

은닉 마르코프 모델 기반의 교통량 예측 기법 연구

김민재[○], 윤희용^{*}

[○]서원대학교 컴퓨터교육과

^{*}성균관대학교 정보통신대학

e-mail: kmj01300@hanmail.net[○], youn@ece.skku.ac.kr^{*}

A Study of Traffic Prediction Method Based on Hidden Markov Model

Min-Jae Kim[○], Hee-Young Youn^{*}

[○]Dept. of Computer Education, Seowon University

^{*}College of Information and Communication Engineering, Sungkyunkwan University

● 요약 ●

최근 급증하는 교통 혼잡으로 인해 시간적/물질적 손실이 크게 발생하고 있다. 이러한 교통난 해소는 시설투자만으로는 근본적인 해결책이 될 수 없다는 판단 하에 지난 수년간 보다 정확한 교통량을 예측하기 위해 시계열 기반의 다양한 교통량 예측 모델들이 개발 되어 왔다. 그러나 시계열 기반의 모델들은 회귀분석을 통해 과거 교통량을 분석하고 과거의 교통패턴이 미래에도 지속적으로 연장된다는 가정 하에 연구되었기 때문에 실시간으로 급변하는 불규칙한 교통 패턴에 대한 예측의 신뢰성을 떨어트린다. 또한 시계열 기반의 예측 기법은 어떠한 회귀분석 모델을 사용하는지에 따라 성능의 차이가 많이 나타나기 때문에 회귀분석 모델 선택이 중요하다. 이러한 제약을 극복하기 위해 본 논문에서는 은닉 마르코프 모델(Hidden Markov model)을 이용해 동적인 교통 패턴에 따라 현재 상황에 맞는 회귀분석 모델을 선택하는 신뢰도 높은 교통량 예측 시스템을 제안한다.

키워드: 회귀분석(regression), 은닉 마르코프 모델(Hidden Markov Model)

I. 서론

최근 급증하는 교통 혼잡으로 인해 시간적/물질적 손실이 크게 발생하고 있다. 이를 위해 지난 수년간 보다 정확한 교통량을 예측하기 위해 다양한 교통량 예측 모델들이 개발되어 왔다. 그러나 기존 분석법은 교통패턴이 미래에도 지속된다는 가정 하에 이루어지고 있기 때문에 실시간으로 급변하는 불규칙한 교통 패턴에 적용하기에는 신뢰성이 많이 떨어진다. 따라서 본 논문에서는 기존의 시계열 기반의 교통량 예측 모델링에서 발생하는 실시간 상황에 대한 부정확성 문제를 보완하기 위해 은닉 마르코프 모델을 적용한 교통량 예측 시스템을 제안한다. 제안하는 은닉 마르코프 모델기반의 교통량 예측 시스템은 기존에 수집된 관측치와 실시간으로 수집한 관측치의 값에 따라 전이 행렬과 관측행렬을 가지게 되고 이를 통해 관측치들의 동적 패턴을 분석해 현재 상황에 가장 적합한 회귀분석 모델을 결정하게 된다.

II. 관련 연구

회귀분석은 변수들의 상관관계뿐만 아니라 하나 혹은 그 이상의

변수가 다른 변수에 미치는 영향의 정도를 파악하는데 그 목적이 있다. 따라서 회귀분석은 원인 혹은 결과의 인과관계를 이용하여 하나 혹은 그 이상의 변수들의 변화에 따른 다른 변수들의 값을 예측하고자 할 때 유용하게 쓰인다. 회귀분석은 설명 또는 예측하고자 하는 종속변수가 무엇인가를 결정해야하고 종속변수의 변화를 설명 혹은 예측하는데 이용한 설명 변수를 선택해야한다. 독립변수의 선택은 기존의 이론이나 경험적 관계에 의해 결정될 수 있으며 이들 종속 변수와 독립변수와의 관계를 구체적 함수관계로 표현하고 이를 근거로 설명 변수의 값으로부터 종속변수의 값을 예측하기 위해 사용된다.

III. 본론

본 연구에서는 수집된 교통정보와 교통량에 영향을 준다고 판단되는 요소를 은닉 마르코프 모델에 적용하여 실시간 교통 상황에 맞춰 적합한 회귀분석을 선택할 수 있는 교통량 예측 모델을 구현하였다. 기존의 시계열 기반 모델이 통계 기반의 교통량 정보를 반영했다면 본 연구에서는 구간에 교통량을 바탕으로 도로 구간 내에 있는 각종 환경요소들을 선정하였고 요일과 시간 및 강수 유무에 따라

변화하는 도로 상황을 반영하였다.

1. 회귀분석 모델 선정

본 논문에서는 실시간 교통상황 예측에 적용시킬 수 있는 회귀분석 선정을 위해 선정된 예측변수를 바탕으로 각 회귀 분석 모델의 적합도를 식1을 통해 판정 하였다.

$$S_e = \sqrt{\frac{\sum(Y_i - \hat{Y}_i)^2}{n - k - 1}} = \sqrt{\frac{\sum e_i^2}{n - k - 1}} \quad [식 1]$$

선정한 회귀분석 모델은 1차, 2차, 3차 함수의 모델로서 그림 1과 같다. 이 모델들은 은닉 마르코프 모델에 적용 되어 상태열로 활용되며, 이후 학습된 관측치와 실시간으로 수집한 관측치의 값에 따라 전이 행렬과 관측행렬을 가지게 된다. 이를 통해 관측치들의 동적 패턴을 분석하게 되고 선정된 3개의 모델 중 현재 상황에 가장 적합한 회귀분석 모델을 사용하게 된다.

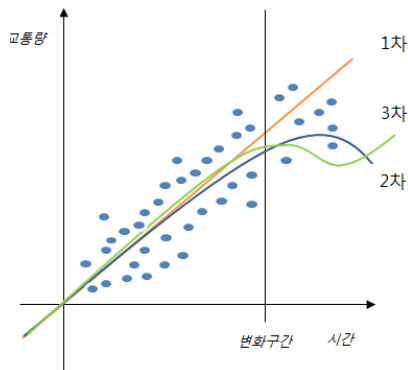


그림 1. 선정된 회귀분석 모델
Fig. 1. Regression model

2. 은닉 마르코프 모델 적용

은닉마르코프 모델 적용 단계는 입력된 기존의 관측치열과 현재 관측열을 통해 동적인 관측열의 패턴을 분석하여 15분 단위로 회귀 분석 모델을 변경시켜준다. 모델의 관측 입력으로 각 데이터의 상태 열은 식1에 의해 선정된 회귀분석 모델이 사용된다.

각각의 훈련된 은닉 마르코프 모델이 학습된 후, 주어진 교통 상황 관측열을 생성할 확률이 가장 높은 회귀모델을 선택함으로써

관측열의 소속 클래스를 결정하게 되며, 수학적으로는 식2 와 같이 나타낸다.

$$\max_i P(X|H_i) \quad [식 2]$$

본 논문에서는 이 과정을 수행하기 위해 동적 프로그래밍 기법중 하나인 viterbi알고리즘을 사용한다. 이를 통해 15분단위로 현재 교통상황에 적합한 회귀분석 모델을 선택함으로써 교통량 예측의 정확도를 향상시킬 수 있다.

IV. 결론

본 연구는 시계열 기반의 교통량 예측 모델링에서 발생하는 실시간 상황에 대한 부정확성 문제를 보완하기 위해 은닉 마르코프 모델을 적용한 교통량 예측 모델을 제안하였다. 이를 위해 교통상황에 영향을 줄 수 있는 예측 변수를 선정하여 전처리 과정을 거쳤고 이러한 변수를 선정하여 전처리 과정을 거쳤고 이러한 변수를 바탕으로 현재 교통 상황에 적합한 회귀모델을 선택함으로써 예측의 정확도를 향상 시켰다.

ACKNOWLEDGEMENT

본 연구는 한국산학연합회(C0017380), BK21+사업, 한국연구재단 기초 연구 사업 (2013R1A1A2040257), (2013R1A1A2060398), 미래부가 지원 한 2013 년 정보통신-방송(ICT) 연구개발 사업 (1391105003)의 지원을 받아 수행되었음.

참고문헌

- [1] Ching, Wai Ki, and Michael K. Ng. Markov chains: Models, algorithms and applications. Springer, 2006.
- [2] Hastie, Trevor, Robert Tibshirani, and Jerome Friedman. Linear Methods for Regression. Springer, 2009.
- [3] Chen, Zhe, and Robert C. Qiu. "Prediction of channel state for cognitive radio using higher-order hidden Markov model." IEEE SoutheastCon 2010, Proceedings of the. IEEE, 2010.