

Algoritmos para controle de admissão de chamadas e escalonamento em redes IEEE 802.16

Eduardo Castilho Rosa
Faculdade de Engenharia Elétrica, UFU
Uberlândia, Brasil
castilhorosa@gmail.com

Paulo Roberto Guardieiro
Faculdade de Engenharia Elétrica, UFU
Uberlândia, Brasil
prguardieiro@ufu.br

Resumo – Neste trabalho apresenta-se uma nova proposta de escalonamento para o tráfego *uplink* juntamente com um mecanismo de CAC baseado na predição do atraso na rede. O escalonamento é feito para priorizar as estações cujas filas são maiores e o algoritmo de CAC decide a entrada de uma estação na rede através da predição do atraso médio geral da rede. A predição do atraso é calculada com base no estado de ocupação dos *buffers* das estações-assinante, que são enviados à estação-base periodicamente, através do mecanismo de requisição de banda. Foram realizados estudos baseados em modelagem e simulação para analisar a proposta, em um ambiente dinâmico, e os resultados obtidos foram suficientes para a provisão de QoS às aplicações.

Palavras-Chave: CAC, Escalonamento, Qualidade de Serviço, IEEE 802.16, WiMAX.

Abstract – In this paper we present a new scheduling algorithm for uplink traffic and an admission control scheme based on delay prediction. The algorithm scheduling prioritizes the stations which queues are larger and the admission control scheme determines whether a station enters in the network by predicting the overall network average delay. The prediction delay is calculated based on the queues current size in the subscriber stations, which are sent to the base station periodically, through a bandwidth request mechanism. We analyze the proposed scheme by means of modeling and simulation and the results were sufficient for QoS provision to applications.

Keywords: CAC, Scheduling, Quality-of-Service, IEEE 802.16, WiMAX.

I. INTRODUÇÃO

O padrão IEEE 802.16 [1], conhecido também pelo acrônimo WiMAX (*Worldwide Interoperability for Microwave Access*), define uma infraestrutura física que permite o acesso em banda larga sem fio para usuários fixos e móveis. Essa infraestrutura é composta por um conjunto de estações-assinantes (*SSs* – *Subscriber Stations*), seu provedor de serviços de Internet (*ISP* – *Internet Service Provider*) e uma estação-base (*BS* – *Base Station*). A arquitetura definida por esses três elementos é um complemento para as redes de acesso cabeadas, tais como, as baseadas em “*cable modem*” e xDSL, no que se refere à última milha, e vem sendo uma opção de acesso bastante promissora em razão dos baixos custos de implantação e as altas taxas de transmissão que se consegue.

A principal característica das redes WiMAX, e o que a faz uma forte candidata para implementar os sistemas 4G, é o

tratamento diferenciado do tráfego gerado pelas aplicações, fundamental para a provisão de qualidade de serviço (*QoS* – *Quality of Service*). Esse tratamento é garantido através de mecanismos de escalonamento de pacotes nas *SSs* e *BSs*, políticas de policiamento de tráfego e controle de admissão de chamadas (*CAC* – *Call Admission Control*). Para possibilitar maior competitividade entre os fabricantes de equipamentos de redes, esses mecanismos não são definidos pelo padrão. Sendo assim, o padrão possibilita a proposta de soluções que atendam os requisitos de QoS das aplicações e ainda faz com que os equipamentos apresentem características diversificadas, o que permite ao usuário a escolha com base no desempenho requerido.

Tendo em vista que o padrão não especifica as políticas que devem ser usadas na implementação dos mecanismos de escalonamento e CAC, este trabalho propõe um mecanismo para a provisão de QoS, que consiste em um algoritmo de escalonamento de pacotes para o tráfego *uplink*, com um esquema simples de reserva de banda e um algoritmo de CAC dinâmico. O escalonamento é realizado de forma a priorizar as estações que possuem a maior quantidade de pacotes em suas respectivas filas. Já o algoritmo de CAC proposto visa garantir que a entrada de uma nova conexão na rede não prejudique a QoS das conexões já existentes. Esse algoritmo baseia-se na predição do atraso médio da rede. Para isso, foi proposto um mecanismo de predição de atraso que utiliza como parâmetro o comprimento atual das filas nas *SSs*. Foram realizadas simulações para avaliar os cenários considerados utilizando diferentes classes de aplicações.

O restante desse trabalho está organizado da seguinte forma: A Seção II apresenta a estrutura geral do padrão IEEE 802.16. A Seção III identifica os trabalhos relacionados, seguido pela Seção IV que descreve o mecanismo proposto. A Seção V define o cenário de rede e os parâmetros utilizados na simulação e a Seção VI traz a análise dos resultados obtidos. Finalmente a Seção VII traz as conclusões desse trabalho.

II. PADRÃO IEEE 802.16

O padrão IEEE 802.16 define uma camada física (*PHY*) e uma camada de controle de acesso ao meio (*MAC*), que juntas criam, para uma determinada área de cobertura, um enlace lógico sem fio compartilhado, que possibilita a intercomunicação entre *SSs* e *BS*. O padrão define duas arquiteturas para a comunicação: Modo PMP (*Point-to-multipoint*) e modo *Mesh* (Fig. 1). No modo PMP, toda a comunicação entre as *SSs* se faz por meio da *BS* (na forma *SS-BS-SS*). Essa arquitetura facilita o projeto da rede, centralizando o gerenciamento da comunicação na *BS*. No

modo *Mesh*, as SSs podem trocar informações diretamente entre si sem intervenção da BS. No entanto, a complexidade das SSs nesse modo de operação é maior, em razão de cada SS possuir um módulo de controle adicional responsável pelo gerenciamento da comunicação ponto-a-ponto (SS para SS).

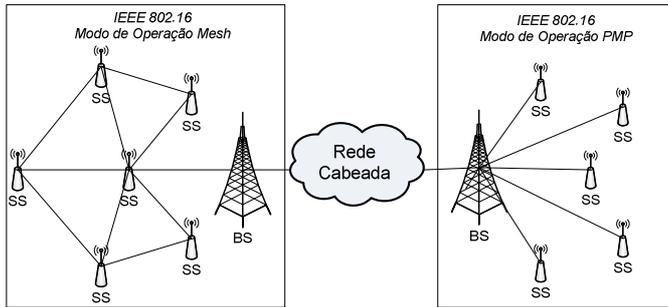


Fig. 1. Modo de operação numa rede IEEE 802.16.

A camada MAC é responsável pelo suporte a diferentes classes de tráfego e, conseqüentemente, toda a arquitetura de provisão de QoS. Também é função da camada MAC prover conectividade com outros tipos de redes, como por exemplo, IP, ATM, etc. além de dar suporte à mobilidade.

A camada física, por sua vez, trata da transmissão de bits sobre o canal sem fio. Ela opera na faixa de frequência de 10 a 66 GHz, com linha de visada e entre 2 e 11 GHz sem linha de visada, com taxas de transmissão que variam de 32 a 130 Mbps dependendo da largura de banda do canal e da técnica de modulação empregada [2].

O controle sobre o enlace sem fio criado, de acordo com o padrão, é realizado na camada física através dos modos de duplexação TDD (*Time Division Duplexing*) e FDD (*Frequency Division Duplexing*), uma vez que o meio é compartilhado. No modo TDD, como pode ser visto na Fig. 2, um quadro de transmissão, ou *frame*, é dividido, no domínio do tempo, em subquadro de *downlink* e subquadro de *uplink*.

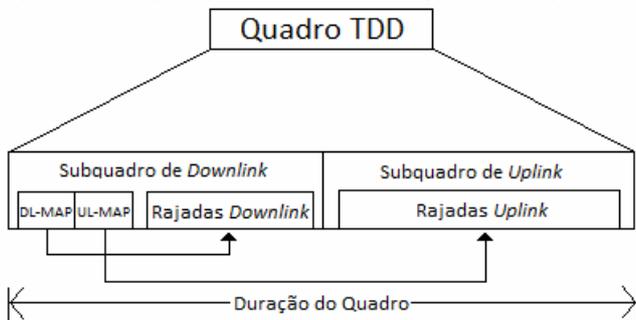


Fig. 2. Estrutura geral do quadro TDD numa rede IEEE 802.16. [3]

No intervalo de tempo dedicado ao *downlink*, a BS transmite em *broadcast* para as SSs utilizando todo o espectro de frequência disponível no meio. Os dados transmitidos nessa direção são distribuídos a cada SS através de um mapa contendo o instante de tempo em que cada SS deve receber os dados. Esse mapa é chamado de DL-MAP. O padrão define ainda um mapa de *uplink*, o UL-MAP que indicará quais SSs transmitirão no subquadro de *uplink* subsequente. No modo de operação FDD, o espectro de frequência é dividido em duas

partes, destinadas a *downlink* e *uplink*, só que nesse caso a transmissão pode ser realizada simultaneamente.

A. Provisão de QoS

Partindo da necessidade de oferecer suporte a uma ampla variedade de aplicações, o padrão IEEE 802.16 foi desenvolvido para prover QoS às aplicações. Para tal, o padrão IEEE 802.16 é orientado à conexão, ao nível da camada MAC. Cada conexão é identificada por um CID (*Connection Identifier*), que por sua vez é associado a um fluxo de serviço (conjunto de parâmetros de QoS, como, por exemplo, atraso máximo fim-a-fim, variação do atraso e vazão mínima). O estabelecimento da conexão se dá através do envio de mensagens de controle DSA-REQ, DSA-RSP e DSA-ACK, conforme ilustrado na Fig. 3, e não é possível transmitir e receber dados sem que haja uma conexão para essa finalidade.

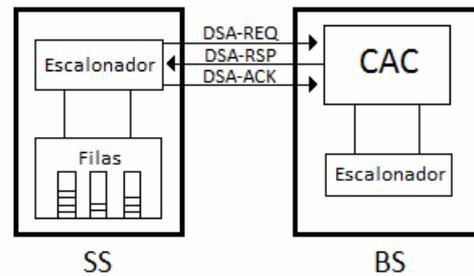


Fig. 3. Estabelecimento de uma conexão no padrão IEEE 802.16.

O principal elemento que atua no processo de estabelecimento da conexão é o CAC. Ele receberá como entrada os parâmetros de QoS dentro da mensagem DSA-REQ e decidirá, em conjunto com o escalonador na BS, se aceita ou não a conexão, com base nos recursos já alocados às conexões correntes e a largura de banda do canal. Essa decisão é importante para que a admissão de uma nova conexão não prejudique a QoS das conexões já aceitas na rede.

O padrão IEEE 802.16, conforme já mencionado, trata de maneira diferenciada os dados do usuário, com o intuito de prover QoS às suas aplicações. Essa diferenciação é realizada através de quatro classes de serviço: UGS (*Unsolicited Grant Service*), rtPS (*Real-time Polling Service*), nrtPS (*Non Real-time Polling Service*) e BE (*Best Effort*). O escalonamento de pacotes, que determina a ordem de transmissão dos mesmos, atuará sobre essas classes de acordo com níveis variados de prioridade: UGS > rtPS > nrtPS > BE. Essas classes são especificadas como segue:

- **UGS** – Essa classe é destinada a aplicações de tempo real que requerem largura de banda fixa. O escalonador na BS aloca *grants* (concessões) periódicos para as SSs a fim de atender esse requisito de taxa constante.
Ex. VoIP sem supressão de silêncio. (*CBR - Constant Bit Rate*).
- **rtPS** – Essa classe é destinada a aplicações de tempo real que requerem largura de banda variável. Para atender esse requisito, o escalonador na BS permite que, periodicamente, as SSs individualmente solicitem

banda para transmitirem, através de um mecanismo conhecido como *Polling Unicast*.

Ex. Vídeo *Streaming* com codificação MPEG. (*VBR – Variable Bit Rate*).

- **nrtPS** – Essa classe é destinada a aplicações de não tempo-real mas que requerem largura de banda mínima garantida. Da mesma forma que na classe rtPS, as SSs aqui solicitam banda para transmitir os dados, porém em intervalos de *polling unicast* menor. Além de solicitar banda em cada intervalo pré-definido de *polling unicast*, nessa classe, a BS permite também que as SSs solicitem banda por *piggyback* e/ou mecanismos de contenção ou disputa.

Ex. Tráfego FTP – *File Transfer Protocol*.

- **BE** – Essa classe é destinada a aplicações que não possuem requisitos de atraso nem requerem largura de banda garantida. Nessa classe é realizada solicitação de largura de banda somente por *piggyback* e/ou mecanismos de contenção.

Ex. Tráfego Web (HTTP).

III. TRABALHOS RELACIONADOS

Em [4], é apresentado um esquema de CAC que utiliza as informações de largura de banda e atraso. O controle da largura de banda é realizado de acordo com o critério de alocação fixa, reservando-se a taxa mínima de cada classe. Já o controle sobre o atraso máximo é realizado fazendo-se uma predição numérica do mesmo, em relação a certa carga na rede, e uma posterior comparação entre essa predição e o requisito de atraso máximo referente a requisição de conexão. Esse trabalho apresenta uma boa contribuição e serve como referência primária.

Um estudo detalhado sobre algoritmos de escalonamento *uplink* é apresentado em [5]. Nesse estudo, são apresentadas várias disciplinas de escalonamento, isoladas e/ou híbridas, bem como uma análise comparativa minuciosa. Em [6], também é proposto um algoritmo de escalonamento *uplink*, mas em conjunto com um esquema de CAC. Nesse trabalho, os autores mostram a relação entre as características de tráfego com seus respectivos parâmetros de QoS. Os resultados mostram que a solução proposta é capaz de prover QoS, em relação a largura de banda e atraso, para as aplicações.

Em [7], uma proposta de CAC utilizando *Token Bucket* é apresentada. Nesse trabalho, um modelo de estimação de atraso é fornecido utilizando os parâmetros taxa de *tokens* (r_i) e tamanho do *bucket* (b). O método descrito nesse trabalho também utiliza um modelo de estimação de atraso, porém baseado no comprimento das filas (mecanismo de requisição de largura de banda).

Em [8][9], é proposto um esquema de CAC baseado na reserva fixa de recursos. O aceite ou rejeição da conexão é feito com base em *thresholds* para cada uma das classes. Entretanto, o processo de admissão considera apenas o requisito de largura de banda e não realiza nenhum tipo de predição quanto ao atraso. O método descrito nesse trabalho, no entanto, leva em conta o requisito de latência máxima para aplicações de tempo-real.

IV. PROPOSTA DE ALGORITMOS DE ESCALONAMENTO E CONTROLE DE ADMISSÃO DE CHAMADAS

A provisão de QoS no padrão IEEE 802.16 é garantida através de mecanismo de escalonamento de pacotes, policiamento de tráfego tanto nas SSs como na BS e controle de admissão de chamadas (CAC) [10]. Esse trabalho focaliza o escalonamento *uplink* e o CAC na BS. Além disso, propõe-se também um mecanismo simples de reserva de largura de banda para separar os tráfegos de tempo-real e não tempo-real. Esses mecanismos são apresentados nas Seções seguintes.

A. Esquema de Reserva de Largura de Banda

No modelo proposto foi definido um esquema simples de reserva de largura de banda, necessário para evitar *bandwidth starvation* (inanição do tráfego), quando esta disputa com conexões de tempo-real, e também evitar que a rede fique completamente saturada, a ponto de inviabilizar o envio de mensagens de controle, como por exemplo, a mensagem DSA-REQ. A Fig. 4 ilustra o esquema de reserva proposto.



Fig. 4. Esquema de reserva de largura de banda proposto.

Esse mecanismo de reserva aloca uma parcela fixa, relativamente grande, de largura de banda para as conexões que possuem requisitos de tempo-real, como UGS e rtPS (W bps). Essa reserva é alocada sob demanda para as conexões que chegam ao longo do tempo.

Uma segunda parcela é reservada exclusivamente para as conexões de não tempo-real nrtPS e BE (T bps). A reserva T é utilizada como segue: se existem conexões BE disputando com conexões de tempo-real mas sem conexões nrtPS na rede, T é dividida entre a quantidade de estações BE na rede e, independentemente do número de estações UGS ou rtPS na rede, T ficará reservada a todo instante às conexões BE. Isso evitará o *bandwidth starvation* dessa classe, pois toda estação BE receberá, em qualquer instante, uma fatia fixa T/N_{be} (N_{be} = número de sessões BE's na rede). Por outro lado, se existem estações nrtPS e BE disputando com conexões de tempo-real, toda a largura de banda reservada T será destinada às conexões nrtPS, pois estas possuem maior prioridade, e assim as estações BE sofrerão *bandwidth starvation*.

Por fim, foi definida ainda uma parcela pequena de largura de banda de R bps que é utilizada exclusivamente para sinalização *uplink*. Essa reserva é necessária porque, como o mecanismo de CAC proposto é dinâmico, ou seja, as conexões podem entrar na rede em qualquer instante de tempo, quando a rede estiver congestionada e uma SS qualquer quiser criar uma conexão de dados ela não poderá enviar a mensagem DSA-REQ, pela falta de recursos. Garantindo-se essa reserva, independentemente do nível de saturação *uplink* do canal, cada SS poderá enviar suas mensagens DSA-REQ para um eventual estabelecimento de conexão.

B. Escalonamento

O mecanismo de escalonamento proposto visa atender os requisitos de largura de banda mínima bem como de atraso máximo limitado para as estações que tem aplicações de tempo-real, no sentido de prover QoS. Para tal, o mecanismo foi projetado para atender as classes de serviço do padrão IEEE 802.16 de acordo com seus níveis de prioridade, sendo $UGS > rtPS > nrtPS > BE$.

Deste modo, em cada ciclo de escalonamento, as estações que possuem conexões UGS são atendidas em primeiro lugar, em razão de requererem largura de banda constante. Em seguida, são atendidas as estações com conexões rtPS, priorizando aquelas que possuem filas maiores e depois as estações nrtPS. Finalmente, as estações BE são atendidas. Com o mecanismo de reserva, as conexões BE não sofrerão *bandwidth starvation* se não houver conexões nrtPS na rede.

Para o tráfego rtPS, o escalonamento proposto ordena as estações em relação ao tamanho de suas filas e atende integralmente a requisição de largura de banda solicitada. A predição do atraso médio da rede é realizado para as estações que não puderam ser atendidas no quadro corrente, em relação à quantidade de bytes presentes em suas filas. No mecanismo proposto, cada requisição rtPS que chegar à BS (R_i) será armazenada, juntamente com sua respectiva conexão, em uma fila Q . Tanto o algoritmo de escalonamento quanto o algoritmo de CAC preditivo acessarão o conteúdo dessa fila Q para determinar a ordem de transmissão e a decisão de aceite da conexão, respectivamente.

Algoritmo de Escalonamento rtPS Proposto

```

1. início
2.    $N$  = Número de elementos de  $Q$  (número de conexões rtPS).
3.    $Nb$  = Número de símbolos OFDM no subframe de uplink.
4.   para ( $j=1$  até  $N-1$ ) faça
5.     para ( $k=j+1$  até  $N$ ) faça
6.       se ( $Q_k.req > Q_j.req$ ) então
7.         Temp =  $Q_j.conexao$ ;
8.          $Q_j.conexao = Q_k.conexao$ ;
9.          $Q_k.conexao = Temp$ ;
10.    fim se
11.  fim para
12. fim para
13.  $P$  = Símbolos para o pré-âmbulo.
14. para ( $i=1$  até  $N$ ) faça
15.    $BW_i$  = Símbolos necessários para transmitir  $Q_i.req$  bytes.
16.   se ( $BW_i + P > (Nb - T - R)$ ) então
17.     se ( $Nb - T - R < P$ ) então
18.       Sai do algoritmo, pois os símbolos não são suficientes.
19.     fim se
20.      $BW_i = Nb - T - R$ 
21.   fim se
22.   UL-MAP = Anexa  $BW_i$  referente a conexão  $Q_i.conexao$ 
23.    $Nb = Nb - BW_i$ 
24. fim para
25. retorna (UL-MAP)
26. fim

```

C. CAC preditivo

No algoritmo de CAC proposto, a BS ao receber a mensagem DSA-REQ proveniente de uma SS arbitrária, solicitando uma nova chamada, realiza uma predição do atraso médio que essa chamada poderá sofrer no sistema, se for admitida. Se esse valor for menor ou igual a um limite pré-estabelecido para cada

classe, a chamada é aceita. Caso contrário ela é rejeitada. Uma vez que o atraso médio calculado esteja dentro de um limite pré-

Algoritmo de CAC Preditivo Proposto

```

1. início
2.    $A_i$  = Busca o atraso predito na rede atual.
3.   se ( $A_i \leq threshold(rtPS)$ ) então
4.     Aceita conexão
5.   senão
6.     Rejeita Conexão
7.   fim se
8. fim

```

fixado, a largura de banda é implicitamente garantida para a conexão.

O algoritmo proposto, ao admitir uma nova chamada, não define um valor de alocação máxima para cada conexão rtPS. Ao invés disso, o esquema proposto aloca a quantidade de símbolos OFDM (símbolo OFDM é a menor unidade de alocação de largura de banda existente no quadro TDD) necessária para transmitir a requisição inteira, de acordo com o esquema de modulação e codificação (*MCS – Modulation Coding Scheme*) empregado. Sendo assim, o cálculo da largura de banda alocada para todas as conexões na rede varia estocasticamente, uma vez que depende dos valores das requisições presente na fila Q .

Considerando que haja, em um determinado instante de tempo t , no máximo K conexões que podem ser escalonadas no quadro corrente e que existem N conexões na rede, a alocação média para cada estação rtPS (b_i) é calculada de acordo com (1)(2)(3):

$$b_i = \frac{\sum_{i=1}^K R_i}{K} \quad (1)$$

sujeito a:

$$K * b_i \leq B_u - T - R, e \quad (2)$$

$$K \leq N \quad (3)$$

O mecanismo de predição proposto é realizado de acordo com a arquitetura mostrada na Fig. 5. Para cada SS que possua conexão rtPS ou nrtPS, as requisições de largura de banda enviadas periodicamente (R_i), cada uma refletindo a ocupação atual do *buffer* nas SSs, são armazenadas em duas filas na BS: uma para rtPS e outra para nrtPS. O módulo de predição aces-

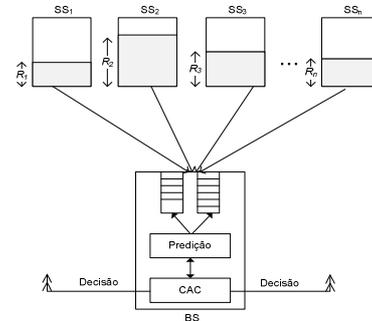


Fig. 5. Arquitetura do mecanismo de CAC proposto.

sará o conteúdo das filas de requisições rtPS e nrtPS e, periodicamente, realizará a predição, levando-se em conta os valores correntes de R_i presente nas filas. Assim que chegar uma chamada na BS, proveniente de uma SS qualquer, o valor atual da predição será utilizado no algoritmo de CAC proposto.

Em (4) descreve-se como é modelado a dinâmica do tamanho médio do *buffer* das estações (B_t) que não conseguiram transmitir no quadro corrente, considerando o instante de tempo t , com N conexões e no máximo K estações podendo ser escalonadas no mesmo quadro:

$$B_t = \frac{\sum_{i=1}^{N-K} R_i}{N-K} \quad (4)$$

Quando a rede está levemente congestionada, ou seja, quando $N = K$, intuitivamente percebe-se que B_t deve ser zero, o que consequentemente sugere um atraso médio na rede baixo. Por outro lado, quando N aumenta com a entrada de novas chamadas, $N - K$ tende a um número grande e K tende a um.

De acordo com a estimativa do tamanho médio do *buffer* B_t , a predição do atraso na rede pode então ser calculada, segundo (5):

$$A_t = \frac{B_t}{r} * T_s * f \quad (5)$$

onde r é a eficiência da MCS empregada, em bits por símbolo, T_s o tempo útil do símbolo em *ms* e f é a duração do quadro TDD, também em *ms*.

Através dessa equação, percebe-se que o atraso predito é diretamente proporcional à B_t , uma vez que r , T_s e f são constantes ao longo do tempo.

V. MODELAGEM E SIMULAÇÃO

Os algoritmos de escalonamento e CAC propostos foram avaliados por meio de experimentos de simulação com a ferramenta de simulação NS-2 [11]. Essa ferramenta inclui algoritmos de escalonamento para as classes UGS, rtPS (baseado em *deadline* dos pacotes) e nrtPS, entretanto ela não possui nenhum esquema de CAC implementado. Para implementar o modelo proposto, foi necessário também incorporar ao NS-2 um módulo de simulação para redes IEEE 802.16.

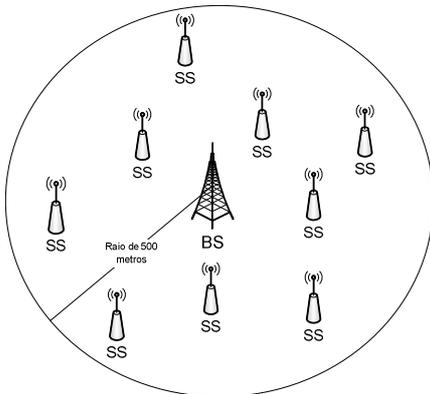


Fig. 6 – Cenário de rede considerado na simulação.

Os cenários considerados envolvem uma BS e várias SSs entrando na rede sucessivamente, em intervalos regulares. A distância máxima permitida de uma SS até a BS é de 500 metros, o que possibilita a utilização de um esquema de modulação e codificação mais eficiente [12]. A Fig. 6 ilustra o cenário considerado na simulação.

O modelo de simulação considera uma conexão por estação e o modo GPSS (Grant per Subscriber Stations) é usado na concessão de largura de banda. Os principais parâmetros de simulação são apresentados na Tabela I. Esses parâmetros foram escolhidos em razão deles serem utilizados na maioria dos trabalhos relacionados.

TABELA I. PRINCIPAIS PARÂMETROS DA SIMULAÇÃO.

Parâmetro	Valor
Frequência de Operação	3,5 GHz
Largura de Banda	5 MHz
Duplexação	TDD
Antena	Omnidirecional
Modelo de Propagação	Two Ray Ground
Duração do Quadro	20 ms
Cyclic Prefix (CP)	0.25
Modulação Empregada	OFDM 64QAM 3/4
Taxa de Transmissão Uplink	7,70 Mbps
Tráfego UGS	CBR (Tamanho do pacote=40 bytes; Intervalo=0.02s). Taxa = 16 Kbps
Tráfego rtPS	Streaming de Vídeo MPEG (Tamanho do pacote=[200:1000]; Intervalo=0.01s). Taxa média = 480 Kbps
Tráfego nrtPS	FTP (Taxa mínima = 160 Kbps; Taxa máxima = 800 Kbps)
Tráfego BE	Tráfego Web (Taxa média = 75 Kbps)
Latência máxima	20 ms para UGS; 200 ms para rtPS.
Razão W:T:R (Reserva de largura de banda)	(W=0.85; T=0.1; R=0.05)
Intervalo de Interchegadas	5 segundos.

VI. APRESENTAÇÃO E ANÁLISE DOS RESULTADOS

Os resultados obtidos nessa Seção referem-se a uma rodada de simulação para cada cenário. A vazão e o atraso são calculados em intervalos periódicos ($t, t+k$) de acordo com as equações (6) e (7):

$$vazao_{t,t+k}^z = \frac{\sum_{i=1}^{Tot} size_pac_{i,z}^{t,t+k}}{Tot} \quad (6)$$

$$atraso_{t,t+k}^z = \frac{\sum_{i=1}^N \left(\sum_{j=1}^{P_{cont,i}} (Rx_{j,z}^{t,t+k} - Tx_{j,z}^{t,t+k}) / P_{cont,i} \right)}{N} \quad (7)$$

onde z = classe de serviço; Tot = total de pacotes recebidos no intervalo $[t, t+k]$; $size_pac_{i,z}^{t,t+k}$ = tamanho do i -ésimo pacote recebido no intervalo $[t, t+k]$; $P_{cont,i}$ = quantidade de pacotes recebidos pela estação i ; $Rx_{j,z}^{t,t+k}$ and $Tx_{j,z}^{t,t+k}$ = tempo de

recepção e transmissão do j -ésimo pacote, respectivamente.; k = intervalo de amostra (2 ou 5 segundos); N = quantidade total de estações na rede no intervalo $[t, t+k]$.

O primeiro cenário foi modelado para analisar o desempenho do algoritmo de escalonamento proposto, para a classe de serviço rtPS. Nesse cenário, as SSs com conexões rtPS registram-se na rede dinamicamente, uma por vez, em intervalos fixos de 5 segundos, iniciando no instante 15 segundos. As SSs começam a transmitir dados assim que o processo de entrada na rede finaliza e continuam a transmissão até o fim da simulação, no instante de tempo 130 segundos. A intenção de fazer as SSs permanecerem na rede até o fim da simulação é que, desta forma, a adição periódica de conexões, de alguma forma, irá interferir nas conexões já existentes. Considerou-se como métrica de avaliação para esse cenário o atraso médio, que é calculado também em intervalos fixos de 5 segundos.

Conforme pode ser observado pelo gráfico apresentado na Fig. 7, o algoritmo proposto apresenta um desempenho superior ao escalonamento RR (Round Robin), quando a rede está saturada, isto é, quando a utilização da largura de banda reservada W está próxima do valor pré-fixado (85%). Essa situação acontece a partir do instante de tempo 80 segundos. Isso ocorre devido ao fato de que, nessa situação, as conexões rtPS já não podem mais ser totalmente escalonadas no mesmo quadro TDD, o que provoca um aumento nas filas das SSs que não puderam ter suas conexões escalonadas naquele instante, pois no mínimo serão escalonadas no quadro seguinte. A justificativa para o desempenho superior do algoritmo proposto é que ele prioriza as SSs cujas filas são maiores, o que não acontece na disciplina RR, e o principal fator que contribui para o atraso é exatamente o acúmulo de pacotes nos buffers das SSs. Quando a rede está levemente congestionada (do instante 15 seg. até 75 seg.), ou seja, quando todas as conexões podem ser escalonadas no quadro TDD corrente, não ocorre acúmulo de pacotes nas filas das SSs, e, conseqüentemente, o desempenho dos dois algoritmos são semelhantes.

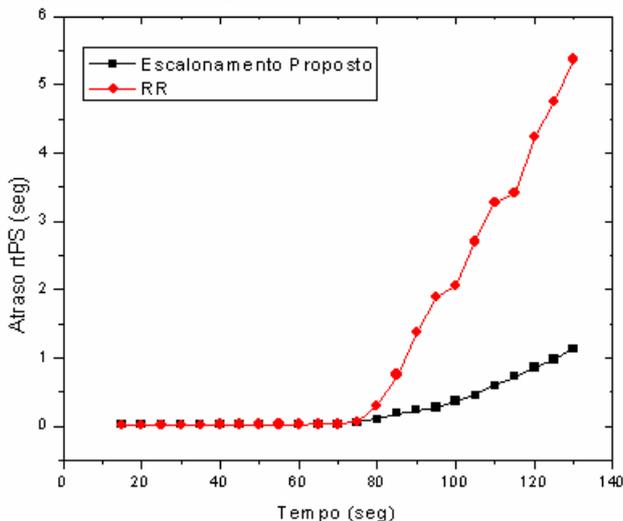


Fig. 7. Desempenho do escalonador proposto, com amostras de 5 segundos.

O segundo cenário procura avaliar o desempenho do escalonador em relação às classes de tempo-real UGS e rtPS.

Foram adicionadas 8 estações UGS na rede no instante 15 segundos e esse número permaneceu constante ao longo do tempo. Periodicamente, em intervalos de 5 segundos, foram adicionadas SSs com conexões rtPS a fim de verificar a influência dessas estações no atraso UGS. O número máximo de estações rtPS na rede foi fixado em 20. Conforme pode ser observado no gráfico da Fig. 8, a inserção de conexões rtPS na rede não interferiu no atraso das conexões UGS. Isso ocorreu porque, como as estações UGS têm recursos garantidos, somente quando há largura de banda suficiente na rede é que as conexões das outras classes são servidas. Nesse cenário, quando a utilização da largura de banda reservada W aproximou-se de 0,85 (instante 90 segundos) as conexões rtPS não puderam mais ser escalonadas no quadro TDD corrente, o que provocou um aumento das filas, e conseqüentemente o atraso.

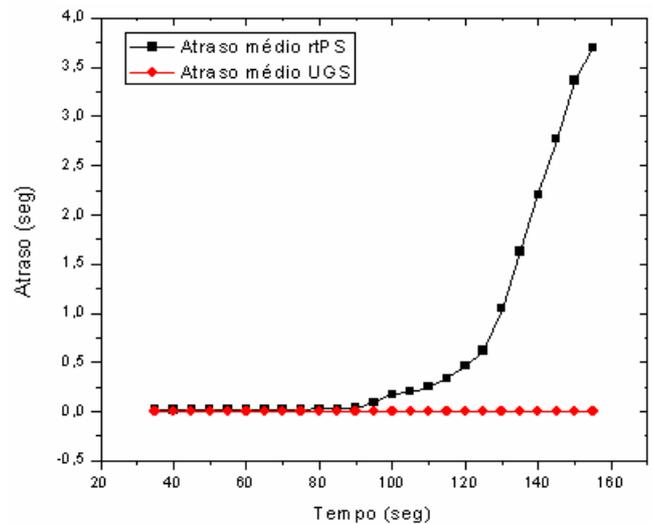


Fig. 8. Desempenho do escalonador para fluxos de tempo-real, com amostras de 5 segundos.

No cenário seguinte, procurou-se verificar o desempenho do mecanismo de CAC preditivo proposto, para a classe de serviço rtPS. Para tal, foram consideradas apenas estações rtPS, entrando na rede em intervalos de 5 segundos, até um limite máximo de 20 estações. O *threshold* do atraso rtPS utilizado pelo algoritmo foi de 200ms. O gráfico apresentado na Fig. 9 mostra que a utilização do esquema de CAC proposto fixou o atraso em 200ms, ou seja, evitou que a entrada de novas chamadas, quando a rede está saturada, prejudicasse a QoS das conexões já existentes.

O próximo cenário foi criado com o intuito de verificar a influência do mecanismo de reserva de largura de banda para as conexões de não tempo-real, nrtPS e BE. Foram adicionadas 8 estações nrtPS e 8 estações BE no início da simulação, no instante 15 segundos, e a cada 5 segundos, a partir do instante 45 segundos, foi acrescentada uma estação rtPS. A taxa média de dados nrtPS foi de 500 Kbps e a taxa média BE foi de 70 Kbps. Fixou-se um limite mínimo para a taxa nrtPS, ou seja, a taxa mínima sustentável, para 200 Kbps. Pelo gráfico apresentado na Fig. 10 observa-se que, quando todas as estações rtPS são escalonadas no mesmo quadro TDD, ou seja, quando o número de estações rtPS é baixo (tempo < 90 segun-

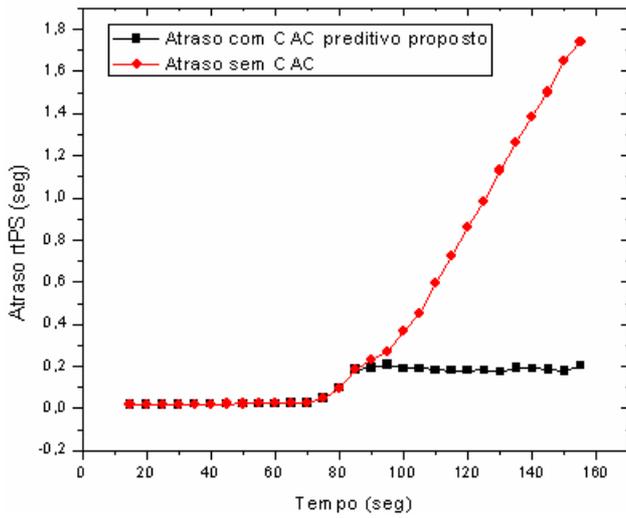


Fig. 9. Desempenho do CAC preditivo para rtPS, com amostras de 5 segundos.

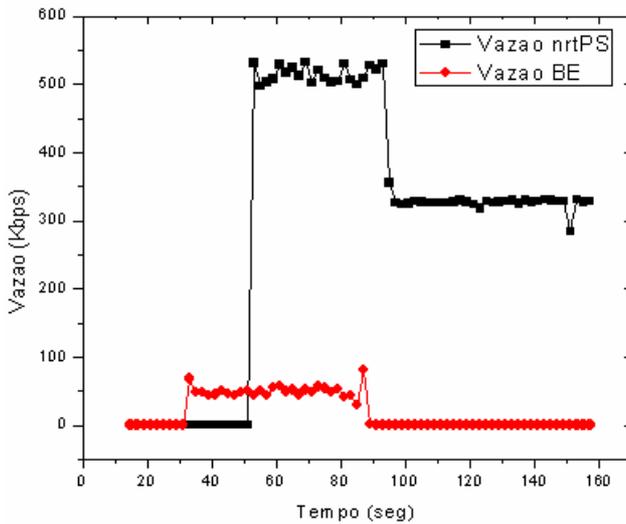


Fig. 10. Desempenho do mecanismo de reserva de banda para fluxos de não tempo-real, com amostras de 2 segundos.

dos), a rede consegue dar vazão para ambas as classes nrtPS e BE de acordo com suas respectivas taxas. No entanto, a partir do momento em que o número de estações rtPS atinge um valor onde as mesmas não podem mais ser totalmente escalonadas no mesmo quadro TDD (tempo > 90 segundos), percebe-se que a vazão das estações nrtPS fixa em aproximadamente 300 Kbps e a vazão BE cai para 0. Isso ocorre devido a dois fatores: O primeiro é que, nessa situação, não foi aplicado o CAC para controlar a entrada das conexões rtPS na rede. Isso provocou a queda nas vazões nrtPS e BE. Já o segundo fator está relacionado ao mecanismo de reserva de largura de banda proposto. Devido essa reserva, a partir do instante de congestionamento (tempo > 90 segundos) as estações BE e nrtPS disputaram a reserva T alocada exclusivamente para elas. Como a classe nrtPS possui prioridade maior, ela ficou com 100% da reserva a partir desse momento, o que provocou *bandwidth starvation* da classe BE.

Finalmente, no cenário seguinte, aplicou-se o CAC sobre as conexões rtPS no sentido de avaliar seu efeito sobre a vazão das conexões de não tempo-real, especificamente a classe nrtPS. Assim como no cenário anterior, aqui considerou-se também 8 estações nrtPS entrando na rede no instante 15 segundos e, sucessivamente, a partir do instante 20 segundos e a cada período de 5 segundos, foram adicionadas estações rtPS na rede. Comparando-se, pelo gráfico da Fig. 11, a vazão das conexões nrtPS percebe-se claramente que, quando se aplica o CAC a vazão permanece inalterada do início ao fim da simulação, o que não ocorre quando não se aplica (a vazão cai para o mínimo). A justificativa para tal comportamento é que o CAC bloqueia a entrada de conexões quando a rede já não pode mais escalonar todas as conexões rtPS no mesmo quadro TDD, através do mecanismo de predição do atraso.

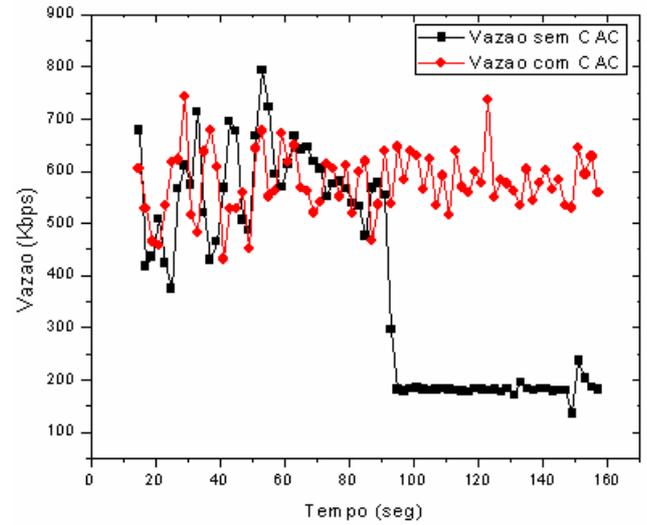


Fig. 11. Desempenho do CAC preditivo para fluxos de não tempo-real, com amostras de 2 segundos.

VII. CONCLUSÕES

Neste trabalho apresentou-se uma proposta para a provisão de QoS no padrão IEEE 802.16, consistindo de um algoritmo de escalonamento, um mecanismo simples de reserva de largura de banda e um CAC preditivo com base no atraso da rede. O algoritmo de escalonamento e o mecanismo de CAC dinâmico utilizaram informações referentes aos tamanhos das filas nas SSs a fim de determinar a ordem de transmissão e a decisão de aceite da conexão, respectivamente. Os resultados apresentados demonstraram que a proposta de escalonamento é satisfatória para o controle do atraso dos fluxos rtPS e o mecanismo de CAC também se mostrou adequado tanto para as aplicações de tempo-real quanto as de não tempo-real. O mecanismo de reserva de largura de banda também foi testado, e se mostrou adequado para a manutenção da taxa mínima nrtPS mesmo quando a rede está congestionada.

Como trabalho futuro pretende-se desenvolver um modelo de reserva dinâmica de largura de banda, com base na carga de tráfego na rede e analisar o desempenho da proposta apresentada considerando-se, além da vazão e atraso, também perdas, taxa de utilização do canal e *jitter*.

REFERÊNCIAS

- [1] IEEE Standard for Local and Metropolitan Area Networks - Part 16: Air Interface for Fixed Broadband Wireless Access Systems. IEEE Std., Rev. IEEE Std 802.16-2004, 2004.
- [2] A. Bacciola, et al. "IEEE 802.16: History, status and future trends". *Computer Communications*, vol. 33, pp. 113-123, 2009.
- [3] M. Maode, et al. "Current technology developments of WiMAX systems". *Springer Publishing Company*, 2009.
- [4] M. Castrucci et al. "Connection admission control in WiMAX networks". *The International Conference on Information and Communication Technologies, ICT'08*, Paris, France, 2008.
- [5] P. Dhrona, et al. "A performance study of uplink scheduling algorithms in point-to-point WiMAX networks". *Computer Networks*, vol. 32, pp. 511-521, 2009.
- [6] K. Wongthavarawat et al. "Packet scheduling for QoS support in IEEE 802.16 broadband wireless access systems". *International Journal of Communications Systems*, vol. 16, pp. 81-96, Feb, 2003.
- [7] T. Tsai et al. "CAC and packet scheduling using token bucket for IEEE 802.16 networks". *Journal of Communications*, vol. 1, pp. 30-37, 2006.
- [8] H. Fen et al. "Performance analysis of a reservation based connection admission scheme in 802.16 networks". *IEEE GLOBECOM*, San Francisco, USA, 2006.
- [9] K. Tsang et al. "Admission control scheme for mobile WiMAX networks". *The International Symposium on Consumer Electronics, ISCE'07*, Dallas, USA, 2007.
- [10] A. Mohamed et al. "Call admission control in wireless networks: a comprehensive survey". *IEEE Communications Surveys*, vol. 7, pp. 50-68, 2005.
- [11] *The Network Simulator - ns-2*. <http://www.isi.edu/nsnam/ns/> - 2002.
- [12] C. Hoymann. "Analysis and performance evaluation of the OFDM-based metropolitan area network IEEE 802.16". *Computer Networks*, vol. 49, pp. 341-363, June, 2005.