

무선 네트워크에서 QOS를 보증하는 스케줄링 정책

이정환, 박연희, 민승현, 이현정, 신규철, 김명준
충북대학교 컴퓨터 과학과
{cdevil, foreverg, imturtle, junghyun, kiuchool,
mjkim}@rtlab.chungbuk.ac.kr

SCHEDULING SCHEME FOR GUARANTEEING QOS IN WIRELESS/MOBILE NETWORKS

Jung-Hwan Lee, Yeon-Hee Park, Seung-Hyun Min, Hyun-Jung Lee,
Kiu-Cheol, Shin, Myung-Jun Kim
Dept. of Computer Science, Chungbuk National University

요 약

현재 많은 개인용 휴대통신은 지역적 제약에 벗어났다는 편리함을 사용자에게 제공하고 있다. 사용자는 지역적 제약을 벗어난 장점을 이용하여, 화상통신, 인터넷, FAX전송, 실시간 전자거래등과 같은 많은 서비스를 요구하게 되었다. 그러나 무선통신 유선통신에 비하여 대역폭이 매우 작다. 이러한 작은 대역폭을 가지고 사용자들에게 좀더 질이 좋은 서비스를 제공하기 위해서는 실시간성을 요하면서 무선통신에 적은 대역폭을 최대한으로 활용할 수 있는 정책이 필요하다. 본 논문에서는 이러한 요구를 만족하기 위한 새로운 스케줄링 기법을 제시한다. 본 논문에서 제시하는 스케줄링 기법은 데이터를 실시간과 비실시간으로 나눈 뒤 실시간데이터 전송에 사용되고 남은 나머지 대역폭에 가용성을 계산하여 비 실시간 대역폭을 전송하는 방법을 제시한다. 이러한 대역폭에 나머지 가용성을 최대한으로 활용함으로써 무선통신에 대역폭 낭비를 줄일 수 있다.

1. 서론

현재 많은 개인용 휴대통신은 지역적(locality) 제약에 벗어났다는 편리함을 사용자에게 제공하고 있다. 사용자들은 이러한 장점을 넘어서서 더욱더 다양한 서비스를 요구하게 되었다. 이러한 다양한 서비스를 제공하기 위해서는 네트워크에 전송 데이터들 또한 다양해질 수 밖에 없다. 그러나 무선통신(mobile/wireless network)은 유선통신(wired network)보다 적은 대역폭, 폭주제어 문제등 여러 가지 면에서 많은 한계를 가지고 있다.

본 논문에서는 네트워크 전송 데이터를 실시간성을 요하는 비디오, 오디오, 음성같은 실시간 데이터와 실시간을 요하지 않는 ftp, telnet같은 비 실시간 데이터로 나눈다. 무선통신의 적은 대역폭을 최대한으로 활용하고 질 높은 서비스를 제공하기 위해서 실시간성을 요하는 실시간 데이터에 지연한계를 보장하고 실시간성을 요하지 않는 데이터들은 실시간 데이터가 사용하고 남은 대역폭에 가용성을 계산해서 데이터 전송을 해야 한다.

2장에서는 uplink와 downlink에 대해서 자세히 설명한다. 그리고 3장에서는 무선통신 서비스 유형에 대해서 알아본다. 4장에서는 무선 네트워크 실시간 스케줄링 정책을 제시하고 5장에서는 스케줄링 가능성을 검증하기 위한 admission 컨트롤을 제시한다. 6장에서는 비실시간 데이터 전송을 위한 대역폭 가용성 계산을 제시하고 마지막으로 7장에서 결론을 맺는다.

2.Uplink 와 Downlink

Downlink는 기지국(Base Station)에서 mobile(PDA나 PCS 같은 휴대통신)로 데이터를 전송하는 것을 말한다. 이 경우에 실시간 데이터들은 BS에서 스케줄링 후 우선 순위에 따라서 mobile로 데이터를 전송한다. 그러므로 유선통신에서 사용하는 Virtual Clock[5], EDF(Earliest Deadline First)[6]와 같은 실시간 스케줄링 기법을 그대로 적용시킬 수 있다. 본 논문에서는 downlink를 위해서 실시간 데이터 전송은 비선점 EDF를 비실시간 데이터 전송은 FIFO를 채택한다.

Uplink은 mobile에서 기지국(Base Station)으로 데이터를 전송하는 것을 말한다. 이 경우에 기지국은 mobile에 데이터가 있는지 없는지를 알수 없다.[3] 만약 주기적으로 mobile에 실시간 데이터가 있는지 없는지를 확인하지 않고 데이터를 전송한다면 실시간 데이터가 없는 경우에는 그 만큼 빈 데이터를 가지고 오는 것이 되므로 이는 대역폭의 큰 손실이다. 그러므로 본 논문에서는 uplink스케줄링을 위해서 먼저 RMS(Rate-Monotonic Priority Scheduling)[2]로 폴링(polling)을 한 뒤에 데이터가 있을 경우에만 데이터 전송을 하는 비선점 RBOP(RMS based on polling)기법을 사용한다. 비실시간 데이터의 경우에는 실시간 데이터에 사용되고 남은 대역폭에 가용성을 계산하여 전송을 시도한다.

3. 무선통신 서비스 유형

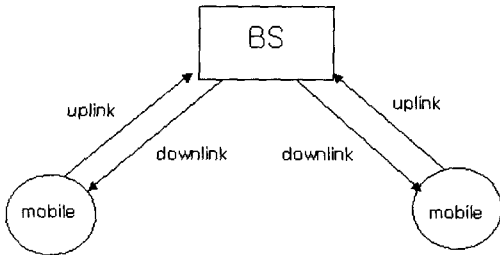


그림 1. mobile-to-mobile

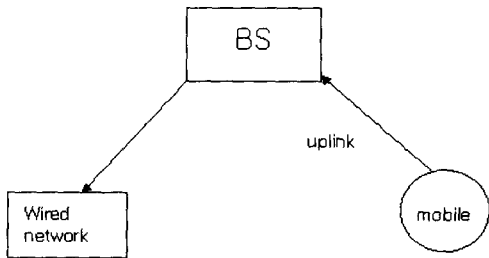


그림 2. mobile-to-wired network

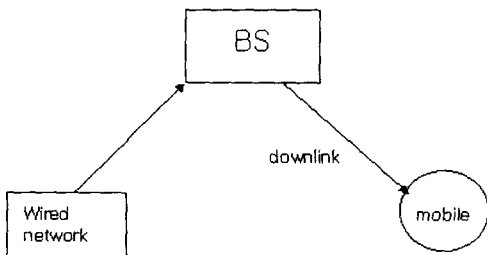


그림 3. wired network-to-mobile

무선통신 서비스 유형은 크게 세 가지로 분류할 수 있다.

(1) mobile-to-mobile 통신(그림 1)

(2) mobile-to-유선 통신(그림 2)

(3) 유선-to-mobile 통신(그림 3)

(4) 기지국에서 일방적으로 mobile에 데이터를 전송

유형(4)는 Downlink만 이루어 지는 형태로서 기존의 유선 네트워크에서 많이 사용되는 EDF스케줄링을 사용한다. 유형(3)에서 네트워크는 유선 통신으로 이루어지므로 기존의 EDF스케줄링을 사용하고 기지국에서 mobile로 데이터를 전송 형태는 유형(4)와 같으므로 EDF스케줄링을 사용한다. 유형(1)과 유형(2)는 uplink를 필요로 하므로 RBOP 스케줄링을 사용하고 Downlink나 기지국과의 유선통신은 EDF를 사용한다. 그러므로 본 논문에서 유형(1)에 대해서 주로 다룰 것이다. 나머지 유형들은 유형(1)에서 사용하는 스케줄링 기법을 그대로 적용시키면 된다.

4. 무선 네트워크 실시간 스케줄링 정책

연결을 설정할 때 입력 트래픽을 위해 일반적으로 많이 사용되는 leaky bucket 모델을 채택하고 leaky bucket 모델은 (a_i, p_i) 로 표현할 수 있다.[1] 버킷에 들어올 수 있는 셀들의 전체 크기는 a_i 로 나타내고 네트워크 진입 속도는 p_i 로 나타낸다.

네트워크에서 패킷은 M 로 나타내고 M 의 크기는 ATM 네트워크에서 셀(Cell)과 같이 고정적이다.

uplink와 downlink 스케줄링 단계

주기 = $T_i = M/P_i$

P_o = 폴링 데이터 크기

(1) mobile은 P_i 의 속도를 요청하고 Admission Test를 거쳐서 BS에 연결을 설정한다.

(2) 기지국은 mobile에 P_i 정보를 가지고 폴링 주기 T_{pi} 를 설정한다.

(3) 폴링순서는 RMS로 정한다.

(4) 기지국이 mobile을 폴링할 때, 패킷 M 이 있다면 mobile에 패킷 M 을 BS로 보낸다.

(5) 기지국의 Packet M 들은 EDF로 우선순위를 재 설정한다.

(6) 우선순위로 기지국에 패킷 M 을 mobile로 downlink시킨다.

(7) 연결이 새로이 설정되거나 삭제될 때, Network에 가용성을 계산해서 비 실시간 데이터를 전송한다.

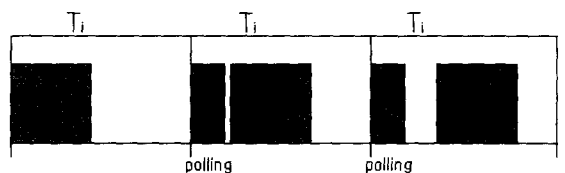


그림 4. uplink 스케줄링

5. Admission control

다른 실시간 채널에 영향을 주지 않으면서 새로운 실시간 채널을 설정하기 위해서 admission test를 거쳐야만 한다.

본 장에서는 Kandlur의 방법[4]의 방법을 이용해서 링크의 스케줄 가능도를 검증한다. 어떤 링크에 N개의 실시간 연결이 공동으로 존재할 때 실시간 연결에 집합은 $\{i, i=1, \dots, N\}$ 으로 나타낸다.

Uplink admission test

C=한 패킷 M의 전송시간

우선순위 인덱스 $i < j$ 이면, $p_i \geq p_j$

$i=1, 2, \dots, N$ 인 경우에,

$$\sum_{k=1}^i C \lceil \frac{(M+p_o)/p_i}{(M+p_o)/p_k} \rceil \leq \frac{M+p_o}{p_i} \tag{1}$$

풀링 시 패킷이 없으면 mobile에서 기지국으로 데이터 전송이 이루어지지 않으므로 기지국에서 mobile로 데이터를 전송하는 downlink를 위해 우선 순위를 재설정 해야 한다 또한 기지국에서 mobile로 일방적으로 데이터를 전송하기 위해 연결을 설정할 수가 있으므로 downlink admission test도 해야만 한다.

Downlink admission test

네트워크 전체 대역폭=Cu

BS에 새로 설정된 연결=i

uplink에 사용되는 연결 인덱스=j

downlink에 사용되는 연결 인덱스=q

풀링시 데이터가 없어서 패킷을 전송하지 않은 연결 인덱스=r

$j=1, 2, \dots, J,$

$q=1, 2, \dots, Q,$

$r=1, 2, \dots, R$ 인 경우에,

$$\sum_{k=1}^i C \lceil \frac{M+p_o}{p_k} \rceil + \sum_{k=1}^q C \lceil \frac{M+p_o}{p_k} \rceil + C \frac{M+p_o}{p_i} \leq C_u + \sum_{k=1}^r C \lceil \frac{M+p_o}{p_k} \rceil \tag{2}$$

6. 무선 네트워크 비 실시간 스케줄링 정책

새로운 연결이 설정되거나 해제될 때 대역폭에 전체 가용성은 변한다. 본 논문에서는 새로운 연결이 설정되거나 해제될 때 변하는 대역폭에 전체 가용성을 계산해서 나머지 가용성을 비실시간 데이터 전송에 이용한다. 비실시간에 사용될 대역폭 가용성 계산은 식(2)를 이용해서 할 수 있다.

전체 대역폭 계산을 RMS나 EDF중 어디에 초점을 맞추는가에 따라서 hard real-time 인지 soft real-time인지가 결정된다.

Hard real-time을 위한 대역폭 가용성 계산

N_r =비주기 가용성

$$\sum_{k=1}^i C \lceil \frac{M+p_o}{p_k} \rceil + \sum_{k=1}^q C \lceil \frac{M+p_o}{p_k} \rceil + N_r \leq I_n \tag{3}$$

$$N_r \leq I_n - \sum_{k=1}^i C \lceil \frac{M+p_o}{p_k} \rceil - \sum_{k=1}^q C \lceil \frac{M+p_o}{p_k} \rceil$$

Soft real-time을 위한 대역폭 가용성 계산

$$\sum_{k=1}^i C \lceil \frac{M+p_o}{p_k} \rceil + \sum_{k=1}^q C \lceil \frac{M+p_o}{p_k} \rceil + N_r \leq C_u + \sum_{k=1}^r C \lceil \frac{M+p_o}{p_k} \rceil$$

$$N_r \leq C_u + \sum_{k=1}^r C \lceil \frac{M+p_o}{p_k} \rceil - \sum_{k=1}^i C \lceil \frac{M+p_o}{p_k} \rceil - \sum_{k=1}^q C \lceil \frac{M+p_o}{p_k} \rceil \tag{4}$$

network이 엄격한 실시간성을 요구한다면 식(3)을 이용해서 대역폭에 나머지 가용성인 N_r 값 만큼을 비주기의 대역폭으로 사용하고 엄격한 실시간성이 아닌 soft real-time 이라면 식 (4)에 N_r 값을 사용할수 있다.

만약 식(4)에 가용성을 이용한다면 네트워크는 대역폭을 모두 사용할 것이다.

7. 결론 및 향후 연구 과제

본 논문에서 제시 하는 스케줄링 정책은 무선통신에서 실시간 데이터에 지연한계를 보장하면서 실시간에 사용되고 남은 가용대역폭을 비 실시간 데이터 전송에 활용함으로써 서비스의 질을 높이고 대역폭 또한 최대로 활용할 수 있다. 그러나 폭주제어나 핸드오프처리등 무선통신에서 중요한 많은 제약 사항등을 고려하지 않았다. 그러므로 앞으로는 무선통신에 폭주제어나 핸드오프처리등을 고려한 스케줄링 정책 연구가 필요하다.

참고 문헌

[1]E. Knightly, D. Wrege, J.Liebeherr, and H. Zhang "Fundamental limits and tradeoffs of providing deterministic guarantees to VBR video traffic," in Proc. of ACM SIGMETRICS, pp. 98-107, 1995

[2]Liu, C. L. and Layland J. W. Scheduling Algorithms for Multiprogramming in a Hard Real Time Environment J. ACM 20(1):46-61, 1973.

[3]Sunghyun Choi and Kang G. Shin, "A Comparative Study of Bandwidth Reservation and Admission Control Schemes in QoS-Sensitive Cellular Networks," to appear in ACM Wireless Networks, fourth quarter, 2000.

[4]D. D. Kandlur, K. G. Shin, and D. Ferrari, "Real-time communication in multi-hop networks," in Proc. 11-th Int'l Conf. Distributed Comput. Systems, pp. 300-307, May 1991.

[5]Lixia Zhang. "A New Architecture for Packet Switched Network Protocols", PhD dissertation, Massachusetts Institute of Technology, July 1989.

[6]Chengzhi Li, Riccardo Bettati, Wei Zhao. "static Priority Scheduling for ATM Networks". In Proc. of the Real-Time Systems Symposium, San Francisco, CA, December 1997.