

자기조직형 신경망 이론을 이용한 국도 통행시간 추정 알고리즘

Development of Travel Time Estimation Algorithm for National Highway by using Self-Organizing Neural Networks

도명식* · 배현숙**

Do, Myungsik · Bae, Hyunesook

Abstract

The aim of this study is to develop travel time estimation model by using Self-Organized Neural network(in brief, SON) algorithm. Travel time data based on vehicles equipped with GPS and number-plate matching collected from National road number 3 (between Jangji-IC and Gonjjam-IC), which is pilot section of National Highway Traffic Management System were employed. We found that the accuracies of travel time are related to location of detector, the length of road section and land-use properties. In this paper, we try to develop travel time estimation using SON to remedy defects of existing neural network method, which could not additional learning and efficient structure modification. Furthermore, we knew that the estimation accuracy of travel time is superior to optimum located detectors than based on existing located detectors. We can expect the results of this study will make use of location allocation of detectors in highway.

Keywords : travel time estimation, self-organization neural network, detectors, outlier filtering

요 지

본 연구의 목적은 수도권 남부 국도 ITS 시범구간인 국도 3호선의 장지IC~곤지암IC구간에서 수집되는 교통자료를 기반으로 자기조직형 신경망 이론을 도입하여 국도구간의 통행시간 추정모형을 개발하는 방안을 제시하는 것이다. 지점 검지기 적정 설치위치와 구간의 연장 및 연도의 토지이용특성이 단속류의 구간통행시간에 영향을 미침을 확인하였으며, 구간 통행시간 추정을 위해 기존의 인공신경망 모형이 가지는 추가학습이 불가능하다는 단점과 신경망 구조의 최적구성이 어려운 점 등을 고려하여 자기조직형 인공신경망 구조방법을 도입하였다. 통행시간 추정결과 기존 검지기에서 수집된 자료와 최적위치에서 수집된 자료를 이용하여 모형을 검증한 결과 통행특성을 가장 잘 반영하는 지점자료를 활용한 모형의 추정력이 우수한 것으로 나타났다. 이러한 시도는 향후 국도 ITS 사업의 설계에서 검지기의 설치 위치 선정에 응용할 수 있을 것으로 기대된다.

핵심용어 : 통행시간 추정, 자기조직형 뉴럴네트워크, 검지기, 이상치 제거

1. 서 론

도시화와 자동차화(motorization)가 보편화되면서 이동에 따른 시간가치에 많은 관심을 가지고 있는 추세이며 이동 및 통행에 소요되는 시간을 최소화하기 위해 많은 노력들이 이루어지고 있다. 이를 위해 최근 활발히 진행되고 있는 것이 지능형교통체계(ITS; Intelligent Transportation System)이다. ITS 사업은 교통정보를 이용자에게 제공해 줌으로써 기존 교통시설의 이용을 극대화 하고, 통행정보를 제공함으로써 기중점간을 통행하는 차량의 통행시간을 최소화하는데 그 목적이 있다.

한편, 국내의 수도권 국도 ITS 사업의 경우 영상검지기에 의해 수집되는 지점교통정보를 기반으로 통행시간을 추정하여 이를 운전자에게 제공할 예정이나, 사업 초기의 교통특성

검지를 위한 검지기의 위치가 시간이 경과함에 따라 토지이용, 기하구조 및 교통수요 등의 변화로 인하여 동질구간 설정의 재조정이 요구되며, 동질구간에서의 검지기 자료의 누락 및 이상치(outlier)에 대한 보정에 대한 연구도 통행시간 추정을 위해 선결되어야 할 과제이다.

통행시간 정보를 제공하기 위해서는 무엇보다도 현장에서 자료 수집이 우선되어야 하며, 수집된 자료는 보정과 가공의 단계를 거쳐 자료의 제공이 이루어지게 된다. 제공되는 자료는 통행시간 및 최단경로, 대체경로, 통행량, 돌발상황, 기상정보 등이며 최근에는 주변의 시설물과 관광정보까지 포함하고 있다. 이러한 정보제공에 있어서 가장 중요한 것은 수집자료의 신뢰성이며, 이와 같은 신뢰성을 높이기 위한 연구가 지속적으로 진행되고 있으나, ITS 검지기의 적정 설치 위치에 대한 공학적 연구는 거의 이루어지지 않고 있다.

*정회원 · 교신저자 · 국립한밭대학교 토목환경도시공학부 부교수 (E-mail : msdo@hanbat.ac.kr)

**한국산업관계연구원 선임연구원 (E-mail : bhslemon@hanmail.net)

즉, 국내 ITS 사업의 경우 검지기가 지리적인 여건과 전기 공사의 용이성 등을 고려하여 구간 교통특성을 반영하는 입지와는 상당한 차이가 있는 위치에 설치되어 있는 현실이다.

한편, 통행시간의 추정 및 예측에 대한 연구는 매우 다양한 분야에서 이루어져 왔다. 통행시간과 교통량과의 관계에 대한 연구와 교통량과 점유시간을 동시에 이용하는 방법(김영찬과 김태용, 2001)에 대한 연구 및 지점검지기 자료를 이용하여 회귀모형을 작성하고, 구간검지기 방식인 번호판대조법(AVI)으로 구한 매칭율을 근거로 통행시간 융합모형을 개발(이현재, 2005)한 연구도 있다.

한편, 교통량과 속도를 이용한 방법이나 지점속도를 독립 변수로 하는 지점속도 모형, 지점교통량과 점유시간을 이용한 V/C모형, 이 두 개를 결합한 결합모형(Zhang, 1997)도 있다.

그러나 대부분의 연구가 연속류를 대상으로 하고 있으며, 단속류를 대상으로 한 연구에서도 신호의 영향 분석이나 집락간격의 결정 및 검지기의 성능을 점검하는 등의 연구에 집중되어 있다(김재진 등, 2006; 장진환 등, 2005; 임형석 등, 2005; Nam and Drew, 1996; Park et al., 2002).

따라서 본 연구에서는 GPS 시스템을 이용한 실제 국도3호선 단속류 구간의 구간 통행특성을 현지 조사하여 구간검지기 자료에서 수집되는 교통특성과 비교하고, 기존 검지 위치에 따른 구간 통행속도의 설명력 분석 및 적정 위치를 제안하고자 한다.

나아가, 통행시간 추정 및 예측을 위해 기존 연구에서의 방법론을 고찰하고 최근 IT분야에서 활발한 연구가 진행되는 자기조직형 인공신경망 구조(Self-organization Neural network structure) 방법을 도입하여 검지기 자료를 활용한 통행시간 추정모형을 제시하는 것을 목적으로 한다.

2. 기존연구 고찰

통행시간을 추정하거나 예측하는 기법은 사용된 통계적 기법에 따라 구분되는데, 대표적인 예로 1)이동평균법, 2)지수평활법, 3)자기회귀모형, 4)칼만필터링법, 5)신경망 이론(Neural Networks) 및 6)회귀분석을 이용한 기법 등이 있다(장준석, 1995; 김영찬과 김태용, 2001).

이동평균법에는 단순이동평균법, 가중이동평균법이 있으며, 이동 평균법의 한계점은 과거의 패턴이 선형이 아닌 경우 정확성이 감소되는데 과거 자료가 상승적인 추세인 경우 과소 추정을, 하락적인 추세를 가진 경우에는 과대추정의 경향이 있으며, 이동평균에 포함될 가중치 부여에 대한 명확한 근거를 제시할 수 없다는 특징이 있다.

한편, 지수평활법에는 단순지수평활법, 추세조정지수평활법 등이 있으며, 이동평균법이 지닌 결점들을 제거한 예측법이 다. 이 방법은 가중이동평균법과 마찬가지로 최근의 자료에 보다 큰 가중치를 주지만, 가중치의 부여가 능률적이며 컴퓨터를 쉽게 이용할 수 있다는 장점이 있다.

자기회귀모형은 시계열 분석기법의 하나로서 알려져 있는 자기회귀과정을 가정해서 사용되는 모형이다. 또 이것은 시계열을 확률과정으로 다루는 것이며, 시계열이 정상상태인 전제가 필요하다(이지연 등, 2003).

자기회귀모형을 통행시간의 예측에 적용할 경우의 문제점은 통행시간의 시계열은 이전 시점에서의 자료에 의존하는 부분이 적고, 그 자체가 독립적인 부분이 많다는 점이다. 자기회귀 모형은 통행시간의 추이가 자기회귀과정에 있는 것으로 가정된 모형이기 때문에 변곡점 이하에서는 과대 추정이나 과소 추정되는 어떤 발산이 일어나서 오차요인이 되고 있다. 나아가 지점, 시간대에 따라서 적합한 모형이 달라지기 때문에 모형의 적용이 어려운 문제점을 가지고 있다.

칼만필터링(Kalman filtering)은 무작위변동을 발생시키는 동적 체계에 적용되는 최적상태 예측과정으로써 이산적 실시간에 얻어진 잡음 섞인 정보로부터 동적 체계의 미지 상태에 대한 최적 예측을 위한 선형, 불편(unbiased), 최소 오차 분산의 특성을 지닌 반복적 알고리즘이다(김성현 등, 2005).

한편, 신경망이론은 입력과 출력 패턴을 분류하고 연관시킬 수 있는 학습기능을 가지고 있으므로 서로 다른 소스(source)로부터 자료를 이용할 수 있다. 즉 서로 다른 검지기로부터 받는 교통류 패턴을 이용하여 실시간 통행시간 추정 등에 효율적으로 사용할 수 있다. 공학 분야에 많이 사용되는 MLP(Multi Layer Perceptron) 신경망에 있어서, 출력 뉴런들은 입력된 출력 가중치(Weight)들에 의해서 다양하고 복잡한 결정 영역을 형성한다. 이러한 가중치들은 출력 노드들을 모든 은닉 노드들과 연결한다. 결정 영역의 복잡한 정도에 의해, 모든 출력 노드는 목표 패턴을 정확하게 학습하기 위해 일정수의 은닉 노드들을 필요로 한다. 다시 말하자면, 결정 영역의 복잡도가 높아질수록 출력층이 필요로 하는 은닉 노드들이 증가한다는 것이다.

또한, 네트워크를 구성하는 과정에서 상당수의 은닉 노드들이 불필요하게 많아질 수 있으므로, 학습 이후에 네트워크의 크기를 줄이기 위하여 은닉 노드와 출력 가중치를 제거하는 다양한 알고리즘이 제안되었다. 그러나 대부분의 알고리즘의 경우에, 먼저 학습이 수행되어 지고, 제거 과정에서 기존의 가중치들을 바꾸기 위해 다시 학습이 이루어지므로 복잡하고 시간이 걸리는 문제점이 있다(강정규와 남궁성, 2002; 김영찬과 김태용, 2001; 이의은과 김정현, 2002; 홍은주와 김영찬, 2002).

구간통행시간 예측을 위한 기존의 인공신경망(Neural Networks)의 적용에는 Input Vector로 주로 과거의 구간통행시간 또는 통행량, 속도, 점유율 등의 단일 파라미터를 사용하였다. 뿐만 아니라 인공신경망 모델을 사용하기 위해서는 결정해야 할 많은 요소가 있으나, 대부분의 적용에서는 인공신경망 모델 자체에 대한 고려는 미흡한 실정이다(홍은주와 김영찬, 2002, 장진환 등, 2005). 한편, 자기조직형 신경망을 이용한 연구는 교통분야에서는 아직 보이지 않으며, 주로 IT 분야의 연구에서 활발하게 연구되고 있다(신위재와 김동희, 2001; 이현관 등, 2001; Chang and Hwang, 1999).

따라서 본 연구에서는 신경망학습 방법을 통해 국도구간의 통행시간 추정을 위해 기존의 신경망 학습을 위한 최적 구조의 설정에서의 문제점을 해결하기 위해 자기조직형 신경망이론을 제안하고자 한다. 즉, 입력-은닉-출력층의 구조로 된 구조의 최적구성을 스스로 구성하는 방안을 제시하고 이

를 이용하여 국도 구간의 통행시간 추정에 활용하는 방안을 제시하고자 한다.

3. 국도통행특성을 고려한 구간통행시간 추정

3.1 자료 수집

본 연구에서는 단속류 특성을 보이는 국도를 대상으로 교통데이터를 수집하였으며 이를 기반으로 정보제공 대상구간의 통행시간을 추정하는 알고리즘을 개발하였다.

대상구간은 NHTMS(National Highway Traffic Management System)을 위한 인프라가 구축되어있는 국도 3호선 중 곤지암 IC~장지 IC간의 10.72km 구간으로 최근 아파트단지 개발 및 시가화로 인하여 국도기능을 상실한 구간이 많이 존재하며 상당구간에 걸쳐 그린벨트 및 녹지와 접하고 있다.

대상구간은 교차로 및 신호에 의해 단속되는 결절점과 결절점을 단일링크로 구성하였으며, 롯데아파트앞을 기점으로 하여 단속류의 특성과 연속류의 특성으로 구별할 수 있다. 표 1에서 보는 바와 같이 대상구간의 검지기는 총 9개로 평균 1.38km 간격으로 설치되어 있다.

서로 다른 검지기 체계 즉 지점검지기와 구간검지기를 통해 수집되는 자료는 서로 다른 특성을 가지고 있다, 지점검지기는 개별 링크에 설치되어 특정 시간단위의 교통량과 속도 점유율 정보가 실시간으로 수집되며, 구간정보는 구간의 시점과 종점을 통행하는데 소요되는 개별차량의 통행시간을 직접 수집한다.

VDS를 통하여 수집한 자료는 교통량, 속도, 점유율 등이 있으며, 일반적으로 두 지점 사이의 통행시간을 추정하는 방법은 AVI를 이용하는 형태와 자료수집 기술방식에 따라 시험차량(test vehicle)방식, 번호판 대조법(license plate matching)방식, ITS 프로브차량(probe vehicle)방식으로 분류한다.

국내의 경우, 장비를 이용하여 구간검지기를 상시 운영할 경우에 있어서는 이미지 방식에 의한 번호판 인식 기술을 이용한 구간검지기를 이용하고 있으며, 특정시점에 대한 조사를 실시할 경우 시험차량 운행이나 번호판 대조법을 이용

하고 있다.

이중 번호판 대조법의 경우 구간 통행시간이나 통행속도를 산출하기 위한 구간의 시종점에서 번호판을 판독하여 각 지점에서의 시간차와 거리를 이용하여 통행시간, 통행속도를 산출하는 방법으로 국내에서 상시운영중인 구간검지기의 방법과 동일한 방법이다.

현장 수집자료는 차량검지기에 의해 이루어지는데, 검지기는 검지방식에 따라 매설형과 비매설형으로 구분되며, 정보 수집 방식에 따라 지점검지 방식과 구간검지 방식으로 구분된다. 지점검지 방식은 검지센서 종류, 검지원리, 매설 여부 등에 따라 구분할 수 있으며, 구간검지 방식은 차량 번호판 자동 인식방법, GPS, Probe 차량방식, Tag 방식 등이 있다.

한편, GPS란 인공위성에서 발사된 전파를 GPS 수신기로 수신하여 위성에서 GPS 수신기까지 전파가 도달되는 시간을 이용하여 이동물체의 위치를 측정할 수 있는 시스템이다.

본 연구에서는 지점검지기에서 수집된 교통량-속도자료를 GPS 수신기를 장착한 차량의 운행결과로 얻은 통행시간(참값)과 매칭하여 구축된 분석자료를 신경망알고리즘을 통해 통행시간을 추정하는 로직을 적용하였다.

이때, 지점검지기 자료는 1분 및 5분주기의 교통량 속도 점유율이 수집되었으며, 구간검지기 자료는 7월 23일~7월 26일(평일 2일, 휴일 2일) 4일에 걸쳐서 침투별 2시간동안 4대의 GPS차량을 운행하여 얻은 (x, y)좌표 및 1초 단위의 지점속도 자료를 GIS 도면을 이용하여 10m 간격으로 분할하여(교통량, 속도)자료로 변환하여 구하였다.

한편, 구간별 통행시간의 자료는 평균주행시간으로 획득하였으며, 결측된 시간대의 데이터의 보정은 과거 이력자료를 이용하였다.

3.2 지점검지기 적정설치지점 선정

지점검지기의 설치지점 선정시에는 교통류를 관리하기 위해 교통량의 변화가 잦거나 해당구간을 대표할 수 있는 교통자료를 획득할 수 있는 지점, 돌발상황을 즉각적으로 감지할 수 있는 사고다발지점, 지하 지장물 등 장애 요소가 없는 지점 등이 전반적으로 함께 고려되어야 한다(오기도와 김영찬, 2000; 변상철과 이승재, 2006).

표 1. 구간 현황

구간	시 점	종 점	거리(m)	설치검지기	접속도로
1	곤지암 IC	삼리 2리	500		
2	삼리 2리	신이리	460	D-130	지방도337 교차
3	신이리	레미콘 공장	640		
4	레미콘 공장	쌍동 1리	550	D-120	
5	쌍동 1리	롯데아파트 앞	502		
6	롯데아파트 앞	용수리	1728	D-115, D-110	
7	용수리	선린촌	510	D-100	지방도337 교차
8	선린촌	도평리	400		
9	도평리	쌍령2리	1490	D-90	
10	쌍령2리	경안 IC	1540	D-80, D-70	구국도43 교차
11	경안 IC	장지 IC	2400	D-60	
계	총 12개 교차로, 11개구간		10723		

또한 지점검지기의 적정위치를 산정하기 위해서는 자료수집 구간내 모든 지점에서의 통행특성 자료를 수집하여야 한다. 그러나 이러한 자료수집은 시뮬레이션을 통한 방법 이외에는 완벽하게 구축하기 어려우며, 구간검지기 자료는 구간별 통행자료만을 갖는 한계로 적정 설치위치를 분석하는데 한계가 있다.

한편 국내 ITS 사업의 구축단계에서 검지기의 설치 위치에 대한 사전 연구가 이루어지지 않은 채, 평지 등 지리적인 여건과 전기공사의 용이성 등을 기준으로 설치되어 구간교통특성을 반영하는 입지와는 상당한 차이가 있다.

따라서 본 연구에서는 현재 설치된 ITS 검지기의 위치선정의 적정성을 검토하기 위해 GPS 장비를 차량에 장착하여 10m 단위로 지점속도를 산출하여 대상구간의 통행시간과 상관분석을 실시하였다(표 2 참조).

이는 이청원(2005)의 연구에서와 같은 방법으로 실시간 좌표 및 속도자료를 수집할 수 있는 GPS 장비를 이용하여 실시간으로 수집된 데이터를 이용하여 각 구간별로 각 10m 지점에 대하여 평일 오전, 오후, 비침투(주간), 휴일 오전, 오후, 비침투(주간) 등 6개 시간대로 분류하여 상관분석을 하였다.

상관분석을 토대로 적정 위치선정은 각 시간대별로 산출된 상관관계수에 시간대별 가중치와 자료의 질에 따른 별점을 적용하여 이를 합한 값이 가장 높은 것을 선택하였다.

표 2. 대상구간내 검지기 및 수집자료

구분	지점검지기		구간검지기
	VDS	NC-97	GPS
설치 및 운행대수	9	36(상/하행 9개 지점의 차로별 설치)	4
수집 자료	교통량, 속도, 점유율	교통량, 속도 (5분평균)	속도, 구간통행시간

가중치는 침투시간대의 통행에 중점을 두어 오전과 오후 시간대에 비침투 보다 2배의 가중치를 두었으며, 휴일 오전의 경우 교통특성 분석에서와 같이 침투교통이 발생하지 않는바 가중치를 두지 않았다. 이는 침투교통이 발생하여 혼잡한 상황하에서의 통행시간 추정 및 예측력을 높이기 위함이다. 그리고, 별점은 상관관계수가 0.5 미만 양의 상관관계를 갖는 것에 대해 상관관계수를 "0"으로 처리하여 별점을 부과하였다. 이는 모든 시간대에서 통행시간을 잘 반영하는 지점을 선정하기 위함이다.

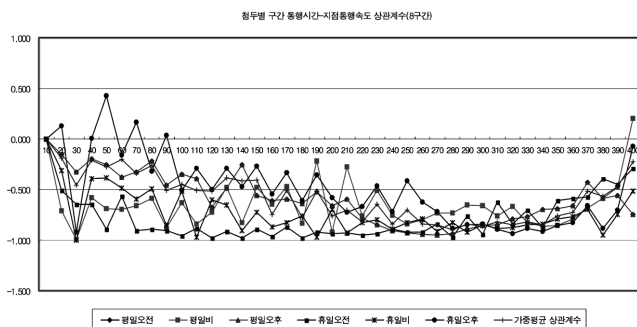


그림 1. 시간대별 상관계수 분포(8구간)

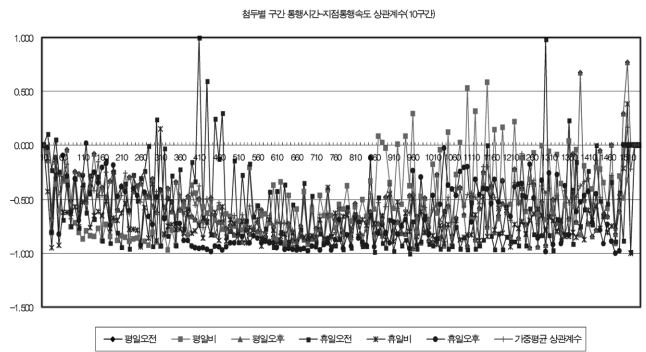


그림 2. 시간대별 상관계수 분포(10구간)

자료수집 구간별로 침투별, 위치별 상관관계수를 가중평균 상관관계수를 기준으로 볼 때, 6번, 9번, 10번, 11번구간을 제외하고는 각 자료수집 구간의 시점에서부터 상관관계수가 좋아지다가 일정시점에서 나빠지는 U자형 형태를 띠고 있다.

그리고 가장 높은 순위의 위치로 볼 때, 자료수집 구간내 중간지점보다는 구간내에서 상류부나 하류부측에 위치함을 알 수 있으며, 그림 1과 그림 2에는 제8구간 및 10구간의 구간통행시간과 지점속도간의 상관관계수를 도식화 한 것이다. 가로축에는 각 구간에서의 거리를 세로축에는 상관관계수를 나타내고 있다.

그림 1과 그림 2에서 알 수 있는 바와 같이 단속류의 특성을 가지는 국도의 경우 구간통행시간의 추정에 검지기의 설치위치가 매우 중요한 역할을 하고 있음을 알 수 있다. 즉, 속도분포의 폭이 상대적으로 좁은 연속류와는 달리, 경사 등 지형적인 여건과 신호 및 횡단보도 등과 같은 주변 환경의 영향으로 검지기의 설치위치가 수집되는 데이터에 큰 영향을 미쳐 검지기의 설치 전에 위치의 적정성을 반드시 점검해 보아야 할 것이다.

상관분석 결과, 각 구간별로 지점검지기의 최적위치 및 적정구간을 산정하면, 표 3과 같다. 최적선정위치는 각 구간(11개)에서 구간통행시간과 지점 통행시간의 상관관계수가 가장 높은 지점을 나타내며, 현 검지기의 설치위치는 각 구간에서 기존 검지기가 위치한 지점을 나타낸다.

구간별로 상관관계수는 -0.5866~-0.9038로 나타났으며, 1구

표 3. 구간별 최적위치 및 적정 구간

구간	거리 (m)	현 검지기 설치위치(m)	최적선정위치 (m)	상관계수
1구간	500		370	-0.9038
2구간	460	230	310	-0.6832
3구간	640		190	-0.7213
4구간	550	280	290	-0.8156
5구간	502		240	-0.8700
6구간	1,728	450, 860	570	-0.5866
7구간	510	337	110	-0.8475
8구간	400		280	-0.9018
9구간	1,490	487	160	-0.7354
10구간	1,540	282, 1,132	660	-0.8956
11구간	2,400	1,092	620	-0.8458

간이 가장 높은 -0.9038이며, 6구간이 가장 낮은 -0.5866을 보이고 있다.

이러한 상관계수는 구간연장이 1km 이상인 구간과 미만인 구간을 구분하여 볼 때, 상대적으로 자료수집 구간의 길이가 1km 이상인 6, 9, 10, 11구간에서 불규칙한 변동을 보임을 확인하였다. 나아가 시간대별로 상관계수 값의 차이 또한 1km 이상의 구간에서 많은 차이를 보이고 있었다.

3.3 자기조직형 신경망 학습구조

인공신경망(Artificial Neural Network) 혹은 신경망이론은 사람과 같은 유연한 정보처리 시스템을 구성하기 위하여 뇌의 구조를 모방한 모형으로 1943년 맥컬리치(McCulloch)와 피츠(Pitts)의 뉴런 모형에 대한 연구에서부터 시작되어 발전하게 되었다. 신경망은 병행적으로 상호작용 하는 여러 개의 계산요소들로 이루어져 있으며, 각 계산요소는 가중치 합(weighted sum)과 같은 단순한 계산만을 수행한다.

예측 및 추정기법에 많이 사용되는 일반적인 다층 신경회로망은 선형 노드들로 구성된 입력층과 출력층 사이에 하나 이상의 중간층, 즉 비선형 노드층으로 구성된 은닉(hidden)층을 갖는 다층(multi-layer) 구조의 회로망이다. 일반적인 다층 퍼셉트론의 학습과정은 입력층의 각 유닛에 입력 데이터를 제시하면 이 신호는 각 유닛에서 변환되어 중간층에 전달되고 최종적으로 출력층으로 나오게 된다. 이 출력값과 원하는 출력값을 비교하여 그 차이를 감소시키는 방향으로 연결강도를 조정하는 것이다.

신경망 구조의 특성은 이러한 비선형 연속함수를 근사화하는 능력을 가지고 있으므로 비선형 시스템을 모델링 하는데 사용되며, 많은 입력과 출력을 가질 수 있기 때문에 다변수 시스템에도 쉽게 응용된다.

통행시간 추정에서와 같이 교통류의 특성 변수인 교통량, 속도, 점유율 등이 통행시간과 선형관계를 가지는 경우가 드물기 때문에 이러한 비선형 함수를 이용하여 예측 기법에 활용할 수 있어 교통 분야에서 많이 이용되고 있다.

본 연구에서 제안하고자 하는 방식은 모든 출력의 오차의 평균 자승 합(MSE: mean square error)과 오차의 평균 자승합의 변화율(DMSE)을 관찰하고 이것을 이용하여 만약 출력 노드들의 수렴하지 않는 방향으로 학습이 진행된다면 수렴하지 않는 출력 노드에 대하여 은닉 노드를 추가하면서 최적 신경망 구조를 구성하는 방법이다.

한편, 은닉 노드들과 출력 노드들 간의 가중치가 학습과정에서 자동으로 추가되는 기준은 오차의 평균 자승합과 오차의 평균 자승합의 변화율을 산출하게 된다.

본 연구에서와 같이 국도를 대상으로 한 통행시간의 추정을 위해서는 여러 구간에서 많은 입력변수를 입력해야 하고, 각각의 입력 변수 별로 출력값과의 비선형 관계로 구성되기 때문에 이를 해결할 수 있는 가장 적합한 모형으로 인공신경망 이론으로 선택하였으며, 기존의 인공신경망 모형이 가지는 추가학습이 불가하다는 단점과 신경망 구조의 최적구성이 매우 어려운 점 등을 개선하기 위해 자기조직형 신경망(Self-Organizing Networks: SON) 학습방법을 도입하였다.

본 연구에서 제안한 자기조직형신경망 모델은 기존 연구에

서 사용되었던 사전에 은닉층의 개수 및 뉴런수 등을 결정하거나 임의로 정하는 방식과는 상이한 방법으로써 복잡한 비선형 시스템의 최적 추정기법이다. 따라서 복잡한 비선형 계의 예측이 가능하고 입력변수의 선택방법, 각 네트워크 구조의 선택방법 등에 따른 최적구조 설계를 통하여 주어진 시스템 특성에 맞는 구조로 적응능력을 향상시킨 모델구조를 구축할 수 있다.

신경회로망 학습에 영향을 미치는 요인으로는 초기가중치, 히든레이어의 수, 뉴런의 수를 들 수 있다. 초기가중치는 인위적인 조정의 어려움으로 인하여 초기에 임의의 값을 부여하여 주고 학습과정에서 변경하는 값이다.

히든레이어의 수와 뉴런의 수는 일반적으로 시행착오적인 방법으로 수를 결정하며, 본 연구와 같이 여러 구간의 다른 데이터로 학습을 시킬 경우에 나올 수 있는 모든 경우에 대하여 시행착오적으로 히든레이어와 뉴런의 수를 산정해야 한다.

이러한 시행착오적인 방법의 문제점을 극복하기 위해 본 연구에서 고안한 자기조직형 알고리즘의 개요는 아래와 같다.

- 1) 히든레이어 및 뉴런 수 결정 과정
 - (1) 초기 설정 : 히든레이어 1개, 뉴런 11개
 - (2) 최대 설정 : 히든레이어 2개, 뉴런 77개
- 2) 학습 방법
 - (1) 학습데이터를 초기 설정(히든레이어 1개, 뉴런11개)부터 시작
 - (2) 총 100회 학습시킨 후 뉴런수를 1개씩 증가해서 최대 설정(77개)까지 증가
 - (3) 최대 설정 뉴런까지 도달하면 히든레이어수를 1개 증가한 후 뉴런수는 초기 설정(11개)로 재 조정 후 (2)의 과정을 반복
 - (4) 최대설정 히든레이어 및 뉴런수에 도달하면 순서가 늦어서 이미 학습되어진 가중치를 사용하는 문제점과 많은 히든레이어와 뉴런수를 사용하면 학습이 늦어지는 문제점을 해결하기 위해 (2) (3)의 과정을 역순으로(최대 설정부터 최소 설정으로) 재 학습시킴
 - (5) (2),(3),(4)의 과정 중 오류합이 최저인 히든레이어수, 뉴런수, 가중치를 파일에 기록
 - (6) (5)에서 결정된 히든레이어수, 뉴런수, 가중치를 가지고 오류합이 MIN ERROR이내에 도달할 때까지 학습
- 3) 학습과정은 Back Propagation 알고리즘을 적용

한편, 히든레이어의 수나 뉴런의 수는 가변적으로 구성할 수 있으며, 본 연구에서 제시한 구조는 효율적인 수렴의 속도나 용이성을 고려하여 하나의 예에 불과함을 밝혀둔다.

나아가 은닉층의 수를 결정하는 학습횟수(100회)의 설정도 데이터의 신뢰성에 따라 가변적으로 설정가능하며, 학습과정의 종료의 판단은 오류합이 원하는 기준을 만족하는 범위에 도달할 때까지 반복되기 때문에 충분한 학습을 통해 추정 단계로 넘어가는 구조로 설계하였다.

한편, 본 연구에서 제안하는 방식은 역전파 알고리즘에 의해서 일정회수 동안 반복 학습되며, 반복되는 모든 순간에 출력 노드들에 대한 MSE와 DMSE를 계산한다. 정해진 회

수의 반복 학습이 끝난 뒤 각 출력 뉴런들의 MSE와 DMSE에 퍼지 논리를 적용하여 은닉층 노드의 증가 유무와 은닉층 노드들간의 가중치 첨가 유무를 결정하게 되면 이 과정은 전체 MSE가 목표로 설정한 오차 이하로 감소 할 때 까지 계속된다.

역전과 알고리즘에서, 각각의 가중치는 식 (1)과 같이 변한다.

$$\Delta w_{ij}(t) = \eta \delta_j a_i + a \Delta w_{ij}(t-1) \quad (1)$$

여기서, a_i 는 노드 i 의 출력값이고 η 는 학습률, a 는 모멘텀, δ_j 는 노드 j 로부터 역으로 전파되어 온 오차항이 된다. 이때 출력 노드 j 에 대한 δ_j 는 식 (2)와 같이 주어진다.

$$\delta_j = o_j(1-o_j)(t_j-o_j) \quad (2)$$

여기서, o_j 는 출력층 노드 j 의 출력이고 t_j 는 j 번째 목표값이다. 은닉노드 j 에 대한 δ_j 는 식 (3)과 같이 주어진다.

$$\delta_j = a_j(1-a_j) \sum_k \delta_k w_{jk} \quad (3)$$

여기서, a_j 는 은닉층의 j 번째 노드의 출력이고, δ_k 는 출력층의 k 번째 노드에서 전파되어 온 값이고, w_{jk} 는 은닉층의 j 번째 노드와 k 번째 노드를 연결하는 가중치이다. 학습도중에 모든 출력 노드의 MSE가 목표로 설정한 오차 ϵ 보다 작아지면 학습을 종료하게 되며 n 번째 출력노드에 대한 MSE_n 은 식 (4)와 같다.

$$MSE_n = \frac{1}{P} \sum_p (t_{pn} - o_{pn})^2 \quad (4)$$

여기서, P 는 입력패턴의 개수이고 t_{pn} 는 p 번째 패턴의 n 번째 목표값, o_{pn} 은 p 번째 패턴에 대한 n 번째 출력노드의 출력값이 된다. 출력 노드의 목표값으로의 수렴속도를 나타내는 MSE의 변화율, DMSE는 식 (5)와 같이 정의한다.

$$DMSE_n = \frac{|MSE_{n-1} - MSE_n|}{MSE_n} \quad (5)$$

여기서, MSE_n 은 n 번째 노드의 현재의 반복학습후 발생한

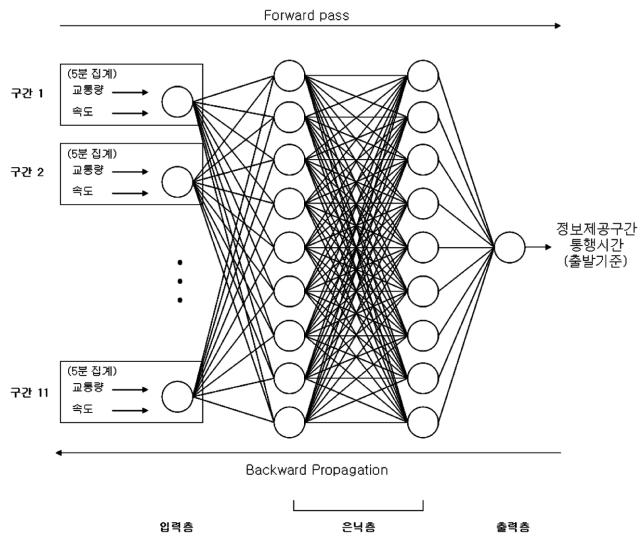


그림 3. 개발 모형의 네트워크 구성도

MSE가 되고, MSE_{n-1} 은 n 번째 노드의 이전의 반복학습에서 발생한 MSE가 된다. 일정횟수의 반복학습이 끝난후, 은닉노드를 추가할 것인지, 혹은 이미 존재하는 은닉노드와 출력노드사이의 가중치를 설정할 것인지 결정한다. 최종적으로 모든 출력 노드의 MSE가 목표로 설정한 에러값 보다 작아지면 학습을 종료하게 된다.

또한 모형의 적용에 있어서 다층 신경회로망 모형의 최적 구성을 위해 제안된 알고리즘은 각 구간에서 수집되는 자료에 따라 전체 수집되는 자료를 가중치를 적용하여 모두 사용하거나 대표검지기를 선정하여 하나의 자료만을 사용할 수 있는 방법을 자유롭게 선택 가능하도록 고안되었으며, 그림 3과 같이 11개의 각 구간의 교통자료를 입력하고 해당 정보 제공 구간(1구간: 곤지암IC-용수리, 2구간: 용수리-장지IC)의 종점도착시각기반 자료를 시점출발시각기준으로 변환시킨 통행시간을 출력층에 입력하여 학습을 시키는 구조로 이루어져 있다.



그림 4. SON 신경망프로그램의 구현

또한 본 연구에서는 구간별 수집되는 자료수에 따라 다양한 방법으로 접근하는 방법론을 적용하였으며 그림 4에서와 같이 각 구간에서 수집되는 자료에 따라 전체 수집되는 자료를 가중치를 적용하여 모두 사용하거나 대표검지기를 선

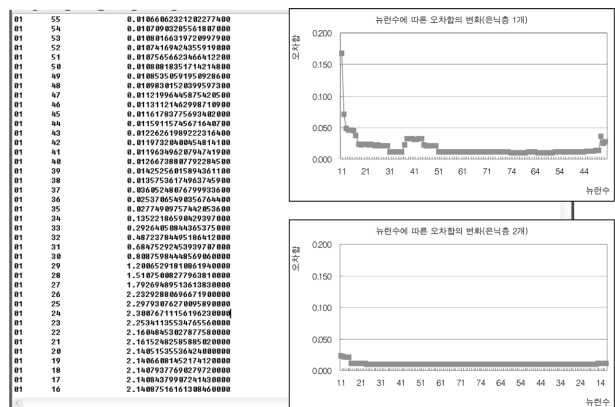


그림 5. 은닉층수에 따른 뉴런 오차합의 변화

정하여 하나의 자료만을 사용할 수 있는 방법을 자유롭게 선택 가능하도록 하였다.

또한 학습과정 및 결과가 파일로 기록되어 초기설정부터 히든레이어와 뉴런수 변화과정을 파악할 수 있다. 그림 5는 은닉층수에 따른 뉴런수의 변화와 오차합의 변화를 보여주고 있으며, 기존 연구에서와 같이 은닉층의 수를 고정하여 학습을 반복하는 구조가 아닌 오류합이 원하는 기준에 도달하는 신경망의 구조가 이루어진다면 어떤 구조이든 상관없이 학습을 종료할 수 있게 개발되었다.

즉, 자기조직형 신경망 이론은 도입하여 학습대상이 되는 데이터의 양과 질(분산 정도 등)에 대응하여 보다 쉽고 유동적인 최적 학습환경을 스스로 결정하는 시스템을 구축하였다.

4. 현장 적용 및 검증

4.1 모형의 설정

현재 국도 구간을 대상으로 구축 운영중인 국도 교통관리 시스템의 경우 구간 검지기인 AVI는 짧은 구간내의 교통상황이나 교통량과 속도의 관측보다는 통행시간 추정의 검증 등을 위한 목적으로 주로 사용되고 있어 그 활용도가 낮은 실정이다.

따라서, 본 연구에서는 앞에서 설정한 자료수집 구간에서의 지점검지기 최적 설치지점과 현재 지점검지기가 설치된 지점에서의 통행시간 추정력 검증을 위해 모형을 구분하여 적용하였다.

모형은 기존 지점검지기 위치에서의 교통량, 속도를 이용하는 경우(Model 1)와 최적위치 지점에서의 교통량, 속도를 이용하는 경우(Model 2)로 구분하고 각각의 경우마다 11개 자료수집 구간별 모형을 구축하였다(표 4 참조).

표 4. 모형별 학습데이터 유형

구분			Model 1 (현 지점 검지 위치)	Model 2 (최적 설치 선정지점)
1구간	기존 검지기 지점	Q	■	
		V	■	
	최적지점	Q		■
		V		■
.....	
11구간	기존 검지기 지점	Q	■	
		V	■	
	최적지점	Q		■
		V		■
GPS차량 구간통행시간(T)			■	■

통행시간 추정을 위한 신경망에서 데이터 셋은 학습 셋과 검증 셋으로 분리하였다. 학습을 위한 데이터 셋에 획득한 모든 자료를 학습시키고 그 중 일부를 검증을 위한 자료로 활용하는 경우와 처음부터 데이터 셋을 학습용과 검증용으로 구분하는 경우가 있다.

본 연구에서는 학습용 자료 셋과 검증용 자료 셋을 분리하고 검증용 자료는 각 시간대별 자료에서 2개씩(26일 오후

침두 자료는 1개 사용)을 추출하여 총 23개의 자료를 사용하였다. 이렇게 분리하는 이유는 모형의 추정력이 어느 시간대에 치우치는 것을 막기 위함이다.

4.2 모형의 적용 및 검증

검지기의 위치의 차이에 따른 각 모형별 통행시간의 추정력을 비교해 본 결과 각 구간의 통행특성을 가장 잘 반영하는 지점에서 구한 자료를 이용하여 통행시간을 추정하는 Model 2의 경우가 현 검지기(VDS 혹은 NC-97) 위치에서 얻은 자료로 추정한 Model 1보다 추정력이 우수한 것으로 나타났다.

통행시간 추정결과와의 정확도를 평가하기 위한 지표로는 일반적으로 사용되는 평가지표인 RMSE(Root mean square error), MARE(Mean absolute relative error), RRSE(Root relative square error)를 이용하여 평가를 시행하였으며, 각 지표별 식은 다음과 같다.

(1) MARE(Mean Absolute Relative Error)

$$\frac{1}{N} \sum_{t=1}^n \frac{|x(t) - \hat{x}(t)|}{x(t)}$$

(2) RRSE(Root Relative Square Error)

$$\sqrt{\frac{1}{N} \sum_{t=1}^n \left[\frac{|x(t) - \hat{x}(t)|}{x(t)} \right]^2 x(t)}$$

(3) RMSE(Root Mean Square Error)

$$\frac{1}{N} \sqrt{\sum_{t=1}^n |x(t) - \hat{x}(t)|^2}$$

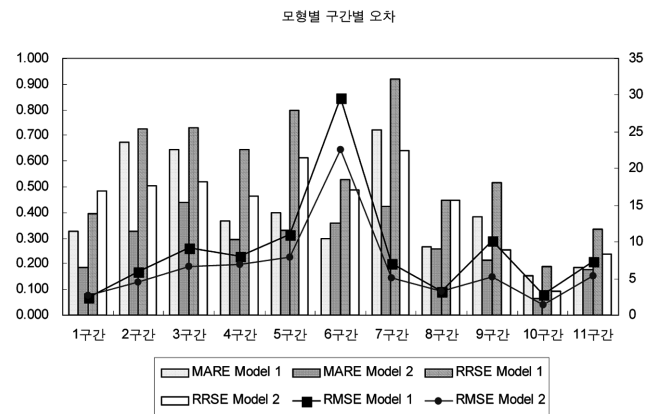


그림 6. 모형별 구간별 오차

각 구간별로 평가지표를 구분하여 나타내면 그림 6과 같으며, 적정위치에서 추정한 예측값이 그 외 지역에서 예측한 값보다 전반적으로 우수한 추정력을 보이고 있다.

오차율·값을 살펴보면 몇몇 구간을 제외하고 전반적으로 적정지점에서 자료를 수집하는 것이 통행시간 추정 신뢰도에 좋은 것으로 분석되었다.

표 5의 오차는 단속류를 대상으로 한 기존 연구와 비교해 보아도 유의한 수준이며(김성현 등, 2005; 홍은주와 김영찬,

2002; 장진환 등, 2005; 오기도와 김영찬, 2000; 김영찬과 김태용, 2001), 본 연구가 통행시간 추정치의 정도 개선보다는 자기조직형 신경망 이론의 도입 가능성과 국도 교통 특성을 반영한 검지기의 위치 선정과의 관계에 초점을 맞추었음을 밝혀둔다.

이 분석 결과를 통해 통행시간의 추정에는 통행특성을 가장 잘 반영하는 지점에서 얻은 자료를 활용하는 것이 유효함을 알 수 있었다. 이는 지금까지 일반국도의 ITS 사업 진행시 별도의 기준 없이 지점검지기를 설치하여 운영중에 있었으나, 현지 조사를 통한 적정위치에서의 지점검지기 설치가 필요하다. 따라서 본 연구 성과를 통해 향후 국도 ITS의 구축에 있어 교통특성을 고려한 검지기의 위치 선정 등에 활용될 수 있을 것으로 판단된다.

표 5. 모형별 · 구간별 오차

구분	RMSE		MARE		RRSE	
	Model 1	Model 2	Model 1	Model 2	Model 1	Model 2
1구간	2.359	2.749	0.327	0.185	0.396	0.482
2구간	5.925	4.562	0.672	0.325	0.726	0.506
3구간	9.187	6.633	0.645	0.440	0.730	0.521
4구간	8.084	6.864	0.368	0.296	0.646	0.464
5구간	11.011	7.957	0.399	0.329	0.798	0.611
6구간	29.646	22.646	0.300	0.359	0.529	0.486
7구간	7.064	5.048	0.720	0.423	0.920	0.642
8구간	3.277	3.305	0.268	0.259	0.448	0.446
9구간	10.092	5.187	0.382	0.212	0.516	0.255
10구간	2.889	1.381	0.155	0.064	0.190	0.093
11구간	7.308	5.419	0.185	0.176	0.333	0.239

5. 결론 및 향후 과제

본 연구의 목적은 국도 3호선 구간의 장지 IC ~ 곤지암 IC 구간에서 수집되는 교통정보를 기반으로 통행특성을 고려하여 통행시간 추정모형의 개발을 통해 정보 제공 시스템의 효율성과 경제성, 정확성을 높이는 방안을 제시하는 것이다.

국도 구간의 통행특성을 살펴보기 위해 대상구간에 설치된 검지기로부터 수집된 자료를 근거로 지점 검지기 적정 설치 위치에 대한 분석결과 구간의 연장에 따라 상이한 결과가 나타났으며, 이는 보다 안정적인 구간통행시간의 자료 수집을 위해서는 자료수집 구간의 연장을 길게 하지 않는 것이 유리하다고 판단된다.

또한 기존 검지기에서 수집된 자료(Model 1)와 최적검지기 설치위치에서 수집된 자료(Model 2)를 이용하여 모형을 검증한 결과 통행특성을 가장 잘 반영하는 지점 자료를 활용한 모형 즉, Model 2가 추정력이 우수한 것으로 나타났다.

이는 현재 별도의 기준 없이 지점검지기를 설치 운영하고 있는 ITS사업의 진행에 있어서 필요한 연구라고 할 수 있으며, 이러한 시도는 향후 국도 ITS 사업의 설계에서 검지기의 설치 위치 선정에 응용할 수 있을 것으로 기대된다.

단속류 특성을 보이는 구간 통행시간 추정을 위한 기존

연구들의 문제점을 살펴보고 자기조직형 신경망 이론을 도입하여 통행시간 추정방법을 제시하였다.

향후 과제로서, 적정 검지기 위치 선정에 있어서는 지점 및 구간검지기에서 구한 수집자료의 신뢰성 확보를 위한 방안 및 보다 세밀한 조사 패턴 및 수집 자료의 양에 대한 검정이 필요할 것으로 판단된다. 적절한 수집 자료의 수를 확보하기 위해서는 적정위치의 선정 방법 뿐 아니라 적정 설치간격 및 차로배치에 대한 연구가 필요할 것으로 판단된다.

통행시간 추정에 있어서는 통행시간에 영향을 미치는 변수는 다양하게 존재하므로 본 연구에서 제시된 모형이외에 다양한 상황 및 변수에 대한 고려가 필요할 것으로 판단된다.

감사의 글

본 연구는 2005년도 한밭대학교 교내학술연구비의 지원을 받았음

참고문헌

- 강정규, 남궁성(2002) 고속도로 통행료수납자료를 이용한 통행시간 예측모형 개발, **대한교통학회지**, 대한교통학회, 제20권 제4호, pp. 151-162.
- 김성현, 임강원, 이영인(2005) 일반국도의 지점 및 구간검지기 자료의 융합을 통한 통행시간 추정알고리즘 개발, **대한교통학회지**, 대한교통학회, 제23권 제5호, pp. 135-146.
- 김영찬, 김태용(2001) 검지자료합성을 통한 도시간선도로 실시간 통행시간 추정모형, **대한교통학회지**, 대한교통학회, 제19권 제6호, pp. 171-182.
- 김재진, 노정현, 박동주(2006) 구간검지체계를 이용한 On-Line 출발시각기준 링크 통행시간 추정(연속류를 중심으로), **대한교통학회지**, 대한교통학회, 제24권 제2호, pp. 157-168.
- 변상철, 이승재(2006) 일반국도의 통행특성을 고려한 지점검지기의 적정설치지점 선정에 관한 연구, **대한교통학회지**, 대한교통학회, 제24권 제2호, pp. 19-30.
- 신위재, 김동희(2001) 자기구성 신경회로망을 이용한 매니플레이터의 역기구학해에 관한 연구, *Theses Collection*, Vol. 19, pp. 57-63.
- 오기도, 김영찬(2000) 단속류 피지 통행시간 추정기의 개발, **대한교통학회지**, 대한교통학회, 제18권 제5호, pp. 57-67
- 이지연, 도명식, 김성현, 류승기(2003) 교통량 데이터의 실시간 보정모직 - 국도 3호선을 중심으로, **응용통계연구**, 제16권 2호, pp. 203-215.
- 이의은, 김정현(2002) 시간차집현상을 고려한 장거리구간 통행시간 예측모형 개발, **대한교통학회지**, 대한교통학회, 제20권 제4호, pp. 51-61.
- 이정원(2005) 교통류분석을 위한 GPS자료의 활용, **대한토목학회 논문집**, 대한토목학회, 제25권 제2D, pp. 241-246.
- 이현관, 이정훈, 엄기환(2001) 피지규칙을 이용한 신경회로망의 자동구성, **한국해양정보통신학회논문지**, 한국해양정보통신학회, 제5권 제4호.
- 이현재(2005) VDS와 AVI 기반의 최적 통행시간 추정모형 개발, 박사학위논문, 아주대학교
- 임형석, 이승환, 이현재(2005) 단속류 통행시간 추정을 위한 적정 집락간격 결정에 관한 통계적 방법론 비교 연구, **대한교통학회지**, 대한교통학회, 제23권 제3호, pp. 109-123.
- 장준석(1995) 검지기 자료를 이용한 신호화된 도시부 가로에서의 링크 통행시간 추정에 대한 연구, 석사학위논문, 서울대학교
- 장진환, 변상철, 백남철, 김성현(2005) AVI 자료 필터링 알고리즘 개발-일반국도를 중심으로-, **대한토목학회논문집**, 대한토목

- 학회, 제25권 제2D, pp. 233-240.
- 홍은주, 김영찬(2002) 지점검지체계를 이용한 남산1호 터널 구간
통행시간 추정, **한국ITS학회논문지**, 한국ITS학회, Vol. 1,
No. 1, pp. 41-51.
- Chang, Fi-John and Hwang, Yuan-Yih (1999) A self-organization
algorithm for real-time flood forecast, *Hydrol. Process.* Vol. 13,
pp. 123-138.
- Dion, F. and Rakha, H. (2003) *Estimation Spacial Travel Time
using Automatic Vehicle Identification Data*, TRB
- Nam D. H. and Drew R. D. (1996) Traffic dynamics: Model for
estimation freeway travel time in real time from flow measure-
ment, *ASCE J. of Transportation Engrg.*, Vol. 122, pp. 185-
191.
- Park, D., Rillett, L.R. and Pattanamekar, P. (2002) Estimating travel
time summary statistics of larger internals from smaller inter-
vals without storing individual data, *Transportation Research
Record 1804*, pp. 39-47.
- Zhang (1999) Link-journey - Speed Model for Arterial Traffic,
T·R·R 1676, Paper No. 99-0876, pp. 109-115.
- (접수일: 2008.3.26/심사일: 2008.4.21/심사완료일: 2008.4.21)