

기상상황에 따른 서울시 대중교통 이용 변화 분석: 폭설을 중심으로

원민수* · 천승훈** · 신성일*** · 이선영****

Won, Minsu*, Cheon, Seunghoon**, Shin, Seongil***, Lee, Seonyeong****

Analysis of Public Transport Ridership during a Heavy Snowfall in Seoul

ABSTRACT

Severe weather conditions, such as heavy snowfall, rain, heatwave, etc., may affect travel behaviors of people and finally change traffic patterns in transportation networks. To deal with those changes and prevent any negative impacts on the transportation system, understanding those impacts of severe weather conditions on the travel patterns is one of the critical issues in the transportation fields. Hence, this study has focused on the impacts of a weather condition on travel patterns of public transportations, especially when a heavy snowfall which is one of the most critical weather conditions. First, this study has figured out the most significant weather condition affecting changes of public transport ridership using weather information, card data for public transportation, mobile phone data; and then, developed a decision-tree model to determine complex inter-relations between various factors such as socio-economic indicators, transportation-related information, etc. As a result, the trip generation of public transportations in Seoul during a heavy snowfall is mostly related to average access times to subway stations by walk and the number of available parking lots and spaces. Meanwhile, the trip attraction is more related to business and employment densities in that destination.

Key words : Weather condition, Public transport ridership, Card data, Mobile phone data, Decision-tree model

초 록

기상상황(폭염, 폭우, 한파, 폭설)은 대중교통 통행 및 이용 패턴에 영향을 미치는 중요한 변수 중의 하나이며, 시스템의 예측가능성과 안정성을 중시하는 교통분야에서 이러한 기상의 영향을 이해하는 것은 매우 중요한 요소중의 하나이다. 그러므로 본 연구에서는, 서울시를 대상으로 기상 상황에 따른 대중교통 이용 변화를 분석하고 해석하고자 하였다. 먼저, 기상, 모바일폰통신, 대중교통카드 자료를 이용하여 각 기상 상황별 서울시 대중교통 이용 변화를 살펴보고, 가장 영향이 큰 폭설 상황을 기준으로 대중교통 이용패턴을 지역별로 분석하였다. 또한, 의사결정모델 (Decision-tree Model)을 활용하여 각 영향 변수들 간의 복잡한 관계를 밝히고자 하였다. 분석결과, 폭설 시, 전체 통행에 대한 잠재수요는 감소하고, 대중교통으로의 수단 전환이 일어나는 것을 확인할 수 있었다. 또한, 강동 및 송파 지역과 강서, 구로, 양천, 영등포 지역은 대중교통 이용이 증가하였으며, 관악, 금천, 동작 지역은 상대적으로 큰 변화가 없는 것을 확인할 수 있었다. 마지막으로, 폭설 시 대중교통 출발량의 변화는 도보 접근통행시간, 정류장 근처 주차 가용성 등이 중요한 역할을 하며, 도착량의 변화는 해당 도착지의 종사자 및 사업체 밀도와 밀접한 연관이 있다는 것을 알 수 있었다.

검색어 : 기상, 대중교통수요, 통신데이터, 카드데이터, 의사결정모델

* 정회원 · 한국교통연구원 부연구위원 (The Korea Transport Institute · minsuwon@koti.re.kr)

** 정회원 · 교신저자 · 한국교통연구원 연구위원 (Corresponding Author · The Korea Transport Institute · sh1000@koti.re.kr)

*** 서울연구원 연구위원 (The Seoul Institute · ssi@si.re.kr)

**** 정회원 · 한국교통연구원 연구위원 (The Korea Transport Institute · sylee708@koti.re.kr)

Received November 7, 2019/ revised November 15, 2019/ accepted November 17, 2019

1. 서론

대중교통 이용 패턴은 승용차보유대수, 유가, 요금, 서비스 수준, 도로 현황, 기상 상황 등 다양한 요소와 관계가 있으며, 그 중 기상 상황은 짧은 시간 내에 통행 의사결정에 영향을 줄 수 있는 주요 변수 중의 하나이다(Choi et al., 2013). 특히, 지구 온난화와 함께 급변하는 기상 상황과 그에 따른 일별 통행 패턴의 불확실성은, 예측 가능하고 안정화된 시스템을 유지해야 하는 교통 분야에서 매우 중요한 이슈 중의 하나이다. 그럼에도 불구하고, 기존 많은 연구들은 지역적 특수성, 자료의 부족 등을 이유로 대중교통 이용 패턴과 기상 상황의 관계를 효과적으로 분석하지 못하고 있다(Choi et al., 2013).

그러므로 본 연구에서는, 서울시를 대상으로 폭우, 폭설, 폭염, 한파와 같은 다양한 기상 상황에 따른 대중교통 이용의 변화 패턴을 분석하고, 그 관계를 보다 명확히 이해하고자 한다. 먼저, 1) 통행 패턴 및 대중교통 이용 패턴을 파악할 수 있는 모바일통신자료와 대중교통카드자료, 기상상황을 파악할 수 있는 기상청 기상자료 등을 이용하여, 대중교통 일별 수요에 가장 큰 영향을 주는 특정 기상 상황을 파악하고, 2) 해당 주요 기상 상황에 따른 서울시 지역별 대중교통 이용 변화를 분석하여, 3) 다양한 사회경제적 및 교통관련 지표들을 이용하여 서울시 대중교통 이용 변화의 원인들을 분석하고자 한다(Fig. 1 참조).

2. 선행 연구

대중교통의 통행특성, 수요변화, 이용패턴 등에 대한 연구는 기상조건을 포함한 다양한 영향변수들을 활용하여 국내외에서 폭 넓게 연구되고 있다. 먼저, Park and Lee(2012)은 강우량 수준에

따른 통행패턴 분석을 위해 단독통행, 환승 통행으로 나누어서 선형회귀모형(Linear Regression)을 적용하여 분석하였다. 주요변수로는 강우량 수준, 연평균 일 통행량 평균 통행 금액이며 강우일의 평균통행건수 및 평균 통행금액이 비강우일에 비해 적은 것으로 나타났다. 또한 강한 비일수록 버스 통행이 줄어드는 것을 확인하였고, 특히 주말과 평일의 통행이 많은 차이를 보이는 것을 확인하였다. Choi et al.(2013)는 강우, 불쾌지수, 강설, 체감온도 등 4가지 기상조건이 대중교통 수요 변화량의 차이에 대한 분석을 회귀모형을 적용하여 분석하였다. 분석 시 시내버스, 마을버스, 지하철 수단으로 분류하여 그 차이를 확인하였으며 불쾌지수, 체감온도의 기상 조건은 대중교통 수요를 약 2~7 % 감소시키는 것으로 나타났고, 지하철보다는 버스가 높은 상관관계를 보이는 것을 확인하였다. Shin and Choe(2014)는 강수량이 대중교통 수단별 승차 승객 수 변화에 미치는 영향을 SUR모형을 적용하여 분석하였다. 이 연구에서는 통행 수단별 통행량은 영향이 있다는 연구 가설로 시작하였다. 주요 변수로는 버스, 도시철도, 마을버스의 일 승객수와, 일 강수량이다. 분석 결과 일 강수량이 10 mm 이상일 때는 대중교통수단의 승객 수는 감소하는 것으로 나타났다. Lee et al.(2014)은 강우일과 비강우일을 구분하여 1주일치 대중교통카드 자료를 활용하여 버스, 지하철에 대해 행정구 단위로 지역을 구분하여 대중교통 수요 영향을 분석하였다. 주요변수로는 주말, 평일, 버스 유형(광역버스, 간선버스, 지선 마을버스)별 대중교통 일일 승차인원이다. 강우일에 지선버스, 마을버스보다 광역, 간선버스의 경우 대중교통 이용 감소가 더 많은 것을 확인하였으며, 강우량은 장거리 버스 통행에 더 큰 영향을 주는 것으로 나타났다. Lee et al.(2015)는 서울지하철 역세권을 대상으로 각 지하철역의 침투 시간, 비침투시간의 승하차 이용자 수로 역세권을 6개로 유형화하여 통행 행태를 파악하였다. 주요변수로는 승하차량, 인구밀도,

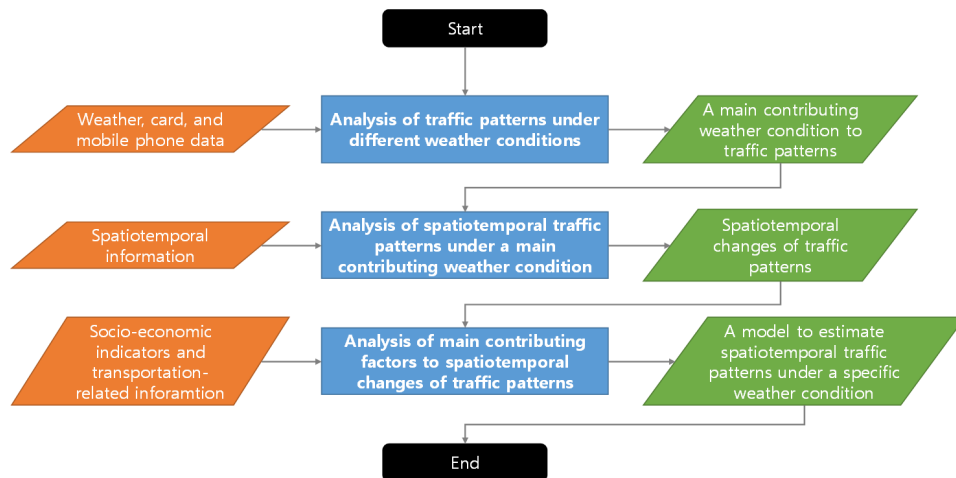


Fig. 1. Organization of This Study

가구밀도, 사업체 밀도, 토지이용(상업시설, 업무시설, 주택 등)정보이다. 분석 결과 역세권 유형에 따라 오전, 오후의 첨두 특성이 다르게 나타났고 역 주변 토지이용 및 도시 특성 요소는 대중교통 통행에 매우 밀접한 관계가 있음을 확인하였다. Kim et al.(2017)은 기상에 따른 서울시 대중교통 이용률 예측 분석을 위해 19가지 Regression을 비교하여 분석하였다. 주요 변수는 기온, 강수량, 미세먼지, 버스노선별, 지하철역별 승하차 인원 정보다. 분석을 위해 지하철역 주변 버스데이터를 이용하여 Regression 함수의 평균 오차율을 가지는 Gaussian Process Regression 함수를 선택하였으며 대중교통 이용률 예측에 따른 평균 오차율은 15.49 %으로 나타났다. Lee and Jung(2018)은 기상조건이 대중교통 수단별 통행량에 미치는 영향을 분위회귀모형(Quantile Regression)의 분위수를 0.1, 0.3, 0.5, 0.7, 0.9으로 설정하여 분석하였다. 주요 변수는 기온, 습도, 풍속, 강수량, 단일 통행량, 환승 통행량이며, 통행 수단별, 평일 주말여부에 대해 통행량 차이가 있는 것을 확인하였다. 평일보다는 주말의 경우, 환승 통행보다는 단일 통행이 경우에 기상조건에 따라 분위 계수 변동 폭이 큰 것으로 나타났다.

국외 연구로는 먼저, Stover and McCormack(2012)는 기상조건이 대중교통 수요 관계를 선형회귀모형(Linear Regression)을 적용하여 계절별로 분석하였다. 주요 변수는 대중교통 승객 수, 풍속, 강수량, 적설량, 요일, 실업률, 취업률이다. 분석 결과 강한 바람은 봄, 가을, 겨울에 대중교통 수요를 감소시키는 영향을 주고, 낮은 온도(영하 7°C 이하)는 겨울에 대중교통 수요를 감소시킨다는 것을 확인하였다. Kashfi et al.(2013)는 버스 이용률과 강수량의 관계를 선형회귀모형(Linear Regression)으로 분석하였다. 주요 변수는 계절과 요일을 구분한 시간당 대중교통 승객 수, 강수량, 기온이다. 아침 첨두시간, 주말 승객은 강수량이 많아질수록 버스 통행이 감소하는 것을 확인하였다. 특히 다른 계절과 비교했을

때, 여름에 강우량과 버스 통행에 대한 영향이 더 민감한 것으로 나타났다. Arana et al.(2014)는 기상조건에 따른 대중교통 이용자 통행 패턴 분석을 위해 통행 목적을 고려한 버스 통행에 미치는 영향을 다중회귀모형(Multiple Regression)을 적용하여 분석하였다. 주요 변수는 바람, 강수량, 기온, 버스 승차 승객 수, 통행목적이다. 여가 목적으로 버스를 이용할 경우 바람, 강우가 증가할수록 버스 통행 횟수를 줄이고 기온 상승은 통행 횟수를 증가시키는 것을 확인하였다. 또한, 해당 연구에서는 기상 조건은 규칙적인 버스 통행자보다는 불규칙적인 버스 이용자에게 더 민감한 영향을 준다고 밝혔다. Zhou et al.(2017)는 규칙적인 대중교통 이용자를 대상으로 기상조건과의 영향을 분석하였다. 주요 변수는 시간당 버스, 지하철 승차 승객 수, 습도, 풍속, 강수량이다. 분석결과 습도, 바람 및 강우량이 증가하면 특정 수준에서 대중교통 이용률 감소와 영향이 있는 것을 확인하였다.

기존 연구들을 통해, 대중교통 통행 및 이용 패턴에 영향을 주는 다양한 변수들이 있으며, 그 중 기상 조건도 중요한 역할을 한다는 것을 알 수 있었다. 하지만, 이러한 기상 조건에 따른 대중교통 이용 변화는 지역별, 시간대별 차이에 따라 다양한 패턴을 보이며, 사용하는 모델링 방법과 가용 데이터의 종류에 따라 다양하게 해석되는 것을 알 수 있었다. 특히, 기존 연구들은 대중교통의 이용량을 추정할 수 있는 자료만을 이용하여 대중교통의 변화를 분석하였다. 그렇기 때문에 각 기상상황에 따른 대중교통의 통행패턴 변화가 잠재 수요의 변화에서 비롯되는 것인지, 통행 수단 간 전환에 의한 것인지를 파악하는데 한계가 있었다. 그러므로 본 연구에서는 대중교통의 통행패턴 변화를 파악할 수 있는 자료는 물론, 전체 통행량의 변화를 추정할 수 있는 모바일통신자료를 이용하여 이러한 통행 수단 간 전환과 잠재 수요의 변화를 보다 면밀하게 밝히고자 하는데, 그 차별성이 있다고 할 수 있다(Table 1).

Table 1. Analysis of Literature

Reference	Approach	Main contributing factors
Park and Lee(2012)	Linear Regression	Rainfall, # of daily trips, # of transfers
Choi et al.(2013)	Linear Regression	Discomfort index, sensible temperature, snowfall
Shin and Choe(2014)	Seemingly Unrelated Regression	Daily precipitation, # of daily trips of bus and subway
Lee et al.(2014)	Simple Comparison	Type of transport mode, district levels, type of passengers, time periods
Lee et al.(2015)	Cluster Analysis	Peak and non-peak hours, land use
Kim et al.(2017)	Gaussian Process Regression	Temperature, rainfall, fine dust
Stover and McCormack(2012)	Linear Regression	Wind, temperature, rain, snowfall, season
Kashfi et al.(2013)	Linear Regression	Temperature, rain, time periods, season
Arana et al.(2014)	Linear Regression	Wind, rain, temperature, trip purposes
Zhou et al.(2017)	Multivariate Modeling Approach	Humidity, wind, temperature, trip purposes

3. 자료 구축

본 연구에서는 기상에 따른 대중교통 이용 변화를 분석하기 위하여, 기상청에서 제공하는 기상특보 정보(Weather report)와 서울시 대중교통카드자료(Card data for public transportation), 모바일폰통신자료(Mobile phone data), 사회경제지표(Socio-economic indicators), 그 외 각종 교통관련지표(Other transportation-related information) 등을 활용하였다(Table 2 참조).

본 연구에서는 보유하고 있는 데이터의 가용성을 고려하여, 공간적 범위는 서울시 내부로 한정하고, 시간적 범위는 2016년 4월 1일부터 2017년 3월 31일까지 1년으로 한정하였다. 또한, 필요에 따라 주요 기상상황에 따른 대중교통 이용 변화를 부가적으로 분석하기 위하여, 추가적인 데이터를 더 구성하였다(‘4. 자료 분석 및 결과’ 부분 참조).

모바일폰통신자료는 사람들이 휴대폰을 사용하거나, 사용하지 않을 때에도 인근 기지국과 주기적으로 통신하는 신호 데이터(Mobile Phone Signaling Data, MPSD)의 축적이다. 이러한 신호 데이터는 데이터 생성일, 단말기 식별번호, 통신 기지국 위치정보 등을 포함한다. 이렇게 수집된 모바일폰통신자료는 다른 데이터들과 시공간 단위가 다르기 때문에 분석을 위해서는 각 시공간 단위를 이용하여 표준화하는 작업이 필요하다. 이를 위하여 브르노이 다이어그램 기법(Voronoi diagram method)을 사용하여 각 기지국별 통신폴리곤(polygon)을 구성하였다. 대중교통카드 이용 내역은 각 이벤트 단위로 수집되며, 승차 및 하차 정류장 ID 정보를 포함하고 있다. 이러한 정보는 버스정류장정보(Bus Information System, BIS)와 매칭되어 위치정보를 계산하고, 모바일폰통신자료의 수집 단위인 통신폴리곤 단위로 집계하였다. 마지막으로 본 연구에서는, 서울 열린데이터광장(Seoul Metropolitan Government, 2018)과 한국교통연구원 내부 자료(The Korea Transport Institute, 2015)를 이용하여, 종사자수, 사업체수, 지하철역 및 버스정류장 수, 지하철역 접근 통행 시간 등, 각종 사회경제적 및 교통관련 지표를 구축하여, 기상 상황별 대중교통 이용 변화 분석에 활용하였다.

4. 자료 분석 및 결과

4.1 기상 상황별(폭염, 폭우, 한파, 폭설) 대중교통 이용 변화

기상청에서는 해당 및 주변일의 온도, 습도, 풍속, 강우량 정보 등을 이용하여 크게 폭염, 폭설, 폭우, 한파 등 4가지 형태로 기상정보를 발효한다(Chung et al., 2009). 단, 폭설은 24시간 신적설이 5 cm 이상 예상될 때, 폭우는 3시간 강우량이 60 mm 이상 예상되거나 12시간 강우량이 110 mm 이상 예상될 때, 폭염은 일최고기온이 33°C 이상인 상태가 2일 이상 지속될 것으로 예상될 때, 마지막으로 한파는 10월~4월 중 아침 최저기온이 전날보다 10°C 이상 하강하여 3°C 이하이고 평년값보다 3°C가 낮을 것으로 예상될 때, 아침 최저기온이 -12°C 이하가 2일 이상 지속될 것이 예상될 때, 또는 급격한 저온현상으로 중대한 피해가 예상될 때로 정의하고 있다. 본 연구에서는 이렇게 제공되는 기상 정보를 바탕으로 차량 및 대중교통 이용자들의 통행 의사결정이 변경될 것이라는 가정하에, 대중교통카드자료 및 모바일폰통신자료를 이용하여 그 사실을 검증하고자 한다.

먼저, Fig. 2(a)는 모바일통신자료를 이용하여 추정된 서울시 전체의 통행량(출발+도착) 차이를 기상 상황별로 보여주고 있으며, Table 3은 이러한 통행량 차이의 유의성을 판단하기 위한 독립표본 t검정의 p-value 값을 보여주고 있다. 결과를 살펴보면, 평상시 대비, 폭염과 폭설 시 통행량이 유의미하게 줄어든 것을 확인할 수 있다. 하지만, 폭우와 한파 시에는 통행량이 다소 증가하는 것처럼 보이지만, 통계적으로는 유의하지 않다. 특히, 폭설 시에는 평상시 대비 약 5.1 %의 통행량이 줄어든 것을 확인할 수 있다.

두번째로, Fig. 2(b)와 Table 3의 두번째 부분은 대중교통카드이용자료를 이용하여 추정된 서울시 전체의 대중교통 이용률 차이를 기상 상황별로 보여주고 있다. 결과를 보면, 폭염과 폭설 시, 평상시 대비 대중교통의 이용률이 오히려 유의미하게 증가한 것을 확인할 수 있다. 특히, 폭설 시에는 평상시 대비 약 9.9 %으로 가장 많이 대중교통 이용률이 증가한 것을 확인할 수 있다.

Fig. 3은, 앞에서 분석된 Fig. 2와 Table 3을 바탕으로, 각 기상상황 및 대중교통 수단별 상대적 이용률을 보여주고 있다. 분석 결과를 살펴보면, 폭설 시 대중교통 이용률이 상대적으로

Table 2. Descriptions of Datasets

Dataset	Temporal scope	Spatial scope	Note
Weather report	2016.01.01~2017.12.31	Korea	Weather information, per every day
Card data for public transportation	2016.04.01~2017.03.31	Seoul	# of public transportation trip information, per every day
Mobile phone data	2016.04.01~2017.03.31	Korea	# of trip information, per every 15 minutes,
Socio-economic indicators	2016.01.01~2017.12.31	Korea	Business density, Employment density, etc.
Other transportation-related information	2016.01.01~2017.12.31	Korea	# of parking lots, access time to a subway, etc.

가장 크게 증가한 것을 확인 할 수 있으며, 지하철과 버스 또한 폭설 시 상대적 이용률이 가장 높은 것으로 나타났다. 그러나 지하철과 버스 이용률 변화 패턴에는 별다른 차이가 없는 것으로 나타났다.

정리하면, 서울시 대중교통 이용은 폭염과 폭설 시 유의미하게 변화한다. 특히, 모바일폰통신자료로 추정된 통행량은 감소하며, 대중교통카드자료로 추정된 대중교통 이용률은 증가한다. 이러한 현상은 폭염과 폭설 시, 일부 사람들은 굳은 날씨로 인해 통행을 포기하거나, 기존 자동차 및 도보 등을 이용하던 사람들이 대중교통을 이용하여 이동한다는 것을 보여준다. 즉, 전체 통행에 대한 잠재수요의 감소와 대중교통으로의 통행 수단 전환이 동시에 일어난다는 것이다. 특히, 이러한 현상은 폭설 시 두드러지게 나타나는 것을 확인할 수 있었다.

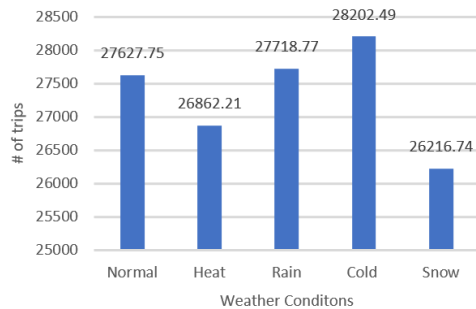
4.2 주요 기상 상황(폭설)에 따른 지역별 대중교통 이용 변화

앞 절의 주요 기상 상황별 대중교통 이용 변화 분석을 통해, 서울시 대중교통 이용률은 폭설 시 가장 많이 변한다는 사실을 Figs. 2 and 3을 통해 확인할 수 있었다. 그러므로 본 연구에서는,

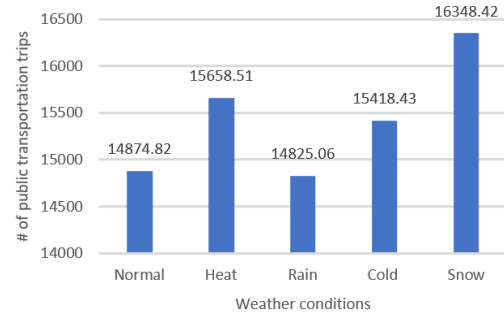
대중교통카드자료를 이용하여 추정된 대중교통 이용 변화를 지역적으로 분석하기 위하여, 폭설 시 서울시 권역별 대중교통 이용 변화와 각 권역의 사회경제적 지표의 특성을 Fig. 4와 Table 4와 같이 살펴보았다.

Fig. 4는 대중교통 통행 특성을 바탕으로 분류된 서울시 8개 권역을 보여주고 있으며(Sung and Kim, 2005), Table 4는 이러한 각 권역의 대중교통 통행 기종점별 이용 변화와 각종 사회경제적 및 교통관련 지표를 보여주고 있다. 분석결과를 살펴보면, 폭설 시, 서울의 동쪽에 해당하는 강동 및 송파(권역-4)의 대중교통 이용률이 가장 많이 증가한 것을 확인할 수 있다(기점 및 종점 기준 각각 16.7 %와 15.7 % 증가). 또한, 서울의 서쪽으로 대표되는 강서, 구로, 양천, 영등포(권역-7) 지역도 대중교통 이용률이 많이 증가한 것을 확인할 수 있다(기점 및 종점 기준 각각 12.4 % 증가). 반면에, 관악, 금천, 동작(권역-6) 지역은 상대적으로 대중교통 이용률이 작게 증가한 것을 확인할 수 있다.

해당지역들의 특성을 살펴보면, 대중교통 이용률이 상대적으로 작게 증가한 권역-6은 종사자 밀도(1999.3)와 폴리곤별 평균 지하



(a) # of Trips Estimated by Mobile Phone Data

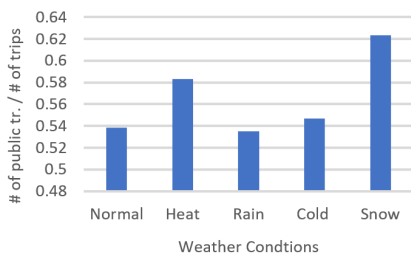


(b) # of Public Transportation Trips Estimated by Card Data

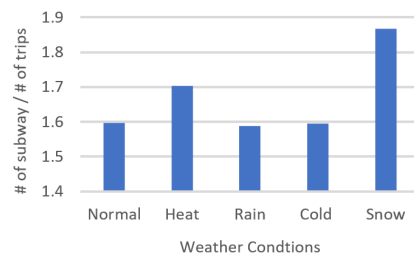
Fig. 2. # of Trips According to Weather Conditions

Table 3. Unpaired Two-Samples T-Test for # of Trips Based on Weather Conditions

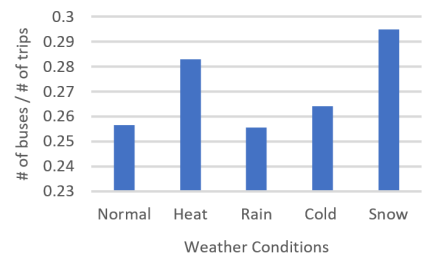
Weather conditions (Normal vs.)		Heat	Rain	Cold	Snow
p-values	(a) # of trips by mobile phone data	0.0009	0.7497	0.1533	0.0101
	(b) # of public transportation trips by card data	0.0128	0.8917	0.2965	0.0679



(a) Ratio of Public Transportation Trips



(b) Ratio of Subway Trips



(c) Ratio of Bus Trips

Fig. 3. Ratio of Public Transportation Trips by Modes and Weather Conditions

철역 수(1.1)가 상대적으로 작으며, 도보 및 다른 대중교통을 이용한 평균 지하철 접근 시간(각각, 13.3분과 10.4분)은 상대적으로 긴 것으로 나타났다. 한편, 상대적으로 대중교통 이용률이 크게 증가한 권역-4의 경우, 폴리곤별 평균 버스 정류장 수(50.6)가 작았지만, 그 외 유의미한 특징을 찾을 수 없었다. 즉, 폭설 시, 서울시 대중교통



Fig. 4. Classification of Area

이용 변화를 지역별로 살펴보면, 종사자 밀도, 지하철역 수 및 접근 시간 등이 유의미한 영향을 미친다는 사실을 확인할 수 있다.

4.3 대중교통 이용 변화의 주요 원인 분석

해당 단락에서는 앞 단락에서 발견된 사실들을 바탕으로 대중교통 이용 변화를 보다 면밀하게 분석하기 위하여, 통신 폴리곤 단위와 그에 해당하는 각종 사회경제적 및 교통관련 지표들(사업체 수, 종사자 수, 버스정류장 수, 지하철역 수, 주차장 수, 주차면 수, 지하철역 도보 접근 시간, 지하철역 대중교통 접근 시간)을 이용하여, 폭설 시 대중교통 이용 변화를 설명하는 모형을 만들하고자 한다. 먼저, 설명 변수로 사용될 수 있는 각종 사회경제적 및 교통관련 지표들은 서로 높은 상관관계를 가지고 있으며(Kim, 2009), 대중교통 이용 변화와 선형 관계를 담보할 수 없다. 특히, 평균적 대중교통 이용률의 변화가 아닌, 분포의 극단에 있는 상대적으로 두드러지는 대중교통 이용 변화를 찾아내기 위해서는 분류 모형(Classification Model)이 유리하다(Aggarwal, 2013). 왜냐하면, 의사결정트리모형(Decision-tree Model)과 같은 분류 모형은 각 변수의 분포 가정에 자유로우며, 분포의 극단의 값을 설명하는데 더욱 유용하기 때문이다(Han et al., 2011; Aggarwal, 2013). 또한, 폭설 시 예상되는 급격한 대중교통의 이용 변화를 분석 및 예측하는 것은, 시스템의

Table 4. Trip Changes (Estimated by Card Data) According to Regions and Their Socio-Economic Conditions

Classification of area		1	2	3	4	5	6	7	8
Origin	Trips under normal conditions	896,062	818,737	750,595	516,898	1,032,450	730,494	976,339	768,644
	Trips under heavy snowfall	982,217	903,123	831,257	603,145	1,128,231	794,579	1,097,499	843,999
	Rate of change (%)	9.6 %	10.3 %	10.7 %	16.7 %	9.3 %	8.8 %	12.4 %	9.8 %
Destination	Trips under normal conditions	898,449	803,913	755,128	530,133	1,023,043	725,879	989,043	764,601
	Trips under heavy snowfall	983,691	889,184	833,025	613,424	1,117,362	789,332	1,111,490	843,692
	Rate of change (%)	9.5 %	10.6 %	10.3 %	15.7 %	9.2 %	8.7 %	12.4 %	10.3 %
Business density		1077.8	346.3	468.7	480.2	338.9	405.8	545.6	462.5
Employment density		5712.0	1404.1	2171.4	2827.1	3083.0	1993.3	2807.1	2335.7
Average number of bus stations by polygon*		64.9	87.2	64.1	50.6	65.7	76.8	85.7	91.7
Average number of subway stations by polygon		3.5	1.3	2.4	1.8	2.1	1.1	1.7	2.2
Average number of parking lots by polygon		118.7	153.7	178.1	209.6	148.8	232.4	204.1	279.6
Average number of parking spaces by polygon		1902.7	2070.3	2008.0	2875.9	2894.0	2175.3	3088.9	2803.9
Average access time to subway stations by the other transit (minutes)		9.8	13.0	10.6	9.6	10.7	13.3	11.6	10.9
Average access time to subway stations by walk (minutes)		7.2	9.8	7.0	7.2	7.8	10.4	8.0	7.7

* Polygon: a standard unit space to collect mobile phone data in this study

신뢰성과 안정성을 유지해야 하는 교통 운영 측면에서 매우 중요한 목표 중의 하나이다. 그러므로 본 연구에서는, 의사결정트리모형의 대표적인 알고리즘인 CART (Classification And Regression Trees) 알고리즘(Breiman, 2017)을 활용하여 분류 모형을 구축하였다.

본 연구에서는 분류 모형에서 발생할 수 있는 각 분류 카테고리별 불균형 문제(Unbalance Problem)를 해결하고, 각 터미널 노드(Terminal Node)의 신뢰성을 확보하기 위하여, 다음과 같은 과정을 통해 의사결정트리모형을 구축하였다(Won et al., 2018).

- Step-1: 대중교통 이용 증감률을 기종점별 ‘Origin’과 ‘Destination’으로 구분하고, 각 카테고리별 데이터 빈도를 고려하여, 각 증감률을 ‘감소’, ‘0~5 %증가’, ‘5~10 %증가’, ‘10~15 %증가’, ‘15~20 %증가’, ‘20 %이상 증가’로 구분함(Table 5 참조).

Table 5. Frequency of Each Category of Trip Changes

Category of trip changes	Origin	Destination
Decreased	178 (11.3 %)	152 (9.6 %)
0~5 % increased	165 (10.4 %)	162 (10.2 %)
5~10 % increased	319 (20.2 %)	348 (22.0 %)
10~15 % increased	317 (20.1 %)	329 (20.8 %)
15~20 % increased	223 (14.1 %)	244 (15.4 %)
≥ 20 % increased	379 (24.0 %)	349 (22.0 %)
Total	1581 (100 %)	1584 (100 %)

- Step-2: Step-1에서 구성된 기종점별 대중교통 이용 증감률과 각종 사회경제적 및 기초 교통 지표를 CART 알고리즘을 이용하여 의사결정트리(Decision-tree)를 구성함.

- Step-3: 도출된 의사결정트리의 각 터미널 노드(Terminal Node)의 신뢰성을 약 60 % 이상 확보하기 위하여, 주변 카테고리들과 결합하여 각 터미널 노드의 최종 결과를 도출함.

Figs. 5 and 6은 위에서 언급된 모델링 과정을 통해 산출된 의사결정트리이다. 결과를 살펴보면, 폭설 시 기종점별 대중교통 이용률의 변화는 ‘평균 버스정류장 수(bus)’, ‘평균 지하철역 수(subway)’, ‘평균 지하철역 도보 접근 시간(walk, minutes)’, ‘평균 지하철역 대중교통 접근 시간(transit, minutes)’, ‘평균 주차장 수(parking lots)’, ‘평균 주차면 수(parking spaces)’, ‘사업체 밀도(business)’, ‘종사자 밀도(employment)’와 이들의 조건 조합(Combination of Conditions)으로 설명되는 것을 확인 할 수 있다. 예를 들어, Fig. 5에서 평균 지하철역 도보 접근 시간이 3.7분 이상이고, 평균 주차장 수가 485개 이상이면, 폭설 시 대중교통 통행 발생량은 약 15 %이상 증가한다는 것을 의미한다.

특히, 폭설 시, 기점 기준 대중교통 이용 변화에 영향을 미치는 첫번째 조건은 해당 대중교통 수단으로 접근하기 위한 도보 통행 시간(walk)이다. 또한 해당 대중교통 정류장 근처의 가용한 주차장 수 및 주차가능면 수가 중요한 역할을 한다는 것을 알 수 있다. 반면에, 종점 기준 대중교통 이용 변화에 영향을 미치는 첫번째

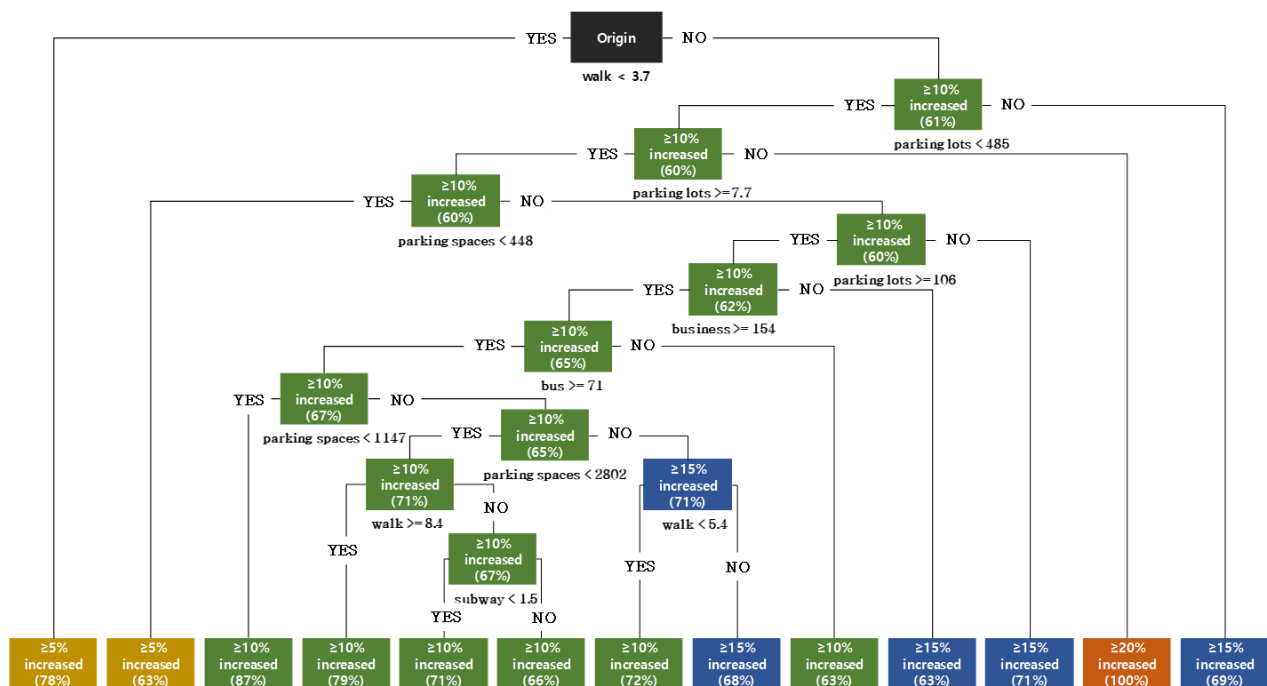


Fig. 5. Classification Rules for Trip Changes of Public Transportation During Heavy Snowfalls (Based on Origin)

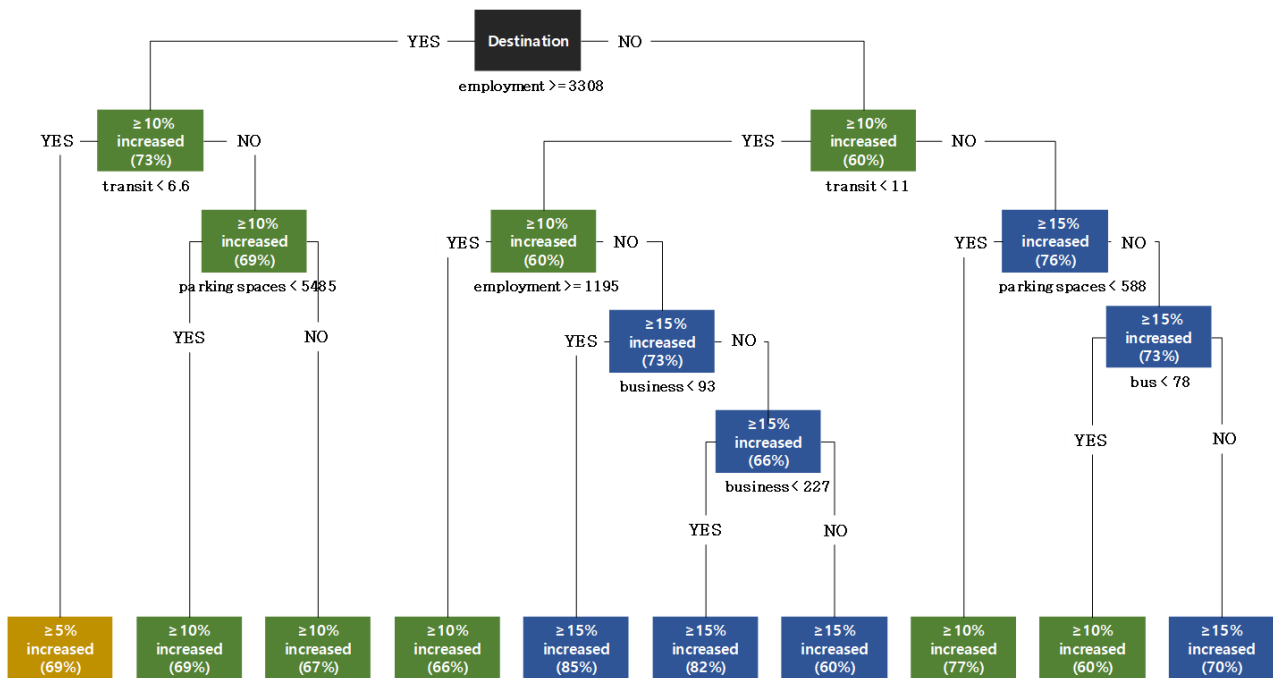


Fig. 6. Classification Rules for Trip Changes of Public Transportation During Heavy Snowfalls (Based on Destination)

조건은 해당 종점의 종사자 수 밀도이다. 또한 해당 종점의 사업체 수도 대중교통 도착량을 결정하는 중요한 역할을 한다는 것을 확인할 수 있다.

5. 결론

본 연구는 서울시를 대상으로 기상상황에 따른 대중교통 이용 변화를 분석하였다. 먼저, 기상, 모바일폰통신, 대중교통카드 자료를 이용하여, 각 기상 상황별(폭염, 폭우, 한파, 폭설) 서울시 통행패턴과 대중교통 이용 변화를 살펴보고, 가장 영향이 큰 기상 상황을 설정하였다. 폭염과 폭설인 경우에 모바일폰통신자료로 추정된 서울시 통행량은 줄어들고, 대중교통카드자료로 추정된 대중교통 이용은 증가하였다. 이러한 현상은 전체 통행에 대한 잠재수요는 감소하고, 대중교통으로의 수단 전환이 동시에 일어났다는 것을 보여준다. 특히 이러한 현상은 폭설 시 두드러지는 것을 확인할 수 있었다. 다음으로, 이렇게 설정된 폭설 상황을 기준으로 대중교통 이용 변화를 서울시 8개 권역별로 분석하였다. 서울의 동쪽에 해당하는 강동 및 송파 지역과 서쪽에 해당하는 강서, 구로, 양천, 영등포 지역은 대중교통 이용이 증가하였으며, 관악, 금천, 동작 지역은 상대적으로 변화가 크지 않은 것으로 확인되었다. 또한, 각 지역의 사회경제적 및 교통관련 지표들의 특성과 지역적 위치를 고려해보면, 강동 및 송파 지역은 상대적으로 대중교통으로의 접근이 용이하였기 때문이고, 강서, 구로, 양천, 영등포 지역은 이러한

대중교통으로의 접근성이 용이하지는 않았지만, 서울의 서쪽인 김포 및 인천지역에서 접근하는 기존 승용차 통행자들의 수단 전환이 일어났기 때문인 것으로 보인다. 마지막으로 이러한 대중교통 이용 변화를 보다 면밀하게 분석하기 위하여, 통신 폴리곤 단위와 그에 해당하는 각종 사회경제적 및 교통관련 지표들을 활용하여, 폭설 시 서울 대중교통 이용 변화를 분석 및 추정할 수 있는 의사결정 트리를 구성하였다. 분석결과, 폭설 시 대중교통 출발량의 변화는 도보접근통행시간, 정류장 근처 주차 가용성 등이 중요한 역할을 하며, 도착량의 변화는 해당 도착지의 종사자 및 사업체 밀도와 밀접한 연관이 있다는 것을 알 수 있었다. 즉, 폭설 시 전체적인 통행량이 줄어들어도 불구하고, 대중교통으로의 접근성이 용이한 지역을 중심으로 출퇴근을 위한 목적통행이 수단 전환되었다는 것을 유추할 수 있었다.

본 연구는 기상정보, 대중교통카드자료, 모바일폰통신자료, 사회경제적 및 교통관련 지표들을 이용하여 기상 상황별 서울시 대중교통 이용 변화를 지역적으로 분석하고, 주요 원인들의 모호하고 복잡한 관계를 분류 모형을 활용하여 보다 명확하게 밝혔다는 것에 그 의미가 있다. 그럼에도 불구하고, 체감 기상에 대한 영향, 통신 및 카드 자료의 전수화와 대표성에 대한 이슈, 폴리곤 단위의 사회경제적 지표의 부재 등은 여전히 풀어야할 숙제로 남아있다. 하지만, 본 연구에서의 발견들이 향후 보다 정밀하고 정확한 기상 상황별 대중교통 이용 패턴 예측 모형을 만드는 데 활용될 수 있을 것이라고 기대된다.

감사의 글

이 연구는 기상청이 출연하고 한국기상산업기술원에서 발주하여 시행한 ‘미래유망 민간기상서비스 성장기술개발사업(KMI 2018-04910)’의 지원으로 수행되었습니다.

References

- Aggarwal, C. C. (2013). *Outlier analysis*, Springer, Berlin.
- Arana, P., Cabezudo, S. and Peñalba, M. (2014). "Influence of weather conditions on transit ridership: A statistical study using data from smartcards." *Transportation research part A: policy and practice*, Vol. 59, pp. 1-12.
- Breiman, L. (2017). *Classification and regression trees*, Routledge.
- Choi, S. G., Rhee, J. H. and Oh, S. H. (2013). "The effect of weather conditions on transit ridership." *J. Korean Soc. Civ. Eng.*, KSCE, Vol. 33, No. 6, pp. 2447-2453 (in Korean).
- Chung, W. Y., Jung, S. J., Kim, J. J. and Kwon, T. H. (2009). "A study on local area weather condition monitoring system in WSN and CDMA." *The journal of the Korea Institute of Maritime Information & Communication Sciences*, Vol. 13, No. 8, pp. 1713-1720 (in Korean).
- Han, J., Kamber, M. and Pei, J. (2011). *Data mining: concepts and techniques third edition*, The Morgan Kaufmann Series in Data Management Systems, Morgan Kaufmann Publishers, Massachusetts, pp. 83-124.
- Kashfi, S. A., Lee, J. and Bunker, J. (2013). "Impact of rain on daily bus ridership: a Brisbane case study." *Australasian Transport Research Forum 2013 Proceedings 2-4 October 2013*, Brisbane, Australia.
- Kim, H. J., Oh, S. and Kim, U. M. (2017). "A study on the prediction of public transportation consumption in seoul by weather." *Proceedings of the Korea Information Processing Society Conference*, Korea Information Processing Society, pp. 656-659 (in Korean).
- Kim, J. (2009). "An analysis of the changes in the cause-and-effect relationships between socio-economic indicators and the road network of seoul using structural equation model." *Journal of the Korean Geographical Society*, Vol. 44, No. 6, pp. 797-812 (in Korean).
- Lee, J. H. and Jung, H. Y. (2018). "The impact of weather conditions on transit ridership using quantile regression analysis." *Journal of Korea Planning Association*, Vol. 53, No. 4, pp. 95-106 (in Korean).
- Lee, J., Go, J. Y., Jeon, S. and Jun, C., (2015). "A study of land use characteristics by types of subway station areas in Seoul analyzing patterns of transit ridership." *The Korea Spatial Planning Review*, Vol. 84, pp. 35-53 (in Korean).
- Lee, K. S., Eom, J. K., Min, J. H. and Yang, K. Y. (2014). "The Impact of rain on public transit ridership in Seoul." *Journal of the Korean Society of Railway*, Vol. 2014, No. 5, pp. 252-257 (in Korean).
- Park, K. and Lee, S. (2012). "A study on the effect of adverse weather conditions on public transportation mode choice." *J. Korean Soc. Civ. Eng.*, KSCE, Vol. 32, No. 1, pp. 23-31 (in Korean).
- Seoul Metropolitan Government (2018). *Seoul data center*, Available at: <http://data.seoul.go.kr> (Accessed: November 4, 2019) (in Korean).
- Shin, K. and Choe, G. J. (2014). "Analyzing the relationship between precipitation and transit ridership through a seemingly unrelated regression model." *Journal of Korean Society of Transportation*, Vol. 32, No. 2, pp. 83-92 (in Korean).
- Stover, V. W. and McCormack, E. D. (2012). "The impact of weather on bus ridership in Pierce County, Washington." *Journal of Public Transportation*, Vol. 15, No. 1, pp. 95-110.
- Sung, H. and Kim, T. H. (2005). "A study on categorizing subway station areas in Seoul by rail use pattern." *Journal of Korean Society of Transportation*, Vol. 23, No. 8, pp. 19-29 (in Korean).
- The Korea Transport Institute (2015). *Unpublished report* (in Korean).
- Won, M., Kim, H. and Chang, G. L. (2018). "Knowledge-based system for estimating incident clearance duration for maryland I-95." *Transportation Research Record: Journal of the Transportation Research Board*, Vol. 2672, No. 14, pp. 61-72.
- Zhou, M., Wang, D., Li, Q., Yue, Y., Tu, W. and Cao, R. (2017). "Impacts of weather on public transport ridership: Results from mining data from different sources." *Transportation Research Part C: Emerging Technologies*, Vol. 75, pp. 17-29.