

화물자동차 투어유형 선택행태 분석

An Analysis of Choice Behavior on Tour Type of Commercial Vehicle

김 한 수

(서울시립대, 교통공학과, 박사수료)

김 찬 성

(한국교통연구원, 국가교통DB센터, 연구위원)

박 동 주

(서울시립대, 교통공학과, 교수)

김 경 수

(서울시립대, 교통공학과, 석사과정)

Key Words : tour type, tour-based approach, decision tree, logit model, logistic regression

목 차

- I. 서론
- II. 화물자동차 투어유형
 - 1. 통행과 투어의 정의
 - 2. 투어유형 기존연구 고찰
- III. 연구방법
 - 1. 연구자료
 - 2. 분석모형
 - 3. 모형비교방법
- 4. 변수선정
- IV. 연구결과
 - 1. 의사결정나무 분석결과
 - 2. 로짓모형 분석결과
 - 3. 분석결과 비교
- V. 결론
- 참고문헌

I. 서론

최근 화물교통 부문에서는 화물수요모형에 대한 패러다임 변화가 나타나고 있다(Tavasszy, 2008; Liedtke et al., 2009; 김한수 외, 2009). 여객수요모형 기반의 전통적 4단계 접근방법에서 출발한 화물수요모형이 갖는 한계를 지적하고, 이를 극복하기 위한 연구들이 진행 중이다. 이러한 변화 중 화물자동차가 다수의 도착지를 방문하는 이동행태인 화물자동차 투어를 반영하기 위한 시도가 주목 받고 있다. 전통적 4단계 접근방법에서 화물자동차 이동은 화물이동과 방향이 같으며, 화물적재물 만큼 화물을 운송하는 것으로 단순히 가정하고 있다. 그러나 화물자동차 이동을 결정하는 의사결정주체(운송인, 운전자)는 화물이동을 결정하는 의사결정주체(화주, 소비자)와 다르고, 그들 간의 이동에 차이가 있기 때문에 화물자동차의 이동을 단순히 화물이동과 같다고 가정하기 어렵다는 것이다. 특히 화물자동차는 다수의 도착지를 방문하는 이동행태를 보이고 있기 때문에, 이를 모형화 하기 위한 방법으로 투어기반 접근방법이 제시되었다(Hunt and Stefan, 2007).

화물자동차 이동을 투어기반 접근방법으로 모형화 하기 위해서는 화물자동차 투어와 투어유형에 대한 이해가 필요하다. 본 연구는 화물자동차 투어유형을 왕복형 투어와 체인형 투어로 구분하여 이들 투어유형 선택행태를 분석하고자 한다. 투어유형 선택행태를 분석하기 위한 방법으로 중속변수가 범주형 자료일 때 적합한 의사결정나무(decision tree)와 로짓모

형(logit model)을 이용하였다.

본 연구는 2장에서 화물자동차 투어유형에 대한 기존연구를 고찰하고, 3장에서 연구자료에 대한 설명과 함께 분석모형, 모형비교방법, 변수선정에 대해 기술한다. 4장에서는 의사결정나무와 로짓모형을 이용한 투어유형 선택행태 분석결과를 제시하고, 두 모형의 분석결과를 비교한다. 마지막으로 연구의 시사점을 제시하고, 연구한계와 향후 연구과제에 대해 논한다.

II. 화물자동차 투어유형

1. 통행과 투어의 정의

통행(trip), 통행체인(trip chain), 투어(tour)에 대한 개념 정의는 중요한데, 통행은 특정 목적을 위해 출발지에서 하나의 도착지까지의 개별적인 화물자동차의 이동을 의미한다. 투어에 대한 정의는 연구자에 따라 차이가 있다. Holguín-Veras는 통행체인과 투어는 동일 개념으로 화물자동차가 다수의 도착지를 순차적으로 이동하여 만들어진 물리적 이동의 순서이며, 화물자동차의 전체 여정(entire journey)인 투어는 여러 개의 통행으로 구성된다고 하였다(Holguín-Veras and Patil, 2005; Wang and Holguín-Veras, 2008). 또한 Figliozzi(2007)는 화물자동차가 창고(depot) 또는 물류센터를 출발하여 2개 이상의 다른 도착지를 순차적으로 방문한 뒤 창고 또는 물류

센터로 복귀하는 여정으로 정의한다고 하였다. 본 연구에서는 화물자동차 투어를 “출발지에서 하나 이상의 도착지를 방문하고 출발지로 복귀하는 통행으로 구성된 여정”으로 정의한다.

2. 투어유형 기존연구 고찰

화물자동차 투어유형을 구분한 기존연구는 대부분 화물수요모형에 투어특성을 반영하기 위함을 목적으로 하고 있다. 투어유형을 구분하는 기준은 투어횟수, 도착지수, TL(truckload)/LTL(less than truckload), 기타 특성(거리제약, 시간제약, 정기/비정기)을 적용하고 있다(Holguín-Veras and Patil, 2005; Maruyama and Harata, 2005; Figliozzi, 2007; Raothanachonkun et al., 2007).

Raothanachonkun et al.(2007)은 본 연구에서 주목하는 바와 같이 투어유형에 따라 화물자동차 이동에 차이가 발생하는 것을 인식하고, 이를 반영하여 화물자동차 O/D를 추정하는 방법을 제시하였다. 그러나 화물자동차 투어유형이 어떤 메커니즘에 의해 선택되는지에 대한 설명이 없다. 본 연구는 화물자동차 이동을 모형화 하는데 중요한 투어유형이 어떤 메커니즘에 의해 결정되는지 설명하는 연구로서 의의를 갖는다.

본 연구에서는 화물자동차 투어유형을 <그림 1>과 같이 왕복형 투어와 체인형 투어 두 가지로 구분하였다. 이와 같이 구분한 이유는 화물자동차 투어에는 다음과 같은 두 가지 특성이 존재하기 때문이다. 첫째, 화물 이동과 화물자동차 이동의 관계이다. 왕복형 투어의 경우 화물 이동과 화물자동차 이동이 같다. 반면, 체인형 투어는 첫 번째 통행만 같을 뿐 나머지 통행은 다르다. 둘째, 화물자동차 공차통행의 존재와 이동방향의 차이이다. 공차통행은 투어목적에 따라 첫 번째 통

행 또는 마지막 통행에서 발생할 수 있으며, 보통 한 투어는 모두 적재통행 이거나 하나 이상의 공차통행으로 구성된다. <그림 1>과 같이 동일한 화물 이동이라 하더라도 화물자동차 투어 유형에 따라 음영부분과 같이 화물자동차 이동에는 차이가 있다. 이러한 차이는 체인형 투어가 많아질수록 증가하게 된다.

III. 연구방법

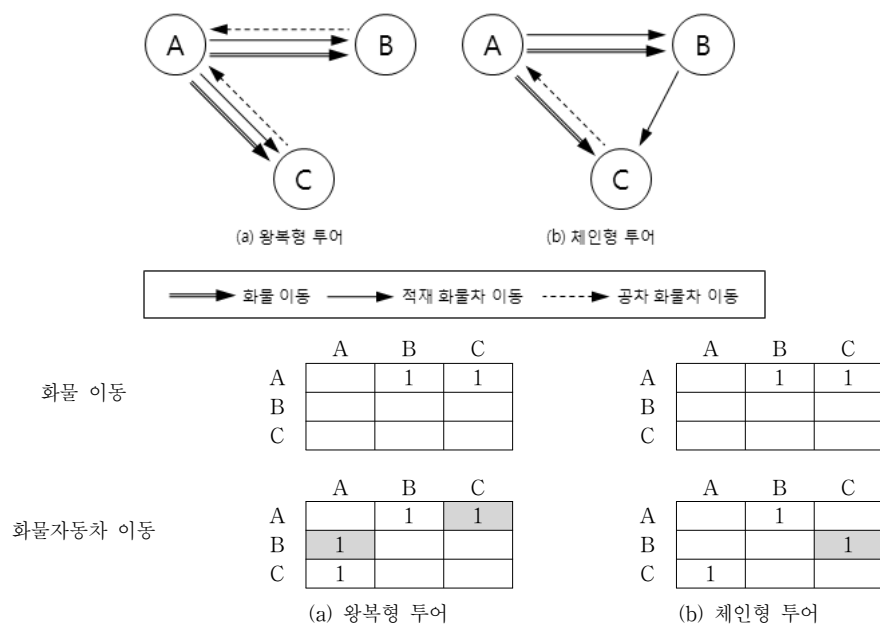
1. 연구자료

본 연구는 2005년 한국교통연구원에서 국가교통DB구축을 위해 수행한 화물자동차 통행실태 조사자료를 이용하였다(한국교통연구원, 2006). 화물자동차 통행실태조사는 화물자동차 운전자로부터 한 차량이 하루 동안 통행한 내용의 다이어리 형태로 되어 있으며, 주요 설문항목은 아래와 같다.

- 화물자동차 업종 : 사업용(용달, 개별, 일반), 비사업용(자가용, 관용)
- 출발지, 출발지 유형, 출발시간
- 도착지, 도착지 유형, 도착시간
- 적재톤수, 화물품목, 적재중량, 통행거리

조사자료 중 다음과 같은 자료는 이상치로 판단하였다.

- 출발지와 도착지가 시도만 기재되어 있고, 시군구는 기재되지 않은 통행이 포함된 자료
- 조사된 거리, 시간을 이용하여 산출한 속도가 5 km/h 미만, 120 km/h 초과인 통행이 포함된 자료
- 단위통행거리가 500km 이상 통행이 포함된 자료



<그림 1> 화물과 화물자동차 이동 : (a) 왕복형 투어, (b) 체인형 투어

- 적재능력이 0톤 또는 30톤을 초과한 자료
- 적재중량이 30톤을 초과한 통행 또는 적재능력의 2배를 초과한 통행이 포함된 자료

본 연구의 공간적 범위는 수도권(79개)¹⁾을 대상으로 한다. 화물자동차 차종에 따른 투여유형 선택행태를 분석하기 위해 적재톤수에 따라 구분²⁾하였으며, 이들 자료를 투여유형과 도착지 수로 구분하면 <표 1>과 같다. 대형 화물자동차는 자료가 적어 제외하였다. 투여 당 평균 도착지 수는 2.22로 도착지 3개 이상인 체인형 투여가 상대적으로 적었다. 도시지역에 대한 기존연구에서 Calgary는 대략 6(Hunt and Stefan, 2007), Denver는 5.6(Holguín-Veras and Patil, 2005), Amsterdam는 6.2(Vluge and Janic, 2004)로 투여 당 평균 5개 이상의 도착지 수를 갖는 것으로 보고되고 있다. 이와 같이 도착지 수가 차이나는 이유는 교통존의 크기 때문이다. 본 연구에서는 시군구 단위의 교통존으로 조사된 자료를 이용하였으나, 기존 연구에서는 보다 작은 행정구역 단위의 교통존으로 조사된 자료를 이용하여 보다 많은 도착지 수를 갖는 것으로 나타났다.

<표 1> 투여유형별 투여건수

구분		소형 화물자동차 투여건수(비율)	중형 화물자동차 투여건수(비율)
왕복형 투여	도착지 2개	1,128 (82.3%)	444 (81.6%)
	도착지 3개	177 (12.9%)	77 (14.2%)
체인형 투여	도착지 4개	65 (4.7%)	23 (4.2%)
	합계	1,370	544

2. 분석모형

본 연구는 투여유형 선택행태 분석이 목적이므로 종속변수가 왕복형 투여와 체인형 투여로 구분된 선택행태를 분석해야 한다. 종속변수가 범주형 자료인 경우 로짓모형, 프로빗모형과 같은 로지스틱 회귀모형이 많이 이용되어 왔다. 로지스틱 회귀모형 대신 대안적인 방법으로 패턴인식문제(pattern recognition problem)로 다루는 방법이 있으며, 패턴인식에 사용하는 방법은 의사결정나무(decision tree), 신경망(neural networks), 베이저안 네트워크(bayesian network) 등을 이용하고 있다(Wets et al., 2000; Xie et al., 2003). 이 중 의사결정나무는 의사결정규칙(decision rule)을 도표화하여 관심대상이 되는 집단을 분류하거나 예측하는 방법으로 유용하다. 통행행태분석에 의사결정나무를 적용한 사례는 대부분 의사결정나무와 다른 모형(로짓모형, 베이저안 네트워크, 신경망)을

동시에 분석하여 비교결과를 제시하고 있다. 비교결과 로짓모형 보다는 패턴인식방법이 정확도 측면에서 보다 우수한 결과를 제시하고 있다. 최근 Janssens et al.(2006), Arentze and Timmermans(2007)는 의사결정나무의 한계를 보완하여 새로운 모형을 제안하기도 하는 만큼 의사결정문제에 있어 효과적인 방법으로 인식되고 있다.

1) 의사결정나무

의사결정나무는 의사결정규칙을 도표화하여 관심대상이 되는 집단을 몇 개의 소집단으로 분류(classification)하거나 예측(prediction)을 수행하는 분석방법이다. 의사결정나무는 뿌리마디(root node)로부터 시작하여 각 가지(branch)가 끝마디(terminal node)에 이를 때까지 자식마디(child node)를 계속적으로 형성해 나감으로써 완성된다. 의사결정나무를 형성하는 알고리즘은 CHAID(Chi-square Automatic Interaction Detection; Kass, 1980), CART(Classification and Regression Trees; Breiman et al., 1984), QUEST(Quick, Unbiased, Efficient, Statistical Tree; Loh and Shih, 1997) C4.5(Quinlan, 1993) 등을 이용하여 나무구조를 형성하는데, 이들은 분리기준(splitting criterion), 정지규칙(stopping rule) 및 가지치기(pruning) 등에서 서로 다른 형성과정을 가지고 있다(최종후 외, 1998).

2) 로짓모형

로짓모형은 종속변수가 범주형 자료의 응답변수(response variable)를 갖는 경우 사용된다. 로짓모형의 독립변수(independant variable)는 두 가지 유형으로 구분할 수 있다. 첫째, 응답변수가 변화함에 따라 변화하지 않고 오직 개개의 관찰치에 의해서만 변화하는 변수이며, '응답범주에 따라 변화하지 않는 변수(invariant variables to response categories)'라고 한다. 둘째, 응답범주가 다르면 동일한 개인에 대해서도 값이 달라지는 변수이며, '선택한정변수(choice specific variables)'라고 한다. 전자로 구성된 모형을 다항로짓모형(multinomial logit model)이라 하고, 후자는 다항로짓모형을 보다 일반화 시킨 것으로 조건부 로짓모형(conditional logit model)이라 한다. 연구자에 따라 이를 구분하기도 하며, 함께 묶어 다항로짓모형으로 서술하기도 한다(이성우 외, 2005). 본 연구에서는 이항로짓모형을 이용하는데, 투여유형 y를 선택할 확률(Prob(y=1))과 선택하지 않을 확률(Prob(y=0))은 아래 식 (1)에 의해 산출한다.

$$\text{Prob}(y = 1) = \frac{\exp\left(\sum_{k=1}^K \beta_k x_k\right)}{1 + \exp\left(\sum_{k=1}^K \beta_k x_k\right)}$$

- 1) 수도권은 서울(25개), 인천(10개), 경기(44개)의 행정구역으로 구성된 대도시권임.
- 2) 화물자동차 적재톤수에 따라 소형(2.5톤 미만), 중형(2.5톤 이상~8.5톤 이하), 대형(8.5톤 초과)으로 구분함.

$$\text{Prob}(y=0) = \frac{1}{1 + \exp(\sum_{k=1}^K \beta_k x_k)} \quad (1)$$

3. 모형비교방법

모형비교를 위해 가장 많이 사용하는 척도는 정확도 (accuracy)이다. 정확도는 모형 예측값이 실제값과 얼마나 동일한지를 측정하는 것이다. 정확도 외의 기준으로 민감도 (sensitivity)는 연구자가 관심을 가진 데이터를 가장 잘 분류하는 능력을 나타내는 척도이며, 특이도 (specificity)와 정밀도 (precision)은 분류능력을 나타내는 척도이다 (Polat and Güneş, 2009). 각 척도는 식 (2)-(5)에 의해 산출한다.

<표 2> 투어유형 분류표의 구조

구분		예측	
		왕복형 투어	체인형 투어
실제	왕복형 투어	참 : TP	거짓 : FN
	체인형 투어	거짓 : FP	참 : TN

주 : TP(true positive) - 왕복형 투어를 잘 맞춘 경우,
 FP(false positive) - 왕복형 투어를 체인형 투어로 예측,
 TN(true negative) - 체인형 투어를 왕복형 투어로 예측,
 FN(false negative) - 체인형 투어를 잘 맞춘 경우

$$\text{정확도}(accuracy) = \frac{TP + TN}{TP + TN + FP + FN} \quad (2)$$

$$\text{민감도}(sensitivity) = \frac{TP}{TP + FN} \quad (3)$$

<표 3> 종속변수와 독립변수

구분	변수	정의	값	
종속변수	TourType	투어유형	0 : 왕복형 투어, 1 : 체인형 투어	
독립변수	화물 특성	ItemType	화물품목 유형	0 : 소비재, 1 : 중간재
		SumTon	총화물량	연속형 자료 (단위 : 0.1톤)
		AvgTon	평균화물량	총화물량/(도착지수-1)
	화물 자동차 특성	Clload	적재톤수	연속형 자료 (단위 : 0.1톤)
		ClloadType	적재톤수 유형	0 : 소형, 1 : 중형
		Cbiz	업종	0 : 사업용, 1 : 비사업용
		LoadRate	화물적재율	총화물량/적재톤수
	출발지 특성	DType	출발지 유형	0 : 차고지, 1 : 공장, 2 : 도소매업체, 창고
		DIsestFlag	출발지 산업단지 여부	0 : 아님, 1 : 산업단지
		DTemp	출발지 전체종사자수	연속형 자료 (단위 : 인)
		DMemp	출발지 제조업종사자수	연속형 자료 (단위 : 인)
	첫번째 도착지 특성	AType	도착지 유형	0 : 차고지, 1 : 공장, 2 : 도소매업체, 창고
		AIsestFlag	도착지 산업단지 여부	0 : 아님, 1 : 산업단지
		ATemp	도착지 전체종사자수	연속형 자료 (단위 : 인)
		AMemp	도착지 제조업종사자수	연속형 자료 (단위 : 인)
	통행 및 투어 특성	PriDist	첫 번째 통행거리	연속형 자료 (단위 : km)
TourDist		투어통행거리	연속형 자료 (단위 : km)	
AvgDist		평균통행거리	투어거리/(도착지수-1) (단위 : km)	
PriTime		첫 번째 통행시간	연속형 자료 (단위 : 분)	
TourTime		투어통행시간	연속형 자료 (단위 : 분)	
AvgTime		평균통행시간	투어거리/(도착지수-1) (단위 : 분)	

$$\text{특이도}(specificity) = \frac{TN}{TN + FP} \quad (4)$$

$$\text{정밀도}(precision) = \frac{TP}{TP + FP} \quad (5)$$

4. 변수선정

1) 종속변수

종속변수는 투어유형이다. 본 연구에서는 투어유형을 왕복형 투어와 체인형 투어 두 가지로 구분하였다.

2) 독립변수

투어유형 선택에 영향을 미칠 것으로 예상되는 변수들로 독립변수를 구축하였다. 독립변수는 화물, 화물자동차, 출발지, 도착지, 통행 및 투어 특성을 고려하여 <표 3>과 같은 변수들을 선정하였다. 화물 특성은 화물품목 유형, 총화물량, 도착지 평균화물량으로 구성되어 있다. 화물품목 유형인 소비재와 중간재는 조사항목에 없어 화물품목, 출발지 유형, 도착지 유형으로 추정하였다. 화물자동차 특성은 적재톤수, 적재톤수 유형, 업종, 화물적재율로 구성되어 있다. 적재톤수 유형은 소형-2.5톤 미만, 중형-2.5톤 이상~8.5톤 이하로 구분하였으며, 대형은 자료부족으로 제외하였다. 화물적재율은 총화물량을 적재톤수로 나누어 추정하였다. 출발지와 도착지 특성은 시설유형, 사회경제지표들로 구성되어 있으며, 통행 및 투어 특성의 통행거리와 통행시간에 대해 첫 번째 통행, 전체 투어, 평균으로 독립변수를 구성하였다.

IV. 연구결과

본 연구에서는 체인형 투어와 왕복형 투어를 선택하는 행태를 분석하기 위해 소형 화물자동차(1,370건)와 중형 화물자동차(544건)의 투어자료를 훈련데이터(training data set)와 검증데이터(validation data set)로 구분하였다. 구분비율은 67 : 33으로 하였으며, 투어유형을 기준으로 층화 단순무작위 추출을 적용하였다. 분석툴은 SAS 9.1.3의 Enterprise Miner 4.3을 이용하였다.

1. 의사결정나무 분석결과

본 연구의 의사결정나무 분리기준은 엔트로피 감소(entropy reduction)를 이용한 C4.5 알고리즘을 적용하였다. 의사결정나무 분석결과 <그림 2, 3>과 같이 6개 끝마디를 가진 나무가 생성되었다. 소형 화물자동차의 끝마디는 노드 5, 6, 8, 9, 12, 13이며, 중형 화물자동차의 끝마디는 노드 5, 6, 8, 9, 10, 11이다. 각 노드의 왼쪽은 훈련데이터, 오른쪽은 훈련데이터로부터 생성한 의사결정나무 규칙에 검증데이터를 적용한 결과를 의미한다. 또한 각 노드의 음영이 짙을수록 왕복형 투어, 백색인 경우 체인형 투어로 분류된다. 의사결정나무 각 노드의 신뢰값(confidence value)은 식 (6)에 의해 계산된다. 여기서, n은 각 노드의 관측치 수이며, m은 오분류 수이다. 예를 들면, 소형 화물자동차의 노드 6은 화물적재율(LoadRate) 1.05 이상, 평균화물량(AvgTon) 1.15톤 미만으로 체인형 투어로 분류되는 노드이며, 훈련데이터 56개 중 4%(2개)가 오분류 되어 이 노드의 신뢰값은 0.96이다.

$$CV = 1 - m/n \quad (6)$$

소형 화물자동차의 투어유형을 결정하는 설명변수는 화물 적재율, 평균화물량, 총화물량으로 나타났다. 노드의 확률값 50%를 기준으로 왕복형 투어와 체인형 투어로 구분하였을 때, 왕복형 투어 집단은 노드 5, 8, 12이며, 체인형 투어 집단은 노드 6, 9, 13이다. 추정된 나무는 다음과 같은 해석이 가능하다.

- 1) 화물적재율이 적재능력(1.05) 미만인 경우 왕복형 투어 확률이 높다(노드 5, 8). 다만, 이 중 평균화물량이 0.85톤 미만, 총화물량이 0.85톤 이상인 경우는 체인형 투어 확률이 높다(노드 9).
- 2) 화물적재율이 적재능력(1.05) 이상인 경우 체인형 투어 확률이 높다(노드 6, 13). 다만, 이 중 평균화물량이 1.15톤 이상, 총화물량이 2.25톤 미만인 경우는 왕복형 투어 확률이 높다(노드 12).

중형 화물자동차의 투어유형을 결정하는 설명변수는 평균 화물량, 총화물량으로 나타났다. 평균화물량은 하위구조에서

다시 분류기준으로 사용되기 매우 중요한 변수로 인식되었다. 왕복형 투어 집단은 노드 6, 8, 9, 11이며, 체인형 투어 집단은 노드 5, 10이다. 추정된 나무는 다음과 같은 해석이 가능하다.

- 1) 평균화물량이 1.85톤 이상이면 왕복형 투어 확률이 높다(노드 6, 11). 다만, 이 중 총화물량이 3.75톤 이상이고, 평균화물량이 3.9톤 이하이면 체인형 투어 확률이 높다(노드 10).
- 2) 평균화물량이 1.85톤 미만이고 총화물량이 1.9톤 이상이면 체인형 투어 확률이 높다(노드 5).
- 3) 평균화물량이 1.85톤 미만이고 총화물량이 1.9톤 미만이면 왕복형 투어 확률이 높다(노드 9).

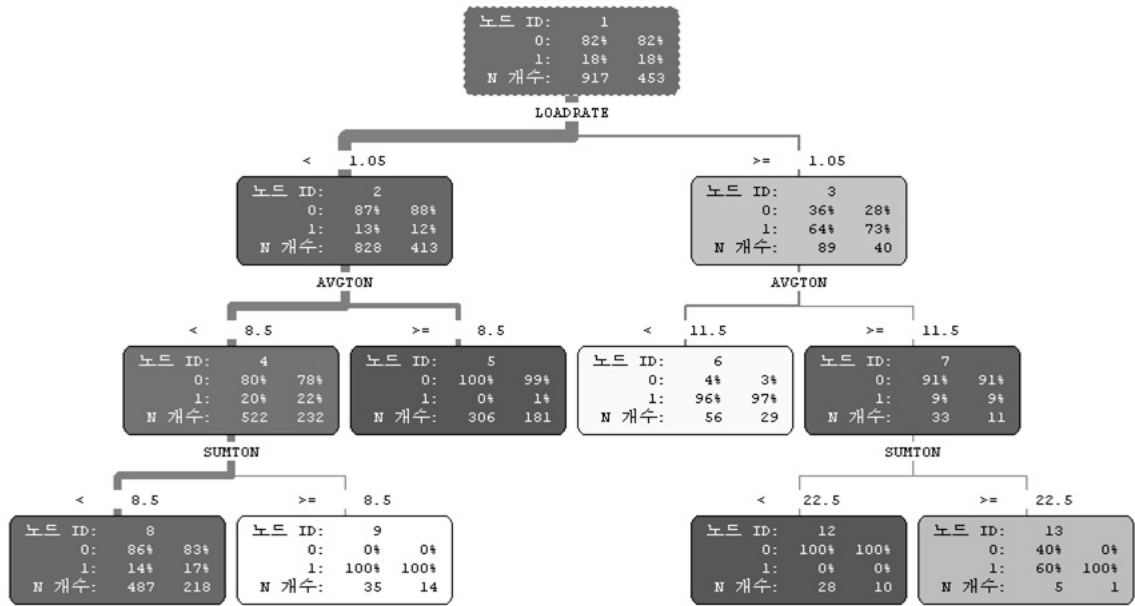
투어유형을 설명하는 독립변수로 화물특성 뿐만 아니라 화물자동차, 출발지, 도착지, 통행 및 투어 특성을 입력하였는데도 불구하고 화물적재율, 평균화물량, 총화물량이 투어유형을 분류하는 설명변수로 선정되었다. 이러한 이유는 화물자동차 투어를 계획함에 있어 화물을 어떻게 적재 할 것인지가 가장 중요하다는 것이다. 앞에서 해석한 내용을 다시 정리해보자. 소형 화물자동차는 적재능력을 넘어 다소 과적하더라도 운송할 화물량이 다수이고, 적재가 가능하면 체인형 투어를 계획하는 반면, 운송할 화물의 도착지가 한 곳이거나, 다수의 도착지를 방문할 수 있도록 적재하는 것이 불가능한 경우 왕복형 투어를 계획하는 경향이 있다는 것을 알 수 있다. 중형 화물자동차는 화물적재율에는 크게 영향을 받지 않고, 한 번에 다수의 도착지를 갖는 화물을 적재하는 것이 가능한 경우 체인형 투어를 계획하는 경향이 있다. 이러한 결과를 볼 때 화물자동차 투어유형을 결정 하는데 화물 출하단위(shipment size), 출하빈도(frequency), 출하량과 함께 운행계획을 수립하는 메커니즘이 중요한 영향력을 끼친다고 할 수 있다.

2. 로짓모형 분석결과

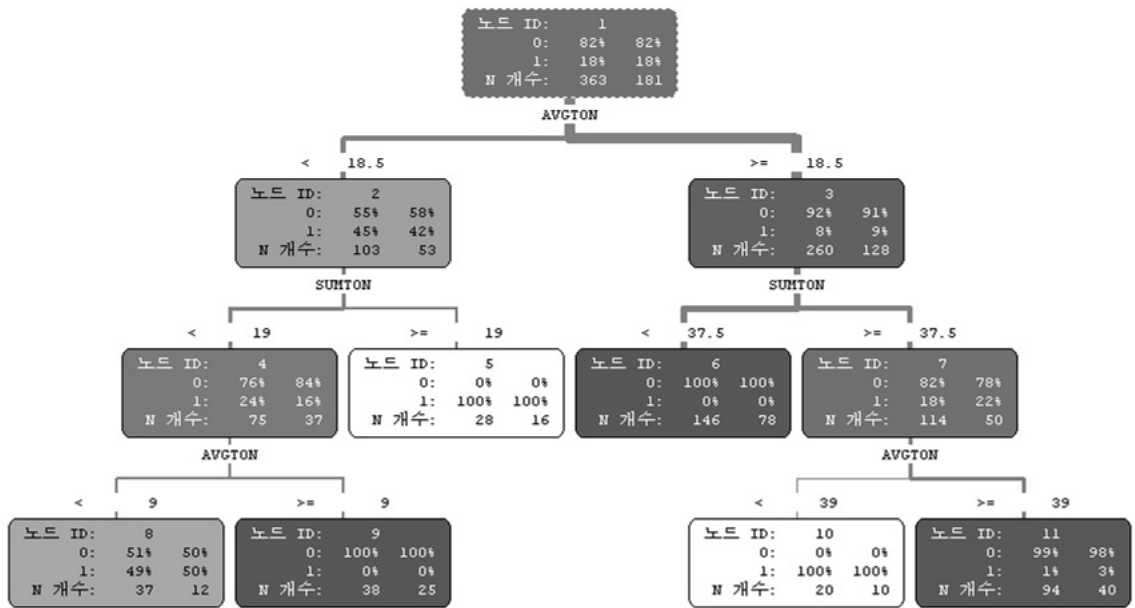
로짓모형 분석시 변수투입방식은 단계별 투입법(stepwise)을 사용하였다. 변수 투입(entry)과 유지(stay)에 적용된 유의수준은 0.05이다. 범주형 자료(화물품목 유형, 업종)는 가변수

<표 4> 로짓모형 추정결과

구분	변수	계수	t값	표준화 계수
소형 화물자동차	상수	-2.624	-14.33	-
	화물적재율	1.366	7.47	0.361
	L(c)			427.59
	L(β)			397.42
	χ ²			60.35(df=1)
중형 화물자동차	상수	5.582	4.71	-
	평균화물량	-0.088	-5.51	-0.957
	출발지 제조업종사자수	-0.520	-4.50	-0.344
	L(c)			173.61
	L(β)			132.20
	χ ²			82.82(df=2)



<그림 2> 소형 화물자동차의 투어유형선택 의사결정나무



<그림 3> 중형 화물자동차의 투어유형선택 의사결정나무

(dummy variable)를 생성하여 추정하였다. 로짓모형 분석 결과는 <표 4>와 같다. 소형과 중형 화물자동차 모두 모형의 적합도를 나타내는 우도비에 대한 χ^2 통계량 검정결과 유의수준 0.01에서 유의한 것으로 나타났다. 상수 및 개별계수 또한 0.01 유의수준에서 유의한 것으로 나타났다.

소형 화물자동차의 투어유형을 결정하는 설명변수는 화물적재율 하나로 나타났다. 화물적재율의 부호가 양(+)의 값으로 화물적재율이 증가하면, 체인형 투어 확률이 증가하는 것을 의미한다. 이러한 결과는 의사결정나무와 마찬가지로 화물적재율이 높은 경우 체인형 투어를 계획하는 것으로 해석하는 것이 가능하다. 그러나 다른 설명변수는 통계적으로 유의

하지 않았다. 중형 화물자동차의 경우 평균화물량과 출발지 제조업종사자수가 설명변수로 선정되었다. 이 두 변수의 부호가 모두 음(-)의 값을 가지므로, 평균화물량과 출발지 제조업종사자수가 증가하면, 왕복형 투어 확률이 증가하는 것을 의미한다. 표준화 계수를 보면 출발지 제조업종사자수에 비해 평균화물량이 미치는 영향이 큰 것을 알 수 있다. 이러한 결과는 의사결정나무와 마찬가지로 평균화물량이 크면 왕복형 투어를 계획하는 것으로 해석하는 것이 가능하다. 그러나 소형 화물자동차의 로짓모형 결과처럼 다른 설명변수는 통계적으로 유의하지 않았다.

3. 분석결과 비교

화물자동차 투어유형 선택에 대한 모형 간의 예측력 비교를 위한 분류표(contingency table)는 <표 5>와 같다. 이 분류표를 이용하여 의사결정나무와 로짓모형의 예측력을 비교하면 <표 6>과 같다. 정확도, 민감도, 특이도, 정밀도의 모든 지표에서 의사결정나무가 로짓모형에 비해 상대적으로 우수한 결과를 제시하고 있다. 특히, 체인형 투어가 잘 분류되는지 여부를 나타내는 특이도에서 로짓모형은 0.113, 0.364로 낮은 수치를 보이고 있다.

이러한 결과가 나타나는 이유는 로짓모형은 설명변수의 선형적 결합에 의해 선택 행위를 구분하는 방법인 반면, 의사결정나무는 규칙기반모형(rule-based model)으로 의사결정자의 운송조건에 따라 화물자동차 투어유형을 구분하기 때문이다. 즉, 화물자동차 투어유형을 분류하는 문제는 설명변수의 선형적 결합에 의한 설명 보다는 다수 설명변수들의 규칙조합으로 설명하는 것이 효과적임을 의미한다.

<표 6> 의사결정나무와 로짓모형의 예측 비교

구분		정확도	민감도	특이도	정밀도
소형 화물자동차	의사결정나무	0.956	0.997	0.763	0.951
	로짓모형	0.834	0.989	0.113	0.839
중형 화물자동차	의사결정나무	0.989	0.993	0.970	0.993
	로짓모형	0.867	0.980	0.364	0.873

V. 결론

화물자동차는 다수의 도착지를 방문하는 이동행태를 보이고 있는데, 이를 모형화 하기 위한 방법으로 투어기반 접근방

법이 화물수요를 추정하는 방법으로 연구되고 있다. 본 연구는 투어기반 접근방법의 화물수요모형을 위해 화물자동차 투어유형 선택행태를 분석하였다. 기존연구에서 이와 같은 분석 사례가 없기 때문에 다양한 설명변수를 구축하였으며, 분석방법 또한 종속변수가 범주형 자료일 때 적합한 의사결정나무와 로짓모형을 적용하여 모형간 비교를 실시하였다.

화물자동차 이동은 화물 이동에 의한 유발수요라는 측면에서 서로 관계가 있으나, 이들 이동을 결정하는 의사결정주체가 다르기 때문에 이동에 차이가 발생하고 있다. 따라서 화물자동차 투어를 왕복형 투어와 체인형 투어로 구분하여 화물자동차 투어를 계획하는 단계에서 어떤 유형의 투어를 선택하는지를 분석하는 것이 의미가 있다. 본 연구에서는 화물자동차 적재톤수를 기준으로 소형(2.5톤 미만)과 중형(2.5톤 이상~8.5톤 이하)로 구분하여 선택행태를 분석하였다. 그 결과 화물자동차 투어유형을 분류하는 설명변수로 화물적재율, 평균화물량, 총화물량이 선정되었으며, 의사결정나무와 로짓모형이 유사한 결과를 도출하였다. 또한 소형과 중형 화물자동차의 투어유형을 분류하는 설명변수가 큰 차이를 보이지 않음에 따라 화물자동차 투어를 계획함에 있어 화물을 어떻게 적재할 것인가가 가장 중요한 것으로 나타났다. 의사결정나무와 로짓모형의 예측력을 비교한 결과는 의사결정나무가 로짓모형에 비해 상대적으로 우수한 결과를 제시하였다. 특히, 체인형 투어가 잘 분류되는지 여부를 나타내는 특이도에서 로짓모형은 소형 0.113, 중형 0.364로 낮은 수치를 나타냈다. 이러한 결과는 화물자동차 투어유형을 분류함에 있어 로짓모형과 같이 설명변수의 선형적 결합에 의한 설명 보다는 의사결정나무와 같이 다수 설명변수들의 규칙조합으로 설명하는 것이 효과적임을 의미한다.

본 연구는 화물자동차 투어유형 선택행태를 분석함에 있어 시군구 단위의 교통존으로 조사된 자료를 이용하였다. 기존연구가 시군구 단위 보다는 작은 행정구역 단위의 교통존으

<표 5> 의사결정나무와 로짓모형의 분류표

구분	의사결정나무				로짓모형			
	훈련데이터		예측값		훈련데이터		예측값	
소형 화물 자동차			왕복형투어	체인형투어			왕복형투어	체인형투어
	실제값	왕복형투어	751	4	실제값	왕복형투어	741	14
		체인형투어	42	120		체인형투어	145	17
	검정데이터		예측값		검정데이터		예측값	
			왕복형투어	체인형투어			왕복형투어	체인형투어
	실제값	왕복형투어	372	1	실제값	왕복형투어	369	4
		체인형투어	19	61		체인형투어	71	9
	중형 화물 자동차	훈련데이터		예측값		훈련데이터		예측값
		왕복형투어	체인형투어			왕복형투어	체인형투어	
실제값		왕복형투어	295	1	실제값	왕복형투어	283	13
		체인형투어	2	65		체인형투어	47	20
검정데이터		예측값		검정데이터		예측값		
		왕복형투어	체인형투어			왕복형투어	체인형투어	
실제값		왕복형투어	147	1	실제값	왕복형투어	145	3
		체인형투어	1	32		체인형투어	21	12

로 조사된 자료를 이용한 것에 비하면 투어 당 평균 도착지수가 2.22로 상대적으로 낮았다. 따라서 본 연구의 결과는 시군구 단위의 교통존간 화물자동차 이동에 대해서만 의미가 있다고 할 수 있다. 시군구 단위의 교통존간 화물자동차 이동이 배송통행에 비해 본선통행이 많은 특징을 가지고 있기 때문에 읍면동 단위의 보다 작은 교통존간의 화물자동차 이동으로 확대 해석하는 것은 한계가 있다. 이와 같은 한계는 물류현황조사와 화물자동차 운행실태조사를 읍면동 단위로 조사함으로써 해결할 수 있다. 추후 추가적인 연구가 수행되어야 할 부분이라 하겠다.

참고문헌

- 김한수·박동주·김찬성·권선아(2009), “화물수요분석을 위한 화물자동차 이동의 모형화 방법론 비교 : 전통적 4단계 접근방법과 투이기반 접근방법”, 대한교통학회 제61회 학술발표회, 2009. 11. 5-6, 인천, 송도.
- 이성우·민성희·박지영·윤성도(2005), 로짓·프라빗모형 응용, 박영사, 서울.
- 최종후·한상태·강현철·김은석(1998), AnswerTree를 이용한 데이터마이닝 의사결정나무분석, 자유아카데미, 서울.
- 한국교통연구원(2006), 2005년도 「국가교통DB구축사업」 전국 지역간 화물 기종점통행량 조사.
- Arentze, T. and H. Timmermans(2007), “Parametric Action Decision Trees: Incorporating Continuous Attribute Variables into Rule-based Models of Discrete Choice”, Transportation Research Part B, Vol.41, No.7, pp.772-783.
- Breiman, L., J.H. Friedman, R.A. Olshen, and C.J. Stone(1984), Classification and Regression Trees, Wadsworth, Belmont, CA.
- Figliozzi, M.A.(2007), “Analysis of the Efficiency of Urban Commercial Vehicle Tours: Data Collection, Methodology, and Policy Implications”, Transportation Research Part B, Vol.41, No.9, pp.1014-1032.
- Holguín-Veras, J. and G.R. Patil(2005), “Observed Trip Chain Behavior of Commercial Vehicles”, Transportation Research Record, No.1906, pp.74-80.
- Hunt, J.D. and K.J. Stefan(2007), “Tour-Based Microsimulation of Urban Commercial Movements”, Transportation Research Part B, Vol.41, No.9, pp.981-1013.
- Janssens, D., G. Wets, T. Brijs, K. Vanhoof, T. Arentze, and H. Timmermans(2006), “Integrating Bayesian Networks and Decision Trees in a Sequential Rule-based Transportation Model”, European Journal of Operational Research, No.175, No.1, pp.16-34.
- Kass, G.V.(1980), “An Exploratory Technique for Investigating Large Quantities of Categorical Data”, Applied Statistics, Vol.29, No.2, pp.119-127.
- Liedtke, G.T., L.A. Tavasszy, and W. Wisetjindawat(2009), “A Comparative Analysis of Behavior-Oriented Commodity Transport Models”, CD-ROM, Transportation Research Board, National Research Council, Washington, D.C.
- Loh, W.Y. and Y.S. Shin(1997), “Split Selection Methods for Classification Trees”, Statistica Sinica, Vol.7, No.4, pp.815-840.
- Maruyama, T. and N. Harata(2005), “Incorporating Trip-chaining Behavior into Network Equilibrium Analysis”, Transportation Research Record, No.1921, pp.11-18.
- Polat, K. and S. Güneş(2009), “A Novel Hybrid Intelligent Method Based on C4.5 Decision Tree Classifier and One-against-all Approach for Multi-class Classification Problems”, Expert Systems with Applications, Vol.36, No.2, pp.1587-1592.
- Quinlan, J.R.(1993), C4.5 Program for Machine Learning, Morgan Kaufmann Publishers, San Mateo, CA.
- Raothanachonkun, P., K. Sano, W. Wisetjindawat, and S. Matsumoto(2007), “Estimating Truck Trip Origin-destination with Commodity-based and Empty Trip Models”, Transportation Research Record, No.2008, pp.43-50.
- Tavasszy, L.(2008), “Freight Demand Modeling Tools for Public-Sector Decision Making”, Conference Proceedings, edited by Hancock, K.L., Transportation Research Board, Washington, D.C., 25-27 September 2006, pp.47-55.
- Vluegel, J., and M. Janic(2004), “Route Choice and the Impact of Logistic Routes”, In: Taniguchi, E., Thompson, R. (Eds.), Logistics Systems for Sustainable Cities, Elsevier.
- Wang, Q. and J. Holguín-Veras(2008), “Investigation of Attributes Determining Trip Chaining Behavior in Hybrid Microsimulation Urban Freight Models”, Transportation Research Record, No.2066, pp.1-8.
- Wets, G., K. Vanhoof, T. Arentze, and H. Timmermans(2000), “Identifying Decision Structures Underlying Activity Patterns: An Exploration of Data Mining Algorithms”, Transportation Research Record, No.1718, pp.1-9.
- Xie, C. J. Lu, and E. Parkany(2003), “Work Travel Mode Choice Modeling with Data Mining: Decision Trees and Neural Networks”, Transportation Research Record, No.1854, pp.50-61.