

LTE 시스템에서 지연에 대한 QoS 보장을 위한 하향링크 패킷 스케줄링 알고리즘

준회원 최 범 곤*, 종신회원 정 민 영*, 정회원 이 혜 경**, 김 태 석**, 강 지 웅**

Delay-Aware Packet Scheduling (DAPS) Algorithm in 3GPP LTE System

Bum-Gon Choi* *Associate Member*, Min Young Chung*° *Lifelong Member*,
Hye Kyung Lee**, Tai Suk Kim**, JeeWoong Kang** *Regular Members*

요 약

3GPP(3rd Generation Partnership Project)에서 표준화 중인 LTE(Long Term Evolution) 시스템은 고품질의 다양한 서비스를 제공하기 위한 차세대 이동통신 기술로써, 최근 LTE 시스템의 무선자원 사용효율을 향상시키기 위한 하향링크 스케줄링 기법 연구가 활발하게 수행되고 있다. 본 논문에서는 무선 채널 상태나 공정성뿐만 아니라 다양한 트래픽 종류의 지연에 대한 QoS 요구사항을 보장하기 위해서 지연 가중치(delay weight) 함수를 적용한 DAPS(Delay-Aware Packet Scheduling) 알고리즘을 제안한다. DAPS 알고리즘은 패킷의 큐잉 지연이 허용되는 최대 지연에 가까울수록 지연 가중치를 증가시켜 높은 우선순위를 제공함으로써 최대 허용 지연 이내에 처리되는 패킷의 수를 증가시킨다. 모의실험 결과 DAPS 알고리즘은 기존 스케줄링 알고리즘과 비교하여 전송되는 패킷의 큐잉 지연이 최대 허용되는 지연을 초과할 확률을 최소화하여 패킷 지연에 관한 성능을 개선하였다.

Key Words : LTE, QoS, Downlink Packet Scheduling, PF, M-LWDF

ABSTRACT

In wireless mobile environments, large bandwidth and high QoS (Quality of Service) are recently required to support the increased demands for multimedia services. LTE (Long Term Evolution) is one of the promising solutions for the next generation broadband wireless access systems. To efficiently use downlink resource and effectively support QoS, packet scheduling algorithm is one of the important features in LTE system. In this paper, we proposed DAPS (Delay-Aware Packet Scheduling) algorithm to consider QoS requirements of delays for various traffic classes as well as channel condition and fairness. To reflect delay experiences at scheduling instance, DAPS observes how queue waiting time of packet is closed to maximum allowable delay. The simulation results show that the DAPS algorithm yields better performance for delay experience by increasing the number of transmitted packets with satisfying the required delay time compared with existing scheduling algorithms.

* 본 연구는 삼성전자 및 지식경제부 및 정보통신연구진흥원의 대학 IT연구센터 지원사업(IITA-2009-C1090-0902-0005)의 연구결과로 수행되었습니다.

* 성균관대학교 정보통신공학부 ({gonace, mychung}@ece.skku.ac.kr) (° : 교신저자)

** 삼성전자 DMC연구소 ({hk1103.lee, kts71, jw75.kang}@samsung.com)

논문번호 : KICS2009-02-042, 접수일자 : 2009년 2월 3일, 최종논문접수일자 : 2009년 5월 6일

I. 서 론

WCDMA(Wideband Code Division Multiple Access) 기반의 3세대 UMTS 시스템은 패킷 전송 속도를 향상시키고, 시스템의 처리율을 개선하는 목적을 두고 하향링크 고속 패킷 전송 기술인 HSDPA(High Speed Downlink Packet Access), 상향링크 고속 패킷 전송 기술인 HSUPA(High Speed Uplink Packet Access) 기술로 진화해왔다^{[1][3]}. 하지만 기존 시스템을 유지하면서 일부 기능을 개선하는 단기적인 진화는 서비스 개선 폭에 한계가 있기 때문에 보다 빠른 속도로 좋은 품질의 서비스를 제공할 수 있는 기술의 필요성이 대두되었으며, 이를 위해 3GPP(3rd Generation Partnership Project)에서는 LTE(Long Term Evolution) 표준화 작업을 진행 중이다. LTE 시스템은 커버리지 확장 및 시스템 용량 개선뿐만 아니라 낮은 전송 지연과 높은 전송률 지원을 통해 사용자와 사업자의 비용 절감과 서비스의 품질 개선을 목표로 하고 있다^{[1][2]}.

LTE는 기존 WCDMA/HSPA(High Speed Packet Access)와 다른 네트워크 구조를 갖는다^[4]. WCDMA/HSPA의 RAN(Radio Access Network)은 NB(NodeB)와 RNC(Radio Network Controller)로 구성되어 있다. NB는 하나 또는 다수의 셀에서 UE(User Equipment)와의 데이터 및 제어 신호의 송수신 역할을 담당하며, RNC는 RAN을 코어 네트워크와 연결하는 역할을 한다. 반면에 LTE에서는 NB를 관리하는 RNC를 생략하고 eNB(enhanced NB)만으로 RAN를 구성한다. eNB는 NB와 동일한 기능을 수행 할뿐만 아니라 WCDMA/HSPA의 RNC에서 제공되던 대부분의 기능을 지원한다. 이처럼 LTE는 WCDMA/HSPA에 비해 간소화된 네트워크 구조를 통해 시그널링과 데이터 전송에서 발생하는 지연을 최소화한다.

다양한 대용량 멀티미디어 서비스를 제공하기 위해 LTE에서는 OFDMA(Orthogonal Frequency Division Multiple Access)방식을 물리계층 전송방식으로 채택하였다^[4]. OFDM은 직렬로 입력되는 데이터 열을 N개의 병렬 데이터 열로 변환하여 각각 분리된 부반송파에 실어 전송함으로써 데이터 전송률을 높이는 방식이고, OFDMA는 다수의 사용자가 동시에 서비스를 받을 수 있도록 고려한 방식이다. LTE에 적용된 OFDMA는 PRB(Physical Resource Block)를 최소 자원 할당 단위로 정의하여, 기지국이 현재 시스템의 필요 요건에 따라 대역폭 사용을 동적으

로 조정 가능하도록 한다^{[1][4]}.

무선 통신 환경에서 사용자의 QoS(Quality of Service)를 제공하기 위해서는 무선 자원 관리(RRM: Radio Resource Management)가 매우 중요하다. 특히 RRM의 기능 중 스케줄링 기능은 각각의 서비스에 대한 QoS 보장과 밀접한 관련이 있다. LTE에서는 UE로부터 보고받은 채널 상태 정보를 사용하여 신속하고 효율적인 무선 자원 할당을 위해 MAC(Medium Access Control) 계층에서 스케줄링 기능을 정의하고 있다^[5]. 무선 자원을 효율적으로 사용하고, 사용자에게 공정하게 자원을 할당하며, 수락된 호에 대해서는 서비스 요구사항을 만족시키기 위한 것이 스케줄러의 역할이다. 이를 위해 다양한 하향링크 스케줄링 기법이 연구되고 있다.

LTE 시스템에서의 MAC 스케줄러는 시간과 위치에 따라 변하는 무선 채널 특성, 사용자 간의 공평성, 서비스의 QoS 요구사항 등을 고려하여 자원을 할당한다. 무선 채널 상태가 좋은 사용자에게 높은 전송기회를 부여할 경우 시스템 효율을 높일 수 있고, 평균 지연을 낮출 수 있다. 그러나 채널 상태가 좋은 사용자만 스케줄러에 의해 선택될 가능성이 높아지기 때문에 사용자 간의 자원 사용에 대한 공정성이 보장되기 어려울 수 있다. 따라서 이를 위해 스케줄러 정책 결정은 무선채널 특성, QoS 요구사항 뿐만 아니라 공정성 측면이 고려되어야 한다.

본 논문에서는 채널 상태와 공정성뿐만 아니라 다양한 트래픽 종류의 지연에 대한 QoS 요구사항을 보장하기 위해서 지연 가중치(delay weight) 함수를 적용한 DAPS(Delay-Aware Packet Scheduling) 알고리즘을 제안한다. DAPS 알고리즘은 패킷의 큐잉 지연이 허용되는 최대 지연에 가까울수록 지연 가중치를 증가시켜 높은 우선순위를 제공함으로써 최대 허용 지연을 초과하지 않고 처리되는 패킷의 수를 증가시킨다. 모의실험 결과 DAPS 알고리즘은 기 제안된 PF, M-LWDF 알고리즘과 비교하여 패킷 지연이 최대 허용 지연을 초과할 확률을 최소화하여 패킷 지연에 관한 성능을 개선하였다.

본 논문은 다음과 같이 구성되어 있다. 서론에 이어 2장에서는 LTE 하향링크 스케줄러 모델의 구조에 대해 서술하고, 3장에서는 기존 무선통신에서의 스케줄링 알고리즘에 대해 소개한다. 4장에서는 지연에 관한 QoS를 보장하기 위해 제안하는 DAPS 알고리즘을 설명한다. 5장에서는 제안 알고리즘의 성능평가를 위한 시뮬레이션 환경 및 성능평가 결과에 대해 서술하고, 마지막으로 6장에서 결론을 맺는다.

II. 하향링크 스케줄러 모델

패킷 스케줄러의 역할은 수락된 호의 서비스 특성과 UE의 채널 상태 정보를 고려하여 가용 자원을 사용자에게 효율적이고 공정하게 할당하기 위한 것이다. 즉, 스케줄러는 시간에 따라 변화하는 무선 채널을 통해 음성, 데이터, 비디오 등의 서비스를 포함하는 광대역 무선 서비스를 효율적으로 제공해야 한다. 그림1은 하나의 eNB에 속한 UE의 수가 $N(0 \leq i < N)$, eNB에서 지원하는 트래픽의 종류의 수가 $M(0 \leq j < M)$ 개일 때 하향링크 스케줄러 모델의 구조를 나타낸다. eNB에는 N개의 UE와 M개의 트래픽 종류를 위한 논리적인 큐가 존재하고, 하향링크 트래픽은 패킷 classifier에 의해 분류되어 해당 큐에 적재된다. 스케줄러는 지원하는 스케줄링 정책을 기반으로 1ms의 TTI(Transmission Time Interval)마다 패킷의 전송 우선순위를 결정한다^[5]. eNB는 결정된 우선순위로 패킷을 LTE 시스템의 물리계층 전송 단위인 PRB에 할당하여 패킷 전송을 수행한다^{[1][4]}.

트래픽 종류에 따른 서비스 요구사항은 크게 지연, 패킷 손실률, 대역폭의 세 가지로 분류할 수 있다. 인터넷 전화, 원격회의, 인터랙티브 게임(interactive game)과 같이 상호작용하는 RT(Real Time)서비스는 효과적인 데이터 전송을 위해 엄격한 시간 조건을 요구한다. 전자메일, 파일 및 웹 문서 전송, e-commerce와 같은 서비스는 신뢰적인 데이터 전송, 즉 데이터 손실이 없는 전송을 요구한다. RT 오디오/비디오와 같은 서비스는 어느 정도의 데이터 손실을 허용하지만, 파일 데이터 또는 재무 거래 데이터는 패킷 손실률에 제약을 받는다. 대역폭에 민감한 서비스는 일정량의 대역폭을 요구하는 반면 전자메일, 파일전송 등과 같은 융통성 있는 서비스는 가용 대역폭을 적게 사용할 수 있다. 3GPP TS 23.203에서는 패킷을 지연, 데이터 전송률, 패킷 손

실량에 따라 9가지의 QCI(QoS Class Identifier)로 분류한다^[6]. eNB는 코어 네트워크에서 들어오는 패킷을 분류된 9가지의 QCI에 따라 사용자의 트래픽 종류별로 구분된 논리 큐에 적재한다.

LTE 하향링크 스케줄러는 데이터 송신을 위해 무선 자원을 1ms의 빠른 주기로 할당하는 동적(dynamic) 스케줄링 방식을 기본으로 사용하고, VoIP(Voice over Internet Protocol) 서비스의 효율적 지원을 위해 SPS(Semi-Persistent Scheduling) 모드를 제공한다^[5]. UE가 주기적으로 피드백(feedback) 한 채널 상태 정보는 스케줄링 결정시 반영되는 중요한 요소 중 하나이다. 하나의 PRB에는 UE가 eNB에게 채널 상태를 보고하기 위해 사용하는 4개의 레퍼런스 심볼이 삽입되어 있으며, UE는 다수의 레퍼런스 심볼의 정보를 이용하여 주기적으로 eNB에게 채널 상태 정보인 CQI(Chanel Quality Indicator)를 보고한다^{[1][4][5]}.

스케줄러는 UE의 채널 상태 정보를 포함한 스케줄링 정책에 고려되는 요소들을 반영하여 패킷의 우선순위를 결정하고, 결정된 우선순위로 HOL(Head-of-Line) 패킷을 PRB에 할당한다. PRB는 OFDMA 기반 LTE 시스템의 물리계층 전송 단위로서 하나의 슬롯 동안 12개의 부반송파로 구성된다^{[1][4]}. 각 슬롯은 6 또는 7개의 OFDM 심볼들로 이루어져 있으며, 각 심볼은 CP(Cyclic Prefix)를 포함하여 다중 경로 페이딩에 의해 발생하는 채널 간 간섭을 완화한다. DC(Direct Current) 부반송파를 포함하여 전송되는 하향링크 부반송파의 총 개수 N_{sc} 는 $12 \cdot N_{RB} + 1$ 과 같으며, 여기서 N_{RB} 는 PRB의 개수를 의미한다. LTE 물리계층 표준에서는 N_{RB} 를 6개부터 100개까지 동적으로 구성할 수 있도록 정의하고 있어 하향링크 전송 대역폭은 1MHz 부터 20MHz까지 적용 가능하다^{[2][4]}.

III. 무선통신에서의 하향링크 스케줄링 알고리즘

스케줄러는 사용자에게 공정하게 자원을 할당하되 수락된 호의 서비스 요구사항을 만족시키기 위해 무선 자원을 효율적으로 분배하는 역할을 수행한다. 이를 위해 다양한 하향링크 스케줄링 알고리즘이 연구되고 있으며, LTE 시스템에서도 효율적인 전송 제어를 위해 채널 상태와 서비스의 QoS를 만족시키기 위한 스케줄링 연구가 수행되고 있다. 스케줄링 알고리즘은 크게 다음과 같은 요소로 구분

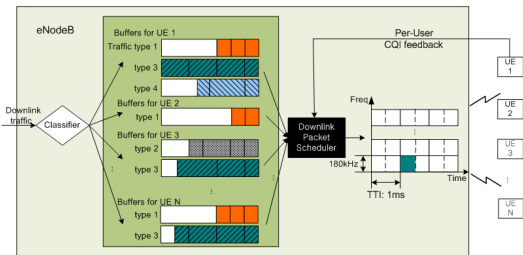


그림 1. LTE 하향링크 스케줄러 모델

할 수 있다. 첫 번째는 스케줄러가 지원하는 서비스 종류이다. 제안된 스케줄링 정책이 RT 트래픽의 최대 허용되는 지연을 만족시키는지의 여부, NRT(Non-RT) 트래픽이 요구하는 최소 데이터 전송률을 만족시키는지의 여부, 또는 RT와 NRT의 QoS를 동시에 보장해줄 수 있는지의 여부로 구분된다. 두 번째는 무선 채널 상태에 대한 고려이다.

무선 채널 상태를 반영하지 않고 스케줄링 하는 대표적인 알고리즘으로는 RR(Round Robin)이 있고, 시간에 따라 변하는 무선 채널 상태를 반영하여 시스템 효율을 높이는 대표적인 알고리즘으로는 Max C/I(Maximum Carrier-to-Interference Ratio)가 있다³¹. 그러나 좋은 채널 상태를 갖는 UE만을 우선적으로 고려할 경우 무선 자원 점유의 공정성(fairness)이 떨어지기 때문에 사용자 간의 공정성이 보장되는지에 대한 고려도 스케줄링 정책을 결정하는 중요한 요소로 사용된다.

무선채널 환경과 사용자 간의 공정성을 보장하기 위한 대표적인 스케줄링 알고리즘으로는 PF를 들 수 있다^{31,71}. PF 알고리즘은 모든 UE $i(0 \leq i < N)$ 의 트래픽 $j(0 \leq j < M)$ 에 대하여 n 번째 스케줄링 시점까지의 평균 전송률($DR_{ij}[n]$)과 n 번째 스케줄링 시점에서의 UE i 의 현재 채널 상태($CR_i[n]$)의 비율을 계산한다. PF는 UE i 의 트래픽 $j(UE_i: 0 \leq i < N, 0 \leq j < M)$ 에 대한 $DR_{ij}[n]$ 를 갱신하기 위해 식 (1)을 사용한다.

$$DR_{ij}[n+1] = (1-w) \cdot DR_{ij}[n] + \Delta_{ij} \cdot w \cdot CR_i[n] \quad (1)$$

만약 UE_{ij} 가 전송 기회를 부여 받았다면 Δ_{ij} 는 1이 되고, 그렇지 않다면 Δ_{ij} 는 0이 된다. w 는 $DR_{ij}[n]$ 을 측정하기 위한 윈도우 개수의 역수를 나타내며 일반적으로 1/1000이 사용된다⁷¹. PF알고리즘은 UE 간의 공정성을 보장하기 위해 제안되었지만 모든 UE가 비슷한 채널 상태를 갖는 경우에만 공정한 자원할당이 가능하다³¹. 또한 RT 트래픽에 대한 지연 요구사항을 고려하지 않기 때문에 RT 트래픽의 QoS 제약사항을 만족시키기 어렵다.

RT 트래픽의 QoS를 보장하기 위하여 M-LWDF (Modified-Longest Weighted Delay First)과 EXP rule 알고리즘이 제안되었다^{18,91}. M-LWDF 알고리즘은 두 가지 조건에 대해 QoS를 정의하고 있다. 한 가지는 지연에 관한 것으로 다음과 같다.

$$Prob\{W_{queue}^{ij}[n] > D_{MAX}^j\} \leq \delta_{ij} \quad (2)$$

여기서 $W_{queue}^{ij}[n]$ 는 n 번째 스케줄링 시점에서 UE_{ij} 의 HOL 패킷의 큐잉 지연, D_{MAX}^j 은 트래픽 종류 j 의 최대 허용되는 지연, δ_{ij} 는 아웃지(outage) 확률에 대한 임계 값을 나타낸다. 다른 한 가지는 수율에 관한 정의로 식 (3)처럼 $DR_{ij}[n]$ 은 최소 보장돼야 하는 데이터 전송률(DR_{MIN}^j)보다 커야 함을 나타낸다.

$$DR_{ij}[n] \geq DR_{MIN}^j \quad (3)$$

M-LWDF 스케줄러는 스케줄링 시점에서 $PV_{M-LWDF}^{ij}[n]$ 가 가장 큰 UE_{ij} 의 순서대로 UE를 정렬하고, 정렬된 순서대로 PRB에 해당 UE의 HOL 패킷을 할당하여 전송한다. n 번째 스케줄링 시점에서의 $PV_{M-LWDF}^{ij}[n]$ 은 다음과 같이 결정된다.

$$PV_{M-LWDF}^{ij}[n] = \gamma_{ij}[n] \cdot W_{queue}^{ij}[n] \cdot CR_i[n] \quad (4)$$

여기서 γ_{ij} 는 임의의 양수으로써 식 (2)와 (3)에서 정의된 두 가지 QoS 요구사항을 만족시키기 위해 사용된다. 식 (2)에서 정의된 지연에 대한 QoS를 만족시키기 위해서 M-LWDF에서는 γ_{ij} 를 $a_{ij}/\overline{CR_i[n]}$ 로 정의한다. 여기서 $\overline{CR_i[n]}$ 은 n 번째 TTI까지 UE_{ij} 의 평균 채널 상태를 의미하고, a_{ij} 는 $-(\log \delta_{ij})/D_{MAX}^j$ 이며 서로 다른 QoS 요구사항을 반영하기 위해 적용되었다. 즉, M-LWDF 알고리즘은 패킷 지연 또는 큐 길이, 채널 상태, QoS 제약사항을 반영하여 전송 순서를 결정하는 방식이다. 그러나 다양한 종류의 트래픽이 혼재된 환경에서 각 트래픽 별로 상이한 QoS 요구사항을 반영하기 어렵다는 한계를 갖는다.

EXP rule 알고리즘은 RT 서비스의 지연에 대한 성능을 향상시키기 위해 제안되었으며, 식 (5)와 같이 정의된다.

$$PV_{EXP}^{ij}[n] = \exp\left(\frac{a_i W_{queue}^{ij}[n] - a W_{queue}^j}{1 + \sqrt{a W_{queue}^j}}\right) \cdot \frac{CR_i[n]}{CR_i[n]} \quad (5)$$

이때,

$$\overline{aW_{queue}^j} = \frac{1}{N} \sum_{i=0}^{N-1} a_i W_{queue}^{*j} [n] \quad (6)$$

이며, a_i 는 서비스간 가중치 부여를 위한 임의의 상수이다. EXP rule 알고리즘은 지연이 작은 경우에는 PF와 같이 동작하고 지연이 큰 경우에는 지연요구 값 대비 상대적인 지연이 큰 트래픽에 높은 우선순위를 부여한다. 그러나 M-LWDF와 같이 입력 트래픽이 과도하게 유입되는 경우 모든 트래픽의 지연요구가 동시에 나빠지는 특성이 있다.

ATAS(Adaptive Traffic Allocation Scheduling)와 UEPS(Urgency- and Efficiency- based Packet Scheduling) 알고리즘은 RT 트래픽의 QoS 제약사항뿐만 아니라 NRT 트래픽의 수율을 향상시키기 위해 제안되었다^{[10][11]}. ATAS 알고리즘은 RT 및 NRT 패킷을 D_{MAX}^j 에 따라 여러 개의 그룹으로 나누고, 각 그룹은 D_{MAX}^j 이 작은 순서대로 높은 우선순위를 부여한다. 임의의 한 그룹의 데이터 전송률이 DR_{MIN}^j 을 만족하는 경우 다음 우선순위를 갖는 그룹에게 전송기회를 보장해준다. ATAS 알고리즘에서는 각 그룹의 트래픽 종류에 따라서 서로 다른 스케줄링 정책을 사용한다. RT 트래픽 그룹의 경우 RR 알고리즘으로 전송 우선순위를 결정하고, NRT 트래픽 그룹의 경우 PF 알고리즘을 사용하여 우선순위를 결정한다. ATAS는 우선순위가 높은 RT 트래픽 그룹의 수율이 일정량 이상을 만족하는 경우에만 NRT 트래픽 그룹에 전송기회가 부여되기 때문에 RT 트래픽 부하가 과도하게 발생하는 경우 NRT 트래픽의 QoS 보장이 어렵다.

UEPS 알고리즘은 스케줄링 시점에 두 가지 요소를 고려한다. 하나는 무선 자원 사용에 대한 효율성이고 다른 하나는 각 패킷의 지연에 대한 긴박성(urgency)에 관한 요소이다. 무선 자원의 효율성을 위해서 UEPS 알고리즘은 PF 알고리즘처럼 평균 데이터 전송률에 대한 스케줄링 시점에서의 현재 전송 가능한 데이터율의 비율을 사용한다. 패킷 지연에 대한 긴박성은 time-utility 함수를 사용하여 결정한다. 각 패킷은 최대 허용되는 지연 시간에 대한 QoS 요구사항을 만족시키기 위해 스케줄링 받아야 하는 최대 시점을 deadline으로 정의하고 RT 트래픽이 deadline에 가까운 경우 우선적으로 스케줄링 된다. 그 외의 경우에는 NRT 트래픽이 높은 전송 기회를 부여 받게 된다. 즉, 무선 자원 사용에

대한 값이 같고 RT 트래픽의 긴박성이 NRT 트래픽 보다 작은 경우 UEPS 스케줄러는 NRT 트래픽을 전송한다. UEPS의 경우 스케줄링 시점마다 각 패킷의 긴박성을 계산해야 하기 때문에 이에 따른 복잡도가 증가한다. 따라서 시스템의 복잡도를 크게 증가시키지 않도록 적은 매트릭으로 다양한 종류의 트래픽의 QoS 요구사항을 보장할 수 있는 MAC 스케줄링 알고리즘의 개발이 필요하다.

IV. Delay-Aware Packet Scheduling (DAPS) 알고리즘

시스템 효율을 높이기 위한 방법으로 무선 채널 상태가 좋은 UE에게 높은 전송기회를 부여해 주기 위해서 채널 상태를 고려하는 것은 우선순위 결정의 중요한 요소 중 하나이다. LTE에서는 UE의 채널 상태 정보에 기반한 MCS(Modulation and Coding Scheme)에 따라 데이터 전송률이 달라기 때문에 무선 채널 상태가 좋을수록 하나의 PRB에 더 많은 데이터를 할당하여 전송할 수 있게 된다. 따라서 스케줄링 요소로 채널 상태를 고려하는 것은 시스템 효율을 높일 뿐 아니라 평균 지연을 줄일 수 있다. 그러나 채널 상태만을 고려할 경우 채널 상태가 좋은 UE만 스케줄러에 의해 선택될 가능성이 높아지기 때문에 사용자 간의 공정성이 보장되기 어렵다. 이를 위해 스케줄링 결정 요소로 공정성 측면이 고려되어야 한다.

3GPP TS 23.203에서는 RT 뿐만 아니라 NRT 트래픽의 최대 허용되는 지연을 정의하고 있다^[6]. DAPS 알고리즘은 채널상태, 공정성뿐만 아니라 QoS 제약사

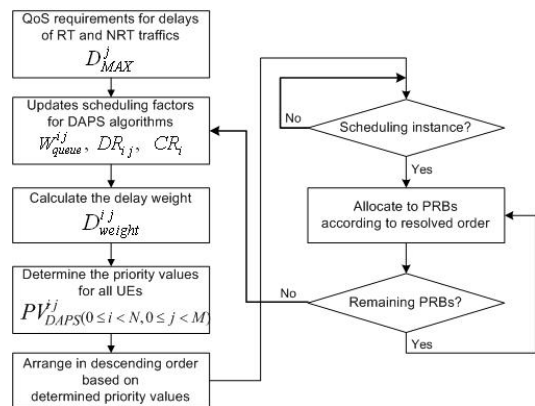


그림 2. DAPS 알고리즘의 순서도

항 중 허용되는 최대 지연($D_{MAX}^j(0 \leq j < M)$)을 보장하기 위한 방식이며 그림2는 DAPS 알고리즘의 순서도를 나타낸다. DAPS 알고리즘에서는 다양한 트래픽 종류의 지연에 대한 QoS 요구사항을 보장하기 위해 지연 가중치(delay weight) 함수를 적용하며 지연 가중치는 패킷의 큐잉 지연이 최대 허용되는 지연에 얼마나 가까운가를 반영한다. 스케줄링 시점에서 패킷의 지연 경험을 반영하는 지연 가중치 함수는 식 (7)과 같이 정의된다.

$$D_{weight}^{ij}[n] = \exp\left(-\left|1 - \frac{W_{queue}^{ij}[n]}{D_{MAX}^j}\right|\right) \quad (7)$$

DAPS 알고리즘에서는 정의된 지연 가중치와 채널 상태, 평균 데이터 전송률을 반영하여 HOL 패킷들의 우선순위 값을 식 (8)과 같이 계산한다.

$$PV_{DAPS}^{ij}[n] = \frac{CR_i[n]}{DR_{ij}[n]} \cdot D_{weight}^{ij}[n] \quad (8)$$

DAPS 알고리즘은 계산된 우선순위 값에 따라 UE를 내림차순으로 정렬하고, 정렬된 순서대로 TTI 동안 사용될 수 있는 PRB에 해당 UE의 HOL 패킷들을 할당한다.

그림 3은 임의의 트래픽의 최대 허용 지연이 100ms일 때 지연 가중치 함수의 예를 나타낸다. 지연 가중치는 패킷의 큐잉 지연이 허용되는 최대 지연인 100ms에 가까워질수록 1까지 증가하고, 큐잉 지연이 허용되는 최대 지연보다 커지게 되면 시간이 지남에 따라 감소한다. 즉, 허용되는 최대 지연과 패킷의 큐잉 지연의 값이 가까울수록 지연 가중치는 증가한다. 따라서 DAPS 알고리즘은 패킷의 전송이 최대 허용 지연 이내에 이루어질 확률을 증가시킨다.

UE₁과 UE₂가 각각 RT와 NRT 트래픽을 서비스 중이라고 가정한다. UE_{1,RT}와 UE_{2,NRT}의 현재 채널

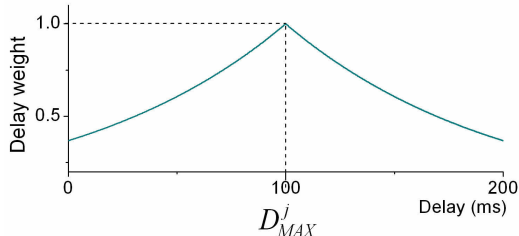


그림 3. 최대 허용 지연이 100ms일 때 지연 가중치 함수의 예

상태와 평균 데이터 전송률이 같다면 UE_{1,RT}와 UE_{2,NRT}의 우선순위 값은 지연 가중치에 의해 결정된다. 만약 UE_{1,RT}의 HOL 패킷이 큐에서 기다린 시간이 최대 허용 지연 시간의 1/10이고, UE_{2,NRT}의 HOL 패킷이 큐에서 기다린 시간이 최대 허용 지연 시간과 같다면 UE_{1,RT}와 UE_{2,NRT}의 지연 가중치는 각각 0.4와 1이 된다. 따라서 UE_{2,NRT}는 UE_{1,RT}보다 더 높은 우선순위를 갖는다. DAPS 알고리즘은 지연 가중치를 적용함으로써 RT 트래픽만 전송 우선순위를 갖는 것이 아니라 NRT 트래픽도 지연에 대한 요구사항을 충족시키기 위해 높은 우선순위로 처리될 수 있음을 나타낸다.

기존의 패킷 지연을 고려한 스케줄링 알고리즘들은 패킷의 지연이 커질수록 높은 가중치를 부여하는 정책을 사용하였다. 그러나 네트워크에 과부하가 걸렸을 경우에 이러한 정책을 사용한다면, 해당 네트워크를 사용하는 모든 사용자들의 서비스 품질이 저하되는 문제점이 발생한다. 그러나 DAPS 알고리즘은 해당 패킷의 큐잉 지연이 D_{MAX}^j 를 초과하였을 경우 우선순위 결정을 위한 가중치를 감소시켜 해당 사용자의 스케줄링 우선순위를 낮추어 지연요구 조건을 만족하는 다른 사용자의 트래픽을 처리해 주므로, 기존의 스케줄링 알고리즘과 비슷한 수율 및 지연 특성을 제공하면서도 서비스 품질의 저하를 경험하는 사용자 수를 최소화할 수 있다.

V. 성능평가

DAPS 알고리즘의 성능 검증을 위한 모의실험을 위해 C언어를 이용하여 몬테 카를로 이벤트 드리븐 방식의 3GPP LTE 하향링크 스케줄러를 모델링 하였다. 표 1는 DAPS 알고리즘의 성능 검증을 위한 시뮬레이션 환경을 나타낸다. 셀의 형태는 한 변의 길이가 250m인 정육각형으로 가정하였고, 셀에서는 전 방향 안테나를 사용한다. 스케줄러는 1ms 마다 동작하며, 각 사용자의 트래픽 종류 별로 10Mbytes의 버퍼 크기를 갖는다. 중심 주파수는 2.3GHz, 전송 대역폭은 5MHz로 TTI동안 전송되는 총 PRB의 수는 50개이다. PRB에는 레퍼런스 심볼과 CP를 제외하고 데이터 심볼만이 할당된다고 가정했다. PRB는 시간 평면에서 한 쌍이 최소 단위로 할당되기 때문에 하나의 TTI 동안 최대 25대의 UE까지 동시에 수용 가능하다. 무선 채널 환경은 인접 셀 간섭은 고려하지 않고, 다중 경로 페이딩과 전파 경로

손실만을 가정하였다. 전파 경로 손실은 modified COST231-Hata 모델을 사용하였다^[12]. UE의 이동 패턴을 모델링하기 위해서 랜덤 워크 모델을 사용하였으며 랜덤 워크 모델에서 각 UE는 이동 속도, 방향, 이동 시간을 결정하여 다음 위치로 이동한다. 모든 UE의 이동 속도는 4km/h로 설정하였고, 방향은 0도와 360도 사이에서, 이동 시간은 10초와 20초 사이에서 균등 분포로 선택된다. 생성되는 트래픽은 FTP, web browsing, video streaming, VoIP의 4가지 종류의 트래픽 모델을 가정하였고, VoIP 서비스도 동적 스케줄링 기법을 적용된다고 가정하였다^[12]. 액세스 네트워크에서 최대 허용 지연과 패킷 손실률은 3GPP TS 23.320에서 정의한 QCI를 바탕으로 FTP와 web browsing은 100ms와 10^{-6} , video streaming은 50ms와 10^{-3} , VoIP는 30ms와 10^{-2} 을 기준으로 하였다^[6]. 성능평가는 Video와 VoIP의 사용자 수는 각각 50으로 고정시키고, FTP와 web browsing의 사용자 수(n)를 각각 4대에서 40대까지 4씩 증가시키며 실시하였다. 스케줄러의 성능을 평가하기 위해서 CAC는 고려하지 않았으며, 제안된 DAPS 알고리즘을 PF, M-LWDF 스케줄링 알고리즘과 비교하였다. 시뮬레이션은 10,000초 동안 수행하였고, 처음 2,000초 동안의 결과를 제외한 8,000초 동안의 성능 수치를 기록하였다.

그림 4는 PF, M-LWDF, DAPS 알고리즘에서의 평균 데이터 전송률을 나타내고, 그림 5는 평균

표 1. 시뮬레이션 환경

항목	값	비고
셀의 수	1개	-
UE 수	x:x:50:50	FTP: Web: Video: VoIP (x는 4대에서 40대까지 4씩 증가시킴)
사용 대역	5MHz	TTI당 PRB 50개 사용 가능 (PRB는 2의 배수 단위로 할당)
시뮬레이션 시간	10,000sec	0~2,000 초 데이터는 반영 안 함
큐 길이	10Mbytes	UE의 서비스 별
평균 서비스 지속 시간	180sec	지수분포, 하나의 UE는 동시에 하나의 서비스만 사용할 수 있음
TTI	1msec	-
유선망 데이터 전송속도	1Mbps/UE	-
이동성 모델	Random walk	이동속도: 4km/h 이동시간: Uniform[10, 20] 이동방향: Uniform[0, 360]

PRB 사용률을 나타낸다. 평균 데이터 전송률은 초당 전송된 평균 비트 수, PRB 사용률은 매 TTI 동안 사용된 평균 PRB 개수를 측정하였다. 평균 데이터 전송률과 PRB 사용률은 서로 다른 스케줄링 정책 보다는 생성되는 트래픽 양이나 UE의 이동 속도에 따른 채널 상태 변화에 영향을 받기 때문에 PF, M-LWDF, DAPS 알고리즘이 거의 비슷한 결과를 보임을 알 수 있다.

그림6의 (a), (b)는 PF, M-LWDF, DAPS 알고리즘에서의 평균 지연을 나타낸 그래프이다. NRT 트래픽의 경우 FTP의 평균 지연은 PF, M-LWDF, DAPS 알고리즘이 각각 $n \geq 20$, $n \geq 20$, $n \geq 16$ 일 때 최대 허용 지연인 100ms를 초과하고, web browsing의 평균 지연은 각각 $n \geq 28$, $n \geq 20$, $n \geq 24$ 일 때 100ms를 초과하는 결과를 보인다. RT 트래픽의 경우 PF 알고리즘은 NRT 트래픽 UE 수에 상관없이 VoIP 서비스의 지연에 대한 QoS를 만족

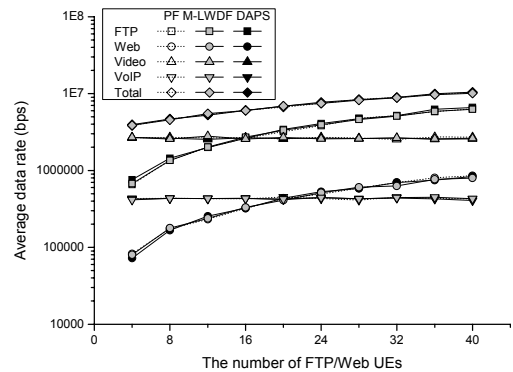


그림 4. PF, M-LWDF, DAPS 알고리즘에서의 평균 데이터 전송률

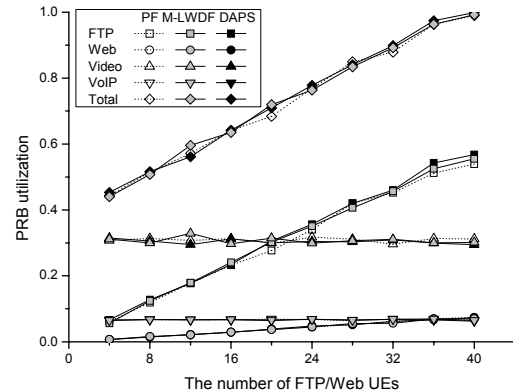
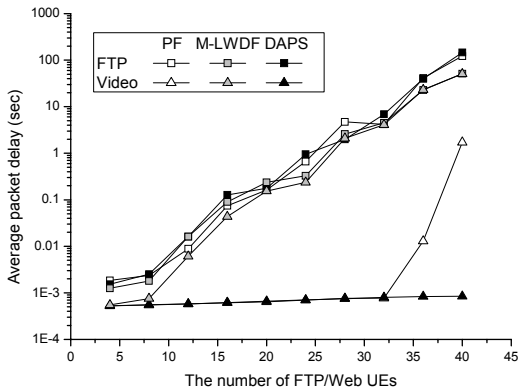


그림 5. PF, M-LWDF, DAPS 알고리즘에서의 평균 PRB 사용률

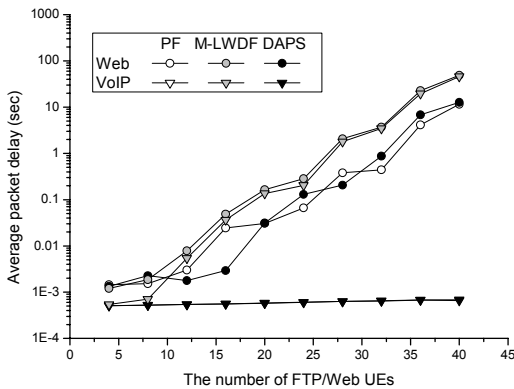
시키는 반면 video streaming의 평균 지연은 n 이 40이 되면 최대 허용 지연인 50ms보다 커지는 결과를 나타낸다. M-LWDF 알고리즘은 VoIP와 video streaming 서비스의 평균 지연이 각각 $n \geq 16$, $n \geq 20$ 일 때 최대 허용 지연인 30ms와 50ms를 초과한다. 하지만, 제안된 DAPS 알고리즘은 n 이 40까지 증가하더라도 RT 트래픽을 최대 허용 지연 이내에 서비스 해줄 수 있음을 보여준다. 이는 PF 알고리즘이 UE의 채널 상태와 평균 데이터 전송률만으로 전송 우선 순위를 결정하고 지연에 민감한 RT 트래픽의 QoS를 보장하기 위한 요소를 스케줄링 정책에 반영하지 않았기 때문이다. M-LWDF 알고리즘은 이를 위해 최대 허용 지연과 패킷 손실률을 스케줄링 요소에 적용하였지만 두 가지 요소가 실제 패킷이 경험하는 지연이나 해당 트래픽의 패킷 손실률이 아닌 QoS 제약사항으로 정의된 값을 반영하고 있기 때문에 다양한 트래픽 종류가 혼재된 환경에서 RT 트래픽의 지연에 대한 QoS를 보

장하기 어렵다. 반면 지연 기중치가 적용된 DAPS 알고리즘은 NRT 트래픽의 평균 지연을 PF, M-LWDF과 비슷하게 유지하면서 NRT 트래픽의 UE 수가 증가하더라도 RT 트래픽의 지연에 대한 QoS를 보장해 줄 수 있음을 보여준다.

그림 7의 (a), (b)는 PF, M-LWDF, DAPS 알고리즘에서의 전송되는 패킷의 큐잉 지연이 최대 허용 지연을 초과할 확률이다. 세 가지 알고리즘에서 NRT 트래픽의 평균 지연은 비슷한 결과를 나타내지만, 전송 되는 패킷이 최대 허용 지연을 초과할 확률은 제안된 알고리즘인 DAPS가 가장 낮음을 확인할 수 있다. RT 트래픽의 경우 DAPS 알고리즘은 최대 허용 지연 이내에 패킷 전송이 이루어지는 반면, PF 알고리즘은 $n \geq 36$ 인 경우 video streaming 서비스의 최대 허용 지연을 초과할 확률이 10^{-3} 보다 크고, M-LWDF 알고리즘은 $n \geq 12$ 인 경우 VoIP와 video streaming 서비스에 허용 되는 최대 지연을

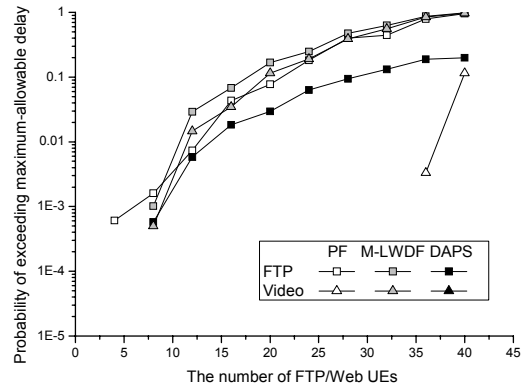


(a) FTP, Video의 평균지연

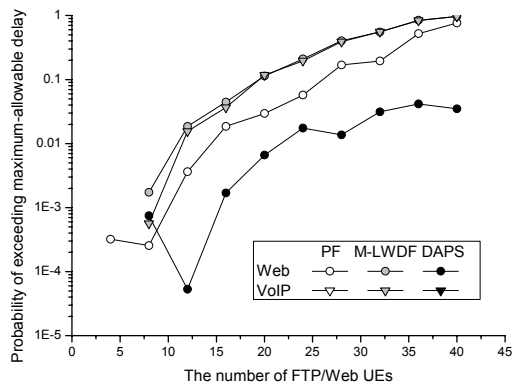


(b) Web, VoIP의 평균지연

그림 6. PF, M-LWDF, DAPS 알고리즘에서의 평균 지연



(a) FTP, Video이 최대 허용 지연을 넘을 확률



(b) Web, VoIP가 최대 허용 지연을 넘을 확률

그림 7. PF, M-LWDF, DAPS 알고리즘에서의 최대 허용 지연을 넘을 확률

초과할 확률이 각각 10^{-2} 과 10^{-3} 보다 큰 결과를 나타낸다. 즉, 제안하는 DAPS 알고리즘은 최대 허용 지연 이내에 전송되는 패킷의 수를 증가시켜 지연에 대한 QoS를 보장하며, NRT 트래픽의 UE 수가 증가하더라도 RT 트래픽 UE에게 신뢰적인 데이터 전송이 가능하게 함을 알 수 있다.

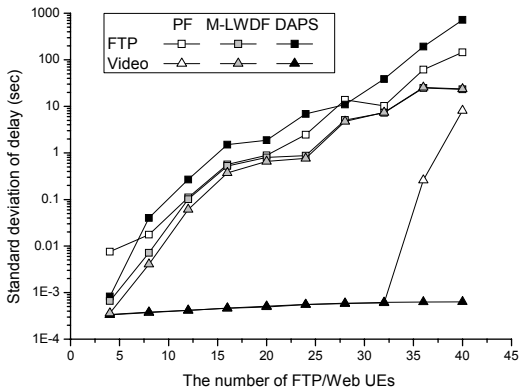
지연에 대한 QoS를 보장하기 위해 적용된 지연 가중치를 사용하는 DAPS 방식에서는 전송되는 패킷이 최대 허용 지연을 초과할 확률을 낮출 수 있지만, 그림 8의 (a), (b)와 같이 NRT 트래픽의 경우 지연에 대한 표준편차는 PF, M-LWDF 알고리즘에 비해 큰 결과를 보인다. DAPS 알고리즘에서 사용되는 지연 가중치는 HOL 패킷의 큐잉 지연이 최대 허용 지연을 초과할 경우 지연 가중치가 감소하게 되므로 해당 패킷은 큐에서 오래 대기한 후에 전송될 수 있다. 이에 따라 최대 허용 지연 이내에 패킷을 전송할 확률은 높아지는 반면 NRT 트래픽

의 지연에 대한 표준편차는 다른 알고리즘에 비해 커지게 되는 현상이 나타난다. 그러나 DAPS 알고리즘에 적용되는 스케줄링 요소 중 지연 가중치가 감소함에 따라 평균 데이터 전송률에 대한 값도 작아지게 되기 때문에 두 가지 요소가 서로 보상되어 결과적으로 해당 HOL 패킷은 높은 우선순위로 처리 될 수 있다. 따라서 DAPS 방식에서 HOL 패킷이 오랜 시간 동안 처리되지 않는 교착상태는 발생하지 않는다. 일반적으로 CAC의 역할은 요구되는 QoS를 보장하기 위해서 신규 호나 핸드오버 호의 수락 여부를 결정하는 것이고, 스케줄러의 역할은 수락된 호에 대해서 QoS를 보장해 주기 위해 전송 우선순위를 결정하는 것이다. 서로 다른 스케줄러의 성능을 평가하기 위해서 CAC를 고려하지 않았기 때문에 트래픽 부하가 과도한 상황에서는 DAPS 알고리즘이 일정 시간 동안 높은 지연 가중치를 부여했음에도 불구하고 최대 허용 지연 보다 오랜 시간 큐에서 기다린 뒤에 전송되는 패킷이 발생할 수 있다. 또한 모의실험에서 패킷의 지연이 최대 허용 지연을 초과한 패킷을 폐기(drop)하는 정책이 고려되지 않았기 때문에 이러한 정책이 적용된다면 제안 방식은 지연에 대한 표준 편차도 커지지 않으면서 패킷이 최대 허용 되는 지연 이내에 전송될 확률도 높이는 결과를 보일 수 있을 것으로 기대된다.

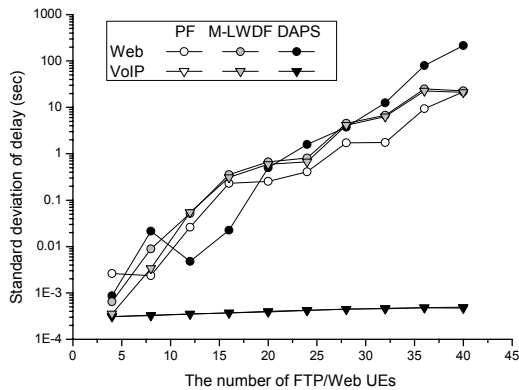
VI. 결론

본 논문에서는 채널 상태나 공정성뿐만 아니라 다양한 트래픽 종류의 지연에 대한 QoS 요구사항을 보장하기 위해서 DAPS 알고리즘을 제안하였다. DAPS 알고리즘은 지연에 대한 QoS 보장을 위해 HOL 패킷의 큐잉 지연이 최대 허용 지연에 가까울 수록 높은 가중치를 부여하는 지연 가중치를 사용한다. 모의실험 결과 DAPS 알고리즘은 PF, M-LWDF 알고리즘과 비교했을 때 NRT 서비스에 대해 평균 데이터 전송률, PRB 사용률, 그리고 평균 지연은 비슷한 성능을 유지하면서 최대 허용되는 지연 이내에 전송되는 패킷의 수를 증가시키는 결과를 나타내었다. 또한 NRT 서비스의 UE 수가 증가되더라도 RT 서비스 사용자의 패킷을 최대 허용 지연 이내에 처리함으로써 RT 사용자에게 안정적이고 신뢰적인 서비스를 제공할 수 있음을 보였다.

DAPS 알고리즘은 전송되는 패킷이 최대 허용 지연 이내에 처리될 확률은 높아지는 반면 NRT 트래픽의 지연에 대한 표준 편차가 커지는 결과를 보



(a) FTP, Video의 지연에 대한 표준편차



(b) Web, VoIP의 지연에 대한 표준편차

그림 8. PF, M-LWDF, DAPS 알고리즘에서의 지연에 대한 표준편차

였다. 이는 모의 실험에서 서로 다른 스케줄러들의 성능을 비교하기 위해서 CAC를 생략하였고, 최대 허용 지연을 초과한 패킷을 폐기하는 정책이 생략되었기 때문이다. 따라서 향후 지연 가중치 함수를 사용하는 DAPS 알고리즘과 연관된 CAC 알고리즘에 대한 연구 및 패킷의 폐기를 반영한 환경에서의 성능 검증이 요구된다.

참 고 문 헌

[1] 김학성, 김봉희, 김기영, 최진성, “3GPP RAN LTE 기술 개요 및 표준화 현황,” *한국통신학회지 (정보와통신)*, 제24권, 제3호, pp. 97-111, Mar. 2007.

[2] 이현우, 지형주, “3GPP 표준 로드맵 및 LTE 기술 개요,” *한국통신학회지 (정보와통신)*, 제25권, 제9호, pp. 3-8, Aug. 2008.

[3] B. Al-Manthari, H. Hassanein, and N. Nasser, “Packet Scheduling in 3.5G High-Speed Downlink Packet Access Networks: Breath and Depth,” *IEEE Network*, vol. 21, no. 1, pp. 41-46, Jan.-Feb. 2007.

[4] E. Dahlman, S. Parkvall, J. Skold, and P. Beming, “3G Evolution: HSPA and LTE for Mobile Broadband,” *Academic Press*, Aug. 2007.

[5] 송평중, D. H. Tri, 임선배, “LTE에서의 Medium Access Control 표준기술,” *한국통신학회지 (정보와통신)*, 제25권, 제9호, pp. 32-40, Aug. 2008.

[6] 3GPP TR 23.203 v8.2.0, “Policy and charging control architecture,” 2008.

[7] R. Padovani, A. Jalali, and R. Pankal, “Data Throughput of CDMA HDR a High Efficiency-High Data Rate Personal Communication Wireless System,” in *Proc. of the 51th IEEE Vehicular Technology Conference (VTC2000)*, vol. 3, pp. 1854-1858, May 2000.

[8] M. Andrews, K. Kumaran, K. Ramanan, A. Stolyar, and P. Whiting, “Providing Quality of Service over a Shared Wireless Link,” *IEEE Communications Magazine*, vol. 39, no. 2, pp. 150-154, Feb. 2001.

[9] S. Shakkottai and A. Stolyar, “Scheduling Algorithms for a Mixture of Real-Time and Non-Real-Time Data in HDR,” in *Proc. of the 17th International Teletraffic Congress (ITC-17)*, Brazil, Sep. 2001.

[10] J.-Y. Hwang and Y. Han, “An Adaptive Traffic Allocation Scheduling for Mobile WiMAX,” in *Proc. of the 18th IEEE International Symposium on Personal, Indoor and Mobile Radio Communications*, pp. 1-5, Sep. 2007.

[11] S. Ryu, B.-H. Ryu, H. Seo, M. Shin, and S. Park, “Wireless Packet Scheduling Algorithm for OFDMA System Based on Time-Utility and Channel State,” *ETRI Journal*, vol. 27, no. 6, pp. 777-787, Dec. 2005.

[12] WiMax Forum, “WiMax System Evaluation Methodology,” v1.0, 2007.

최 범 곤 (Bum-Gon Choi)

준회원



2007년 8월 성균관대학교 정보통신공학부 학사
 2009년 2월 성균관대학교 전자전기컴퓨터공학과 석사
 2009년 3월~현재 성균관대학교 휴대폰학과 박사과정
 <관심분야> 무선 LAN, 무선 메쉬 네트워크, LTE

정 민 영 (Min Young Chung)

종신회원



1990년 2월 KAIST 전기 및 전자공학과 학사
 1994년 2월 KAIST 전기 및 전자공학과 석사
 1999년 2월 KAIST 전기 및 전자공학과 박사
 1999년 1월~2002년 2월 ETRI 선임연구원
 2002년 3월~현재 성균관대학교 정보통신공학부 부교수
 <관심분야> 유/무선 홈네트워크, 무선 LAN/PAN, 이동통신 네트워크, IP 라우터 시스템, 광 네트워크

이 혜 경 (Hye Kyung Lee)

정회원



2007년 2월 경희대학교 전파통신공학과 학사
 2009년 2월 성균관대학교 휴대폰학과 석사
 2009년 2월~현재 삼성전자 DMC 연구소 연구원
 <관심분야> EDGE/HSPA/LTE

시스템

김 태 석 (Tai Suk Kim)

정회원



1993년 2월 KAIST 전기 및
전자공학과 학사

1995년 2월 KAIST 전기 및
전자공학과 석사

1999년 8월 KAIST 전기 및
전자공학과 박사

1999년 9월~2000년 9월

Stanford 대학 박사후연구원

2000년 10월~현재 삼성전자 DMC연구소 수석연구원

<관심분야> CDMA, WCDMA, HSPA, LTE 스케
줄러, 모델

강 지 응 (JeeWoong Kang)

정회원



1998년 2월 서울대학교 전기공
학부 학사

2000년 2월 서울대학교 전기공
학부 석사

2005년 8월 서울대학교 전기공
학부 박사

2005년 9월~현재 삼성전자 DMC

연구소 책임연구원

<관심분야> 통신이론, 통신신호처리, 다중 안테나
통신