

Mecanismos para aliviar o problema de assimetria de retardos em Redes IEEE 802.11e

Juliana Freitag^{1*}, Nelson L. S. da Fonseca¹, José Ferreira de Rezende²

¹Instituto de Computação – Universidade Estadual de Campinas
Caixa Postal 6176 – 13084-971 Campinas, SP

²GTA – COPPE – Universidade Federal do Rio de Janeiro
Caixa Postal 68504 – 21945-970 Rio de Janeiro, RJ

juliana@ic.unicamp.br, nfonseca@ic.unicamp.br, rezende@gta.ufrj.br

Abstract. *The IEEE 802.11e standard aims at supporting applications with QoS requirements. However, it does not define how this support should be given under specific network conditions. When the load of the downlink traffic is high, the queues in the access point build up faster than the ones in the stations, resulting in higher delays for the downlink traffic. This paper introduces a new control mechanism which dynamically tunes parameters of the 802.11e contention-based access method. The proposed mechanism aims at providing QoS as well as ameliorating the problem of delay asymmetry.*

Resumo. *A extensão 802.11e introduz provisão de Qualidade de Serviço (QoS) no padrão IEEE 802.11 para redes locais sem fio. Entretanto, ela não define como essa funcionalidade deve ser utilizada em situações particulares de carga na rede. Quando a carga gerada pelo tráfego na direção do ponto de acesso para as estações é alta, as filas no ponto de acesso crescem mais rapidamente do que nas estações resultando em altos retardos para o tráfego downlink. Este artigo aborda o desenvolvimento e a avaliação de um mecanismo de controle que ajusta dinamicamente os parâmetros de acesso ao meio do 802.11e. O mecanismo proposto tem como objetivo prover QoS, bem como resolver o problema da assimetria no retardo.*

1. Introdução

O padrão IEEE 802.11, para redes locais sem fio (WLAN - *Wireless Local Area Networks*), tem sido amplamente utilizado para o acesso móvel à Internet. Dado que este padrão oferece apenas serviço de melhor esforço, a extensão 802.11e vem sendo desenvolvida para prover Qualidade de Serviço (QoS) às aplicações. Esta extensão inclui dois novos métodos de acesso na subcamada de acesso ao meio: um baseado em contenção e outro sem contenção, os quais permitem a diferenciação de serviços.

Embora a extensão 802.11e forneça diferenciação de serviços, ela não define como essa funcionalidade deve ser utilizada em situações particulares de carga na rede. Dessa forma, faz-se necessária a adoção de novos mecanismos para monitoramento e controle dos níveis de serviço como complemento à funcionalidade de QoS.

O ponto de acesso pode utilizar os parâmetros que regulam o acesso ao meio no método de acesso com contenção da extensão 802.11e para adaptar o acesso ao meio às mudanças na carga oferecida à rede, bem como para prover QoS [6]. Em [7], propôs-se

*Auxílio financeiro do CNPq.

um mecanismo de controle de carga com os objetivos de possibilitar que o tráfego de melhor esforço explore ao máximo a largura de banda disponível na rede, bem como evitar que o retardo máximo do tráfego de tempo real ultrapasse um determinado limite. O mecanismo regula a carga gerada pelas fontes associadas à uma classe de tráfego através do ajuste de parâmetros de acesso ao meio, tais como o tamanho mínimo da janela de contenção e o intervalo entre quadros na subcamada de acesso ao meio (MAC). Em [11], Romdhani *et al* apresentaram um esquema chamado AEDCF (*Adaptive Enhanced Distributed Coordination Function*), que torna o novo método de acesso com contenção do 802.11e mais adaptativo através do ajuste dinâmico do tamanho da janela de contenção após cada transmissão realizada com sucesso e após cada colisão. Busca-se, assim, reduzir o número de colisões e aumentar a utilização do canal. Cali *et al* [3] derivaram um método que utiliza a probabilidade de colisão na rede para calcular o tamanho da janela de contenção mais apropriado e, dessa forma, elevar a vazão no enlace sem fio.

Embora estes mecanismos de ajuste dinâmico dos parâmetros da MAC apresentem bons resultados quanto a eficiência na utilização do canal, eles não tratam do problema da assimetria entre o retardo no sentido de subida (*uplink*) e o retardo no sentido de descida (*downlink*), que ocorre quando se utiliza o método de acesso com contenção do 802.11e. Quando a rede é exposta à altas cargas compostas por tráfego *uplink* e *downlink* o ponto de acesso torna-se um gargalo na rede, pois é responsável pela entrega de todo o tráfego *downlink*, ou seja, pelo tráfego proveniente da rede fixa (Internet, por exemplo) para as estações. Dado que o padrão 802.11e utiliza o protocolo de acesso ao meio CSMA/CA, o ponto de acesso tem a mesma probabilidade de acesso ao meio que todas as outras estações, ou seja, uma estação que tem apenas uma fonte ativa tem a mesma chance de acessar o meio que o ponto de acesso que pode ter várias fontes ativas. Esse comportamento do protocolo de acesso ao meio resulta em um rápido crescimento das filas no ponto de acesso e, conseqüentemente, em altos valores de retardo para o tráfego *downlink*.

Este artigo enfoca o desenvolvimento e avaliação de dois mecanismos de ajuste dinâmico dos parâmetros de QoS da MAC da extensão 802.11e. Os mecanismos têm como objetivo auxiliar na provisão de QoS, além de resolver o problema da assimetria no retardo dos tráfegos *downlink* e *uplink*. O artigo está estruturado da seguinte forma: a Seção 2 descreve conceitos básicos sobre redes 802.11. A Seção 3 apresenta a motivação para o desenvolvimento deste trabalho. A Seção 4 apresenta o mecanismo para ajuste dos parâmetros AIFS e CW_{min} . A Seção 5 descreve o mecanismo para ajuste dos parâmetros TXOP e CW_{min} . Finalmente, a Seção 6 apresenta as conclusões.

2. O Padrão IEEE 802.11

O padrão IEEE 802.11 [5] define duas funções de acesso ao meio: a função de coordenação distribuída (*Distributed Coordination Function* - DCF) e a função de coordenação em um ponto (*Point Coordination Function* - PCF). No modo PCF, um Ponto de Coordenação, localizado no ponto de acesso (PA), controla o acesso ao meio consultando periodicamente as estações sem fio (STAs). O modo DCF provê acesso ao meio com contenção através do protocolo CSMA (*Carrier Sense Multiple Access*).

No modo DCF, as STAs devem observar o meio livre durante um período de tempo igual a DIFS (*Distributed Inter Frame Space*) antes de competir pelo canal. Quando uma STA decide transmitir e detecta o canal ocupado, a transmissão é adiada. A grande diferença entre o CSMA/CD utilizado em redes cabeadas e o protocolo de acesso ao meio do 802.11 está no fato de que no ambiente sem fio não é possível detectar colisão. Esta funcionalidade é substituída por mecanismos para prevenção de colisão (CSMA/CA) e reconhecimento positivo (ACKs) dos quadros enviados. O receptor deve enviar um ACK para cada quadro recebido. Além disso, as STAs devem aguardar durante intervalos de

tempo randômicos (tempo de backoff) antes de transmitir um quadro de forma que colisões sejam evitadas. O tempo de backoff é randomicamente escolhido em um intervalo de 0 até o tamanho da janela de contenção (CW). Um tamanho mínimo de CW (CW_{min}) é utilizado após uma transmissão bem sucedida. STAs envolvidas em uma colisão têm sua janela de contenção aumentada para uma próxima potência de 2 menos 1 até que um valor máximo seja alcançado (CW_{max}).

A extensão 802.11e [6] introduz uma nova função de coordenação, a HCF (*Híbrida Coordination Function*), que possui um Coordenador Híbrido responsável pela funcionalidade de QoS. A Função Híbrida é composta por dois modos de acesso: HCF e EDCF (*Enhanced DCF*). O EDCF é utilizado no período com contenção, enquanto o HCF é utilizado em ambos os períodos, com contenção e livre de contenção.

O EDCF provê diferenciação de serviços utilizando diferentes categorias de acesso (ACs - *Access Categories*). Cada AC possui uma fila e valores diferentes para os parâmetros de acesso ao meio: CW_{min} , CW_{max} e intervalo entre quadros (AIFS - *Arbitration IFS*). A priorização é obtida através da atribuição de valores menores para os parâmetros das categorias de maior prioridade. Para aumentar o desempenho e melhorar o nível de utilização do canal, as STAs realizam a transmissão de quadros em rajadas durante intervalos de tempo chamados TXOPs (*Transmission Opportunity*). Neste intervalo, uma STA pode enviar vários quadros sem a necessidade de disputar o meio a cada quadro.

3. Motivação

Avaliações de desempenho mostram que a extensão 802.11e oferece melhor suporte à QoS do que os modos DCF e PCF em condições de baixa e média carga [1]. Entretanto, o aumento da carga provoca quedas na vazão das fontes de dados, enquanto o aumento no número de estações resulta em um rápido crescimento da taxa de colisão o que prejudica o desempenho do tráfego multimídia. Resultados de simulação apresentados por Mangold *et al* [16] mostram que se existe um pequeno número de estações competindo pelo canal, um valor baixo de CW_{min} possibilita o acesso com prioridade. No entanto, se o número de estações aumenta, faz-se necessário um esquema adaptativo para ajustar o valor de CW_{min} de acordo com o número de estações ou com a carga presente na rede [11].

Este comportamento revela que as funcionalidades de QoS introduzidas pela extensão 802.11e não são suficientes para atender aos requisitos de QoS das diferentes classes de tráfego em situações de alta carga na rede. Esta deficiência motiva o desenvolvimento de novos mecanismos para monitoramento e controle dos níveis de serviço. Um destes mecanismos, o controle de admissão, restringe o número de usuários simultaneamente presentes na rede de forma a evitar a saturação do enlace sem fio, o que resulta em retardos elevados. O controle de admissão é um mecanismo de controle de tráfego extensivamente estudado nas redes fixas ([9], [12], [15]). Em mecanismos de controle de admissão baseado em parâmetros, os fluxos de dados informam seus requisitos de QoS através de parâmetros. Com base nesta informação, um algoritmo de controle de admissão determina se a Qualidade de Serviço requisitada pode ser oferecida sem prejudicar os demais fluxos. Os esquemas de controle de admissão baseado em medições (MBAC - *Measurement-Based Admission Control*) não necessitam de uma caracterização precisa do tráfego, dado que as decisões de admissão baseiam-se em medições do tráfego realizadas periodicamente. Os esquemas MBAC permitem uma maior utilização da rede do que os algoritmos de controle de admissão baseado em parâmetros [15], assim sendo, a escolha de um algoritmo MBAC parece ser a mais apropriada dada a escassez de recursos das redes sem fio. Contudo, em situações nas quais a rede está sujeita a uma alta carga composta, principalmente, por tráfego *downlink* este mecanismo não é suficiente para manter o retardo das aplicações inferior ao valor requisitado.

Este fato pode ser observado através de um experimento¹ envolvendo tráfego de tempo real (voz e vídeo) e de melhor esforço (dados). As fontes são admitidas por um MBAC que utiliza envelopes de tráfego para gerar as estimativas da rede [8]. A técnica consiste em medir e controlar envelopes de chegada e de serviço para cada uma das classes. O PA é responsável pelo controle dos envelopes e pela admissão de novas fontes.

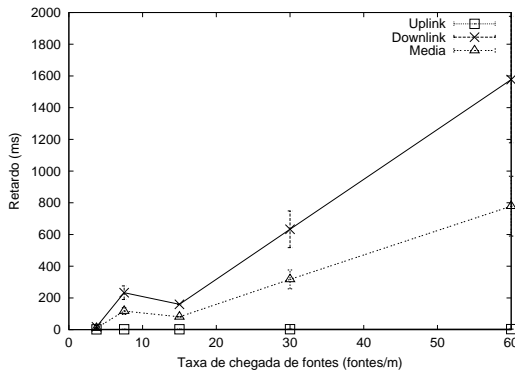


Figura 1: Retardo das fontes de voz

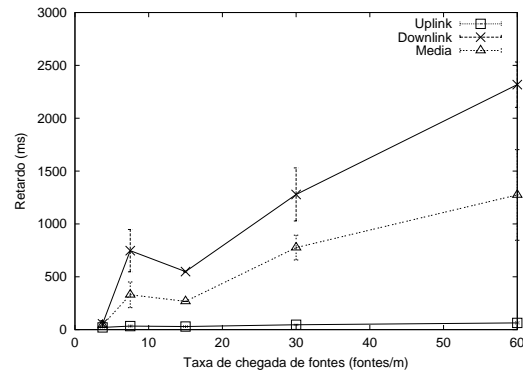


Figura 2: Retardo das fontes de vídeo

As Figuras 1 e 2 apresentam, respectivamente, o retardo obtido pelos tráfegos de voz e vídeo no sentido *uplink* e *downlink*, bem como a média entre estes retardos. Nota-se que o retardo do tráfego *uplink* mantém-se sempre abaixo do valor requisitado, 100 ms para voz e vídeo. Por outro lado, com o aumento na taxa de chegada de fontes, o retardo do tráfego *downlink* apresenta um crescimento rápido, resultando em violação do requisito de Qualidade de Serviço. Este fenômeno ocorre porque em uma rede local sem fio o ponto de acesso é responsável pela entrega de todo o tráfego *downlink*. Assim sendo, há N instâncias *uplink* (estações) competindo pelo canal com uma única instância *downlink* (ponto de acesso). Dado que, no protocolo de acesso ao meio do 802.11e, a probabilidade de acesso ao canal é igualmente distribuída entre instâncias pertencentes a mesma classe de tráfego, as filas no ponto de acesso crescem mais rapidamente do que nas estações resultando em altos retardos para o tráfego *downlink*.

Note que há duas alternativas para a solução deste problema de assimetria no retardo: empregar um mecanismo de controle de admissão mais conservador, o que resultaria em uma queda na utilização da rede, ou adotar um mecanismo adicional que realize a distribuição dos recursos da rede de acordo com a necessidade das estações e do ponto de acesso, sem, no entanto, reduzir a utilização do canal.

As seções seguintes mostram que com o auxílio de um mecanismo de ajuste dos parâmetros que regulam o acesso ao meio na extensão 802.11e, é possível prover QoS tanto para o tráfego *uplink* quanto para o tráfego *downlink* e, ainda, explorar ao máximo a largura de banda disponível na rede.

4. Ajuste dinâmico dos parâmetros AIFS e CW_{min}

O mecanismo proposto nesta seção realiza o ajuste dinâmico dos parâmetros AIFS e CW_{min} e tem como objetivo auxiliar o controle de admissão a garantir que os requisitos de QoS sejam atendidos, bem como manter elevada a utilização da rede. Este mecanismo estende a proposta apresentada em [7] para que o problema da assimetria no retardo também seja considerado.

O ajuste do CW_{min} deve levar em consideração o fato de que o tamanho da janela de contenção (CW) está fortemente relacionado ao número de estações presentes na rede.

¹A metodologia do experimento segue a descrição apresentada na Seção 4.3

Quando há poucas estações na rede, uma janela de contenção com valor pequeno reduz o tempo ocioso do canal e possibilita uma melhor utilização da largura de banda. Quando o número de estações aumenta, é preferível um valor maior para CW para reduzir a probabilidade de colisão entre as estações [2, 3, 4].

A relação entre os valores de AIFS e CW_{min} das classes de acesso também influencia o funcionamento do ajuste dinâmico dos parâmetros. Se o AIFS de uma classe $(i + 1)$ for igual ao $AIFS + CW_{min}$ usados na classe de acesso i , mais prioritária do que a classe $(i + 1)$, uma separação total entre as classes é garantida. Neste caso, não há sobreposição entre as janelas dos diversos níveis de prioridade, reduzindo a probabilidade de que um quadro menos prioritário seja transmitido antes de outro mais prioritário. Entretanto, se a classe mais prioritária apresentar carga alta, de forma que haja sempre algum quadro a transmitir nas filas desse nível, a classe menos prioritária sofrerá inanição por banda passante (*starvation*). Além disso, o número de níveis de serviço suportáveis fica reduzido, dado que o nível mais baixo utilizará valores de AIFS e CW_{min} elevados.

Por outro lado, se para a classe $(i + 1)$ o $AIFS(i + 1)$ é menor que $AIFS(i) + CW_{min}(i)$, a inanição das classes de menor prioridade pode ser evitada e uma maior granularidade de níveis de serviço pode ser suportado. Em compensação, ocorre uma probabilidade de inversão de prioridade proporcional ao tamanho da região de sobreposição, que é dado por $AIFS(i) + CW_{min}(i) - AIFS(i + 1)$.

Com base nestas observações, propõe-se um mecanismo de ajuste dinâmico dos parâmetros AIFS e CW_{min} , que atribui a $AIFS(i)$ e a $CW_{min}(i)$ os menores valores possíveis tal que não provoquem violação do retardo. Se não há carga na classe i , o mecanismo minimiza a parte determinística da espera das classes desde $(i + 1)$ até a de mais baixa prioridade, reduzindo os valores do AIFS dessas classes. À medida que o número de estações com fontes da classe i aumenta, o mecanismo aumenta o CW_{min} desta classe, o que provoca a necessidade de aumentar o CW_{min} das classes menos prioritárias que i .

Para tratar do problema da assimetria no retardo, um mesmo tipo de tráfego é mapeado para classes diferentes nas estações e no ponto de acesso. No ponto de acesso um tipo de tráfego específico pertence a uma classe de maior prioridade do que nas estações. O parâmetro AIFS é utilizado para prover a diferenciação entre as classes e o CW_{min} é ajustado de acordo com o número de estações com fontes de um mesmo tipo de tráfego. Assim, por exemplo, o tráfego de voz, que em uma rede 802.11e pertence a classe 0 (maior prioridade), no esquema proposto pertence a classe 0 no ponto de acesso e a classe 1 nas estações, sendo que a classe 0 possui AIFS inferior ao AIFS da classe 1 e CW_{min} igual ao CW_{min} da classe 1. Caso existam três tipos de tráfego na rede, o ponto de acesso utiliza as classes 0, 2 e 4, enquanto as estações utilizam as classes 1, 3 e 5. Dessa forma, pretende-se fornecer maior probabilidade de acesso ao ponto de acesso e, então, equilibrar os valores do retardo do tráfego *downlink* e do tráfego *uplink*.

4.1. Restrições

O mecanismo de ajuste dinâmico dos parâmetros de QoS deve observar algumas restrições na determinação dos valores de AIFS e CW_{min} .

- Restrição 1: o valor de AIFS de qualquer classe i deve ser sempre maior ou igual a 2, para evitar falhas no funcionamento da MAC. Como o intervalo entre-quadros da classe é definido por $SIFS + AIFS(i) * SlotTime$, o valor de $AIFS(i)$ não pode ser 0, para evitar que este intervalo fique igual a SIFS, nem pode ser 1, para não se igualar ao intervalo entre-quadros do PCF (PIFS), que é igual a $SIFS + SlotTime$.
- Restrição 2: em qualquer situação, o valor de $AIFS(i + 1)$ deve ser maior ou igual ao valor de $AIFS(i)$, para preservar a diferenciação. Se $AIFS(i + 1)$ for igual a $AIFS(i)$, então $CW_{min}(i + 1)$ deve ser obrigatoriamente maior que $CW_{min}(i)$.

- Restrição 3: $AIFS(i + 1)$ deve ser menor que $AIFS(i) + CW_{min}(i)$ se a carga da classe i for alta, para evitar inanição da classe $(i + 1)$.
- Restrição 4: o CW_{min} de uma classe não pode ser maior que o valor de CW_{max} .
- Restrição 5: nas situações de pouca carga, o CW_{min} das classes não deve ser menor do que um valor mínimo. Este valor pode ser estabelecido pelo administrador do sistema, sendo que para a classe de maior prioridade o CW_{min} não deve ser menor do que 7 [5], dado que este valor é um valor mínimo razoável para permitir um compromisso entre baixo retardo e prevenção de colisão.

4.2. Algoritmos

O mecanismo de ajuste dos parâmetros AIFS e CW_{min} é executado periodicamente de acordo com um intervalo de monitoração. A cada execução, o mecanismo verifica se o tamanho do CW_{min} usado nas classes está adequado. Em seguida, realiza o ajuste da carga imposta pelas classes, aumentando ou diminuindo os valores de AIFS e CW_{min} , sempre tendo em vista as restrições e os objetivos de diferenciação. Após ter executado as duas operações, o mecanismo emite um novo conjunto de parâmetros de QoS.

```

ALGORITMO ajustarParametros /*  $n$  é o número de classes */
1. para cada classe ímpar  $i$ ,  $i$  de 1 até  $(n - 1)$  /* classes das estações */
2.   ajustarContencao( $i$ )
3. para cada classe par  $i$ ,  $i$  de 0 até  $(n - 4)$  /* classes do ponto de acesso */
4.   ajustarCarga( $i$ )
5. emitirConjuntoDeParametrosDeQoS()

```

Figura 3: Módulo principal do ajuste de parâmetros para n classes

```

ALGORITMO ajustarContencao( $i$ )
1. se  $estacoes(i) > CW_{min}(i)$ 
2.   para cada classe  $j$ ,  $j$  de  $(i - 1)$  até  $(n - 1)$ 
3.      $CW_{min}(i) = CW_{min}(i) * 2 + 1$ 
4. se  $estacoes(i) < CW_{min}(i)/2$ 
5.   para cada classe  $j$ ,  $j$  de  $(i - 1)$  até  $(n - 1)$ 
6.      $CW_{min}(j) = (CW_{min}(j) - 1)/2$ 
7. para cada classe  $j$ ,  $j$  de  $(i - 1)$  até  $(n - 1)$ 
8.   se  $CW_{min}(j) < CW_{Minimo}(j)$ 
9.      $CW_{min}(j) = CW_{Minimo}(j)$ 

```

Figura 4: Ajuste de contenção nas classes

As Figuras 3, 4, 5 e 6 mostram os algoritmos utilizados para o caso de n classes, explorando a coexistência de tráfego de tempo real com tráfego de melhor esforço. As classes $(n - 1)$ e $(n - 2)$ representam a classe de melhor esforço nas estações e no ponto de acesso respectivamente.

Na Figura 3 está o módulo principal do mecanismo. O módulo chama o procedimento $ajustarContencao(i)$ para as classes das estações, iniciando pela de maior prioridade. Este procedimento verifica se o tamanho mínimo da janela de contenção das classes deve ser maior para reduzir o número de tentativas de transmissão. Para tanto, compara-se o $CW_{min}(i)$ com o número de estações que possuem fontes aceitas na classe i . O procedimento verifica também se o $CW_{min}(i)$ pode ser reduzido, comparando a metade desse valor com o número de estações. O número de estações é obtido subtraindo-se do número de fontes admitidas na classe i , informado pelo controle de admissão, o número de fontes

ALGORITMO *ajustarCarga(i)* /* verifica os limites de qualidade */

1. se $(\text{RetardoAtual}(i) > \text{RetardoMaximo}(i))$
2. $\text{reduzirCarga}(n - 1)$ /*reduz a carga da classe de melhor esforço nas estações*/
3. $\text{reduzirCarga}(n - 2)$ /*reduz a carga da classe de melhor esforço no PA*/
4. $j = n - 4$
5. enquanto $(\text{RetardoAtual}(j) < \text{RetardoMaximo}(j)*\delta)$ e $(\text{CargaAtual}(j) \leq \text{CargaAnterior}(j) + \text{CargaMediaDaFonte}(j))$ e $(j > i)$
6. $\text{reduzirCarga}(j)$ /*reduz a carga da classe j do ponto de acesso*/
7. $\text{reduzirCarga}(j + 1)$ /*reduz a carga da classe $(j + 1)$ das estações*/
8. $j = j - 2$
9. $\text{reduzirCarga}(j + 1)$
10. senão
11. para cada classe par j , j de $(n - 4)$ até 0 /*classes do ponto de acesso*/
12. se $(\text{RetardoAtual}(j) < \text{RetardoMaximo}(j)*\delta)$ e $(\text{CargaAtual}(j) \leq \text{CargaAnterior}(j) + \text{CargaMediaDaFonte}(j))$
13. $\text{aumentar} = \text{verdadeiro}$
14. senão
15. $\text{aumentar} = \text{falso}$
16. break
17. se $\text{aumentar} = \text{verdadeiro}$
18. para cada classe j , j de $(i + 1)$ até $(n - 1)$
19. $\text{aumentarCarga}(j)$

Figura 5: Ajuste da Carga

que estão no sentido *downlink*. Assume-se que o ponto de acesso tem controle sobre esta informação e que cada estação possui uma única fonte ativa.

Caso o tamanho do CW_{min} da classe i seja alterado, o ajuste é realizado desde a classe $(i - 1)$ até a classe $(n - 1)$. Dessa forma, mantém-se o $CW_{min}(i)$ igual ao CW_{min} da classe $(i - 1)$, que é a classe do ponto de acesso responsável pelo mesmo tipo de tráfego da classe i . Além disso, mantém-se a diferenciação dos valores desse parâmetro para os diferentes níveis de serviço. Por fim, o procedimento verifica se o valor do CW_{min} das classes não foi alterado para um valor abaixo do mínimo permitido. Caso esta situação se confirme, o CW_{min} recebe o valor mínimo.

Em seguida, o módulo principal chama o procedimento $\text{ajustarCarga}(i)$, mostrado na Figura 5, para cada classe i do ponto de acesso, iniciando pela de maior prioridade. Esse procedimento verifica se a carga das classes pode ser ajustada comparando o valor atual do retardo no ponto de acesso com o valor requisitado pela classe. Para que este procedimento seja chamado para as classes das estações, estas devem monitorar o retardo em suas filas e enviar esta informação para o ponto de acesso, o que implica em maior sobrecarga para a rede. Dado que os resultados apresentados na Seção 3 mostram que o retardo do tráfego *uplink* está sempre abaixo do valor máximo requisitado e do valor do retardo do tráfego *downlink*, optou-se por chamar este procedimento apenas para as classes do ponto de acesso.

Se o retardo da classe i estiver acima do valor máximo requisitado, o procedimento reduz a carga das classes de melhor esforço (do ponto de acesso e das estações) e, então, verifica se é possível reduzir a carga das classes do ponto de acesso menos prioritárias que i , verificando se o retardo destas classes é menor que o produto do valor requisitado por um fator δ , que varia de 0 a 1. A utilização do fator δ é importante pois caso o retardo atual da classe seja menor e muito próximo do valor requisitado pequenas alterações no retardo,

ALGORITMO *reduzirCarga(j)* /* reduz a carga da classe j */

1. se $AIFS(j) < (AIFS(j-1) + CW_{min}(j-1))$
2. $AIFS(j) = AIFS(j) + SlotsReducao$
3. senão se $(CW_{min}(j) * 2 + 1) < CW_{max}$
4. $CW_{min}(j) = CW_{min}(j) * 2 + 1$

ALGORITMO *aumentarCarga(j)* /* aumenta a carga da classe j */

1. se $((CW_{min}(j)-1)/2 > CW_{Minimo})$ e $((CW_{min}(j)-1)/2 > estacoes(j))$
 2. $CW_{min}(j) = (CW_{min}(j) - 1)/2$
 3. senão
 4. se $(AIFS(j) - SlotsAumento) > 2$
 5. $AIFS(j) = AIFS(j) - SlotsAumento$
-

Figura 6: Reduzir/Aumentar Carga

provocadas pela redução de carga da classe, podem resultar em violação do requisito de QoS. Para evitar instabilidade do mecanismo, uma condição adicional ($CargaAtual(j) \leq CargaAnterior(j) + CargaMediaDaFonte(j)$) é verificada. Essa condição sugere que houve a saída de uma fonte no período anterior. A carga atual e o retardo das classes podem ser obtidos monitorando-se periodicamente as filas de transmissão no PA.

Se o retardo da classe i estiver abaixo do valor máximo requisitado, o procedimento verifica se há sobra de recursos para todas as classes de tempo real. Caso a condição seja verdadeira as classes menos prioritárias que i podem ter sua carga aumentada. Note que caso uma determinada classe j não satisfaça a condição, as outras também não podem ter sua carga aumentada, pois prejudicariam ainda mais a situação da classe j . Sempre que o mecanismo decide reduzir ou aumentar a carga de uma classe j (do ponto de acesso), a mesma alteração é aplicada à classe $(j + 1)$ (das estações).

A Figura 6 apresenta os procedimentos para ajustar a carga das classes. A redução da carga é feita inicialmente somando-se um número de *slots* a AIFS (*SlotsReducao*), até o limite da Restrição 3. A partir desse ponto, o valor de CW_{min} passa a ser dobrado, até o limite da Restrição 4. A elevação é realizada de modo inverso, respeitando as Restrições 1 e 5 e utilizando um número de *slots* (*SlotsAumento*) diferente. Utilizar *SlotsAumento* menor que *SlotsReducao* faz com que o controle seja menos agressivo na elevação do que na redução de carga, o que também evita instabilidade do mecanismo.

Depois que esses procedimentos são executados, um novo conjunto de parâmetros de QoS é composto e deve ser transmitido para as estações.

4.3. Avaliação do Mecanismo

Utilizou-se simulação de eventos discretos para obter os resultados de desempenho do mecanismo proposto. As simulações foram realizadas utilizando a ferramenta *ns* [10].

Cada experimento de simulação consiste em um processo randômico de chegada de fluxos que requisitam o serviço da rede e são admitidos ou rejeitados de acordo com a decisão do mecanismo de controle de admissão. Um fluxo rejeitado deixa a rede sem enviar qualquer pacote de dados. Um fluxo aceito envia pacotes de dados durante um tempo de vida escolhido a partir de uma distribuição exponencial. Os pacotes são gerados de acordo com um modelo de tráfego.

O cenário de simulação consiste em um ponto de acesso conectado a um nó fixo por meio de um enlace de 100 Mbps, com retardo de 2 ms. O PA está localizado no centro de uma área de 350 x 350 metros, de forma que todas as estações contidas na área estejam a uma distância menor que 250 metros do ponto de acesso, que é o alcance padrão do *ns*

para estações sem fio. A taxa de dados utilizada é de 11 Mbps e a taxa básica é de 1Mbps.

Foram utilizados três modelos de tráfego nos experimentos para representar tráfego de voz, vídeo e dados. As fontes de voz geram bytes a uma taxa de 64Kbps, com pacotes de 256 bytes e têm períodos de atividade e de silêncio que seguem distribuições exponenciais com médias 1,2 e 1,8 segundos, respectivamente [14]. Para simular o padrão conversacional de cada conexão de voz, foi empregada uma fonte exponencial (*onoff*) na direção da estação para o nó fixo (*uplink*) e outra na direção do nó fixo para a estação (*downlink*), que são iniciadas com uma diferença de até 1 s. Essa bidirecionalidade é importante para tornar a simulação mais próxima da situação real dado que, nas redes 802.11, o tráfego *downlink* compete pelo meio sem fio com o tráfego *uplink*.

O tráfego de vídeo é obtido a partir de um arquivo de trace produzido por uma codificação MPEG do programa de televisão ARD News [13]. As fontes de vídeo geram pacotes de 512 bytes a uma taxa média de 720 Kbps e taxa de pico de 3,4 Mbps.

As fontes de dados geram pacotes de 1024 bytes de acordo com o modelo Pareto *onoff* com os seguintes parâmetros: período de atividade e de silêncio com média de 250 ms e taxa de pico de 400 Kbps. O valor do parâmetro *shape* é 1,9 [17].

O tempo de vida das fontes segue uma distribuição exponencial com média de 300 s para fontes de voz e de dados e 180 s para fontes de vídeo. O sentido de cada fonte é determinado randomicamente. O requisito de QoS, retardo máximo, para o tráfego de tempo real (voz e vídeo) é 100 ms. As fontes de dados, apesar de fazerem parte do tráfego de melhor esforço, são submetidas a admissão para que o ponto de acesso possa estar informado sobre o número de fontes na rede. No entanto, o valor utilizado para o requisito de retardo é alto (5 s) de forma que o número de fontes rejeitadas seja pequeno.

O protocolo de transporte utilizado é o UDP, com tamanho de pacote igual ao usado pela fonte para evitar fragmentação e remontagem na camada de transporte. Por simplicidade, não foi utilizado um protocolo de sinalização de chamadas específico, uma vez que é possível, no *script* de controle da simulação, iniciar a fonte apenas se a conexão for admitida. No entanto, deve ficar claro que tal protocolo é indispensável em uma implementação real.

A avaliação do mecanismo de ajuste dinâmico dos parâmetros foi realizada após 10 rodadas do cenário de simulação, sendo que cada rodada utilizou um valor de semente diferente escolhido randomicamente. Os dados coletados durante um período inicial de *warm up* foram descartados. A duração de cada rodada foi 600 s.

Classe	Tipo do Tráfego	AIFS	CW_{min}
0 (PA)	voz	2	7
1 (estação)	voz	4	7
2 (PA)	vídeo	6	15
3 (estação)	vídeo	8	15
4 (PA)	dados	10	31
5 (estação)	dados	12	31

Tabela 1: Configuração dos parâmetros AIFS e CW_{min}

O intervalo de monitoração utilizado nas simulações é de 100 ms. A avaliação do impacto do intervalo de monitoração sobre o retardo do tráfego *downlink* foi realizada variando-se esse intervalo de 50 ms até 3 s. Valores inferiores a 100 ms não apresentaram ganhos significativos, enquanto valores superiores apresentaram um crescimento mais rápido no valor do retardo de acordo com o aumento da taxa de chegada de fontes.

No procedimento ajustarCarga são utilizados 2 *slots* para *SlotsAumento* e 8 para

SlotsReducao. O parâmetro δ , usado neste mesmo procedimento, é igual a 0,5 e os valores iniciais para os parâmetros AIFS e CW_{min} , para cada uma das classes, são apresentados na Tabela 1.

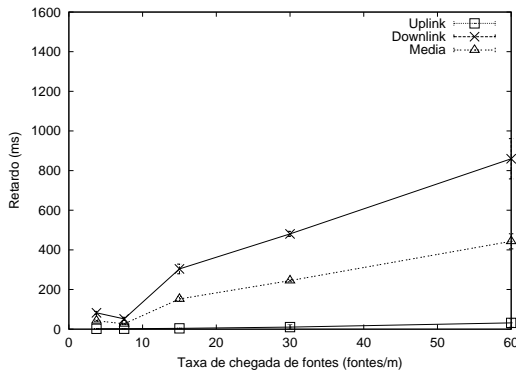


Figura 7: Retardo das fontes de voz

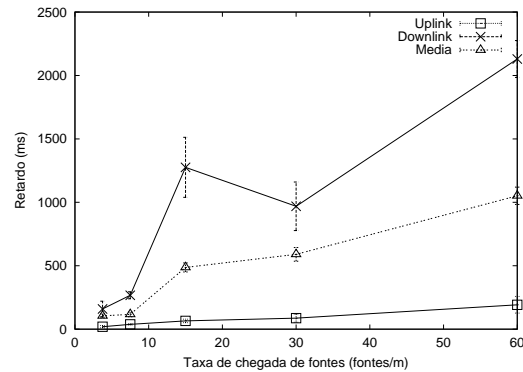


Figura 8: Retardo das fontes de vídeo

Os resultados apresentados nas Figuras 7 e 8 mostram o retardo obtido pelos tráfegos de voz e vídeo. Se comparados aos resultados apresentados nas Figuras 1 e 2, no qual se utilizou apenas o mecanismo de controle de admissão, observa-se que em algumas situações houve uma queda no retardo do tráfego *downlink*, enquanto em outras situações houve um aumento, sendo que o tráfego de voz foi mais beneficiado pelo mecanismo de ajuste dos parâmetros do que o tráfego de vídeo.

Embora o mecanismo tenha proporcionado, em alguns casos, uma melhora no retardo do tráfego *downlink*, os resultados não satisfazem a expectativa, ou seja, o mecanismo não foi capaz de prover ao tráfego *downlink* um retardo semelhante ao retardo do tráfego *uplink*. Este resultado deve-se, principalmente, ao fato dos parâmetros AIFS e CW_{min} fornecerem recursos limitados quando se deseja aumentar a vazão de algumas classes. Pequenas diferenças entre o AIFS e o CW_{min} das classes não proporcionam aumento expressivo na vazão das classes cujo parâmetro têm valor inferior. Enquanto diferenças maiores fazem com que as classes de menor prioridade tenham valor de AIFS muito alto e conseqüentemente retardo elevado. Além disso, o fato deste parâmetro ser utilizado para promover a diferenciação entre os níveis de serviço, não permite ajustá-lo de acordo com a necessidade de vazão das classes. Por exemplo, se, em determinado momento, uma classe $(i + 1)$ necessitar de maior vazão do que uma classe i , de prioridade superior, não é possível atribuir a $(i + 1)$ valor menor de AIFS do que aquele utilizado por i , pois violar-se-ia a regra de que classes com menor prioridade devem ter AIFS menor do que classes de prioridade superior.

Assim sendo, acredita-se que os parâmetros AIFS e CW_{min} , embora eficientes no estabelecimento de níveis de serviço, não são a melhor escolha quando o ajuste de seus valores tem como objetivo controlar a carga imposta pelas classes, bem como solucionar o problema da assimetria entre o retardo dos tráfegos *downlink* e *uplink*. Além disso, é possível concluir que fornecer prioridade mais alta para o PA do que para as STAs não é uma solução eficiente para o problema de assimetria.

5. Ajuste dinâmico dos parâmetros TXOP e CW_{min}

Dado que os resultados obtidos com o mecanismo de ajuste dinâmico dos parâmetros AIFS e CW_{min} não foram satisfatórios, nesta seção propõe-se um mecanismo de ajuste dinâmico dos parâmetros TXOP e CW_{min} . O parâmetro TXOP controla o período durante o qual uma estação pode utilizar o canal após a contenção. Neste período, a estação pode enviar vários quadros sem a necessidade de competir pelo canal e, portanto, o ajuste do TXOP pode resultar em aumento ou redução na vazão de uma classe de acesso.

A grande causa da assimetria no retardo está no fato de que o ponto de acesso deve dar vazão à várias fontes em cada classe, o que não acontece nas estações (assume-se que as estações têm uma fonte ativa em uma classe). Logo, deve-se fornecer a cada classe do ponto de acesso uma vazão condizente com o número de fontes que elas possuem. Pode-se ajustar a vazão do ponto de acesso da seguinte forma: se o TXOP de uma classe i possibilita o envio de q quadros e se esse valor é suficiente para que as estações atendam ao requisito de QoS, então, se o ponto de acesso tem k fontes na classe i , ele deve possuir um TXOP que possibilite enviar $k * q$ quadros para que proporcione a cada fonte a mesma vazão fornecida por uma estação e, conseqüentemente, um retardo semelhante.

Da mesma forma, o valor do TXOP deve permitir que as classes tenham uma vazão de acordo com a carga imposta pelas aplicações, ou seja, o valor do TXOP deve ser ajustado segundo a necessidade da classe e não de acordo com sua prioridade. Nem sempre aplicações que demandam maior vazão têm prioridade sobre aquelas que necessitam vazão menor. Normalmente, o que define a prioridade de uma aplicação é o retardo que ela exige, ou seja, aplicações que exigem retardo menor têm prioridade sobre aquelas que exigem um retardo maior. Aplicações de voz, por exemplo, requerem menor retardo e menor vazão do que aplicações de vídeo, que têm prioridade inferior.

Tendo em vista os aspectos mencionados, o mecanismo proposto ajusta o valor do parâmetro TXOP atribuindo à classe com menor taxa de chegada de pacotes o valor zero, o que significa que a classe poderá enviar um único quadro quando obtiver o canal. As classes com taxa de chegada de pacotes maior recebem um TXOP proporcional à razão entre o número de pacotes que chegaram em suas filas e o número de pacotes que chegaram na fila da classe com menor taxa de chegada. Em seguida, o mecanismo ajusta o TXOP no ponto de acesso de acordo com o número de fontes *downlink* ativas em cada classe. O parâmetro CW_{min} é ajustado conforme o número de estações com fontes ativas em cada classe, de forma que o número de colisões seja reduzido. A diferenciação entre as classes é promovida pelos parâmetros CW_{min} e AIFS, sendo que este é mantido com valor constante.

5.1. Algoritmos

O mecanismo de ajuste dos parâmetros TXOP e CW_{min} é executado periodicamente de acordo com um intervalo de monitoração. A cada execução, o mecanismo verifica se o valor do CW_{min} utilizado pelas classes está adequado. Em seguida, realiza o ajuste do TXOP segundo a carga imposta pelas aplicações e o número de fontes *downlink*. Após ter executado as duas operações, emite-se um novo conjunto de parâmetros de QoS.

ALGORITMO *ajustarParametros* /* n é o número de classes */

1. para cada classe i , i de 1 até $(n - 1)$
 2. *ajustarContencao*(i)
 3. *ajustarTXOP_Classes*()
 4. para cada classe i , i de 0 até $(n - 1)$
 5. $TXOP_{PA}(i) = TXOP_{PA}(i) * fontesPA(i)$
 6. *emitirConjuntoDeParametrosDeQoS*()
-

Figura 9: Módulo principal do ajuste de parâmetros para n classes

A Figura 9 apresenta o módulo principal do mecanismo. O módulo chama o procedimento *ajustarContencao* para todas as classes, iniciando pela de maior prioridade. O processo realizado por esse procedimento é idêntico aquele apresentado na Seção 4.3 para o ajuste do parâmetro CW_{min} .

No passo seguinte, o módulo principal chama o procedimento *ajustarTXOP_Classes*, mostrado na Figura 10. O procedimento verifica qual classe recebeu o

ALGORITMO *ajustarTXOP_Classes*

1. $minimo = \infty$
 2. para cada classe i , i de 1 até $(n - 1)$
 3. $numPacotes(i) = \lfloor cargaAtual(i) / tamPacote(i) \rfloor$
 4. se $numPacotes(i) < minimo$
 5. $minimo = numPacotes(i)$
 6. para cada classe i , i de 1 até $(n - 1)$
 7. se $(numPacotes(i) = minimo)$
 8. /*a classe pode enviar um quadro*/
 9. $TXOP_{PA}(i) = 0$
 10. $TXOP_{STA}(i) = 0$
 11. senão
 12. $TXOP_{PA}(i) = tempoTx(i) * \lfloor numPacotes(i) / min \rfloor$
 13. $TXOP_{STA}(i) = tempoTx(i) * \lfloor numPacotes(i) / min \rfloor$
-

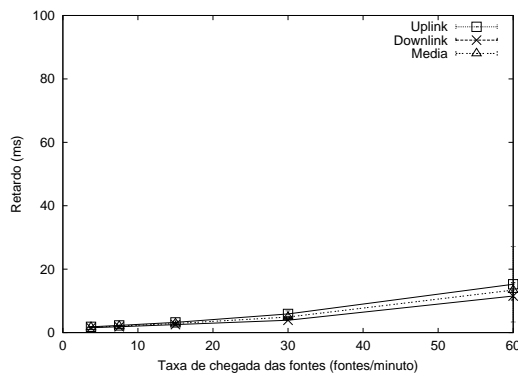
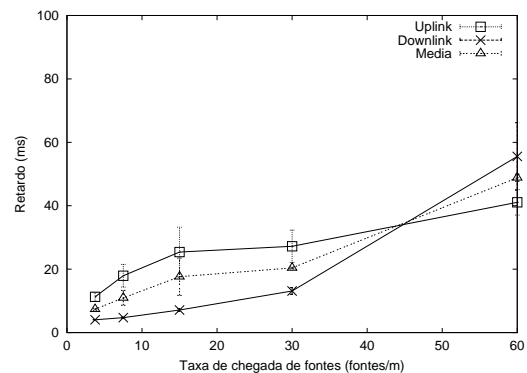
Figura 10: Módulo para ajuste do parâmetro TXOP para as n classes

menor número de pacotes no último intervalo de monitoramento, dividindo a carga atual da classe (número de bytes que chegaram no último intervalo de monitoração) pelo tamanho dos pacotes recebidos por ela. A classe, cujo número de pacotes recebidos é mínimo, recebe TXOP igual a zero e todas as outras classes recebem TXOP igual ao produto do tempo de transmissão de um pacote pela razão entre o número de pacotes recebidos pela classe e o número mínimo de pacotes recebidos. Informações sobre a carga atual em cada classe e o tamanho dos pacotes podem ser obtidas monitorando-se periodicamente as filas de transmissão no PA.

Em seguida, o módulo principal ajusta, no ponto de acesso, o TXOP para cada classe, multiplicando o valor atribuído no procedimento *ajustarTXOP_Classes* pelo número de fontes *downlink* ativas na classe.

Ao final, um novo conjunto de parâmetros de QoS é composto e deve ser transmitido para as estações.

5.2. Avaliação do Mecanismo

**Figura 11: Retardo das fontes de voz****Figura 12: Retardo das fontes de vídeo**

Os resultados apresentados, nesta seção, evidenciam a eficiência do mecanismo proposto em uma situação idêntica aquela apresentada na Seção 3. As Figuras 11 e 12 apresentam, respectivamente, o retardo dos tráfegos de voz e vídeo. Nota-se que, para qualquer que seja a carga na rede, o mecanismo de ajuste dinâmico dos parâmetros TXOP e CW_{min} é capaz de fornecer retardo semelhante e inferior ao valor máximo requisitado tanto para o tráfego *downlink* quanto para o retardo do tráfego *uplink*.

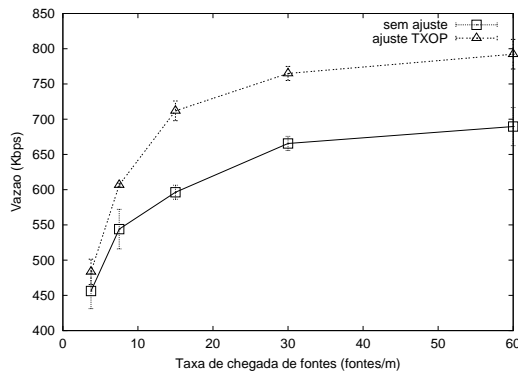


Figura 13: Vazão das fontes de voz

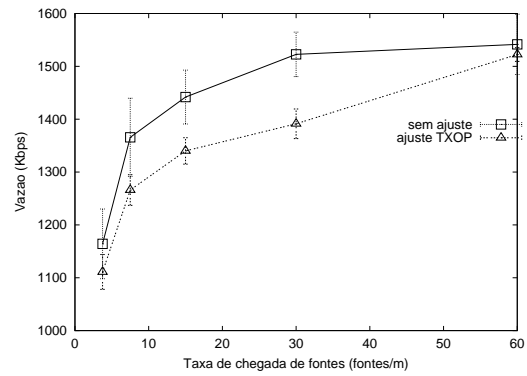


Figura 14: Vazão das fontes de dados

A partir destes resultados, pode-se concluir que com o ajuste dinâmico dos parâmetros TXOP e CW_{min} é possível realizar a manutenção da QoS requisitada pelas aplicações de tempo real e, ainda, fornecer retardo equivalente nos dois sentidos do tráfego mesmo em situações de alta carga na rede. É importante observar que o mecanismo proposto permite atingir os objetivos almejados sem, no entanto, provocar inanição do tráfego de melhor esforço ou reduzir a utilização da rede.

As Figuras 13 e 14 mostram, respectivamente, a vazão obtida pelos tráfegos de voz e dados nos experimentos realizados apenas com controle de admissão (sem ajuste) e com o mecanismo de ajuste dinâmico dos parâmetros TXOP e CW_{min} (ajuste TXOP). Quando se utilizou o mecanismo de ajuste dos parâmetros TXOP e CW_{min} , a vazão do tráfego de voz foi um pouco maior, dado que o retardo do tráfego *downlink* foi reduzido resultando em um maior número de fontes aceitas (em média 2 fontes de voz a mais foram aceitas) e, conseqüentemente, em um aumento na vazão. Por outro lado, a vazão do tráfego de dados sofreu uma pequena queda, sendo que o número de fontes aceitas foi o mesmo com e sem ajuste dos parâmetros. Como a taxa de chegada de pacotes do tráfego de dados é inferior a taxa de chegada do tráfego de tempo real, o mecanismo de ajuste dos parâmetros de acesso ao meio atribui valores menores para o TXOP da classe de dados, o que resulta em um aumento na taxa de perda de pacotes desta classe. Para o tráfego de vídeo, a vazão e o número de fontes aceitas foi semelhante em ambos os experimentos. A vazão agregada da rede não foi penalizada, nota-se, porém, que o mecanismo proposto permitiu que os recursos fossem alocados de acordo com a necessidade de cada classe para que o retardo requisitado pudesse ser atendido.

6. Conclusão

Este artigo apresentou duas propostas de mecanismos de ajuste dinâmico dos parâmetros de acesso ao meio da extensão 802.11e. Os mecanismos têm como objetivos auxiliar na tarefa de provisão de Qualidade de Serviço e resolver o problema da assimetria no retardo dos tráfegos *downlink* e *uplink*, dado que se observou, através de experimentos de simulação, que esta é a principal causa da violação dos requisitos de QoS quando a rede é exposta à altas cargas.

A primeira proposta, de ajuste dos parâmetros AIFS e CW_{min} , não apresentou resultados satisfatórios, dada a limitação imposta por estes parâmetros quanto a faixa de valores que pode ser utilizada para prover diferenciação entre a vazão das classes. Por outro lado, a partir da avaliação do segundo mecanismo proposto, pôde-se concluir que o ajuste dinâmico dos parâmetros TXOP e CW_{min} , aliado à um mecanismo de controle de admissão, permite fornecer ao tráfego de tempo real a QoS desejada e ao tráfego *downlink* um retardo semelhante ao retardo do tráfego *uplink*.

Referências

- [1] A. Lindgren, A. Almquist, e O. Schelen. *Evaluation of quality of service schemes for IEEE 802.11 wireless LANs*. In Proceedings of the 26th Annual IEEE Conference on Local Computer Networks, pg. 348-351, 2001.
- [2] F. Cali, M. Conti, e E. Gregori. *IEEE 802.11 Protocol: Design and Performance Evaluation of an Adaptive Backoff Mechanism*. IEEE Journal on Selected Areas in Communications, vol 18, pg. 1774-1786, 2000.
- [3] F. Cali, M. Conti, e E. Gregori. *Dynamic tuning of the IEEE 802.11 protocol to achieve a theoretical throughput limit*. IEEE/ACM Transactions on Networking, vol 8, pg. 785-799, 2000.
- [4] G. Bianchi, L. Fratta, e M. Oliveri. *Performance Evaluation and Enhancement of the CSMA/CA MAC Protocol for 802.11 Wireless LANs*. In Proceedings PIMRC, pg. 392-396, 1996.
- [5] Institute of Electrical and Electronics Engineers. *Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications*. IEEE Standard 802.11, 1999.
- [6] Institute of Electrical and Electronics Engineers. *Medium Access Control (MAC) enhancements for quality of service (QoS)*. IEEE Draft for Standard 802.11e/D3.2, 2002.
- [7] J. Filho. *Mecanismos de Controle de Qualidade de Serviço em Redes 802.11*. Tese de Mestrado, NCE/UFRJ, 2003.
- [8] J. Freitag, N. L. S. da Fonseca e J. F. de Rezende. *Mecanismos de Controle de Admissão Baseado em Medições em Redes IEEE 802.11e*. Anais do XXI Simpósio Brasileiro de Telecomunicações, 2004.
- [9] J. Qiu, e E. Knightly. *Measurement-based admission control with aggregate traffic envelopes*. IEEE/ACM Transactions on Networking, vol 9, pg. 199-210, 2001.
- [10] K. Fall, e K. Varadhan. *The NS Manual*. Relatório Técnico, UC Berkeley, The VINT Project, 2002.
- [11] L. Romdhani, Q. Ni e T. Turletti. *AEDCF: Enhanced Service Differentiation for IEEE 802.11 Wireless Ad-Hoc Networks*. In Proceedings of the IEEE Wireless Communications and Networking Conference, pp. 1373-1378, 2003
- [12] M. Grossglauser, e D. Tse. *A framework for robust measurement-based admission control*. Computer Communications Review, vol 27, pg. 237-248, 1997.
- [13] MPEG-4 and H.263 Video Traces for Network Performance Evaluation. URL: <http://trace.eas.asu.edu/TRACE/trace.html>, 2000.
- [14] P. Brady. *A technique for investigating on-off patterns of speech*. Bell Systems Technical Journal, vol. 44, pg. 1-22, 1965.
- [15] S. Jamin *et al.* *A measurement-based admission control algorithm for integrated services packet networks*. ACM SIGCOMM Computer Communication Review, vol 25, pg. 2-13, 1995.
- [16] S. Mangold *et al.* *IEEE 802.11e wireless LAN for quality of service*. In Proceedings of European Wireless, pg. 32-39, 2002.
- [17] W. Willinger, M. Taqqu, R. Sherman, e D. Wilson. *Self-similarity through high-variability: Statistical analysis of Ethernet LAN traffic at the source level*. IEEE/ACM Transactions on Networking, vol 5, pg. 71-86, 1997.