

모바일 멀티미디어 시스템에서 버퍼 수준을 이용하는 통신망 대역폭 할당 기법

김진환*

*한성대학교 멀티미디어공학

e-mail:kimjh@hansung.ac.kr

A Network Bandwidth Allocation Scheme using Buffer Level in a Mobile Multimedia System

Jinhwan Kim*

*Dept of Multimedia Engineering, Hansung University

요 약

모바일 멀티미디어 시스템에서 클라이언트의 재생 버퍼 수준을 이용하여 통신망 대역폭을 동적으로 할당하는 기법이 본 논문에서 제시된다. 이 기법에서 비디오 요청을 서비스하기 위해 할당된 대역폭의 양은 요구한 클라이언트의 버퍼 수준에 따라 결정된다. 즉 클라이언트의 버퍼 수준이 일시적으로 낮아지거나 높아지면 재생 상황을 개선하기 위한 목적으로 더 많은 대역폭 또는 더 적은 대역폭이 할당되는 것이다. 제시된 기법의 주 목적은 전송이 중단되는 프레임의 수를 최소화하여 비디오 재생 성능과 서비스 품질을 향상시키며 동시에 진행되는 모든 비디오 요청 작업을 공평하게 서비스하는 것이다. 제시된 기법의 성능은 시뮬레이션을 통하여 다른 정적 대역폭 할당 기법과 비교 분석되었다.

1. 서론

지난 수년동안 모바일 통신망의 급격한 성장과 발전은 대역폭과 서비스 품질(QoS; Quality of Service) 보장 등 효과적이며 효율적인 자원 활용에 대한 요구를 증가시켜왔다[1]. 모바일 멀티미디어 시스템에서 모바일 클라이언트들은 모바일 셀룰라 통신망이나 무선 LAN과 같은 무선 통신망으로 비디오 서버에 연결된다[2].

유선 통신망의 분산 비디오 재생장치 시스템의 성능을 개선하기 위해 제시된 비디오 프레임 버퍼링, 피드백 제어 메카니즘, 비디오 스트림 스무딩 기법[3] 등은 모바일 시스템의 특성으로 인하여 모바일 비디오 재생장치 시스템에 동일하게 적용하기에는 여러 가지 제약 사항들이 있다[4]. 특히 휴대폰 등 모바일 단말기의 버퍼는 아직 노트북이나 데스크탑 컴퓨터에 비하면 버퍼 용량이 크지 않다. 이러한 모바일 단말기에 재생될 비디오 스트림을 위하여 모바일 통신망의 대역폭을 할당하는 기존의 스케줄링 기법들은[4, 5] 비디오 트래픽 크기의 변동성과 일시적인 과부하 현상을 효과적으로 해결하기 어려운 면이 있으며 또한 반드시 여분으로 사용할 수 있는 대역폭을 확보해야 한다는 문제점이 있다.

현재 무선 통신망은 비디오, 음성, 데이터 등 멀티미디어 트래픽에 대한 서비스 품질을 효율적으로 제공하기 위한 효과적인 방법으로 중앙 스케줄러를 사용하고 있다[2]. 이 중앙 스케줄러는 트래픽의 양과 채널 조건에 따라 대역폭을 동적으로 할당할 수 있으며 스케줄링 서비스를 공

평하게 수행할 수 있는 장점이 있다. 본 논문에서도 모바일 통신망 기반의 클라이언트-서버 시스템[6]에서 서버가 중앙 스케줄러의 역할을 수행하며 비디오 스트림을 효율적으로 전송할 수 있도록 클라이언트의 버퍼 수준에 기반한 동적 대역폭 할당 기법이 제시된다.

2. 모바일 멀티미디어 시스템

본 논문에서는 모바일 통신망을 기반으로 하는 모바일 멀티미디어 시스템을 고려한다. 모바일 통신망에서는 기지국과 모바일 클라이언트들이 셀(cell) 단위를 기반으로 연결되며 기지국은 고속 통신망을 통하여 비디오 서버와 연결된다. 모바일 클라이언트들은 기지국을 통하여 비디오 서버에 비디오 재생을 요청하게 된다. PDA, 휴대폰, 포켓 PC 등을 사용하는 thin 클라이언트들은 노트북 컴퓨터를 사용하는 thick 클라이언트에 비하여 비디오 버퍼 크기가 작으며 제한이 있게 된다[7].

모바일 클라이언트는 서버로부터 수신한 프레임들을 자신의 버퍼에 저장하며 버퍼 수준이 사전에 정의된 수치에 도달하면 비디오의 재생을 시작하게 된다. 즉 복원 과정이 시작되는 것이며 실제 비디오의 재생은 복원된 프레임들을 비디오 재생장치에 한 개씩 전송함으로써 수행된다. 그러나 비디오 프레임의 재생 시간을 놓치게 되면 해당 프레임은 즉시 재생이 취소되며 폐기된다. 이러한 현상은 끊김없는(seamless) 비디오 스트리밍(streaming) 유지에 심각한 영향을 미치게 되며 재생율(1초간 전송되는 프레

임들의 수)에 관한 서비스 품질과 실시간적 성능도 저하시키게 된다[8].

본 논문에서 비디오 서버는 MPEG-4 표준으로 압축된 비디오를 가지고 있는 것으로 가정하며 각 비디오 스트림은 GOP(Group of Picture) 시퀀스로 구성되어 있다. 각 GOP는 I, P, B 프레임들로 구성된다.

3. 버퍼 수준을 이용하는 동적 대역폭 할당 기법

모바일 통신망의 비디오 서버는 클라이언트가 요청한 MPEG-4 비디오 스트림을 전송하기 전 수락 제어(admission control) 절차를 수행한다. 클라이언트들이 요청한 비디오 스트림 n 개의 평균 대역폭들의 합은 수식 (1)에서 M_{sum} 으로 정의된다. 임의의 스트림에 대한 평균 대역폭은 비트 수로 나타나는 프레임의 평균 크기를 재생율과 곱한 값으로 결정된다.

모바일 통신망에서 기지국과 클라이언트들 간에 사용가능한 최대 대역폭을 B_{total} 이라 하며 실제 사용가능한 통신망 대역폭을 B_{avail} 로 정의한다. 즉 B_{avail} 은 B_{total} 에서 오류 발생시 재전송을 위한 대역폭과 비디오 스트림 전송 이외의 목적으로 사용되는 대역폭을 제외한 대역폭을 의미한다. M_{sum} , B_{avail} , B_{total} 들에 대해서는 수식 (1)의 관계가 적용된다.

$$M_{sum} \leq B_{avail} \leq B_{total} \quad (1)$$

새로운 클라이언트가 요청한 $n+1$ 번째 비디오 스트림의 평균 대역폭이 추가된 대역폭의 합이 수식 (1)의 관계를 계속 만족하면 서버가 이를 수용하지만 그렇지 않은 경우에는 이를 거부할 수 있다. 그리고 거부된 비디오 스트림은 나중에 서버가 재수용할 수 있으나 이에 관한 구체적인 과정은 본 논문에서 기술하지 않는다.

수락 제어 이후 각 비디오 스트림의 실제 대역폭은 VBR 특성에 따라 크기가 가변적이며 일정 주기마다 i ($1 \leq i \leq n$)번째 스트림이 요청한 프레임들의 크기를 비트로 표시한 실제 대역폭 B_i 와 이들의 합인 B_{sum} 은 수식 (2)의 관계로 성립된다.

$$B_{sum} = \sum_{i=1}^n B_i \quad (2)$$

비디오 서버는 매초마다 각 스트림이 요구한 프레임들을 저장장치에서 검색 후 서버의 버퍼에 저장하게 되며 B_{sum} 은 B_{avail} 을 초과하는 경우가 발생할 수 있다. 이 경우 통신망 대역폭의 한계로 인해 서버가 전송을 중단시켜야 하는 프레임들이 발생하게 된다. 전송이 중단되는 프레임들의 수가 증가할수록 해당 클라이언트는 끊임없는 비디오 스트리밍 유지가 어렵게 되고 재생율 및 서비스 품질(QoS)이 저하되는 현상이 발생한다. 본 논문에서는 이러한 현상을 최소화시키고자 서버가 각 클라이언트의 버퍼 수준을 고려하여 매주기마다 대역폭을 동적으로 할당한다.

버퍼 수준이란 클라이언트의 단말기가 일정 시간 동안 전송된 프레임들을 저장할 수 있는 버퍼의 크기를 의미한다. 모바일 클라이언트는 단말기의 특성상 기억장치의 크

기가 상이하며 실제 버퍼의 수준도 상이하게 된다. 버퍼 수준이 높을수록 오랜 시간 동안 재생할 수 있는 많은 수의 프레임들을 버퍼에 유지하게 되며 통신망의 일시적 장애나 오류 발생 시에도 덜 민감하게 반응할 수 있는 장점이 있다[7]. 그러나 버퍼 수준이 높을 경우 비디오 재생이 시작되는 시간이 늦어지고 기억장치 활용도가 저하되는 문제가 있으며 모바일 단말기에 대용량의 버퍼를 유지하기 어려운 면이 있다.

i 번째 비디오 스트림을 위한 단말기의 버퍼 수준을 L_i 라 하며 시간 단위인 초를 이용하여 크기를 설정한다. 본 논문에서 임의의 비디오 스트림을 재생하기 위한 모든 단말기의 버퍼 수준은 초기치가 1보다 크거나 같은 값을 가지는 것으로 가정한다. 서버는 각 스트림의 버퍼 수준을 일정 주기(본 논문에서는 1초로 가정함)마다 파악하며 버퍼 수준의 합 L_{sum} 을 다음과 같이 정의한다.

$$L_{sum} = \sum_{i=1}^n L_i \quad (3)$$

서버는 전송하고자 하는 실제 대역폭의 합 B_{sum} 이 사용가능한 통신망 대역폭의 합 B_{avail} 보다 클 경우 단말기 버퍼 수준을 고려하여 버퍼 수준이 높은 클라이언트일수록 비디오 스트림을 가급적 다음번에 전송하도록 대역폭을 조정한다. 실제 대역폭의 합 B_{sum} 이 B_{avail} 보다 클 경우의 동적 대역폭 할당 방법이 의사 언어로 그림 1에서 기술되었다.

B_{sum} 과 B_{avail} 의 차는 B_{lack} 으로 그림 1의 3행에서 기술되었으며 B_{lack} 은 결국 부족한 대역폭을 의미한다. B_{lack} 이 0보다 클 경우(4행) 클라이언트들의 버퍼 수준 중 가장 큰 값인 L_{max} 를 찾는 과정이 반복된다(6행). L_{max} 가 1.0을 초과하는 경우(7행) L_{max} 를 가진 클라이언트의 대역폭 B_i 는 B_{lack} 과 비교된다. B_i 가 B_{lack} 보다 작거나 같을 경우(9행) 해당 B_i 는 전송이 보류되며 이 대역폭을 P_i 로(10행) 설정하고 L_{max} 는 1.0이 차감된다(11행). 해당 클라이언트는 1초간 서버로부터 단말기 버퍼에 있는 프레임들을 재생함으로써 버퍼 수준이 감소되는 것이다. 이와 같이 서버는 버퍼 수준이 낮은 다른 클라이언트들의 프레임들이 더 많이 전송될 수 있도록 버퍼 수준이 높은 클라이언트의 프레임들의 전송을 보류하게 된다. B_{lack} 은 전송이 보류된 B_i 만큼 차감하게 되며(11행) 이는 필요한 통신망 대역폭의 크기를 감소시킨 것을 의미한다. 9행과는 달리 L_{max} 를 가진 클라이언트의 대역폭 B_i 가 B_{lack} 보다 클 경우(12행) B_i 에서 B_{lack} 을 차감한 대역폭은 전송이 가능하며 B_{lack} 만큼만 전송이 보류되며 P_i 로 설정된다(13행). 이 경우 해당 클라이언트는 B_{lack}/B_i 시간(단위 초)동안은 단말기 버퍼에 있던 프레임들을 재생하게 되므로 버퍼 수준도 이 비율만큼 차감된다(14행). 서버가 해당 클라이언트에게 전송하는 프레임들의 재생 시간은 $1.0 - B_{lack}/B_i$ 이 되는 것이다. 서버는 더 이상 보류할 수 있는 대역폭이 없으므로 B_{lack} 은 0이 되고(14행) 4행의 반복문을 벗어나므로 새로운 L_{max} 를 찾는 과정은 중지된다.

```

1: if (Bsum > Bavail)
2:   begin
3:     Black = Bsum - Bavail;
4:     while (Black > 0)
5:       begin
6:         Li가 가장 큰 값을 Lmax로 설정함;
7:         if (Lmax > 1.0)
8:           begin
9:             Lmax에 해당하는 Bi가 Black보다 작거나 같으면
10:            Bi는 전송보류되며 이 대역폭은 Pi로 설정함;
11:            Lmax를 1.0 차감하고 Black에서 Bi를 차감함;
12:            Lmax에 해당하는 Bi가 Black보다 크면
13:            Black만큼 전송보류되며 이를 Pi로 설정함;
14:            Lmax에서 Black/Bi 차감하고 Black은 0이 됨;
15:          end
16:        else // Lmax ≤ 1.0
17:          begin
18:            Bi에서 전송보류된 Pi를 빼고 Bi'로 설정함;
19:            Bsum' = ∑i=1n Bi';
20:            Black' = Bsum' - Bavail;
21:            for i=1 to n do
22:              begin
23:                Pi' = Black' × Bi' / Bsum';
24:                Di = Bi - Pi - Pi';
25:              end
26:            Black은 0으로 설정함;
27:          end
28:        end

```

(그림 1) B_{sum} > B_{avail} 경우의 동적 대역폭 할당 알고리즘

한편 7행과 달리 L_{max}가 1.0이하인 경우(16행) 서버는 각 클라이언트 i 마다 실제 대역폭 B_i 에서 전송이 보류된 대역폭 P_i (10행과 13행에서 설정됨)를 차감한 대역폭의 합 B_{sum}'을 구하고(19행) B_{sum}'에서 B_{avail}을 차감한 대역폭을 B_{lack}'으로 정의한다(20행). B_{lack}'은 B_{sum}'을 전송하기에 여전히 부족한 대역폭을 의미한다. 따라서 서버는 각 클라이언트마다 전송을 보류해야 하는 대역폭을 한번 더 산정하게 된다. 스트림 i 마다 B_i' 대역폭을 B_{sum}'으로 나눈 비율과 B_{lack}'을 곱하여 전송을 보류해야 하는 대역폭 P_i'가 결정된다(24행). 각 스트림 i 의 실제 대역폭 B_i에서 전송이 1차로 보류된 P_i와 2차로 보류된 P_i'를 모두 차감한 대역폭은 D_i로 설정되며(24행) 이는 실제로 전송이 중단되는 대역폭을 의미한다. 전송이 보류된 대역폭에 해당하는 프레임들은 다음 주기에 전송될 수 있으나 전송이 중단된 대역폭에 속한 프레임들은 서버에 의해 폐기되는 것을 의미한다. 본 논문에서는 전송이 중단되는 프레임들에 대하여 우선순위 정책을 적용한다. 즉 중요도가 가장 낮은 B 프레임들이 먼저 중단되며 이후 P 프레임, I 프레임 순으로 결정됨으로써 중요도가 큰 프레임일수록 전송이 중단되는 경우가 최소화되도록 하였다. 그리고 26행에서 서버는 새

로운 L_{max}를 찾는 과정을 중지하고자 B_{lack}을 0으로 설정하여 4행의 반복문을 벗어나게 된다.

그리고 B_{sum}이 B_{avail}보다 작은 경우에는 서버가 모든 클라이언트의 프레임들을 전송한 후 여분의 대역폭을 버퍼 수준이 가장 낮은 클라이언트부터 추가로 할당하여 다음 주기에 전송할 프레임들을 이번 주기에 미리 전송함으로써 해당 클라이언트의 버퍼 수준이 추가로 증가된다. 이에 대한 알고리즘은 본 논문에서 생략한다.

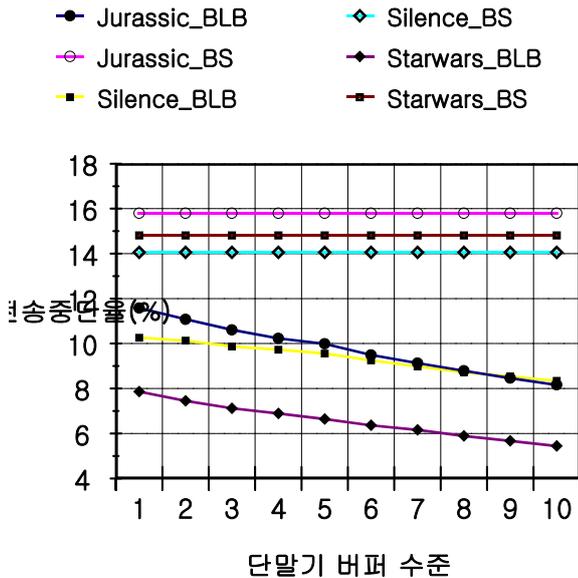
4. 성능 분석

본 논문에서는 압축률이 우수한 VBR(Variable Bit-Rate) 방식으로 부호화된 MPEG-4 비디오 스트림 "Jurassic Park One", "Silence of the Lambs", "Star Wars IV" 등[10]을 대상으로 실험을 수행하였다. 이 비디오 스트림들은 모두 QCIF 형식이며 GOP내 프레임 수는 12이고 GOP 패턴은 IBBPBBPBBPBB이다. 세 비디오 스트림들의 총 프레임 수는 각각 108000개이며 프레임의 평균크기는 각각 276.85, 214.0, 247.5 바이트로 분석되었다. 비디오 스트림의 재생율은 24이며 20개 클라이언트들을 대상으로 실험이 수행되었다.

단말기 버퍼 수준을 사전에 파악하여 비디오 스트림별로 통신망 대역폭을 정적으로 할당하는 BS(Buffer Sensitive) 대역폭 할당 기법[4]과 제시된 BLB(Buffer Level-based) 대역폭 할당 기법의 성능이 비교 분석되었다. BS 대역폭 할당 기법은 이용가능한 통신망 대역폭을 스트림의 수로 나눈 후 이 값을 각 스트림의 평균 대역폭과 비교하여 더 작은 값을 해당 스트림의 실제 대역폭으로 결정한다. 이후 모든 스트림의 실제 대역폭의 합이 이용가능한 통신망 대역폭보다 작을 경우에만 여분의 통신망 대역폭이 발생하며 이 여분의 대역폭을 스트림별로 단말기 버퍼 수준을 고려하여 추가 할당하게 된다. 그러나 BS 대역폭 할당 기법은 스트림별로 통신망 대역폭을 정적으로 할당한 후 전송 중 발생하는 단말기 버퍼 수준을 별도로 고려하지 않기 때문에 이러한 사항을 고려하여 동적으로 대역폭을 조절하는 BLB 대역폭 할당 기법과 성능을 직접 비교하기는 사실상 어렵다. 본 실험에서는 BS 기법과 BLB 기법의 수락 제어 단계에서 통신망의 여유 대역폭은 전혀 없는 것으로 가정한다. 두 기법은 동일한 실험 환경에서 시뮬레이션이 4200초간 수행되었으며 이 결과를 토대로 두 기법의 성능이 비교 분석되었다.

본 논문에서는 두 기법의 성능을 비교하기 위하여 프레임 전송중단율이 측정된다. 즉 서버가 각 클라이언트 단말기로 전송할 때 대역폭 부족으로 전송을 중단시켜야 하는 프레임의 수를 서버가 각 클라이언트 단말기로 전송해야 하는 프레임의 수로 나눈 값이 프레임 전송중단율로 정의된다. 이 비율이 클수록 클라이언트들이 재생하지 못하는 프레임들의 수가 증가하는 것이다. 모든 단말기 버퍼 수준의 초기치를 1.0부터 10.0까지 1.0씩 차례로 증가시킨

상태에서 두 기법의 프레임 전송중단율을 각각 비교한 결과가 그림 2에 나타나 있다. "Jurassic Park One"의 경우 BS 기법은 여분의 통신망 대역폭이 없으므로 버퍼 수준의 크기에 상관없이 프레임 전송중단율이 동일하며 약 15.80%로 분석되었다. 그러나 BLB 기법의 경우에는 버퍼 수준이 1.0인 경우 프레임 전송중단율이 약 11.58%로 나타났으며 버퍼 수준의 초기치를 1.0씩 증가시킬 때마다 약 0.2%에서 0.4% 정도로 프레임 전송중단율이 감소되었으며 버퍼 수준이 10.0인 경우에는 프레임 전송중단율이 약 8.37%로 분석되었다.



(그림 2) 프레임 전송 중단율

여분의 통신망 대역폭이 없는 상태에서 버퍼 수준에 따라 통신망 대역폭을 동적으로 조정하여 스트림의 프레임 전송하는 BLB 기법의 경우 대역폭을 정적으로 할당하는 BS 기법보다 프레임 전송중단율이 감소되는 결과가 나타났으며 버퍼 수준의 초기치를 증가시킬수록 근소한 차이로 프레임 전송중단율이 감소되는 결과도 나타났다. "Silence of the Lambs"와 "Star Wars IV" 경우 BS 기법의 프레임 전송중단율은 각각 14.06%와 14.81%로 나타났으며 BLB 기법의 프레임 전송중단율은 버퍼 수준에 따라 10.26%에서 8.34%로 감소된 결과와 7.86%에서 5.43%로 감소된 결과가 각각 나타났다. 프레임의 최대값을 평균값으로 나눈 비율은 "Star Wars IV"가 17.29로 가장 크지만 프레임 전송중단율이 가장 낮게 나타났으며 이는 평균대역폭을 초과하는 빈도수가 다른 비디오 스트림들에 비하여 상대적으로 낮게 발생한 것이 원인으로 분석되었다.

5. 결론

모바일 멀티미디어 시스템에서 클라이언트가 요청한 비디오 스트림을 서버가 전송할 때 단말기 버퍼 수준을 활용하여 모바일 통신망 대역폭을 동적으로 할당하는 기

법이 본 논문에서 제시되었다. 제시된 BLB 기법은 단말기 버퍼 수준을 고려하여 제한된 통신망 대역폭을 효율적으로 사용함으로써 서버가 대역폭 부족으로 인하여 전송을 중단시켜야 하는 프레임 수를 최소화한다. 서버는 이를 위하여 버퍼 수준이 큰 클라이언트의 프레임들은 다음 주기에 전송할 수 있도록 보류하고 보류로 인해 확보된 대역폭은 버퍼 수준이 작은 클라이언트가 활용할 수 있도록 동적으로 조정하게 된다. 그리고 전송 중인 스트림의 대역폭 합이 이용가능한 통신망 대역폭보다 작을 경우에는 모든 클라이언트들의 프레임을 전송하고 남은 대역폭은 버퍼 수준이 작은 클라이언트들에게 추가로 할당하여 더 많은 프레임들이 전송되도록 함으로써 버퍼 수준별 프레임 전송중단율의 차이를 최소화할 수 있는 공평성이 유지된다. 본 논문의 BLB 기법은 성능 비교 결과 비디오 스트림의 서비스 품질과 끊김없는 비디오 스트리밍 효과를 향상시킬 수 있는 것으로 분석되었다.

참고문헌

[1] M. Ravichandran, P. Sengottuvelan, and A. Shanmugam, "An Approach for Admission Control and Bandwidth Allocation in Mobile Multimedia Network Using Fuzzy Logic," *Int'l Journal of Recent Trends in Engineering*, Vol. 1, No. 1, May 2009.

[2] L. Xu, X. Shen and J. W. Mark, "Dynamic Fair Scheduling With QoS Constraints in Multimedia Wideband CDMA Cellular Networks", *IEEE Trans. on Wireless Communications*, Vol. 3, No. 1, pp. 60-73, 2004.

[3] S. Rao and A. M. K. Cheng, "Scheduling and Routing of Real-Time Multimedia Traffic in Packet-Switched Networks," *IEEE International Conference on Multimedia*, July -Aug. 2000.

[4] J. Yuen, K. Y. Lam, and E. Chan, "A Fair and Adaptive Scheduling Protocol for Video Stream Transmission in Mobile Environment," *IEEE International Conference on Multimedia and Expo*, Lausanne, Switzerland, August 2002.

[5] K. Y. Lam, J. Yuen, S. H. Son, and E. Chan, "Scheduling Video Stream Transmissions for Distributed Playback over Mobile Cellular Networks," *Int'l Conf. on Parallel and Distributed Systems*, Dec. 2002.

[6] M. Burza, J. Kang, and P. van der Stok, "Adaptive Streaming of MPEG-based Audio/Video Content over Wireless Networks," *Journal of Multimedia*, Vol. 2, No. 2, pp. 17-27, 2007.

[7] X. Hei, Y. Liu, and K. W. Ross, "Inferring Network-Wide Quality in P2P Live Streaming Systems," *IEEE Journal on Selected Areas in Communications* Volume 25, Issue 9, pp. 1640-1654, Dec. 2007.

[8] X. Cheng, P. Mohapatra, S. Lee, and S. Banerjee, "Performance Evaluation of Video Streaming in Multihop Wireless Mesh Networks," *Network and Operating Systems Support for Digital Audio and Video*, May, 2008.

[9] H. Espeland, and et al, "Transparent Protocol Translation and Load Balancing on a Network Processor in a Media Streaming Scenario," *Network and Operating Systems Support for Digital Audio and Video*, May, 2008.

[10] <http://trace.eas.asu.edu/cgi-bin/main.cgi>