■ 論 文 ■

다항로짓모형을 이용한 지역간 철도통행 연구

Multinomial Logit Modeling: Focus on Regional Rail Trips

김 경 태
(한국철도기술연구원 선임연구원)

이 진 선
(우송대학교 철도경영학부 조교수)

목 차

Ⅰ. 서론
1. 연구의 배경 및 목적
2. 기존 연구 고찰
Ⅱ. 지역간 철도의 특징 및 차종별 분담률
1. 지역간 철도의 특징
2. 차종별 수요분담률
Ⅲ. 다항로짓모형의 이론적 검토
Ⅳ. 다항로짓모형의 추정 및 검증
1. 분석 자료
2. 다항로짓모형을 이용한 경험적 모형 추정
3. 모형의 검증
Ⅴ. 결론 및 향후 연구 주제
참고문헌

Key Words : 지역간 철도통행, 차종별 수요분담률, 열차운행횟수, 차내시간, 다항로짓모형
regional rail trips, modal share, frequency, in-vehicle time, multinomial logit

요 약

지역간 철도통행은 열차종별로, 요일별로 또는 거리구간대별로 다양한 특성을 가지고 있으며, 각 구간별로 제공되는 교통 서비스나 경쟁요인에 따라 이용수요에 차이를 보이게 된다. 따라서 지역간 철도내부의 서비스 및 경쟁요인의 변화는 결과적으로 차종별 분담률에 영향을 주는 요인이 된다. 본 연구에서는 지역간 철도통행에서 역간 수요의 차종별 분담률에 영향을 주는 요인으로서 열차운영계획과 밀접한 관련이 있는 빈도, 차내시간 그리고 운임변수를 고려하여 다항로짓모형을 추정하고, 모형을 이용한 예측치와 실제 통행량을 비교함으로써 모형의 적정성을 검토해 보았다. 지역간 철도에서 역간 수요의 차종별 분담률은 빈도와 차내시간이 차종선택에 중요한 영향을 미치는 것으로 분석되었으며 열차운영계획의 매우 중요한 고려대상임을 확인할 수 있었다. 또한 일반적으로 철도수요가 요일별 특성이 뚜렷하지만 역간 수요의 차종별 분담률은 요일별 특성을 보이지 않는 것으로 분석되었다. 본 연구의 결과는 효율적인 열차운영계획의 수립과 철도서비스 구축을 위한 기반을 제공할 수 있을 것으로 기대되며, 다양한 전략적·정책적 해석이 가능할 것으로 판단된다.

Increasingly, the emphasis in regional passenger rail planning is finding ways to more efficiently use existing facilities, with particular attention being paid to policies designed to spread peak-period travel demand more evenly throughout the week with consideration of train classification. In this context the individual's choice of time to travel is of crucial significance. This paper investigates the use of multinomial logit analysis to model ridership by rail classification using data collected for travel from Seoul to Busan during the one week in October 2004. The particular model form that was successfully calibrated was the multinomial logit (MNL) model: it describes the choice mechanism that will permit rail systems and operations to be planned on a more reliable basis. The assumption of independently and identically distributed (IID) error terms in the MNL model leads to its infamous independence from irrelevant alternatives (IIA) property. Relaxation of the IID assumption has been undertaken along a number of isolated dimensions leading to the development of the MNL model. For business and related rail travel patterns, the most important variables of choice were time and frequency to the chosen destination. The calibrated model showed high agreement between observed and predicted market shares. The model is expected to be of use to railroad authorities in planning and determining business strategies in the increasingly competitive environment of regional rail transport.

Ⅰ. 서론

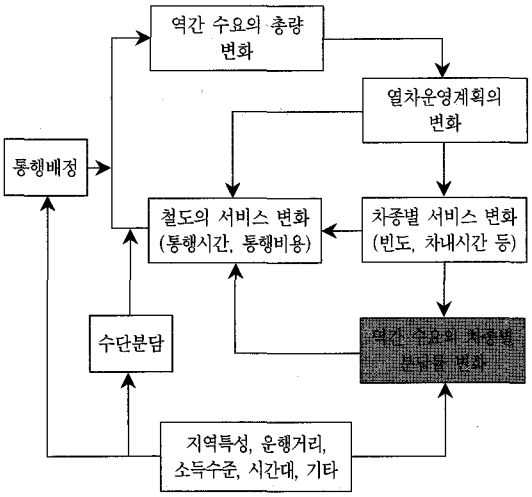
1. 연구의 배경과 목적

2004년 말 현재 전국에서 1일 발생하는 지역간 통행인구는 총 1,309만명(통행)으로 전년도에 비해 지속적으로 증가하고 있는 것으로 나타났다. 통행거리를 고려한 수송분담률은(인키로 기준) 도로가 87.5%(승용차 75.7%, 버스 11.8%), 철도 10.7%, 항공 1.7%, 해운 0.2% 순으로 나타났다. 지역간 여객 통행에서 철도의 비율은 도로수단에 비해서 낮지만 고속철도 개통 이후 그 비율은 점진적으로 증가하고 있다.

우리나라의 지역간 철도는 수도권을 중심으로 하여 방사형 형태를 취하고 있으며, 이로 인해 서울~시흥 구간과 같이 수요가 집중되는 특정 구간에서 용량이 부족한 현상이 발생하고 있다. 이러한 현상은 철도수요가 지속적으로 증가할 것으로 예상되기 때문에 앞으로도 심화될 것으로 판단된다. 특정 구간의 용량부족 현상을 해결하기 위해서는 신선 건설이 가장 큰 효과가 있으나 예산, 도심통과로 인한 부지확보 등의 문제로 장기적인 관점에서 접근할 수 밖에 없다. 다른 대안으로서는 기존의 시설을 효율적으로 활용하면서 신규열차의 도입, 열차운영계획의 조정 등을 통해서 용량부족 문제를 완화시킬 수 있다. 특히 효율적인 열차운영계획 수립을 통한 방법은 많은 예산을 필요로 하지 않은 효과적인 방법이 될 수 있지만 열차운영계획과 철도수요와의 관계에 대한 명확한 규명이 선행되어야 한다.

〈그림 1〉에서 보는 바와 같이 열차운영계획과 역간 수요는 상호간에 영향을 주는 관계에 있다. 일반적으로 역간 수요의 차종별 분담률은 통행배정의 결과로서 얻어지는 것이 보통이지만 본 연구에서는 통행배정모형을 활용하지 않고, 각 차종별 서비스 수준을 직접 활용하여 역간 수요의 차종별 분담률을 구하고자 하였다.

본 연구에서는 지역간 철도에서 역간 수요의 차종별 분담률에 영향을 주는 요인으로서 열차운영계획과 밀접한 관련이 있는 빈도, 차내 시간을 포함한 운임을 고려하여 다항 로짓모형을 이용하여 분석하였다. 이러한 차종별 수요에 영향을 주는 여러 가지 변수를 이용하여 모형을 추정하고, 모형을 이용한 예측치와 실제 통행량을 비교함으로써 모형의 적정성을 검토해 보았다. 본 연구의 결과는 효율적인 열차운영계획의 수립과 철도서비스 구축을 위한 기반을 제공할 수 있을 것으로 기대된다.



〈그림 1〉 열차운영계획과 역간수요의 개념도

2. 기존 연구 고찰

대중교통에서 노선별 수요분담률의 분석은 통행배정모형의 결과로서 도출되는 것이 일반적이다. 즉, 존 단위로 분할된 네트워크에서 대중교통의 기종점 통행량을 기반으로 하여 노선별 통행량을 산정하게 된다. 차종별 수요분담률은 동일한 차종이라도 다른 노선으로 구분될 수 있으므로, 노선별 분석 자료를 이용하여 구할 수 있다.

일반적으로 기종점을 기반으로 하는 대중교통 통행배정모형을 활용하여 노선별 역간수요를 직접 도출하는 것은 한계가 있다. 대중교통 통행배정모형에 많이 활용되는 최적전략 기반의 모형(Spiess & Florian (1989), Spiess(1993), Wu *et al.*(1994), Jayakrishna *et al.*(1995))은 직접적으로 역간수요를 도출하는 데에 한계가 있음이 알려져 있다. 경로의 일부분으로 연속일 필요가 없는 임의의 두 환승노드를 연결한 것으로 정의할 수 있는 경로구간 기반의 통행배정모형(De Cea & Fernández(1989, 1993), Lam *et al.*(1999))은 역간 수요를 도출하는 것이 보다 용이하지만 경로구간을 이용하는 노선을 단순화시켜 지역간 철도의 경우와 같이 다양한 차종이 운행하는 경우에는 각 차종별 특성을 합리적으로 반영하는 데에 어려움이 있다.

지역간 철도에서 모든 정차역은 환승노드가 될 수 있으므로 출발역과 도착역 쌍은 하나의 경로구간으로 표현할 수 있다. 기존의 대중교통 통행배정모형이 지역간 철도의 경로구간의 차종별 수요분담률을 현실적으로 예측

하는 데에 어려움을 겪는 가장 큰 이유는 모든 노선의 통행비용을 일반화비용이라는 하나의 지표로만 활용하기 때문에 각 노선의 분담률에 영향을 줄 수 있는 요인 중에서 일반화비용에 포함되지 않거나, 포함되더라도 노선별 분담률에 독자적인 영향을 줄 수 없기 때문이다. 본 연구에서는 지역간 철도에서 경로구간의 차종별 수요 분담률에 영향을 주는 요인들 각각의 영향력을 확인할 수 있는 보다 효과적인 접근을 통하여 차종별 분담률을 계산하고자 하였다.

Ⅱ. 지역간 철도의 특징 및 차종별 분담률

1. 지역간 철도의 특징

일반적인 대중교통의 특징은 다음과 같다. 첫째, 대중교통은 사전에 결정된 일정한 노선을 가지고 운행되며, 노선은 하나의 기종점 쌍을 운행하는 차량의 집합으로 정의된다. 같은 노선에 속하는 차량은 크기, 용량, 운행패턴 등이 동일하며 항상 동일한 노드 및 링크를 경유해서 운행한다. 둘째, 대중교통은 동일한 경로를 운행하는 노선이 다수 존재한다. 셋째, 대중교통은 사전에 결정된 일정한 운행계획에 따라 운행된다. 넷째, 대중교통을 이용하는 통행자의 기종점간 통행은 대중교통의 차내시간 뿐만 아니라 접근, 대기, 환승 등 복잡한 통행요소로 구성된다. 다섯째, 대중교통은 요금이 존재한다.

지역간 철도는 일반적인 대중교통의 특징을 가지고 있지만 몇 가지 차별화되는 점도 있다. 지역간 철도는 노선버스나 도시철도에 비해서 동일한 경로를 운행하는 차량의 수는 적지만 대중교통에서 정의하는 노선의 개념에서 본다면 노선의 수는 많아진다. 즉, 지역간 철도는 출발역과 도착역 사이의 기종점 중심이 아닌 차량별로 고유번호를 가지면서 각기 다른 운행 특징을 가지기 때문이다. 예를 들어 2004년 10월 20일 기준 전국의 지역간 철도를 운행하는 차량을 대상으로 하여 노선을 구분해 보면 총 632개 차량이 운행하였으며, 노선으로 구분할 경우 총 463개 노선이 되어 1개 노선당 차량수는 평균 1.37개로서 매우 낮은 수준이다. 즉, 동일한 노선에 속하는 차량은 매우 적으면서 서로 다른 노선들은 운임의 특성이나 통행시간의 특성이 유사한 노선들이 많으며, 이러한 현상은 노선버스나 도시철도 등 도시권 대중교통과는 매우 다른 지역간 철도만의 특징인 것이다.

지역간 철도의 다른 특징 중의 하나는 복합열차라는 개념이 존재한다는 것이다. 복합열차는 특정 구간에서는 하나의 차량으로 묶어서 운행하고 나머지 구간에서는 분리하여 운행하는 것을 말한다. 예를 들면, 새마을의 경우 서울~울산, 서울~포항의 서로 다른 경로를 운행하는 차량을 서울~경주 구간은 하나의 차량으로 묶어서 운행하고, 경주~포항, 경주~울산 구간은 각각 분리하여 운행하고 있다. 복합열차라는 개념이 도입된 것은 혼잡한 구간에서는 선로용량의 여유가 없기 때문에 선로용량을 최대한 효율적으로 이용할 수 있도록 하기 위해서이다. 각각의 차량이 따로 운행할 경우에는 시간당 10대가 운행할 수 있다고 할 때 이 중에서 2대를 복합열차로 편성하게 되면 선로용량이 1회 여유가 생겨 다른 구간에 차량을 투입할 수 있게 되는 것이다. 위에서 예를 든 복합열차를 모형에 반영하기 위해서는 서울~경주 구간은 1회의 차량이 운행하는 것으로 하고, 경주~포항, 경주~울산 구간에서는 각각 1회의 차량이 운행하는 것을 반영할 수 있어야 한다.

지역간 철도 통행의 또 하나의 특징은 환승수요가 매우 적다는 것이다. 이것은 비단 지역간 철도에만 적용되는 것이 아니라 지역간 통행의 일반적인 특성이기도 하다. 2003년 지역간 1일 총 통행량을 보면 목적 통행량은 12,252,974명/일, 수단 통행량은 12,269,273인으로 조사되어 목적 통행량 대비 수단 통행량의 비율은 1.00133배에 불과하여, 최대 0.133%가 수단간 환승을 한다는 것을 의미한다. 이것은 지역간 통행에서는 접근수단을 제외한 수단의 선택이 단일수단으로 한정된다는 것을 의미하며, 다시 말하면 지역간 통행에서는 수단간 환승이 거의 발생하지 않는다는 것을 의미한다. 이러한 현상은 2003년 수도권의 목적 통행량이 47,756,792통행/일, 수단 통행량은 59,985,497통행/일로서 목적 통행량 대비 수단 통행량의 비율 1.256배와 대조적이다.

지역간 철도에서 전체 통행에서 차지하는 환승수요의 비율은 〈표 1〉에서 보는 바와 같이 1.6% 수준에 불과하다. 환승수요 중에서도 KTX와 일반철도간의 환승이

〈표 1〉 지역간 철도의 환승수요

구분	총수요 (A)	환승수요								합계 (B)
		동일 차종간			다른 차종간					
		무-무	기타	소계	K-새	K-무	K-통	기타	소계	
수요	322,419	755	122	877	921	2,432	465	416	4,234	5,111
비율 (%)	A 대비	0.2	0.0	0.3	0.3	0.8	0.1	0.1	1.3	1.6
	B 대비	14.8	2.4	17.2	18.0	47.6	9.1	8.1	82.8	100.0

주: 2004년 10월 자료의 일평균 수요

75% 수준이며, 일반철도간의 환승은 직결 서비스가 제공되지 않는 기종점 쌍 사이에 일부 발생하는 것으로 매우 낮은 수준이다. 이것은 지역간 철도가 거의 대부분의 지역에 직결 서비스를 제공하고 있기 때문이기도 하다.

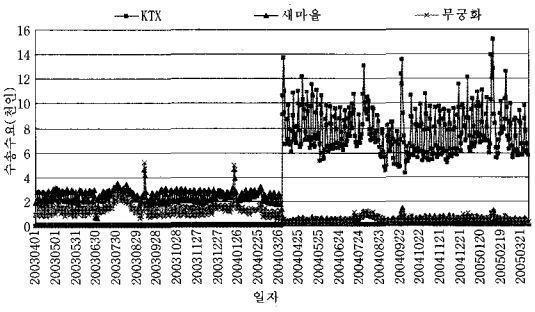
지역간 철도의 특징으로부터 지역간 철도의 경로구간에 대해서 차종별 분담률을 결정하는 것은 기존의 대중교통 통행배정모형을 활용하는 것은 효율적이지 못하다는 것을 알 수 있다. 또한 지역간 철도는 경로구간을 운행하는 노선이 매우 많으므로 경로구간을 운행하는 노선별 분담률을 산정하는 것은 비효율적이며, 보다 집합화하여 차종별 분담률을 산정하고 동일한 차종에 속하는 노선의 분담률은 빈도에 비례하는 것으로 설정하는 것이 합리적이다. 그리고 지역간 철도에서 환승수요는 큰 의미가 없으므로 환승수요는 환승역을 기준으로 하여 통행을 구분하여 분석하는 것이 합리적이다.

2. 차종별 수요분담률

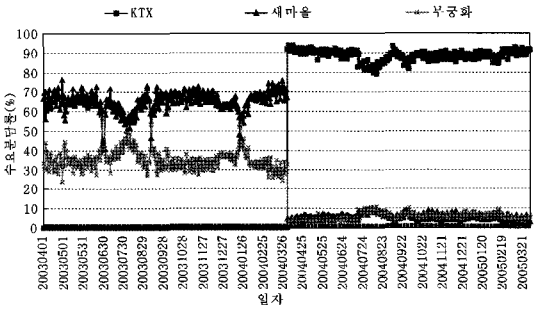
〈그림 2〉는 고속철도 개통 전후(2003년 4월 1일~2005년 3월 31일) 서울~부산 구간의 차종별 수요를 2년 동안의 자료로 분석한 것이다. 본 연구에서는 상·하행 방향별로 별다른 차이점이 없으므로 하행을 기준으로 분석하였다. 〈그림 2〉에서 보는 바와 같이 요일별 변동이 뚜렷한 것으로 분석되었다.

반면에 지역간 철도 통행에 있어 경로구간의 차종별 통행량 분담률은 상당히 안정적인 경향을 나타낸다. 〈그림 3〉에서 보는 바와 같이 각 차종별 분담률은 일률적으로 평균값에 몰리는 경향을 볼 수 있다. 일반적으로 열차운행횟수의 변화, 명절, 휴일 등 대수송기간의 특성, 수해 등 천재지변으로 인한 열차운행 중단, 요일별 특성 또는 휴가, 방학과 같은 계절적 특성 등 다양한 요인들이 분담률에 영향을 주게 되는데, 이러한 변동 요인들이 포함되지 않는 일상적(routine) 통행시기를 조사한다면, 경로구간의(철도역간) 차종별 분담률은 더욱 안정적인 값을 가질 것으로 예상할 수 있다.

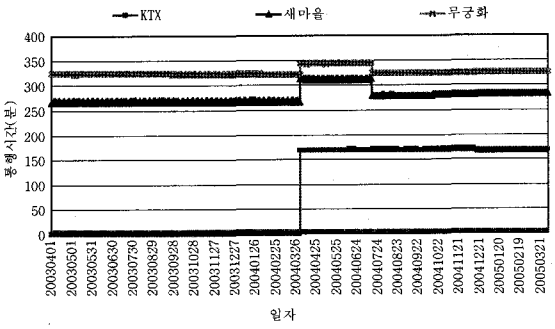
〈그림 4〉와 〈그림 5〉에서는 고속철도 개통 전후 동일기간의 지역간 철도의 차종별 수요 분담률에 크게 영향을 줄 것으로 판단되는 차종별 통행시간과 열차운행횟수를 분석하였다. 일상적 통행에 영향을 주지 않는 즉, 변화가 일어나지 않는 기간을 추려본 결과, 통행시간과 열차운행횟수가 어느 정도 일정한 양상을 유지하는 2003년 10월 6일부터 12월 14일(고속철도 개통 전)까지,



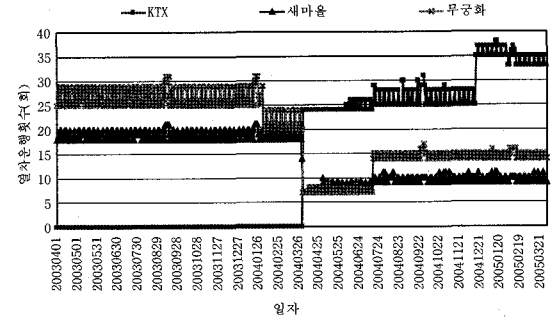
〈그림 2〉 고속철도 개통 전후 2년간의 서울-부산 구간의 차종별 수송수요 분석



〈그림 3〉 고속철도 개통 전후 2년간의 서울-부산 구간의 차종별 수요분담률 분석



〈그림 4〉 서울-부산 구간의 차종별 통행시간



〈그림 5〉 서울-부산 구간의 차종별 열차운행횟수

또한 2004년 10월 4일부터 12월 12일(고속철도 개통 후)까지 각각 10주에 대해서 분석하였다.

〈표 2〉는 고속철도 개통 전후 차종별 수요 분담률의 통계치를 나타내는데 자료수 70개는 10주간의 요일별 데이터를 총합한 표본의 수를 의미한다. 고속철도 개통 전에는 새마을과 무궁화가 모두 평균에 비해서 표준편차가 매우 작은 것을 알 수 있으며, 고속철도 개통 후에는 새마을과 무궁화의 절대적인 분담률의 크기가 작아서 상대적으로 평균에 비해 데이터의 범위가 넓게 분포하는 것을 나타내고 있다.

〈표 2〉의 데이터를 이용하여 차종별 수요분담률에 대한 밀도함수를 그려본 결과는 〈그림 6〉 및 〈그림 7〉에 제시하였다. 전반적으로 평균값을 기준으로 매우 밀집된 형태를 보이고 있으며, 정규분포와 유사한 형태를 가지는 것으로 분석되었다.

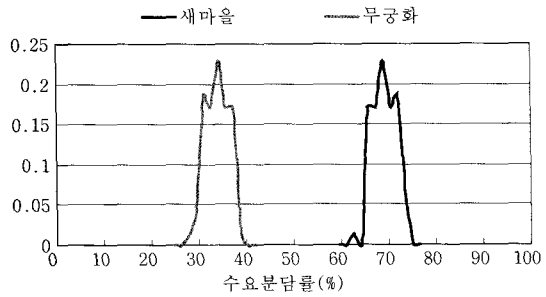
실질적으로 요일별로 열차운행횟수가 변하기 때문에 〈표 3〉에서 제시된 것처럼 비슷한 특성을 가지는 요일별(월·금/화·수·목/토·일)로 나눠서 분석하면 표준편차의 차이를 보다 명확히 알 수 있다. 일반적으로 차종별 수요분담률에 영향을 주는 요인들이 변하지 않는 조건에서 10주 이상의 자료를 이용하여 분석한다면 보다 더 명확한 자료가 제시될 수 있지만 현재까지의 결과로도 수요분담률은 안정적인 것으로 가정하는 것은 타당하다고 판단된다.

결과적으로 수요 분담률은 특정한 구간에 대해서는 평균값이나 표준편차 또는 밀도함수로 표현하게 되면 〈그림 6, 7〉에서 보는 바와 같이 정규분포의 양상을 띠게 된다. 본 연구에서는 서울-부산 구간에 대한 분석만을 제시하였으나, 기타의 구간에 대한 분석에서도 동일한 결과를 얻을 수 있었다. 이상의 결과로부터 경로구간의 차종별 분담률의 분석은 적정한 확률선택모형을 활용하는 것이 가능한 것으로 분석되었다.

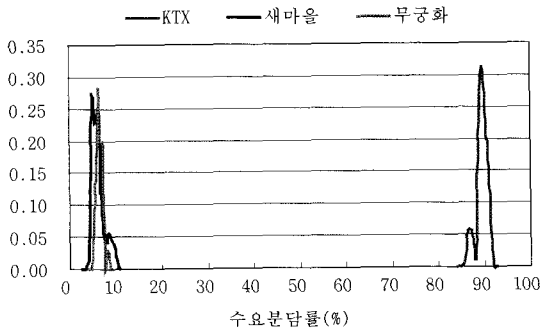
〈표 2〉 고속철도 개통 전후 10주간의 차종별 수요분담률의 통계치(서울-부산)

구분	차종	자료수	평균	표준편차	최대값	최소값	범위
고속철도 개통 전	새마을	70	67.6	2.3	72.8	61.9	10.9
	무궁화	70	32.4	2.3	38.1	27.2	10.9
고속철도 개통 후	KTX	70	88.6	1.3	90.7	85.3	5.4
	새마을	70	5.4	1.3	8.9	3.7	5.2
	무궁화	70	6.0	0.7	7.9	4.7	3.2

주 : 고속철도 개통 전 2003.10.6~12.14, 개통 후 2004.10.4~12.12의 70일 자료를 기준으로 분석



〈그림 6〉 고속철도 개통 전 수요분담률에 대한 차종별 밀도함수(서울-부산)



〈그림 7〉 고속철도 개통 후 수요분담률에 대한 차종별 밀도함수(서울-부산)

〈표 3〉 고속철도 개통 전후 요일별, 차종별 수요분담률 통계치 (서울-부산)

(a) 화, 수, 목요일

구분	차종	자료수	평균	표준편차	최대값	최소값	범위
고속철도 개통 전	새마을	30	69.6	1.6	72.8	65.9	7.0
	무궁화	30	30.4	1.6	34.1	27.2	7.0
고속철도 개통 후	KTX	30	89.4	0.8	90.7	87.7	3.0
	새마을	30	4.3	0.4	5.2	3.7	1.6
	무궁화	30	6.3	0.6	7.9	5.3	2.5

(b) 월, 금요일

구분	차종	자료수	평균	표준편차	최대값	최소값	범위
고속철도 개통 전	새마을	20	66.9	1.3	68.4	64.6	3.8
	무궁화	20	33.1	1.3	35.4	31.6	3.8
고속철도 개통 후	KTX	20	87.4	1.3	89.1	85.3	3.8
	새마을	20	6.4	1.5	8.9	4.5	4.4
	무궁화	20	6.2	0.6	7.7	5.1	2.6

(c) 토, 일요일

구분	차종	자료수	평균	표준편차	최대값	최소값	범위
고속철도 개통 전	새마을	20	65.4	1.6	68.6	61.9	6.7
	무궁화	20	34.6	1.6	38.1	31.4	6.7
고속철도 개통 후	KTX	20	88.8	0.8	90.2	86.4	3.8
	새마을	20	5.9	0.6	7.9	4.9	3.0
	무궁화	20	5.4	0.4	6.0	4.7	1.2

주 : 고속철도 개통 전 2003.10.6~11.30, 개통 후 2004.10.6~12.12

Ⅲ. 다항로짓모형의 이론적 검토

모든 통행자가 동일한 가치기준과 완전한 정보를 가지고 있다는 결정적 모형의 가정을 보다 일반화하여, "개별 통행자가 링크의 통행비용을 서로 다른 방식으로 인지한다."고 가정하는 것이 확률적 모형이다.

본 연구에서는 경로구간의 차종별 분담률이 정규분포의 의미를 가지게 되어 모형구축이 힘들지만 이를 Gumbel 분포로 가정할 경우 로짓모형 구축을 통해 영향변수에 대한 분석이 가능할 것으로 판단하였다. 확률적 효용은 실제 관측되지 않거나 설명할 수 없는 효용부분으로 로짓모형의 확률적 부분을 구성하는 요소이다. 확률적 효용에 대한 Gumbel 분포 가정으로부터 비관련 대안간 독립성(I.I.A. property) 문제가 제기되고, 이는 모형의 가장 큰 한계점이다. 그러나 본 연구에서는 전혀 특성이 다른 차종에 대한 분석으로서 로짓모형의 한계가 별다른 영향을 주지 않을 것으로 판단하였다.

다음은 경로구간의 각 차종별 효용을 구성하는 관측불가능한 효용에 대한 Gumbel 분포 가정으로부터 로짓모형식을 유도하는 과정을 기술한 것이다.

경로구간의 차종 m의 인지통행비용(C_m)이 다음과 같이 관측 가능한 비용과 관측 불가능한 비용으로 구성된다고 하고, 차종 m의 효용(U_m)은 인지통행비용과 음의 관계에 있다고 하자.

$$C_m = c_m + \epsilon_m$$
$$U_m = V_m + \epsilon_m = -\theta \cdot C_m = -\theta \cdot c_m - \theta \cdot \xi_m$$

여기서, V_m : 관측 가능한 효용
ϵ_m : 관측 불가능한 효용

다수의 차종 중에서 특정 차종 m이 선택된다는 것은 차종 m의 효용이 다른 어떤 차종의 효용보다도 크다는 것을 의미한다. 따라서 차종 m을 선택할 확률($P(m)$)은 다음과 같이 계산된다.

$$P(m) = \Pr(U_m > U_1) \cdot \Pr(U_m > U_2) \cdots \Pr(U_m > U_N)$$
$$= \prod_{j \neq m} \Pr(\epsilon_j < V_m - V_j + \epsilon_m)$$

관측 불가능한 효용은 I.I.A. 속성을 갖고 Gumbel 분포를 따르는 것으로 가정한다. 일반적인 Gumbel 분포의 분포함수와 확률밀도함수는 다음과 같다.

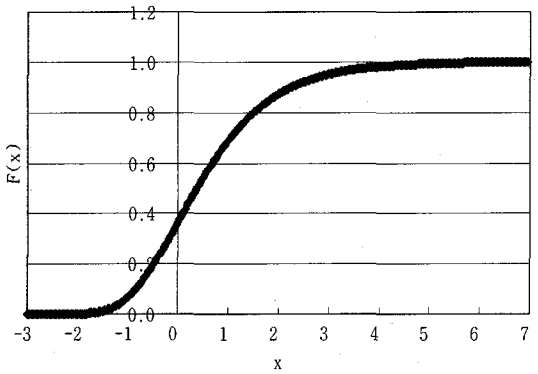
$$F(x) = \exp(-e^{-\frac{x-\mu}{\beta}})$$
$$f(x) = \frac{1}{\beta} \cdot e^{-\frac{x-\mu}{\beta}} \cdot \exp(-e^{-\frac{x-\mu}{\beta}})$$

위의 식에서 $\mu=0$이고, $\beta=1$인 경우가 표준 Gumbel 분포인데, 이를 활용한다면 표준 Gumbel 분포의 분포함수와 확률밀도함수는 다음과 같다.

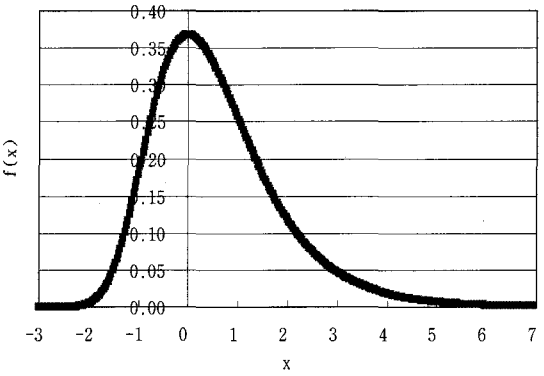
$$F(x) = \exp(-e^{-x}), \quad f(x) = e^{-x} \cdot \exp(-e^{-x})$$

이제 차종 m을 선택할 확률은 다음의 식과 같이 나타낼 수 있다.

$$P(m) = \prod_{j \neq m} \exp(-e^{-(V_m - V_j + \epsilon_m)})$$



〈그림 8〉 표준 Gumbel 분포의 분포함수



〈그림 9〉 표준 Gumbel 분포의 확률밀도함수

관측 불가능한 효용에 대한 확률밀도함수를 이용하여 차종 m을 선택할 확률의 기대값은 다음과 같이 계산할 수 있다.

$$p_m = \int_{-\infty}^{\infty} \exp[-e^{-x}\sum_{j \neq m} e^{(V_j - V_m)}]e^{-x}\exp(-e^{-x})dx$$

위의 식을 단순화시키기 위해서 $y=e^{-x}$로 치환하면 다음이 성립한다.

$$\frac{dy}{dx} = -e^{-x},$$
$$x \to \infty \Rightarrow y \to 0, \quad x \to -\infty \Rightarrow y \to \infty$$

따라서 위의 식은 다음과 같이 풀 수 있다.

$$\begin{aligned} p_m &= \int_{-\infty}^{\infty} \exp[-e^{-x}\sum_{j \neq m} e^{(V_j - V_m)}]e^{-x}\exp(-e^{-x})dx \\ &= -\int_{\infty}^{0} \exp(-y \cdot \sum_{j \neq m} e^{(V_j - V_m)}) \cdot \exp(-y)dy \\ &= -\int_{\infty}^{0} e^{-y \cdot [\sum_j e^{V_j - V_m}]} dy \\ &= \frac{1}{\sum_j e^{V_j - V_m}} = \frac{e^{V_m}}{\sum_j e^{V_j}} \end{aligned}$$

위에서 도출된 p_m은 하나의 경로구간(출발역과 도착역 쌍)을 운행하는 모든 차종 중에서 특정 차종 m을 이용하는 확률인데, 이를 모든 경로구간에 대해서 일반화할 수 있다. 이 때 관측 가능한 효용은 관측 가능한 비용으로 표현할 수 있다. 즉, $V_m = -\theta \cdot c_m$이 된다. 그리고 선택확률 p_{od}^m에 의한 차종별 수요 v_{od}^m은 다음의 식과 같이 산출할 수 있다.

$$p_{od}^m = \frac{\exp[-\theta \cdot c_{od}^m]}{\sum_l \exp[-\theta \cdot c_{od}^l]}$$

$$v_{od}^m = v_{od} \cdot \frac{\exp[-\theta \cdot c_{od}^m]}{\sum_l \exp[-\theta \cdot c_{od}^l)]}$$

$$\forall o \in O, \ \forall d \in D, \ \forall m \in M_{od}$$

여기서, O, D : 출발역 및 도착역 집합
M_{od} : od를 운행하는 차종의 집합
θ : 인지분산계수

Ⅳ. 다항로짓모형의 추정 및 검증

1. 분석자료

통행자가 경로구간에서 차종을 선택할 확률은 각 차종이 가지고 있는 특성인 내부적인 요인(통행시간, 운임, 빈도, 좌석수, 운행시간대 등의 특성)과 외부적인 요인(소득수준, 통행거리, 계절적 변동 등)에 의해서 결정될 수 있다. 본 연구에서는 자료 취득의 한계로 인하여 차종별 선택확률은 빈도와 통행시간, 운임에 의해서 결정되는 것으로 가정하고, 선택확률은 다항로짓모형을 이용하여 계산하였다.

〈표 3〉에서 보는 바와 같이 지역간 철도통행에서는 요일별 변동이 뚜렷하게 관측되기 때문에 요일별로 다항로짓모형을 구축하였다. 분석자료는 고속철도 개통 이후의 자료를 선정하였다. 열차운영계획이 안정적인 것으로 분석된 바 있는 2004년 10월 6일부터 12월 12일까지의 10주 자료 중에서 2004년 10월 18일~2004년 10월 24일의 일주일 자료를 활용하였다.

분석에 사용된 자료는 첫째, 분석기간의 일별 역간수요, 둘째, 분석기간의 일별 열차별 정차역, 정차역간 차내시간 등의 열차운행실적, 셋째, 분석기간 중의 운임체계를 활용하였다. 다항로짓모형 구축을 위한 표본자료 추출방법은 특정일에 총 지역간 통행수요 중 무작위추출(random)을 이용하여 총 자료수의 10%를 임의로 추출하였다.

2. 다항로짓모형을 이용한 경험적 모형 추정

빈도와 통행시간, 운임을 변수로 하는 효용함수는 다음과 같은 형태를 취하며, 통근열차의 더미를 0으로 설정하였다.

$$U_a^m = \alpha_1 \cdot FREQ_a^m + \alpha_2 \cdot TIME_a^m + \alpha_3 \cdot COST_a^m + D^m$$

여기서, U_a^m : 경로구간 a의 차종 m의 효용
$FREQ_a^m$: 경로구간 a의 차종 m의 빈도
$TIME_a^m$: 경로구간 a의 차종 m의 통행시간
$COST_a^m$: 경로구간 a의 차종 m의 운임
D^m : 차종 m의 더미

〈표 4〉 요일별 다항로짓모형 구축 결과 I

구분		빈도 (α_1)	차내시간 (α_2)	운임 (α_3)	상수 (D^m)			자료수	ρ^2
					KTX	새마을	무궁화		
월	계수	0.0812	-0.0238	0.000006	-1.6689	-1.3983	-0.4315	27,882	0.402
	t-값	33.9	-18.2	0.5	-15.9	-16.8	-6.8		
화	계수	0.0802	-0.0249	-0.000012	-1.4952	-1.3136	-0.3393	25,241	0.444
	t-값	31.6	-17.4	-1.0	-13.6	-15.2	-5.3		
수	계수	0.0817	-0.0244	-0.000006	-1.5792	-1.3930	-0.4348	25,275	0.434
	t-값	31.9	-17.2	-0.6	-14.1	-15.6	-6.4		
목	계수	0.0809	-0.0248	-0.000005	-1.4429	-1.2687	-0.2764	27,225	0.417
	t-값	32.9	-18.0	-0.4	-13.6	-15.1	-4.5		
금	계수	0.0788	-0.0251	-0.000001	-1.2857	-1.0004	-0.0453	35,833	0.412
	t-값	37.1	-35.1	-0.8	-17.1	-18.6	-1.0		
토	계수	0.0664	-0.0279	-0.000002	-1.6330	-1.2689	-0.2475	42,289	0.313
	t-값	38.6	-40.1	-0.7	-22.3	-23.2	-4.9		
일	계수	0.0660	-0.0263	0.000001	-1.4424	-1.1396	-0.1122	43,576	0.358
	t-값	38.7	-40.5	0.4	-20.2	-21.0	-2.2		
1주일	계수	0.0745	-0.0255	-0.000001	-1.5031	-1.2324	-0.2350	227,321	0.389
	t-값	93.0	-85.6	-0.9	-48.2	-52.9	-11.2		

〈표 5〉 요일별 다항로짓모형 구축 결과 II

구분		빈도 (α_1)	차내시간 (α_2)	상수 (D^m)			자료수	ρ^2
				KTX	새마을	무궁화		
월	계수	0.0811	-0.0243	-1.6399	-1.3721	-0.4246	27,882	0.400
	t-값	33.9	-31.5	-18.4	-20.6	-6.8		
화	계수	0.0803	-0.0237	-1.5598	-1.3704	-0.3552	25,241	0.445
	t-값	31.7	-29.1	-17.1	-20.3	-5.7		
수	계수	0.0818	-0.0237	-1.6134	-1.4233	-0.4432	25,275	0.436
	t-값	32.0	-29.5	-17.2	-20.2	-6.7		
목	계수	0.0809	-0.0242	-1.4682	-1.2909	-0.2827	27,225	0.418
	t-값	33.0	-30.6	-16.7	-19.6	-4.7		
금	계수	0.0787	-0.0249	-1.2939	-1.0073	-0.0471	35,833	0.412
	t-값	37.1	-36.2	-17.4	-19.0	-1.0		
토	계수	0.0664	-0.0276	-1.6415	-1.2764	-0.2495	42,289	0.314
	t-값	38.6	-42.6	-22.7	-23.8	-4.9		
일	계수	0.0659	-0.0263	-1.4384	-1.1362	-0.1113	43,576	0.357
	t-값	38.7	-42.6	-20.4	-21.2	-2.2		
1주일	계수	0.0744	-0.0253	-1.5089	-1.2375	-0.2364	227,321	0.39
	t-값	93.0	-93.1	-49.4	-54.6	-11.3		

일반적으로 운임수준과 통행시간이 증가할수록 분담률은 감소하고, 빈도 즉 열차운행횟수가 증가하면 분담률은 늘어날 것으로 추정하였다. 따라서 추정한 계수가 이와 부합하면 채택하고, 계수의 부호가 다르면 그 변수는 제외하고 분석하고자 하였다. 모형 구축과정의 초기에는 빈도(FREQ), 차내시간(TIME), 운임(COST)을 변수로 하여 모형을 구축하였으며 그 결과는 〈표 4〉에서 제시하였다. 구축한 모형 중 월요일 모형과 일요일 모형에서 운임(COST) 변수에 대한 계수가 양의 값을 가지고, 모든 요일에서 추정계수가 통계적으로 유의하지 않은 것으로 분석되었다.

그 결과 운임(COST) 변수를 제외하고 빈도와 통행시간 변수만으로 모형을 구축하였으며, 그 결과는 〈표 5〉에서 제시하였다. 〈표 5〉에서 제시된 것처럼 빈도와 통행시간이 각각 (+) 및 (-)부호를 가지면서 통계적으로 유의성이 있는 것으로 나타나 열차운행횟수 즉 빈도(frequency)

와 통행시간이 차종선택에 중요한 영향을 미치는 것으로 나타났다. 결과적으로 열차운행횟수의 증감을 통한 열차운영계획의 변화는 요일별 철도수요의 현저한 영향을 주는 중요한 고려대상임을 확인할 수 있었다.

3. 모형의 검증

〈표 5〉에서 구축된 모형을 이용하여 관측통행량과 예측통행량을 Theil의 부등계수와 RMSE(Root Mean Square Error)를 이용하여 비교하였다. RMSE란 실험이나 관측에서 나타나는 오차(Error)를 제곱(Square)해서 평균(Mean)한 값의 제곱근(Root)을 뜻하며, 표준편차(Standard Deviation)개념으로 인식할 수 있다.

또한 Theil의 부등계수(Inequality Coefficient)란 최적모형으로 선택된 모형하에서 미래의 예측값을 구했을 때 그 결과의 정확성을 알아보는 기준통계량으로서 일반적으로 0에 근접할수록 모형의 예측력이 좋은 것이고, 1에 가까울수록 예측값과 관측값이 서로 관련이 없는 값이 되는 것을 의미한다.

$$RMSE = \sqrt{\sum_{a=1}^{N}(\hat{V}_a - V_a)^2 / N}$$

$$Theil's\ U = \frac{\sqrt{\sum_{a=1}^{N}(\hat{V}_a - V_a)^2 / N}}{\sqrt{\sum_{a=1}^{N}(V_a)^2 / N} + \sqrt{\sum_{a=1}^{N}(\hat{V}_a)^2 / N}}$$

여기서, V_a : 관측 통행량

$\hat{V}_a$: 예측 통행량

N : 비교대상 자료 수

요일별로 구축한 모형을 이용하여 분석한 결과를 〈표 6〉에서 제시하였다. 분석결과에 의하면, 차종별로 KTX와 통근열차가 관측통행량과 예측통행량 간에 오차가 가장 적은 것으로 분석되었고, 새마을이 가장 낮은 수준인 것으로 분석되었으나 전체적으로 모형의 예측력은 상당히 우수한 것으로 판단된다.

결과적으로 〈표 6〉에서 제시되 것처럼 Theil의 부등계수가 대부분의 요일별 모형에 대해 0.05 이내의 수치를 나타냄으로써 모형의 안정성이 높은 것으로 판명되었다. 요일별 자료 분석에 의하면, 요일별로 편차가 크지 않기 때문에 일주일 자료를 모두 취합하여 구축한 모형을 이용하여 분석한 결과를 〈표 7〉에서 제시하였다. 그 결과 각각의 일자별로 요일별 모형을 구축한 것과 1주일 자료 전체를 가지고 분석한 것과의 차이는 특별히 나타나지 않았다. 따라서 일주일치 자료로 구축한 모형을 활용하는 것이 적용성에서 보다 유리한 것으로 판단되었다.

〈표 6〉 관측통행량과 예측통행량의 비교 1안

구분		KTX	새마을	무궁화	통근열차	전체
월	관측수	22,849	22,849	22,849	22,849	91,396
	RMSE	9.08	9.89	10.65	3.38	8.73
	Theil's U	0.0325	0.2476	0.0557	0.0358	0.0493
화	관측수	22,062	22,062	22,062	22,062	88,248
	RMSE	4.60	5.46	6.33	2.56	4.94
	Theil's U	0.0261	0.3004	0.0542	0.0341	0.0439
수	관측수	22,101	22,101	22,101	22,101	88,404
	RMSE	4.91	5.49	6.15	1.67	4.87
	Theil's U	0.0252	0.2916	0.0530	0.0248	0.0411
목	관측수	22,052	22,052	22,052	22,052	88,208
	RMSE	4.70	6.11	6.98	2.23	5.32
	Theil's U	0.0244	0.3041	0.0562	0.0303	0.0441
금	관측수	22,274	22,274	22,274	22,274	89,096
	RMSE	7.75	9.09	9.90	3.98	8.01
	Theil's U	0.0308	0.3033	0.0682	0.0530	0.0531
토	관측수	23,032	23,032	23,032	23,032	92,128
	RMSE	9.57	10.11	10.88	3.18	8.98
	Theil's U	0.0333	0.2608	0.0600	0.0320	0.0504
일	관측수	22,849	22,849	22,849	22,849	91,396
	RMSE	9.08	9.89	10.65	3.38	8.73
	Theil's U	0.0325	0.2476	0.0557	0.0358	0.0493

〈표 7〉 관측통행량과 예측통행량의 비교 2안

구분		KTX	새마을	무궁화	통근열차	전체
월	관측수	21,949	21,949	21,949	21,949	87,796
	RMSE	5.86	6.35	7.66	2.14	5.87
	Theil's U	0.0304	0.2735	0.0600	0.0305	0.0484
화	관측수	22,062	22,062	22,062	22,062	88,248
	RMSE	4.76	5.29	6.38	2.70	4.97
	Theil's U	0.0270	0.2810	0.0548	0.0359	0.0441
수	관측수	22,101	22,101	22,101	22,101	88,404
	RMSE	5.13	5.34	6.28	1.85	4.94
	Theil's U	0.0264	0.2747	0.0543	0.0277	0.0418
목	관측수	22,052	22,052	22,052	22,052	88,208
	RMSE	4.88	5.88	7.12	2.35	5.35
	Theil's U	0.0255	0.2828	0.0575	0.0319	0.0445
금	관측수	22,274	22,274	22,274	22,274	89,096
	RMSE	7.74	8.97	10.10	4.05	8.04
	Theil's U	0.0308	0.2987	0.0697	0.0538	0.0534
토	관측수	23,032	23,032	23,032	23,032	92,128
	RMSE	9.87	10.67	10.88	3.00	9.20
	Theil's U	0.0343	0.2838	0.0598	0.0303	0.0516
일	관측수	22,849	22,849	22,849	22,849	91,396
	RMSE	9.82	10.47	10.70	3.08	9.08
	Theil's U	0.0350	0.2698	0.0559	0.0325	0.0511

V. 결론 및 향후 연구 주제

지역간 철도에서 경로구간의 차종별 분담률은 빈도와 차내시간이 각각 (+) 및 (-)부호를 가지면서 통계적으로 유의성이 있는 것으로 나타나 차종선택에 중요한 영향을 미치는 것으로 분석되었으며 신중한 열차운영계획의 매우 중요한 고려대상임을 확인할 수 있었다. 또한 일반적으로 철도수요가 요일별 특성이 뚜렷하게 관측되기 때문에 경로구간의 차종별 분담률을 일주일치 자료를 이용하여 각각 요일별로 분석한 결과, 각 요일별로 모형을 구축한 것과 일주일치 총합 자료 전체 모형과 차이점이 작은 것으로 나타났다. 이러한 영향 변수를 이용한 모형 구축을 통해 철도수요의 예측치와 실제 통행량을 비교함으로써 모형의 적정성을 검토한 것은 의미있는 것으로 이러한 경험적 모형의 추정으로 보다 다양한 전략적·정책적 해석이 가능할 것으로 판단된다.

본 연구에서는 빈도, 차내시간, 운임만을 고려하였으나 차종별 분담률에 영향을 줄 것으로 예상되는 통행거리, 좌석수, 운행시간대, 지역특성 등에 대한 요인과 시계열적 자료를 필요로 하는 국민소득, 계절 변동 등에 대한 요소를 고려한 분석이 요구된다. 또한 비록 그 영향력은 크지 않지만 환승수요를 고려한 모형에 대한 연구도 필요하다.

참고문헌

1. De Cea J., E. Fernández (1989) Transit Assignment to Minimal Routes: an Efficient New Algorithm, Traffic Engineering and Control, pp. 491~494.
2. De Cea J., E. Fernández (1993) Transit Assignment for Congested Public Transport Systems: An Equilibrium Model, Transportation Science Vol. 27, No. 2, pp. 133~147.
3. Dial R. B (1971), A Probabilistic Multipath Traffic Assignment Model Which Obviates Path Enumeration, Transportation Research, Vol. 5, No. 2, pp.83~111.
4. Jayakrishna R., Mcnally M. G., Marar A. G. (1995) Recursive Structure for Exact Line Probabilities and Expected Waiting Times in Multipath Transit Assignment, TRR 1493, pp.78~187.
5. Lam W. H. K., Z. Y. Gao, K. S. Chan, Yang H. (1999) A Stochastic User Equilibrium Assignment Model for Congested Transit

Networks, Transportation Research B(33), pp.351~368.
6. Spiess, H., Florian, M. (1989) Optimal Strategies : A New Assignment Model for Transit Networks, Transportation Research B, Vol. 23, pp.83~102.
7. Spiess H. (1993) Transit Equilibrium Assignment Based on Optimal Strategies: An Implementation in EMME/2, EMME/2 Support Center.
8. Wu J. H., Florian M., Marcotte P. (1994) Transit Equilibrium Assignment: A Model and Solution Algorithms, Transportation Science Vol. 28, No. 3, pp.193~203.
9. 건설교통부(2005), 2004년 국가교통DB구축사업 전국 지역간 여객 기종점 통행량 자료의 현행화, pp.31~34.
10. 건설교통부(2005), 2004년 국가교통DB구축사업 수도권 및 지방5개 광역권 여객 기종점 통행량 자료의 현행화, pp.29~34.
11. 이재섭·김익기(2001), 다이알 알고리즘을 이용한 다수단 대중교통 노선배정기법에 관한 연구, 대한교통학회지, 제19권 제2호, 대한교통학회지, pp.53~60.
12. 이성모·유경상·전경수(1996), 차량용량을 고려한 대중교통 통행배정모형구축에 관한 연구, 대한교통학회지, 제 14권 제 3호, 대한교통학회지, pp.27~44.

♧ 주 작 성 자 : 김경태
♧ 교 신 저 자 : 이진선
♧ 논문투고일 : 2006. 10. 13
♧ 논문심사일 : 2006. 12. 1 (1차)
2006. 12. 26 (2차)
♧ 심사판정일 : 2006. 12. 26
♧ 반론접수기한 : 2007. 6. 30