,251.32
163번(월계동-목동)	257,425.03
108번(양주 덕정-종로5가)	249,416.79
262번(중랑차고지-여의도)	234,900.08
9403번(성남분당-을지로5가)	203,796.27
150번(도봉산-석수역)	198,182.15
7212번(은평차고지-옥수동)	197,088.45
3412번(강동공영차고지-우면동)	191,154.14
146번(상계주공7단지-강남역)	183,131.82
149번(하계동-서빙고)	182,684.17
7211번(기차촌-신설동)	178,551.01
143번(정릉-개포동)	177,410.79
571번(가산동-기차촌)	172,925.32
2016번(중랑차고지-효창동)	169,721.85
강서05(강서면허시험장-발산역)	167,400.48
363번(송파차고지-여의도)	167,354.91
6515번(양천차고지-삼막사거리)	162,740.16
6411번(구로동-개포동)	162,481.63
760번(금촌-영등포)	156,462.66

분석 결과 노선이 길고 정류장 수가 많은 노선이 대체로 중요도 상위를 차지하였다. 특히 (표 1)에서는 중심성 감소량의 크기를 곱한 영향으로 매개 중심성이 높은 도봉로, 종로, 강남대로, 통일로, 한강대로 등 주요 대로 연선을 운행하는 노선들이 중요도가 높은 노선으로 나타났다. 그러나 (표 2)에서는 중심성 감소비율을 기준으로 나타난 순위로 노선이 지나가는 노드의 중심성이 대체로 낮더라도 중심성이 높은 주요 지점을 연결해주면서 다른 노선과 중복도가 낮아 대체하기 어려운 노선이 비교적 높은 순위를 나타냈다. 461번, 146번, 7211번, 571번, 2016번, 363번, 6515번 버스 등은 주요 간선도로보다는 지선도로를 위주로 운행하면서 중요한 지점을 연결해주는 노선의 성격에 가까운데, (표 1)에서는 상대적으로 낮거나 순위권 외에 있었으나 (표 2)에서는 높은 순위를 나타냈다.

5. 요약 및 결론

대중교통 체계가 발전하면서 다수단 환승이동을 분석할 필요성이 높아졌다. 복잡한 대중교통 네트워크를 표현하기 위해 다층 네트워크로 버스 네트워크를 구축하였으며 노선 중요도 분석 사례를 통하여 다층 네트워크를 이용한 대중교통 분석 방법을 제시하였다.

그러나 이번 분석에서 활용한 네트워크는 가중치가 없는 네트워크(unweighted networks)로, 모든 연결성이 동등한 가중치를 가지고 있기 때문에 노선별 이동시간이나 환승시간이 제대로 고려되지 않았다는 한계점이 있다. 향후에는 수도권 대중교통에 대하여 이동시간과 환승시간 등을 포함한 가중치 네트워크로 분석을 시행할 경우 더 정확한 결과가 도출될 것으로 기대된다.

참고문헌

- [1] Y. Dong, X. Yang, G. Chen, "Robustness Analysis of Layered Public Transport Networks Due to Edge Overload Breakdown", *Information Technology and Computer Science*, Vol. 2014-3, pp. 30-37, 2014.
- [2] J. Park, S. Gang, "A Model for Evaluating the Connectivity of Multimodal Transit Networks", *Korean Society of Transportation*, Vol. 28(3), pp. 85-98, 2010.
- [3] De Domenico et al., "Ranking in Interconnected Multilayer Networks Reveals Versatile Nodes", *Nature Communications* 6, 6868, 2015
- [4] S. Boccaletti et al., "The Structure and Dynamics of Multilayer Networks", *Physics Reports* 544, pp.1-122, 2014
- [5] M. Kivelä et al., "Multilayer Networks", *Journal of Complex Networks* Vol. 2014-2, pp. 203-271, 2014
- [6] S. Mishra, T. F. Welch, M. K. Jha, "Performance Indicators for Public Transit Connectivity in Multi-modal Transportation Networks", *Transportation Research Part A*, Vol. 46, pp. 1066-1085, 2012
- [7] L. Page, S. Brin, R. Motwani, T. Winograd, "PageRank Citation Ranking: Bringing Order to the Web", 1999
- [8] M. Barthélemy, "Betweenness Centrality in Large Complex Networks", *The European Physical Journal B*, Vol. 38(2), pp. 163-168, 2004
- [9] J.W. Joubert, K.W. Axhausen, "A Complex Network Approach to Understand Commercial Vehicle Movement", *Transportation*, Vol. 40(3), pp. 729-750, 2013
- [10] J.A.M. Gonçalves, L.S. Portugal, C.D. Nassi, "Centrality Indicators as an Instrument to Evaluate the Integration of Urban Equipment in the Area of Influence of a Rail Corridor", *Transportation Research Part A*, Vol. 43, pp. 13-25, 2009
- [11] U. Brandes, "A Faster Algorithm for Betweenness Centrality", *Journal of Mathematical Sociology* Vol. 25(2), pp. 163-177, 2001
- [12] A. Solé-Ribalta, M. De Domenico, S. Gómez, A. Arenas, "Centrality Rankings in Multiplex Networks", *WebSci '14 Proceedings of the 2014 ACM conference on Web Science*, pp. 149-155, 2014