

# 주변 도로의 교통량 Pattern을 학습 및 적용한 버스도착시간 예측<sup>1)</sup>

유종빈\*, 이찬근\*, 강현철\*, 박호현\*  
 \*중앙대학교 공과대학 전자전기공학부  
 e-mail:rjbkings@wm.cau.ac.kr

## Predict a bus arrival time from traffic volume of surrounding roads

Jong-Bin Ryu\*, Chan-Gun Lee\*, Hyun-Chul Kang\*,  
 Ho-Hyun Park\*

\*School of Electrical and Electronics Engineering,  
 Chung-Ang University

### 요약

BMS(Bus Management System)의 핵심인 버스도착예정시간을 산출하는 데 있어서 기존 대부분의 도시에서는 시계열 모형의 이동평균법, 칼만필터링 등으로 버스도착예정시간을 예측하고 있으나 이는 급격한 통행량의 변화 또는 급작스러운 사고, 신호체계 등에 적용 할 수 없다.

따라서 본 논문에서는 주변 도로의 통행량에 따른 버스의 정류장 도착시간을 예측하는 방법을 제안한다.

주변 도로의 통행량과 실제 버스의 통행시간을 실측하여 기록, 학습하고 모델링하여 미래의 버스의 운행시간을 예측하는 방법이다.

또, 이동평균법에 의한 버스도착시간 예측결과와 본 논문에서 제안하는 결과와 비교, 분석하였다.

### 1. 서론

U-City 구축의 일환으로 BMS(Bus Management System)의 연구·개발이 활발히 이루어지고 있으며, 이미 많은 도시에서 BMS를 구축하여 이용하고 있다.

버스의 도착 시간, 배차 간격, 노선 구성에 대한 정보를 언제, 어디서나 시민들이 이용 할 수 있도록 BMS를 구축하고 있다.

버스는 지하철과는 달리 도로를 이용하기 때문에 운행시간과 배차 간격을 사전에 정하고 정확한 시간에 서비스를 할 수 없다. 도로의 교통량, 신호체계, 운전기사의 성향 등 정확하게 예측할 수 없는 변수들이 많기 때문이다.

따라서 가장 정확한 방법으로 버스의 운행시간을 예측하고, 이것을 토대로 해당 버스의 정류장 도착시간을 산출하는 것이 BMS의 핵심이 된다.

[표 1] 과 같이 현재의 BMS에서 사용되고 있는 버스도착시간 예측 알고리즘은 시계열모형을 지닌 이동평균법과 칼만필터링을 사용하고 있다.

이러한 방법은 예측이 쉽고 구현이 용이한 장점이 있지만 도로의 급작스러운 통행량변화에 적용할 수 없으며, 가까운 과거

의 데이터로 예측을 하기 때문에 정확한 예측이 될 수 없다.

이 두 가지의 방법은 ‘가까운 과거에 X분의 시간이 소요되었기 때문에 현재도 통행시간이 X분이 될 것이다.’ 라는 가정을 하고 있다.

하지만 이는 추측에 불과하며 논리적 연관성이 없다. 따라서 본 논문에서는 ‘현재 도로의 통행량이 Y이기 때문에 Y에 의해서 버스의 통행시간이 f(Y) 될 것이다.’ 라고 제안하였다.

[표 1] 기존 BMS의 버스도착시간 예측 알고리즘

구분	개념	국내 도입 지역
이동평균법	이전시간동안 해당구간의 차량통행시간을 가중 이동평균을 적용	안양시 대전시 서울시 (종로구)
칼만필터링	특정구간 주행시간 시계열 데이터의 통계처리를 통해 미래의 주행시간을 예측	부천시 울산시

1) 이 논문은 서울시 산학연 협력사업(SJ080658)의 지원으로 수행하는 유비쿼터스 버스 관리시스템 개발(서울시 기술기반 구축사업)의 일환으로 수행되었습니다.

## 2. 관련연구

본 연구와 관련하여 많은 연구가 이루어 졌으며, 앞서 서론에서 밝혔듯이 버스도착시간을 예측하기 위한 알고리즘은 크게 이동 평균법과 칼만필터링 등의 시계열 모형과 본 논문에서 제안하는 방식인 Pattern을 바탕으로 하는 신경망모형으로 나눌 수 있다.

시계열 모형을 이용하는 경우 가까운 과거자료로 현재 버스의 도착예정시간을 예측하는 방법을 사용하였으며, Van Arem B에 의해 ARMA모형(이동평균혼합모형)을 적용한 연구([1])가 이루어 졌으며, 신호교차로를 감안한 실시간 버스도착시간 예측기법 개발에 관한 연구 (강태구)([2])에서는 시계열 모형에 교통신호의 주기를 버스통행시간 예측에 이용하였다.

이에 더불어 Pattern Recognition, 신경망을 활용한 방법도 많이 연구가 되었으며, Dynamic Bus Arrival Time Prediction with Artificial Neural Networks(Steven I)[3]에서는 버스구간을 링크로 나누어 각 링크의 운행시간을 신경망방법을 이용해 운행시간을 예측하였다.

시계열 모형은 가까운 과거와 가까운 미래의 시계열 적인 특성이 같아야 한다는 조건이 따르며, Pattern Recognition과 신경망의 경우에는 모델링의 어려움과 정확한 결과를 위해서는 많은 양의 데이터가 필요하다는 단점이 있다.

## 3. 연구방법

임의로 구성된 도로에서 버스를 포함한 교통량이 랜덤하게 유입되는 상황을 시뮬레이션 하였으며 시간대 별, 도로 별 유입되는 교통량을 달리 하였다.

그리고 각 차량은 교통량에 따라 속도를 다르게 하여 도로를 통과하도록 하였으며 교통량외의 모든 상황(신호체계, 사고 등)은 모두 무시 하였다.

t시간에 i도로의 실제 통행시간  $T_{t,i}$ 는 t시간에 도로의 시작점에 있는 차량이 도로의 끝지점에 도달하는데 걸린 실제시간을 측정함 값이며 t시간에 각 도로의 교통량  $L_{i,t}$ 와 함께 데이터베이스에 저장하였다.

t의 주기는 1초로 하였으며 매 초마다 각 End Point에서 랜덤하게 교통량을 발생시켰다.

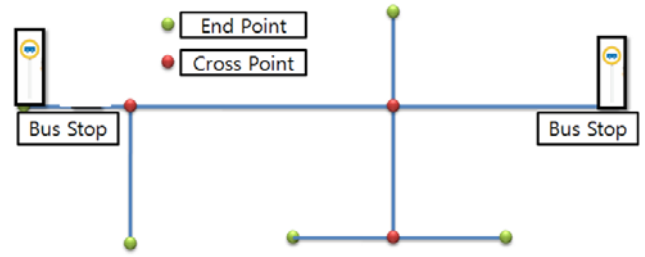
$$T_{t,i} = t\text{시간에 } i\text{도로의 실제 통행시간}$$

$$L_{i,t} = t\text{시간에 } i\text{도로의 교통량}$$

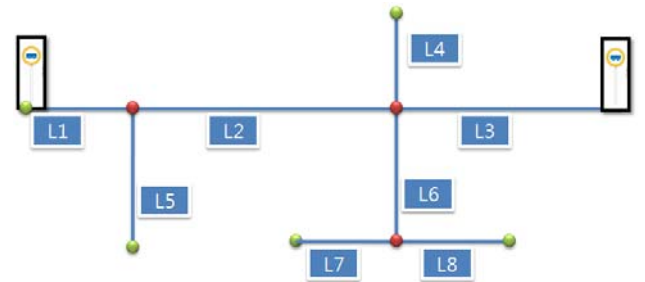
### 3.1 시뮬레이션 Tool

Matlab 을 통해 시뮬레이션 환경을 구축하였으며 My-SQL 에 데이터를 저장하였다.

### 3.2 시뮬레이션 설계 및 학습



[그림 1] 시뮬레이션 환경 I



[그림 2] 시뮬레이션 환경 II

버스 통행구간은 정류장과 정류장 사이인 L1~L3가 되며 교통량은 End Point에서 시간에 따라 Random하게 발생하고, 각 Cross Point에서 일정 확률로 주변 도로로 분산되도록 설계하였다.

또 L1~L3 사이의 도로를 운행하는 버스의 통행시간을 데이터베이스에 저장하였으며 통행량 중 버스의 비율을 1/10로 하였다.

그리고 각 End Point( $E_i$ ) 마다 고유 값 범위 내에서 통행량이 발생하도록 하였으며, Cross Point 에서도 각 도로로 분산되는 고유값을 가지도록 하였다.

또한, 예측은  $E_{E_i}$ ,  $E_{L_{i,k}}$  값을 모른다는 가정 하에 수행하였다.

$$E_i = E_{E_i} + \text{Rand}(-E_{E_i}/2, E_{E_i}/2)$$

$$L_{i,t} = L_{i,t-1} - \sum(L_{i,t-1} * E_{L_{i,k}}) + E_k + \sum(L_{k,t-1} * E_{L_{i,k}})$$

$$E_k = L_i \text{의 Endpoint}$$

$$L_{k,t} = t\text{시간의 } L_i \text{와 인접한 도로 } L_k \text{의 통행량}$$

$$E_{E_i} = \text{End Point } E_i \text{의 고유 통행량}$$

$$E_{L_{i,k}} = \text{차량이 } L_k \text{ 도로에서 } L_i \text{ 도로로 이동할 고유 확률}$$

위의 변수들에 의해 매초마다  $L_{i,t}$ 를 계산하고 버스통행시간  $T_{t,i}$  ( $T_{t,1} + T_{t,2} + T_{t,3}$ )와 함께 데이터베이스에 저장하였다.

또 교통량이 3만대, 10만대, 30만대, 100만대, 300만대가 될 때 까지 단계별로 저장하였다.

### 3.3 Pattern에 따른 통행시간예측 모델링

시뮬레이션을 통한 데이터를 분석하여 통행량에 따른 버스통행 시간을 모델링 하였다.

먼저 도로 L1~L8의 통행량에 따라 버스통행시간 예측을 모델링 했으며 8개의 도로마다 가중치  $F_i$ 를 구하였다.

$$F_i = \frac{\frac{d \sum_{k=1}^3 T_{k,t}}{dt}}{\frac{dL_{i,t}}{dt}}$$

if (Fi<0) Fi = Fi  
else Fi = 1/Fi

버스구간 L1~L3의 통행시간의 합인 시간에 대한 변화량을 Li 도로의 시간에 대한 통행량 변화량으로 나누어 Li 도로의 통행량과 버스구간의 통행시간의 상관관계를 구하였다.

Li도로의 통행량의 변화량과 버스구간 통행시간의 합의 변화량이 비슷할 수록 1에 가까운 수가 되어 가중치가 커지게 된다.

그리고 각 도로의 통행량에 따른 버스 통행시간 TTi,k(i 도로의 통행량이 s일 때의 버스 통행시간)는

$$TT_{i,k} = \frac{\sum T_{i,n}}{s(n)}$$

n = i도로의 통행량이 k와 비슷한 값일 때의 시간  
s(n) = 시간 n의 총 개수

로 구하였다.

즉, 통행량이 k일 때 각 8개의 도로마다 과거에 k와 비슷한 통행량을 보일 때의 통행시간 패턴을 참고하였다.

\* 실험에서는 통행량 k와 가장 비슷한 순서대로 5개의 n을 선택했다. 따라서 s(n)=5 이다.

최종목표인 버스통행예측 시간 Rt(t시간의 통행예측시간)는

$$R_t = \frac{\sum_{k=1}^8 TT_{k,s} * F_k}{\sum_{k=1}^8 F_k}$$

가 된다.

#### 4. 결과분석

학습량 별로 데이터베이스에 저장을 하고, Fi를 구하였으며 각 학습량 별로 도로의 교통량을 달리하여 총 100번의 예측을 하였고, 100번의 예측마다 실제 시뮬레이션을 한 실측치결과와 비교하여 오차를 계산해 보았다.

또 이동평균법을 이용해 같은 100번의 상황에서 예측을 하고 실측치와 비교한 오차를 구하였다.

##### 4.1 학습 교통량 별 결과

[표 2] 교통량 별 패턴인식에 의한 예측과 실측치 대비 오차결과 (실측치-예측치)

학습교통량 (대)	오차평균 (초)	오차분산
3만	8.10	30.5
10만	6.19	18.7
30만	6.05	23.3
50만	5.91	20.7
100만	5.96	19.2
300만	6.03	18.1

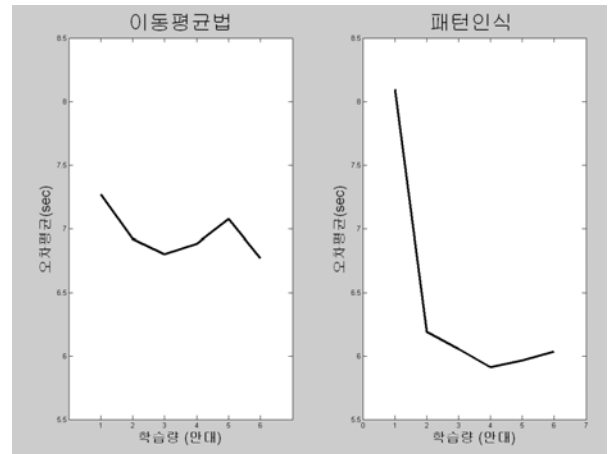
학습교통량이 많아 질수록 정확도가 높아졌다.

특히, 3만대에서 10만대로 학습량을 늘렸을 때 정확도의 개선이 뚜렷했으며, 그 이상의 학습량에서는 미세하게 개선이 되었으며 오차평균은 6초로 수렴하는 특성을 보였다. 실측치 평균은 151.2초로 측정되었으며 오차가 6초일 경우 정확도는 96.03% 가 된다.

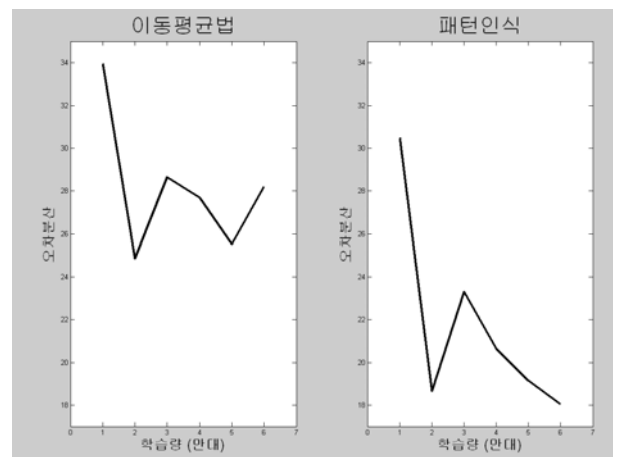
##### 4.2 이동평균법과 비교

본 논문에서 제안하는 방법의 우수성을 증명하기 위해 이동평균법에 의한 예측과 비교하였다.

이동평균법은 예측시점 이전의 5번 실측치를 기준으로 측정하였다.



[그림 3] 학습량별 오차평균 변화



[그림 4] 학습량별 오차분산 변화

먼저, 학습량에 따른 오차의 변화를 살펴보면 [그림3],[그림4]와 같이 이동평균에 의한 예측은 학습량의 변화에 영향을 받지 않았지만 본 논문에서 제안하는 패턴인식에 의한 예측방법은 학습량이 많아 질수록 오차가 줄어드는 것을 확인 할 수 있다.

[표 3] 교통량 별 이동평균법의 실측치 대비 오차결과

학습교통량 (대)	오차평균 (초)	오차분산
3만	7.27	34.0
10만	6.92	24.9
30만	6.80	28.6
50만	6.88	27.7
100만	7.08	25.5
300만	6.77	28.2

[표 4] 이동평균법[표3]과 패턴인식[표2]의 비교 (표[2]-표[3])

학습교통량 (대)	오차평균 (초)	오차분산
3만	-0.83	3.5
10만	0.73	6.2
30만	0.75	5.3
50만	0.97	7
100만	1.12	6.3
300만	0.74	10.1

정확도에 있어서 [표 4]에서 확인 할 수 있듯이 학습교통량이 3만대 이상에서는 모두 패턴인식에 의한 방법이 오차가 적었다.

또 오차의 분산에서도 개선정도가 학습량이 많아질수록 뚜렷하게 나타났다.

### 5. 결론 및 향후 연구

본 논문에서는 도로의 통행량에 의거하여 버스의 정류장 도착시간을 예측하였다.

그리고 예측의 정확도는 4.1절에서와 같이 96.03%였으며 이동평균법의 정확도보다 우수하게 평가되었다.

하지만 실제 버스의 통행에 있어 도로의 교통량만으로 정확한 예측을 할 수는 없다. 도로의 신호체계, 차량의 돌발 사고, 운전자의 성향 등도 고려해야 할 사항들이다.

따라서 본 연구를 바탕으로, 버스의 통행에 영향을 주는 위의 요소들을 포함하여 패턴인식을 통해 학습·예측할 수 있는 알고리즘을 연구하도록 하겠다.

### 참고문헌

- [1] Van Arem B, "Travel time estimation in the GERDIEN project", International Journal of Forecasting 13 (1997) 73-85, 3월, 1997.
- [2] 강태구, "신호교차로를 감안한 실시간 버스도착시간 예측 기법 개발에 관한 연구", 경기대 대학원 석사 학위논문, 2004.
- [3] Steven I, "Dynamic Bus Arrival Time Prediction with Artificial Neural Networks", Journal of Transportation Engineering, 9월, 2002.
- [4] 이영호, "기회손실비용을 고려한 버스 운행시격과 링크 통행시간 예측 알고리즘", 대한교통학회지 제 18권 제3호, 6월, 2000.
- [5] 김지홍, "교통수요 기반의 도착예정시간 산출 알고리즘 개발", 대한교통학회지 제23권 제 2호, 4월, 2005.