

# SISTEMA FUZZY APLICADO À ALOCAÇÃO DE RECURSOS DE RÁDIO EM REDES LTE, UTILIZANDO ESTIMAÇÃO DE BANDA EFETIVA E MODELAGEM MULTIFRACTAL $\beta$ MWM DE FLUXOS DE TRÁFEGO

Diego Cruz Abrahão  
Flávio Henrique Teles Vieira

Universidade Federal de Goiás  
Escola de Engenharia Elétrica, Mecânica e de Computação  
e-mails: diegoabrahao@outlook.com e flavio@eee.ufg.br  
Av. Universitária, nº 1488, qd. 86, bloco A, Setor Leste Universitário, Goiânia-Goiás

**Resumo-** Neste trabalho, é proposto um esquema de alocação de blocos de recurso para sistemas de comunicação LTE (*Long Term Evolution*) em que são calculadas prioridades para os usuários, através de lógica *fuzzy* e estimação de banda efetiva para os fluxos de tráfego. A banda efetiva do fluxo de tráfego de cada usuário é usada de forma a atender a parâmetros de Qualidade de Serviço (QoS), sendo estimada em tempo real, através dos parâmetros da modelagem multifractal  $\beta$ MWM adaptativa ( *$\beta$ -Multifractal Wavelet Mode*). O esquema de alocação proposto visa o atendimento de parâmetros de QoS dos usuários e das restrições do esquema de modulação e código (MCS – *Modulation and Coding Scheme*) da transmissão de *downlink* LTE. O algoritmo proposto considera a qualidade média do canal e a estimativa da banda efetiva em tempo real, através da modelagem multifractal  $\beta$ MWM adaptativa de fluxo de tráfego de cada usuário, para tomar decisões sobre o escalonamento dos recursos de rádio disponíveis. É verificada a eficiência do esquema proposto, comparando parâmetros, como: vazão total do sistema, taxa de perda de dados, tempo de retardo e critério de justiça, com os de outros algoritmos da literatura.

**Palavra Chave-** LTE, Escalonamento, *Fuzzy*, QoS, Banda Efetiva, Modelagem  $\beta$ MWM.

**Abstract-** In this paper, it is proposed a scheme to allocate resource blocks for LTE (Long Term Evolution) communication systems where priorities are calculated for users, through fuzzy logic and estimation of effective bandwidth for traffic flows. The effective bandwidth of each user's traffic flow is used to meet the Quality of Service (QoS) parameters, estimated in real time through the parameters of the adaptive  $\beta$ MWM multifractal modeling. The proposed allocation scheme aims at attaining the QoS parameters of users and the constraints of modulation and code (MCS - Modulation and Coding Scheme) scheme of the LTE downlink transmission. The proposed algorithm considers the average channel quality and the estimation of effective bandwidth in real time, through the adaptive multifractal  $\beta$ MWM modeling of traffic flows to each user to decide on the scheduling of radio resources available. The efficiency of the proposed scheme is verified by comparing parameters such as throughput of the system, loss of data rate, delay time and criterion of fairness, to those of other algorithms in the literature.

**Keywords-** LTE, Scheduling, Fuzzy, QoS, Bandwidth Effective, Modeling  $\beta$ MWM

## 1 Introdução

O sistema LTE (*Long Term Evolution*) tem como objetivo fornecer altas taxas de dados, baixa latência, acesso via rádio otimizado por pacote e flexibilidade de implementação com diferentes larguras de banda (3GPP TSG RAN TR 25.913 v8.0.0, 2008). O LTE emprega a Multiplexação por Divisão de Frequência Ortogonal (OFDM- *Orthogonal Frequency-Division Multiplexing*) na transmissão de *downlink*, o que permite maior liberdade no escalonamento de canais e no gerenciamento flexível dos recursos de rádio. Um desafio do sistema LTE é suportar um grande número de usuários e satisfazer suas diferentes exigências de taxas, dado um recurso de rádio limitado (Guan, et al., 2011).

A alocação de recursos de rádio para o sistema de comunicação LTE tem sido extensivamente estudada, como exemplo temos os algoritmos apresentados em (Su & Ping Wang, 2012) e (Guan, et al., 2011). Estes algoritmos apresentam soluções para o problema de alocação de recursos de rádio para redes LTE, considerando o ganho médio do canal e uma taxa de dados mínima requerida de cada usuário. O primeiro utiliza o algoritmo PSO (*Particle Swarm Optimization*) e o segundo propõe um novo esquema de alocação de blocos de recurso. Os algoritmos de alocação de recursos de rádio para rede LTE possuem restrições de utilização do esquema de modulação e código (MCS - *Modulation and Coding Scheme*). No LTE, o MCS adotado por um usuário, durante um intervalo de tempo de transmissão (TTI - *Transmission Time Interval*), deve ser o mesmo para todos os blocos de recurso alocados a ele, em uma configuração de antena única.

É proposto neste trabalho, um esquema de alocação de blocos de recurso para o sistema de *downlink* LTE, visando maximizar a vazão total do sistema, garantir parâmetros de QoS dos usuários e diminuir a taxa de perda de dados da rede. Tendo em vista a complexidade matemática para estabelecer um algoritmo que alcance estes objetivos e não infrinja as restrições da rede, é

utilizada a lógica nebulosa, ao invés do raciocínio *crisp*. É empregado um sistema de inferência *fuzzy* para calcular a prioridade dos usuários e tomar decisões de escalonamento, tendo como variáveis de entrada o ganho médio do canal e a estimativa da banda efetiva dos fluxos de tráfego dos usuários. A banda efetiva do fluxo de tráfego de cada usuário é usada de forma a atender a parâmetros de QoS, em especial a uma determinada taxa de perda de bytes, sendo estimada em tempo real, através dos parâmetros da modelagem multifractal  $\beta$ MWM adaptativa (Gonçalves, et al., 2013).

Foram realizadas simulações, comparando os resultados obtidos pelo esquema proposto, com os resultados obtidos pelos algoritmos apresentados em (Su & Ping Wang, 2012) e (Guan, et al., 2011). Diferentemente de (Su & Ping Wang, 2012) e (Guan, et al., 2011), o esquema proposto leva em consideração a estimativa da banda efetiva do fluxo de tráfego de cada usuário em tempo real, ao invés da taxa mínima requerida, além de utilizar lógica nebulosa. A teoria de banda efetiva é utilizada, para calcular os parâmetros de QoS a serem atendidos e modelar o tráfego dos usuários.

O trabalho está organizando da seguinte forma. Na seção 2 é apresentada a estrutura de um *frame* LTE e o modelo do sistema. Na seção 3 é apresentado o algoritmo PSO. Na seção 4 é apresentada a modelagem multifractal  $\beta$ MWM adaptativa e o cálculo para estimação da banda efetiva dos usuários. Na seção 5 é proposto um esquema para alocação de blocos de recurso para o sistema de *downlink* LTE, baseado em lógica *fuzzy*. Na seção 6, é validado o esquema proposto, apresentando os resultados de simulações realizadas. Na seção 7, são apresentados alguns trabalhos relacionados à alocação de recursos de rádio em sistemas LTE. Finalmente, é concluído o trabalho na seção 8.

## 2 Estrutura de um *frame* LTE e o modelo do sistema LTE

Nesta seção, é apresentada a estrutura de um *frame* e a composição dos blocos de recurso e de escalonamento na transmissão de *downlink* LTE. É introduzido também o modelo básico do sistema LTE e o problema de otimização para maximização da vazão total do sistema, tendo em vista suas restrições.

### 2.1 Estrutura de um *frame*

A estrutura de um *frame* de transmissão de *downlink* LTE é mostrada na Figura 1. Cada *frame* de rádio ocupa 10ms, que são divididos em dez *subframes* de 1ms (Dahlman, et al., 2007). Cada *subframe* por sua vez é dividido em dois *slots* de tempo de 0,5ms. Há sete ou seis símbolos OFDM para cada *slot* de tempo, dependendo da utilização de prefixo ciclo normal ou estendido, respectivamente (Dahlman, et al., 2007).

No domínio da frequência, os recursos são agrupados em 12 subportadoras de 15KHz, totalizando uma largura de banda de 180KHz. Um bloco de recurso (RB – *ResourceBlock*) é definido como uma unidade de 12 subportadoras durante um *slot* de tempo (3GPP TSG RAN TR 25.913 v8.0.0, 2008). No sistema LTE, os blocos de recurso são escalonados sempre em pares de RBs, chamados assim de blocos de escalonamento (SB – *SchedulingBlock*), com duração de 1ms.

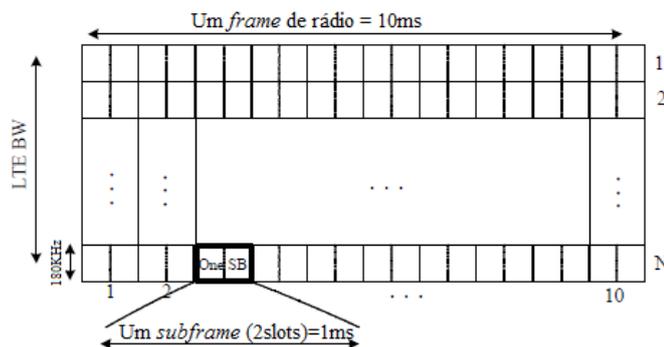


Figura 1: Estrutura de recurso tempo-frequência básico do LTE.

### 2.2 Modelo do Sistema LTE

Considerando a transmissão de *downlink* do sistema LTE com  $N_{SB}$ s, onde cada SB é alocado com a mesma potência. Assumindo que  $K$  usuários são servidos por uma estação base (BS - *Base Station*) e que a taxa mínima exigida pelo  $k$ -ésimo usuário seja  $R_k$  bits/s. Define-se um bloco de escalonamento (SB - *SchedulingBlock*) com  $N_s$  símbolos OFDM consecutivos no domínio do tempo e  $N_{sc}$  subportadoras consecutivas no domínio da frequência. Considerando que existem sinais pilotos e de controle nos blocos de escalonamento, apenas  $N_{sc}^{(d)}(s)$  das  $N_{sc}$  subportadoras podem ser utilizadas para transferência de dados no  $s$ -ésimo símbolo OFDM, onde  $s \in \{1, 2, \dots, N_s\}$  e  $N_{sc}^{(d)}(s) \leq N_{sc}$ . Seja  $R_j^{(c)}$  a taxa de código associada com o MCS  $j \in \{1, 2, \dots, J\}$ , onde  $J$  é o número total de MCS suportados na transmissão,  $M_j$  é o tamanho da constelação do MCS  $j$  e  $T_s$  é a duração do símbolo OFDM. Então, a taxa de bits  $r^{(j)}$  alcançada por um único SB com o MCS  $j$  é dada por:

$$r^{(j)} = \frac{R_j^{(c)} \log_2(M_j)}{T_s N_s} \sum_{s=1}^{N_s} N_{sc}^{(d)}(s) \quad (1)$$

No LTE, o índice CQI (*Channel Quality Indicator*) é definido em termos da taxa de código e do esquema de modulação, que possui a informação de qual MCS deve ser adotado para o usuário  $k$  durante um TTI (*Transmission Time Interval*). Considerando  $g_{k,n}$  como o indicador da qualidade do canal do  $k$ -ésimo usuário no  $n$ -ésimo SB, então o CQI deste usuário para os NSBs pode ser representado como  $g_k = [g_{k,1}, g_{k,2}, \dots, g_{k,N}]^T$  e o CQI para os  $K$  usuários nos NSBs é dado por  $G = [g_1, g_2, \dots, g_k]^T$ . O máximo CQI do usuário  $k$  para todos SBs é dado por:

$$n^* = \operatorname{argmax}(g_{k,n}), \text{ onde } n \in N \quad (2)$$

Assim  $q_{k,max}(g_{k,n^*}) \in \{1, 2, \dots, J\}$  é o maior índice MCS alcançado pelo usuário  $k$  no  $n^*$ -ésimo SB. Considerando que cada SB é atribuído exclusivamente a um único usuário, durante a duração de um TTI. Definimos  $\rho_{k,n}$  como o indicador do recurso atribuído para o usuário  $k$  no  $n$ -ésimo SB. Então, quando  $\rho_{k,n} = 1$ , o  $n$ -ésimo SB é alocado ao usuário  $k$  e  $\rho_{k',n} = 0$ , para  $\forall k' \neq k$ . Seja  $b_{k,j}$  o indicador da escolha do MCS do usuário  $k$  para todos SBs alocados em um *subframe*. Assim,  $b_{k,j} = 1$  é a escolha do MCS  $j$  para o usuário  $k$  e  $b_{k,l} = 0$  para  $\forall l \neq j$  em um *subframe*. A taxa de dados (bits/s) em um *subframe* pode ser dada por:

$$r_k = \sum_{n=1}^N \rho_{k,n} \sum_{j=1}^{q_{k,max}(g_{k,n^*})} b_{k,j} r^{(j)} \quad (3)$$

onde,  $q_{k,max}(g_{k,n^*}) \in \{1, 2, \dots, J\}$  é o maior índice MCS alcançado pelo usuário  $k$  no  $n$ -ésimo SB.

A alocação de recursos de rádio multiusuária ótima tem como objetivo maximizar o seguinte somatório:

$$\max \sum_{k=1}^K \sum_{n=1}^N \rho_{k,n} \sum_{j=1}^{q_{k,max}(g_{k,n^*})} b_{k,j} r^{(j)} \quad (4)$$

Sujeito a:

$$r_k \geq R_k, \quad \forall k \quad (4.1)$$

Se  $\rho_{k,n} = 1$ , então  $\rho_{k',n} = 0$ , para  $\forall k' \neq k$

$$\sum_{j=1}^{q_{k,max}(g_{k,n^*})} b_{k,j} = 1 \quad (4.2)$$

que representa a taxa de dados total alcançada em um TTI sob certas restrições, que neste caso é o atendimento à taxa mínima exigida pelos usuários para satisfazer parâmetros de QoS. Pode-se observar que (4) é um problema de otimização com complexidade exponencialmente crescente com o número de restrições e variáveis. Com objetivo de encontrar uma solução sub-ótima para este problema, atendendo aos requisitos de QoS dos usuários, dado por (4.1), e não infringindo as restrições da rede, dado por (4.2), propomos um novo esquema de alocação de recursos de rádio, baseado em lógica *fuzzy*.

### 3 Algoritmo PSO

Uma forma de encontrar uma solução para o problema de maximização (equação 4) é utilizando o algoritmo PSO (*Particle Swarm Optimization*), como proposto em (Su & Ping Wang, 2012). A otimização PSO é estocástica, sub-ótima, baseada em população e é de fácil implementação. A população é chamada de enxame e cada indivíduo, que corresponde a uma solução para o problema, é chamado de partícula. No algoritmo PSO padrão, cada partícula possui posição, velocidade e memoriza a melhor posição da partícula encontrada até o momento (também chamada de melhor posição local). A melhor posição da partícula na população, ou seja, a solução com menor custo, também é memorizada. Os vetores velocidade e posição são variáveis contínuas. Na inicialização, cada partícula possui posição e velocidade aleatórias. O algoritmo procura a

solução ótima através de atualizações das posições e velocidades de cada partícula, levando em conta as velocidades, as melhores posições das partículas e a melhor posição da população, até um critério de parada. As posições e as velocidades das partículas são atualizadas segundo as equações:

$$v_{t+1} = wv_t + r_1c_1(P_t - X_t) + r_2c_2(G_t - X_t)$$

$$X_{t+1} = X_t + v_{t+1}$$

onde  $w$  é o peso de inércia;  $c_1$  e  $c_2$  são taxas de aprendizagem;  $r_1$  e  $r_2$  são dois números aleatórios gerados segundo uma distribuição uniforme  $[0;1]$ ;  $v_t$ ,  $X_t$  e  $P_t$  são, respectivamente, a velocidade, a posição e a melhor posição da partícula, no instante de tempo  $t$ ; e  $G_t$  é a melhor posição da população neste instante.

No caso de alocação de blocos de recurso no sistema LTE, o vetor solução, que representa o usuário alocado a cada bloco de escalonamento, é inteiro, sendo adequado utilizar a versão modificada da PSO, que discretiza a posição e a velocidade das partículas segundo a equação (Wang & Yin, 2008):

$$INT(r) = \begin{cases} floor(r), & \text{se } rand > r - floor(r) \\ ceil(r), & \text{caso contrário} \end{cases}$$

onde  $floor(r)$  e  $ceil(r)$  são funções de arredondamento para o maior inteiro menor que  $r$  e menor inteiro maior do que  $r$ , respectivamente, e  $rand$  é um número aleatório gerado segundo uma distribuição uniforme  $[0,1]$ .

A PSO padrão não possui restrições: assim, a restrição de taxa mínima dos usuários para garantir QoS é convertida em uma função de penalidade (5). A função de penalidade transforma um problema de otimização com restrição em uma otimização sem restrição (Su & Ping Wang, 2012).

$$Penalidade = \sum_{k=1}^K [\min(0, r_k - R_k)]^2 \quad (5)$$

O algoritmo PSO realiza a otimização avaliando os custos de cada solução (partícula) através da função objetivo (6). Os menores custos são memorizados por partícula da população e utilizados no algoritmo.

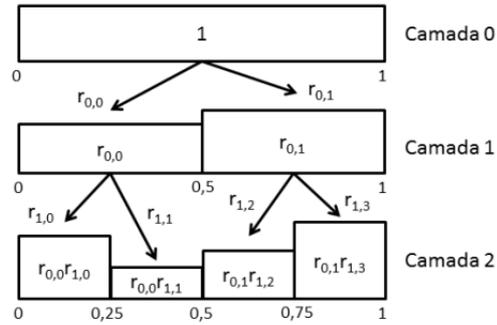
$$F = \sum_{k=1}^K r_k(t) - \sum_{k=1}^K [\min(0, r_k - R_k)]^2 \quad (6)$$

#### 4 Modelagem Multifractal $\beta$ MWM e Banda Efetiva

O MWM (*Multifractal Wavelet Model*) é um modelo multifractal muito utilizado na modelagem de fluxos de tráfego de rede (Riedi, et al., 1999). O MWM utiliza uma cascata multiplicativa no domínio *wavelet* (Chui, 1992), para modelar os fluxos de tráfego. Através dessa modelagem é possível capturar a dependência de longa duração presente nos dados de tráfego, bem como capturar outras características multifractais. Uma das variações do modelo MWM é o  $\beta$ MWM, que utiliza uma distribuição beta simétrica para modelar os coeficientes *wavelet* da cascata multiplicativa.

O processo de modelagem do  $\beta$ MWM realiza a transformada discreta de *wavelet* de Haar, para um número fixo de camadas da cascata multiplicativa binomial (Rocha & Vieira, 2009), considerando a série de tráfego completa em uma única etapa. A partir dos coeficientes *wavelet* e dos coeficientes de escala gerados por camada, os parâmetros MWM são estimados.

A cascata multiplicativa binomial divide um processo estocástico  $X(t)$ , definido no intervalo  $I = [0,1]$ , com massa unitária, em  $2^n$  subintervalos disjuntos diádicos  $I_n^k = [k2^{-n}, (k+1)2^{-n}]$ , onde  $k = 0, 1, \dots, 2^n - 1$  é o índice do subintervalo e  $n$ , chamada de resolução ou camada, indica o número de subintervalos. A Figura 2 apresenta a cascata multiplicativa binomial, onde  $r_{n,j}$  são os multiplicadores da cascata, cujo valores estão entre  $[0,1]$ .



**Figura 2:** Cascata multiplicativa binomial

A transformada *wavelet* discreta é empregada para representação multiescala de sinais da seguinte forma:

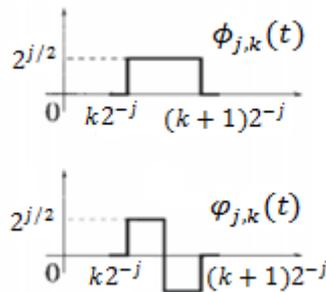
$$f(t) = \sum_k U_{j_0,k} \phi_{j_0,k}(t) + \sum_{j=j_0}^{\infty} \sum_k W_{j,k} \varphi_{j,k}(t) \quad (7)$$

onde,  $W_{j,k}$  e  $U_{j_0,k}$  são os coeficientes *wavelet* e de escala, respectivamente, dados por:

$$W_{j,k} = \int f(t) \varphi_{j,k}(t) dt \quad (8)$$

$$U_{j,k} = \int f(t) \phi_{j,k}(t) dt \quad (9)$$

e  $\phi_{j,k}(t)$  e  $\varphi_{j,k}(t)$  são as funções de escala e *wavelet* de Haar, respectivamente, representadas na Figura 3.



**Figura 3:** Funções de escala  $\phi_{j,k}(t)$  e *wavelet* de Haar  $\varphi_{j,k}(t)$ .

Os coeficientes de escala dados pela equação (9), podem ser recursivamente calculados utilizando a *wavelet* de Haar  $\varphi_{j,k}(t)$  através das seguintes equações:

$$U_{j,2k} = 2^{-\frac{1}{2}}(U_{j-1,k} + W_{j-1,k}) \quad (10)$$

$$U_{j,2k+1} = 2^{-\frac{1}{2}}(U_{j-1,k} - W_{j-1,k}) \quad (11)$$

Este cálculo recursivo é repetido até que se alcance a resolução desejada ou até que se obtenha o número desejado de amostras, formando assim, uma árvore binária de coeficientes de escala. Os coeficientes  $U_{j,k}$  representam a média local do processo em escala e deslocamento de tempo diferente. A fim de se assegurar a não-negatividade da série de tráfego sintética, gerados pelo modelo MWM, deve se impor algumas condições a seus coeficientes *wavelet* de escala. A condição  $X(t) \geq 0, \forall t$ , impõe que,  $U_{j,2k+1} \geq 0, \forall j$  e  $k$ . Impondo a condição  $U_{j,k} \geq 0, \forall j$  e  $k$ , pode-se afirmar que  $|W_{j,k}| \leq U_{j,k}, \forall j$  e  $k$ . Os coeficientes *wavelet* são gerados a partir da equação (Riedi, et al., 1999):

$$W_{j,k} = U_{j,k} A_{j,k} \quad (12)$$

onde  $A_{j,k}$  é uma variável aleatória, cujo valor está entre o intervalo  $[-1,1]$ . Os multiplicadores  $A_{j,k}$  são independentes e identicamente distribuídos (i.i.d) dentro de cada escala, são também independentes de  $U_{j,k}$  e simétricos em torno de zero. No caso particular do  $\beta$ MWM, os multiplicadores  $A_{j,k}$  são modelados por uma distribuição beta simétrica, cuja função densidade de probabilidade (pdf–probabilitydensityfunction) é dada por (Spiegel & Liu, 2011):

$$f(x) = \frac{(1+x)^{p-1}(1-x)^{p-1}}{B(p,p)2^{2p-1}} \quad (13)$$

onde  $B(.,.)$  é a função beta e  $p$  é o parâmetro que determina a forma da distribuição. Os multiplicadores  $A_{j,k}$  são escolhidos de maneira a controlar as energias dos coeficientes *wavelet*. Assim, o decaimento de energia ( $n_j$ ) dos coeficientes *wavelet* é dado por (Riedi, et al., 1999):

$$n_j = \frac{E(W_{j-1,k}^2)}{E(W_{j,k}^2)} = \frac{2p_j + 1}{2p_{j-1} + 1} \quad (14)$$

e

$$(2p_0 + 1)E(W_{0,0}^2) = E(U_{0,0}^2) \quad (15)$$

Nota-se que  $p_j$  é usado para obter o decaimento da energia dos coeficientes *wavelet* em escala. O coeficiente  $U_{0,0}$  na maior escala, ou seja, escala mais fina, é modelado como sendo uma variável aleatória normal com a média ( $\mathbb{E}_c$ ) e a variância ( $\sigma_c^2$ ) iguais aos dos coeficientes de escala dos fluxos de tráfego reais.

É possível relacionar no  $\beta$ MWM, o decaimento de energia ( $n_j$ ) dos coeficientes *wavelet*, por camada  $j$  (escala), com valores dos parâmetros  $p_j$  das distribuições beta simétrica, que é utilizada para modelar os multiplicadores  $A_{j,k}$ . Com isso, é possível estimar os parâmetros  $p_k$  recursivamente pela seguinte equação (Riedi, et al., 1999):

$$p_j = \frac{n_j}{2}(p_{j-1} + 1) - \frac{1}{2} \quad (16)$$

Considerando um tempo discreto  $k$ , o processo discreto de tráfego MWM é obtido pelos coeficientes de escala  $U_{j,k}$  na maior escala  $j$ , como se segue:

$$X[k] = 2^{-j} U_{j,k} \quad (17)$$

O processo estocástico a partir do modelo  $\beta$ MWM, na camada  $n$ , é dado por (Riedi, et al., 1999):

$$C^n[k] = 2^{-n} Norm(\mathbb{E}_c, \sigma_c^2) \prod_{j=0}^{n-1} (1 + \beta(p_j, p_j)) \quad (18)$$

Onde  $p_j$ ,  $\mathbb{E}_c$  e  $\sigma_c^2$  são os parâmetros do modelo  $\beta$ MWM;  $Norm(\mathbb{E}_c, \sigma_c^2)$  é uma variável aleatória normal, com média  $\mathbb{E}_c$  e variância  $\sigma_c^2$ ; e  $\beta(p_j, p_j)$  é uma variável aleatória beta com p.d.f (função densidade de probabilidade) dada pela equação (13).

Uma desvantagem do MWM é o número de parâmetros a serem estimados, que para isso, faz uso de toda a série de tráfego. Desta forma, o MWM se torna inadequado para aplicações em tempo real (Gonçalves, et al., 2013). Neste trabalho é utilizado o modelo  $\beta$ MWM Adaptativo, proposto por (Gonçalves, et al., 2013), para estimar em tempo real a banda efetiva do tráfego dos usuários, afim de tomar decisões de escalonamento dos recursos de rádio de um sistema de comunicação LTE.

#### 4.1 Modelagem $\beta$ MWM Adaptativa

Em (Gonçalves, et al., 2013) é proposto um algoritmo para cálculo adaptativo dos parâmetros do modelo multifractal  $\beta$ MWM. Ao invés de processar todos os dados da série de tráfego em uma única etapa, os dados são processados de forma iterativa em janelas de tamanho fixo de  $2^J$  amostras, onde  $J$  é o número de camadas da cascata. Não há a necessidade de armazenar uma grande quantidade de dados sobre o fluxo, uma vez que apenas algumas variáveis são armazenadas no processo de modelagem.

O algoritmo funciona da seguinte forma (Gonçalves, et al., 2013):

**1º passo-** As variáveis do modelo são inicializadas. Assume-se que o segundo momento dos coeficientes *wavelet* seja nulo, i.e.,  $E[W_{j,k}^2](0) = 0$ ; assim como a média e a variância dos coeficientes de escala,  $\mu_c(0) = 0$  e  $\sigma_c^2(0) = 0$ , e o contador de janela  $n = 0$ ;

**2º passo-** Realiza-se a transformada de Haar na janela de dados. A transformada de Haar em cada janela de  $2^j$  amostras de tráfego, gera  $2^j$  coeficientes *wavelet*, nomeados de  $\tilde{W}_{j,k}$ , pra cada camada  $j$  e um coeficiente de escala, nomeado de  $\tilde{U}_{0,0}$ , na camada  $j = 0$ ;

**3º passo-** Atualiza-se o segundo momento  $E[W_{j,k}^2]$  dos coeficientes *wavelet* através da equação:

$$E[W_{j,k}^2](n+1) = E[W_{j,k}^2](n) \frac{n}{n+1} + \frac{\sum_{i=0}^{2^j-1} \tilde{W}_{j,i}^2}{(n+1)2^j} \quad (19)$$

**4º passo-** As taxas de energia  $n_j$  são recalculadas segundo a equação:

$$n_j = \frac{E(W_{j-1,k}^2)}{W_{j,k}^2} = \frac{2p_j + 1}{2p_{j-1} + 1} \quad (20)$$

e os parâmetros  $p_j$  são recalculados segundo a equação:

$$p_j = \frac{n_j}{2} (p_{j-1} + 1) - \frac{1}{2} \quad (21)$$

**5º passo-** As estatísticas dos coeficientes de escala são atualizadas segundo as equações:

$$\mu_c(n+1) = \mu_c(n) \left( \frac{n}{n+1} \right) + \frac{\tilde{U}_{0,0}}{n+1} \quad (22)$$

$$\sigma_c^2(n+1) = (\sigma_c^2(n) + \mu_c^2(n)) \left( \frac{n}{n+1} \right) - \mu_c^2(n+1) + \frac{(\tilde{U}_{0,0})^2}{n+1} \quad (23)$$

Os passos 2, 3, 4 e 5 são repetidos a cada nova janela de dados de  $2^j$  amostras, incrementando o valor da variável  $n$  em 1. Com isso, obtém-se os parâmetros do modelo  $\beta$ MWM que são:  $p_j, \mu_c$  e  $\sigma_c^2$ .

Na Tabela 1 é mostrada uma comparação de parâmetros estatísticos de duas séries de tráfego reais TCP/IP, entre a Universidade de *Waikato* e o resto do mundo, coletadas em 07/04/2011 a 05/11/2011 e podem ser encontradas em <http://wand.net.nz/wits/waikato/8/>. Os dados das séries de tráfego foram agregados em intervalos de 10ms. As séries sintéticas geradas através da modelagem  $\beta$ MWM tradicional e adaptativa possuem 500 amostras, foram utilizadas 5 camadas para a cascata multiplicativa binomial. Nota-se que os parâmetros: média e desvio-padrão da modelagem  $\beta$ MWM adaptativa apresentaram erros menores do que 1%, em relação aos das séries reais. Esses parâmetros são próximos dos obtidos com a modelagem  $\beta$ MWM tradicional. Assim, conclui-se que a modelagem  $\beta$ MWM adaptativa pode ser aplicada para nossos propósitos.

<b>Tabela 1:</b> Estatísticas de séries reais e séries sintéticas geradas pelo modelo $\beta$ MWM adaptativo				
<b>Série</b>	<b>Waikato10ms-1</b>		<b>Waikato10ms-2</b>	
	Média	Desvio Padrão	Média	Desvio Padrão
Real	2.5968	1.9498	3.3987	2.2775
$\beta$ MWM tradicional	2.5468	1.9549	3.3842	2.1916
$\beta$ MWM adaptativa	2.5839	1.9541	3.4095	2.2934
Erro (entre $\beta$ MWM adaptativa e Real)	0,50%	-0,22%	-0,32%	-0,7%
Erro (entre $\beta$ MWM tradicional e Real)	1,93%	-0,26%	0,43%	3,77%

## 4.2 Determinação da Banda Efetiva dos Usuários

A banda efetiva representa a taxa de transferência de dados mínima, para atender a parâmetros de QoS exigida para um fluxo de tráfego. Um requisito de QoS frequentemente associado à teoria de banda efetiva é a probabilidade de transbordo do *buffer*. As aplicações reais que utilizam esta teoria, geralmente fazem uso da modelagem do fluxo de tráfego para estimar a banda

efetiva ou a estimam de forma direta a partir do fluxo de tráfego.

Seja  $X[0,t]$  o tráfego acumulado durante o intervalo de tempo  $[0,t]$ , para um fluxo de tráfego, e tendo  $X[0,t]$  incremento estacionário, ou seja, incrementos no processo são independentes do tempo de observação, mas dependentes do tamanho do intervalo observado. A banda efetiva para um determinado fluxo de dados é definida por Kelly (Kelly, 1996) como:

$$\alpha(s, t) = \frac{1}{st} \ln(E[e^{sX[0,t]}]), \quad s > 0 \text{ e } t < \infty \quad (24)$$

onde  $s$  e  $t$  são parâmetros de espaço e tempo, respectivamente, que determinam os requisitos de QoS e as características do tráfego. A banda efetiva pode ser calculada, através da função geradora de momentos do processo  $X[0,t]$ , dado por:

$$\Lambda(s, t) = E[e^{sX[0,t]}] \quad (25)$$

Quando fluxos de tráfego são simultaneamente servidos com taxas equivalentes as suas bandas efetivas, requisitos de QoS são atendidos para estes fluxos (Duffield & Connell, 1993). Quando há apenas uma fonte ( $N = 1$ ) e a capacidade da rede for igual à banda efetiva da fonte, pode-se aproximar a probabilidade de transbordo do *buffer* por (Vieira, et al., 2004):

$$\ln P(Q_N > B) \approx Bs \quad (26)$$

$$P(Q_N > B) \approx \exp(Bs) \quad (27)$$

onde  $Q_N$  é a quantidade estacionária de dados na fila,  $N$  é o número de fontes multiplexadas e  $B$  é o tamanho do *buffer*.

A partir da equação (18) e (24), a banda efetiva pode ser definida como (Gonçalves, et al., 2013):

$$\tilde{\alpha}(s, 2^n) = \frac{1}{st} \ln(E[e^{[2^{-J+n} \text{Norm}(\varpi_c, \sigma_c^2) \prod_{j=0}^{J-n} (1 + \beta(p_j, p_j))]]]) \quad (28)$$

onde  $J$  é número de camadas da cascata multiplicativa no domínio *wavelet* (Chui, 1992) e  $t = 2^n$ ,  $0 \leq n \leq J - 1$ , pois a cascata multiplicativa do MWM é diádica.

Os parâmetros do modelo  $\beta$ MWM,  $p_j$ ,  $\varpi_c$  e  $\sigma_c^2$ , são obtidos a partir do algoritmo apresentado na seção 4.1, para estimar a banda efetiva dos usuários em tempo real, através da equação (28). A estimativa da banda efetiva adaptativa é utilizada no esquema de alocação de recursos de rádio proposto, para garantir o atendimento de parâmetros de QoS dos usuários. Um dos métodos para estimação de banda efetiva, que é uma simplificação da equação (24), é o estimador direto (29) (Gibbens, 1996). O estimador direto da banda efetiva pode ser calculado utilizando a média temporal das amostras de tráfego, ao invés do valor esperado na equação (24), ou seja:

$$\tilde{\alpha}(s, t) = \frac{1}{st} \ln(E[\frac{1}{N-t} \int_0^{N-t} e^{s \sum_{i=1}^N x_i I(\tau \leq t_i \leq \tau+t)} d\tau]) \quad (29)$$

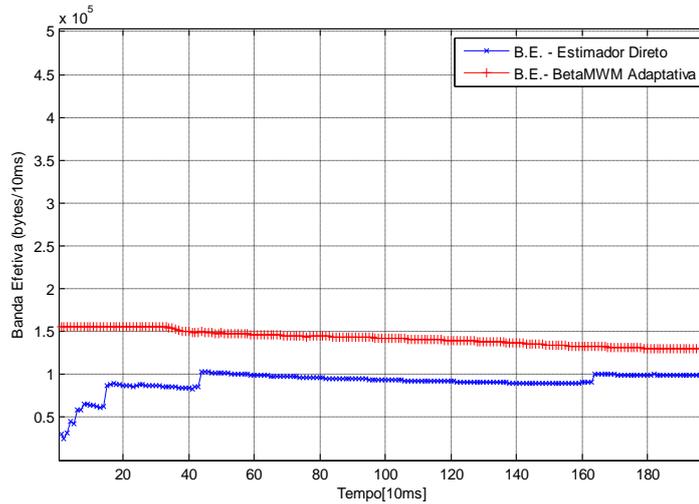
onde  $N$  é o tamanho da série;  $t_i$  representa o tempo da amostra  $x_i$ ; e  $I(\tau \leq t_i \leq \tau + t)$  é 1 se  $t_i$  estiver no intervalo  $\tau \leq t_i \leq \tau + t$  e 0 caso contrário.

Na Figura 4 é mostrado o resultado do estimador de banda efetiva direto e do método  $\beta$ MWM adaptativo, para uma série de tráfego<sup>1</sup> TCP/IP, com volume de dados agregados em intervalos de 10ms. Foi utilizado na simulação o parâmetro de espaço  $s$  igual a  $-\ln(0,01)/(60 \times 1024)$ , ou seja, para atender uma probabilidade de transbordo de 1% para um tamanho de *buffer* de 60 kB.

Analisando a Figura 4, nota-se que o valor de banda efetiva calculado pelo estimador  $\beta$ MWM adaptativo, superestima o valor dado pelo estimador direto. Portanto, não viola os requisitos de QoS (probabilidade de perda), que são efetivamente atendidos pela banda do estimador direto.

A estimação de banda efetiva sendo adaptativa, diminui a quantidade de dados armazenados para efetuar a atualização dos parâmetros do modelo  $\beta$ MWM e ainda consegue acompanhar melhor as variações bruscas das características dos fluxos de tráfego.

<sup>1</sup>Tráfego TCP/IP entre a Universidade de Waykato e o resto do mundo, coletada em 07/04/2011 a 05/11/2011. A série de tráfego pode ser encontrada em <http://wand.net.nz/wits/waikato/8/>.



**Figura 4:** Comparação entre o estimador direto e o estimador  $\beta$ MWM adaptativo para a série Waykato10ms-3.

## 5 Esquema de Alocação de Blocos de recurso Usando Prioridade Inteligente Baseada em Lógica Fuzzy

O esquema de alocação de blocos de recurso proposto adota uma disciplina de escalonamento baseada em prioridade e um algoritmo de alocação de blocos de recurso (RB), considerando as restrições do Sistema LTE, com relação à utilização do esquema de modulação e código (MCS). No sistema LTE, o MCS adotado por um usuário em um intervalo de tempo de transmissão (TTI - *Transmission Time Interval*), deve ser o mesmo para todos SBs atribuídos a ele, em uma configuração de antena única.

Para cada usuário é determinada uma prioridade, calculada de forma adaptativa, através do sistema de inferência *fuzzy*, considerando a qualidade média do canal, a estimativa da banda efetiva do tráfego de dados, para o atendimento de parâmetros de QoS, e a taxa de dados alcançada por cada usuário. A alocação de cada bloco de escalonamento (SB) é feita após a determinação da prioridade dos usuários. As prioridades dos usuários são atualizadas após cada alocação de recurso, uma vez que são recalculados a qualidade média do canal e a taxa de dados alcançada por cada usuário, considerando apenas os blocos de escalonamento ainda disponíveis.

### 5.1 Algoritmo para Determinação de Prioridade Baseado em Lógica Fuzzy

O algoritmo de determinação de prioridade baseado em lógica *fuzzy* determina a prioridade de cada usuário, considerando o indicador de qualidade do canal (CQI) e a estimativa da banda efetiva do tráfego de dados a cada TTI, calculada através da modelagem  $\beta$ MWM adaptativa (Gonçalves, et al., 2013), para atendimento de QoS (taxa de perda). Denotamos por  $QM_k$  a qualidade média do canal do usuário  $k$ , calculada da seguinte forma:

$$QM_k = \frac{1}{N} \sum_{n=1}^N g_{k,n} \quad (30)$$

onde  $g_{k,n}$  é o indicador de qualidade de canal do  $k$ -ésimo usuário no  $n$ -ésimo SB e  $N$  é a quantidade de SBs disponíveis naquele instante.

O valor de  $QM_k$  é atualizado após a alocação de cada SB, considerando apenas os canais disponíveis (ainda não alocados). Quanto maior o valor de  $QM_k$  do usuário  $k$ , maior é a sua taxa de dados, assim, maior deve ser a prioridade dada ao usuário. Denotamos por  $QFI_k$ , a informação de atendimento de QoS do usuário  $k$ , calculada da seguinte forma:

$$QFI_k = \frac{(Be_k - r_k)}{Be_k} \quad (31)$$

onde  $Be_k$  é a estimativa da banda efetiva do usuário  $k$ , para atendimento de parâmetros de QoS, e  $r_k$  é a taxa de dados alcançada pelo usuário  $k$ . O valor de  $QFI_k$  é atualizado após a alocação de cada SB. Quanto maior o valor de  $QFI_k$  do usuário  $k$ , maior é a distância do atendimento de QoS, assim, maior deve ser a prioridade dada ao usuário.

São utilizados  $QM_k$  e  $QFI_k$  como variáveis linguísticas de entrada de um sistema de inferência *fuzzy* para gerar a prioridade de cada usuário. O conjunto de termos *fuzzy* para  $QM_k$  é definido como  $T(QM_k) = \{\text{Ruim}(R), \text{Bom}(B), \text{Ótimo}(O)\}$ . O conjunto de termos *fuzzy* para  $QFI_k$  é definido como  $T(QFI_k) = \{\text{Muito Baixo}(MB), \text{Baixo}(B), \text{Médio}(M), \text{Alto}(A)\}$ . É utilizada uma distribuição gaussiana para modelar as funções de pertinência  $\mu_X(QM_k)$  e  $\mu_Y(QFI_k)$ . A variável linguística de saída é o  $\alpha_k$ , cujos termos são definidos como  $T(\alpha_k) = \{A, B, C, D, E\}$ . A função de pertinência  $\mu_Z(\alpha_k)$  é definida por valores constantes, escolhidos heurísticamente,  $Z = 0,1; 0,3; 0,6; 0,9$  ou  $1,2$ . Esses valores foram escolhidos como resultado de experimentos e conhecimento prévio do problema de maximização (equação 4).

Para a redefinição de determinação de prioridade, foi utilizado um sistema de inferência baseado no modelo *fuzzy* TSK (Takagi, Sugeno e Kang) (Sugeno & Kang, 1988) (Takagi & Sugeno, 1985), com a utilização da técnica Max-Prod (Lee, 1990). Assim, a função de pertinência resultante da regra  $j$ , para o usuário  $k$ , é dada por  $\mu_{k,j} = \mu_X(QM_k) \cdot \mu_Y(QFI_k)$ , e a função de pertinência na saída da rede *fuzzy* é dada por,  $\mu_Z(\alpha_k) = \max_j(\mu_{k,j})$ , sendo  $1 \leq j \leq$  (quantidade de regras *fuzzy*),  $X = R, B$  ou  $O$  e  $Y = MB, B, M$  ou  $A$ . O valor de  $\alpha_k$  do usuário  $k$ , é obtido pela média ponderada das funções de pertinência resultantes das regras *fuzzy*. Para obter os parâmetros das funções de pertinência (premissas) e os parâmetros do consequente do sistema de inferência *fuzzy*, foi utilizado um algoritmo de treinamento híbrido, baseado no estimador de mínimos quadrados (LSE – *Least-square Estimator*) e no método da descida mais íngreme, que atualizam os parâmetros lineares e os parâmetros não-lineares, respectivamente, a cada iteração do algoritmo. Os conjuntos de dados de treinamento utilizados são gerados de forma autônoma, através de um algoritmo descrito a seguir.

O algoritmo para geração dos dados de treinamento da variável  $QM_k$  gera três conjuntos de valores, considerando os termos Ruim ( $R$ ), Bom ( $B$ ) e Ótimo ( $O$ ). Estes conjuntos são gerados com valores dentro do intervalo compreendido entre o maior e o menor valor da qualidade dos canais de todos usuários, medidos naquele instante, sendo o primeiro intervalo referente ao termo  $R$ , o segundo intervalo referente ao termo  $B$  e o último intervalo referente ao termo  $O$ . Para a variável  $QFI_k$ , o algoritmo gera quatro conjuntos de dados, considerando os termos Muito Baixo ( $MB$ ), Baixo ( $B$ ), Médio ( $M$ ) e Alto ( $A$ ). Estes conjuntos são gerados com valores dentro do intervalo compreendido entre 0 e 1, sendo o primeiro intervalo referente ao termo  $MB$ , o segundo intervalo referente ao termo  $B$ , o terceiro intervalo referente ao termo  $M$  e o quarto intervalo referente ao termo  $A$ . Assim, a rede *fuzzy* gera as funções de pertinência de forma autônoma, para tomar as decisões de escalonamento dos recursos de rádio.

As escolhas das regras *fuzzy* mostradas na Tabela 2, são justificadas pelo seguinte: os usuários com maior  $QFI_k$ , ou seja, mais distantes de atingirem os requisitos de QoS, devem ter maior prioridade, independente da condição de canal, o que é representado pelas regras 4, 8 e 12. Se os usuários possuem o mesmo  $QFI_k$ , o usuário com melhor condição média de canal, ou seja, com maior  $QM_k$ , deve ter maior prioridade, o que é representado pelas regras 2, 6 e 10.

**Tabela 2: Regras *fuzzy*.**

Regras	1	2	3	4	5	6	7	8	9	10	11	12
$QM_k$	R	R	R	R	B	B	B	B	O	O	O	O
$QFI_k$	MB	B	M	A	MB	B	M	A	MB	B	M	A
$\alpha_k$	0,1	0,3	0,6	1,2	0,3	0,6	0,9	1,2	0,6	0,9	1,2	1,2

## 5.2 Algoritmo para Alocação de Blocos de recurso Baseado em Prioridade Inteligente

O algoritmo proposto para alocação de blocos de recurso baseado em prioridade inteligente, aloca os SBs tendo em vista a determinação das prioridades calculadas pelo sistema de inferência *fuzzy*. Assim, é escolhido o usuário com o maior valor de  $\alpha_k$ , para receber, naquele instante, um SB. Após cada decisão do algoritmo de alocação de blocos de recurso, são atualizados os valores de  $QM_k$  e  $QFI_k$ , através das equações (30) e (31), respectivamente, considerando os valores de condição média de canal, taxa de dados alcançada e banda efetiva estimada de cada usuário naquele instante, através da modelagem multifractal/MWM adaptativa.

Após a alocação de todos SBs, o MCS a ser adotado por cada usuário é escolhido considerando o SB atribuído com pior valor de CQI do usuário. Para garantir uma taxa de erros de bloco menor ou igual a 10%, é utilizada a Tabela 3 (Kawser, et al., 2012), que define o esquema de modulação e código a ser utilizado, tendo em vista a SNR (*signal-to-Noise-Ratio*) dos usuários.

**Tabela 3:** SNRs associados ao MCS(Kawser, et al., 2012)

Nível MCS	SNR (dB)	MCS	
		Modulação	Taxa de Código x 1024
1	1,95	QPSK	78
2	4	QPSK	120
3	6	QPSK	193
4	8	QPSK	308
5	10	QPSK	449
6	11,95	QPSK	602
7	14,05	16QAM	378
8	16	16QAM	490
9	17,9	16QAM	616
10	19,9	64QAM	466
11	21,5	64QAM	567
12	23,45	64QAM	666
13	25	64QAM	772
14	27,3	64QAM	873
15	29	64QAM	948

## 6 Resultados e Simulações

### 6.1 Parâmetros de Simulação

Foram realizadas simulações considerando os dados da Tabela 4 e quatro séries reais de tráfego TCP/IP, referente ao tráfego entre a Universidade de Waikato e o resto do mundo, coletadas em 07/04/2011 a 05/11/2011 (<http://wand.net.nz/wits/waikato/8/>). As séries de tráfego consideradas foram: *waikatoVIII-20110520-000000-0*; *waikatoVIII-20110623-230233-4*; *waikatoVIII-20110921-000000-0* e *waikatoVIII-20111029-110001-1*. Foram utilizadas nas simulações 16 usuários sendo atendidos por uma estação base. Nas simulações, cada série real corresponde ao tráfego de 4 usuários diferentes. O sistema LTE e as simulações foram implementados no *software* Matlab.

A SNR (*Signal-to-Noise-Ratio*) dos usuários para cada bloco de recurso da rede, variando a cada TTI, foi calculada conforme equação:

$$SNR = 10 \log \left( \frac{H^2 Pr}{No} \right) \quad (32)$$

onde,  $H$ ,  $Pr$  e  $No$  são a resposta de frequência do canal de *Rayleigh*; a potência recebida na estação base; e a potência de ruído branco; referente a um bloco de escalonamento de 180KHz, respectivamente.

Foi considerado um modelo de perda de percurso de  $128,1 + 37,6 \log(R)$  dB, com sombreamento log-normal com média 0 e desvio padrão de 10dB (3GPP TR 36.942 version 10.2.0, 2011), sendo  $R$  igual a 1,2Km, que representa a distância dos usuários em relação a estação base. É assumido que as informações da qualidade do canal (CQI) são totalmente conhecidas na estação base e que todas as subportadoras são utilizadas para a transmissão de dados.

Para avaliar o esquema de escalonamento proposto em termos de justiça na alocação de recursos, foi utilizado o índice de justiça. O índice de justiça, considerando a banda efetiva dos usuários como restrição, foi calculado pela seguinte equação (Jain, et al., 1999):

$$\text{Índice Justiça} = \frac{[\sum_{i=1}^k x_i]^2}{k \sum_{i=1}^k x_i^2} \quad (33)$$

onde  $x_i = \frac{r_k}{Be_k}$  é a vazão normalizada; e  $r_k$  e  $Be_k$  são a taxa de dados alcançada e a estimativa de banda efetiva do usuário  $k$ , respectivamente.

Nas próximas subseções é apresentada uma comparação de desempenho em termos de vazão, taxa de perda de dados, tempo de retardo e índice de justiça, para o algoritmo PSO (Su & Ping Wang, 2012), para o algoritmo apresentado em (Guan, et al., 2011), que denominamos de QoS Garantido, e para o esquema de alocação proposto.

O algoritmo de alocação de blocos de recurso QoS Garantido é dividido em duas etapas. Na primeira etapa é estimada a quantidade de SBs necessárias para cada usuário. Esta estimativa é calculada através da relação entre a taxa mínima requerida e a condição média de canal de cada usuário. Na segunda etapa, os SBs são alocados conforme suas prioridades. Esta prioridade é calculada, considerando o usuário com maior condição média de canal e menor taxa mínima requerida, com maior prioridade. Assim, se o usuário tiver a maior média de condição de canal, terá maior prioridade. Se tiverem a mesma média de condição de canal, o usuário com menor taxa mínima requerida terá maior prioridade. A alocação de blocos de recurso é realizada na sequência das prioridades definidas na segunda etapa e atendendo a quantidade de SBs estimadas na primeira etapa. Se o usuário não alcançar a taxa mínima requerida, com a quantidade de SBs estimadas na primeira etapa, é alocado mais SBs até atingir a este objetivo.

<b>Tabela 4: Configuração do Sistema LTE</b>	
Número total de subportadoras	300
Número de SBs	25
Tamanho do <i>subframe</i>	1ms
Símbolos OFDM por <i>slot</i>	7
Quantidade de usuários na rede	16
Tamanho do canal	5MHz
Frequência da Portadora	2GHz
Modo de Transmissão	Antena única
Ganho da antena do eNodeB	15dBi
Ganho da antena do equipamento do usuário	0dBi
Potência de transmissão do eNodeB	43 dBm
Densidade Espectral de Potência do ruído branco	-174dBm/Hz
Modelo do canal	<i>Rayleigh</i>
Máximo <i>doppler</i>	30Hz
Dispersão temporal	5us

\*(3GPP TR 36.942 version10.2.0, 2011)

Em todos os algoritmos simulados foi utilizado o valor estimado da banda efetiva, calculada de forma adaptativa, ao invés da utilização de taxa mínima requerida, que foi utilizado em (Su&Ping Wang, 2012) e (Guan, et al., 2011).

A otimização PSO inteira foi realizada considerando uma população de 30 indivíduos, critério de parada de 100 iterações, peso de inércia  $w = (100 - i) / 100$  (onde  $i$  é o número da iteração, iniciando em 0) e parâmetros  $c_1 = 1$  e  $c_2 = 3$ .

## 6.2 Comparações de Desempenho dos Escalonadores

Na Figura 5 são apresentadas as estimativas de banda efetiva dos fluxos de tráfego dos usuários a cada TTI, calculadas de forma adaptativa, utilizando a modelagem multifractal  $\beta$ MWM. Nas Figuras 6, 7, 8 e 9 são mostradas a vazão total do sistema, o tempo de retardo, a taxa de perda de dados e o índice de justiça, respectivamente, em função do TTI. A vazão total do sistema é o resultado da soma das taxas de dados de todos os usuários atendidos na estação base, após a alocação dos recursos de rádio. A taxa de perda de dados representa, em termos percentuais, a quantidade da banda efetiva estimada total que não foi atendida pela rede, considerando todos os usuários, ou seja, a taxa de bytes de dados perdidos, tendo em vista o atendimento de parâmetros de QoS dos usuários. Já o tempo de retardo calculado se refere ao tempo gasto pela rede, para escoar o tráfego de todos os usuários (foi considerado um buffer de tamanho 60KB para cada usuário).

Na Figura 6 e 7 pode-se observar que o esquema de alocação de blocos de recurso proposto apresenta, em geral, maior vazão do sistema e menor tempo de retardo, quando comparado com os algoritmos PSO e o QoS Garantido. Nota-se também, que o esquema proposto consegue manter praticamente constante, ao longo do tempo observado, a vazão do sistema e o tempo de retardo, sendo o menos susceptível às variações do canal. É possível verificar que a lógica nebulosa, juntamente com a estratégia de alocação de recursos proposto, consegue prover uma solução sub-ótima para o problema de otimização (4), apresentando um ganho com relação aos algoritmos comparados.

Na Figura 8 é apresentada a taxa de perda de dados do sistema para todos os usuários. Podemos observar que o esquema de alocação proposto consegue manter, em quase todo o período de observação, perdas em torno de 1%, sendo, em geral, menores do que os outros algoritmos comparados. Os algoritmos PSO e QoS Garantido, apresentam picos de perda de aproximadamente 9,5% e 4%, respectivamente, que são maiores que a do esquema proposto.

Na Figura 9, nota-se que o algoritmo QoS Garantido possui o maior índice de justiça, em comparação aos três algoritmos apresentados. Já o esquema de alocação proposto supera o algoritmo PSO, na maior parte do tempo de observação. Embora o algoritmo QoS Garantido consegue alcançar um bom índice de justiça, este algoritmo apresenta um desempenho inferior, em

relação a vazão do sistema, ou seja, ao problema de otimização (4), quando comparado aos demais algoritmos apresentados.

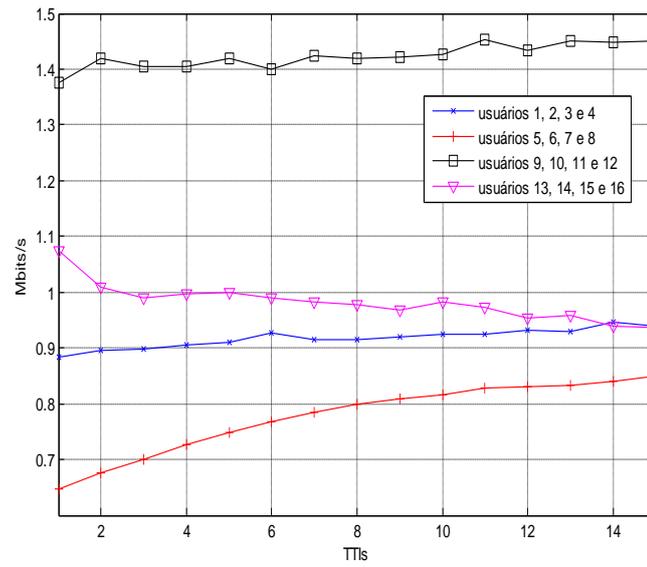


Figura 5: Estimativa da banda efetiva dos usuários, através da modelagem multifractal  $\beta$ MWM adaptativa.

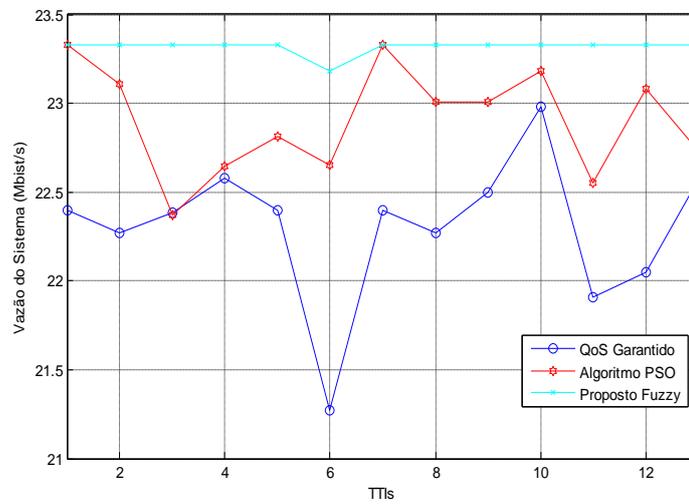
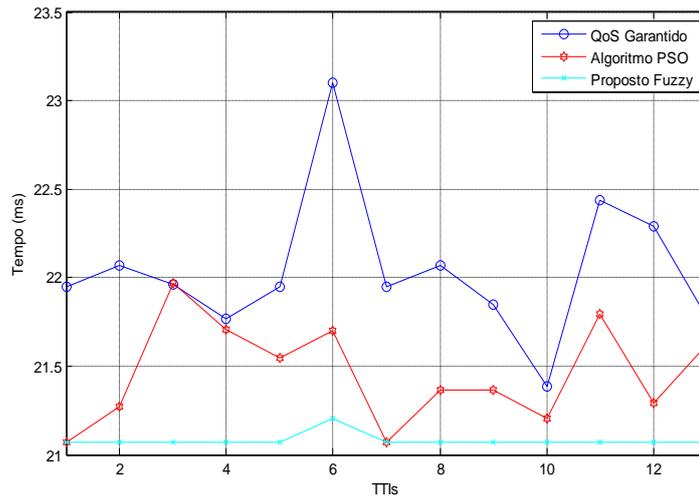
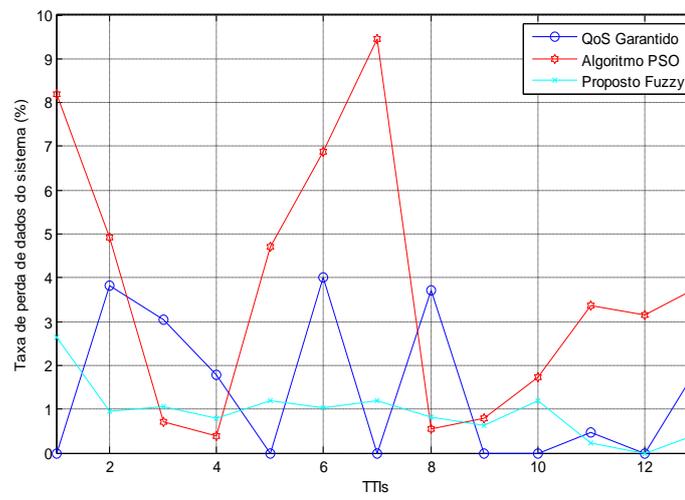


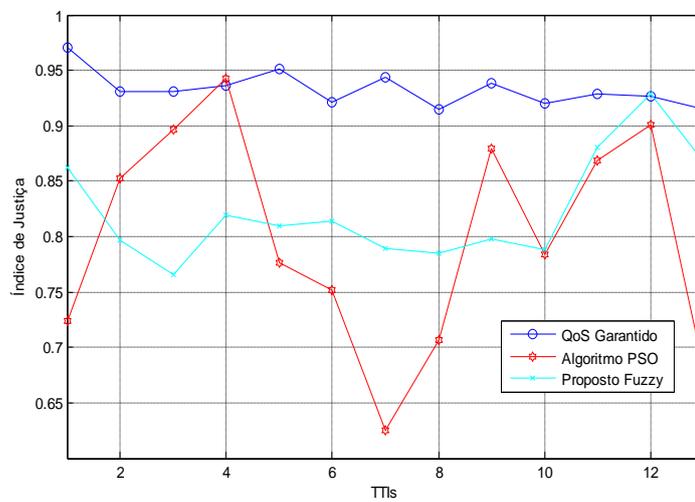
Figura 6: Vazão total do sistema.



**Figura 7:** Tempo de retardo do sistema.



**Figura 8:** Taxa de perda de dados do sistema.



**Figura 9:** Índice de justiça.

## 7 Trabalhos Relacionados

Os algoritmos de escalonamento são muito importantes para as redes de comunicação sem fio, uma vez que permitem o compartilhamento justo de recursos de rádio em um sistema multiusuário e a garantia de parâmetros de desempenho e de QoS. Além dos algoritmos abordados neste trabalho, PSO(Su & Ping Wang, 2012) e QoS Garantido(Guan, et al., 2011), existem outros algoritmos conhecidos na literatura, com o objetivo de aumentar a QoS dos sistemas móveis. Como exemplo, temos: *Round Robin* (RR)(Shreedhar & Varghese, 1996), *Maximum Carrier Interference* (Max-C/I) e *Proportional Fair* (PF)(Jalali, et al., 2000). Outros trabalhos baseados em lógica fuzzy, para o escalonamento de recursos de rádio em sistemas de *downlink* LTE conhecidos na literatura, são: (Chung, et al., 2012) e (Khan, et al., 2013).

No algoritmo Max-C/I, a cada intervalo de tempo de transmissão da rede (TTI), todos os blocos de recurso disponíveis são alocados ao usuário com a melhor condição de canal instantânea. Esta característica do escalonador assegura ao sistema móvel altas taxas de dados. Por outro lado, os usuários em piores condições de canal são penalizados com maiores atrasos e indisponibilidade de serviço, como por exemplo, os que estão na borda da célula. Em situações reais é comum terminais móveis experimentarem diferentes condições de canal, devido às diferentes distâncias e desvanecimentos entre a estação base e os terminais móveis. Neste caso, um terminal móvel pode experimentar uma condição de canal pior do que a de outros terminais, por um tempo relativamente longo, fazendo com que este terminal não receba recursos de rádio.

O escalonador de dados RR se baseia no revezamento da utilização dos blocos de recurso de rádio disponíveis e não leva em consideração a condição de canal dos usuários. Esta estratégia é considerada justa, no sentido de que a mesma quantidade de recursos de rádio é atribuída aos usuários. Por outro lado, o escalonador não é justo, no sentido de proporcionar a mesma qualidade de serviço a diferentes usuários, com diferentes necessidades de QoS. Neste caso, mais recursos de rádio deveriam ser atribuídos aos usuários com pior condição de canal. Como o escalonador RR não leva em consideração as condições de canal instantâneas dos usuários no processo de alocação de recursos de rádio, tem-se um menor desempenho geral do sistema, mas a qualidade de serviço experimentada pelos diferentes terminais móveis será mais igualitária.

No escalonador PF os recursos de rádio são atribuídos ao usuário com a maior razão entre a taxa de dados instantânea e a taxa média de dados transmitidas, isto é, a cada intervalo de tempo de transmissão (TTI) é escolhido o usuário para receber todos os blocos de recurso disponíveis, conforme equação:

$$k = \operatorname{argmax} \frac{R_i}{\bar{R}_i} \quad (34)$$

onde  $R_i$  é a taxa de dados instantânea para o usuário  $i$  e  $\bar{R}_i$  é a média da taxa de dados do usuário  $i$ . Esta média é calculada por um período definido  $T_{PF}$ . O  $T_{PF}$  deve ser maior que a constante de tempo das variações de curto prazo do canal e menor do que o intervalo de tempo em que as variações de canal não são fortemente percebidas pelo usuário. Tipicamente  $T_{PF}$  é na ordem de 1 segundo (Dahlman, et al., 2007).

Nos algoritmos Max-C/I, RR e PF assume-se que todos os recursos de rádio no *downlink* são atribuídos a um único usuário de cada vez. O escalonamento é feito puramente no domínio do tempo usando TDM (*Time Division Multiplexing*) entre os usuários. No entanto, em várias situações, o TDM pode ser complementado pelo FDM (*Frequency-Division Multiplex*). Em princípio, existe uma razão para não depender exclusivamente de TDM no *downlink*. Em caso de carga insuficiente, onde a quantidade de dados a transferir para um usuário não é suficientemente grande para utilizar a capacidade total do canal, assim uma fração de recursos pode ser atribuído a outro usuário, alcançando assim, maiores desempenhos da rede.

Em (Chung, et al., 2012) é proposto um esquema de alocação de recursos em sistemas de *downlink* LTE-A, com tráfego multimídia. Este esquema é baseado na determinação de prioridades para os usuários, através de um sistema de inferência fuzzy, que considera os diferentes tipos de classes de serviço previstos no LTE-A, que são: serviço sensível ao atraso (DS – *Delay Sensitive*), serviço sensível a taxa (RS – *Rate Sensitive*) e serviço de melhor esforço (BE – *Best Effort*). Para o serviço DS, os requisitos de QoS considera a máxima BER (*Bit Error Ratio*), o atraso máximo e a taxa de perda de pacotes máxima. Para o serviço RS, os requisitos de QoS considera a máxima BER e a taxa mínima requerida. Para o serviço BE, os requisitos de QoS leva em consideração a máxima BER. Este esquema tem como objetivo maximizar a vazão da rede, com alocação justa e sem violar os diferentes requisitos de QoS dos usuários. Diferente do esquema proposto neste trabalho, em (Chung, et al., 2012), a prioridade de cada usuário é baseada em um constante padrão, que varia de acordo com a classe de serviço de cada usuário, podendo ser DS, RS ou BE, e uma prioridade inteligente. A prioridade inteligente é calculada por um sistema de inferência fuzzy que utiliza como variáveis de entrada, a qualidade do canal e a informação de atendimento do parâmetro de QoS, que utiliza uma equação heurística para cada tipo de classe de serviço.

Em (Khan, et al., 2013) é proposto um esquema de escalonamento de *downlink* para tráfegos sensíveis ao atraso, em sistemas sem fio OFDMA (*Orthogonal Frequency Division Multiple Access*), como o LTE/LTE-A. Este esquema utiliza a qualidade média do canal, juntamente com o atraso de pacote e a taxa de perda de pacote, para reduzir a violação de QoS dos usuários com baixa qualidade de canal e assim proporcionar um sistema justo. O esquema faz uso de uma estratégia de escalonamento,

com a combinação de um componente de prioridade no domínio do tempo, baseado em controladores *fuzzy*, e um componente de prioridade no domínio da frequência, baseado na informação da qualidade de canal por bloco de recurso físico. Diferente do esquema proposto neste trabalho, em (Khan, et al., 213), é utilizado como variáveis da rede *fuzzy*, a informação do atraso máximo de pacote e a soma ponderada do atraso normalizado com a média temporal da qualidade do canal. As prioridades calculadas pela rede *fuzzy* tem como objetivo não violar o atraso máximo permitido dos usuários. A arquitetura da rede *fuzzy* utilizado em (Khan, et al., 213), possui duas entradas, ambas com a mesma função de pertinência, uma saída e é baseada em 4 regras *fuzzy*.

No esquema proposto neste trabalho, diferentemente de (Khan, et al., 213) e (Chung, et al., 2012), o sistema de inferência *fuzzy* determina a prioridade de cada usuário, tendo como variáveis de entrada: a qualidade média do canal e a estimativa da banda efetiva do tráfego de dados, calculada através da modelagem  $\beta$ MWM adaptativa (Gonçalves, et al., 2013), para atendimento do parâmetro de QoS. É utilizado uma arquitetura para a rede *fuzzy*, com duas entradas, sendo composta por três e quatro conjuntos *fuzzy*, uma saída e doze regras *fuzzy*, diferentemente da utilizada por (Khan, et al., 213) e (Chung, et al., 2012).

O esquema de alocação de blocos de recurso proposto neste trabalho, bem como os algoritmos PSO (Su & Ping Wang, 2012) e o QoS Garantido (Guan, et al., 2011), utilizam o escalonamento no domínio do tempo e da frequência, com o objetivo de alcançar um maior desempenho do sistema e garantir a qualidade de serviço dos usuários.

## 8 Conclusão

Foi proposto neste trabalho, um esquema de alocação de blocos de recurso para redes de comunicação LTE, baseado em lógica *fuzzy*, banda efetiva e modelagem multifractal  $\beta$ MWM adaptativa. O esquema proposto tem como objetivo garantir que parâmetros de QoS dos usuários sejam atendidos, aumentar a vazão de dados e diminuir o tempo de retardo da rede, considerando as restrições do sistema. O esquema proposto foi comparado com algoritmos apresentados na literatura como, PSO (Su & Ping Wang, 2012) e o QoS Garantido (Guan, et al., 2011). Diferente destes dois algoritmos, o esquema proposto faz uso da estimativa da banda efetiva do fluxo de tráfego, calculada em tempo real, para atender a parâmetros de QoS dos usuários, ao invés da utilização de uma taxa mínima requerida.

Os resultados mostram que o esquema de alocação proposto apresenta, em geral, melhor desempenho no atendimento de parâmetros de QoS dos usuários, uma maior vazão para o sistema e menor tempo de retardo da rede, quando comparado com os algoritmos PSO e QoS Garantido. Podemos concluir, que o esquema proposto promove um bom equilíbrio entre o atendimento dos objetivos e as restrições do problema de alocação de recursos em sistemas LTE, utilizando estratégias de escalonamento e a lógica nebulosa.

## 9 Referências

- 3GPP TR 36.942 version 10.2.0, 2011. Evolved Universal Terrestrial Radio Access (E-UTRA); Radio Frequency (RF) system scenarios.
- 3GPP TSG RAN TR 25.913 v8.0.0, 2008. Requirement for Evolved Universal Terrestrial Radio Access (UTRA) and Universal Terrestrial Radio Access Network (UTRAN).
- Chui, C. K., 1992. *An Introduction to Wavelets*. San Diego: Academic Press.
- Chung, W.-C., Chung-Ju & Wang, L.-C., 2012. An Intelligent Priority Resource Allocation Scheme for LTE-A Downlink Systems. *IEEE Wireless Communications Letters*, 1(3), pp. 241-244.
- Dahlman, E., Parkvall, S., Sköld, J. & Beming, P., 2007. *3G Evolution HSPA and LTE for Mobile Broadband*. Oxford: Elsevier.
- Duffield, N. & Connell, N., 1993. Large Deviations and Overflow Probabilities for the General Single-Server Queue, With Applications. *Dublin Institute for Advanced Studies-Applied Probability Group, DIAS-STP-93-30*.
- Gibbens, R. J., 1996. Traffic Characterisation and Effective Bandwidths for Broadband Network Traces. Em: *Stochastic Networks: Theory and Application*, Royal Statistical. s.l.:Oxford University Press, pp. 169-179.
- Gonçalves, B. H. P., Vieira, F. H. T. & Costa, V. H. T., 2013. Alocação Dinâmica de Slots de Tempo Multiusuário para Redes OFDM/TDMA baseado em Banda Efetiva e Modelagem BMWM. *XXXI Simpósio Brasileiro de Telecomunicações - SBrT2013*, Setembro.
- Gonçalves, B. H. P., Vieira, F. H. T. & Costa, V. H. T., 2013. Modelagem Multifractal BetaMWM Adaptativa para Tráfego de Redes de Computadores. *X Encontro Anual de Computação*.
- Guan, N. et al., 2011. QoS Guaranteed Resource Block Allocation Algorithm for LTE Systems. *IEEE 7th International Conference on Wireless and Mobile Computing, Networking and Communications*.

- Jain, R., Durrresi, A. & Babic, G., 1999. Throughput Fairness Index: An Explanation. *Department of CIS, The Ohio State University, ATM\_Forum/99-0045*.
- Jalali, A., Padovani, R. & Pankaj, R., 2000. Data throughput of CDMA-HDR a High Efficiency-High Data Rate Personal Communication Wireless System. *Vehicular Technology Conference*, p. 1854–1858.
- Jang, J.-S., 1993. ANFIS: adaptive-network-based fuzzy inference system. *IEEE Systems, Man, and Cybernetics Society*, 23(3), pp. 665-685.
- Kawser, M. T. et al., 2012. Downlink SNR to CQI Mapping for Different Multiple Antenna Techniques in LTE. *International Journal of Information and Electronics Engineering*, September .Volume 2.
- Kelly, F., 1996. Notes on effective bandwidths. In *Stochastic Networks*. Oxford University Press.
- Khan, N., Martini, M. G. & Staehle, D., 2013. Opportinistic QoS-Aware Fair Downlink Sheduling for Delay Sensitive Applications using Fuzzy Reactive and Proactive Controllers. *Vehicular Technology Conference (VTC Fall), 2013 IEEE 78th*, 2-5 Setember, pp. 1-6.
- Lee, C. C., 1990. Fuzzy Logic in Control Systems: Fuzzy Logic Controller, Part II. *IEEE Transactions on Systems, Man, and Cybernetics*, 20(2), pp. 419-435.
- Riedi, R. H., Crouse, M. S., Ribeiro, V. J. & Baraniuk, R. G., 1999. A Multifractal Wavelet Model with Application to Network Traffic. *IEEE Transactions on Information Theory*, 45(3), pp. 992-1018.
- Rocha, F. G. C. & Vieira, F. H. T., 2009. Modelagem de tráfego de vídeo MPEG-4 utilizando cascata multifractal com distribuição autorregressiva dos multiplicadores. *I2TS*.
- Shreedhar, M. & Varghese, G., 1996. Efficient Fair Queuing Using Deficit Round-Robin. *IEEE/ACM Transactions on Networking*, 4(3), pp. 375-385.
- Spiegel, M. R. & Liu, J., 2011. *Manual de Fórmulas e Tabelas Matemáticas*. Coleção Schaum. 3 ed. s.l.:Makron Books.
- Sugeno, M. & Kang, G., 1988. Structure identification of fuzzy model. *Fuzzy Sets and Systems*, 28(1), pp. 15-33.
- Su, L. & Ping Wang, F. L., 2012. Particle Swarm Optimization Based Resource Block Allocation Algorithm for Downlink LTE Systems. *The 18th Asia-Pacific Conference on Communications*.
- Takagi, T. & Sugeno, M., 1985. Fuzzy identification of systems and its applications to modeling and control. *IEEE Systems, Man, and Cybernetics Society*, 15(1), pp. 116-132.
- Vieira, F. H. T., Bianchi, G. R., Ling, L. L. & Lemos, R. P., 2004. Estimação de banda efetiva dinâmica em redes de computadores utilizando uma modelagem auto-regressiva nebulosa. *XXI Simpósio Brasileiro de Telecomunicações (SBrT)*.
- Wang, J. & Yin, Z., 2008. A ranking selection-based particle swarm optimizer for engineering design optimization problems. *Structural and Multidisciplinary Optimization*, Volume 37, pp. 131-147.