

KTDB 기반 노선배정의 예측오차 원인과 분석결과 해석

김익기*

한양대학교 교통물류공학과

Practical Interpretation and Source of Error in Traffic Assignment Based on Korea Transport Database(KTDB)

KIM, Ikki*

Department of Transportation and Logistics Engineering, Hanyang University, Gyeonggi 15588, Korea

*Corresponding author: ikkikim@hanyang.ac.kr

Abstract

This study reviewed factors and causes that affect on reliability and accuracy of transportation demand forecasting. In general, the causes of forecasting errors come from variety and irregularity of trip behaviors, data limitation, data aggregation and model simplification. Theoretical understanding about the inevitable errors will be helpful for reasonable decision making for practical transportation policies. The study especially focused on traffic assignment with the KTDB data, and described the factors and causes of errors by classifying six categories such as (1) errors in input data, (2) errors due to spacial aggregation and representation method of network, (3) errors from representing values for variations of traffic patterns, (4) errors from simplification of traffic flow model, and (5) errors from aggregation of route choice behavior.

Keywords: demand forecasting error, OD trip data, traffic assignment, traveler's behavior, transportation network data

초록

이 연구에서는 교통수요예측의 신뢰성에 영향을 미치는 요소와 원인을 검토하였다. 통행의 다양성과 불규칙성, 입력자료 한계, 자료의 집합화, 모형의 단순화가 포괄적 의미에서 교통예측 오차원인이 된다. 또한 불가피하게 존재하는 예측 오차의 이론적 배경을 정확히 규명함으로써 예측결과를 실무적 정책결정에 활용할 시에 합리적 판단을 하는데 도움이 되도록 하였다. 본 연구에서는 특히 노선배정모형의 예측 오차의 요인에 초점을 두고, KTDB 자료기반 분석오차를 6개 항목으로 나누어 설명하였다. 즉, (1) 입력 자료의 오차, (2) 공간 집합화와 네트워크 표현방식에 따른 오차, (3) 교통패턴 변동에 대한 대푯값 설정에 따른 오차, (4) 교통류모형 단순화에 따른 오차, (5) 노선선택 행태 집합화에 따른 오차로 구분하여 설명하였다.

주요어: 수요예측오차, OD 통행자료, 노선배정, 통행자행태, 교통 네트워크자료

교통 분석 상 예측오차의 원인과 불가피성

우리나라의 경우 노선배정분석의 입력 자료가 되는 OD 통행량을 기준년도에서 예측년도에 이르기까지 국가에서 정기적으로 조사 추정하고 또한 현행화하여 제공하고 있다. 즉 방대한 가구통행실태 표본조사를 통해 전수화 과정 및 예측모형의 정산과정과 예측을 거쳐 구축된다. 이와 같은 과정은 매우 기술적이며 또한 시간과 비용이 많이 소요되므로 분석자별로 구축하기가 어렵기 때문에 국가가 대표하여 구축하고 배포하는 것이다. 그 뿐 아니라 정부의 여러 사업추진에 있어 객관적 비교지표와 합리적 의사결정을 위한 표준적 자료가 필요하기 때문이기도 하다. 자료와 분석방법의 표준화는 예측신뢰성 확보라는 교통수요예측의 근본적 목표와 일치하는 방향이어야 할 것이다. 즉 교통수요예측의 목표는 과거 관측 자료와 과학적 모형을 통해 장래에 가장 있음직한 상태를 예측하는 것이다. 이와 같은 교통수요예측의 신뢰성에 영향을 미치는 요소와 원인을 검토하고, 오차를 불가피하게 포함하는 분석 결과를 실무적으로 어떻게 해석할 것인가가 이 연구의 주제인 것이다.

교통수요예측에 있어 수학적 모형 예측치와 관측치간 차이의 요인은 크게 3가지 범주로 나누어 설명될 수 있다. 즉 (1) 통행자 행태의 다양성과 불규칙성 특성으로 인한 요인, (2) 자료 집합화 및 모형 단순화에 따른 요인, 그리고 (3) 모형의 입력자료 구축 시 부정확성에 따른 요인으로 구분하여 설명될 수 있다. 첫 번째 요인범주로부터 설명하고자 한다. 통행자의 통행의사결정 행태는 오직 5가지만 존재한다. 즉 통행빈도의 선택(trip frequency choice), 통행목적지의 선택(destination choice), 교통수단의 선택(mode choice), 노선의 선택(route choice) 그리고 출발시각의 선택(departure time choice)의 5가지 통행선택행태이다. 현재 국제적으로 적용되는 모든 수학적 교통예측모형은 이러한 5가지 통행자 행태를 수학적 모형 또는 시뮬레이션 모형으로 설명하고자 하는 것이다. 가장 보편적인 전통적 교통수요4단계 예측모형은 출발시각 선택을 제외한 앞의 4가지 행태를 순차적 모형구조(Sequential approach)로 설명하고자 한 모형이다(Kanafani, 1983). 순차적 선택행태 가정 하의 각 단계별로는 독립적으로 다양한 모형 적용이 가능한 모형구조이다. 통행행태의 기본 단위를 독립적인 통행으로 분석한 것이 통행기반모형(trip-based model)이라면, 좀 더 예측의 정확성을 높이기 위해 활동과 활동의 연관성을 모형에 반영하고자 하는 것이 활동기반모형이다. 즉 활동과 연계하며 통행간의 상관성을 모형에 반영 분석하고자 하는 것이 활동기반모형(activity-based model)이다. 즉 집에서 나와 다시 집으로 되돌아가기까지의 여행(하나의 tour로 정의되고 있다)을 구성하는 통행고리(trip chain)의 통행간 관계성을 활동과 관련지으며 설명하고자 하는 노력이다. 활동기반모형의 경우도 통행기반의 정통적4단계분석과 같이 순차적인 통행의사결정 모형구조가 적용될 수 있다. 또 다른 예측 정확성 향상을 위한 노력으로 교통패턴의 시간대별 변화를 모형에 반영하고자 하는 노력이다. 즉 출발시각 선택행태를 추가 고려하여 정적모형(static model)에서 동적모형(dynamic model)로 확장하는 것이다. 이와 같이 통행 의사결정행태의 다양한 요소를 현실적으로 묘사하기 위해 꾸준히 모형을 발전시켜 왔지만, 자료수집한계와 모형분석의 용이성을 고려해서 모형의 단순화가 필수적이라는 것도 피할 수 없는 현실적 제약인 것이다. 즉 현실에서는 통행자나 통행관련 의사결정이 다양하고, 또한 교통 환경도 위치와 시간에 따라 다양할 뿐 아니라, 같은 통행자도 매번 통행관련 의사결정이 일관되지 않고 변화하고 있다. 반면에 분석의 용이성과 수학적 완결성을 가져야 함축된 모형으로 현실 교통패턴을 완벽하게 예측하는 것은 원천적으로 한계가 있는 것이다. 이와 같은 통행행태의 다양성과 불규칙적 특성으로 인해 모형의 현실적 설명에 한계가 존재하는 것이며, 이러한 한계는 예측 오차의 불가피성을 의미하는 것이다.

두 번째 예측 오차 요인은 집합화에 기인하는 오차이다. 공간적 측면에서는 개별 통행자보다는 오히려 공간으로 집합화 한 교통존(Traffic Analysis Zone)을 기준으로 통행관련 의사결정 행태를 설명하고자 하는 것이 일반적 교통수요모형이다. 시간적 측면으로는 하루를 주기로 일어나는 통행패턴을 설명하고자 대표적 하루 기간의 집합적 통행패턴을 예측하거나, 혹은 좀 더 세분화하여 오전 첨두, 오후 첨두 및 비첨두로 하루를 3개의 대표 시간대로 구분 집합화하여 분석하기도 한다. 통행자 속성과 환경측면에서 다양한 통행자들을 일일이 별도 분석을 할 수가 없으므로 통행자들의 사회경제적 속성과 교통적 환경을 집합적으로 묶어 대표적 통행자의 행태모형으로 교통패턴을 설명하

고 있다. 모형 구축 측면에서는 통행에 영향을 주는 모든 요인을 설명변수로 포함시킬 수가 없고 또한 정확하게 파악하기도 어렵다는 한계성으로 변수의 수나 수학적 함수형태를 단순화하여 모형을 구축하고 정산을 하고 있다. 따라서 현재 개발된 모든 교통예측모형들은 단순화함에 따른 누락되는 요인들에 의한 오차와 집합화로 인한 집합화 오차(aggregation error)를 가질 수밖에 없는 것이다. 또한 이렇게 구축된 수학적 모형이 공간적 시간적 전이성(transferability)이 있어 한 공간에서 잘 맞는 모형이 다른 공간에서도 잘 현상을 설명한다는 보장이 없으며, 또한 과거와 현재의 교통현상을 기반으로 정산된 모형이 예측년도에도 잘 현상을 설명한다는 보장도 확신하기 어렵다는 점이다. 이와 같이 모형의 단순화 및 자료의 집합화를 통해 계산과 적용 용이성을 고려하였음에도 불구하고 공간적 범위와 시간적 범위가 방대하여 정확한 수학적 해(solution)를 찾기 어려울 때가 있어 계산상의 오차도 발생하기도 한다.

셋째 예측 오차의 요인은 입력자료의 오차에 기인하는 예측 오차이다. 즉 모형정산에 사용되는 기초 입력 자료와 예측시의 설명변수 입력 자료 자체에 내포되어 있는 오차를 의미한다. 이러한 오차는 조사과정 및 예측과정에서 발생하게 된다. 장래의 토지이용 체계나 교통체계의 변화에 따른 교통패턴의 변화를 예측하기 위해서는 과거와 현재의 관측 자료를 통해 교통패턴에 영향을 주는 요소와 교통패턴과의 상관관계 원칙을 모형으로써 정산하여 구축하는 일이 우선되어야 한다. Ortuzar and Willumsen(2011)이 표본오차의 내용을 구분하여 설명하였듯이 표본조사 과정에서 sampling bias and sampling error가 존재하게 된다. 즉 조사 설계, 조사 집행, 자료집계 및 코딩 과정에서의 연구자 오류에 따른 sampling bias는 철저히 준비하고 집행함으로써 줄일 수 있는 오류이지만, 모집단 자료가 아닌 표본 자료이기 때문에 발생하는 sampling error는 표본율을 높이지 않고는 오차를 줄일 수 없는 내용인 것이다. 즉 sampling bias는 분석자들의 노력으로 오차를 최소화 할 수는 있어도 sampling error는 전수를 조사하지 않고서는 항상 표본오차는 존재한다는 것이다. 이것은 모형 정산에 필요한 입력 자료 자체에서부터 오차가 발생한다는 것을 의미하는 것이다. 또한 예측 시의 입력 자료인 설명변수 값 예측 역시 미래의 불확실성으로 인해 예측 상의 오차가 발생할 여지가 매우 높다. 그 결과 이렇게 구축된 모형에 의한 예측 교통패턴에도 오차가 발생하게 되는 것이다.

이상에서 설명한 3가지 요인범주로 인해 교통패턴 예측에는 불가피하게 오차가 존재할 수밖에 없다. 하지만 이와 같은 예측방법이 가장 과학적이고 합리적 예측으로 불확실한 미래를 정확히 예측할 수 있는 다른 방법도 현재는 없는 상태이다. 그러므로 이와 같이 오차를 품고 있는 예측에 대한 이해와 현명한 해석을 통해 정책 의사결정과정에서 적절히 활용할 수 있는 능력을 갖추는 것이 더욱 중요할 것이다. 특히 KTDB 기준년도와 예측년도의 OD자료와 네트워크 자료가 제공되고 있는 우리나라의 경우는 경직되게 정해서 준 자료로 예측된 결과이므로 오차에 의한 왜곡된 정책 결정에 이르더라도 아무도 의심 없이 획일적인 결론에 도달하게 하는 위험이 있다. 이론적 이해와 결과에 대한 해석적 능력에 따라 탄력적으로 다양하게 분석하는 것보다 원칙적으로 제약된 상황에서는 오히려 예측오차가 더 커질 수도 있다. 이렇게 조작 우려로 공식 배포된 자료에 대한 융통적 수정이 제약된 환경에서는 오차의 요인들을 명확하게 이해함으로써 정책결정에 실수를 하지 않기 위한 해안을 갖추는 것이 더욱 필요할 것이다.

본 연구에서는 교통수요분석 가운데에서도 노선배정분석에만 초점을 두고자 한다. 이것은 KTDB에서 통행목적별 및 교통수단별 OD자료가 주어져 있기 때문에 개발계획 및 사업으로 인한 교통수단 전환이 없는 한 통행발생, 통행분포 및 교통수단선택 분석은 이루어지지 않고 노선배정 분석만 이루어지는 경우가 많기 때문이다. KTDB 자료에 기반 한 노선배정 분석 상에 오차를 6개 항목으로 나누어 설명하고자 한다. 즉 (1) 입력 자료의 오차, (2) 공간 집합화와 네트워크 표현방식에 따른 오차, (3) 교통패턴 변동에 대한 대푯값 설정에 따른 오차, (4) 교통류모형 단순화에 따른 오차, (5) 노선선택 행태 집합화에 따른 오차로 구분하여 설명하고자 한다.

노선배정 입력 자료의 오차

현재 우리나라에서 중앙정부 및 지방정부의 교통정책 분석 및 교통시설투자정책의 타당성 분석에 보편적으로 적용하고 있는 노선배정 분석기법은 와드롭의 제1법칙(Wardrop's first principle)을 만족시키는 결정적 사용자균형(Deterministic User equilibrium) 상태로 경로선택이 이루어진다는 원칙을 바탕으로 하고 있다. 즉 특정 출발점에

서 도착점 사이의 선택된 경로 간에는 통행시간이 동일하며, 선택되지 않은 경로는 선택된 경로보다 느리며, 통행자 누구도 자신의 경로를 바꾸어 더 빨라질 수 없는 상태인 와드롭의 제1법칙에 따라 노선을 배정하게 되는 것이다. 이 모형은 확률적 모형이 아닌 결정적 모형이므로 대안 경로가 조금이라도 빠르면(예로 1초라도) 노선 전환이 일어나게 됨에 따라 중국에는 선택된 경로 간의 통행시간이 동일할 때까지 조종되면서 균형상태에 이르게 하는 모형이다. 그리고 Sheffi(1985)에서 서술되었듯이 특정한 가정을 할 경우 수학적 완결성을 만족시키게 되므로 수많은 링크 교통량이라는 모르는 변수 값(value of unknown variables)에 대해 유일해(unique solution)가 존재하고 있음이 수학적으로 증명되어 있다. 그 의미는 동일한 조건 하에서 어떤 분석가가 분석하여도 항상 동일한 분석 결과가 나온다는 것을 뜻하는 것이다. 이와 같은 수학적 완결성에 따른 유일해 존재라는 특성으로 인해 현재에 이르기까지 국제적으로 가장 보편적으로 많이 사용되고 있는 분석 기법인 것이다.

결정적 이용자균형모형은 2개 입력 자료를 이용하며 분석이 이루어진다. 그 하나는 pcu(passenger car equivalent unit) 단위의 차종별 OD 통행량이며, 다른 하나는 각 링크의 VDF(volume delay function) 함수와 그 계수 값을 포함하고 있는 네트워크 자료인 것이다. 한국교통연구원은 통행목적별, 차종별로 기준년도와 예측년도의 OD 통행량 자료를 정기적으로 자료를 갱신하며 공식적으로 제공하고 있다. 노선배정 분석의 입력자료인 KTDB OD 자료의 오차 원인에 초점을 두고 이론적 쟁점 사항에 대해 논의하고자 한다.

교통예측모형의 정산을 위한 기초적 자료의 수집과 분석 과정에서 발생할 수 있는 오차 요인은 생각하고, 통행발생, 통행분포 및 교통수단선택의 예측모형 정산 상에서 발생할 수 있는 오차 원인의 쟁점 사항에 대해 초점을 두고 논의하고자 한다. 현재 KTDB의 통행발생 예측모형은 가구통행실태 표본자료를 전수화한 P-A(통행생성-통행유인, trip production-trip attraction) 자료와 O-D(통행유출-통행유입, origin-destination)를 입력 자료로 하여 모형정산을 수행 분석하고 있다. 다만 광역권의 경우 교통존 단위로 총량 통행생성량을 추정하는 통행생성모형(trip production model)과 통행유인모형(trip attraction model)을 정산하고, 전국 모형에 있어서는 통행발생모형(여기서는 통행유출모형이라고 표현함)과 통행도착모형(여기서는 통행유입모형으로 표현함)을 정산하여 통행발생 예측모형을 구축했으며, 광역권의 경우는 2010년부터 통행생성모형과 통행유인모형을 정산하여 통행발생 예측모형을 구축하였다. 즉 교통존 단위의 총 통행량과 교통존의 집합적 사회경제지표 간의 회귀분석모형에 의해 예측모형이 정산된 것이다(교통존기반 총량통행모형, zonal-based trip total model). 하지만 가구(household)단위의 자료에 의해 각 가구의 통행생성량을 각 가정의 사회경제지표 자료를 가지고 회귀분석모형에 의해 예측모형을 정산하는 가정기반의 비집계모형(household-based disaggregate model)이 가능하고, 또 다른 방법으로는 교통존 기반으로 하되 모든 변수를 해당 교통존의 가구 수로 나누어 평균적 가구당 통행생성량과 평균적 사회경제지표로 정산한 예측모형(교통존기반 평균통행모형, zonal-based trip rate model)이 가능하다.

다양한 입력자료 형태에 따른 예측모형 정산방법에 대해 Ortuzar와 Willumsen(2011)과 Douglas와 Lewis(1970a, 1970b, 1971a, 1971b)의 4개 연속 논문에서 이론적으로 가정기반의 비집계모형이 이론적으로 가장 우수하며, 집계모형(aggregate model)인 교통존 기반의 총량통행모형이나 평균통행모형 모두 통계학적인 이론적 단점이 있음을 지적하고 있다. 즉 Washington et al.(2003) 및 여러 통계학 책에서 설명되고 있듯이 회귀분석의 오차항이 이분산(Heteroscedasticity)이거나 설명변수간의 다중공선성(multicollinearity)이 존재할 경우 추정된 계수의 분산이 커져서 참 값으로부터 오차가 큰 계수 값으로 추정될 확률이 높아짐을 지적하고 있다. 교통존 기반의 집계자료를 사용할 경우 회귀식에서 인과관계적 설명보다는 교통존의 크기에 따른 높은 상관관계로 설명변수간의 다중공선성이 존재할 수가 있으며 또한 교통존 크기에 따라 오차항 크기에 차이가 생기는 이분산 문제가 있게 된다. 집계자료를 가지고 회귀분석 시 설명변수와의 인과관계로 설명되기 보다는 교통존의 크기에 의해 설명되는 부분이 많이 설명됨에 따라 높은 R^2 값이 나오게 되어 착시적인 실수를 야기할 수가 있게 된다. 오차가 큰 계수 값의 통행발생예측모형이 적용한다는 점은 당연히 통행발생량 예측의 오차 원인 중에 하나가 될 수 있는 것이다.

국제적으로 비집계모형이 보편적으로 사용하고 있음에도 우리나라에서는 아직도 교통존기반의 집계모형이 지속

적으로 적용되고 있는 이유는 예측년도의 설명변수 예측이 비집계로 이루어질 수 없다는 이유가 그 하나인데 Manheim(1979) 및 Ben-Akiva와 Lerman(1985)에서 비집계자료에 의해 모형 정산된 예측모형을 집계자료 형태로 예측할 때의 방법으로 통합대표값 적용방법(Nave approach) 외에도 분류 적용방법(classification approach) 등 다양한 분석방법을 제시하고 있어 교통존 단위로의 예측은 가능하다. 특히 선형회귀분석과 같은 선형모형을 적용할 경우는 Anas(1982)에서 설명한 비선형함수일 때 젠슨의 부등관계(Jensen's inequality)로 발생하는 집합화 오차(aggregation error)가 없이 집계형태의 예측이 가능하므로 더욱이 큰 문제가 되지 않는 것이다. 비집계자료를 이용하여 예측모형을 정산할 경우 회귀분석에서 R^2 값이 적게 나와 의사결정자를 통계학적으로 설득하는 것이 어렵다는 것도 또 다른 이유가 될 것이다. 하지만 가구단위의 자료에 의한 R^2 값이 아니라 교통존 단위로 R^2 값을 산출할 경우는 교통존기반의 집계모형에서 계산된 R^2 값처럼 높은 값이 나올 것이므로 이것도 문제가 되지 않을 것이라 생각된다.

통행분포모형의 경우 KTDB에서는 광역권 지역에 대해서는 중력모형(Gravity model)을 전국에 대해서는 프라타모형(Fratar model)이 적용되고 있다. 전국 지역에 대한 프라타모형의 경우는 정산이 필요 없이 기준년도의 통행분포패턴을 기준으로 예측된 통행유출량과 통행유입량에 따른 기준년도 대비 증가율에 의해 계산되므로 특별히 정산 과정이 필요하지가 않다. 하지만 광역권에 대한 중력모형은 통행목적별로 적합한 저항함수(impedance function)을 선정하여 정산하는 과정이 필요하게 된다. 광역권의 중력모형 통행분포는 각 통행자들의 목적지선택모형에 기반을 두지 않고 거리(통행시간)에 반비례하고 통행유인량과 통행생성량에 비례하게 집합적 개념으로 공간에 분포 배분하는 모형이다. 따라서 개별통행자의 활동체계에 기반을 둔 행태모형이 아니다보니 인과관계 설명이 비교적 부족한 모형이다. 그리고 기준년도에 정산된 모형을 예측년도에 그대로 적용하여 예측할 경우 기존의 통행패턴이 완전히 무시되고 새롭게 수학적으로 배분시켜서 기존 통행의 공간적 분포패턴과 많이 다른 패턴이 나올 수도 있다는 한계가 있다. 따라서 기존 통행의 공간적 분포패턴을 가능한 유지하고자 하는 프라타모형과 새로운 토지이용 및 교통 체계 변화에 따른 영향을 반영할 수 있는 중력모형을 조합적으로 활용하여 예측하는 모형을 적용하는 것도 고려해 볼 필요가 있다. 이와 같이 통행분포 모형의 정산과 예측 적용 방법도 OD자료 구축 상의 오차 원인 중에 하나가 될 수 있다.

KTDB 자료의 교통수단선택모형 정산은 목적통행(Linked trip) 개념의 통행단위를 기준으로 교통존 기반의 수단별 OD자료를 이용하여 로짓모형(Multinomial Logit model)을 정산하여 예측모형을 구축하였다. 비집계자료를 모형정산에 적용할 경우에는 통행자의 사회경제지표를 모형에 포함시킬 수가 있으나 교통존 단위의 집계자료를 적용할 경우에는 통행자의 속성 적용이 어렵게 되고 다만 교통수단별 서비스 속성만이 모형에 포함된다는 한계가 있게 된다. 또한 로짓모형은 선형회귀분석과는 달리 비선형함수 형태를 갖고 있으므로 비집계자료에 의해 모형 정산을 하고 집계단위로 예측할 경우 젠슨의 부등관계로 인한 집합화 오차가 발생하게 된다. 따라서 비집계모형을 적용하여 통행자의 속성이 포함된 교통수단선택 모형정산을 통해 예측모형을 구축하고, 예측시에는 통행자들의 범주를 구분하여 그 범주의 대푯값을 적용하는 분류 적용방법에 의한 예측을 수행하는 것이 집합화 오차를 줄일 수 있다. 이와 같이 교통수단에서의 모형정산과 예측방법의 차이도 역시 OD 구축 상 오차 원인 중 하나가 될 수 있다.

이상에서 노선배정 분석의 가장 기본이 되는 입력 자료인 기준년도와 예측년도의 OD 자료 구축에 있어 통행발생, 통행분포 및 교통수단선택 모형정산과 예측 상에서의 방법론 차이로 오차가 발생할 수가 있음을 이해할 필요가 있다. 또한 이와 같은 기초입력자료 구축 상의 오차 외에도 2절에서 설명하였듯이 통행자 행태의 다양성과 불규칙성 그리고 자료 집합화와 모형 단순화에 따른 오차도 함께 융합되어 있다는 점도 고려하며 노선배정 분석결과에 대한 해석에 주의할 필요가 있다.

공간적 집합화와 네트워크 표현의 단순화

통행자 각 개인 행태를 기준으로 분석모형이 구축되어야 하겠으나, 모든 통행자의 출발점과 도착점을 개인별로 모형화하기가 어렵기 때문에 거주 공간 및 활동 공간을 집합화한 교통존(traffic zone)으로 영역을 구분한다. 그리고 교통존을 하나의 점인 존중심점(centroid)으로 대표시켜서, 모든 통행이 존중심존에서 출발하고 도착하는 것으로 가정하고 있다. 공간의 집합화는 특정 교통정책을 분석하기에 적절한 규모로 교통존 크기를 결정해야 한다는 과제

를 남기게 된다. 즉 공간을 집합화하여 하나의 점으로 대표시킴에 따라 통행은 존간통행(inter-zonal trip)과 존내통행(intra-zonal trip)으로 구분되어진다. 존내통행의 경우는 같은 교통존에 출발/도착 지점이 함께 있어 같은 점에서의 이동으로 표현되어 네트워크 상 교통량이 배정되지 않게 된다. 이것은 가깝게 이동하며 하위기능의 도로를 이용하는 통행은 분석 상에서 제거하여 분석을 단순화하고자 하는 의미를 갖는 것이다. 즉 정책분석의 대상이 되는 상위기능 도로를 이용하게 되는 먼 거리 통행만 분석 대상에 포함시킴으로써 분석을 단순화시키면서도 분석목적인 상위기능 도로의 교통량 패턴 파악에는 크게 오류를 범하지 않는 방법을 추구한 것이다. 따라서 존내 통행이 이용하는 하위기능 도로는 교통존의 크기에 따라 네트워크에서 적합하게 제거되어야 할 것이고, 또한 존간통행이 이용할 것으로 기대되는 상위기능 도로는 네트워크에 포함되어야 할 것이다. 그리고 존중심점에서 네트워크에 연결시키는 링크인 존중심연결선(centroid connector)은 제거된 하위기능의 도로망을 대표하는 가상적 링크로 해석될 수 있다. 이와 같이 교통존 크기와 네트워크의 상세성 정도를 조화롭게 구성해야 하는 교통존과 네트워크의 호환성(zone and network compatibility)를 만족시키지 못할 경우 노선배정에서 잘못된 교통량 추정 분석결과를 야기하게 된다(U.S. DOT, 1983). 예로 교통존 크기가 큰 반면에 상대적으로 네트워크가 상세하게 구축될 경우 존간통행은 적어지는 반면 도로망은 조밀하여 처리 용량이 상대적으로 커지게 됨으로써 각 링크의 추정 교통량은 과소 추정되는 결과를 낳게 된다. 이와 같이 교통존과 네트워크 호환성의 적합성 여부도 노선배정의 오차 원인 중 하나가 된다. 하지만 교통존의 크기와 네트워크의 상세성 정도에 대한 명확한 지침이 있지 않고 각 지역의 상황에 따라 조절되어야 하는 내용이기 때문에 분석자의 전문적 지식과 경험에 의해 적합하게 교통존과 네트워크가 구축될 필요가 있다. 우리나라 교통연구원의 KTDB 자료는 자료 구득의 용이성으로 전국자료는 시군구 단위의 교통존으로 구성하고, 광역권 자료는 읍면동 단위의 교통존으로 구성하여 자료를 제공하고 있으며, 네트워크도 전국권과 각 광역권에 적합하다고 판단되는 수준으로 구축되어 있다. 교통존 설정은 자료 구득의 용이성과 시계열적 일관성을 위해 행정단위로 구축된 것이라 고려된다. 반면에 GIS 및 내비게이션 지도가 점차 정교하게 구축되어짐에 따라 네트워크 상세성은 분석 목적에 따라 얼마든지 상세하게 구축할 수가 있다. 따라서 행정단위의 교통존 크기에 비해 네트워크가 너무 상세하게 구축됨으로써 노선배정결과에 오류를 야기할 가능성도 있는 것이다.

네트워크 구축에 있어 존중심점(centroid)의 위치와 존중심연결선(centroid connector)의 개수와 위치 그리고 속성설정 또한 교통존에 의한 공간적 집합화와 관련된 내용이다. 즉 공간적으로 넓게 퍼져있는 출발/도착지점들을 존중심점 한 점으로 대표됨에 따라 현실과 다르게 교통량이 한 점에 집중하게 된다는 문제가 있게 된다. 그리고 현실에서는 존재하지만 생략된 하위기능 도로들을 대표하는 존중심연결선이 실제와 다른 노선배정을 야기하는 문제도 있게 된다. 그러므로 노선배정 분석결과와의 오차를 최소화 할 수 있도록 존중심점과 존중심연결선의 위치와 개수를 적절히 설정하는 네트워크 구축방법이 필요하다. 한 예로 존중심연결선에서의 통행시간(혹은 속도)를 교통량 크기와 링크 길이와 무관한 고정된 값을 적용할 경우에는 존중심점의 위치가 중요하지 않고 존중심연결선의 속성만 중요하지만, 만일 교통량 크기와 링크 길이에 비례한 통행시간을 적용할 경우에는 존중심연결선의 속성 뿐 아니라 존중심점의 위치도 중요하게 된다. 이와 같이 분석 오차를 야기하는 요인임에도 불구하고, 네트워크 구축 작업이 많은 시간과 노력이 필요한 단순 작업으로 구성되다보니 교통전공 초보자들에 의해 작업이 이루어지는 경우가 있어 KTDB 네트워크에 불합리한 존중심연결선 개수와 위치로 구축된 경우가 발견되기도 한다. 그 뿐 아니라 Easa(1991)에서 설명하였듯이 KTDB 네트워크 대부분에서 적용되고 있는 거시적 네트워크표현(macroscopic network representation) 방식은 미시적 네트워크표현(microscopic network representation) 방식과는 달리 비논리적 경로(illogical movement)가 발생할 수 있어 네트워크 구축시 주의가 필요하다. 이와 같은 네트워크 구축 상의 오류도 노선배정 분석결과와의 오차 원인이 되기도 하는 것이다.

교통패턴 변동에 대한 대푯값

교통수요의 기간적 단위는 반복적 주기가 비교적 명확한 하루 단위를 기준으로 분석하는 경우가 많다. 하지만 하루 통행량이라 할지라도 계절별, 월별, 요일별로 변화하며 일정한 값을 갖지 않는다. 따라서 평균적 개념인 연평균

통행량(AADT, Annual Average Daily Traffic) 혹은 연평균평일통행량(AAWDT, Annual Average Weekday Daily Traffic)을 기준치로 조사 및 예측하고, 계절별, 월별, 요일별, 시간대별 변동은 변동계수값(factor)을 곱하여 특정 하루의 교통량을 추정하여 사용하기도 한다. 그러므로 특정 하루의 교통량을 알아야하는 상황을 제외하고는 일반적으로 AADT 혹은 AAWDT를 기준으로 교통수요를 예측하고 있다. 노선배정 분석도 AADT 혹은 AAWDT 교통량과 비교 검증하며 네트워크 정산(Network calibration)이 이루어지고 있다. 즉 상시조사, 보정조사, 전역조사 및 수시조사를 통해 AADT 혹은 AAWDT 개념으로 추정된 관측교통량과 비교하여 모형을 정산한다. 이와 같은 네트워크 정산에 내재되어 있는 가정은 대표적 수요일로 하루 조사된 KTDB OD와 각 조사링크의 연평균적 통행량과 일치한다고 가정하는 것이다. 하지만 업무지역, 상업지역, 주거지역, 위락지역 등에서의 활동패턴의 차이로 계절별 및 요일별 변동패턴에 차이가 있게 마련이다. 그러므로 가구통행실태 조사일의 대표적 하루 통행패턴이 토지이용패턴에 따라 다른 모든 지점에서의 연평균 통행패턴과 일치한다는 것은 불가능하다. 이와 같이 하루 조사에 기초한 KTDB OD 자료를 이용한 노선배정 분석결과로써 모든 링크의 연평균 교통량을 예측하는 데에는 한계가 있을 수밖에 없으며 이것도 노선배정 예측의 오차 원인이 되는 것이다.

교통패턴 변동에 대한 대푯값 설정의 또 다른 방식으로 대표적 하루 단위로 분석하기 보다는 하루 중 오전첨두(morning peak), 오후첨두(evening peak) 그리고 비첨두(non-peak)로 나누어 대표적 1시간의 교통패턴을 분석하는 것이다. 즉 첨두시의 반복적 체증에 대해서는 교통개선대책이 필요하게 되므로 첨두시를 별도로 분석을 하게 된다. 또한 교통시설 투자 타당성분석에서는 사회적 편익산출 시에 교통패턴에서 차이가 나는 첨두시와 비첨두시에 대해 각각 별도의 구분분석이 필요하기도 하다. 하지만 KTDB 자료는 다만 첨두비율 및 비첨두비율을 제공하고 있다. 그러므로 첨두와 비첨두 시간대 분석이 요구될 때에는 분석자가 하루 OD에 첨두율 혹은 비첨두율을 곱하여 첨두 OD 및 비첨두 OD를 별도로 구축하게 된다. 현재 우리나라에서는 첨두시와 비첨두시를 구분하여 링크 교통량을 추정하는 방법으로 두가지가 지침에서 제시되고 있다. 하나는 OD자체를 첨두와 비첨두 OD로 구분하여 분석하는 방법이고, 또 다른 방법은 지점별 특성에 따라 하루 링크 교통량에 첨두와 비첨두 비율을 곱하여 구하는 방법이다. 첫 번째 방법은 오전 및 오후 첨두 OD 자료와 비첨두 OD 자료가 기준년도와 예측년도에 비교적 정확하게 제공될 때 적합한 방법이다. 두 번째 방법은 첨두 및 비첨두 OD 구분 자료가 없고 지역별 주요지점별 관측 교통량 자료 사용이 가능할 때 적합한 방법이다.

OD자료 자체를 오전 및 오후 첨부와 비첨두로 구분 구축하기 위해서는 Easa(1991)와 Kim(1997, 2006)가 설명하였듯이 PA 자료에서 OD 자료로 전환할 때 PA-to-OD 전환계수(λ_{ij}^p)를 적용하여 통행목적별로 첨두방향이 옹골게 반영 구축되도록 세심한 배려가 필요하다. 하지만 현재 우리나라 KTDB 자료에서는 전환계수 λ_{ij}^p 값은 제공하고는 있으나, 오전 및 오후 첨두 그리고 비첨두로 구분하여 제공하고 있지는 않다. 단지 OD의 첨두비율과 비첨두율만 제공하고 있으므로 하루 OD에 첨두율과 비첨두율을 곱하여 첨두 OD와 비첨두 OD를 구축하게 된다. 하지만 이와 같은 방법은 통행의 방향성도 제대로 반영하지 못하게 된다. 또한 OD별로 출발시간대의 첨두시간대에 차이가 있음에도 불구하고, 전체 지역의 모든 하루 OD에다가 획일적으로 첨두 비율을 적용하는 것을 의미하는 것으로 현실의 통행패턴과 차이가 있게 된다. 그러므로 이와 같은 KTDB 자료의 제약 하에서 첨두 및 비첨두 구분한 분석은 불가피하게 오차를 내포할 수밖에 없게 된다.

두 번째 방법인 노선배정 후 하루 링크 교통량에 첨두 및 비첨두 비율을 곱하는 방법은 관측 시간대별 링크 교통량 자료가 필요하게 된다. 하지만 네트워크상에 상시조사지점의 수는 한계가 있으므로 자료 수집에 제약이 있을 수밖에 없다. 따라서 인접에 위치한 유사 도로위계의 관측값 혹은 동일 지역범위의 평균값을 적용하고 있다. 다만 각 노선배정이 완료된 이후 각 도로위계 혹은 지역별로 첨두비율을 적용함에 따라 첨두 시점에서 상류부와 하류부 교통량 간의 보존법칙(conservation rule)이 만족되지 않는 한계점을 갖고 있다. 관측 자료의 한계 그리고 노선배정결과와 보존법칙이 깨진다는 단점은 오차 원인을 의미하는 것이지만, 그래도 정확한 별도의 첨두 및 비첨두 OD가 제공 안 되는 상황에서는 노선배정 후 첨두 및 비첨두 비율을 하루 링크 교통량에 적용하는 것이 그 중에서는 오차를 줄이는 방안이라고 고려된다.

교통류모형 단순화

우리나라에서는 주로 교통량-지체함수(VDF, Volume-delay function)라고 불리는 교통량과 통행시간과의 관계식은 여러 문헌들에서 다양한 명칭으로 사용되고 있다. 즉 Branston(1976), Boyce et al.(1981), Sheffi(1985), Suh(1990, 1991), Hall et al.(1992), Regueros et al.(1993) 등의 연구에서와 같이 링크성능함수(link performance function), 링크용량함수(link capacity function) 및 교통량-속도관계함수(speed-flow relationship function) 등으로 다양하게 불리고 있다. VDF함수의 명칭에서도 그 의미를 알 수 있듯이 통행자의 노선선택행태와 직접적인 관계가 없는 함수이다. 즉 통행자의 노선선택 행태에 영향을 주는 통행시간을 산출하기 위한 것이 유일하게 노선배정 분석에 적용되는 이유이다. 따라서 VDF함수의 계수 값을 수정하며 노선배정 분석결과와 관측 링크교통량 간의 차이를 줄이는 네트워크 정산은 논리적으로 옳지 않다. 노선배정 정산 시 VDF함수의 계수 값을 수정하는 경우 교통량과 통행시간 간의 관계를 왜곡시켜 기준년도의 관측교통량은 비슷하게 맞출 수는 있어도 예측년도 노선배정 분석결과에 왜곡을 가져와 예측 오차를 야기할 수가 있다. 따라서 네트워크 정산에서 같은 위계 해당의 VDF함수 적용 시 구배, 선형, 노상주차 유무, 차로폭원, 터널 혹은 교량 등 링크의 현지 특성에 따라 도로용량 및 자유류 속도(free flow speed) 속성 조정에 국한하여 조율하고 계수 값은 교통량-속도함수 정산 시 추정된 값을 유지하는 것이 옳다고 고려된다. 그리고 어떠한 VDF함수도 교통량과 통행시간간의 관계를 완벽하게 설명하지 못하고 있는 것 또한 한계이다. 그러므로 적용되고 있는 VDF함수의 한계적 설명력과 정산 시 오류적인 조율이 또 다른 노선배정의 오차원인이 된다.

앞의 절에서 설명한 하루 또는 첨두시와 같은 노선배정분석 시간단위와 VDF함수에 포함된 링크교통량 및 용량의 시간단위 간의 관계에 대한 명확한 이해가 필요하다. 가장 보편적으로 사용하고 있는 노선배정모형은 Sheffi(1985)에서 정리한 백크만의 변환수식(Backmann's transformation formula)으로 정의되는 정적 사용자균형 노선배정 모형(static user equilibrium traffic assignment model)이다. 이것은 정적모형(static model)으로서 시간과 무관하며 모든 통행은 분석기간 내에 출발지에서 도착지까지의 통행이 완료된다는 가정이 내포되어 있다. 즉 하루단위로 분석할 경우 대푯값으로써 평균 1시간의 교통류패턴이 적용된다는 가정을 내포한 정적모형의 특성으로 1시간 단위이든 5시간 단위이든 하루 단위이든 상관없이 하루 단위의 링크 교통량으로 확장하면 노선배정분석 결과가 동일하게 된다. 다시 말해서 첨두 혹은 비첨두 1시간과 같이 별도의 구분된 OD패턴을 적용하지 않고 대푯값 1시간 OD패턴으로 분석한다면 노선배정의 입력 OD의 시간단위는 문제가 되지 않는다. 이것은 1시간 대푯값 OD로 분석하여도 앞 시간대의 교통류패턴과 뒷 시간대의 교통류패턴이 대표 1시간 패턴과 동일하다는 정적분석의 가정과 노선배정이 모든 OD통행량은 출발점에서 도착점까지의 분석시간 내에 통행이 완성된다는 가정 하에서 분석한 것이므로 우리나라 전국을 분석하는데 굳이 5시간 OD를 입력할 필요가 없이 1시간 첨두시 OD 혹은 하루 OD를 입력 자료로 넣고 노선배정분석을 하여도 논리적 문제가 되지 않는다. 다만 각 시간대별로 OD 패턴에 차이가 나는 현실적 분석을 하고자 한다면 시간대별로 다른 OD패턴을 입력 자료로 하는 동적모형(Dynamic model)으로 분석되어야 할 것이다. Branston (1976)은 1시간 단위로 집합화한 조사 자료로 교통량과 통행시간의 관계식을 표현할 때가 가장 안정적인 관계식을 나타내고 있음을 조사결과를 가지고 설명하였다. 그러므로 대표 1시간 단위의 교통량과 용량 간의 관계 즉 V/C 비율을 이용하여 통행시간을 추정하고 있다. 여기서 OD의 하루 단위 통행량 혹은 첨두시 1시간 단위 통행량과 같은 노선배정 상의 분석기간 단위와 VDF함수에 포함된 링크교통량 및 용량의 시간 단위 그리고 그 입력 자료에 의한 노선배정 분석결과에 대한 의미적 해석을 명확하게 하고자 한다. BPR함수가 가장 보편적으로 사용되는 VDF이므로 Equation 1을 가지고 설명하고자 한다.

$$t_l = t_{0l} \left[1 + \alpha \left(\frac{V_l}{C_l} \right)^\beta \right] \quad (1)$$

여기서 t_l : 교통량 V_l 일 때 링크 l 를 통과할 때 소요되는 통행시간

- t_{0l} : 링크 l을 자유류 속도로 통과할 때 소요되는 통행시간
- V_l : 링크 l의 대표 1시간의 교통량 (단위 PCU)
- C_l : 링크 l의 최대 처리 교통량인 1시간 단위 용량 (단위 PCU)
- α, β : VDF 함수의 계수 값

하루 단위의 OD가 입력될 경우 집합적 개념의 경로 선택원칙에 적용될 대푯값 통행시간기준이 정해야 한다. 하루 단위의 OD pcu통행량을 노선배정을 할 경우 Equation 2의 링크 교통량 V_l^w 도 하루단위의 링크교통량이 된다. Equation 2는 하루 단위 집합적 경로선택 행태를 대표적으로 가장 잘 묘사할 수 있는 k_h 값을 찾아서 적용하여야 함을 의미하는 것이다.

$$t_l = t_{0l} \left[1 + \alpha \left(\frac{V_l^w / k_h}{C_l} \right)^\beta \right] = t_{0l} \left[1 + \alpha \left(\frac{V_l^w}{C_l k_h} \right)^\beta \right] \tag{2}$$

여기서 V_l^w : 분석시간 시간단위 w의 OD 입력 자료로 노선배정에 의한 링크 l의 교통량 (단위 : PCU)

k_h : 하루 OD 통행량 중 대표 1시간 통행량이 차지하는 비율 ($0 \leq k_h \leq 1$)

Equation 2에서 좌측 k_h 값은 w 기간 동안 노선선택 행태를 대표할 수 있는 대표적 통행시간이 되게 하는 시간단위 w (예: 하루)의 링크 교통량에 대한 비율 $1/k_h$ 을 의미한다. 따라서 Equation 2 수식에서 우측 k_h 값은 단순히 그 비율을 분모에 옮겨 놓은 것이다. 간혹 우리나라에서는 링크의 시간당 용량 C_l 에 k_h 시간을 곱하여 w 기간의 용량 (예: 하루 용량)으로 전환시키는 계수(factor) 값으로 해석하는 경우가 있는데 이것은 잘못된 해석이다. Horowitz(1993)는 하루 시간대별 통행시간의 기댓값(expected value)을 하루 OD의 노선선택 행태에 적용하는 수식을 제안하기도 하였다. 즉 Equation 3과 같은 VDF함수를 제시하였다.

$$E(t_l) = t_{0l} \left[1 + \alpha \left(\frac{V_l}{C_l} \right)^\beta \left(\sum_{h=1}^H p_h^{\beta+1} \right) \right] = t_{0l} \left[1 + \sum_{h=1}^H \alpha \left(\frac{V_l p_h}{C_l} \right)^\beta p_h \right] \tag{3}$$

여기서, p_h : 통행이 하루 중 h시간대에 네트워크 상 노선배정 될 확률

Equation 3의 p_h 는 실질적으로 Equation 2의 각 시간대별 $1/k_h$ 값과 같은 의미를 갖는다. 하지만 Horowitz가 제시한 방법은 링크별로 모든 시간대의 p_h 값을 알아야 하는데 이것은 현실적으로 매우 어려워 우리나라에서는 사용되고 있지 않다. 그렇다면 하루의 집합적 경로선택행태를 가장 잘 설명할 수 있는 통행시간을 산출하는 k_h 값을 찾아야하나 이에 대한 기초연구는 없는 실정이다. 그러므로 하루 OD를 적용할 때 일반적으로 첨두시의 k_h 값을 적용하고 있다. 이것은 “하루 시간대로 집합화 된 통행자들 모두가 첨두시 통행시간을 기반으로 노선선택을 한다.”라는 가정 하에 분석을 하는 것이다. 노선배정의 대표 시간대의 설정 또한 노선배정의 오차를 야기하는 한 요인이 된다.

Sheffi(1985)가 수학적으로 증명 설명한 바와 같이 정적 사용자균형 노선배정 모형이 유일해(unique solution)를 갖기 위한 조건 중 하나는 VDF함수가 단조증가함수(monotonically increasing function)이어야 한다. 하지만 현실은 단위시간 용량 이상의 교통량 처리는 가능하지 않고 대기행렬 현상이 일어나게 된다. 이와 같은 대기행렬이 형성되게 되는 교통류의 불안정상태에서는 Figure 1에서 표현한 후술현상이 관측된다. Branston(1976)이 설명하였듯이 링크의 유입지점의 용량과 링크 유출지점의 용량 그리고 시간대별 유입 교통량의 변화에 따라 대기행렬이 생성하고 소멸하는 과정이 형성되고, 각 과정별 상황에 따라 대기행렬에서 소요되는 통행시간이 차이가 난다. 정적 노선배정

모형은 시간대별 분석이 가능하지가 않으므로 동적모형과는 달리 대기행렬의 생성과 소멸 현상은 설명이 가능하지 않다. 따라서 정적모형에서는 용량이상 교통량 상태 시 대기행렬에 의한 지체를 간접적으로 노선선택 행태에 반영하고자 일종의 벌점(penalty) 개념으로 통행시간을 가중시키는 방법을 적용하고 있다. 즉 Figure 1 VDF 함수곡선의 용량이상에서 지수곡선 형태로 급격히 증가하는 형태를 갖는다. 용량 이전의 곡선은 링크 교통량과 속도의 관측치에 의해 모형이 정산된 값으로 교통량과 통행시간간의 관계를 적합하게 설명한다고 볼 수가 있지만, 용량 이후의 곡선은 매우 임의적인 벌점 가중치에 해당하므로 대기행렬 상에서의 지체시간을 묘사하기가 어렵다. 이와 같은 용량 이후의 벌점 가중치 증가패턴에 의한 통행시간 증가가 통행자의 노선선택 행태에서 현실과 같이 묘사하는데 한계가 있게 됨에 따라 노선배정 분석의 오차 원인이 될 수가 있는 것이다.

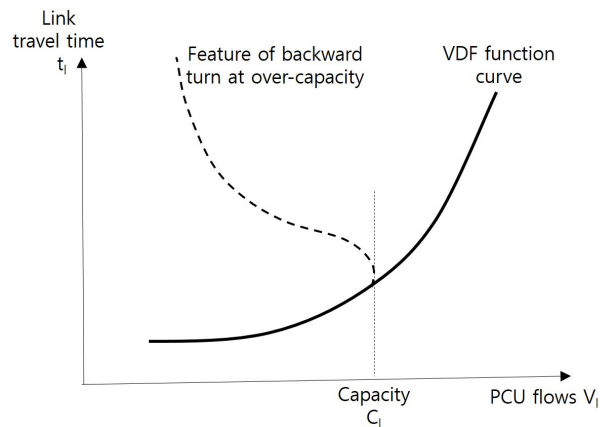


Figure 1. Conceptual graph of VDF function and real phenomena

노선선택 행태의 집합화 및 계산 수렴의 정확성

OD기반 정적 사용자균형 노선배정모형은 모든 통행자들이 동일한 원칙 하에서 노선선택을 한다는 가정이 내포되어 있다. 예로 네트워크의 정보를 많이 갖고 있는 사람이나 길을 잘 모르는 사람이나 모두 완벽한 정보를 가지고 합리적 행태결정을 한다고 가정한 것이다. 하지만 다양한 요소가 경로선택에 영향을 주고 있는 것이 현실이다. 즉 경관, 신호등, 도로 표지판, 내비게이션, 네트워크 정보, 경험적 교통정보, 통신매체 교통정보 등에 대해 통행자가 반응하는 정도가 성별, 연령, 교육수준, 성격, 소득 등에 따라 다르다. 따라서 이와 같은 완벽한 정보 하의 동일한 경로선택원칙이라는 가정도 노선배정의 오차 원인 중에 하나가 된다.

통행자 특성 및 교통수단별 특성을 고려하지 않고 집합화한 한계점을 극복하기 위해서 제약적이기는 하지만 통행자 속성별로 혹은 차량 종류별로 구분하여 분석하는 다분류 다수단 노선배정(Multi-class multi-modal traffic assignment) 기법이 적용되기도 한다. 이 기법은 수학적 특성 상 유일해(unique solution)를 갖지 못하고 다수해(multiple solution)를 갖게 된다. 하지만 이러한 다수해의 문제점을 극복하기 위해 비례적 배분(proportionality) 원칙을 가정함으로써 유일해를 갖도록 분석이 이루어지고는 있다. 현재 우리나라에서 사용되고 있는 모든 상용 프로그램은 이와 같은 원칙을 기초로 한 알고리즘으로 다분류 다수단 노선배정 프로그램을 제공하고 있다. 하지만 이와 같은 비례적 배분 원칙이 현실적 현상과 차이가 있을 수가 있다. 한 예로 트럭과 승용차로 교통수단을 분류 분석할 경우 유료도로 및 도로위계에 따라 경로선택행태에서 차이가 있을 수가 있다. 따라서 트럭의 경우를 승용차와 분류를 달리하여 다수단 노선배정을 수행하는 경우가 많은데, 트럭의 경우 통행료를 회사에서 지불을 하고 있으므로 지방도 혹은 국도보다는 대형차 주행이 편한 고속도로를 이용하는 경향이 있으므로 대안 경로 간에 비례적 배분 원칙을 적용할 경우 현실과 차이가 나는 분석 결과가 나올 수 있다.

노선선택의 집합화된 동일한 원칙을 적용하는 것은 앞에서 설명한 단조증가함수의 VDF함수를 가정한 것과 함께 노선배정 모형이 유일해를 갖기 위한 가정이라고 할 수 있다. 이와 같이 유일해를 갖는다는 것은 어떤 분석가가 분석하던 동일한 입력자료를 사용하면 동일한 분석 결과가 나온다는 것을 의미한다. 따라서 의사결정자 입장에서는 일정한 기준이 제시될 수 있어 정책판단에 혼선을 주지 않는다는 장점이 있게 된다. 이와 같은 일관성은 정확한 계산에 의한 유일해가 보장될 때 가능한 것이다. 하지만 2000년대 초반부터 발달된 컴퓨터 용량과 속도를 활용하면서 Origin-based 혹은 Path-based 알고리즘을 상용 프로그램에 적용하기 시작하면서, 기존의 Frank-Wolf 알고리즘(혹은 Convex combination method이라고도 불림)의 계산 수렴정도에 대한 문제점이 지적하기 시작하였다. 즉 Jayakrishnan et al.(1994), Kang et al.(2000), Bar-Gera(2002), Bar-Gera와 Boyce(2003), Dial(2006), Slavin et al.(2006), Kim et al.(2015) 등의 연구에서 기존 알고리즘과 새로운 알고리즘과의 수렴정도 및 계산 속도를 비교 분석한 결과를 발표하였다. 계산의 수렴정도는 일반적으로 Equation 4와 같은 사용자 균형상태의 수렴기준 측정치(Duality gap measure 혹은 Relative gap measurement)가 사용되고 있다.

$$GM = \frac{\sum_l f_l C_l(f_l) - \sum_i \sum_j OD_{ij} C_{ij}^{\min}}{\sum_i \sum_j OD_{ij} C_{ij}^{\min}} \quad (4)$$

여기서 GM : relative gap measurement

f_l : 링크 l의 교통량

$C_l(f_l)$: 교통량이 f_l 일 때 링크 l의 통행시간

OD_{ij} : 출발존이 i이고 도착존이 j인 OD 통행량

C_{ij}^{\min} : 출발존 i에서 도착존 j까지의 최단경로 통행시간

비교적 큰 네트워크에서 Frank-Wolf 알고리즘으로는 수백번의 반복계산(iteration)을 하여도 $10^{-3} \sim 10^{-4}$ 정도 수렴기준(relative gap measurement) 이하로는 잘 내려가지 않던 것이 비교적 적은 반복계산 횟수로 더 적은 값의 수렴기준으로 계산이 가능하게 되기 시작했다. 더군다나 수렴기준 측정치 값이 큰 값으로 나올 경우 링크의 교통량 해가 매우 불안정하게 나와 정책분석에 왜곡을 줄 정도로 심각할 수 있음을 실질적인 분석을 통해 보여주기 시작하였다. 즉 신규 또는 확장 도로구간과는 전혀 무관한 지역에서의 사업시행시와 사업미시행시의 교통량 변화가 없을 것으로 예상되는데도 불구하고 큰 값의 변화가 발생할 수 있음을 실제 분석에서 보여 주었다. 수렴이 잘 안됨에 따라 이와 같이 사업지와 무관한 지역에서 사업시행시 다른 교통량이 나올 경우에는 편익 산출에도 왜곡된 값을 제시하게 되어 정책결정에 오류의 원인이 될 수가 있다. 이와 같은 문제점을 인식하여 우리나라에서는 직접영향권과 간접영향권을 설정하여 편익산출에 적용되는 지역과 링크를 설정하도록 지침에서 제시하고는 있지만 이것 또한 완벽한 문제해결 방법이 못된다. 새로운 알고리즘으로 적용하여 $10^{-7} \sim 10^{-8}$ 정도 이하로 수렴정도를 낮추면 사업지와 무관한 지역에서 사업시행시 교통량 변화가 생기는 현상이 없어지는 것을 여러 연구에서 제시되고 있다. 이것은 사용자 균형상태로 수렴 정도를 매우 낮게 한다면 영향권 설정에 상관없이 편익 산출은 동일하게 됨을 의미하는 것이다. 이와같이 노선배정분석 수렴 정도에 따라 예측 교통량에 차이가 있다는 점도 또 다른 오차 원인이 된다.

맺는말

본 연구에서는 교통수요예측 상의 전반적인 오차 원인을 간략하게 정리하고, 특히 노선배정분석 상에서 링크 교통량 예측오차의 원인이 되는 요소들을 정리하였다. 즉 노선배정 분석의 입력 자료의 오차, 공간 집합화와 네트워크 표현방식에 따른 오차, 교통패턴 변동에 대한 대푯값 설정에 따른 오차, 교통류모형의 단순화에 따른 오차, 노선선

택 행태 집합화에 따른 오차로 구분하여 각 오차의 요인별로 구체적으로 설명을 하였다. 이러한 오차의 원인 중에는 전문적이고 정교한 분석을 통해 오차의 크기를 줄일 수 있는 요인도 있겠지만, 자료 및 수학적 모형의 한계로 불가피한 요인도 있다. 하지만, 이와 같은 예측오차가 존재한다고 해서 교통수요 분석 그리고 특히 이 연구의 주제인 네트워크 분석 자체가 의미 없는 것은 아니다. 가장 현실적 교통패턴을 정확하게 묘사하는 것은 대규모 실측 교통자료에 기반한 교통패턴이겠지만, 사회기반 교통시설의 신설, 확장 및 폐쇄 그리고 토지이용의 개발과 변화에 따른 교통패턴에 대한 예측은 실현되어 있지 않은 교통 상황에 대한 예측이므로 조사될 수가 없는 현상이어서 수학적 모형을 통한 시뮬레이션 분석에 의해서만 변화 패턴이 예측 가능한 것이다. 따라서 토지이용 및 교통 시스템의 변화에 따른 장래 교통패턴 변화를 예측할 때에는 여러 가지 오차의 요인을 고려하고 또한 오차 존재 가능성을 인정하면서 네트워크 분석의 결과를 현명하게 해석하고 정책에 활용할 수 있는 전문가적 경험과 통찰력이 중요한 이유가 되는 것이다. 네트워크 분석결과에 있어서도 링크 교통량의 절대 값 그 자체보다는 시스템변화에 따른 영향권 도로망의 교통량 변화패턴과 상대적 변화 크기를 분석하고 논리적 검증을 수행하는 것이 더욱 중요한 이유이기도 한 것이다. 현재 분석기법은 다양한 환경과 다양한 통행자 특성을 단순화하여 하나의 원칙으로 설명할 수 있는 수학적 모형으로 평균적 교통패턴을 예측하는 방법이라고 하다면, 장래에는 휴대폰 및 교통카드 자료와 같이 시시각각 모든 장소에서 자동적으로 수집되는 위치정보 및 개인정보를 활용하여 과거의 유사한 환경과 통행자 유형의 실측 행태를 기반으로 개별적 맞춤형 예측 분석이 가능하게 될 수도 있다고 고려된다. 즉 수학적 완결성으로 경직된 가정과 집합적 행태를 표현하기 위한 평균적 행태예측을 기본으로 하는 수학적 예측모형보다는 현실에서 실현 관측된 통행 자료를 기반으로 예측하는 방법이 더 높은 수준의 정확성을 보장할 것이라 기대된다. 하지만 이와 같은 분석기법의 개발은 막대한 대규모 자료의 축적이 필요할 것이며 또한 사회적 이익에만 사용되는 전제 하에 개인정보보호의 장벽을 풀어가는 노력이 전제되어야 할 것이다. 장래에는 오차의 원인이 되는 요소들이 많이 해소될 수 있는 빅데이터 관측 자료의 활용과 수학적 평균적 행태모형을 융합 활용한 예측 분석기법이 개발되기를 기대하는 바이다.

REFERENCES

- Anas A. (1982), Residential Location Markets and Urban Transportation: Economic Theory, Econometrics, and Policy Analysis With Discrete Choice Models, Academic Press.
- Bar-Gera H. (2002), Origin-Based Algorithm for the Traffic Assignment Problem, 36(4), Transportation Science.
- Bar-Gera H., Boyce D. (2003), Origin-based Algorithms for Combined Travel Forecasting Models, 37, Transportation Research B.
- Ben-Akiva M., Lerman S. (1985), Discrete Choice Analysis: Theory and Application to Travel Demand, The MIT Press.
- Boyce D. E., Janson B. N., Eash R. W. (1981), The effect on Equilibrium Trip Assignment of Different Link Congestion Functions, 15A, Transportation Research A.
- Branston. D. (1976), Link Capacity Function: A Review, 10, Transportation Research, 223-236.
- Dial R. B. (2006), A Path-based User-equilibrium Traffic Assignment Algorithm That Obviates Path Storage and Enumeration, 40, Transportation Research B.
- Douglas A. A., Lewis R. J. (1970), Transportation Generation Techniques, November, Traffic Engineering & Control, 362-365.
- Douglas A. A., Lewis R. J. (1970), Transportation Generation Techniques: 2. Zonal Least-squares Regression Analysis, December, Traffic Engineering & Control, 428-431.
- Douglas A. A., Lewis R. J. (1971), Transportation Generation Techniques: 3. Household Least-squares Regression Analysis, January, Traffic Engineering & Control, 477-479.

- Douglas A. A., Lewis R. J. (1971), Transportation Generation Techniques: 4. Category Analysis and Summary of Trip Generation Techniques, February, Traffic Engineering & Control, 532-535.
- Easa S. (1991), Traffic Assignment in Practice: Overview and Guidelines for Users, Journal of Transportation Engineering, 117(6), 602-623.
- Hall F. L., Hurdle V. F., Banks J. H. (1992), Synthesis of Recent Work on the Nature of Speed-Flow and Flow-Occupancy (or Density) Relationships on Freeways, Transportation Research Record 1365.
- Horowitz A. J. (1993), Reference Manual: Quick Response System II for Windows, Version 3.6, AJH Associates, Milwaukee, Wisconsin.
- Jayakrishnan R., Tsai W. T., Prashker J. N., Rajadhyaksha S. (1994), Faster Path-based Algorithm for Traffic Assignment, Transportation Research Record 1443.
- Kanafani A. (1983), Transportation Demand Analysis, McGraw Hill Book.
- Kang S. M., Kwon Y.S., Park C.H. (2000), A Development of a Path-Based Traffic Assignment Algorithm Using Conjugate Gradient Method, J. Korean Soc. Transp., 18(5), Korean Society of Transportation, 99-107.
- Kim I. K. (1997), Theoretical Comparison of O-D Trips and P-A Trips in Travel Demand Analysis, J. Korean Soc. Transp., 15(1), Korean Society of Transportation, 45-62.
- Kim I. K. (2006), Reconsideration on Calibration and Forecasting Procedure for Transportation Demand Analysis, Transportation Technology and Policy, 3(1), Korean Society of Transportation, 87-106.
- Kim J. Y., Kim J. Y., Park S. J., Lee S. J. (2015), Analysis of Convergence Level and Exit Criteria on Traffic Assignment Algorithms, 14(1), The Journal of The Korea Institute of Intelligent Transport Systems.
- Manheim M. L. (1979), Fundamentals of Transportation Systems Analysis, 1(Basic Concepts), The MIT Press.
- Ortuzar J. D., Willumsen L. G. (2011), Modelling Transport, 4th Edition, John Wiley & Sons Ltd.
- Regueros A., Prashker J., Mahalel D. (1993), Equilibrium Assignment Method for Pointwise Flow-Delay Relationship, Transportation Research Record 1413.
- Sheffi Y. (1985), Urban Transportation Networks: Equilibrium Analysis With Mathematical Programming Methods, Prentice-Hall Inc.
- Slavin H., Brandon J., Rabinowicz A. (2006), An Empirical Comparison of Alternative User Equilibrium Traffic Assignment Methods, Association for European Transport and Contributors.
- Spiess H., Florian M. (1989), Optimal Strategies: A New Assignment Model for Transit Network, 23B(2), Transportation Research B, 83-102.
- Stoper P. R., Meyburg A. H. (1975), Urban Transportation Modeling and Planning, Lexington Books.
- Suh S. D. (1990), Study on Validation of Highway Capacity Function in South Korea, Transportation Studies, The Korea Transport Institute, Sejong, South Korea.
- Suh S. D. (1991), Highway Capacity Function and Congestion Cost, Transportation Information, 62, The Korea Transport Institute, Sejong, South Korea.
- US DOT (1983), UTPS Highway Network Development Guide, U.S. Department of Transportation, January.
- Washington S. P., Karlaftis M.G., Mannering F. L. (2003), Statistical and Econometric Methods for Transportation Data Analysis, A CRC Press Company, 121-141, 155-167.