

ITS 빅데이터를 활용한 도시 교통네트워크 구조분석*

Analysis of Urban Traffic Network Structure based on ITS Big Data

김용연¹ · 이경희² · 조완섭^{3†}

(주)엔코아¹, 충북대학교 빅데이터학과², 충북대학교 경영정보학과³

요 약

ITS(Intelligent Transport Systems)는 시민들의 교통이용 안전과 편의를 도모하고 교통 시스템의 효율적인 운영 및 관리를 위해 대도시를 중심으로 도입되었다. 우리나라의 경우 ITS가 전국적으로 확대되면서 도로소통상황, 교통량, 대중교통운영현황 및 관리상황, 대중교통이용현황 등 다양한 교통정보가 생성되고 있다. 본 논문에서는 ITS에서 수집되는 데이터 중 하나인 DSRC(Dedicated Short Range Communications) 빅데이터를 활용하여 도시 교통구조를 네트워크 분석 기법을 통해 규명한다. 이를 통해 도심에서의 복잡한 교통현상을 단순화시키고, 차량 흐름에 따른 도시 교통의 구조적 특징을 도출한다. 분석 결과는 도시의 교통을 좀 더 쉽게 이해할 수 있도록 도와주고, 향후에 도시교통의 혼잡 해소방안, 도로 확장 계획 등의 교통정책 수립시 기초연구 자료로 활용할 수 있다.

■ 중심어 : 빅데이터, 교통빅데이터, 네트워크 분석, 시각화

Abstract

Intelligent transportation system (ITS) has been introduced to maximize the efficiency of operation and utilization of the urban traffic facilities and promote the safety and convenience of the users. With the expansion of ITS, various traffic big data such as road traffic situation, traffic volume, public transportation operation status, management situation, and public traffic use status have been increased exponentially. In this paper, we derive structural characteristics of urban traffic according to the vehicle flow by using big data network analysis. DSRC (Dedicated Short Range Communications) data is used to construct the traffic network. The results can help to understand the complex urban traffic characteristics more easily and provide basic research data for urban transportation plan such as road congestion resolution plan, road expansion plan, and bus line/interval plan in a city.

■ Keyword : Bigdata, Traffic big data, Network analysis, Visualization

I. 서 론

세계 각국은 도시의 교통 혼잡과 이로 인한
물류비용의 증가 그리고 교통사고를 줄이기 위

하여 최근 들어 빅데이터를 활용하는 연구에 집
중하고 있다. 첨단 정보통신 기술의 발전으로
자동차 내부는 물론 주변의 다양한 교통관련 기
기들에서 교통 빅데이터가 생성되고 있으며, 이

2017년 11월 30일 접수; 2017년 12월 21일 수정본 접수; 2017년 12월 26일 게재 확정

* 본 연구는 “도시 교통 문제 개선을 위한 클라우드 기반 트래픽 예측 시뮬레이션 SW 기술 개발사업”의 위탁과제 “도시대중교통 정보검지 데이터를 이용한 혼잡분석연구”로 수행되었음.

† 교신저자 조완섭 wscho63@gmail.com

를 활용하면 도시의 교통문제를 해결하는데 큰 도움이 될 것이다. 특히, 우리나라 주요 도시들은 이용자의 안전과 편의를 도모하고 교통시스템의 운영 및 이용 효율을 극대화하기 위해 지능형 교통 시스템(Intelligent Transport Systems: 이하 ITS)을 도입하여 활용하고 있다. ITS는 기존의 교통 시설물(도로 신호시스템, 도로 안전 시설 등)에 첨단 전자, 통신, 제어 기술을 접목시켜 기존 도로의 처리용량을 100% 활용토록 하는 시스템이다.

본 논문에서는 청주시 ‘교통 빅데이터 분석시스템’[6]에서 확보한 2015년 3월의 단거리 전용 통신(Dedicated Short Range Communications, 이하 DSRC) 데이터를 활용하여 도시의 교통 네트워크를 생성하고, 기존의 네트워크 분석기법들을 이용해 도심 도로 교통의 구조적 특징을 파악하고자 한다. 또한 네트워크 분석을 통해 특정 지점을 기준으로 차량의 흐름을 도출하여 교통정책 수립에 활용하고자 한다. 분석 결과는 차후에 청주시 교통 발전을 위한 혼잡 해소방안, 도로 확장 계획 등의 청주시 교통계획의 기초연구 자료로 사용될 수 있을 것이다.

본 논문의 의의는 다음과 같다. 먼저, 차량의 흐름을 보기위해 DSRC에서 수집되는 대규모 자료에 대하여 빅데이터 기술을 활용하여 네트워크 분석을 수행한다. 다음으로, 시간대별 다차원으로 네트워크를 구축함으로써 특정 조건에 따른 교통 네트워크의 구조적 특징을 파악하고 비교한다. 또한, 연결 도로가 많은 교차로들을 중심으로 하위 네트워크를 생성하여 차량의 흐름을 파악할 수 있도록 한다. 마지막으로 교차로 별로 중심성 지표를 산출해 특정 교차로가 가지는 성격을 분석한다.

네트워크 분석의 결과는 도시의 교통 발전을 위한 혼잡 해소방안, 도로 확장 계획 등의 교통계획의 기초연구 자료로써 활용될 수 있을 것이다.

본 논문의 구성은 다음과 같다. 제 II 장에서

관련연구를 설명하고, 제 III 장에서는 청주시 교통 네트워크 분석 내용과 시각화에 관해 기술한다. 마지막으로 제 IV 장에서 결론 및 향후연구에 대해 기술한다.

II. 관련 연구

본 장에서는 네트워크 분석기법과 교통 분야에의 활용사례에 관하여 설명한다.

네트워크 분석 기법은 분석 수준에 따라 <표 1>과 같이 6가지 유형으로 구분할 수 있다[5]. 본 논문에서는 위의 분석 기법 중 네트워크 수준 및 내재된 특성 분석과 중심성 분석을 수행하며, 이를 거시적 관점(전체 교통 네트워크)에서 분석한 결과를 기술한다.

<표 1> 네트워크 분석 수준에 따른 주요 분석지표

분석 수준	내용	주요 분석지표
네트워크 수준 분석	네트워크 전체적 특성 파악	네트워크 크기, 네트워크 밀도
노드 수준 분석	노드의 특성, 노드들 간의 링크 관계에 대한 특성 분석	연결정도, 연결강도, 연결거리, 직경, 평균연결거리, 도달가능성, 보행, 경로, 최대흐름
네트워크에 내재된 특성 분석	노드들의 특성 분석	상호성, 이행성, 군집화 계수, E-I지수
중심성 분석	네트워크 내에서 각 노드들의 영향력 크기를 분석	연결 중심성, 근접 중심성, 매개 중심성, 파워 중심성, 아이겐 벡터 중심성, 페이지 랭크
하위집단 분석	네트워크의 하위 집단을 구분함	컴포넌트 분석, 파당 분석, 구조적 등위성 분석 등
에고 네트워크 수준 분석	특정한 에고를 중심으로 형성되는 네트워크의 특성 분석	에고 네트워크 특성 분석, 중개성 분석, 구조적 공백 분석

특정 지역의 DSRC 기반 빅데이터를 활용하여 교통 네트워크 분석 연구한 사례는 이제 시작단계에 있는 연구분야이다. 본 연구에서는 교통 분야에 네트워크 분석을 수행한 연구사례와 해석을 위해 다양한 분야에서 네트워크 분석을 적용한 사례를 참고하였다. 관련 연구는 <표 2>와 같다. 교통 분야의 네트워크 분석 사례의 경우 대부분 대중교통을 대상으로 이루어졌으며, 중심성 지표를 도출하고 비교하는 연구이다. 참고문헌[4]의 경우 대중교통 취약지를 도출하기 위해, 서울시 지하철 네트워크를 구성하여 각 지점마다 중심성 지표를 활용하였다. 각 중심성 지표마다 높은 값을 가지는 지하철역에 대해서 분석하였다. 참고문헌[7]의 경우 지하철을 이용하는 승객의 승하차 태그를 활용해 네트워크를 구축하고, 중심성 지표를 통해 전체 네트워크에서 각 지하철역의 역할에 대해 고찰하였다. 참고문헌[3]의 경우 버스의 승하차 태그를 통해 네트워크를 구축하고, 이를 사회 경제 지표와의 연계를 통해 다차원으로 분석하였다.

<표 2> 교통 분야 네트워크 분석 관련연구

연구자 (연도)	활용 중심성 지표	연구의 특징
양지수 외[4]	근접 중심성, 직결 중심성, 매개 중심성	교통 취약지 도출
한경훈 외[7]	연결 중심성, 근접 중심성, 위세 중심성	각 중심성 지표에 따른 특징 도출
박한솔 외[3]	연결 중심성, 매개 중심성	사회 경제 지표와의 연계 분석

III. 분석 결과

본 장에서는 DSRC 교통 데이터를 분석한 결과를 소개한다. 제3.1절에서는 본 연구에 사용된 데이터를 소개한다. 제3.2절에서는 네트워크 분석

결과를 소개하고, 제3.3절에서 시각화를 소개한다.

3.1 데이터소개

분석의 대상이 되는 데이터는 충북대와 청주시가 공동으로 구축한 ‘교통 빅데이터 분석시스템’ 내에 축적된 DSRC 데이터이다(청주시, 2015). 본 논문에서 사용한 데이터는 2015년 3월 1일부터~2015년 3월 31일까지 1개월간 청주시 전역에 있는 DSRC에서 수집한 데이터로 레코드의 수는 29,886,072건이다.

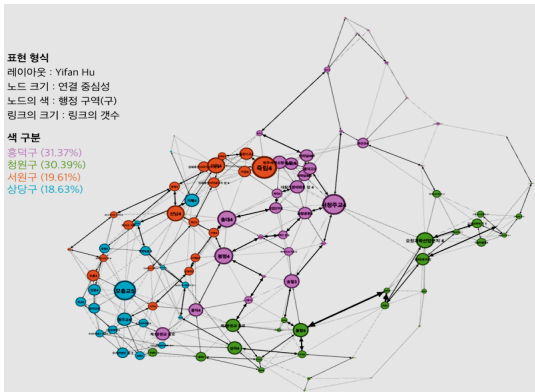
DSRC 데이터는 <그림 1>과 같이 수집차량(OBU_ID)이 수집된 시간(GEN_TIME)에 수집 지점(EM_ID)을 통과할 때 수집된 데이터이며, 추가적으로 수집 지점의 유형(EM_TYPE), 수집 차량의 유형(OBU_TYPE)이 포함되어 있다. 차량에 부착된 단말기의 경우 암호화되어 있기 때문에 개인 식별은 불가능하다.

	GEN_TIME	EM_TYPE	EM_ID	OBU_ID	VEH_TYPE
1	2014-11-22 00:00:00.000	D	1005	88B383998BE398EA3349519814E0FC6C	1
2	2014-11-22 00:00:00.000	D	1028	85E31861D3D783E22FCAFD3B2527B0A2	1
3	2014-11-22 00:00:00.000	D	1070	1A08A616929DF6824ED45C4703BC859A	1
4	2014-11-22 00:00:01.000	D	1003	C965CDFCD9AA96A3D37A1138E3227A7A	1
5	2014-11-22 00:00:01.000	D	1041	F949981887680FB98A6240CFD00459FE	1
6	2014-11-22 00:00:02.000	D	1034	F1CB1F98470E280BD38D46CB6B394B52	1
7	2014-11-22 00:00:03.000	D	1008	A8FAF3082BAAD1C730F4EF66E0DC74A2	1
8	2014-11-22 00:00:03.000	D	1026	240E69F414828F55BF6E85569326DF98	1
9	2014-11-22 00:00:03.000	D	1026	2A4B73CC569BA6877394689D2E53816B	1

<그림 1> DSRC 데이터의 형식

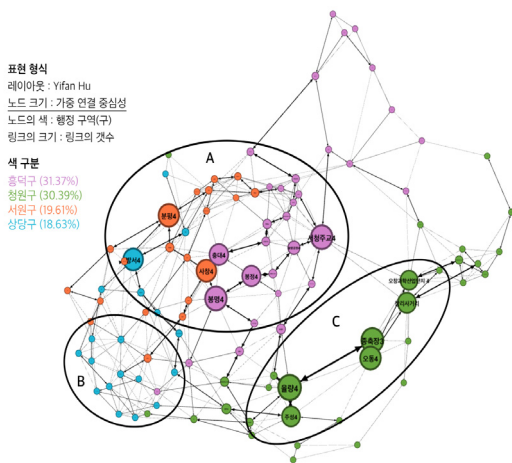
3.2 데이터 분석 및 시각화

<그림 1>의 데이터를 이용하여 각 수집지점별로 교통량과 교통흐름을 분석할 수 있는 교통흐름 네트워크 모델을 구축하였다. 구축된 전체 네트워크는 <그림 2>와 같으며, 노드의 크기는 노드의 연결중심성을 간선의 굵기는 교통량을 의미한다. 노드의 색상은 행정구역의 구분을 의미하는 것으로 각각 흥덕구, 청원구, 서원구, 상당구에 위치한 수집지점의 특성을 구분하여 볼 수 있다.



<그림 2> DSRC 기반 청주시 교통 네트워크

<그림 2>를 보면, 청주시 교통 네트워크의 유형은 현실 세계에 가장 많이 존재하는 허브(hub)형 네트워크로 볼 수 있으며, 세부적으로는 다수의 허브 노드들을 중심으로 국지적으로 결집되어 있는 형태의 다(多)허브형 네트워크로 볼 수 있다[1]. 허브 노드 또는 중심(central) 노드는 전체 네트워크에서 중심이 되는 노드를 의미한다. 하지만, 연결중심성은 수집지점의 구조적 특성을 표현하는 것으로 각 지점의 활동성을 평가할 수 있는 지표를 산출하기 위하여 가중연결중심성을 산출하였고 그 결과는 <그림 3>과 같다.



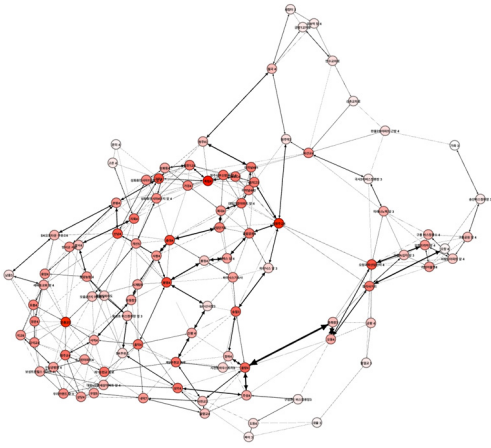
<그림 3> DSRC 기반 청주시 교통 네트워크 (실제 교통량)

<표 3>은 노드의 연결중심성과 교통량의 크기를 함께 고려한 가중 연결 중심성과 연결중심성을 표시한 것이다(상위 5개 지점만 표시함). 표에서 율량 사거리의 가중연결 중심성 값이 가장 높게 나타났다. 노드의 가중 연결 중심성 지표가 높다는 것은 네트워크 내에서 활동량이 크다는 것을 의미한다. 즉, 교통 네트워크에서는 주요 교차로로 연결된 경로가 많고, 그 경로로 차량의 유출입이 많다는 것을 의미한다. <그림 2>와 비교하여 네트워크를 살펴보면 공통적으로 교차로로 연결된 경로가 많은 지점과 실제 차량의 이동이 많은 교차로는 꼭 일치하지 않는다는 것을 알 수 있다. 두 개의 지표를 비교한 표는 아래 <표 3>과 같으며, 값 뒤에 괄호는 순위를 의미한다. 이러한 차이가 발생하는 이유는 운전자는 다른 지역으로 이동하기 위해 가능한 경로가 많은 교차로뿐 아니라 인근 지역의 특성, 이동의 목적, 도로의 크기 등을 고려해 교차로와 도로를 선택해서 운전하기 때문으로 보인다.

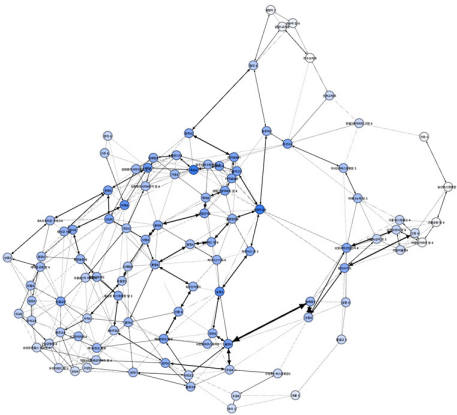
<표 3> 가중 연결 중심성 상위 5 교차로의 중심성 지표 비교

교차로	가중 연결 중심성	연결 중심성
율량 사거리	1401822(1)	19(9)
종축장 사거리	1009293(2)	10(66)
봉명 사거리	937125(3)	22(4)
서청주교사거리	846803(4)	26(2)
오동 사거리	830930(5)	12(55)

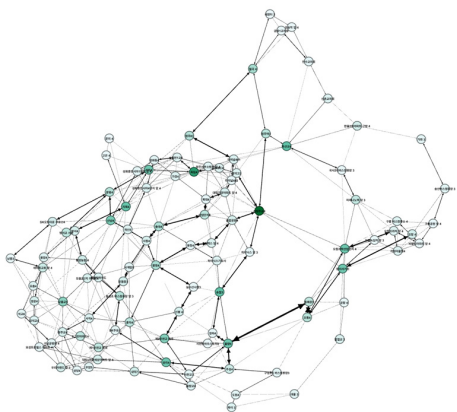
<그림 4>, <그림 5>, <그림 6>은 미시적 관점에서 교차로들의 중심성 지표들을 비교한 것이다. 연결 중심성이 높은 수집 지점은 근접 중심성, 매개 중심성 역시 높게 나왔다. 또한, 다른 행정구역이나 교차로로 이동이 쉬운 지점에 차량이 많이 몰리는 것으로 볼 수 있다.



〈그림 4〉 연결중심성(Degree Centrality)

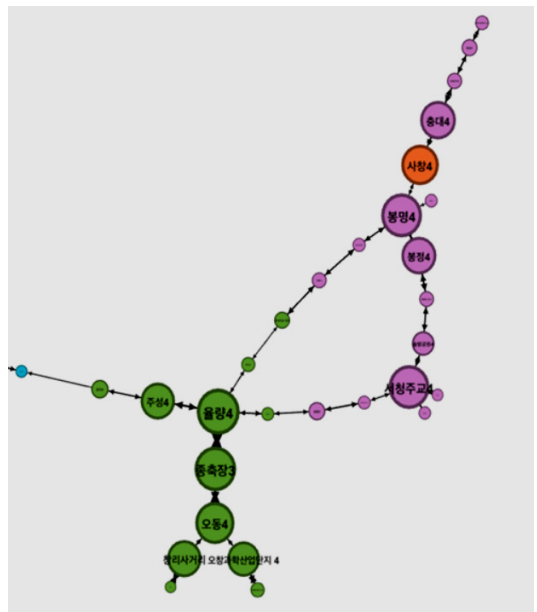


〈그림 5〉 근접중심성(Closeness Centrality)



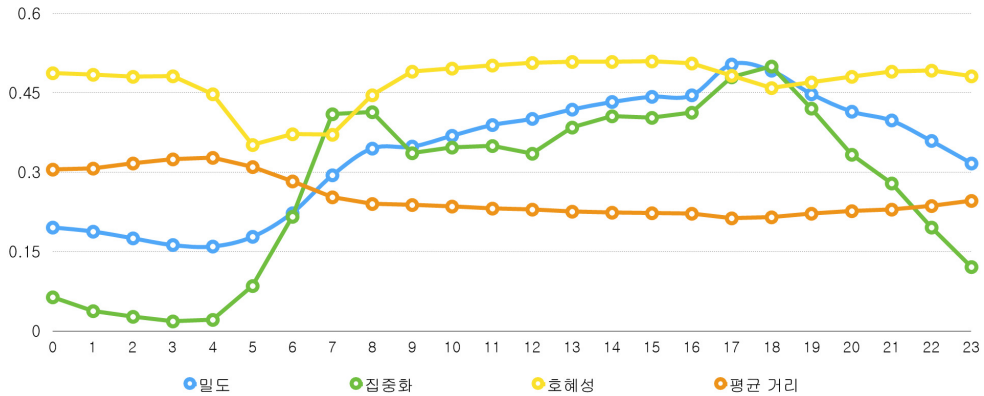
〈그림 6〉 매개중심성(Betweenness Centrality)

<그림 3>의 교통 네트워크에서 링크의 가치가 낮은 것을 제외하고 시각화하면 <그림 6>과 같이 대표 네트워크를 도출할 수 있다. 청주시 교통은 ‘울량’ 사거리, ‘서청주 교차로’ 사거리, ‘봉명’ 사거리로 구성되는 삼각형 형태를 중심으로 실제 교통의 유출입이 많은 교차로들이 연결되어 있다. 이것이 청주시의 대표적인 교통 흐름(무방향)이라고 볼 수 있다.



〈그림 7〉 교통흐름 대표 네트워크

마지막으로 교통 네트워크 수준의 패턴을 보기 위해 시간대별로 네트워크를 구성하여 비교하였다. 현실 세계에 존재하는 대부분의 네트워크는 끊임없이 네트워크의 구조가 변화하는 동적 네트워크(dynamic network)의 특성을 가진다. <그림 8>은 밀도, 집중화, 상호성의 변화를 보기 위한 그래프로써, 밀도를 기준으로 정규화를 시킨 것이다. 교통 네트워크도 시간에 따라 네트워크의 수준이 변화하는 동적 네트워크라는 것을 확인 할 수 있다. 특히 공통적으로 출근(6~9시), 퇴근(17~20시) 시간에 변화량이 크다.



(그림 8) 시간별 밀도, 상호성, 집중성, 평균거리 비교 그래프

<그림 8>로부터 두 가지 사실을 알 수 있다. 첫째, 밀도와 집중화(centralization)의 변화이다. 밀도가 네트워크 전체 노드들이 연결하는 정도의 크기를 말하는 것이라면, 집중성은 소수의 노드들 사이에 존재하는 연결에 따라 집중하는 정도를 표현한다[5]. 교통에서 네트워크의 밀도가 증가할수록 교차로의 복잡도가 높아지고, 집중성이 증가 할수록 다른 교차로에 비해 특정 교차로의 복잡함이 크게 증가한다고 할 수 있다. 밀도와 집중화 모두 사람이 활동하는 시간에 모두 높게 나타나는 것을 볼 수 있다. 이는 교통에서 차량을 운전하는 것은 사람이기 때문이기 때문에 사람이 활동하는 시간에 맞춰 교통도 복잡해지기 시작하는 것이다. 특히 집중화의 경우 출근 시간과 퇴근 시간에 집중적으로 높아지는데, 이는 출퇴근 시간에 사람들이 주로 이용하는 교차로가 있다는 것을 짐작할 수 있다.

둘째, 상호성의 변화이다. 상호성은 노드들 간의 양방향 링크가 많을수록 증가한다. 일반 통행 도로의 경우 0을 나타내게 된다. 전체 교통 네트워크에서 상호성은 0.917로 모든 도로에서 양방향 소통이 차이가 크지 않음을 볼 수 있다. 하지만 동적 네트워크에서 출퇴근 시간대에 상호성이 감소하는 것을 볼 수 있다. 이는 특정 시간대에 일정한 통행의 흐름이 발생하는 것으로 볼 수 있다.

VI. 결 론

본 논문에서는 DSRC 데이터를 활용하여 교통 네트워크를 구축하고, 도심 교통의 구조적 특징을 파악하고, 차량의 흐름을 분석하였다.

첫째, 교통 네트워크가 가지는 구조적 특징을 분석한 결과 청주시의 교통 네트워크는 다수의 허브(3개 지역: 울량사거리, 서청주교차로, 봉명사거리)가 국지적으로 결집되어 있는 삼각형 네트워크를 형성하고 있었다.

둘째, 교통이 시간별로 변하는 동적 네트워크라는 것을 확인하였고, 출근시간대(6~9시)와 퇴근시간대(17~20)에 밀도와 집중성이 급격히 올라감을 확인하였다.

셋째, 연결 도로가 많은 주요 교차로의 차량 흐름을 분석한 결과 도로가 많이 연결되어 있는 교차로 임에도 차량의 흐름이 많지 않은 경우가 발견되었으며, 이는 차량의 흐름을 파악할 때 연결도로 뿐 아니라 그 도로의 성격을 함께 파악해야 함을 의미한다.

본 연구에서 활용한 DSRC 데이터는 단말기를 부착한 차량의 정보만 수집한 것으로 청주의 경우 전체 차량의 약 60% 정도에 해당한다. 또한, 청주시 전체 교차로 중 수집장비가 부착된 100여 개의 교차로는 주요교차로만 포함한 것으로 청주시 전체 교통 네트워크 모델을 구축한

것으로 보기 어렵다. 다만, 이 분석의 결과는 청주시 교통흐름의 특성을 파악하고, 각 교차로별 특징에 맞는 교통신호 체계, 도로 확장 방안 등 도시의 교통으로 인한 사회경제적비용을 줄일 수 있는 교통정책을 수립할 때 기초연구 자료로 활용될 수 있을 것이다.

참 고 문 헌

- [1] 김상배, “네트워크 세계정치이론의 모색”, 국제정치논총, 제48권, 제4호, pp.35-61, 2008.
- [2] 김용연, DSRC 기반 청주시 교통 네트워크 구조 분석, 충북대학교 대학원 석사학위논문, 2017.
- [3] 박한솔, 버스정보시스템과 사회경제지표 기반의 다차원 대중교통 네트워크 분석, 충북대학교 대학원 석사학위논문, 2016.
- [4] 양지수, 김동성, 전경수, “네트워크 중심성 지표를 활용한 대중교통 취약지 분석”, 대한토목학회 학술대회 논문집, pp.21-24, 2009.
- [5] 이수상, 네트워크 분석 방법론, 논형, 2012.
- [6] 청주시, 청주시, 교통정보 빅데이터 분석 시스템 구축 최종 보고회, 발표자료, 2015.
- [7] 한경훈, 서영수, 박근병, “사회 네트워크 분석을 이용한 지하철역 네트워크 구조 분석”, 한국철도학회 학술발표대회논문집, pp.34-40, 2015.

저 자 소 개



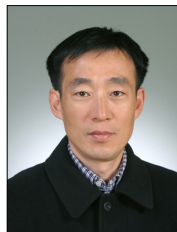
김 용 연(Yong Yeon Kim)

- (주)엔코아
- 관심분야 : 빅데이터, 데이터 마이닝



이 경 희(Kyung-Hee Lee)

- 2004년 : 충북대 컴퓨터과학과 (박사)
- 2008년~현재 : 충북대 빅데이터학과 교수
- 관심분야 : 빅데이터, 알고리즘, 데이터마이닝



조 완 섭(Wan-Sup Cho)

- 1987년 : KAIST 전산학과 (박사)
- 1996년~현재 : 충북대학교 교수
- 관심분야 : 빅데이터, 빅데이터 거버넌스