

예측 기반 동적 자원할당 구조를 이용한 실시간 VBR 비디오 트래픽의 QoS 제어 방법

정희원 유 상 조*, 홍 성 훈**, 강 성 원***, 홍 경 표****

QoS control for real-time VBR video traffic using prediction-based dynamic resource allocation architecture

Sang-Jo Yoo*, Sung-Hoon Hong**, Sungwon Kang***, Kyoung-Pyo Hong****

Regular Members

요 약

본 논문에서는 통신망을 통한 실시간 비디오 트래픽의 전송에 있어, 망 자원을 효율적으로 사용하고 사용자에게는 원하는 레벨의 QoS를 제공할 수 있는 새로운 예측기반 동적 자원할당 방법을 제안한다. 제안된 방법에서는 사전에 사용자로부터의 트래픽 정보 없이, 실시간 예측과 제공되어지는 QoS 정도를 관측하여 각 비디오 소스에 차별적으로 자원을 할당한다. 최대 패킷 지연과 패킷 손실율에 대한 실시간 QoS 관측을 위하여 간단한 구조의 지연 패킷 카운터 갱신과정을 제시하였으며, 현재까지 제공되어진 QoS 정도가 낮은 비디오가 우선적으로 자원을 할당 받을 수 있는 방법을 제안하였다. 모의실험을 통하여 제안된 방법은 높은 망 자원 이용율을 달성하는 동시에 기존의 방법이 제공할 수 없었던 연결별 QoS를 매우 정확히 제공할 수 있음을 보였다.

ABSTRACT

In this paper, we propose an efficient quality providing scheme to satisfy a diversify combination of delay bound and loss ratio requirements from users by a predictive dynamic resource allocation method for real-time video applications. For utilizing the network resources more efficiently while meeting the service requirements, we adjust the resources based on the predicted traffic and the currently provided quality level. We developed a simple delayed-packet counter updating scheme for real time QoS monitoring. Simulation results show that our proposed method can provide an accurate and flexible quality control.

1. 서 론

최근 초고속 인터넷 및 ATM과 같은 통신망의 발달과 다양한 통신망 응용 미디어의 출현으로 사용자는 음성, 데이터, 영상 등의 멀티미디어 서비스에 대해 신뢰성 있는 고품질의 서비스를 요구하고 있다. 특히 실시간 (real-time) 비디오는 향후 멀티미디어 서비스의 핵심 응용으로 그 중요성이 큰 반면, 매우 복잡한 트래픽 특성을 보이기 때문에 통신

망에서의 서비스 품질 (QoS: quality of service) 제 공을 위한 트래픽 제어가 매우 어렵다.

기존의 ATM과 같은 초고속 통신망에서 가변 비트 율 (VBR: variable bit rate) 비디오 트래픽에 대하여 QoS를 보장하기 위한 관리 구조에서는 연결 설정 시 미리 사용자에게 해당 비디오 소스의 트래픽 특성을 망에게 선언 (description) 할 것을 요구하고 이를 기초로 통신망에서는 호 수락 제어와 자원 할당을 수행한다^{[1][2]}. 그러나 인터넷 방송, 대화형 비

* 인하대학교 정보통신대학원 (sjyoo@inha.ac.kr),

*** 한국정보통신대학원대학교 공학부,

논문번호: 010251-0917, 접수일자: 2001년 9월 17일

** 전남대학교 정보통신공학과,

**** 한국통신 통신망연구소

디오 등의 실시간 비디오 응용에서 사용자가 사전에 비디오 트래픽의 특성을 미리 망에 알려 주는 것이 불가능하거나 어렵다는 문제점이 있으며, 망에서는 고정된 트래픽 기술자 조합을 이용하여 망의 자원을 할당함으로써 자원의 과 할당 (over-allocation)과 함께 망의 이용도 (utilization)가 낮아지게 된다.

동적 자원 할당 (dynamic resource allocation) 구조는 입력 트래픽을 관측하여 동적으로 소요 대역 (bandwidth) 및 버퍼를 할당하는 것으로 사용자에게 사전에 트래픽 정보를 요구하지 않으면서, 망의 이용도를 향상시킬 수 있으므로 많은 연구가 이루어지고 있다. 기존의 동적 자원 할당은 일반적으로 선형 또는 비선형 필터를 사용하여 향후 트래픽을 예측하고 이를 기반으로 적용적으로 대역을 할당한다 [3][4][5][6]. VBR 비디오 트래픽의 실시간 on-line 예측을 위해 지금까지 몇 가지 대표적인 방법들이 제안되었다. [7]에서는 RLS (recursive least square)와 TDNN (time delay neural network) 방법을 이용하여 JPEG 및 MPEG 트래픽을 예측하였다. 다중화된 MPEG 비디오 스트림에 RLS 필터를 적용시킨 방법이 [8]에 제시었고 [9]에서는 LMS (least mean square)를 이용한 적응 선형 예측 방법이 제안되었다. 또한 LMS 방법의 느린 수렴속도를 향상시키기 위해 [5][10]에서는 적응 웨이블릿 (wavelet) 예측 방법이 사용되었다.

그러나 지금까지의 연구는 주로 트래픽 예측의 정확성을 높이기 위한 방법에 집중되었고, 동적인 대역할당을 위해서는 예측 트래픽과 현재 큐 길이를 기반으로 한 단순한 구조가 제안되었다. 그러나

이러한 기존의 구조로는 사용자에게 제공 가능한 QoS를 추정할 수 없으므로, 예측을 통한 자원의 효율적 사용과 함께 멀티미디어 트래픽에 적합한 동적이고 유연한 QoS 제어 메커니즘이 필요하다. 본 논문에서는 최대 패킷 지연과 패킷 손실율의 사용자 QoS 요구사항을 준수하면서 동적으로 대역을 할당할 수 있는 새로운 방법을 제안한다. 제안된 방법은 프레임 시간 단위의 QoS 관측을 통하여 각 연결에 제공되어진 QoS 정도에 따라 대역을 자동적으로 할당한다.

본 문의 구성은 다음과 같다. 제 2장에서는 제안된 예측 기반 동적 자원 할당 구조를 살펴본다. 제 3장에서는 QoS 파라미터의 정의 및 측정방법을 기술하고, 제 4장에서는 이를 기반으로 한 동적 자원 할당 방법을 제안한다. 제 5장에서는 모의실험을 수행하고 그 결과를 기존의 동적 자원할당 방법과 비교 분석한다. 마지막으로 제 6장에서 본 논문의 결론을 맺는다.

II. 시스템 구조

본 논문에서 제안하는 그림 1의 예측기반 동적 자원할당 시스템 구조는 통신망의 스위치나 라우터에 구현되어, 개별 비디오 연결에 서로 다른 QoS를 제공할 수 있다. 본 시스템에서 각 연결 별 버퍼의 크기는 최소한 스위치 당 지연한계 내에서 모든 입력 패킷들을 큐잉할 수 있을 정도로 크고, 비디오의 입력 패킷들은 같은 패킷 크기 (고정 길이 패킷)를 갖는다고 가정한다.

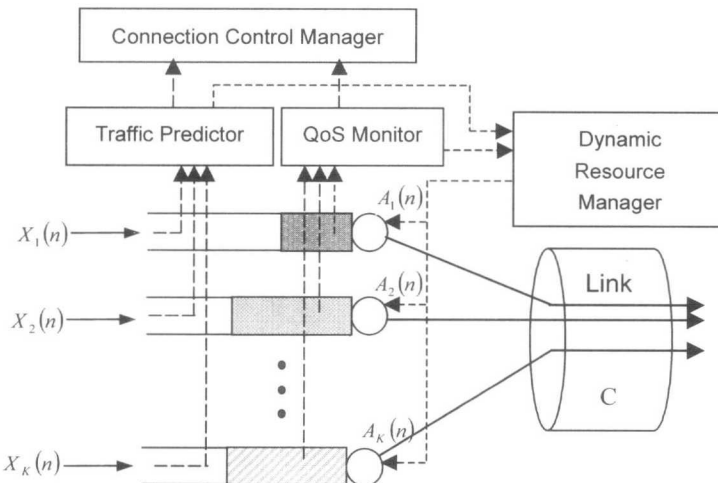


그림 1. 제안된 예측기반 동적 자원할당 시스템 구조

그림 1에서 트래픽의 예측, QoS 측정, 자원할당 등은 매 동일 시간 간격 마다 이루어 지며 본 논문에서는 이러한 시간간격으로 비디오의 한 프레임 시간 (초당 30 프레임의 경우 1/30초)을 사용한다. 제안된 시스템 구조에서는 다음과 같은 기능 블록이 정의된다.

- 트래픽 예측기 (traffic predictor)
각 연결 별 입력 트래픽을 측정하여 다음 시간 간격에 도착할 트래픽을 예측한다.
- QoS 관측기 (QoS monitor)
정의된 QoS 파라미터에 대해 각 연결 별로 서비스 되고 있는 QoS의 정도를 관측하고 사용자 QoS 요구사항에 얼마나 부합되는지를 판별한다.
- 동적 자원 관리기 (dynamic resource manager)
각 비디오 연결 별로 사용자가 요구한 QoS를 제공하기 위해 다음 시간 간격에 필요한 소요 대역을 계산하고 자원을 할당한다.
- 연결 제어 관리기 (connection control manager)
현재 비디오 소스들의 간단한 트래픽 특성과 제공되어지고 있는 QoS를 고려하여 새로운 비디오 소스의 연결 수락 여부를 결정한다.

트래픽의 예측은 비디오 트래픽 예측능성이 좋은 방법을 선택하여 사용할 수 있는데, 본 논문에서는 선형 예측 방법으로 구현이 쉽고 간단하며, 좋은 예측 성능을 보이는 LMS 방법[9]을 사용한다.

III. QoS 파라미터 정의 및 측정 방법

본 논문에서 제안하는 QoS 관리기법은 기존의 ATM 망에서와 같이 사전에 사용자로부터 선언된 트래픽 기술자를 이용하여 필요한 고정대역을 해당 연결에 할당함으로써 연결 기간 (connection duration) 동안 사용자가 요구한 QoS를 엄격하게 보장하는 것과는 달리, 사용자로부터의 트래픽 기술자 없이 관측에 의한 동적 자원할당을 이용한 flexible QoS를 제공한다. 본 논문에서 정의하는 flexible QoS는 연결 기간 중 입력 트래픽의 갑작스러운 증가 등의 이유로 통신망이 혼잡하게 되면 짧은 시간동안은 약속된 QoS를 제공하지 못할 수도 있지만, 긴 시간 동안 평균적으로 안정된 QoS를 제공하는 방식이다. 실시간 비디오 응용을 위하여 각 연결 i 에 대하여 통신망의 각 노드에서 제공하는 QoS 파라미터로 다음 두 가지를 정의하였다. 여기

서 각 파라미터 값은 실제 사용자로부터의 QoS 요구사항을 각 노드 단위로 분리한 것이다.

- i) 패킷 손실율 (packet loss ratio): $lossR_i$
- ii) 최대 패킷 지연 (maximum packet transfer delay): $dmax(i)$

실시간 비디오 통신에서 패킷 손실율과 최대 패킷 지연 범위는 매우 중요한 파라미터이다. 본 논문에서의 최대 패킷 지연은 비디오의 특성을 고려하여 비디오의 프레임 시간 단위로 정의된다. 즉 $dmax(i)=2$ 는 연결 i 의 최대 패킷 지연은 2 프레임 시간 (초당 30 프레임의 경우 2/30초) 이라는 것을 의미한다. 비디오 신호의 복호 (decompression)은 프레임 단위로 이루어 진다고 가정하면 패킷의 프레임 단위 시간 지연은 적절한 지연 시간 스케일 (time scale)로 생각되어 질 수 있다. 제안된 QoS 제어에서는 버퍼에서의 대기시간이 $dmax(i)$ 를 초과한 패킷은 폐기 (discarding)되며 이는 패킷손실로 이어진다.

실시간 비디오 전송에서 현재 제공되고 있는 QoS의 측정은 간단한 구현으로 이루어 져야 한다. 매 프레임 시간 마다 QoS 관측이 이루어 지며 제공되어진 QoS 값이 갱신된다. 첫번째 QoS 파라미터인 패킷 손실율의 측정은 다음 식에서와 같이 현재 n 프레임 시간 까지 연결 i 의 입력 패킷 수 중 손실 (폐기)된 패킷 수의 비로 계산된다.

$$lossM_i(n) = \frac{\sum_{m=1}^n L_i(m)}{\sum_{m=1}^n X_i(m)} \quad (1)$$

여기서 $X_i(m)$ 는 프레임 시간 m 에 연결 i 의 입력 패킷 수이고, $L_i(m)$ 은 이중 손실된 패킷 수이다.

제안된 구조에서, 임의의 패킷이 지연 한계치를 초과 했는가를 알아내는 것은 QoS 측정 및 자원할당을 위해 매우 중요하다. 버퍼에서의 지연시간 계산은 실시간 동작을 위해 간단하게 계산되어야 한다. 동적 대역할당 구조에서는 매 프레임 시간간격마다 자원 관리자에 의해 할당되는 대역폭이 다르므로, 입력된 패킷이 향후 언제 서비스 될지를 예측할 수 없다. 또한 매 패킷 마다 도착 시간과 출발 시간을 관리하고 이를 근거로 해당 패킷의 지연시간을 계산하는 것은 초고속 통신망 환경에서 실시간 구현을 어렵게 만드는 요인이 될 수 있다. 본 논문에서는 지연 프레임 단위의 간단한 카운터를 사용함으로써 버퍼에 프레임 시간 단위로 지연된 패

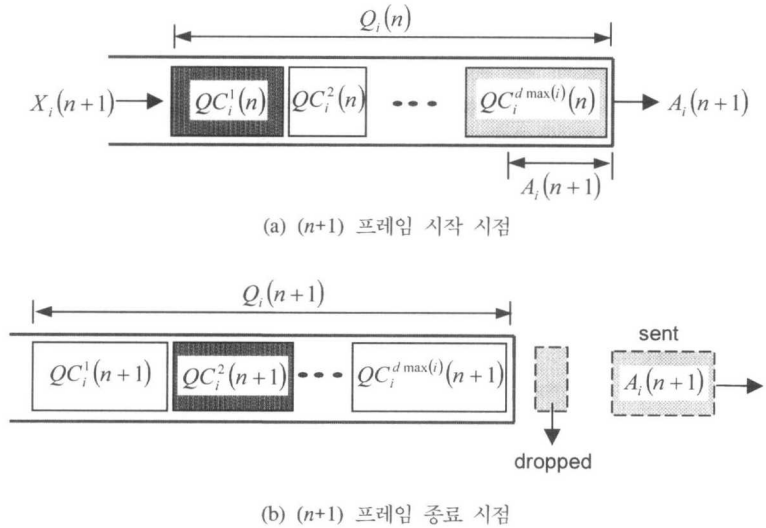


그림 2. 지연패킷카운터 갱신과정.

킷이 얼마나 존재하는 지를 알고, 이를 기초로 QoS 관측, 자원 할당, 패킷 폐기 등을 수행할 수 있는 알고리즘을 제안한다.

그림 2는 제안된 지연 패킷 카운터 (delayed-packet counter)의 운용 방법을 보여준다. 각 비디오 연결 i 는, $d_{max}(i)$ 개의 지연 패킷 카운터 $QC_i^d(n)$ 를 갖는다. $QC_i^d(n)$ 는 n 프레임 종료 시점에 현재 버퍼에 있는 패킷 중 $(d-1, d]$ 프레임 지연을 겪은 패킷 수를 나타낸다. 여기서 $(d-1, d]$ 는 $d-1$ 보다는 크고 d 보다는 작거나 같은 범위를 의미한다. $QC_i^{d_{max}(i)}(n)$ 개의 패킷 중 $(n-1)$ 프레임 시간 동안 서비스 되지 못한 패킷은 최대 지연 한계치 $d_{max}(i)$ 를 초과하게 되고 제안된 방법에서 이들 패킷은 버퍼에서 폐기된다. 앞에서 언급한 바와 같이 제안된 시스템 구조에서의 각 연결 별 버퍼는 모든 입력 패킷들을 최대 허용한계 내에서 큐잉 할 수 있을 정도로 크고, 패킷의 손실은 그림의 예제와 같이 버퍼의 앞 단에서 패킷이 최대 지연 한계 치 내에서 서비스 되

지 못할 때 일어난다. $(n+1)$ 프레임 시간 동안 폐기된 패킷 수 $L_i(n+1)$ 은 다음과 같이 구해진다.

$$L_i(n+1) = \max\{0, QC_i^{d_{max}(i)}(n) - A_i(n+1)\} \quad (2)$$

여기서 $A_i(n+1)$ 은 $(n+1)$ 프레임 시간에 연결 i 에 할당된 대역폭 (단위: 패킷 수) 이다.

그림 2의 예에서는 $QC_i^{d_{max}(i)}(n)$ 가 $A_i(n+1)$ 보다 크기 때문에 $(n+1)$ 종료시점에서 $QC_i^{d_{max}(i)}(n) - A_i(n+1)$ 개의 패킷은 폐기되었다. 매 프레임 시간의 종료시점에서 각 연결의 패킷 지연 카운터는 갱신된다. 제안된 지연 패킷 관측 방법에서는 매 입력 패킷 별로 지연시간을 계산하지 않고 해당 프레임 시간의 입력 패킷 수, 동적으로 할당된 대역폭 (출력 가능 패킷 수), 그리고 각 지연 패킷 카운터의 값 만을 이용하여 프레임 종료시점에서 지연시간 별 대기중인 패킷 수를 다음과 같이 계산할 수 있다.

여기서 $Q_i(n) = \sum_{k=1}^{d_{max}(i)} QC_i^k(n)$ 로 n 프레임 종료시

$$\begin{aligned}
 & d = d_{max}(i) \\
 & QC_i^d(n+1) = \max\{0, QC_i^{d-1}(n) - \max\{0, A_i(n+1) - QC_i^d(n)\}\} \\
 & d = d_{max}(i) - 1 \\
 & QC_i^d(n+1) = \max\{0, QC_i^{d-1}(n) - \max\{0, A_i(n+1) - QC_i^{d_{max}(i)}(n) - QC_i^d(n)\}\} \\
 & \quad \quad \quad \text{M} \qquad \qquad \qquad \text{M} \\
 & d = 2 \\
 & QC_i^d(n+1) = \max\{0, QC_i^{d-1}(n) - \max\{0, A_i(n+1) - QC_i^{d_{max}(i)}(n) - QC_i^{d_{max}(i)-1}(n) - \Lambda - QC_i^d(n)\}\} \\
 & d = 1 \\
 & QC_i^d(n+1) = \max\{0, \min[X_i(n+1), X_i(n+1) + Q_i(n) - A_i(n+1)]\}
 \end{aligned} \quad (3)$$

$$L_i(n+1) = \max\{0, QC_i^{d \max(i)}(n) - A_i(n+1)\} \quad \text{loss}M_i(n+1) = \sum_{m=1}^{m=n+1} L_i(m) / \sum_{m=1}^{m=n+1} X_i(m) \quad (4)$$

$$QC_i^d(n+1) = \begin{cases} \max\left\{0, QC_i^{d-1}(n) - \max\left[0, A_i(n+1) - \sum_{k=d}^{d \max(i)} QC_i^k(n)\right]\right\}, & d = 2, 3, \dots, d \max(i) \\ \max\{0, \min[X_i(n+1), X_i(n+1) + Q_i(n) - A_i(n+1)]\} & d = 1 \end{cases} \quad (5)$$

$$R_i(n+1) = \max\{QC_i^{d \max(i)}(n), (\hat{X}_i(n+1) + Q_i(n)) \times (1 - S_f \times PQD_i(n))\} \\ PQD_i(n) = \log\left(\frac{\text{loss}R_i}{\text{loss}M_i(n)}\right) \quad S_f = \begin{cases} 1 & \text{if } PQD_i(n) \leq 0 \\ \frac{1}{4} & \text{if } PQD_i(n) > 0 \end{cases} \quad (6)$$

점에서 연결 i 의 버퍼에 대기하고 있는 총 패킷 수이다. 일반화된 $(n+1)$ 종료시점에서 패킷 지연 카운터의 갱신과 제공된 QoS 측정은 다음 식과 같다.

IV. 동적 자원 할당

1. 연결별 소요대역 추정

그림 1의 시스템에서 동적 자원 관리기 (dynamic resource manager)는 각 연결들의 향후 입력 트래픽과, 버퍼에 대기중인 지연 패킷 수, 그리고 현재까지 제공된 QoS 정도에 따라 다음 프레임 시간에 각 연결에게 할당할 대역을 결정한다. 동적 자원 관리기는 그림 3과 같이 두개의 기능 블럭으로 구성된다. 소요 대역 추정기(required bandwidth estimator)는 각 연결 별로 향후 안정된 QoS를 지원하기 위해 필요한 대역을 추정하고, 공정 대역 분배기 (fair bandwidth distributor)는 현재 가용한 자원 한계 내에서 각 연결별로 공정한 자원분배를 수행한다.

제안된 소요대역 추정 방법에서는 현재 해당 비디오 연결에 제공된 QoS가 요구 레벨보다 과다하게 제공되는지 또는 부족하게 제공되었는지를 판단하여 동적으로 다음 프레임 시간에 필요한 소요대역 $R_i(n+1)$ 을 식 (6)과 같이 계산한다. $\hat{X}_i(n+1)$ 은 $(n+1)$ 프레임 시간 동안 연결 i 에 입력될 트래픽의

예측치이고, S_f 는 스케일링 factor, $PQD_i(n)$ 은 n 프레임 종료 시점까지 연결 i 에 제공되어진 QoS의 정도 (provided QoS degree)이다.

즉, 식 (6)의 소요대역 추정방법은 기본적으로 다음 프레임 시간 동안의 예측 트래픽과 현재 버퍼에 대기중인 패킷 수를 합한 값에 현재 제공된 QoS 정도 $PQD_i(n)$ 값에 따라 소요 대역폭을 증가시키거나 감소시킨다. 만약 요구된 QoS 보다 낮은 레벨의 품질이 제공되었다면 (즉, $PQD_i(n) < 0$) 다음 프레임 시간의 소요 대역폭은 예측 오류를 보상하고, 다음 절의 공정 대역 분배에서 더 많은 대역을 할당 받을 수 있도록 계산된다. 반대로 요구된 QoS 보다 높은 레벨의 품질이 제공되었다면 (즉, $PQD_i(n) > 0$) 소요대역은 기본 $(\hat{X}_i(n+1) + Q_i(n))$ 값보다 작게 계산된다. S_f 값을 $PQD_i(n)$ 값의 범위에 따라 다르게 설정한 것은 요청한 값보다 높게 QoS가 제공될 경우 큰폭의 소요대역 감소로 갑자기 QoS가 나빠지는 것을 방지하기 위한 것이다.

2. QoS 기반 공정 대역 분배 정책

공정 대역 분배기는 각 연결 마다 요청된 소요대역 추정치를 근거로 가능한 링크용량을 고려하여 연결별 QoS 제공 관점의 공정한 대역 분배를 수행한다. 각 연결별로 제공된 QoS 정도에 따라 두개의

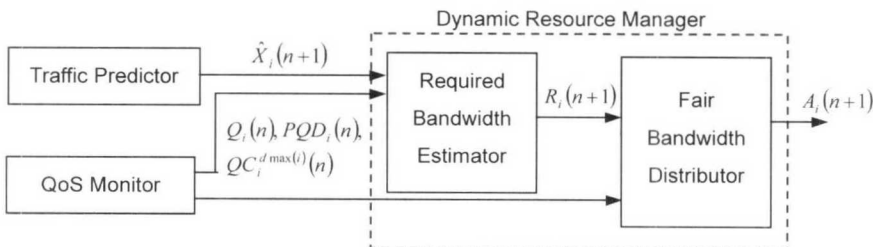


그림 3. 동적 자원 관리기의 기능 모듈.

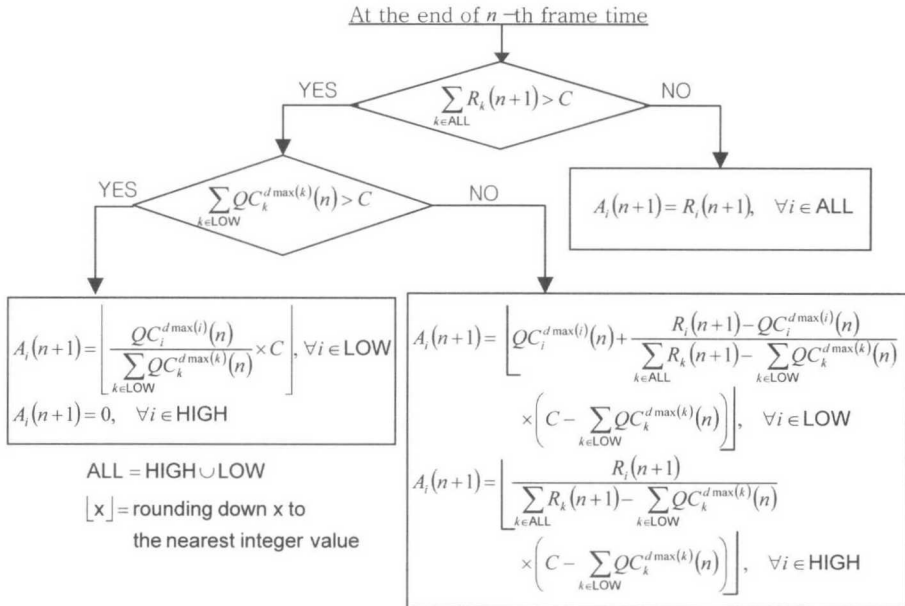


그림 4. 공정 대역 분배 절차.

클래스 집합을 정의한다. LOW 클래스는 $PQD_i(n)$ 값이 0 보다 작은 연결들의 집합이고, HIGH 클래스는 $PQD_i(n)$ 이 0보다 크거나 같은 연결들의 집합이다. 다음 그림 4는 제안된 공정 대역 분배 정책 절차를 설명한다. 만약 각 연결에 대하여 추정된 소요 대역의 합이 링크용량보다 작으면 요청된 소요 대역은 그대로 각 연결의 할당대역이 된다. 소요 대역의 합이 링크용량보다 클 경우 LOW에 속한 연결들의 버퍼에 있는 패킷 중 다음 프레임 시간에 서비스 되지 않으면 지연 요구사항을 위반하게 되는 패킷들을 위해 우선 대역을 분배하고, 나머지 가용 대역을 HIGH와 LOW에 속한 모든 연결들에게 소요대역 추정치를 기준으로 공정 분배 한다. 이러한 정책을 사용함으로써, 제공된 QoS가 낮은 연결들에게는 다음 프레임 시간 이후에는 요청된 QoS를 제공받을 수 있도록 하며, 이미 요청 값 이상을 제공받은 연결들은 다른 LOW 클래스의 연결들을 위해 자원을 덜 사용하게 된다.

다음 프레임 시간에 각 연결들을 위해 할당된 대역의 합이 전체 링크용량에 비하여 작을 경우 나머지 잉여 대역은 실시간 비디오를 제외한 다른 응용을 위하여 사용될 수 있다. 즉 기존의 인터넷 best-effort 서비스나 ATM의 UBR 서비스 클래스를 위해 잉여 대역이 사용된다. 이는 실제 통신망 환경에서 망의 이용도 (utilization)를 크게 향상시킨다.

본 논문의 V장에서 살펴 보겠지만 실시간 비디오 연결을 위해 실제 할당되는 대역은 거의 평균 입력 트래픽에 근사하게 된다.

3. 연결 수락 제어

제안된 flexible QoS 관리 구조는 사용자에게 사전에 트래픽 기술자를 요구하지 않으므로, 임의 순간에 비디오 소스가 갑자기 많은 트래픽을 전송하거나 새로운 비디오 소스가 네트워크에 연결 될 경우 각 사용자 비디오 소스에 제공되는 서비스 품질이 떨어질 수 있다. 새로운 비디오 소스의 추가로 인하여 기존의 비디오가 요청된 레벨 이하로의 QoS 저하를 경험하지 않게 하기 위하여 본 논문에서는 간단한 연결 수락제어 방법을 제안한다. 본 논문에서 제안하는 방법은 간단하면서도 확장성을 갖는다. 네트워크의 각 노드에서는 다음 식 (7)의 두 가지 조건이 만족되면 새로운 비디오 소스의 연결을 수락한다. 본 논문에서는 통신망의 노드간 연결 수락 제어에 필요한 신호 프로토콜이 존재한다는 것을 가정한다. 첫번째 조건은 기존의 K 개의 모든 연결들에 제공되어진 QoS가 모두 원하는 레벨 이상인지를 검사한다. 두 번째 조건은 기존 연결의 평균 입력 레이트의 합이 출력 링크용량 허용레벨 이하인가를 검사한다.

$$\begin{cases} \min\{PQD_i(n), i=1,2,\Lambda K\} \geq 0 \\ \frac{\sum_{i=1}^K Avg_i(n)}{C} \leq f(K) \quad \leftarrow Avg_i(n) = \sum_{m=1}^n X_i(m)/n \end{cases} \quad (7)$$

$f(K)$ 는 새로운 연결을 위한 여유 (marginal) factor로 통신망 운용자에 의해 적당한 K 의 함수로 표현 될 수도 있다. 본 논문에서는 다음과 같은 식을 사용하였다. 아래의 식에서 허용된 비디오 연결 수 K 가 크면 $f(K)$ 은 1에 근접한다. 측정기반 구조에서 적당한 $f(K)$ 함수를 구하는 방법은 계속 연구가 진행되어야 한다.

$$f(K) = \frac{K}{K+1} - 0.3 \exp\{-0.01 * (K-1)\} \quad (8)$$

V. 성능 분석

성능분석은 제안된 연결 수락제어 방법에 의해 비디오 연결이 서비스를 시작함에 따라, 각 비디오 소스의 QoS 요구사항이 어느 정도 잘 지켜지는가에 대해 이루어졌다. 모의실험을 위해 초당 25 프레임으로 미리 MPEG 부호화된 각 30 분 분량의 비디오 스트림[10]을 사용하였고, 트래픽 예측은 I, P, B 프레임 타입별로 LMS 예측 방법을 이용하였다. 비디오 비트 스트림은 고정길이 패킷으로 패킷화 되는데, 본 실험에서는 ATM 셀 크기와 같은 53 바이트의 패킷크기를 적용하였다. 동일한 연결 수락 조건에서의 성능 비교 대상으로 식 (9)에 의해 각 연결 별 할당 대역이 결정되는 기존 방법 [5][9][11]이 사용되었다. 기존 방법에서는 연결 수락제어가 제공되지 않으나, 제안된 방법과의 비교를 위하여 기존 방법의 실험에서도 각 소스의 수락 시점은 제안된 방법을 사용한 수락 시점과 일치 하도록 하였다. 기존의 방법을 통해 서비스된 패킷 중 지연 한계치를 초과한 패킷은 패킷 손실로 처리되었다.

$$A_i(n+1) = \min \left\{ R_i(n+1), \frac{R_i(n+1)}{\sum_{k=1}^K R_k(n+1)} C \right\},$$

$$R_i(n+1) = \hat{X}_i(n+1) + Q_i(n) \quad (9)$$

첫 번째 성능 분석 실험으로 링크 용량 $C=29,820$ packets/sec(Load=0.70) 일때, 제안된 연결 수락 제어를 통해 순차적으로 수락된 25개의 비디오 소스

에 대해 동일한 QoS가 요구되었을 경우에 제안된 방법과 기존의 방법이 어떤 성능을 보이는가를 비교하였다. 표 1은 제일 마지막에 수락된 5개의 비디오 소스에 대한 결과를 보여준다.

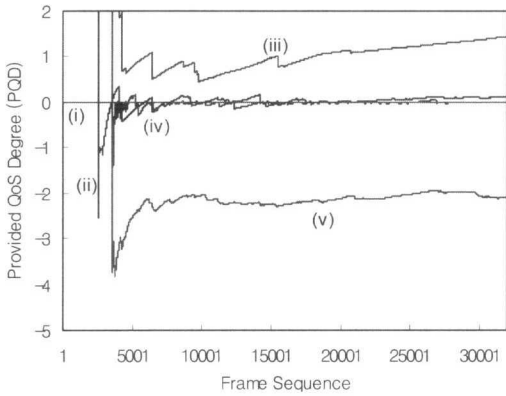
기존의 방법은 표 1에서 알 수 있듯이, 각 연결에 제공된 QoS가 큰 차이를 보인다. QoS가 요구사항보다 높게 제공된 비디오와 낮게 제공된 비디오 소스가 존재한다. 제안된 방법은 모든 비디오 소스가 거의 일정하게 요구된 품질로 서비스 되었음을 알 수있다. 제안 방법에서는 현재 요구된 정도 보다 낮은 QoS를 제공 받고 있는 비디오 소스에 대해서는 보다 많은 소요대역이 추정되고, 공정 대역 할당을 통해 우선적으로 QoS 를 지원 할 수 있도록 처리된다. 반면에 요구된 정도 보다 높은 QoS를 제공 받는 비디오는 적은 대역이 할당되고 이렇게 삭감된 자원은 낮은 QoS를 제공 받는 연결들을 위하여 사용된다.

표 1. 동일 QoS 조건에서의 QoS 제공 성능 분석.

Video Connection	QoS Requirements		Provided QoS			
	Loss Ratio	max Delay	Loss Ratio	PQD	Loss Ratio	PQD
Star Wars	$1.0 * 10^{-5}$	2 frames	$0.95 * 10^{-5}$	0.050	$0.98 * 10^{-5}$	0.009
The Silence of Lambs	$1.0 * 10^{-5}$	2 frames	$2.40 * 10^{-5}$	-0.872	$1.00 * 10^{-5}$	-0.003
Talk Show	$1.0 * 10^{-5}$	2 frames	$4.10 * 10^{-4}$	0.892	$0.99 * 10^{-5}$	0.006
Mr. Bean	$1.0 * 10^{-5}$	2 frames	$1.93 * 10^{-4}$	1.645	$0.90 * 10^{-5}$	0.109
Terminator	$1.0 * 10^{-5}$	2 frames	$7.30 * 10^{-5}$	-1.986	$1.01 * 10^{-5}$	-0.009

그림 5는 “Mr. Bean”과 “Terminator” 비디오 시퀀스에 대하여 시간에 따른 PQD (provided QoS degree)의 변화를 보여준다. 그림에서 볼 수 있듯이 제안된 방법은 전체 연결 시간 동안 사용자가 요구하는 레벨의 품질에 매우 근사하게 서비스를 하고 있다.

그림 6은 제안된 방법과 기존 방법에서의 패킷 지연의 분포도이다. 제안된 방법은 3장에서 설명한 바와 같이 버퍼에서 최대 지연 한계치를 초과하는 패킷은 자동적으로 매 프레임 종료 시점에 폐기되므로 서비스된 패킷들은 지연 한계치 (표 1의 경우 2 프레임 시간)를 초과하여 전달 되지 않는다. 제안된 방법은 각 노드에서 이미 수신단에 전달해도 지연 한계시간을 초과하는 패킷들을 폐기시킴으로서 이러한 절차가 없는 기존의 방법에 비해 큐잉 성능이 향상된다.



(i) desired (PQD=0), (ii) proposed (Mr. Bean), (iii) conventional (Mr. Bean),
(iv) proposed (Terminator), (v) conventional (Terminator)

그림 5. 시간에 따른 제공 QoS 정도 (PQD) 변화
($loss R_i = 10^{-3}$, $d \max(i)=2$).

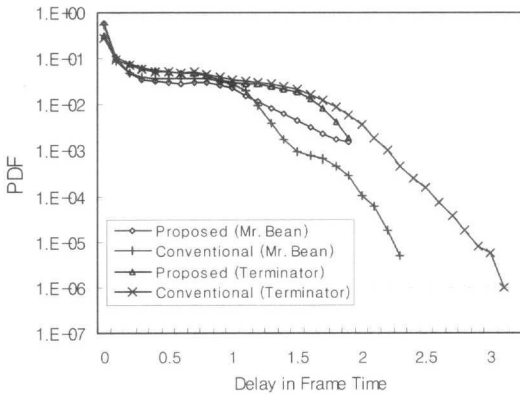


그림 6. 패킷 지연 분포 ($loss R_i = 10^{-3}$, $d \max(i)=2$).

두번째 성능분석은 서로 다른 QoS를 요구하는 비디오 연결에 대해 수행되었다. $\{loss R_i = 10^{-3}, d \max(i)=2\}$ 의 서비스 요구사항을 갖는 20개의 연결이 이미 수락 되어 서비스 되고 있는 상황에서 표 2의 다섯개의 비디오 연결이 순차적으로 연결 수락 되었다. 역시 기존 방법도 같은 수락 시점에서 서비스가 시작될 때 QoS 요구사항과 실제 서비스된 품질간의 비교가 표 2에 나타나있다. 링크용량은 $C=29,820$ packets/sec(Load= 0.70)이다. 표 3은 기존 방법과 제안된 방법에서 실제 각 비디오 연결에 할당된 평균 대역을 보여준다.

표 2에서 볼 수 있듯이, 제안된 방법은 모든 연결에 대해서 요구된 서비스 레벨을 안정되게 서비스 하고 있으나, 기존의 방법은 각 연결 별 제공된

표 2. 서로 다른 QoS 요구사항을 갖는 연결들의 QoS 제공 성능 분석.

Video Connection	QoS Requirements		Provided QoS			
			Conventional		Proposed	
	Loss Ratio	max Delay	Loss Ratio	PQD	Loss Ratio	PQD
Star Wars	1.0×10^{-3}	1 frame	7.53×10^{-2}	-4.322	1.02×10^{-3}	-0.016
The Silence of Lambs	5.0×10^{-4}	1 frame	5.59×10^{-2}	-4.717	5.04×10^{-4}	-0.007
Talk Show	5.0×10^{-4}	2 frames	0.58×10^{-4}	2.157	4.87×10^{-4}	0.027
Mr. Bean	1.0×10^{-4}	2 frames	0.32×10^{-4}	1.134	0.98×10^{-4}	0.017
Terminator	5.0×10^{-3}	3 frames	3.86×10^{-6}	2.562	3.95×10^{-3}	0.234

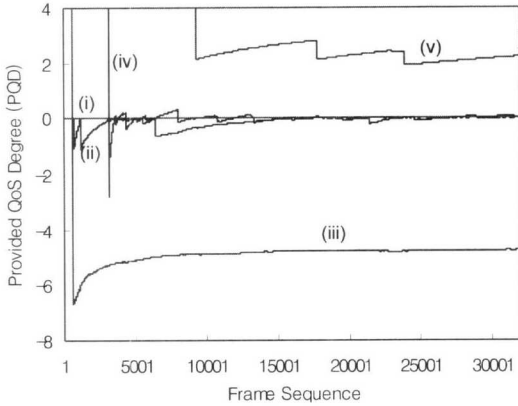
표 3. 서로 다른 QoS에서 각 연결별 평균 할당 대역 비교 (packets/frame).

Video Connection	Conventional			Proposed		
	Average Input	Allocated Bandwidth	Ratio	Average Input	Allocated Bandwidth	Ratio
Star Wars	37.7	50.54	1.342	37.7	51.7	1.373
The Silence of Lambs	18.8	19.69	1.047	18.8	22.4	1.297
Talk Show	36.3	37.41	1.030	36.3	37.6	1.035
Mr. Bean	42.5	43.55	1.025	42.5	52.7	1.241
Terminator	25.2	27.8	1.059	25.2	33.0	1.150

$$\text{Ratio} = (\text{Allocated Bandwidth}) / (\text{Average Input})$$

QoS 준수 정도가 큰 차이를 보인다. 표 3의 결과는 제안된 방법이 QoS를 준수하기 위해 기존의 방법보다 약간 더 많은 대역을 각 연결에게 할당 함을 나타낸다. 그러나 제안된 방법 역시 평균 할당 대역이 평균 입력 레이트에 가까우므로, 각 프레임 시간에서 링크용량에서 각 비디오 연결에 할당된 대역을 제외한 나머지 대역은 다른 서비스 클래스의 응용을 위해 사용될 수 있다. 그림 7은 표 2 연결들의 시간에 따른 PQD의 변화를 나타낸다.

그림 8은 “The Silence of Lambs”와 “Talk Show” 비디오 연결의 패킷 지연 분포를 보여준다. 제안된 방법은 큐에서의 폐기 정책을 따라 항상 최대 지연 바운드 내에서 패킷을 서비스 하며, 지연분포도 기존 방법에 비해 비교적 완만하다. 이는 지연이 허용되면, 다른 서비스가 급하게 이루어져야 할 패킷을 먼저 서비스 하고 해당 패킷은 버퍼에서 대기 시키기 때문이다.



(i) desired (PQD=0), (ii) proposed (The Silence of Lambs), (iii) conventional (The Silence of Lambs), (iv) proposed (Talk Show), (v) conventional (Talk Show)

그림 7. 시간에 따른 제공 QoS 정도 (PQD) 변화 (서로 다른 QoS 요구사항 조건).

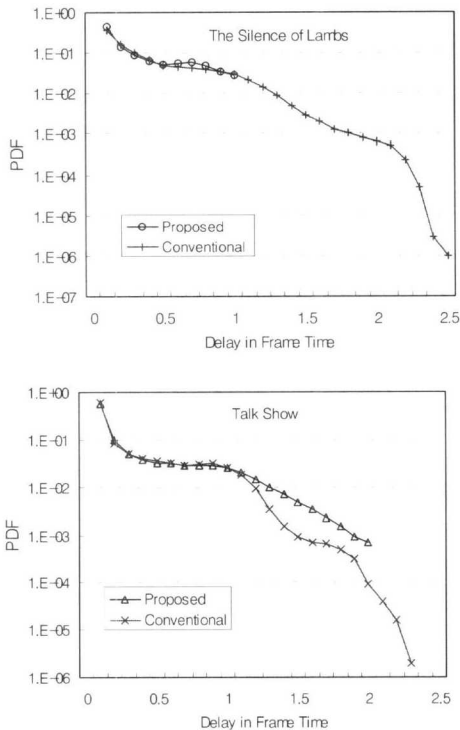


그림 8. 패킷 지연 분포 (서로 다른 QoS 요구사항 조건).

VI. 결론

본 논문에서는 VBR 비디오 소스에 대하여 QoS 요구사항을 고려하여 동적으로 자원을 할당하는 새

로운 구조 및 방법을 제안하였다. 통신망에서 비디오 서비스 시, 사용자에게 트래픽 기술자를 미리 망에게 알려줄 것을 요구하지 않으며 망의 이용도 (utilization)도 기존 ATM 망에서와 같이 고정 대역을 할당할 때 보다 크게 향상시킬 수 있다. 기존의 동적 자원할당 연구에서는 각 비디오 연결 별로 안정되게 QoS를 제어하는 것이 어려운데 비해, 제안된 방법은 간단한 지연 패킷 카운터 갱신 과정과 제공된 QoS 레벨의 관측을 통한 적응적 자원 할당으로 각 연결별 QoS를 매우 정확히 제공할 수 있다. 연결 수락절차는 비디오 서비스 응용들이 연결 기간 동안 안정된 QoS를 제공 받을 수 있도록 한다.

본 논문에서 새롭게 제안된 사항은, i) 간단한 구조의 QoS 모니터링 방법, ii) 제공된 QoS 레벨에 따라 차등적인 대역 할당 정책, iii) 지연 한계치를 초과하는 패킷에 대한 폐기로 전체 통신망에서의 큐잉 특성의 향상, iv) 동적 자원할당 구조에서의 연결 수락제어를 통한 안정된 서비스 제공, 등이며 향후 차세대 인터넷, ATM과 같은 통신망에서 다양한 실시간 멀티미디어 응용의 QoS 제공 방법으로 적용될 수 있을 것이다.

참고 문헌

- [1] "Traffic management specification version 4.0", ATM Forum Technical Committee, 1996
- [2] Marwan Krunz, "Bandwidth allocation strategies for transporting variable-bit-rate video traffic", IEE Communication Magazine, pp. 40-46, January 1999.
- [3] G. Chiruvolu, R. Sankar, N. Ranganathan, "VBR video traffic management using a predictor-based architecture", Computer Communications, Vol. 23, pp. 62-70, 2000.
- [4] Chin-Ling Chen, Ruay-Shiung Chang, "A new dynamic bandwidth allocation scheme for MEG videos in ATM networks", Computer Communication, Vol. 23, pp. 1505-1513, 2000.
- [5] X. Wang, J.S. Meditch, "A predictive bandwidth management scheme and network architecture for real-time VBR traffic", Microprocessors and Microsystems, Vol. 22, pp. 597-604, 1999.
- [6] Suili Feng, Ravi Sanker, "Dynamic bandwidth

allocation scheme for stationary VBR traffic”, in Proceedings of the IEEE GLOBECOM'99, pp. 1478-1482, 1999.

- [7] Song Chong, San-qi Li, Joydeep Ghosh, “Predictive dynamic bandwidth allocation for efficient transport of real-time VBR video over ATM”, IEEE Journal on Selected Areas in Communications, Vol. 13, No.1, pp. 12-23, 1995.
- [8] Tejinder S. Randhawa, R.H.S. Hardy, “Proactive management of MPEG traffic in ATM networks using time sequenced RLS filters”, in Proceedings of the ICATM'99, pp. 507-511, 1999.
- [9] Abdelnaser Mohammad Adas, “Using adaptive linear prediction to support real-time VBR video under RCBR network service model”, IEEE/ACM Transactions on Networking, Vol. 6, No. 5, pp. 635-644, 1998
- [10] ftp-info3.infomaik.uni-uerzbug.de.pub/MPEG
- [11] Y. Afek, M. Cohen, E. Haalman, Y. Mansour, “Dynamic bandwidth allocation policies”, in Proceedings of the IEEE Infocom'96, pp. 880-887, 1996.

유 상 조(Sang-Jo Yoo)



1988년 2월 : 한양대학교
전자통신공학과 (공학사)
1990년 2월 : 한국과학기술원
전기 및 전자공학과
(공학석사)
2000년 8월 : 한국과학기술원
전자전산학과 (공학박사)

- 1990년 3월~2001년 2월 : 한국통신 연구개발본부
전임연구원
- 2001년 3월~현재 : 인하대학교 정보통신대학원
조교수
- <주관심 분야> 인터넷 QoS, 초고속 통신망 구조,
멀티미디어 네트워크, 트래픽엔지니어링

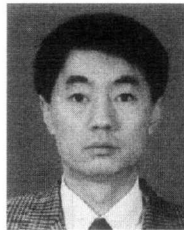
홍 성 훈(Sung-Hoon Hong)



1999년 8월 : 한국과학기술원
전기 및 전자공학과
(공학박사)
1991년 1월~2000년 7월 : LG
전자 DTV 연구소
책임연구원

- 2000년 7월~현재 : 전남대학교 정보통신공학부
전임강사
- <주관심 분야> 영상처리 및 압축 부호화, 멀티미
디어 통신, 디지털TV, 영상처리용 VLSI
설계 등

강 성 원(Sungwon Kang)



1982년 : 서울대학교 사회과학대
학 졸업
1989년 : 미국 University of
Iowa 전산학 석사
1992년 : 미국 University of
Iowa 전산학 박사

- 1993년 12월~2001년 10월 : 한국통신 연구개발본부
신임연구원
- 2001년 10월~현재 : 한국정보통신대학원대학교
공학부 조교수
- <주관심 분야> 프로그램 최적화, 프로토콜공학, 객
체지향시험, 차세대통신망구조, 액티브네트
워크

홍 경 표(Kyoung-Pyo Hong)

- 1984년 : 한국항공대 전자공학과(학사)
- 1987년 : 한국과학기술원 전기 및 전자공학과(석사)
- 1987년~현재 : 한국통신 통신망연구소 통신망구조연
구실장
- 1997년 : 한국과학기술원 전기 및 전자공학과(박사)
- <주관심 분야> 프로토콜 엔지니어링, 분산컴퓨팅,
정책기반 네트워크 기술