

Seleção Dinâmica de Parâmetros de Controle de Qualidade de Serviço em Redes IEEE 802.11 Infra-Estruturadas

José Coelho de Melo Filho^{1,*}, José Ferreira de Rezende²,
Luci Pirmez¹

¹Instituto de Matemática (IM) / Núcleo de Computação Eletrônica (NCE)
Universidade Federal do Rio de Janeiro (UFRJ)

²Grupo de Teleinformática e Automação (GTA) /
Coordenação dos Programas de Pós-Graduação em Engenharia (COPPE)
Universidade Federal do Rio de Janeiro (UFRJ)

jcoelho@posgrad.nce.ufrj.br, luci@nce.ufrj.br, rezende@gta.ufrj.br

Abstract. *This paper describes a control mechanism for dynamic selection of Quality of Service (QoS) parameters used in a IEEE 802.11e Wireless Local Area Network (WLAN). The 802.11e specification, still under development, provides a new contention-based access method, called Enhanced Distributed Coordination Function (EDCF). This method uses distinct values for interframe space and for contention window size to establish priorities between Access Categories (ACs). The proposed mechanism selects these values based on the actual network load condition and on the applications delay requirements. Such mechanism improves the new QoS functionality, aiming to promote better channel utilization while keeping the desired service levels.*

Resumo. *Este artigo descreve um mecanismo de controle que seleciona dinamicamente os parâmetros de Qualidade de Serviço (QoS) empregados em uma rede local sem fio compatível com a especificação IEEE 802.11e. Essa especificação, ainda em desenvolvimento, prevê um acesso com contenção avançado (EDCF) que utiliza valores diferenciados para o tamanho do espaço entre-quadros e o tamanho da janela de contenção a fim de estabelecer prioridades entre categorias de acesso. O mecanismo de controle proposto seleciona esses valores de acordo com a situação de carga da rede e com os requisitos de retardo das aplicações. Esse mecanismo complementa a nova funcionalidade de QoS e promove uma melhor utilização do canal enquanto mantém os níveis de serviço desejados.*

Palavras-chave: Comunicação sem fio, qualidade de serviço (QoS), controle de carga.

*Este trabalho foi realizado com recursos do NCE, CNPq, CAPES, COFECUB e FAPERJ.

1. Introdução

O padrão do IEEE (*Institute of Electrical and Eletronics Engineers*) para as redes locais sem fio (*Wireless Local Area Networks - WLANs*), conhecido como IEEE 802.11 [1], é uma das tecnologias mais promissoras para o cenário de acesso sem fio à Internet. Em relação a outras opções de acesso móvel, esse padrão se destaca por sua elevada popularidade, pelas altas taxas de transmissão de dados que oferece e pelo custo relativamente baixo de instalação e operação. O 802.11 permite o estabelecimento tanto de redes sem infra-estrutura, que utilizam apenas estações sem fio e são denominadas redes *Ad Hoc*, quanto de redes infra-estruturadas, nas quais um nó especial, chamado ponto de acesso (*Access Point -AP*), fornece às estações sem fio o acesso à rede fixa.

Para prover o padrão de um suporte à Qualidade de Serviço (QoS), o IEEE desenvolve a especificação 802.11e [2], denominada *MAC Enhancements for Quality of Service*, como uma extensão à especificação original. Tal extensão adiciona dois novos métodos de acesso aos já disponíveis na subcamada de controle de acesso ao meio (MAC) das WLANs, sendo um com contenção e outro sem contenção, ambos com funcionalidades de QoS. Esses novos métodos permitem aplicar tratamento diferenciado a categorias de tráfego com requisitos distintos de QoS.

No 802.11e, o controle da funcionalidade de QoS tanto no acesso sem contenção quanto no acesso com contenção é responsabilidade de um nó especial chamado coordenador híbrido. Esse coordenador é geralmente o próprio AP, que passa a exercer essa função em adição àquelas que já executa em sua área de cobertura e torna-se capaz de regular o acesso ao meio sem fio seguindo regras de diferenciação de serviços e prioridades pré-estabelecidas. No novo método de acesso com contenção, em particular, o coordenador híbrido pode emitir quadros de controle contendo um conjunto de parâmetros de QoS (*QoS Parameter Set*), usado pelas estações móveis para configurar as regras de acesso ao meio no MAC de suas interfaces de rede. Dessa forma, o coordenador é capaz de alterar a carga imposta à rede pelas estações, adaptando-se a mudanças nas condições de tráfego e estabelecendo ou modificando as políticas de acesso.

A extensão 802.11e cria o suporte à Qualidade de Serviço nas redes locais sem fio, mas não define como esse suporte deve ser utilizado. Questões relacionadas ao gerenciamento de Qualidade de Serviço de forma geral estão fora do escopo dessa especificação. Isso motiva o desenvolvimento de novos mecanismos, voltados para a monitoração e o controle dos níveis de serviço, complementando a funcionalidade de QoS introduzida com a extensão 802.11e.

No novo mecanismo de acesso com contenção, por exemplo, fica em aberto a questão de como o conjunto de parâmetros de QoS (*QoS Parameter Set*) deve ser composto para cada situação particular de carga e quando esse conjunto deve ser alterado.

O trabalho descrito neste artigo aborda o desenvolvimento de um mecanismo de controle de carga que visa capacitar o AP 802.11e a selecionar dinamicamente os parâmetros de QoS a serem empregados pelas estações móveis, durante o acesso com contenção, de acordo com a situação de carga a cada instante. As questões relacionadas ao controle da funcionalidade de QoS no acesso sem contenção estão fora do escopo deste trabalho.

O presente artigo está organizado da seguinte forma. Na Seção 2. são listados al-

guns trabalhos relacionados. A Seção 3. dá uma visão geral do padrão 802.11 e descreve algumas técnicas de diferenciação de serviços nas redes locais sem fio. Na Seção 4. é apresentado o mecanismo de controle de carga. Na Seção 5. são apresentados os resultados de simulações e avaliações do mecanismo de controle proposto. Por fim, a Seção 6. traz as conclusões.

2. Trabalhos Relacionados

Ahn *et al* [3], ao apresentar o modelo de rede SWAN (*Stateless Wireless Ad Hoc Networks*), propõem um mecanismo de controle de carga específico para redes 802.11. O controle de carga do modelo SWAN age diretamente sobre os parâmetros de configuração de baldes de fichas utilizados como modeladores de tráfego nas fontes de dados. Tal controle é executado nas estações, de uma forma distribuída. Embora possa ser adaptado para uso em redes infra-estruturadas, como é o caso das redes de acesso em um cenário de Internet Móvel, o controle de carga do modelo SWAN foi idealizado para redes *Ad Hoc* e adiciona uma complexidade computacional nas estações móveis que não é desejável quando se pretende suportar estações de baixo custo. O modelo SWAN também desconsidera modificações importantes na subcamada MAC, como a nova funcionalidade de QoS criada pela extensão 802.11e.

Romdhani, Ni e Turletti [4] apresentam um esquema chamado *Adaptive EDCF* (AEDCF), que considera o novo método de acesso com contenção do 802.11e (EDCF) e o torna mais adaptativo, obtendo resultados significativos com relação à eficiência na utilização do canal. Esse esquema, no entanto, também foi idealizado para redes *Ad Hoc* e para execução distribuída.

No mecanismo de controle de carga proposto neste trabalho, pretende-se manter restrita ao AP 802.11e toda a complexidade relacionada ao controle da Qualidade de Serviço, tendo em vista os cenários de acesso móvel às redes fixas como a Internet. O AP toma as decisões de controle baseado unicamente no conhecimento que detém sobre os fluxos que entram e saem da rede sem fio e nas restrições de cada tipo de tráfego. Para transmitir às estações os novos parâmetros de acesso é utilizado o próprio *QoS Parameter Set*, criado pela extensão 802.11e. Assim, o controle busca promover a diferenciação de serviços sem comprometer a eficiência na utilização do canal e sem adicionar complexidade às estações móveis, exigindo dessas estações apenas a compatibilidade com a extensão 802.11e.

3. Visão Geral do IEEE 802.11

O padrão 802.11 [1] foi criado pelo IEEE (*Institute of Electrical and Eletronics Engineers*) para definir redes locais sem fio (*Wireless Local Area Networks - WLANs*) compatíveis com os demais padrões de redes locais 802.x. A especificação atual, publicada em 1999, descreve a arquitetura, a camada física e a subcamada de controle de acesso ao meio (MAC) das WLANs.

Neste trabalho, são enfocadas particularmente as questões da subcamada MAC do 802.11, descrita a seguir.

3.1. Controle de Acesso ao Meio (MAC) do 802.11

A subcamada MAC do 802.11 define dois tipos de funções de coordenação para controlar o acesso ao meio: a função de coordenação distribuída (*Distributed Coordination Function* - DCF) e a função de coordenação centralizada (*Point Coordination Function* - PCF). A função PCF é de implementação opcional e está fora do escopo deste trabalho.

A função DCF provê um acesso com contenção e é de implementação obrigatória. Ela utiliza um protocolo de acesso múltiplo com detecção de portadora (CSMA) que usa um mecanismo de prevenção de colisão e reconhecimento positivo (ACK) dos quadros enviados. O método é conhecido como *Carrier Sense Multiple Access with Collision Avoidance* (CSMA/CA).

Na função DCF, o acesso pode ser feito usando dois esquemas distintos: um esquema de acesso básico (obrigatório), que usa apenas CSMA, e um acesso com RTS/CTS (opcional), que usa também quadros especiais de pedidos e permissões para transmitir dados.

No acesso básico, uma estação que quer transmitir "escuta" o meio, e caso ele esteja livre durante todo um intervalo de tempo chamado DIFS (*DCF Interframe Space*), a estação transmite. Caso contrário, a transmissão é adiada e inicia-se um processo de recuo exponencial binário (*Binary Exponential Backoff* ou *backoff*). Ao entrar em *backoff*, a estação escolhe um valor aleatório para dimensionar sua janela de contenção (*Contention Window* - CW), dentro do intervalo uniformemente distribuído que vai de zero a um tamanho mínimo da janela de contenção (*CWMin*) definido na especificação. O valor escolhido é multiplicado pela duração de um slot (*slottime*) para a criação de um temporizador de *backoff*. Esse temporizador passa a ser decrementado periodicamente enquanto o meio estiver livre. O período de decremento é dado pelo parâmetro *slottime* do MAC, que também é definido no padrão. O temporizador é parado quando alguma transmissão é detectada no meio, reiniciando quando o meio fica livre novamente. Quando finalmente o temporizador expira, a estação envia o seu quadro.

A estação receptora recebe o quadro de dados e envia um quadro de reconhecimento positivo (ACK), após um tempo chamado espaço curto entre-quadros (*Short Interframe Space* - SIFS), que é sempre menor que DIFS. Se a transmissão foi bem sucedida e há mais algum quadro a transmitir, o transmissor entra em *backoff*, para evitar a captura do meio. Caso o ACK não seja recebido no tempo esperado, a estação transmissora escalona uma retransmissão e também entra em *backoff*.

O segundo esquema de acesso usado no DCF (Figura 1) inclui quadros especiais de pedidos e de permissões para transmitir (*Request To Send* - RTS e *Clear To Send* - CTS). Nesse tipo de acesso, os quadros RTS e CTS contém os endereços dos nós de origem e de destino e a duração total da transmissão do quadro de dados, incluindo o tempo necessário para a recepção no transmissor do quadro de reconhecimento (ACK) enviado pelo receptor. A estação transmissora envia um RTS ao receptor após DIFS segundos de meio livre, antes da transmissão do quadro de dados, a fim de reservar o meio (Figura 1). O receptor, caso esteja pronto para receber, responde com um CTS após o meio estar livre por um tempo SIFS. Ao receber o CTS, o transmissor espera também por SIFS segundos e inicia o envio do quadro de dados, passando a aguardar o ACK correspondente, como

no acesso básico. Caso não receba o CTS, o transmissor entra em *backoff* e retransmite o RTS.

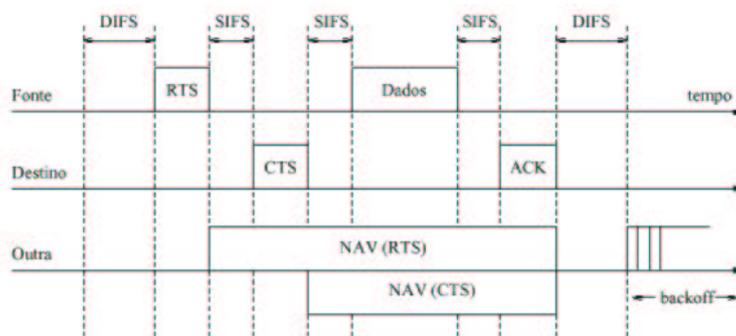


Figura 1: DCF utilizando RTS e CTS.

Todas as estações que escutam o RTS ou o CTS utilizam a informação da duração da transmissão contida nesses quadros para atualizar o seu vetor de alocação de rede (*Network Allocation Vector* - NAV). O NAV é utilizado para a detecção virtual da portadora (Figura 1). Com essa informação, as demais estações não envolvidas na transmissão atual irão evitar fazer acesso ao meio no período de tempo reservado. Desse modo, qualquer terminal escondido poderá adiar a sua transmissão para evitar colisões. O uso de RTS e CTS é controlado por estação através de um limiar de RTS ($RTS_{threshold}$); a estação transmissora usará RTS somente na transmissão de quadros com tamanhos maiores que esse limiar.

Para reduzir a probabilidade de colisões repetidas entre estações tentando transmitir, a cada transmissão não sucedida o valor de CW aumenta para uma próxima potência de 2 menos 1, até que seja atingido um valor máximo pré-definido chamado CW_{Max} (1023 no 802.11b DSSS). Caso um número máximo de tentativas de transmissões seja alcançado (7 no acesso básico ou 4 no acesso com RTS/CTS), o quadro é descartado.

3.2. Diferenciação de Serviços no MAC 802.11

O MAC 802.11 apresenta limitações no que diz respeito à justiça de compartilhamento de banda, desempenho e suporte a aplicações de tempo real. Rubinstein e Rezende [5] resumem alguns dos principais problemas apontados na literatura. Esses problemas, aliados à necessidade imperativa de provisão de QoS nas redes locais sem fio, sinalizaram a necessidade de evolução do MAC 802.11.

Diversas propostas de diferenciação de serviços foram elaboradas. Rubinstein e Rezende [5] descrevem em maiores detalhes algumas dessas propostas. Elas geralmente envolvem um ou mais parâmetros básicos do MAC 802.11 para promover diferenciação, alterando a forma com que cada uma das estações faz acesso ao meio e estabelecendo prioridades de acesso entre elas. Algumas técnicas foram elaboradas para o acesso com contenção (DCF) e outras para o acesso sem contenção (PCF).

Para o acesso com contenção, foco do presente trabalho, as técnicas de diferenciação envolvem basicamente o tamanho do intervalo entre-quadros da função DCF (DIFS), alterações no cálculo do tamanho da janela de contenção (CW) e o tamanho máximo dos quadros de dados.

O tamanho do intervalo entre-quadros da função DCF (DIFS) controla o tempo que uma estação sempre espera antes de tentar iniciar uma transmissão. Uma estação de baixa prioridade, que use um valor de DIFS maior, terá menor probabilidade de encontrar o meio livre comparada com outra de alta prioridade, que esteja usando um DIFS menor.

O procedimento de *backoff* também controla o tempo que uma estação espera antes de tentar uma transmissão. A técnica consiste na escolha de intervalos diferenciados para as janelas de contenção de cada estação, controlados pelos parâmetros CWMin e CWMax do MAC. Uma estação com CWMin menor tenderá a escolher valores para o temporizador menores que outra com CWMin maior. Esperando mais tempo, a segunda estação terá menos chances de obter o meio que a primeira, ficando portanto com prioridade mais baixa. O efeito é similar à alteração do DIFS, com a diferença que o DIFS é uma parte determinística da espera, enquanto o tempo de *backoff* é probabilístico.

O tamanho do quadro também promove diferenciação, uma vez que a disputa pelo acesso é realizada quadro a quadro. Estações com maior prioridade utilizam quadros de maior tamanho nas suas transmissões, garantindo um maior tempo de utilização do meio cada vez que esta estação ganha a disputa. No entanto, quadros maiores estão mais sujeitos a erros, reduzindo a eficácia deste mecanismo em meios com taxas altas de erros [6].

3.3. A Extensão 802.11e

Atento às necessidades de padronização e aos desenvolvimentos recentes na área de QoS em redes 802.11, o IEEE criou o Grupo de Tarefas E [7] dentro do Grupo de trabalho 11 para especificar uma extensão ao padrão que introduzisse mecanismos de provisão de QoS no MAC 802.11. Essa extensão, denominada *MAC Enhancements for Quality of Service*, ainda está em desenvolvimento, de forma que as definições que serão aqui apresentadas estão sujeitas a mudanças.

No 802.11e, as estações (STAs) são chamadas de QSTAs (*QoS Stations*), o ponto de acesso de QAP (*QoS Access Point*) e o BSS de QBSS (*QoS Basic Service Set*). A extensão incorpora uma nova função de coordenação, a Função de Coordenação Híbrida (*Hybrid Coordination Function - HCF*). Nessa função, um Coordenador Híbrido (*Hybrid Coordinator - HC*), normalmente localizado no QAP, é responsável pelo estabelecimento e coordenação da funcionalidade de QoS dentro do QBSS. A função híbrida combina características do PCF e do DCF, introduzindo novas características específicas para QoS. A especificação permite coexistência com as estações legadas (802.11/b/a), de forma transparente para essas estações.

Um conceito importante no 802.11e é o de oportunidade de transmissão, ou TxOP (*Transmission Opportunity*). Uma TxOP é definida como um intervalo de tempo no qual uma estação tem o direito de iniciar transmissões. Ela é marcada por um tempo de início e uma duração. Durante uma TxOP, uma estação pode enviar vários quadros em rajada, separados por SIFS, sem ter que disputar o meio quadro a quadro. A duração de uma TxOP é limitada, para evitar retardo excessivo nas outras estações que estão aguardando o acesso.

A função HCF também fornece dois métodos de acesso: um baseado em consulta, chamado *HCF polled channel access*, que é executado de forma centralizada no HC, e outro, baseado em contenção, chamado *HCF contention-based channel access*, que é executado nas estações.

No acesso baseado em contenção, foco do presente trabalho, as estações competem pelas TxOPs de uma forma distribuída, semelhante ao que ocorre no DCF. Para isso, elas usam uma variante avançada do DCF, chamada *Enhanced DCF* ou EDCF. O EDCF incorpora o conceito de múltiplas filas de transmissão em uma mesma estação, cada uma com um nível de prioridade. O mecanismo é baseado em uma proposta anterior chamada Virtual-DCF ou VDCF [8, 9]. Nessa proposta, dentro de uma mesma estação, várias instâncias do DCF são executadas em paralelo, como se fossem "MACs virtuais", com filas diferentes e parâmetros de QoS (DIFS e CWMin) distintos para cada uma (Figura 2). Essas instâncias, definidas como categorias de acesso (*Access Categories - ACs*), executam o procedimento de *backoff* independentemente umas das outras, e competem entre si pelas oportunidades de transmissão. Cada estação 802.11e pode suportar de 2 a 8 categorias de acesso.

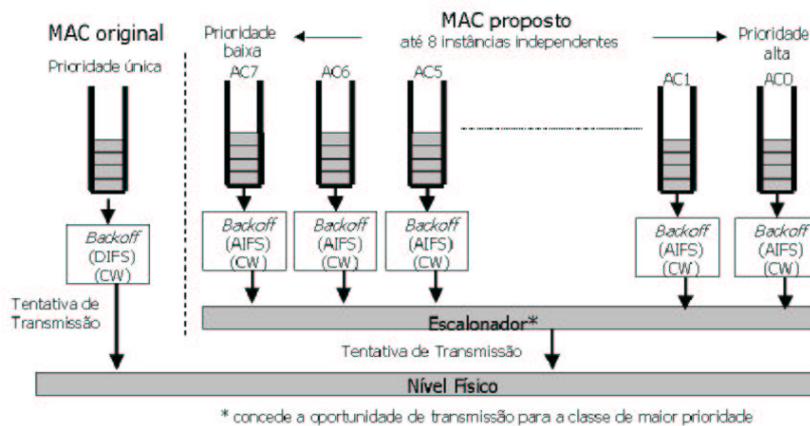


Figura 2: MAC Proposto no EDCF

A diferenciação entre ACs é obtida com o uso conjunto da variação da duração mínima de espera por meio livre e do tamanho da janela de contenção (DIFS e CWMin do DCF). O DIFS de cada categoria é chamado de *AIFS - Arbitration Interframe Space*. As categorias com prioridades mais altas usam valores menores de AIFS. Os conflitos entre essas categorias são resolvidos internamente pela estação, de forma que a categoria de maior prioridade receba a TxOP e a de menor prioridade execute o procedimento de *backoff*.

Os parâmetros de QoS a serem usados são definidos pelo HC sempre com base em 8 prioridades de usuários (*User Priorities - UPs*). Um conjunto de parâmetros de QoS (*QoS Parameter Set ou EDCF Parameter Set*) é usado para informar às QSTAs os parâmetros necessários para o estabelecimento da funcionalidade de QoS. Esse conjunto, que é transmitido pelo HC encapsulado em quadros de controle como os quadros de *Beacon* e *Probe Response* [1], fornece o valor de $AIFS(i)$ e do $CWMin(i)$ a ser usado para as prioridades de 0 a 7. A duração máxima das oportunidades de transmissão obtidas por contenção é definida em um terceiro parâmetro, o *TxOP Limit*. Esse parâmetro é informado em um campo de 2 octetos que indica um número de períodos de $16 \mu s$ correspondente à duração máxima das TxOPs. Ao contrário do $AIFS(i)$ e do $CWMin(i)$, esse valor é configurado para todo o QBSS, e não por prioridade de usuário.

4. Mecanismo de Controle de Carga

A definição dos parâmetros transmitidos no *QoS Parameter Set* é realizada, neste trabalho, por um mecanismo de controle de carga. Esse mecanismo monitora periodicamente a situação de carga da rede, coletando estatísticas nas filas das categorias de acesso (ACs) do AP, como a carga média observada no período e o retardo de acesso dos quadros de dados. Dependendo da situação, o controle altera os valores de $AIFS(i)$ e $CWMin(i)$ de cada categoria, buscando manter limitado o retardo de acesso dos quadros das categorias mais prioritárias e ao mesmo tempo elevar o nível de utilização da rede nas categorias menos prioritárias ou de melhor esforço.

4.1. Efeitos dos Parâmetros AIFS e CWMin

O AIFS e o CWMin controlam o tempo de espera, mas existem algumas diferenças entre esses parâmetros que devem ser consideradas. O AIFS controla uma parte determinística da espera, enquanto o CWMin controla uma parte probabilística. O CWMin também afeta o retardo médio de acesso de uma fonte muito mais gravemente que o AIFS, uma vez que esse parâmetro só assume valores que são potências de 2 subtraídas de 1 unidade. Com AIFS, é possível acrescentar um slot de cada vez ao tempo de espera, o que permite um controle menos agressivo.

Outras diferenças podem ser constatadas estudando-se algumas situações de sobrecarga da rede. As Figuras 3 e 4 mostram o efeito do AIFS sobre a vazão agregada obtida por conjuntos de 1 e 8 fontes, respectivamente. O tamanho do intervalo entre-quadros da categoria é definido por $SIFS + AIFS(i) \cdot SlotTime$, onde $AIFS(i)$ expressa um número de slots. As fontes pertencem a uma mesma categoria de acesso e geram dados a uma taxa de bits constante (*Constant Bit Rate - CBR*). Todas as estações pertencem a um mesmo BSS 802.11b e utilizam a taxa de transmissão de dados máxima de 11 Mbps. Nas duas figuras, cada curva mostra a soma da vazão agregada do conjunto de fontes em um experimento isolado. Em cada experimento é utilizado um determinado tamanho de pacote (p), que varia de 64 a 2304 bytes.

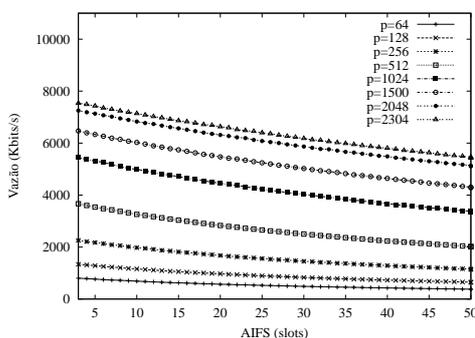


Figura 3: Vazão de 1 Fonte CBR x Tamanho de AIFS

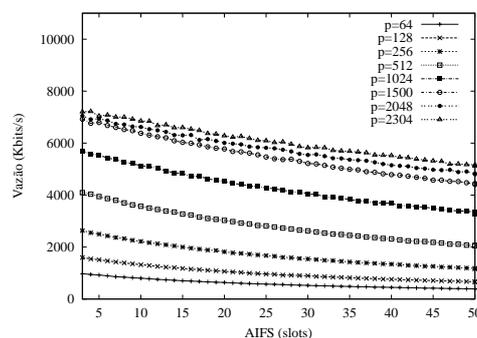


Figura 4: Vazão de 8 Fontes CBR x Tamanho de AIFS

Observando-se as Figuras 3 e 4, percebe-se que o aumento do número de slots utilizados no AIFS provoca uma redução linear na vazão agregada de cada conjunto, independentemente do número de fontes presentes no BSS.

As Figuras 5 e 6 mostram os mesmos experimentos anteriores, realizados com a variação do CWMin. Os valores para CWMin evoluem de 31 a 4095, fazendo-se periodicamente $CWMin = (CWMin * 2) + 1$. O AIFS é mantido fixo e igual a 2.

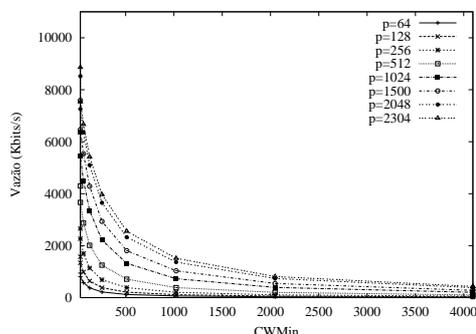


Figura 5: Vazão de 1 Fonte CBR x Tamanho de CWMin

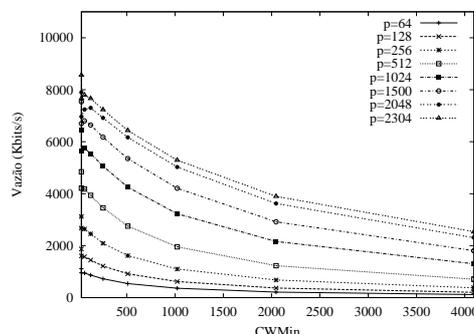


Figura 6: Vazão de 8 Fontes CBR x Tamanho de CWMin

Enquanto para uma única fonte o aumento exponencial do tempo de espera provoca uma grande redução na vazão obtida, para 8 fontes, no entanto, observa-se uma curva de perfil diferente, mais imune às variações de CWMin. Isso pode ser explicado considerando-se a importância que o tamanho da janela de contenção tem na prevenção de colisão do CSMA/CA. Em situações de sobrecarga, como as exploradas nesses experimentos, a probabilidade de colisão aumenta com o número de instâncias do MAC que competem pelo meio. Cada instância que perde uma oportunidade de transmissão é obrigada a entrar em *backoff* usando um valor maior de CW. O tempo médio gasto em *backoff* aumenta com o número médio de tentativas necessárias para uma transmissão bem sucedida. Elevando-se o valor de CWMin, reduz-se a probabilidade de que duas instâncias escolham o mesmo slot, aumentando a chance de transmissão na primeira tentativa e nas subsequentes. A vazão agregada só começa a cair por alteração do CWMin quando o valor desse parâmetro se torna excessivo em relação ao valor ideal para a prevenção de colisão.

4.2. Restrições

As variações de AIFS e CWMin podem ser empregadas em conjunto, de forma complementar. O controle de carga usa o AIFS para definir o grau de separação entre os níveis de prioridade, enquanto o CWMin é alterado de acordo com o número de instâncias do EDCF competindo em uma mesma categoria. No entanto, algumas restrições devem ser observadas na determinação desses valores.

- Restrição 1: $AIFS(i) \geq 2, \forall i$, para evitar falha no funcionamento do MAC. O valor de AIFS(i) não pode ser 0, para evitar que esse intervalo fique igual a SIFS, nem pode ser 1, para não se igualar ao intervalo entre-quadros da função PCF (PIFS), que é definido por $SIFS \cdot SlotTime$.
- Restrição 2: $AIFS(i + 1) \geq AIFS(i)$, para preservar a diferenciação. Se $AIFS(i + 1) = AIFS(i)$, então $CW(i + 1) > CW(i)$ obrigatoriamente. Isso permite que as categorias menos prioritárias explorem ao máximo a capacidade do canal ao mesmo tempo em que mantém a possibilidade de diferenciação caso a atividade nas categorias mais prioritárias reinicie.

- Restrição 3: $AIFS(i+1) < AIFS(i) + CW(i)$ se a carga da categoria i for alta, para evitar o efeito de inanição na categoria $i+1$, que ocorre quando a categoria menos prioritária nunca tem oportunidade de transmitir.
- Restrição 4: $CWMin(i) < CWMax, \forall i$, conforme previsto na especificação do 802.11.
- Restrição 5: $CWMin \geq 7$ para a categoria mais prioritária, considerando-se que esse valor é um mínimo razoável para permitir um compromisso entre baixo retardo e prevenção de colisão. Para a categoria de melhor esforço (menos prioritária), $CWMin \geq 31$, a fim de evitar que tenha prioridade sobre as estações legadas (802.11/11b/11a). Em geral, uma categoria i não pode ter $CWMin$ menor que $2^{3+i} - 1$, para que a restrição 2 também seja atendida.

4.3. Algoritmos

O controle de carga é executado periodicamente, de acordo com um intervalo de monitoração. A cada execução, o controle verifica se o tamanho do $CWMin$ usado nas categorias está adequado, utilizando a função `ajustarContencao(i)`, a partir da prioridade mais alta. Em seguida, realiza o ajuste da carga imposta pelas categorias de menor prioridade, aumentando ou diminuindo os valores de $AIFS$ e $CWMin$ dessas categorias, sempre tendo em vista as restrições e os objetivos de diferenciação, como retardo limitado ou vazão mínima em uma determinada categoria. Após ter executado essas duas operações em todas as categorias, o controle emite um novo conjunto de parâmetros de QoS.

As Figuras 7, 8 e 9 mostram os algoritmos utilizados para um caso específico de duas categorias, enfocando uma situação de coexistência de tráfego de tempo real com tráfego de melhor esforço.

```

controlarCarga() /* executado a cada intervalo de monitoração */
{
    ajustarContencao(i)
    ajustarCarga(i+1)
    emitirConjuntoDeParametrosDeQoS()
}

```

Figura 7: Módulo Principal do Controle de Carga

Na Figura 7 está o módulo principal do controle. No caso de duas categorias, o módulo chama o procedimento `ajustarContencao` para a categoria mais prioritária, mostrado na Figura 8.

Nesse procedimento, é verificado se o $CWMin$ da categoria deve ser maior para reduzir o número de tentativas de transmissão. Isso é feito comparando-se o valor de $CWMin$ com o número de fontes aceitas na categoria, que é informado por um mecanismo de controle de admissão do tráfego de tempo real ¹.

Para decidir se o tamanho do $CWMin$ deve ser alterado, é utilizada uma regra simples: se o número de fontes for maior que o número de slots usados no $CWMin$,

¹O tráfego de tempo real deve estar sujeito a alguma forma de controle de admissão, para evitar que a capacidade máxima do canal seja exaurida por este tráfego. Em um trabalho anterior a este [10], foi desenvolvida a proposta de um mecanismo para essa finalidade, adaptando-se esquemas de controle de admissão estudados em redes fixas.

```

ajustarContencao(i) /* ajusta o CWMin da categoria dependendo do
                    número de instâncias do EDCF nessa categoria */
{
  CWAntigo = CWMin(i)
  se FontesAceitas > CWMin(i) /* FontesAceitas vem do Controle de Admissão */
    CWMin(i) = CWMin(i) * 2 + 1
  se FontesAceitas < CWMin(i)/2
    CWMin(i) = (CWMin(i) - 1) / 2
  se CWMin(i) < 2 * (3 + i) - 1
    CWMin(i) = 2 * (3 + i) - 1 /* restrição 5: CWMin=7 para i=0 */
  se CWMin(i) > CWAntigo /* houve aumento no CWMin da categoria */
    AIFS(i+1) = AIFS(i+1) + [(CWMin(i) - CWAntigo) * δ]
  se CWMin(i) < CWAntigo /* houve redução no CWMin da categoria */
    AIFS(i+1) = AIFS(i+1) - [(CWAntigo - CWMin(i)) * δ]
  se AIFS(i + 1) < 2
    AIFS(i+1) = 2 /* restrição 1 */
}

```

Figura 8: Ajuste do Tamanho do CWMin na Categoria de Tempo Real

então CWMin deve ser maior. Com essa simplificação, busca-se manter a probabilidade de escolha de um slot sempre próxima a $1/FontesAceitas$. O procedimento verifica também se CWMin pode ser reduzido, comparando a metade desse valor com o número de fontes aceitas.

Caso o tamanho de CWMin da categoria de tempo real seja alterado, haverá uma ampliação ou redução da parte determinística da espera da categoria de melhor esforço. Isso é feito somando-se ao AIFS da categoria $i + 1$ a diferença entre o CWMin antigo e o novo da categoria i . É possível utilizar apenas uma fração dessa diferença, multiplicando-a por um fator δ , que varia de 0 a 1.

Em seguida, o módulo principal chama o procedimento ajustarCarga(i+1) (Figura 9), que verifica se a carga da categoria de melhor esforço pode ser elevada ou deve ser reduzida. Se o retardo experimentado pela categoria de tempo real estiver acima de um máximo, a carga do melhor esforço deve ser reduzida. Se estiver abaixo de um limiar inferior ("mínimo"), que representa uma condição de sobra de recursos para o tipo de tráfego de tempo real, a carga do melhor esforço pode ser elevada. Para evitar instabilidade do controle, uma condição adicional é colocada no aumento de carga. Essa condição sugere que houve a saída de uma fonte de tempo real do período anterior para o atual.

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, CWMin passa a ser dobrado, até o limite da restrição 4. A elevação é feita 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 é bom para a estabilidade do mecanismo.

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

```

ajustarCarga(i+1) /* verifica os limites de qualidade */
{
  se RetardoAtual(i) > RetardoMáximo(i)
    reduzirCarga(i+1)
  senão se (RetardoAtual(i) <
RetardoMínimo(i))e(CargaAtual(i) <= CargaAnterior(i) - CargaMediaDaFonte(i))
    aumentarCarga(i+1)
}
reduzirCarga(i+1) /* baixa a carga da categoria i */
{
  se AIFS(i + 1) < (AIFS(i) + CWMin(i)) /* restrição 3 */
    AIFS(i+1) = AIFS(i+1) + SlotsReducao
  senão se CWMin(i + 1) < CWMax /* restrição 4 */
    CWMin(i+1) = (CWMin(i+1) * 2) + 1
}
aumentarCarga(i+1) /* eleva a carga da categoria i */
{
  se CWMin(i + 1) > 31 /* restrição 5 */
    CWMin(i+1) = (CWMin(i+1) - 1) / 2
  senão se AIFS(i + 1) > 2 /* restrição 1 */
    AIFS(i+1) = AIFS(i+1) - SlotsAumento
}

```

Figura 9: Ajuste da Carga da Categoria de Melhor Esforço

5. Simulações

A ferramenta de simulação empregada na avaliação do mecanismo de controle de carga proposto foi o Network Simulator (NS), versão 2.1b8a. Foi empregado um agente de roteamento chamado NOAH (*NOn Ad Hoc routing agent*). Foram necessárias ainda modificações no código do MAC 802.11, a fim de suportar as funcionalidades do EDCF. A implementação para múltiplas filas tomou como ponto de partida o código do Virtual DCF [9].

5.1. Cenário

O cenário de simulação consiste de um AP conectado a um nó fixo por meio de um enlace de 100 Mbps, com retardo de 2 ms. O ponto de acesso (QAP) está localizado no centro de uma área de 350 x 350 metros, de forma que todos os nós móveis contidos na área estejam a uma distância menor que 250 metros. A taxa de dados utilizada é de 11 Mbps e a taxa básica é de 1 Mbps. O cenário, mostrado na Figura 10, simula a situação de uma WLAN 802.11 fornecendo acesso à rede fixa. As simulações são feitas sem o uso de RTS\CTS, ou seja, utilizando o esquema de acesso básico.

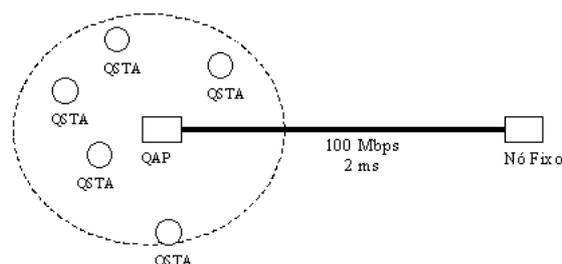


Figura 10: Cenário usado nas simulações.

A avaliação do controle de carga explora a coexistência de tráfego de tempo real com tráfego de melhor esforço. O tráfego de tempo real é modelado por fontes de voz, empregando fontes do tipo *onoff*, com períodos de atividade e de silêncio seguindo distribuições exponenciais com médias 1,2 e 1,8 s, respectivamente. Essas fontes geram pacotes de 210 bytes a cada 26 ms, perfazendo uma taxa de 64 kbps. Tal configuração equivale a um codificador PCM (*Pulse Code Modulation*) com supressão de silêncio. O protocolo de transporte utilizado no tráfego de voz é o UDP. São permitidas até 25 fontes de voz na rede sem fio, utilizando-se um controle de admissão como o descrito em [10].

As fontes de voz têm um intervalo entre chegadas que segue uma distribuição uniforme, indo de 0 a 7 s. Para simular o padrão conversacional de cada conexão de voz, foi empregada uma fonte na direção do nó móvel para o nó fixo (sentido de subida ou *uplink*) e outra na direção do nó fixo para o móvel (sentido de descida ou *downlink*). Essa bidirecionalidade é necessária, já que em uma rede 802.11 o tráfego do sentido de descida compete pelo meio sem fio com o tráfego do sentido de subida.

O tráfego de melhor esforço simula conexões TCP de longa duração, empregando 8 fontes FTP que realizam uma transferência de dados contínua, gerando pacotes de 576 bytes. A origem das conexões FTP está no nó fixo e o destino nos nós móveis. O tráfego de voz usa a categoria de mais alta prioridade (AC0) e o tráfego FTP usa a categoria menos prioritária (AC1). A categoria de acesso para o tráfego de voz (AC0) emprega AIFS=2 e CWMin=7. A categoria de acesso do FTP (AC1) inicia com AIFS=9 e CWMin=31.

As simulações duram 600 s. Para facilitar a observação do efeito da entrada e saída das fontes de voz sobre a vazão obtida pelo tráfego FTP, as fontes só são admitidas até 150 s de simulação e têm duração fixa de 250 s. Assim podem ser observadas 4 fases: uma fase de admissão (Fase 1) entre 0 e 150 s, uma fase estável (Fase 2) entre 150 e 250 s, onde o número máximo de fontes de voz está presente, uma fase de saída das fontes (Fase 3), a partir de 250 s (podendo ir até 400 s, dependendo do tempo de chegada da última fonte) e, por último, uma fase onde somente o tráfego FTP e a primeira fonte de voz estão presentes (Fase 4). Esta primeira fonte é usada, durante toda a simulação, para monitorar o retardo do tráfego de subida.

A fim de estabelecer uma base de comparação para o desempenho do controle de carga, foram realizadas simulações com e sem o uso do controle. Nas simulações com controle de carga, o intervalo de monitoração é de 3 s. Os retardos de acesso máximo e mínimo para os quadros de voz, usados em `ajustarCarga()`, são de 20 ms e 4 ms, respectivamente². Nesse mesmo procedimento, são utilizados 1 slot para *SlotsAumento* e 4 para *SlotsReducao*. O parâmetro δ usado em `ajustarContencao()` é igual a 0.8.

5.2. Resultados

Os resultados obtidos para as situações com e sem controle de carga refletem medidas realizadas em 10 execuções do cenário. Em cada execução são medidas a vazão agregada do tráfego FTP (média durante toda a simulação) e o retardo de acesso médio de todos os pacotes de voz que passam pela fila de transmissão sem fio do AP no sentido de descida (*downlink*).

²O retardo máximo, 20 ms, foi propositalmente fixado em um valor menor que o intervalo de geração dos pacotes pela aplicação de voz.

A Tabela 1 mostra o retardo de acesso dos quadros de voz no sentido de descida. A coluna "retardo" mostra a média e o intervalo de confiança de 95% (IC-95) das médias do retardo observado nas 10 execuções. As colunas "P-90", "P-95" e "P-99" mostram, respectivamente, os percentuais 90, 95 e 99 dos valores de retardo. Os dados obtidos

Tabela 1: Retardo de Acesso dos Quadros de Voz

| | Sem Controle de Carga | | | | Com Controle de Carga | | | |
|-------|-----------------------|------|-------|--------|-----------------------|------|------|------|
| | retardo(ms) | P-90 | P-95 | P-99 | retardo(ms) | P-90 | P-95 | P-99 |
| média | 105.60 | 4.26 | 14.63 | 123.59 | 4.20 | 1.82 | 3.07 | 7.38 |
| IC-95 | 44.84 | 1.06 | 7.97 | 68.90 | 1.53 | 0.15 | 0.24 | 2.01 |

revelam que o controle de carga contribuiu para manter o retardo de acesso limitado, abaixo dos 20 ms pretendidos. Tanto o retardo médio quanto o percentual-99 de retardo ficaram bem abaixo do limite máximo em todas as 10 execuções, o que não ocorreu nas situações sem controle de carga. Nessas situações, a definição estática dos valores de AIFS e CWMin mostrou-se insuficiente para garantir qualquer controle sobre os limites de retardo do tráfego de tempo real na presença do tráfego de melhor esforço. Caso fossem utilizados valores maiores para AIFS e CWMin para a categoria do FTP (AC1), seria possível obter valores de retardo menores para o tráfego de voz mesmo sem controle de carga, porém a vazão do tráfego FTP seria excessivamente reduzida.

O controle de carga, por outro lado, torna possível manter o retardo limitado sem sacrificar excessivamente o tráfego de melhor esforço, pois reduz a carga do FTP apenas nos instantes mais críticos para o tráfego de voz, e a eleva sempre que possível. A Tabela 2 mostra a vazão média agregada do tráfego FTP. A coluna "média" expressa a média das 10 execuções (com e sem controle de carga) e a coluna "IC-95" revela o intervalo de confiança de 95%. A primeira linha mostra as medidas considerando toda a simulação, e as quatro restantes mostram os resultados para cada uma das fases citadas anteriormente.

Tabela 2: Vazão Média do Tráfego FTP

| período | Sem Controle de Carga | | Com Controle de Carga | |
|-----------|-----------------------|-------|-----------------------|--------|
| | média(Kbps) | IC-95 | média(Kbps) | IC-95 |
| 0 a 600 s | 1319.23 | 24.69 | 1206.54 | 47.32 |
| Fase 1 | 1025.74 | 36.96 | 816.75 | 40.29 |
| Fase 2 | 625.47 | 51.40 | 341.05 | 46.51 |
| Fase 3 | 1609.62 | 36.33 | 1533.26 | 125.87 |
| Fase 4 | 2010.43 | 11.72 | 2186.46 | 33.83 |

Na primeira linha dessa tabela observa-se que o controle provocou uma redução de apenas 8,5% na vazão do tráfego de melhor esforço, considerando todo o período de simulação. Esse resultado, bastante significativo, é similar ao atingido por Ahn *et al* com o modelo de rede SWAN [3], onde um retardo 60 a 75% menor para o tráfego de tempo real foi obtido com redução de apenas 2% da vazão do tráfego de melhor esforço.

Nas medidas das fases 1, 2 e 3 verifica-se que a vazão foi sempre menor no caso com controle de carga, especialmente na fase 2, quando há o maior número de fontes de voz presentes na rede. Em contrapartida, na fase 4 a vazão do tráfego FTP chega a

ser maior no caso com controle, porque o procedimento `ajustarCarga()` reconhece que há pouco tráfego de voz e é capaz de selecionar valores menores de AIFS e CWMin até chegar a AIFS=2 e CWMin=31 (valores *default* usados em uma rede 802.11b).

As Figuras 11 e 12 mostram o valor máximo (P-100) do retardo de acesso dos pacotes de voz em duas execuções particulares da simulação, uma sem e outra com controle de carga. Os valores são coletados periodicamente no AP, por monitoração da fila de transmissão da categoria AC(0). Observa-se que, sem controle de carga, o limite de retardo pretendido (20 ms) é violado várias vezes entre 100 e 300 s de simulação, chegando a picos de mais de 600 ms. Essa situação é crítica para o tráfego de voz, considerando-se que para esse tipo de aplicação deve ser garantido um retardo fim-a-fim ideal de 150 ms. Com controle de carga, é possível confinar mais facilmente o retardo entre um limite máximo e um mínimo.

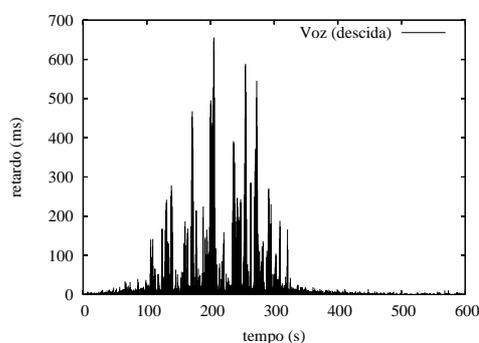


Figura 11: Retardo de Acesso do Tráfego de Voz sem Controle de Carga

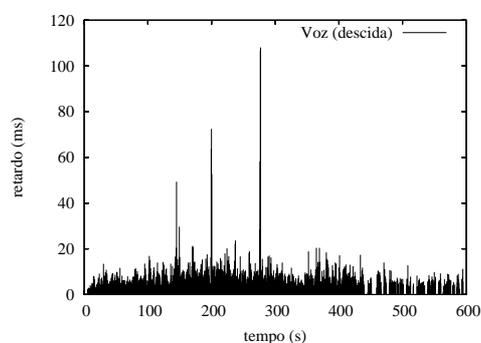


Figura 12: Retardo de Acesso do Tráfego de Voz com Controle de Carga

6. Conclusões

Este artigo descreveu um mecanismo de controle de carga que seleciona os parâmetros de diferenciação empregados na função de coordenação distribuída do padrão 802.11e, de acordo com a situação de carga da rede e com os requisitos das aplicações. O mecanismo proposto foi avaliado por meio de simulações que exploravam cenários de coexistência de tráfego de tempo real com tráfego de melhor esforço.

Os resultados obtidos são compatíveis com os resultados de outros trabalhos [3, 4], e atestam que a seleção dinâmica dos parâmetros de diferenciação é preferível em relação a uma seleção estática, pois permite um maior controle do retardo do tráfego de tempo real em situações de carga elevada e uma maior utilização do canal quando há pouco tráfego de tempo real presente na rede.

O mecanismo tem a vantagem de manter a complexidade de controle de QoS restrita ao AP 802.11e, o que pode ser importante quando se pretende suportar estações móveis de baixo custo. Os algoritmos apresentados podem ser estendidos para o caso de mais de duas categorias de acesso. As condições que disparam a redução e o aumento de CWMin e a redução e o aumento de carga também podem variar, dependendo do objetivo particular de QoS que se quer garantir com o controle.

Trabalhos futuros incluem experimentos com uma maior dinâmica de chegada e partida de fontes, empregando tempos de simulação mais longos, bem como outros modelos de fontes de tempo real e de melhor esforço. Também se pretende estender o mecanismo para mais de duas categorias, estudar as interações do suporte à QoS do 802.11e com as arquiteturas de QoS das redes fixas e desenvolver outros mecanismos que complementem a nova funcionalidade de QoS criada para as redes locais sem fio.

Referências

- [1] Wireless LAN medium access control (MAC) and physical layer (PHY) specifications. IEEE Standard 802.11, 1999.
- [2] MANGOLD, S., CHOI, S., MAY, P., KLEIN, O., HIERTZ, G., E STIBOR, L. IEEE 802.11e wireless LAN for quality of service. *European Wireless (EW'2002)* (Florence, Italy, fevereiro de 2001).
- [3] AHN, G.-S., CAMPBELL, A. T., VERES, A., E SUN, L.-H. Service differentiation in stateless wireless ad hoc networks. *IEEE Joint Conference of Computer and Communications Societies (INFOCOM'2002)* (New York, USA, junho de 2002).
- [4] ROMDHANI, L., NI, Q., E TURLETTI, T. Enhanced service differentiation for IEEE 802.11 wireless ad-hoc networks. *Wireless Communications and Network Conference 2003 (WCNC'03)* (New Orleans, USA, março de 2003) .
- [5] RUBINSTEIN, M. G., E REZENDE, J. F. Qualidade de serviço em redes 802.11. *Simpósio Brasileiro de Redes de Computadores (SBRC'2002)* (Rio de Janeiro, Brazil, maio de 2002).
- [6] AAD, I., E CASTELLUCCIA, C. Differentiation mechanisms for IEEE 802.11. *IEEE Joint Conference of Computer and Communications Societies (INFOCOM'2001)* (Anchorage, Alasca, abril de 2001).
- [7] Medium access control (MAC) enhancements for quality of service (QoS). IEEE Draft for Standard 802.11e/D3.2, março de 2002.
- [8] CHESSON, G., DIEPSTRATEN, W., KITCHIN, D., TEUNISSEN, H., E WENTINK, M. Baseline D-QoS proposal - IEEE document 802.11-00/399. Relatório técnico, 802.11 TGe, 2000.
- [9] CHESSON, G., DIEPSTRATEN, W., HOEBEN, M., SINGLA, A., TEUNISSEN, H., E WENTINK, M. VDCF presentation - IEEE document 802.11-01/132. Relatório técnico, IEEE 802.11 TGe, março de 2001.
- [10] MELO FILHO, J. C., REZENDE, J. F., PIRMEZ, L., E CARMO, L. F.R. C. Controle de admissão e controle de carga para redes IEEE 802.11 infra-estruturadas. *Workshop de Comunicação Sem Fio e Computação Móvel (WCSF'2002)* (São Paulo, Brasil, outubro de 2002).