

■ 論 文 ■

# 시간처짐현상을 고려한 장거리구간 통행시간 예측 모형 개발

Development of a Freeway Travel Time Forecasting Model for Long Distance Section with Due Regard to Time-lag

**이 의 은**

(명지대학교 교통공학과 부교수)

**김 정 현**

((주)포스데이터 SI사업부 교통사업팀)

## 목 차

- I. 서론
    - 1. 연구의 배경 및 목적
    - 2. 연구의 내용 및 방법
  - II. 자료처리방법
    - 1. 자료수집
    - 2. 시간기준 설정
    - 3. 평균값 설정
    - 4. 통행시간 예측간격 설정
  - III. 모형구축
    - 1. 모형구축 방안 및 가정
    - 2. 모형구조 및 자료구조
  - IV. 모형평가
    - 1. 모형평가 시나리오
    - 2. 모형평가 기준
    - 3. 모형평가 결과
  - V. 결론
- 참고문헌

Key Words : 시간처짐, 통행시간, 고속도로, 신경망, TCS

## 요 약

본 연구에서는 운전자 입장에서 원하는 고속도로 다구간의 통행시간을 예측하는 모형을 구축하였다.

현재 지점검지기를 통해 생성되는 예상통행시간 정보는 장거리 통행시 발생하는 시간처짐현상을 반영하지 못하고 있다. 이로 인하여 도로이용자들의 신뢰가 떨어져, 전체적인 ATIS의 효과를 거두지 못하고 있다.

따라서 본 연구에서는 시간처짐 현상과 영업소 지체를 반영한 고속도로 다구간의 통행시간예측을 위하여, 한국도로공사에서 운영중인 검지기의 교통량 자료와 TCS자료를 사용하였다. 또한 실제 시스템에의 적용을 위해 이상치가 섞여있는 자료를 유지하였다.

예측에 사용된 모형은 3개의 입력유니트와 2개의 출력유니트를 가지는 선형신경망의 형태로 구성하였으며, 학습방법은 역전파법을 이용하였다. 또한 학습속도와 예측력에 영향을 주는 학습계수, 은닉층의 유니트수, 반복 횟수에 따라 12개의 대안을 구성하여 예측결과를 토대로 최적대안을 모형으로 채택하였다. 이러한 본 연구의 자료특성에 의해 원하는 구간까지의 통행시간을 구할 수가 있다.

본 연구는 명지대학교 연구지원금으로 이루어져 이에 감사드립니다.

# 1. 서론

## 1. 연구의 배경 및 목적

현재 대다수의 교통관리 시스템에서 생성되는 통행 시간정보는 검지기에 의해 수집된 속도를 기초로 계산된 현시점의 통행시간정보이다. 그러나 검지기 자료에 의해 계산된 통행시간과 실제 구간을 주행하는 속도는 일치하지 않는 경우가 많으며, 더욱이 운전자들이 원하는 구간통행시간은 동일시간에서 해당구간의 평균속도를 시간 단위로 변환한 통행시간이 아닌 해당구간에 진입한 지점에서부터 종점까지의 미래상황을 반영한 통행시간정보이므로 기존의 교통관리 시스템에서 생성되는 통행시간정보는 큰 의미를 가지지 못하며 제공되는 정보에 대한 이용자의 신뢰를 떨어뜨리는 결과를 야기한다.

또한 통행시간예측에 관한 선행 연구를 살펴보면 짧은 주기의 일정한 시간간격에 대한 통행시간을 예측하는 단계에 머물러 있다. 그러나 단거리통행 보다는 장거리통행을 장려해야 하는 고속도로의 경우는 장거리통행이라는 목적상 단기통행시간예측이 아닌 운전자가 원하는 목적지까지의 통행시간 정보를 제공해주어야 한다.

따라서 본 연구에서는 고속도로를 이용하거나, 이용하기를 원하는 운전자들에게 목적지까지의 통행시간정보를 제공할 수 있도록 시간처짐을 고려한 통행시간예측 알고리즘을 개발하고자 한다.

이로서 기존의 통행시간정보 보다 정확도 높은 정보를 제공함으로써 고속도로 이용자의 신뢰를 얻을

수 있고, 이러한 신뢰를 바탕으로 대체도로로의 전이 효과를 높여 전체 도로 네트워크상의 균형유지에 이바지할 수 있을 것이다.

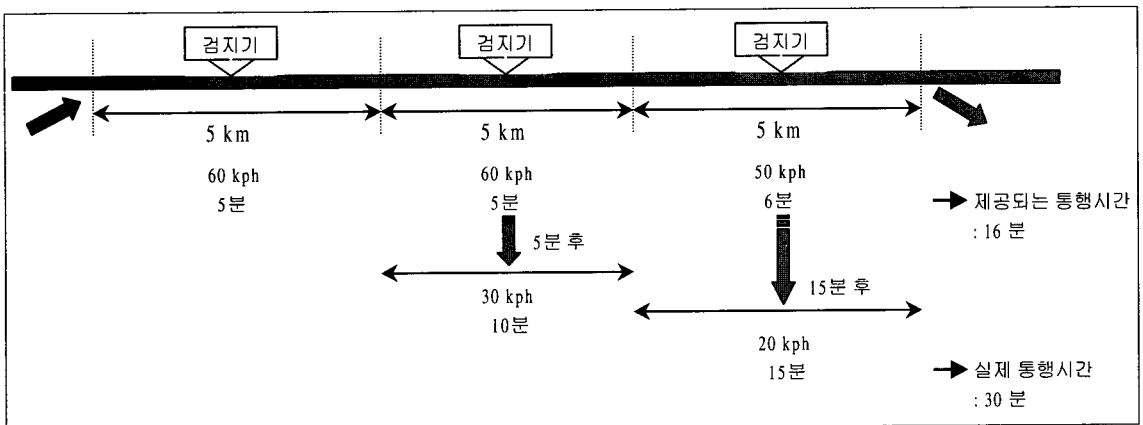
## 2. 연구의 방법

현재 대다수의 교통관리시스템에서 생성되고 있는 통행시간과 실제 운전자가 경험하게되는 통행시간간의 차이가 발생하고 있다. 이러한 실제 통행시간과 가공된 통행시간과의 차이는 여러 가지의 이유에 기인한 것이며 이를 요약하면 다음과 같다.

첫째, 실제로 소요되는 평균통행시간의 관측의 어려움이다. 현재 검지기를 통해 가공되는 통행시간은 기/중점을 검지기 간격에 따라 일정 구간으로 구분한 후, 수집된 속도자료를 통해 동일시점의 통행시간을 계산한다. 이는 주행차량조사, AVI(Automatic Vehicle Identification), 번호판 조사 등의 방법을 통하여 극복될 수 있으며, 특히 폐쇄식 영업소로 운영되는 고속도로상에서는 영업소에서 집계되는 통행권으로부터도 수집될 수 있다.

둘째, 시간처짐 현상의 발생이다. 즉 운전자가 여러 구간을 주행하는 동안 해당구간의 교통상황, 하류부 구간의 교통상황에 따라 통행시간의 변화가 발생한다. 이러한 현상은 장거리 통행이 많은 고속도로의 특성상 구간의 길이와 개수에 따라 그 차이가 더욱 크게 발생할 수 있다.

<그림 1>은 이러한 대다수의 교통관리시스템에서 생성되고 있는 통행시간과 시간처짐에 의한 실제 통행시간의 차이를 설명해준다.



<그림 1> 실제 통행시간과의 차이

〈그림 1〉의 상황을 가정해 볼 때, 현재시점(출발시점)을 기준으로 할 때는 두 번째 구간까지의 통행시간은 10분, 세 번째 구간까지의 통행시간이 16분이 소요되는 것으로 나타난다. 그러나 실제 주행시에는 첫 번째 구간을 통행한 후, 두 번째 구간에서의 교통상황의 변화로 두 번째 구간까지의 통행시간이 15분이 되며, 같은 원인으로 인하여 세 번째 구간까지의 통행시간은 30분이 된다. 이는 시간차집현상을 고려하지 않은 출발시점에서의 통행시간인 16분과는 큰 차이가 나타난다.

따라서 본 연구에서는 ATIS의 연속류구간 예상통행시간제공을 위한 방법 모색의 일환으로 신경망 이론에 의한 통행시간 예측 모형을 개발한다. 신경망 이론은 병렬처리가 가능하고, 복잡한 추론이나 판단을 쉽고 빠르게 해 줄 수 있는 인간의 두뇌 작용에 대한 호기심에서 출발하였으며, 현재 패턴인식, 최적화 문제, 예측 등 다양한 분야에서 연구 개발이 진행되고 있으며 이미 실용화되고 있다.

또한 본 연구에서는 실제 시스템에의 적용성을 고려하여 시스템의 부하를 최소화하기 위하여 smoothing을 제한, 걸러지지 않은 이상치가 섞인 자료를 사용하였다. Smoothing을 통한 자료를 토대로 모형을 구축하였을 때는 보다 정확한 결과를 산출할 수 있으나, 기본적으로 검지기 자료는 1차적인 Validation과 Smoothing 과정을 거치므로 더 이상의 평활화 처리는 자료의 왜곡시킬수도 있으며, 매주기마다 시행되는 Smoothing의 과정을 추가하면 시스템상의 부하를 가중시키는 결과를 초래할 수 있다. 따라서 본 연구는 이러한 이상치가 포함된 자료처리를 위해서는 병렬적인 구조에 기인한 Fault-Tolerance 특성을 지닌 인공신경망을 활용하였다.

신경망 이론에 기초하여 현재 검지기의 교통량 자료와 과거의 구간별 통행시간 자료를 바탕으로 장래 통행시간을 예측할 수 있는 다구간 통행시간 예측모형을 개발한다. 모형의 개발을 위해서 먼저 적절한 수의 뉴런과 은닉층을 가지는 신경망 구조를 설계한 후, 입력과 출력벡터로 구성된 학습정보를 준비한다. 이러한 학습자료중 입력벡터인 과거의 시계열 단일구간, 다구간 통행시간 정보와 검지기의 교통량 정보를 입력 뉴런으로 하고, 신경망을 진행시켜 미래의 다구간 통행시간과 하류부의 교통량을 출력으로 구한다.

신경망의 파라미터에 따른 최적대안 선정을 위해서

는 MSE, MARE, MAE를 사용하고, 선정된 모형의 적합성 평가를 위해서 지점검지기만을 이용하는 통행시간정보 알고리즘과의 비교를 실시한다.

## II. 자료처리방법

### 1. 자료수집

본 연구의 기본자료로 사용되는 교통량 자료와 구간별 통행시간 자료는 한국도로공사의 협조를 받아 -검지기의 신뢰성이 가장 높은- 경부선 시점에서부터 천안 IC까지 2000년 7월 한달간의 TCS(Toll Collection System)자료와 FTMS(검지기)자료를 수집하였다. 구득한 자료는 TCS의 경우 15분 간격으로 차종별 정리되어있으며, FTMS자료는 5분 간격으로 수집되어있다. 수집된 7월 한달간의 자료 중 4일, 11일, 18일, 19일의 자료는 모형구축단계에서 신경망의 학습을 위해 사용되며, 25일의 자료를 이용하여 모형의 테스트 및 검증용을 실시한다.

### 2. 시간기준 설정

TCS에서 수집되는 구간별 통행시간 자료는 차량이 도착한 후에 집계되는 도착기준의 통행시간이다. 그러나 운전자에게 통행시간에 대한 정보를 제공할 때는 정보제공 기점에서부터의 통행시간을 제공해야 하므로 도착기준의 통행시간을 출발기준으로 바꾸어서 '지금 여기서 출발하면 목적지까지 얼마의 시간이 소요될 것으로 예상된다'는 운전자 기준의 예상통행시간을 제공할 필요가 있다. 이를 위해 데이터 구조에서 TCS통행시간은 도착시 경과한 시간을 감하여 정보 제공시점(출발시점)과 동기화 되도록 구성하였다.

### 3. 평균값 설정

TCS자료는 차종별로 분류되어 수집된다. 따라서 통행시간 자료는 차종별로 상이한 결과가 나타나므로 해당 수집주기의 대표값을 선정하여야한다. 대표값 선정에는 평균값(mean)과 최빈값(mode)을 사용할 수 있는데, 본 연구에서는 대표값 선정을 위해서 최빈값을 사용하였다. 이는 평균값에는 장시간 휴게소 이용차량, 화물차량/공공차량 등 일부차량의 外置가,

대표값으로 선정된 전체 여행시간에 미치는 영향이 크게 작용할 수 있으므로 최빈값을 사용하는 것이 타당할 것으로 판단되었기 때문이다. 이는 여행시간 추정 및 예측시 단위 시간 수집 자료군의 대표치 추출 방법으로, 최빈값이 평균값보다 우수하다는 기존의 연구결과에서 참조하였다(이승환외, 2000). 일반적으로 1종 차량이 최빈값을 나타내는 경우가 대부분이나, 시간대와 구간에 따라 대형화물차량이 최빈값을 가지는 경우도 빈번히 발생한다.

#### 4. 통행시간 예측간격 설정

본 연구에 사용된 FTMS자료와 TCS자료의 수집 주기는 각각 5분과 15분으로 서로 동일하지가 않다. 따라서 어떤 기준이 되는 간격 설정이 필요하게 되는데, 이 때 예측결과와 활용 측면에서의 실효성을 적용한다.

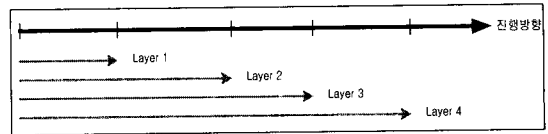
모형이 실제 시스템에 적용 가능하도록 구축되기 위해서는 예측결과에 대한 교통정보를 효과적으로 제공할 수 있는 간격을 확보하여야 한다. 대상구간의 영업소간 통행시간을 고려할 때, 가장 짧은 통행시간을 나타내는 구간은, 서울영업소 ⇒ 수원 I/C 구간으로서 최소값이 약 7분 전/후로 나타나고 있다. 이러한 경우 TCS자료의 수집주기인 15분을 적용하기에는 실효성이 떨어질 것으로 판단되며 따라서 FTMS 자료의 수집주기인 5분을 통행시간 예측간격으로 설정하고, 15분 간격인 TCS자료를 5분 간격으로 만들기 위해서 선형보간법을 사용한다.

### III. 모형구축

#### 1. 모형구축 방안 및 가정

##### 1) 모형구축 방안

TCS의 통행권 데이터는 기점영업소와 종점영업소에서의 지체가 포함되어 있다. 특히 폐쇄형 영업소는 대다수가 본선이 아닌 램프를 통하여 유출한 이후에 설치되어 있다. 그러므로 TCS 데이터는 영업소에서의 지체뿐 아니라 램프부의 통행시간이 포함된다. 따라서 다구간의 경우 운전자의 종점에 따라 지체를 경험하게되는 영업소(기/종점 영업소)와 지체를 경험하



〈그림 2〉 Layer에 따른 구간중복

지 않는 영업소(기/종점의 사이에 위치하는 영업소)가 구분된다. 이를 반영하기 위해서는 〈그림 2〉와 같이 각 구간별로 단계(layer)를 구분한 후 구간을 중복하여 모형을 진행시켜, 종점에 따른 영업소 지체의 유/무를 구분한다.

따라서 시점으로부터 원하는 종점까지의 통행시간을 선택적으로 산출할 수 있다.

이는 또한 신경망의 이식성 문제를 극복할 수 있는 방안으로 사용된다.

기존의 연구결과에 따르면 신경망 이론이 가지는 한계점인, 학습과정을 특정자료에 국한하여 실시하고 사용자가 이해하기 어려운 형태의 매개변수를 산출하기 때문에 다른 장소(구간)으로 이식을 위한 매개변수의 임의수정이 어렵기 때문에 뛰어난 예측력에도 불구하고 이식성이 부족한 것으로 알려졌다(김영찬외, 1997).

그러나 본 연구의 대상범위는, 많은 링크가 혼재되어 있는 도시부가 아닌 고속도로라는 점과 구간별로 단계를 나누어 신경망을 학습시키는 방법으로 이러한 이식성문제는 해결될 것으로 보인다.

#### 2) 모형구축의 가정

본 모형구축을 위해서는 다음과 같은 가정이 요구된다.

첫째, 특정일(주말, 공휴일)을 제외한 고속도로 구간간의 상대적인 통행패턴(O/D 패턴에 따른 IC간 진출입 교통량 비율)은 과거패턴을 유지한다.

둘째, 미래의 하류부 소통상태는 현재의 소통상태와, 상류부의 현재 교통량/소통상태에 의해 결정된다.

셋째, 수집된 자료는 정상시의 고속도로 상황을 반영한 것이다. 즉 유고 등의 돌발상황이 발생하지 않은 고속도로의 상황을 반영한 것이다.

넷째, 단위시간 수집 자료군에서 선정된 통행시간 자료의 대표값에는 휴게소 이용차량, 차량정비등의 경우는 제외된 순수 주행시간과, 영업소에서의 지체, 램프부의 통행시간 만이 포함되어 있다.

## 2. 모형구조 및 자료구조

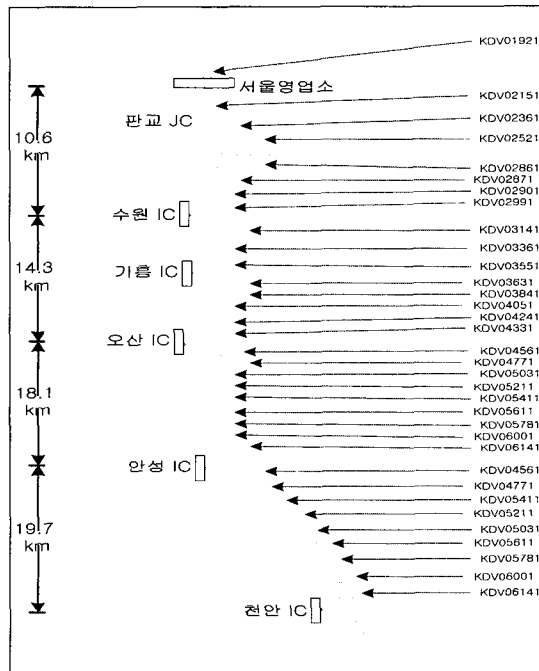
### 1) 모형구조

모형의 개발을 위해서는 우선 모형적용의 가능성을 평가하기 위해 대상구간을 설정하여야 한다. 따라서 수집된 자료(경부선의 TCS, FTMS자료)를 토대로 경부선 하행의 서울영업소에서 천안 IC에 이르는 구간을 연구의 대상범위로 선정하였다.

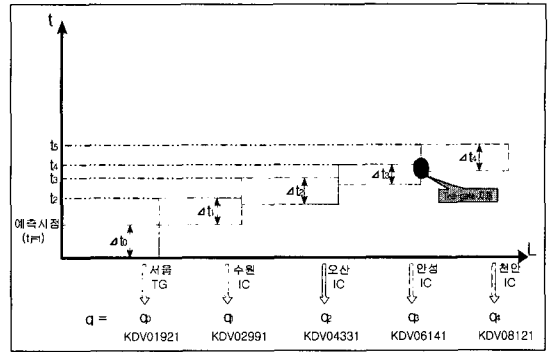
〈그림 3〉과 같이 서울영업소와 천안 IC간에는 수원 IC, 기흥 IC, 오산 IC, 안성 IC가 있으나 TCS 자료를 분석한 결과, 이 중 기흥 IC로의 유출교통량이 너무 적은 관계로 분석에서 제외하였다.

따라서 모형은 총 4개의 layer로 구성되며, 각각의 layer는 『서울영업소⇒수원 IC』, 『수원 IC⇒오산 IC』, 『오산 IC⇒안성 IC』, 『안성 IC⇒천안 IC』로 이루어진다.

〈그림 4〉는 구간과 통행시간을 도식화한 것이다. 이 중 시간이 겹치는 부분(그림에서 원으로 표현된 부분)은 영업소에서의 지체를 나타낸 것이다.



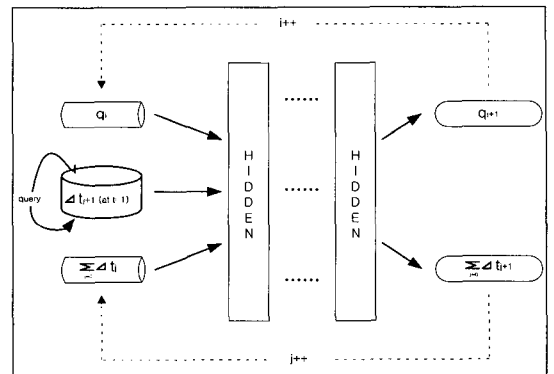
〈그림 3〉 연구 대상구간의 범위



〈그림 4〉 구간과 통행시간의 관계

위와 같은 모형을 구축하기 위해 본 연구에서는 단극성 시그모이드 함수를 변환함수로 채택하였으며, 학습방법으로는 지도학습 방식을 채택하여, Delta rule의 일종인 BP (Back-propagation) 알고리즘을 사용한다. BP 알고리즘은 〈그림 5〉와 같이 순방향 다층 신경망의 학습에 효과적으로 적용할 수 있어서 다양한 분야에 보편적으로 활용되는 매우 중요한 학습 알고리즘으로 교통분야의 동적체계를 설명하는데 가장 유용한 것으로 알려져 있다(Mark Dougherty, 1995).<sup>1)</sup>

BP 알고리즘을 이용한 신경망의 학습은 3단계로 이루어진다. 1단계로 학습입력 패턴을 신경망에 입력하여 출력을 구한다. 2단계는 출력과 목표치의 차이 즉, 오차를 구한다. 3단계는 오차값을 역방향으로 전파시키면서 출력층의 연결강도 및 은닉층의 연결강도를 변경한다.



〈그림 5〉 신경망 모형의 설계

1) Mark Dougherty(1995), "A Review of Neural Networks Applied to Transport". Transportation Research-C, Vol.3, pp.247~260.

〈그림 5〉와 같이 신경망을 구성하여 3개의 입력 뉴런과 2개의 출력 뉴런을 가지는 모형을 설계하였다. 여기서 출력으로 나오는 하류부의 교통량(검지기 자료)과 하류구간까지를 포함한 통행시간은 다음 layer의 입력값으로 사용(return)된다.

2) 자료구조

(1) 입력자료

입력자료는 3가지로 구분된다. 그 중 첫 번째 항목은 대상구간 시점까지의 TCS 통행시간이 입력된다. 알고리즘의 최초 단계에서는 대상구간이 시작되므로, 그 값이 항상 '0'이 되며, layer의 진행에 따라 출력값을 받는다.

두 번째 항목은 대상구간 시점에서 가장 근접한 상류부 검지기의 교통량 자료가 입력된다. 이는 "특정구간의 장래 교통상황이 현재의 그 도로구간 상류부의 교통상황의 결과라는 동적모형의 이론에 따른 것이다. 이 자료는 첫 번째 layer에서는 검지기의 실시간 교통량 자료를 받으며, 두 번째 layer에서 부터는 출력값을 받는다.

마지막 세 번째 입력자료는 대상구간의 과거 TCS 통행시간 자료이다. 이 값은 DB에 저장된 것으로, PK(Primary Key)인 시간값에 첫 번째 자료의 시간이 경과한 후의 값을 query하여 사용한다.

(2) 출력자료

출력자료는 2가지로서, 첫 번째 출력자료로는 대상구간의 종점까지의 통행시간이다. 이 값이 운전자에게로의 제공될 통행시간으로 사용되며 다음 layer의 입력값으로 사용되는데 여기서 중요한 점은 해당구간의 통행시간을 상류구간 통행시간과 단순한 합이 아닌, 신경망 학습과정에서 다구간에 걸친 기/종점을 가진 차량의 통행시간을 목표(target)로 학습시킴으로 해서 기/종점 사이의 거치지 않는 영업소의 지체를 제외시킬 수 있다.

두 번째 출력자료는 대상구간내의 최하류부 검지기, 즉 다음 대상구간의 시점에 가장 근접한 상류부 검지기의 교통량으로서 이 또한 다음 layer로 사용된다.

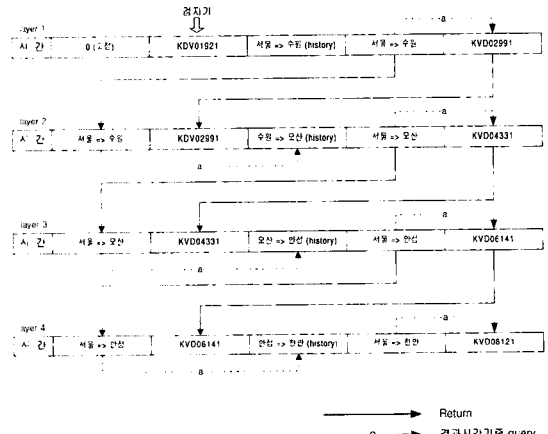
이러한 자료의 구조는 〈표 1〉과 〈그림 6〉에서 설명되고 있다.

본 연구에 사용된 자료들의 상관분석결과 〈표 2〉와 같이, t 통계량이 전체적으로 2이상의 값을 나타냈으며, 이는 각 파라미터가 통계적으로 유의하다는 것을 나타낸다.

〈표 1〉 입/출력 자료

시간 (PK)	입력			Target(출력)	
	$\sum \Delta t_j^{**}$	$q_i$	$\Delta t_j$ (at, t=j) from DB	$\sum \Delta t_{j+1}$	$q_{i+1}$
5분 간격	0	KDV01921	서울 ⇒수원*	서울 ⇒수원	KDV02991
	서울 ⇒수원	KDV02991	수원 ⇒오산*	서울 ⇒오산	KDV04331
	서울 ⇒오산	KDV04331	오산 ⇒안성*	서울 ⇒안성	KDV06141
	서울 ⇒안성	KDV06141	안성 ⇒천안*	서울 ⇒천안	KDV08121

- \* : 과거자료를 query
- \*\* : 여기서  $\sum$ 은 단순히 수학적인 합이 아닌 기/종점 외의 영업소에서의 지체를 제외한 값이다.



〈그림 6〉 자료간의 query, return 관계

〈표 2〉 파라미터별 t 통계량

		t 통계량		
		X1	X2	X3
Layer 1	f4	N/A	0.1875	20.9485
	f5	N/A	32.0067	19.9673
Layer 2	f4	30.8658	9.0325	4.8224
	f5	2.6660	31.2234	4.1435
Layer 3	f4	28.1651	4.8901	12.3225
	f5	5.2435	13.7820	1.1361
Layer 4	f4	48.7609	3.4051	3.4057
	f5	9.9568	11.6410	3.4243

- f4 : 대상구간까지의 누적 통행시간( $\sum \Delta t_{j+1}$ )
- f5 : 대상구간내 최하류부 station 교통량( $q_{i+1}$ )
- X1 : 대상구간 시점까지의 누적 통행시간( $\sum \Delta t_j$ )
- X2 : 대상구간 시점에 가장 인접한 station 교통량( $q_i$ )
- X3 : 대상구간의 과거 통행시간  
(과거자료 :  $\Delta t_j$  (at, t=j) from DB)

## N. 모형평가

### 1. 모형평가 시나리오

수집된 자료로부터 신경망의 학습을 위해 7월 4일, 11일, 18일, 19일의 자료를 사용하고, 구축된 모형의 평가를 위해서는 7월 25일 하루동안의 자료를 사용하여 최종 목적지인 서울영업소에서 천안 IC까지 (61.7km)의 통행시간을 실제 통행시간(TCS 자료)와 비교한다.

신경망 적용에서의 주된 과제는 학습과정에서 나타나는 오차의 진동을 최소화하고 빠른 수렴에 도달할 수 있도록 하는데 있다. 이러한 진동은 신경망의 연결강도와 임계치를 조절해주는 학습계수에 영향을 받을 뿐 아니라 은닉층의 수와 각 층이 갖고 있는 뉴런(PE)의 수에도 영향을 받는다. 따라서 본 연구에서는 <표 3>과 같이 은닉층의 수를 2로 하고 각 은닉층의 유니트가 10, 20인 경우, 학습계수가 0.3, 0.5, 0.7인 경우 그리고 반복횟수<sup>2)</sup>가 10000번, 20000번인 경우에 따라 12개의 대안을 설정하고 이에 따라 최적의 대안을 선택하는 과정을 택하였다.

이러한 신경망 모형은 학습정보의 입력자료에 의한 모형의 출력값과 학습정보의 목표 출력값을 비교하여

이들간의 오차가 최소가 되도록 유니트간의 연결강도를 조정하는 학습과정이 필요하다. 이를 위해 구축된 신경망 모형을 C언어로 작성하여 사용한다.

### 2. 모형평가 기준

일반적으로 구축된 예측모형의 경우 그 수행결과를 평가하는 척도는 다양하며, 아직까지 절대적인 평가 기준은 없고 단지 다음과 같은 지표를 토대로 상대적인 평가에 의존한다.

#### 1) MSE (Mean Square Error)

$$MSE = \frac{\sum_{i=1}^n (y_t - y_t^*)^2}{n} \quad (1)$$

여기서  $y_t$ 는  $t$ 시점에서의 실제값,  $y_t^*$ 는 예측값을,  $n$ 은 표본수를 의미한다.

#### 2) MARE (Mean Absolute Relation Error)

$$MARE = \frac{\sum_{i=1}^n \frac{|y_t - y_t^*|}{y_t}}{n} \quad (2)$$

#### 3) MAE (Mean Absolute Error)

$$MAE = \frac{\sum_{i=1}^n |y_t - y_t^*|}{n} \quad (3)$$

이러한 지표중에 가장 많이 사용되는 지표인 MSE는, 작은 오차는 상대적으로 작게 나타내어지나 오차가 커질수록 값은 지수함수의 형태로 증가한다.

### 3. 모형평가 결과

각 대안별로 수행된 테스트 결과인 <그림 7~18>를 종합하면 다음과 같다.

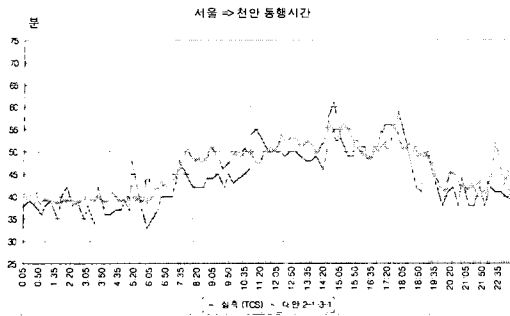
첫째, 학습계수와 예측력의 관계는 반비례의 관계를 나타낸다는 기존의 연구<sup>3)</sup>를 뒷받침하고 있다.

<표 3> 신경망 변수에 따른 모형대안

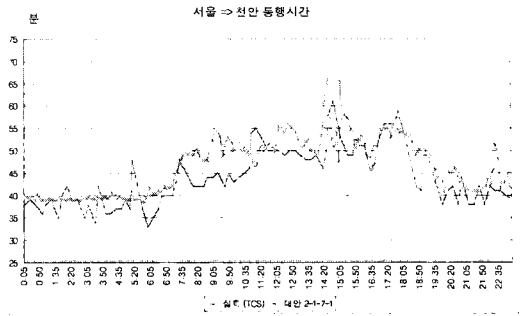
은닉층수	은닉층 뉴런수	학습계수	반복횟수	모형대안
2	10	0.3	10000	2-1-3-1
			20000	2-1-3-2
		0.5	10000	2-1-5-1
			20000	2-1-5-2
		0.7	10000	2-1-7-1
			20000	2-1-7-2
20	0.3	10000	2-2-3-1	
		20000	2-2-3-2	
	0.5	10000	2-2-5-1	
		20000	2-2-5-2	
	0.7	10000	2-2-7-1	
		20000	2-2-7-2	

2) 최소대안 설정시 반복횟수를 1000번, 5000번, 10000번, 20000번으로 설정하였으나, 1000번, 5000번의 경우 수렴하지 않는 대안이 많이 발생하여 10000번, 20000번의 대안만을 비교.

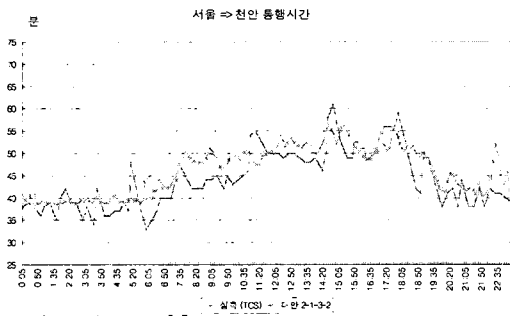
3) Ardeshir Faghri & Aneja Sandeep, (1998), "Analysis of performance of backpropagation ANN with different training parameter" from "Neural Network in Transportation Applications".



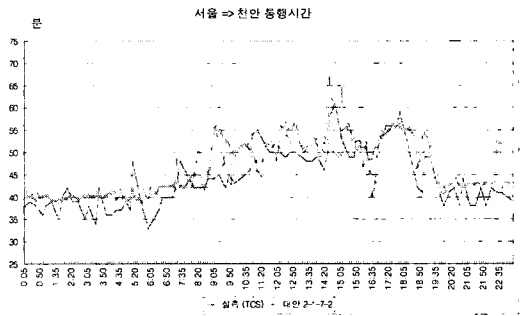
<그림 7> 대안 2-1-3-1



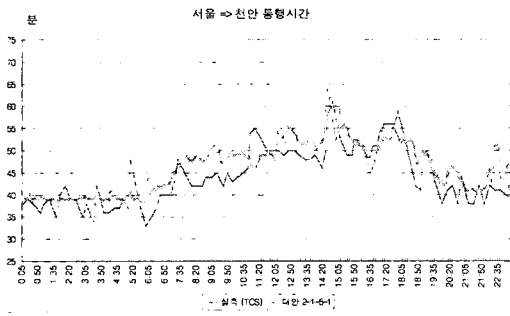
<그림 11> 대안 2-1-7-1



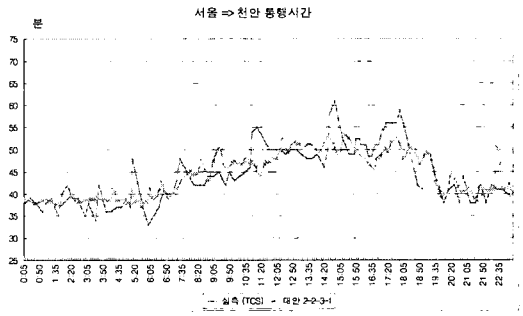
<그림 8> 대안 2-1-3-2



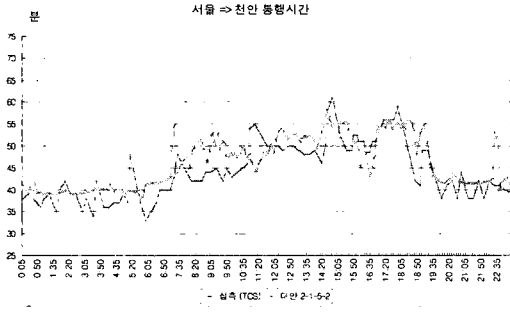
<그림 12> 대안 2-1-7-2



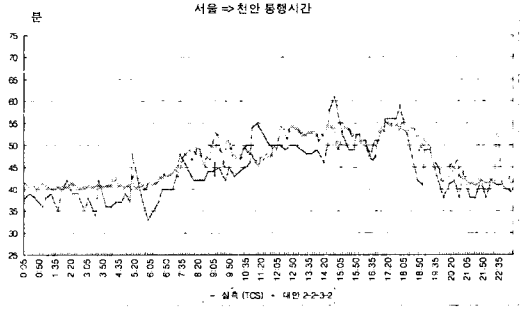
<그림 9> 대안 2-1-5-1



<그림 13> 대안 2-2-3-1

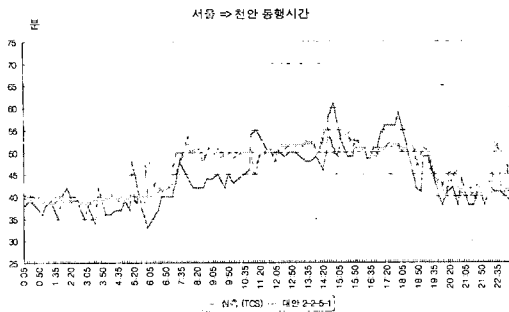


<그림 10> 대안 2-1-5-2

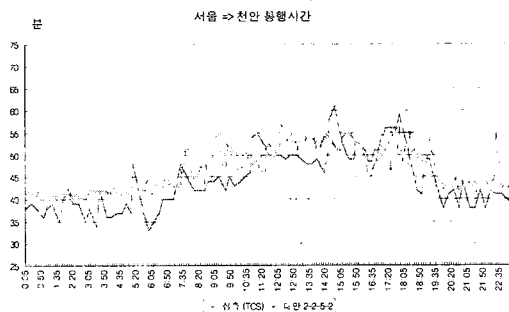


<그림 14> 대안 2-2-3-2

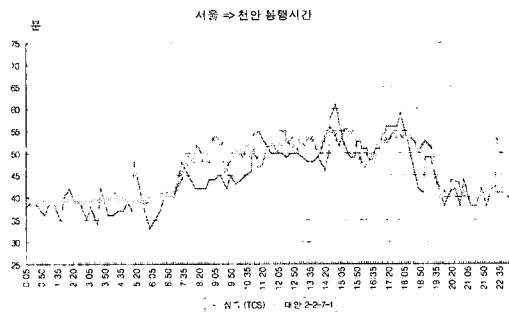




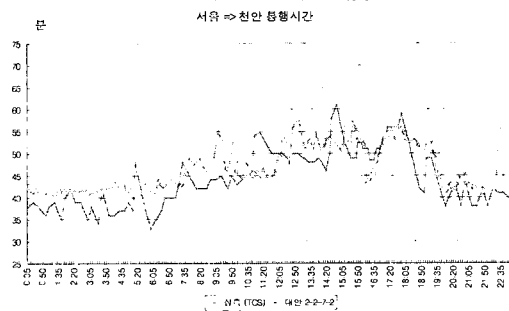
〈그림 15〉 대안 2-2-5-1



〈그림 16〉 대안 2-2-5-2



〈그림 17〉 대안 2-2-7-1



〈그림 18〉 대안 2-2-7-2

둘째, 은닉층의 유니트수에 따른 예측력의 변화는 일정한 패턴을 갖지 못하고 있다. 이는 은닉층의 유니트수가 많을수록, 지역최소(local minimum)에 빠지지 않도록 해준다는 일반적인 생각과는 차이를 보이고 있다. 그러나 BP 알고리즘의 가장 좋은 구조는 은닉층 유니트의 수를 최소화할 때 얻어진다는 최근의 상반된 연구결과와도 병행하여 고려되어야 한다.

셋째, 반복횟수에 의한 차이는 오히려 반복횟수를 적게(반복횟수 10000번) 했을 때 대체적으로 좋은 결과를 도출하였다. 이는 은닉층의 유니트수와 반복횟수의 증가에 따라 학습자료에 지나치게 적합(fitting) 되는 과적합화(over-fitting) 문제에 의한 것으로 보인다. 따라서 튜닝(tuning)과정에서 은닉유니트의 수를 제한하거나, 학습의 반복횟수를 제한할 필요가 있음이 나타났다(Smith & Murray, 1993).

분석결과로 도출된 최적의 대안모형은 2개의 은닉층, 각 은닉층별 20개의 unit, 학습계수 0.3, 반복학습횟수 10000번인 대안(2-2-3-1)으로 나타났다. 대안 2-2-3-1은 〈표 4〉에서 보는 것처럼 MSE, MARE, MAE에서 모두 가장 결과를 보이고 있다.

또한 선정된 최적 대안모형의 실효성을 파악하기 위하여 제1장의 1절에서 설명되어 있는, 지점검지기를 통하여 생성되는 예상통행시간 산출 방식과의 비교를 실시하였다.

참조모형과의 비교평가기준은 대안의 평가에 사용된 MSE, MARE, MAE와 더불어 통계적 검증인 T-test를 실시하였다. T-test는 평균간의 유사/상이성을 검

〈표 4〉 대안의 평가 결과

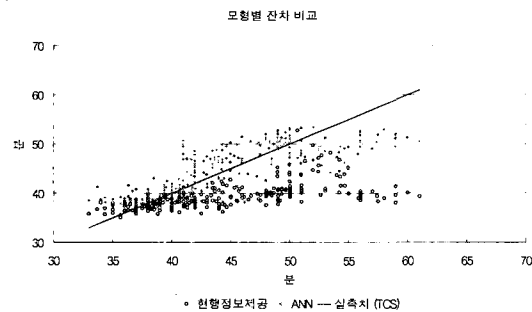
모형대안	MSE	MARE	MAE	비고
2-1-3-1	15.2089	0.0751806	3.2617499	
2-1-3-2	16.4478	0.0756748	3.3109616	
2-1-5-1	15.5348	0.0756075	3.2986693	
2-1-5-2	19.0104	0.0800490	3.4921141	
2-1-7-1	21.0568	0.0831011	3.6462435	
2-1-7-2	21.6018	0.0860577	3.7499645	
2-2-3-1	12.4084	0.0600701	2.7245150	채택
2-2-3-2	16.9274	0.0807090	3.4767056	
2-2-5-1	18.2911	0.0777677	3.4000806	
2-2-5-2	21.6958	0.0922431	3.9670581	
2-2-7-1	16.7107	0.0733718	3.2217718	
2-2-7-2	22.4680	0.0917658	3.9474330	

〈표 5〉 각 모형의 평균과 표준편차

	평균	표준편차
실측값(TCS)	44.21942	6.291036
지점검지기 사용	39.30199	2.492467
ANN	44.36578	4.969324

〈표 6〉 각 모형별 평가지표 비교

	MSE	MARE	MAE	유의수준	
지점 검지기 사용	108.871	0.11040	5.34573	0.1	Reject
				0.05	Reject
				0.01	Reject
ANN	12.4084	0.06007	2.72451	0.1	Accept
				0.05	Accept
				0.01	Accept



〈그림 19〉 각 모형별 전차 비교

증하는 방식으로서 이를 통해 각 모형이 어느 정도의 수준까지 유의한지를 평가, 비교하였다.

T-test에 사용된 가설  $H_0$ 는 귀무가설로서 "상황은 동일하나 우연에 의한 차이가 있으며 같은 모집단에서 추출한 것"으로 설정하였다.

그 결과는 〈표 6〉에 나타나있다.

위의 결과와 같이 인공신경망 모형은 테스트된 유의수준 0.1, 0.05, 0.01에서 모두 유의한 것으로 나타났으나 지점검지기로부터 통행시간을 생성하는 알고리즘은 모든 신뢰수준에서 기각되었다. 또한, 현재 제공되고 있는 예상 통행시간의 평균절대상대오차(MARE)가 11%인 반면 인공신경망에 의해 구축된 모형은 6%로 나타났다.

〈그림 19〉와 같이 지점검지기로부터의 통행시간 산출 알고리즘에 의한 값은 실측값(TCS 자료를 출발기준으로 정리한 통행시간)보다 과소 추정되고 있으며, 이러한 현상은 특히 혼잡도가 증가할수록(지점통행속도

가 떨어질수록) 크게 나타나고 있다. 이는 지점검지기를 사용하여 가공된 통행시간이 정체상황의 변화를 반영하지 못하는 한 단면에서 수집된 정보에 의존함과 함께 시간차집현상 뿐 아니라 교통량 증가에 따라 영업소에서의 지체도 증가하기 때문인 것으로 보인다.

### V. 결론

본 연구에서는 기존에 설치되어 있는 검지기와 지방부 고속도로에 설치된 TCS(Toll Collection System)의 자료를 이용하여, 운전자가 원하는 교통정보서비스(多구간 통행시간)를 제공할 수 있는 알고리즘을 개발하였다.

이미 여러 통행시간예측 알고리즘이 개발되어 있으나, 장거리 통행시 크게 발생하게되는 시간차집 현상을 고려하지 못하고 있다. 따라서 본 논문에서는 시간차집 현상과 영업소에서의 지체를 반영하여 신뢰성 있는 통행시간 정보의 생성에 초점을 맞추었다.

이를 위해 TCS의 도착기준 자료를 출발기준으로 변환하여 운전자의 제공 획득 시점과 동기화 시켰으며 선행신경망을 이용하였다. 또한 구간을 layer로 구분/중첩하여, layer에 따라 원하는 구간까지의 통행시간을 구할 수 있는 모형을 구축하였다.

이렇게 구축된 모형은 대상구간 시점에 가장 가깝게 위치한 상류부의 검지기 자료(교통량)를 시발(trigger)로 하고 DB에 저장한 단일구간 과거 통행시간 자료를 참조하여, 다구간의 통행시간 예측을 수행한다. 본 모형은 정보제공시 운전자의 인지성을 고려하여 영업소간을 1개의 단위로 하는 거시적모형으로 구성된다.

그 예측력은 참조모형으로 선정된, 지점검지기자료로부터 통행시간을 생성하는 알고리즘과의 비교를 수행한 MSE, MARE, MAE, T-test 결과의 모든 측면에서 우수한 것으로 판명되었다. 또한 신경망 이론에 기반을 둔 본 모형은 기존의 가중치에 새로운 교통상황에 대한 추가적인 학습이 가능하며, 오류에 대해 관대한(Fault tolerance) 신경망의 특성상 이상치가 섞여있는 현상의 상황에 가장 잘 부합될 것으로 보인다.

본 모형 구조는 예측값을 다음 layer의 입력으로 사용함으로써, layer가 반복(구간이 길어질수록)될수록 오차의 누적이 발생할 수 있다. 따라서 실제 적용 시에는 허용가능한 오차의 범위를 설정한 후, 그 범

위 내에서 만족되는 구간까지의 정보를 제공하는 방법을 사용해야 할 것으로 보인다. Test에 적용된 대상구간인 '서울영업소 ⇒ 천안 IC'의 경우는 4개의 layer를 중첩(3번의 예측값을 입력으로 사용)하였으나, 그 오차의 범위가 6%인 것으로 나타나 충분한 예측력을 갖는 것으로 판단된다.

그러나 본 연구에 사용한 TCS 데이터에는 고장, 휴식 등의 이유로 실제 주행시간뿐만이 아닌 정지시간이 포함되어 있을 수 있다. 본 연구에서는 수집된 자료를 이용하였기에 실제 주행시간외에 정지된 시간을 제외시키지 못했으며, 이는 향후 probe등을 이용하여 실제 주행시간이 수집되어야만이 가능할 것으로 보이며, 연구에 사용된 자료가 한정된 자료에 의한 결과이므로 보다 광범위한 자료를 통한 일반화가 필요하다.

**참고문헌**

1. 김남선·이승환·오영태(2000), "신경망을 이용한 고속도로 여행시간 추정 및 예측모형 개발", 대한교통학회지, 제18권 제1호, 대한교통학회, pp.47~59.
2. 김영찬·최기주·김도경·오기도(1997), "단일루프검지기를 이용한 간선도로 실시간 통행속도 추정 방법론", 대한교통학회지, 제15권 제4호, 대한교통학회, pp.53~71.
3. 김현주(1996), "인공신경망을 이용한 고속도로 실시간 교통량 예측", 명지대학교 석사논문.
4. 박병규(1994), "신경망 이론에 의한 링크 통행시간 예측모형의 개발", 한양대학교 석사논문.
5. 오창석(1996), "뉴로컴퓨터", 지성출판사.
6. 高速道路調査會(1997), 新東京國際空港關連道路における交通管制に関する研究.
7. Ardeshir Faghri, Aneja sandeep(1998), "Analysis of performance of backpropagation ANN with different training parameters", VTT Communities and Transport Research Finland Report, Neural Networks in Transport Applications.

8. Hiroshi matsui, Montohiro fujita(1998), "Travel time prediction for freeway traffic information by neural network driven fuzzy reasoning", VTT Communities and Transport Research Finland Report, Neural Networks in Transport Applications.
9. Robert A.Hamm, Donald A.Woods, Vergil G.Stover(1993), "Congestion management systems state of the practice review" Federal Highway Administration Report No. FHWA/TX-94/563-1F, Texas Transportation Institute the Texas A&M University System College.
10. Shawn M. Turner, William L. Eisele, Robert J. Benz(1998), "Travel time data collection handbook", Federal Highway Administration Report No. FHWA-PL-98-035, Texas Transportation Institute the Texas A&M University System College.
11. Smith, Murray(1993), "Neural network for statistical modeling", Van Nostrand Reinhold.
12. Takeshi kurokawa, Keiichi ogawa(1998), "A study on travel time prediction method on inter-city expressways", Proceedings of 5th World Congress on Intelligent Transportation Systems.
13. Toru otokita, Kana hashiba(1998), "Travel time prediction based on pattern extraction from database", Proceedings of 5th World Congress on Intelligent Transportation Systems.

♣ 주 작 성 자 : 김정현  
 ♣ 논문투고일 : 2001. 11. 8  
 논문심사일 : 2002. 1. 9 (1차)  
 2002. 5. 21 (2차)  
 2002. 6. 14 (3차)  
 2002. 7. 11 (4차)  
 심사판정일 : 2002. 7. 11  
 ♣ 반론접수기간 : 2002. 12. 31