

# MPEG VBR 트래픽을 위한 GOPARIMA 기반 대역폭 예측 기법

## GOPARIMA based Bandwidth Prediction for Non-stationary VBR Traffic

강성주\*      원유집\*\*  
Sungjoo Kang      Youjip Won

**Abstract** - In this work, we develop on-line traffic prediction algorithm for real-time VBR traffic. There are a number of important issues: (i) The traffic prediction algorithm should exploit the stochastic characteristics of the underlying traffic and (ii) it should quickly adapt to structural changes in underlying traffic. GOP ARIMA model effectively addresses this issues and it is used as basis in our bandwidth prediction. Our prediction model deploy

Kalman filter to incorporate the prediction error for the next prediction round. We examine the performance of GOP ARIMA based prediction with linear prediction with LMS and double exponential smoothing. The proposed prediction algorithm exhibits superior performance against the rest.

**Key Words** : VBR 트래픽, 트래픽 모델링, 트래픽 예측, GOP ARIMA, 칼만 필터

### 1. 서론

네트워크를 이용해 스트리밍 서비스를 하는 경우에 네트워크의 상황에 따라 발생하는 지연에 의해 패킷 손실이 발생하게 되고 이는 사용자 입장에서의 QoS를 떨어뜨리게 된다. 이는 네트워크 상의 여러 라우터로 MPEG VBR 트래픽 같이 버스티(bursty)한 특성의 패킷이 들어올 때 이들 패킷을 처리하고 전달하는 시간이 길어지기 때문에 발생한다.

트래픽 스무딩[1] 등의 기법은 자원 효율을 극대화하고 멀티미디어 서비스의 품질을 극대화 할 수 있다. 저장된 비디오 데이터의 스트리밍 서비스 시 트래픽 스무딩을 위한 트래픽 정보를 알 수 있으므로 트래픽 예측은 무의미하나, 실시간 스트리밍 시에는 미래 트래픽의 정보를 알 수 없으므로 이를 위한 트래픽의 동적 예측이 필요하다.

이를 위한 트래픽 모델링과 트래픽 예측에 관한 연구가 진행되어왔다. Leland[2] 등은 VBR 트래픽의 장기 의존성(LRD, long-range dependency) 특성을 관찰하였고, Beran[3] 등은 그를 증명했다. 또한 VBR의 stationary 특성에 대한 연구도 진행되어, 최근의 연구 결과는 실제 트래픽이 Non-stationary 특성을 가짐을 보였다[4].

이러한 특성을 만족하는 통계적 회귀 모델로 ARIMA (Auto Regressive Integrated Moving Average) 모델이 주목 받았으며, GOP ARIMA 모델 등이 제시되었다[5].

동적 대역폭 할당을 위한 VBR 트래픽 예측에 관한 연구도 활발하다. 초기의 주파수, 시간, 웨이블릿 영역에서

트래픽을 분석하여 예측하는 방법은 계산 복잡도와 장면 전환의 부적응 문제가 있었으며, 퍼지 논리, 신경망 등의 예측 퍼지 논리, 신경망 기반의 예측 방법 등의 지능형 시스템 응용 방법은 장기간 후의 트래픽을 예측 할 경우 큰 오차를 발생시키는 문제를 보였다[6].

본 논문에서는 실시간, 장기간 예측이 가능하고, 장면 전환 적응도 또한 높은 예측 기법을 통해 사용자의 QoS를 극대화시킬 수 있음을 확인하고자 한다.

### 2. VBR 비디오 트래픽의 통계적 모델링

#### 2.1 VBR 비디오 트래픽의 통계적 특성

MPEG VBR 코딩 방식은 연속된 프레임들 사이의 시간적, 공간적 차이를 이용한다. MPEG 압축기는 세 가지 다른 종류의 압축된 프레임, 한 장의 화면을 구성할 수 있는 I (Intra)프레임, 이전 I, P 프레임과 현재 압축할 영상의 차이를 저장하는 P (Predictive)프레임 과 B (Bidirectionally-predictive)프레임으로 압축한다. 이와 같은 압축 방식은 압축률을 높여 디스크의 사용 효율을 높일 수 있으나, 그림1과 같이 I,P,B 프레임의 크기 차이가 생긴다.

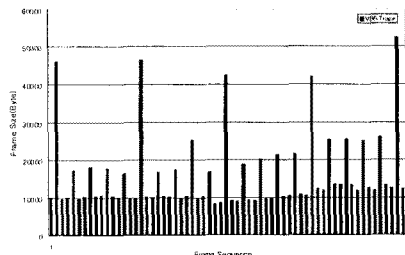


그림1. MPEG VBR 트래픽

저자 소개

\* 姜 聖 柱 : 漢陽大學 電子通信電波工學科 碩士課程

\*\* 元 裕 集 : 漢陽大學 電子電氣學科 助教授

이러한 MPEG VBR 트래픽을 통해 여러 가지 통계적 특성을 추출해 낼 수 있으며, 이를 통한 VBR 트래픽의 특성화가 가능하다. 첫 째, Non-Stationarity 특성이다. 그림1의 VBR 트래픽의 시간에 대한 프레임 크기의 변화에서 프레임간에 주기성이 존재함을 알 수 있다. 매 15 프레임마다 큰 값을 가지는 I 프레임, 3 프레임 간격으로 중간 크기를 갖는 P 프레임, I, P 프레임 사이에 존재하는 작은 값들은 B 프레임을 나타낸다. 임의의 프로세스가 엄격하게 시간 의존적인 행태를 보이며, 그것에 대한 분명한 이유가 있을 때 Non-stationary 특성을 갖는다. VBR 트래픽은 그림1에서 볼 수 있듯이 시간에 의존적이고 결정적인 형태를 보이며, 그 이유가 VBR 트래픽의 고정된 GOP 패턴 때문이란 것을 알 수 있다. 이는 곧 VBR 트래픽의 Non-stationary 특성을 말해준다. 또한 VBR 트래픽에는 LRD(Long Range Dependency) 특성이 있다. 임의의 프로세스  $X_t$ 는 Hurst 파라미터  $H$ 가,  $0.5 < H < 1$  일 때, LRD 특성을 갖는다고 말한다. 본 연구에서는 R/S 플롯 방법을 이용해 구한 VBR 트래픽의 Hurst 파라미터를 구해낸 결과를 통해  $H$  값이 0.89, 즉 LRD 특성이 존재함을 확인하였다.

## 2.2 VBR 트래픽의 GOP ARIMA 모델링

Non-Stationary 프로세스는 ARIMA(Auto Regressive Integrated Moving Average) 모델로 표현될 수 있다. 이 모델은 계절성을 지닌 시계열을 표현할 때 주로 사용된다. 임의의 프로세스를 계절성에 대해 차분한 결과가 ARMA 프로세스일 때, 이 프로세스를 ARIMA 프로세스라고 한다[7]. 이 ARIMA 모델을 사용해, VBR 트래픽의 non-stationary 특성 및 I 프레임 간, 그리고 P 프레임 간 거리의 주기를 가지며, 매우 천천히 감소하는 자기 상관 계수를 갖는 프로세스(VBR 프로세스)를 생성할 수 있다.

본 연구에서는 이와 같은 특성의 VBR 트래픽을 통계적으로 모델링한 GOP ARIMA(ARIMA for Group Of Picture)을 예측 대상 트래픽의 모델로 사용한다[5].

GOP ARIMA는 ARIMA 프로세스의 특별한 형태로 VBR 트래픽의 주기적인 패턴과 자기 상관 함수의 특성을 모델화하였다. 다음은 GOP ARIMA의 정의이다.

### 정의 GOP ARIMA(p,d,q), X(P,D,Q)<sub>s</sub> 프로세스

$s$ 와  $S$ 를 각각 P 및 I 프레임간 거리,  $d$ 와  $D$ 는 음이 아닌 정수라고 하면,  $\{X_t\}$ 를 차분한  $Y_t = (1-B)^d(1-B^S)^D X_t$ 가 ARMA(p,q)이면,  $\{X_t\}$ 를  $s, S$ 에 대한 GOP ARIMA(p,d,q), X(P,D,Q)<sub>s</sub>로 정의한다.

$$\phi(B)\Phi(B^S)Y_t = \alpha(B)\Theta(B^S)\epsilon_t$$

$$\phi(z) = 1 - \phi_1 z - \dots - \phi_p z^p, \Phi(z) = 1 - \phi_{S1} z^S - \dots - \phi_{SD} z^{SD}, \alpha(z) = 1 + \alpha_1 z + \dots + \alpha_q z^q,$$

$$\Theta(z) = 1 + \theta_1 z + \dots + \theta_q z^q \text{이며 } \{\epsilon_t\} \text{는 } N(0, \sigma^2) \text{ iid 난수시퀀스이다.}$$

### GOP ARIMA의 정의

## 3. VBR 비디오 트래픽의 동적 예측

### 3.1 칼만 필터링

칼만 필터는 이전 '상태변수'들을 이용해 1 시점 후의 '상태변수'(state)를 반복적으로 예측 하는 방법이다. 칼만 필터는 다음과 같은 두 방정식, '상태 방정식'과 '측정 방정식'으로 이루어진 '상태-공간 모델'을 이용한다.

$$x_{k+1} = Ax_k + V_k$$

$$z_k = Hx_k + W_k$$

'상태방정식'은 상태 벡터 A를 이용해 현재 '상태변수'의 1시점 후의 '상태'를 예측한다. '측정 방정식'은 측정 벡터 H를 이용해 '상태'를 '측정변수'의 공간으로 전이 시킨다.

칼만 필터는 최소 오차 공분산으로 미래 '상태변수'를 예측하는 칼만 이득 벡터(K, Kalman Gain vector)를 정의한다.

$$K_k = P_k^- H^T (HP_k^- H^T + R_k)^{-1}$$

t-1 시점에서 t 시점의 '상태'를 예측한 '이전 예측값'(priori estimate)은 t 시점에 이르렀을 때, 칼만 이득 벡터와 입력된 '측정변수'와의 오차를 이용해 얻은 보다 신뢰도 높은 '이후 예측값'(posteriori estimate)을 얻게 된다.

$$\hat{x}_k = \hat{x}_k^- + K_k (z_k - H_k \hat{x}_k^-)$$

그림 2는 칼만 필터 알고리즘의 동작 원리를 보여준다. (1) 초기 입력된 '상태변수'와 '오차 공분산'으로 칼만 이득 벡터를 계산하고, (2) 입력된 '측정변수'와 칼만 이득 벡터와의 연산을 통해 신뢰도 높은 예측값을 얻는다. (3) 또한 칼만 이득 벡터를 이용해 '오차 공분산'을 보정하고, (4) '상태 방정식'을 이용해 1단계 및 다음 식과 같이 N 단계 후의 '상태변수'를 예측한다.

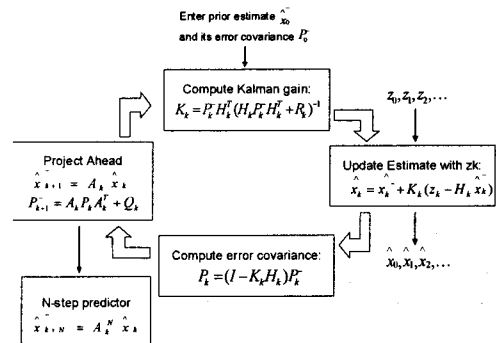


그림2. 칼만 필터 알고리즘

### 3.2 GOP ARIMA의 상태 공간 모델링

GOP ARIMA는 칼만 필터 알고리즘에 적용되기 위해 식(1)의 '상태 공간 모델링'과정을 필요로 한다. 본 연구에서 '상태변수'  $x_k$ 는 VBR 트래픽과 그의 차분 값의 조합으로, '측정변수'  $z_k$ 는 예측 대상인 VBR 트래픽을 사용한다. GOP ARIMA(1,1,1)<sub>3</sub> X (1,1,1)<sub>15</sub>는 차분 및 AR, MA 과정을 거쳐 다음과 같이  $X_t$ 과  $Y_t$ 의 회귀 모형으로 나타난다.

$$X_t = \theta \epsilon_{t-3} + \Theta \epsilon_{t-15} + \Theta \epsilon_{t-18} + \phi Y_{t-3} + \Phi Y_{t-15}$$

$$+ \phi \Phi Y_{t-18} + \epsilon_t + X_{t-3} + X_{t-15} + X_{t-18}$$

## 4. 실험

### 4.1 실험 환경

본 실험은 예측 알고리즘들의 성능을 평가한다. 제시된 칼만 필터 알고리즘 외에 성능 비교를 위해 선형 예측법(linear prediction)[7]과 이중지수평활 예측법(Double Exponential Smoothing based Prediction)[9]을 사용하였다. 콘텐츠 종류에 따른 통계적 특성 차이를 고려해 세 가지 종류의 트래픽(뉴스, 드라마, 스포츠)을 사용하였으며, 세 가지 트래픽은 평균, 최대, 최소 비트율이 각각 4Mbps/sec, 9Mbps/sec, 1.5Mbps/sec로 압축되었다. 또한 GOP의 크기는 15로 제작되었다 (GOP 패턴: IBBPBBPBBPBBPBB). 세 가지 예측 알고리즘의 성능 비교는 '정규화된 평균 제곱 오차(NMSE, normalized mean square error)를 사용하였다.

### 4.2 예측 결과

다음은 실험에 사용된 VBR 트래픽의 GOP ARIMA이다.

$$\begin{aligned} \text{Drama: } (1-B^1)X_t &= (1-0.6292B^1)\epsilon_t, \sigma^2 = (2030.6)^2 \\ \text{News: } (1-B^1)X_t &= (1-0.4081B^1)\epsilon_t, \sigma^2 = (2314.3)^2 \\ \text{Sports: } (1-0.514B^1-0.405B^2)X_t &= (1-0.9775B^1)\epsilon_t, \sigma^2 = (1795.5)^2 \end{aligned}$$

그림 3은 위 모델의 상태-공간 모델을 칼만 필터로 예측한 결과이다.

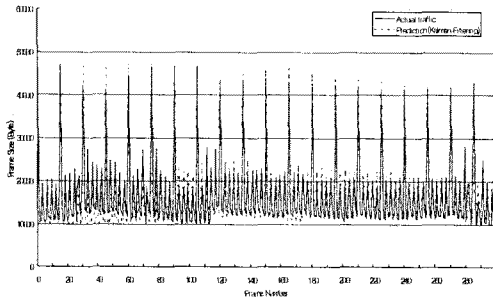


그림3. 칼만 필터를 이용한 예측

그림 4는 위의 세 가지 알고리즘으로 동일한 VBR 트래픽(Drama)을 예측 단계를 증가시키며 예측했을 때의 NMSE의 값을 비교한 것이다.

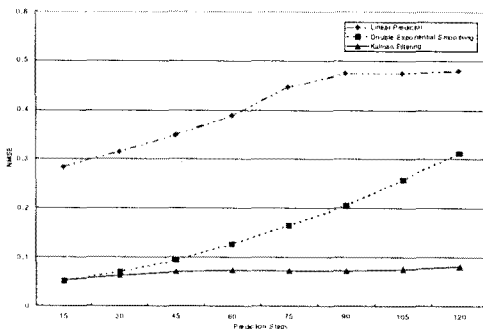


그림4. 세가지 알고리즘의 성능 비교  
예측 단계가 늘어날수록 예측 오차가 증가하는 것을 볼 수

있다. 이는 VBR 트래픽의 특성이 순간마다 변하기 때문이다. 또한 칼만 필터의 경우 오차의 증가율이 다른 알고리즘에 비해 작은 것도 확인할 수 있다. 이는 칼만 필터의 오차 보정 기능, 즉 매 시행마다 실시간 '측정 변수'의 입력으로 '상태 변수'를 보정해 나가면서 예측을 하기 때문이다.

## 5. 결론

본 논문에서는 MPEG으로 압축된 VBR 트래픽의 프레임 크기를 동적으로 예측하는 알고리즘을 제시하였다.

실험적인 VBR 프로세스의 통계적 특성(non-stationary)을 모델화한 GOP ARIMA 모델을 사용하였고, 예측 알고리즘으로 칼만 필터 알고리즘을 응용한 예측 방법을 사용하였다. 칼만 필터의 두 가지 변수, '상태변수'와 '측정변수'를 위해 GOP ARIMA의 상태-공간 모델을 제작하였고, '상태변수'로써 VBR 트래픽과 그 차분 트래픽의 조합을, 그리고 '측정변수'로써 VBR 트래픽을 사용하였다.

칼만 필터 알고리즘을 이용한 예측 방법의 성능 평가를 위해 두 가지 알고리즘(선형 예측 법, 이중 지수 평활 예측 법)을 이용한 결과와 비교하였고, 정량적인 비교를 위해 NMSE를 사용하였다. 칼만 필터 방법은 선형 예측 법에 비해 10%이내의 예측 오차를 보였으며, 다른 알고리즘에 비해 장기간 예측 시에도 작은 오차를 보임을 확인하였다.

## 참 고 문 헌

- [1] J. Rexford and D. Towsley, "Smoothing Variable Bit Rate Video in an Internetwork", IEEE/ACM Trans. on Networking, Vol. 7, Number 2, Apr. 1999
- [2] E. Leland, M. S. Taqqu, W. Willinger, D. V. Wilson, "On the self-similarity nature of Ethernet Traffic", IEEE/ACM Trans. on Networking, Vol. 2, pp. 1-15, Jan. 1994
- [3] J. Beran, R. Sherman, M. S. Taqqu, W. Willinger, "Long-range dependency in variable-bit-rate Video Traffic", IEEE Trans. on Communication, Vol. 43, pp. 1566-1579, Mar. 1995
- [4] Y. Qian, D. Tipper, D. Medhi, "A Nonstationary Analysis of Bandwidth Access Control Schemes for Heterogeneous Traffic in B-ISDN", Journal of VLSI Signal Processing, Vol. 29, pp. 101-113, 2001
- [5] Y. Won, S. Ahn, J. Jeon, "Performance Analysis of Non-stationary model for Empirical VBR Process", Proc. of Globecom 01, 2001
- [6] Y. Liang, "Real-time VBR Video Traffic Prediction for Dynamic Bandwidth Allocation", IEEE Trans. on Systems, Vol. 34, No. 1, pp. 32-47, Feb. 2004
- [7] Peter J. Brockwell, Richard A. Davis, "Introduction to Time Series and Forecasting", Springer
- [8] R. G. Brown, Patrick Y. C. Hwang, "Introduction to Random Signals and Applied Kalman Filtering with Matlab Exercises and Solutions", Wiley
- [9] Joseph Laviola Jr., "Double Exponential Smoothing: An alternative to Kalman Filter-based predictive tracking", The Eurographics Association, 2003