

MODELAGEM FUZZY UTILIZANDO FUNÇÕES DE BASE ORTONORMAIS APLICADA À PREDIÇÃO ADAPTATIVA DE TRÁFEGO DE REDES

Flávio Henrique Teles Vieira e Lee Luan Ling

Departamento de Comunicações (DECOM)
Faculdade de Engenharia Elétrica e de Computação (FEEC)
Universidade Estadual de Campinas
Caixa Postal 6101 13.083-852, Campinas, SP, Brasil
Emails: {flavio;lee}@decom.fee.unicamp.br

Resumo- Neste artigo propomos um modelo *fuzzy* que faz uso de funções de base ortonormais calculadas através de informações providas da análise multifractal de séries temporais. Para obtenção das funções de base ortonormais, utilizamos a função de autocorrelação de processos multifractais e deduzimos uma expressão analítica para o pólo de Laguerre que constitui essas funções. Em seguida, nos focamos no desenvolvimento de um algoritmo adaptativo de treinamento para o modelo *fuzzy*-FBO (Funções de Base Ortonormais) que denominamos de ARFA (Agrupamento Regressivo Fuzzy Adaptativo). Avaliamos então o desempenho de predição de tráfego de redes do modelo *fuzzy*-FBO adaptativo com relação a outros preditores. Tendo como base a predição obtida com o modelo *fuzzy*-FBO treinado com o algoritmo ARFA, apresentamos um novo esquema de alocação de banda para tráfego de redes. Através de simulações mostramos que este esquema de alocação de banda se favorece do eficiente desempenho de predição do modelo *fuzzy* proposto. Comparações com outros esquemas de alocação de banda em termos de taxa de perda de bytes, utilização do enlace, frequência de sinalização e ocupação do *buffer* comprovam a eficiência do esquema proposto.

Palavras Chaves- Modelo *Fuzzy*, Funções de Base Ortonormais, Predição de Tráfego, Cálculo de Banda, Tráfego Multifractal.

1 Introdução

A modelagem *fuzzy* tem sido bastante empregada em vários campos de pesquisa desde que a teoria *fuzzy* foi inicialmente desenvolvida [1]. A razão para estas pesquisas é que os modelos *fuzzy* possuem certas vantagens com relação a determinados sistemas sobre modelos lineares, como por exemplo, na descrição de processos reais desconhecidos, com características não-lineares e variantes no tempo, como o tráfego de redes [2]. Entre os modelos *fuzzy* propostos, particularmente o modelo nebuloso descrito em [3] é capaz de descrever um sistema desconhecido usando um pequeno número de regras nebulosas como o faz o modelo de Takagi e Sugeno [4] e é de fácil de implementação como o modelo de Sugeno e Yasukawa [5]. Porém, estes trabalhos não apresentam algoritmos adaptativos de treinamento para os modelos *fuzzy*, ou seja, para modelagem de uma série temporal são necessários todos as amostras da série em questão. Algoritmos adaptativos são mais adequados para aplicações multimídia em tempo real nas redes atuais devido ao processamento mais rápido das informações [6]. Tendo em vista isso, apresentamos uma nova modelagem *fuzzy* com funções de base ortonormais (FBO) que através do algoritmo proposto ARFA (Agrupamento Regressivo Fuzzy Adaptativo) cria adaptativamente agrupamentos *fuzzy* à medida que dados de tráfego de entrada são disponibilizados, sendo capaz de prever eficientemente o tráfego de redes.

Na busca de uma descrição mais completa do tráfego de redes, modelos multifractais tem sido empregados [7] [8]. A modelagem multifractal generaliza os conceitos envolvidos com processos monofractais. Processos multifractais apresentam além de dependência de longo prazo, diferentes leis de escala, sendo dessa forma mais abrangentes [9]. A dependência de longo prazo, presente nas séries de tráfego e que tem importante impacto no desempenho de redes [10], pode ser constatada pelo decaimento lento da função de autocorrelação dessas séries temporais. Assim, apresentaremos neste artigo uma expressão analítica para a função de autocorrelação de fluxos de tráfego multifractal que será utilizada na modelagem *fuzzy* proposta.

Algumas propostas de alocação de banda para fluxos de tráfego baseadas em predição realizadas por redes neurais e *fuzzy* merecem atenção pois mostram que o controle preventivo de tráfego pode ser bastante eficiente, uma vez que o mesmo tenta reduzir o congestionamento antes que este aconteça e se adapta bem às variações bruscas do tráfego de redes [11][12].

Neste trabalho, apresentamos um novo esquema de alocação de banda onde inserimos o preditor *fuzzy* proposto o qual faz uso de funções de base ortonormais calculadas a partir da função de autocorrelação do processo de tráfego. O artigo está organizado da seguinte forma: na seção 2, descrevemos a modelagem *fuzzy* TSK (Takagi-Sugeno-Kang), fazendo conexão com as funções de base ortonormais de Laguerre. Na seção 3, introduzimos alguns conceitos da análise multifractal, obtemos a função de autocorrelação de um processo multifractal e ainda deduzimos uma equação para o pólo utilizado no cálculo das funções de base ortonormais. Na seção 4, desenvolvemos um algoritmo de treinamento adaptativo para o modelo *fuzzy*-FBO proposto. Mais especificamos na seção 4.2, comparamos o desempenho de predição do modelo *fuzzy*-FBO com outros existentes na literatura. Na seção 5, empregamos o preditor proposto em um novo esquema de alocação de banda, verificando

sua eficiência em relação a outros esquemas que também usam predição em sua composição. Por fim, na seção 6, apresentamos as conclusões obtidas.

2 Modelagem Fuzzy TSK

Modelos *fuzzy* têm sido eficientemente empregados na modelagem de sistemas uma vez que são aproximadores universais. Estes modelos podem ser divididos em 3 classes: Modelos linguísticos (Modelos de Mandani) [13], Modelos relacionais *fuzzy* [13] e Modelos do tipo Takagi-Sugeno [4][14]. Na modelagem *fuzzy*, as informações das séries temporais são divididas em *clusters* (agrupamentos), onde cada *cluster* é descrito por um modelo local. O modelo *fuzzy* TSK (Takagi-Sugeno-Kang) representa um determinado processo pela combinação de modelos locais AR (Auto-Regressivos) via regras *fuzzy* [4]. Tipicamente, um modelo *fuzzy* TSK consiste de regras *fuzzy* do tipo SE-ENTÃO que tem a seguinte forma:

$$\text{Regra } R^i : \begin{cases} \text{Se } x_1 \text{ é } A_1^i(\bar{\theta}_1^i) \text{ e } x_2 \text{ é } A_2^i(\bar{\theta}_2^i), \dots, x_n \text{ e' } A_n^i(\bar{\theta}_n^i) \\ \text{Então } h^i = f_i(x_1, x_2, \dots, x_n; \bar{a}^i) \end{cases} \quad (1)$$

para $i = 1, 2, \dots, C$, onde C é o número de regras, $A_j^i(\bar{\theta}_j^i)$ é o conjunto *fuzzy* da i -ésima regra para a entrada x_j com o conjunto de parâmetros $\bar{\theta}_j^i$ para $j = 1, 2, \dots, n$ e $\bar{a}^i = (a_0^i, a_1^i, \dots, a_n^i)$ é o conjunto de parâmetros das partes conseqüentes das regras R^i . As partes conseqüentes de cada regra R^i , ou seja, a saída de cada regra nebulosa é uma expressão do tipo $h^i = f_i(x_1, x_2, \dots, x_n; \bar{a}^i)$. Neste trabalho, assume-se que as funções de pertinência $A_j^i(\theta_{j1}^i, \theta_{j2}^i)$ são gaussianas dadas pela seguinte equação:

$$A_j^i(\theta_{j1}^i, \theta_{j2}^i) = \exp\left\{-\frac{(x_j - \theta_{j1}^i)^2}{2(\theta_{j2}^i)^2}\right\}, \quad (2)$$

logo θ_{j1}^i e θ_{j2}^i são parâmetros ajustáveis da função de pertinência j da regra *fuzzy* i . O emprego de funções gaussianas como funções de pertinência nos auxiliará na obtenção das equações de atualização dos parâmetros das partes conseqüentes, conforme será ressaltado na seção 4.1. A saída y do modelo *fuzzy* TSK é obtida através das regras *fuzzy* R^i pela seguinte equação:

$$y = \frac{\sum_{i=1}^C h^i w^i}{\sum_{i=1}^C w^i}, \quad (3)$$

onde h^i é a saída da i -ésima regra e $w^i = \prod_{j=1 \dots p} A_j^i(\theta_{j1}^i, \theta_{j2}^i)$ é o ativador da i -ésima regra. Nesta modelagem, os parâmetros das partes de premissa (isto é, $\bar{\theta}_j^i$) e das partes conseqüentes (\bar{a}^i) devem ser determinados.

2.1 Funções de Base Ortonormais em Modelagem Fuzzy

Nesta seção, abordaremos alguns conceitos envolvendo funções de base ortonormais e sua relação com a modelagem *fuzzy*.

Para um sistema não-linear com memória finita, podemos representar sua relação G de entrada-saída $(u(k), y(k))$ através de um modelo não-linear de média móvel (NLMA- *Non-Linear Moving Average*) como [15]:

$$y(k) = G(u(k-1), \dots, u(k-v)). \quad (4)$$

O sinal de entrada $u(k)$ do modelo (4) pode ser considerado como o desenvolvimento da seqüência $u(k)$ em uma base tal que a entrada j seja dada por:

$$u_j(k) = \Phi_{mag,i}(q^{-1})u(k) = q^{-j}u(k) = u(k-j), \quad (5)$$

onde q^{-j} é o operador de deslocamento com $\Phi_{mag,j}(q^{-1}) = q^{-j}$.

No modelo NLMA, a saída de um sistema é modelada com arbitrária precisão de acordo com o número de funções de

base. Muitas vezes é necessário um número elevado de funções para se atingir determinada precisão pelo modelo. Uma maneira de reduzir o número de funções de base, dado um determinado erro desejado, é inserindo conhecimento a priori da dinâmica do sistema no cálculo das funções de base [16] [17]. As funções de base ortonormais de Laguerre tem sido frequentemente utilizadas com esse intuito [15] [18].

A base de Laguerre é usada em vários contextos de identificação e controle de sistemas não-lineares [15] [18]. Neste trabalho, adotamos a base de Laguerre especialmente porque esta é completamente parametrizada por um único pólo, o pólo de Laguerre p . O conjunto de funções de transferência associadas a esta base é dada pela seguinte equação:

$$\Phi_{mag,j}(q^{-1}) = \sqrt{1-p^2} \frac{q^{-1}(q^{-1}-p)^{j-1}}{(1-pq^{-1})^j}, \quad j = 1, \dots, n \quad (6)$$

onde $p \in \{P : -1 < p < 1\}$ é o pólo das funções de Laguerre (base de Laguerre). Pode-se notar que fazendo $p = 0$ em (6), resulta na base $\Phi_{mag,j}(q^{-1}) = q^{-j}$. Portanto, a base $\Phi_{mag,j}(q^{-1}) = q^{-j}$ é um caso especial da base de Laguerre.

A saída do modelo expressa por (4) pode ser reescrita como:

$$y(k) = H(l_1(k), \dots, l_n(k)), \quad (7)$$

onde $l_j(k) = \Phi_{mag,j}(q^{-1})u(k)$ é a j -ésima função de Laguerre no instante de tempo k , n é o número de funções de base utilizadas e H é um operador não-linear. Note que a operação não-linear correspondente a H pode ser realizada através da modelagem *fuzzy*.

As funções de Laguerre $l_j(k)$ são recursivas e podem ser obtidas por equações de estado da seguinte forma [15]:

$$\mathbf{l}(k+1) = \mathbf{A}\mathbf{l}(k) + \mathbf{b}u(k) \quad (8)$$

$$y(k) = H(\mathbf{l}(k)) \quad (9)$$

onde $\mathbf{l}(k) = [l_1(k) \dots l_n(k)]^T$. A matriz \mathbf{A} e o vetor \mathbf{b} dependem da ordem do modelo n e do valor do pólo p como segue [16]:

$$\mathbf{A} = \begin{pmatrix} p & 0 & 0 & \dots & 0 \\ 1-p^2 & p & 0 & \dots & 0 \\ (-p)(1-p^2) & 1-p^2 & p & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ (-p)^{n-2}(1-p^2) & (-p)^{n-3}(1-p^2) & (-p)^{n-4}(1-p^2) & \dots & p \end{pmatrix} \quad (10)$$

$$\mathbf{b} = \sqrt{1-p^2} [1 \quad -p \quad (-p)^2 \quad \dots \quad (-p)^{n-1}]^T \quad (11)$$

O modelo não-linear representado pelas equações (8) e (9) consiste de um mapeamento linear entre a entrada $u(k)$ e as funções de Laguerre $l_j(k)$, mais o mapeamento entre $l_j(k)$ e a saída do sistema $y(k)$. Obtém-se dessa forma, um modelo mais preciso do que o modelo NLMA tradicional com o mesmo número de funções.

Dado um número de funções de Laguerre n (ordem do modelo), a estimação de um valor adequado para o pólo p da base ortonormal acarreta uma melhor representação do sistema a ser modelado. Objetivamos introduzir informações a respeito do comportamento multifractal dos processos de tráfego no cálculo do pólo p . Assim, abordaremos o comportamento multifractal de processos estocásticos na próxima seção.

3 Análise Multifractal

A descrição do comportamento local de medidas e funções em uma forma geométrica e estatística é de interesse da análise multifractal [19]. Na análise multifractal verifica-se o comportamento em escala de momentos estatísticos dos processos para estimar suas regularidades locais [20][21]. Através de ferramentas da análise multifractal algumas propriedades encontradas em processos reais podem ser verificadas. O tráfego de redes, por exemplo, ao ser considerado multifractal significa que possui uma estrutura de forte dependência inerente entre as amostras, com incidência de rajadas em várias escalas [21]. Estas

características podem degradar o desempenho de rede em relação a fluxos de tráfego Gaussianos e de curta-dependência [19]. Definimos processos multifractais a seguir.

Definição 1 Um processo estocástico $X(t)$ é multifractal se satisfaz a equação:

$$E(|X(t)|^q) = c(q)t^{\tau(q)+1} \quad (12)$$

onde $t \in T$ e $q \in Q$, T e Q são intervalos na reta real, $\tau(q)$ e $c(q)$ são funções com domínio Q . Normalmente, assume-se que T e Q tenham comprimentos positivos, e que $0 \in T$, $[0,1] \subseteq Q$. A Definição 1 descreve a ‘multifractalidade’ em termos de momentos estatísticos, onde $\tau(q)$ e $c(q)$ são conhecidos como a função de escala e o fator de momento de um processo multifractal, respectivamente.

Os traços de tráfego reais apresentam suas propriedades multifractais caracterizadas pela função de escala $\tau(q)$ e o fator de momento $c(q)$, conforme se pode observar pela equação (12). Assim, um modelo multifractal deve capturar estas duas propriedades multifractais. Isto pode ser obtido pelo produto de uma cascata multiplicativa $\mu(\Delta t_k)$ e uma variável aleatória i.i.d positiva Y em cada intervalo de tempo Δt_k [22]. A variável Y é independente da medida da cascata $\mu(\Delta t_k)$, então a série obtida denotada por $X(\Delta t_k)$ satisfaz a seguinte equação:

$$E(X(\Delta t_k)^q) = E(Y^q)E(\mu(\Delta t_k)^q) = E(Y^q)\Delta t_k^{\tau_0(q)}, \quad (13)$$

onde $\tau_0(q) = \tau(q) + 1$.

A partir disto, comparando (12) e (13) temos que as variáveis R e Y devem ser escolhidas de forma a atender as seguintes equações:

$$-\log_2(E(R^q)) = \tau_0(q) \quad (14)$$

$$E(Y^q) = c(q) \quad (15)$$

A função de escala pode ser precisamente modelada ao assumirmos que R é uma variável aleatória em $[0,1]$ com distribuição beta simétrica $Beta(\alpha, \alpha)$, $\alpha > 0$ [22]:

$$\tau_0(q) = \log_2 \frac{\Gamma(\alpha)\Gamma(2\alpha + q)}{\Gamma(2\alpha)\Gamma(\alpha + q)} \quad (16)$$

onde $\Gamma(\cdot)$ corresponde à função Gama.

Consideremos Y como sendo uma variável aleatória cuja distribuição seja lognormal e definida pelos parâmetros ρ e γ e momento $E(Y^q) = e^{\rho q + \gamma^2 q^2 / 2}$. Assim sendo, um processo multifractal pode ser descrito por 3 parâmetros (α, ρ, γ) , apresentando média μ e variância σ^2 dadas respectivamente por [22]:

$$\mu = e^{\rho + \gamma^2 / 2} \quad (17)$$

$$\sigma^2 = e^{2\rho + 2\gamma^2} \left(\frac{\alpha + 1}{\alpha + 1/2} \right)^N - e^{2\rho + \gamma^2} \quad (18)$$

Os parâmetros α , γ e ρ para um processo multifractal podem ser atualizados adaptativamente através dos valores das funções $\tau(q)$ e $c(q)$. Um método de estimação destas funções baseado na equação (12) é descrito em [23].

3.1 Função de Autocorrelação para Processos Multifractais

Pode-se constatar pela função de autocorrelação de um processo, a presença ou não de dependência de longo prazo entre as amostras (decaimento lento da função de autocorrelação). Além disso, a função de autocorrelação reflete a estatística de segunda ordem de uma série temporal. A partir das propriedades da modelagem multifractal apresentada na seção anterior podemos obter a função de autocorrelação de um processo multifractal de forma analítica.

Teorema 1 Seja o processo multifractal $X(n)$ discreto no tempo com parâmetros α , ρ e γ . A função de autocorrelação deste processo para os instantes de tempo n e k , é dada pela seguinte equação [22]:

$$E[X(n), X(n+k)] = e^{2\rho+\gamma^2} \left(\frac{\alpha(\alpha+1)^{N-1}}{(\alpha+1/2)^N} k^{-\log_2\left(\frac{\alpha+1}{\alpha+1/2}\right)} \right) \quad (19)$$

onde $N = \log_2(N_a)$ e N_a é a quantidade de amostras do processo.

Prova: Ver Apêndice A

3.2 Pólo do Modelo Obtido a Partir da Função de Autocorrelação para Processos Multifractais

Nesta seção, introduzimos uma expressão analítica para o pólo p utilizado no cálculo das funções de base ortonormais. Propomos a aplicação da modelagem *fuzzy* na implementação do operador não-linear H em (9) que em conjunto com as funções de base de Laguerre constituem a modelagem *fuzzy*-FBO (Funções de Base Ortonormais) proposta.

Utilizamos a função de autocorrelação dada por (19) para se obter o pólo dominante do sistema a partir dos resultados estabelecidos por Levinson e Durbin em seu algoritmo de determinação de pólos [24]. As entradas para o algoritmo de Levinson-Durbin são os valores atualizados da função de autocorrelação e pode-se obter como saída um modelo AR (Auto-Regressivo) de um pólo. Assim, enunciamos a seguinte proposição:

Proposição 1 O pólo p utilizado para o cálculo das funções de base ortonormais para o modelo *fuzzy*-FBO pode ser dado por:

$$p = -\frac{1}{2 \log_2\left(\frac{\alpha+1}{\alpha+1/2}\right)} \quad (20)$$

Prova: Ver Apêndice B

Note que uma vez que o pólo p é dado em função do parâmetro multifractal α , podemos estimá-lo adaptativamente. A Figura 1 mostra o valor do pólo p em função do tempo para um traço de tráfego Internet real.

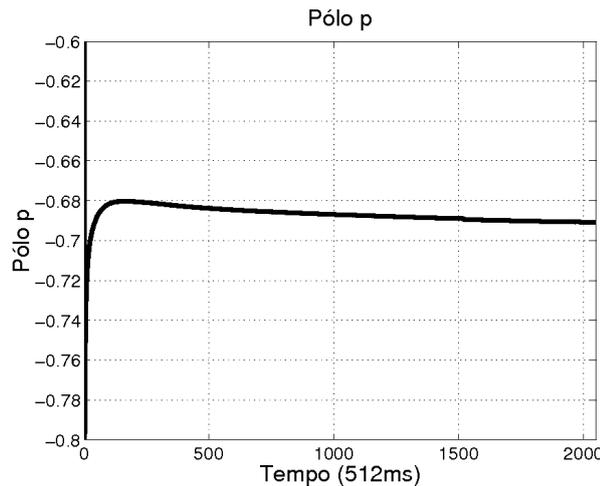


Figura 1. Pólo para o Traço de Tráfego Internet dec-pkt-1

4 Modelo Fuzzy-FBO

A estimação do valor do pólo permite que estendamos a interpolação *fuzzy* de modelos locais, que é a idéia central do modelo TSK, a um contexto de funções de base ortonormais. Nossa proposta se baseia em uma versão em espaço de estados do modelo TSK, ou seja, cada regra R^i do modelo *fuzzy* representa um modelo de espaço de estados diferente:

$$R^i : \text{Se } l_1(k) \text{ e } l_2(k) \text{ e } \dots \text{ e } l_n(k) \text{ e } L_n^i$$

$$\text{Então} \begin{cases} \mathbf{I}_i(k+1) = A_i \mathbf{I}_i(k) + b_i \bar{x}(k) \\ y_i(k) = H_i(\mathbf{I}_i(k)) \end{cases} \quad (21)$$

onde a matriz A_i e o vetor b_i dependem do pólo $p_i(k)$ e $H_i(\mathbf{I}_i(k))$ é o mapeamento que relaciona a saída $y_i(k)$ do modelo local i a seu correspondente estado de Laguerre (funções de base ortonormais) $\mathbf{I}_i(k) = [l_1(k) \ l_2(k) \ \dots \ l_n(k)]$, sendo $\bar{x}(k) = [x_1(k) \ x_2(k) \ \dots \ x_n(k)]$ o vetor de entrada e L_j^i a função de pertinência *fuzzy* para a regra i associada a j -ésima variável de premissa. As variáveis de premissa são os estados de Laguerre do modelo *fuzzy* TSK-FBO resultante. A saída do modelo *fuzzy*-FBO é dada por:

$$y(k) = \frac{\sum_{i=1}^C y_i(k) w_i(\mathbf{I}_i(k))}{\sum_{i=1}^C w_i(\mathbf{I}_i(k))}, \quad (22)$$

onde C é o número de regras (modelos locais) e os pesos $w_i(\mathbf{I}_i(k))$ da regra i são dados por:

$$w_i(\mathbf{I}_i(k)) = \prod_{j=1}^n L_j^i(l_j(k)). \quad (23)$$

Um caso particular é quando se tem pólos iguais para os C modelos locais ($p_1(k) = \dots = p_C(k)$). Com esta condição temos $A = A_i$ e $b = b_i$ para $i = 1, \dots, C$, ou seja, o modelo TSK-FBO pode ser representado conforme as equações (8) e (9) e H dado de acordo com as equações (23) e (22). Esta hipótese é equivalente a se dizer que os modos dominantes do sistema não mudam significativamente em diferentes regiões de operação. No caso dessa hipótese não ser verdadeira, a única consequência é que um maior número de funções pode ser requerido para prover uma modelagem com uma determinada precisão desejada. Entretanto, este caso particular facilita o cálculo de um valor adequado para o pólo de Laguerre p de modo a diminuir o erro de modelagem.

Pode-se verificar que o modelo TSK é capaz de aproximar adequadamente sistemas discretos, causais, invariantes no tempo, estáveis, com entrada e saídas limitadas e que não possuam descontinuidades [4]. Mostraremos posteriormente que, em adição a estas características, o modelo TSK-FBO proposto representa de maneira eficiente processos variantes no tempo (por exemplo, tráfego de rede) com o uso do algoritmo de treinamento apresentado na próxima seção.

4.1 Algoritmo de Treinamento Adaptativo para o Modelo Fuzzy-FBO

Nessa seção, propomos um algoritmo de agrupamento regressivo *fuzzy* adaptativo (ARFA) para o modelo preditivo *fuzzy*-FBO. O algoritmo ARFA possui dois estágios: o primeiro consiste de um estágio não-supervisionado para localização dos centros dos agrupamentos (*clusters*), e no segundo aplica-se um algoritmo baseado no conceito de gradiente descendente para ajustar os parâmetros envolvidos no primeiro estágio.

No algoritmo ARFA, levamos em conta a distribuição espacial dos dados considerando o erro de regressão e a distância entre os dados de entrada e os *clusters*. Seja a função custo do algoritmo ARFA definida como:

$$J = \sum_{i=1}^C \sum_{k=1}^N u_{ik}^2 (r_{ik} d_{ik})^2 \quad (24)$$

sujeito a

$$\sum_{i=1}^C u_{ik} = 1, \quad \text{para } 1 \leq k \leq N \quad (25)$$

onde u_{ik} é o grau de ativação da i -ésima regra para o k -ésimo padrão de treinamento, C é o número de regras *fuzzy* e N é o número total de dados de treinamento. Na equação (24), r_{ik} é o erro entre a k -ésima saída desejada $y(k)$ do sistema modelado e a saída da i -ésima regra $f_i(\bar{x}(k); \bar{a}^i(k))$ com a k -ésima entrada, isto é,

$$r_{ik} = y(k) - f_i(\bar{x}(k); \bar{a}^i(k)), \quad (26)$$

com $i = 1, 2, \dots, C$ e $k = 1, 2, \dots, N$. Na mesma equação (24), d_{ik} é a distância entre o vetor de entrada $\bar{x}(k)$ no instante discreto k e o centro do i -ésimo cluster β_i , ou seja,

$$d_{ik} = \bar{x}(k) - \beta_i. \quad (27)$$

Para minimizar a função custo J em (24), o método dos multiplicadores de Lagrange pode ser aplicado. Para tal, escrevemos a função de Lagrange como [25]:

$$L = \sum_{i=1}^C \sum_{k=1}^N u_{ik}^2 (r_{ik} d_{ik})^2 - \sum_{k=1}^N \lambda_k \left(\sum_{i=1}^C u_{ik} - 1 \right) \quad (28)$$

As condições para minimização desta função fornecem as bases para o desenvolvimento do nosso algoritmo de treinamento. As condições necessárias para que a função custo J seja mínima são:

$$\frac{\partial L}{\partial \bar{a}^i(k)} = \sum_{k=1}^N (u_{ik})^2 (d_{ik})^2 \frac{\partial r_{ik}^2}{\partial \bar{a}^i(k)} = 0 \quad (29)$$

$$\frac{\partial L}{\partial u_{ik}} = 2u_{ik} (r_{ik} d_{ik})^2 - \lambda_k = 0 \quad (30)$$

$$\frac{\partial L}{\partial \beta_i} = \sum_{k=1}^N (u_{ik})^2 (r_{ik})^2 \frac{\partial d_{ik}^2}{\partial \beta_i} = 0 \quad (31)$$

Tomando-se a derivada parcial de r_{ik} em (26) e substituindo em (29), tem-se:

$$\sum_{k=1}^N (u_{ik})^2 (d_{ik})^2 \frac{\partial r_{ik}}{\partial \bar{a}^i(k)} y(k) - \sum_{k=1}^N (u_{ik})^2 (d_{ik})^2 \frac{\partial r_{ik}}{\partial \bar{a}^i(k)} f_i(\bar{x}(k); \bar{a}^i(k)) = 0, \quad (32)$$

$$\frac{\partial r_{ik}}{\partial \bar{a}^i(k)} = \bar{x}(k) \quad (33)$$

Seja $X \in \mathbb{P}^{N \times (n+1)}$ uma matriz onde seus elementos são os valores de $\bar{x}(k)$ em sua j -ésima coluna $j = 1, \dots, n+1$ (a primeira coluna de X é toda composta por 1), $Y \in \mathbb{P}^N$ um vetor onde o k -ésimo elemento é o valor de $y(k)$ e $Q_i \in \mathbb{R}^{N \times N}$ uma matriz diagonal onde a k -ésima diagonal é dada pelo termo $q(k) = u_{ik}^2 d_{ik}^2$. Assim, podemos reescrever a equação (32) na seguinte forma matricial:

$$X^T Q_i(k) Y - (X^T Q_i(k) X) \bar{a}^i = 0. \quad (34)$$

Resolvendo a equação (34) e usando a seguinte notação para a matriz de covariância $P_i(k)$:

$$P_i(k) = (X^T Q_i(k) X)^{-1}, \quad i = 1, 2, \dots, C, \quad (35)$$

resulta que o vetor de parâmetros $\bar{a}^i(k)$ das partes conseqüentes da i -ésima regra para o instante k é obtido pela seguinte equação:

$$\bar{a}^i(k) = P_i(k) X^T(k) Q_i(k) Y(k). \quad (36)$$

Tem-se então um problema de inversão matricial na equação (35) para o cálculo do vetor de parâmetros das partes conseqüentes em um instante de tempo k . Como proposta, tomou-se como solução para este problema a aplicação de um algoritmo recursivo de estimação dos parâmetros representado pelas seguintes equações [26]:

$$\bar{a}^i(k+1) = \bar{a}^i(k) + (P_i(k+1) x(k+1) q(k+1)) \times (y(k+1) - x(k+1)^T \bar{a}^i(k)) \quad (37)$$

e,

$$P(k+1) = P(k) - \frac{q(k+1)P_i(k)x(k+1)x(k+1)^T P_i(k)}{1 + q(k+1)x(k+1)^T P_i(k)x(k+1)}, \quad (38)$$

onde $x(k+1)$ é a $(k+1)$ -ésima linha da matriz X e $q(k+1)$ é o $(k+1)$ -ésimo elemento da matriz diagonal $Q_i(k+1)$.

A partir da segunda condição (equação (30)) de minimização da função J e da condição expressa pela equação (25), obtém-se uma equação para o grau de ativação u_{ik} da i -ésima regra que não depende dos multiplicadores de Lagrange λ_k :

$$u_{ik} = \frac{1/2(r_{ik}^2 d_{ik}^2)}{\sum_{i=1}^C 1/2(r_{ik}^2 d_{ik}^2)}. \quad (39)$$

Resolvendo a equação referente à última condição de minimização (equação (31)), o centro do i -ésimo cluster (β_i) pode ser calculado pela seguinte equação:

$$\beta_i = \frac{\sum_{z=1}^N r_{iz}^2 u_{iz}^2 \bar{x}(z)}{\sum_{z=1}^N r_{iz}^2 u_{iz}^2}. \quad (40)$$

Para completarmos a primeira fase do algoritmo de treinamento ARFA aplicado à modelagem *fuzzy*-FBO, devemos obter os parâmetros θ_{j1}^i e θ_{j2}^i das funções de pertinência gaussianas $A_j^i(\theta_{j1}^i; \theta_{j2}^i)$, isto é:

$$A_j^i(\theta_{j1}^i; \theta_{j2}^i) = \exp\left\{-\frac{(x_j(k) - \theta_{j1}^i)^2}{2(\theta_{j2}^i)^2}\right\}. \quad (41)$$

Os parâmetros θ_{j1}^i e θ_{j2}^i caracterizam a j -ésima função de pertinência da i -ésima regra *fuzzy*, onde $1 \leq j \leq n$ e $1 \leq i \leq C$. Assumindo que as funções de pertinência sejam gaussianas, tais parâmetros podem ser obtidos a partir de u_{ik} e dos elementos $x_j(k)$ do vetor de entrada através das seguintes equações [27]:

$$\theta_{j1}^i = \frac{\sum_{z=1}^N (u_{iz})^2 x_j(z)}{\sum_{z=1}^N (u_{iz})^2} \quad (42)$$

$$\theta_{j2}^i = \sqrt{\frac{\sum_{z=1}^N (u_{iz})^2 (x_j(z) - \theta_{j1}^i)^2}{\sum_{z=1}^N (u_{iz})^2}} \quad (43)$$

Em resumo, na primeira parte do algoritmo ARFA, são encontrados (ou ajustados) o grau de ativação u_{ik} através da equação (39), a melhor posição para o centro dos *clusters* por meio da equação (40), o vetor de parâmetros $\bar{a}^i(k)$ das partes consequentes e o formato mais adequado para as funções de pertinência gaussianas que compõem cada regra *fuzzy* do modelo. Na segunda parte do treinamento, os parâmetros das partes consequentes (\bar{a}^i) e os subespaços *fuzzy* das partes de premissa ($\bar{\theta}_j^i$) são ajustados por um algoritmo de aprendizagem supervisionado para melhorar a precisão da modelagem. Aplicando-se o algoritmo de gradiente descendente [27] para o modelo TSK-FBO cuja saída é dada pela equação (22), obtemos a seguinte equação de atualização para os parâmetros das partes de premissa [28]:

$$\Delta \theta_{jl}^i(k) = \eta (y(k) - \hat{y}(k)) (y^i(k) - \hat{y}(k)) \frac{1}{\sum_{i=1}^c w^i(k)} \frac{\partial w^i(k)}{\partial \theta_{jl}^i(k)}, \quad l = 1, 2 \quad (44)$$

onde η é uma constante de aprendizagem, $y(k)$ é a saída desejada, $\hat{y}(k)$ é a saída do modelo TSK-FBO e $y^i(k)$ é a saída da i -ésima regra do modelo.

De forma análoga, os parâmetros do vetor das partes conseqüentes $\vec{a}^i(k)$ são atualizados com

$$\Delta a_j^i(k) = \zeta (y(k) - \hat{y}(k)) \frac{w^i(k) l_j(k)}{\sum_{i=1}^c w^i(k)}, \quad (45)$$

onde ζ é uma constante de aprendizagem que controla a taxa de atualização dos valores dos parâmetros das partes conseqüentes.

Para exemplificar o funcionamento do algoritmo, a Figura 2 apresenta os agrupamentos formados para uma série de tráfego real utilizando o modelo *fuzzy*-FBO com 2 regras e 1 entrada, e tendo como saída desejada o valor da série a um instante de tempo a frente.

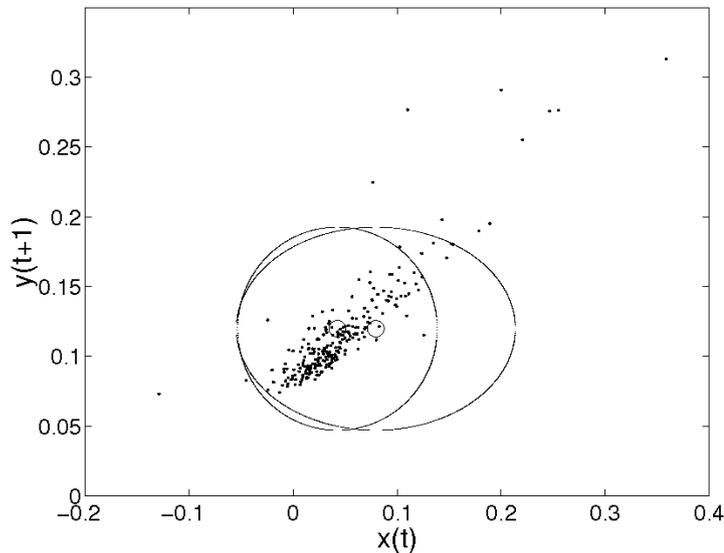


Figura 2. Clusters formados (2 regras e 1 entrada): traço de tráfego Bc-Octext

4.2 Avaliação de Desempenho de Predição do Modelo Fuzzy-FBO

Utilizamos nas simulações, traços de tráfego TCP/IP (*Transmission Control Protocol/ Internet Protocol*) obtidos da Digital Equipment Corporation (DEC)¹, traços de tráfego Ethernet obtidos da Bellcore² e traços capturados entre os anos de 2000 e 2002 na rede Petrobrás através de um analisador de dados DA350 da ActernaTM, com uma resolução de 32 microsegundos [29].

Na presente seção, apresentamos avaliações comparativas entre o desempenho do preditor proposto e o desempenho de outros três diferentes preditores, quando aplicados a traços de tráfego TCP/IP e Ethernet. Os outros preditores levados em consideração foram: o LMS (*Least Mean Square*) [30] [31], o RLS (*Recursive Least Square*) [31] e o Fuzzy FCRM (*Fuzzy Clustering Regression Model*) [3][32]. Utilizamos inicialmente o conceito de o erro quadrático médio (EQM) na análise de desempenho de predição. Avaliamos o preditor proposto utilizando-se duas medidas relativas de erro. Conhecidos como erros quadráticos médios normalizados (EQMN), a primeira medida consiste em normalizar o EQM em relação à variância da série

¹<http://ita.ee.lbl.gov/html/contrib/DEC-PKT.html>

²<http://ita.ee.lbl.gov/html/contrib/BC.html>

predita, enquanto a segunda medida consiste em normalizar o EQM em relação ao erro quadrático médio do preditor ótimo para o processo passeio aleatório (*random walk*) [33]. As definições dos EQMN's são dadas a seguir.

Definição 2 - Seja σ_x^2 a variância do processo X , dada por $\sigma_x^2 = E[(\mu - x)^2]$ onde μ é a média do processo, define-se o erro quadrático médio normalizado do tipo 1 como:

$$EQMN1 = \frac{EQM}{\sigma_x^2} = \frac{E[(\hat{x} - x)^2]}{E[(\mu - x)^2]} \quad (46)$$

Definição 3 - Seja \hat{x}_{pa} o valor predito da amostra do processo X , cujo valor é o mesmo da amostra imediatamente anterior do processo. Define-se o erro quadrático médio normalizado do tipo 2 como:

$$EQMN2 = \frac{E[(\hat{x} - x)^2]}{E[(\hat{x}_{pa} - x)^2]} \quad (47)$$

De acordo com as definições anteriores, um preditor que apresente um valor de EQMN1 igual ou inferior a 1 possuirá desempenho igual ou superior a um preditor que apenas estime o valor futuro como sendo igual a média do processo. Para um EQMN2 próximo de 1, teremos que o preditor analisado apresentará desempenho próximo ao de um preditor que estime o valor futuro como sendo igual ao valor da amostra imediatamente anterior.

Antes de iniciarmos as comparações com outros preditores, apresentaremos alguns resultados de predição do modelo preditor *fuzzy* proposto para a série da Bellcore Bc-octint. A Figura 3.(a) compara os valores preditos e os valores reais através de um gráfico quantil-quantil conhecido como QQplot [34]. Pode-se notar uma relação linear entre os valores preditos e os valores reais, o que denota desempenho de predição adequado. Isso pode ser constatado pela proximidade entre os valores preditos a um passo a frente e valores reais mostrados pelo gráfico da Figura 3.(b).

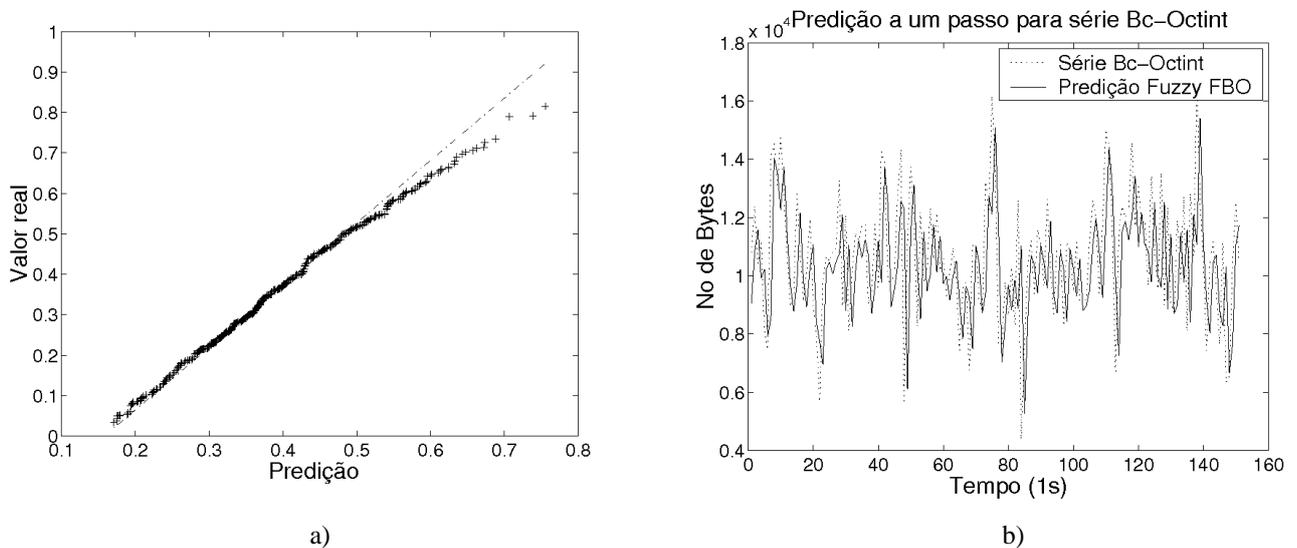
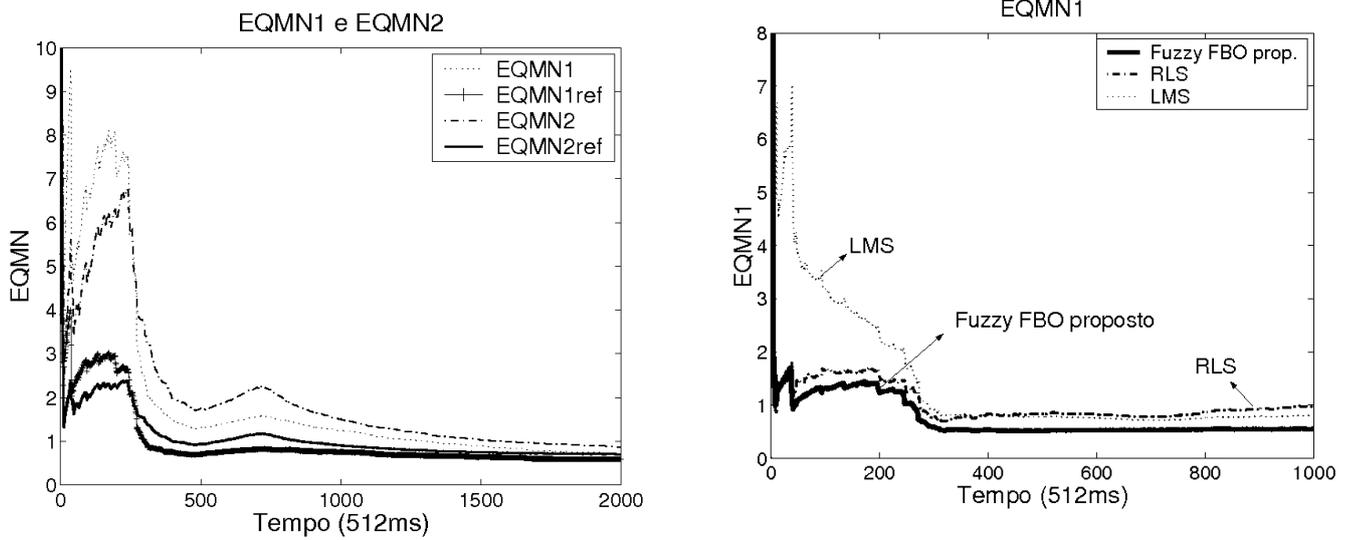


Figura 3. QQ-plot e predição para a série Bc-octint usando o algoritmo *fuzzy*-FBO proposto



a) Comparação entre EQMN1 e EQMN2 antes e depois do refinamento (curvas de aprendizagem)

b) Comparação entre EQMN1

Figura 4. Comparação entre erros quadráticos médios normalizados para o traço de tráfego dec-pkt-2

Tabela 1. Comparação entre EQMN1

Traço de tráfego	Intervalo	Fuzzy-FBO Adapt	Fuzzy Adapt	RLS	LMS	Fuzzy FCRM
Dec-pkt-1	1-2048	0,6564	0,7366	0,8513	0,9304	0,6987
Dec-pkt-2	1-2048	0,5758	0,6704	0,7022	0,7614	0,5836
Bc-Octint	801-1701	0,4102	0,4654	0,4114	0,4817	0,3107
Bc-Octext	1000-2000	0,4144	0,4355	0,4298	0,5010	0,4654

Tabela 2. EQMN1: Aumento do número de regras=número de coeficientes

Número de regras	Fuzzy FBO	Fuzzy Adapt.	RLS	LMS
2	0,6470	0,7973	1,0387	1,1977
3	0,6203	0,7879	0,9587	1,0908
4	0,6050	0,7852	0,9492	1,0718
5	0,6088	0,7945	0,9460	1,0539

Tabela 3. EQMN2: Aumento do número de regras=número de coeficientes

Número de regras	Fuzzy FBO	Fuzzy Adapt.	RLS	LMS
2	0,6121	0,7208	0,9553	1,0779
3	0,5762	0,7114	0,8820	0,9779
4	0,5633	0,7043	0,8739	0,9531
5	0,5846	0,7282	0,8731	0,9378

A segunda fase do treinamento ARFA proporciona reduções dos EQMN1 e EQMN2 como pode ser visto pela Figura 4.(a), uma vez que na primeira fase do treinamento ARFA se obtém um modelo *fuzzy* aproximado. Para realizar a comparação entre os preditores analisados, encontrou-se a configuração para cada preditor de forma a minimizar os EQMN1 e 2. Ou seja, foram escolhidos adequadamente a taxa de aprendizagem, os valores iniciais para os centros, etc. Os resultados apresentados na Tabela 1 apontam que o preditor adaptativo proposto supera em termos de erro de predição os algoritmos adaptativos lineares LMS e RLS para as séries de tráfego consideradas.

Um dos objetivos do treinamento adaptativo é o ajuste do algoritmo ao ambiente dinâmico do tráfego de redes. A fim de realizarmos uma comparação com treinamento do tipo 'on batch', inserimos também na Tabela 1 os valores de EQMN1 obtidos com o preditor *fuzzy* TSK treinado com o algoritmo não-adaptativo FCRM [32]. O preditor *fuzzy* FCRM se utiliza de todas as amostras da série de tráfego para o cálculo de seus parâmetros. Uma vez determinados estes parâmetros, aplica-se o modelo obtido na predição a um passo da série em questão. Pode-se observar que o preditor *fuzzy*-FBO obteve em geral menor EQMN1 para as séries analisadas. Portanto, os resultados comprovam que se pode ter com o conhecimento de poucas amostras do passado (neste teste, utilizou-se 2 amostras), um erro tão baixo quanto ao processamento com todas as amostras da série temporal. Este fato mostra que é possível a implementação de um algoritmo mais rápido, que necessite de pouca capacidade de armazenamento e com convergência acelerada, como é o caso do preditor *fuzzy*-FBO proposto. Além disso, este preditor adaptativo captura mais adequadamente as características do processo de tráfego por não supor de antemão que a 'estrutura' do processo seja invariante, como é feito por alguns modelos com treinamento 'offline'.

Em teoria, à medida que se aumenta o número de funções de base para o preditor *fuzzy*-FBO, se obtém um EQMN de predição menor para determinada série de tráfego. No entanto, o que se observou para todos os preditores testados é que os EQMN1 e EQMN2 diminuem até um certo número de coeficientes, depois disso, nem sempre é possível obter um treinamento eficiente para os modelos preditores. As Tabelas 2 e 3 corroboram esta afirmação, onde para efeito de simplificação, adotamos o número de regras como sendo igual ao número de coeficientes (igual ao número de amostras passadas) dos modelos para predição da série dec-pkt-1. Com 5 regras, o preditor *fuzzy*-FBO começa a apresentar uma deterioração dos EQMN1 e EQMN2. Para os algoritmos RLS e LMS, o mesmo ocorre com um número de coeficientes igual a 7. Note entretanto que, mesmo esses dois algoritmos estando em sua melhor configuração não propiciaram EQMN menor do que o modelo *fuzzy*-FBO com 2 regras e 2 coeficientes.

Ainda com relação ao preditor *fuzzy* proposto é importante avaliar o desempenho de predição dos algoritmos para horizontes maiores de predição. Com esse fim, analisou-se os erros de predição com o teste T [35]. O teste T é um teste de hipótese que pode ser usado para se determinar se uma afirmação sobre uma característica de uma série é verdadeira. Este teste provê a probabilidade do grau de veracidade desta afirmação através da variável conhecida na literatura como p [35]. O valor de p corresponde a probabilidade de se observar determinado resultado dado que a hipótese nula é verdadeira. O teste T foi conduzido nas seqüências de erro produzidas pelos algoritmos de predição adaptativos, observando o valor de p à medida que o passo de predição é aumentado (Figura 5.(a)). A hipótese nula H_0 é de que a média do erro μ seja igual a zero ($\mu = 0$) e a hipótese alternativa $H_1: \mu \neq 0$. O nível de significância de um teste estatístico é a probabilidade de rejeição de uma hipótese verdadeira. Fixamos o nível de significância do teste realizado em 0,05, que corresponde a um intervalo de confiança de 0,95. A Figura 5.(a) apresenta os valores de p dos testes T aplicados às seqüências de erros para diferentes passos de predição para a série dec-pkt-2. Neste caso, fixamos o número de coeficientes dos modelos locais AR como sendo 2 para o modelo *fuzzy* proposto (com 2 regras nebulosas), assim como para os coeficientes do filtro preditor RLS e LMS. Em todos os passos analisados, foram obtidos $h = 0$ para o modelo *fuzzy*-FBO, ou seja, não se deve rejeitar a hipótese nula com um nível de significância 0,05. Apresentamos também na Tabela 4 os valores de p dos testes T aplicados às seqüências de erros de predição a um passo para vários traços de tráfego de redes em comparação a outros preditores adaptativos. Da mesma forma, encontramos o valor de p para a configuração (taxa de aprendizagem, parâmetros iniciais, etc) do preditor que resulte em menor EQMN.

Note que quanto maior é o valor de p para o preditor *fuzzy*-FBO, maior é o grau de segurança em afirmar que as

médias dos erros de predição sejam zero ao se usar esse preditor. De fato, além de valores de p geralmente mais altos, também observamos que os valores de EQMN1 do preditor *fuzzy-FBO* se manteve abaixo dos demais preditores.

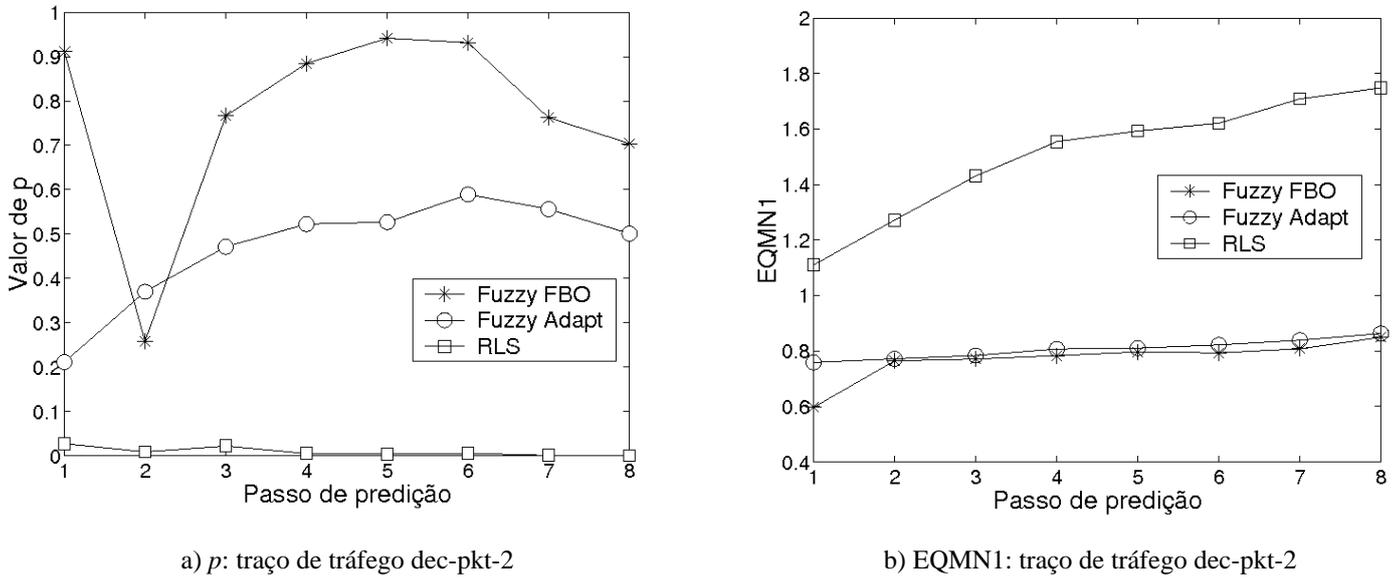


Figura 5. Probabilidade de ocorrência de hipótese nula p e EQMN1 versus passo de predição

Tabela 4. Valor de p

Traço de tráfego	Fuzzy-FBO Adapt	Fuzzy Adapt	RLS	LMS
Dec-pkt-1	0,5204	0,2681	0,0081	$5,0752.10^{-5}$
Dec-pkt-2	0,9109	0,2113	0,0274	$1,08.10^{-4}$
Bc-Octint	0,8103	0,7926	0,0362	$3,0357.10^{-5}$
Bc-Octext	0,4771	0,0256	0,0066	$3,0800.10^{-6}$
lbl-pkt-5	0,9071	0,8713	0,1089	$1,8130.10^{-5}$
10-7-S-1	0,0241	0,0193	0,0211	$2,5657.10^{-9}$

5 Estimativa Adaptativa de Banda Utilizando o Preditor Fuzzy Proposto

Os esquemas de alocação de banda para fluxos de tráfego de redes podem se beneficiar de algoritmos de predição da taxa de tráfego de modo a antecipar as ações para alocação de recursos e controle de congestionamento [36]. Neste sentido, o passo de predição pode ser ajustado de forma a possibilitar que a rede tenha tempo para obtenção e alocação dos recursos necessários. Precisamente nesta seção, propomos um esquema de alocação dinâmica de banda que considera como entrada as predições realizadas pelo modelo *fuzzy* proposto.

Seja $A(\tau, t)$ um processo a tempo discreto correspondente ao tráfego acumulado (neste caso, número de *bytes*) no intervalo de tempo (τ, t) e que chega ao servidor para ser transmitido. Para satisfazer um limite de retardo d_{req} , qualquer pacote (ou dados) deve ser transmitido até o instante de tempo ' $t + d_{req}$ '. A banda ε , ou seja, a taxa necessária para atender a esse critério de retardo e para se ter perda nula de dados de tráfego, deve obedecer a seguinte relação:

$$\varepsilon^*(t - \tau + d_{req}) \geq A(\tau, t) + b_{t-1} \quad \forall \quad t \geq \tau, \quad (48)$$

onde b_{t-1} corresponde ao número de *bytes* não enviados pela rede no instante anterior. A relação acima deve ser satisfeita

para todo $t \geq 0$ e $\tau \geq 0$ onde $t \geq \tau$.

Um meio de se obter um algoritmo eficiente de estimação de banda é estimar o tráfego $A(\tau, t)$ através de seu valor predito $A_p(\tau, t)$, ou seja, prever a taxa com a qual o fluxo será injetado no enlace de modo que este respeite o limite de retardo estipulado.

O procedimento que propomos para mapear o valor predito da intensidade de tráfego com a banda requerida é o seguinte: Seja ΔBW a quantia de banda finita e C a máxima banda disponível. Determina-se o intervalo de banda $[(k)\Delta BW, (k+1)\Delta BW]$ ($k \geq 0$), no qual o valor predito da taxa P_{t+1} se encontra e usa-se o valor superior do intervalo $(k+1)\Delta BW$ ou a banda máxima disponível C no caso de $C < (k+1)\Delta BW$ como banda requerida no instante $t+1$. Ou seja, a banda BW_{t+1} no instante $t+1$ de acordo com a equação (48) é dada por

$$BW_{t+1} = \min \left\{ \left\langle \frac{(P_{t+1} + b_t)/(t + d_{req})}{\Delta BW} \right\rangle \Delta BW, C \right\}, \quad (49)$$

onde o operador $\langle x \rangle$ representa o maior número inteiro mais próximo de x .

Como a banda requerida (49) é o valor superior do intervalo onde se encontra o valor predito da taxa, tem-se uma provisão de banda menos sensível às pequenas variações das predições, dentro dos intervalos de tamanho ΔBW . Ao se aumentar o intervalo de banda, evita-se uma freqüente realocação de banda, reduzindo o custo de sinalização envolvido. Portanto, o valor de ΔBW influencia diretamente o custo de sinalização do esquema de alocação de banda.

5.1 Avaliação de Desempenho do Esquema Proposto de Alocação de Banda

Esta seção apresenta as medidas de desempenho utilizadas para avaliar o esquema de provisão de banda proposto, assim como os resultados de desempenho obtidos para o mesmo. Seja z_t o valor observado de taxa de tráfego e BW_t a banda fornecida no instante de tempo t . As seguintes medidas de desempenho podem ser definidas:

- **Utilização média (u):** A utilização média mede a fração de banda usada para servir o fluxo de dados observados no período de tempo T , calculada por:

$$u = \frac{1}{T} \sum_{t=1}^T \min \left\{ \frac{z_t}{BW_t}, 1 \right\}. \quad (50)$$

- **Taxa de perda (TP):** A taxa de perda mede a quantidade de *bytes* perdidos devido à alocação de banda menor do que a necessária, dada por

$$TP = \frac{1}{T} \sum_{t=1}^T \max \left\{ \frac{(z_t - BW_t) - B}{z_t}, 0 \right\}. \quad (51)$$

Inicialmente avaliaremos o esquema de alocação de banda para o caso em que o *buffer* $B = 0$.

- **Freqüência de sinalização:** A freqüência de sinalização nos ajuda a avaliar quão freqüente o esquema proposto realoca banda. A realocação de banda envolve um custo de sinalização para as redes.

Comparou-se o esquema de provisão de banda proposto com dois outros esquemas, entre eles, o esquema apresentado por Adas et al. [37], que usa o algoritmo adaptativo LMS [30] para a predição de tráfego de redes e estas predições são usadas como taxas exigidas pelos fluxos. Outro trabalho comparado [38], utiliza o algoritmo RLS que possui convergência mais rápida [31]. A fim de se estudar os efeitos do tamanho do intervalo de banda sobre todos os esquemas em questão, a equação (49) é aplicada nos 3 esquemas em comparação, onde a diferença entre eles consiste nos valores de P_{t+1} dados pelos algoritmos de predição *fuzzy-FBO*, RLS e LMS, respectivamente.

Os resultados de desempenho: taxa de perda e utilização para os 3 esquemas de alocação de banda aplicados à série de tráfego 10-7-S-1 da rede da Petrobrás são mostrados na Figura 6. Pode-se observar que o esquema de alocação de banda proposto atinge uma taxa de perda menor a uma freqüência de sinalização semelhante aos demais (Figura 7), porém a custo de uma utilização do enlace ligeiramente menor. Espera-se que isso ocorra quando é maior a precisão do algoritmo de predição utilizado. Observamos que a aplicação do algoritmo de predição proposto na alocação de banda resulta em menores taxas de perda (Figura 8.(a)) e uma ocupação do *buffer* mais bem comportada, indicada pelo seu tamanho médio da fila no *buffer* (Figura 8.(b)). Devido a esses resultados, pode-se concluir que a probabilidade de perda para determinado tamanho de *buffer* é

menor para o esquema de alocação de banda proposto. Ao se plotar a função de distribuição acumulada complementar (Figura 8.(c)) que corresponde à probabilidade do tamanho da fila Q ser maior do que um certo valor de *buffer* x do processo de tráfego no *buffer*, visualiza-se claramente a afirmação anterior.

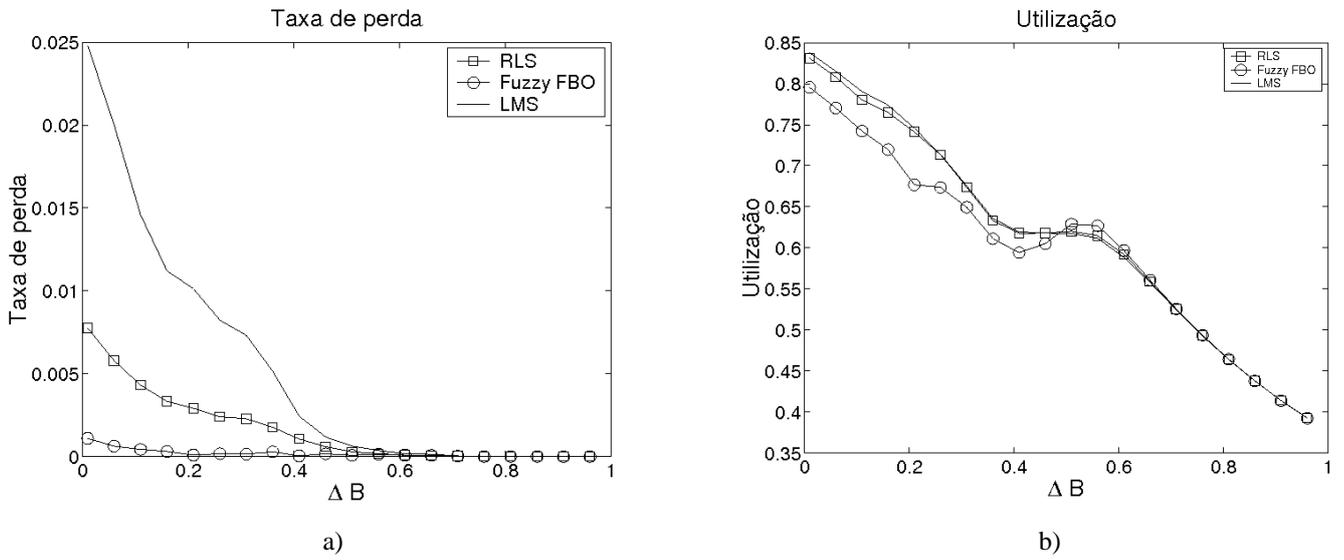


Figura 6. Comparação de desempenho entre esquemas de alocação de banda

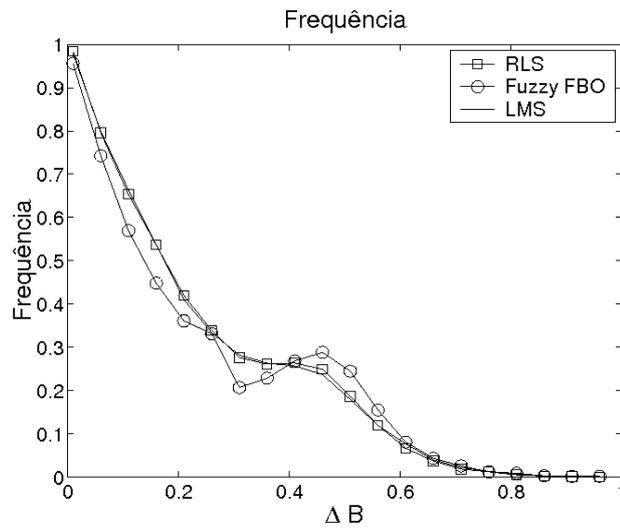


Figura 7. Frequência de sinalização x ΔBW

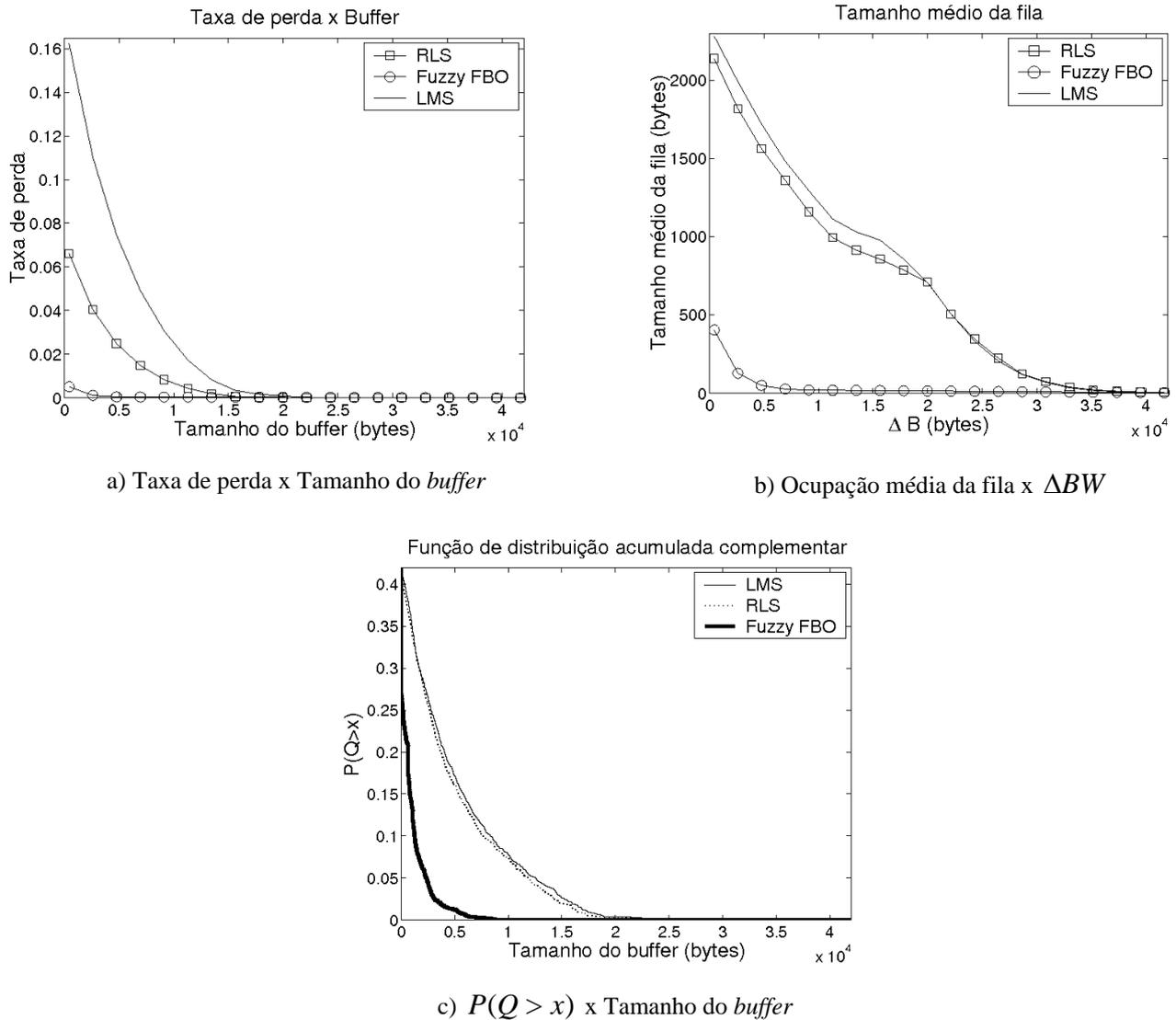


Figura 8. Análise do comportamento de fila dos esquemas de alocação de banda para o traço de tráfego 10-7-S-1

6 Conclusão

As características dos fluxos de tráfego nas rede atuais como dependência de longo prazo e rajadas em múltiplas escalas tornam a modelagem e predição de tráfego tarefas difíceis e desafiadoras. Neste trabalho, foi proposto um modelo *fuzzy-FBO* cujo algoritmo de treinamento adaptativo ARFA permite que a predição adaptativa e em tempo real do tráfego de redes seja realizada com um número reduzido de regras nebulosas. O algoritmo de treinamento ARFA desenvolvido consiste de 2 estágios, ambos adaptativos, onde no primeiro cria-se agrupamentos *fuzzy* e no segundo se faz um ajuste fino dos parâmetros obtidos no primeiro estágio, como posicionamento dos centros e forma geométrica das funções de pertinência.

A fim de se obter funções de base ortonormais para o modelo *fuzzy* em questão através do cálculo do pólo do sistema, apresentamos uma expressão analítica para a função de autocorrelação de processos multifractais que é adequada para descrever várias características do tráfego de redes [22]. Em seguida, introduzimos um procedimento para o cálculo do pólo dominante, o qual é utilizado como pólo de Laguerre para o modelo *fuzzy-FBO*. Comprovamos que há uma melhoria de desempenho de predição do modelo, ou seja, predições mais precisas e robustas são obtidas com o acréscimo das funções de base ortonormais. As comparações realizadas com outros preditores adaptativos, demonstraram na maioria dos casos, um desempenho superior de predição do modelo *fuzzy-FBO* proposto para diferentes horizontes de predição e número de regras consideradas. Constatou-se esse fato por meio do erro quadrático médio normalizado e do teste de hipótese T aplicado a séries de erros obtidos com os preditores comparados.

Quanto à provisão de banda para fluxos de tráfego de redes, apresentamos um novo esquema de provisão de banda baseado em predição. Este esquema relaciona as predições de tráfego realizadas pelo modelo *fuzzy-FBO* com a banda a ser alocada no próximo instante de tempo de modo a obter informação atualizada da taxa necessária para o tráfego de dados em

um enlace. Uma das vantagens do esquema adaptativo de provisão de banda é que a alocação de banda pode ser realizada baseada nas amostras de tráfego disponíveis no instante de tempo atual. Como este esquema de provisão de banda se adapta segundo as mudanças do tráfego, um melhor aproveitamento dos recursos é obtido comparado a uma alocação estática de banda. O método de alocação de banda proposto consegue manter um bom equilíbrio entre perda de dados, frequência de sinalização e utilização. Verificamos que o objetivo de se conseguir uma menor taxa de perda com a aplicação do nosso esquema preditivo de provisão de banda foi atingido. Aliado a isso, observamos uma taxa média de ocupação do *buffer* mais baixa e uma probabilidade de perda de *bytes* menor do que para os outros esquemas. Tal resultado se deve a melhora de desempenho de predição proporcionada ao se utilizar o modelo *fuzzy*-FBO. A partir do modelo de predição proposto, esquemas que considerem também minimização do custo de sinalização e diferentes valores de retardo de pacotes podem ser desenvolvidos. Questões estas que serão tratadas em trabalhos futuros.

Apêndice A

Prova do Teorema 1

Denotemos por Δt_0 a escala de tempo do processo multifractal $X(\Delta t_0)$ obtido pelo produto de um variável lognormal Y e uma cascata multiplicativa binomial $\mu(\Delta t_N)$ em seu número máximo de estágios [21]. A função de autocorrelação $E[X(\Delta t_0)_n, X(\Delta t_0)_{n+k}]$ de um processo multifractal onde $k = 2^b$, $b = 1, 2, \dots, N$ pode ser escrita como:

$$E[X(\Delta t_0)_n, X(\Delta t_0)_{n+k}] = E(Y^2) \{ 2^{2N} E[\mu(\Delta t_N)_n \cdot \mu(\Delta t_N)_{n+k}] - 1 \} + E[X(\Delta t_0)_n] E[X(\Delta t_0)_{n+k}] \quad (52)$$

onde $N = \log_2(N_a)$ é o número de estágios da cascata multiplicativa binomial e N_a é o número de amostras do processo.

As medidas $\mu(\Delta t_N)_n$ e $\mu(\Delta t_N)_{n+k}$ podem dadas em função de $\mu(\Delta t_{N-p-1})$ da seguinte forma:

$$\begin{aligned} \mu(\Delta t_N)_n &= \mu(\Delta t_{N-p-1}) r_{N-p} \prod_{i=N-p+1}^N r_{i, j_1} \\ \mu(\Delta t_N)_{n+k} &= \mu(\Delta t_{N-p-1}) (1 - r_{N-p}) \prod_{i=N-p+1}^N r_{i, j_2} \end{aligned} \quad (53)$$

onde $r_{i,j}$ denota os multiplicadores da cascata multiplicativa no estágio i . Assim, temos:

$$\begin{aligned} E[\mu(\Delta t_N)_n \cdot \mu(\Delta t_N)_{n+k}] &= E[\mu(\Delta t_{N-p-1})^2] E[r_{N-p} (1 - r_{N-p})] E\left[\prod_{i=N-p+1}^N r_{i, j_1} r_{i, j_2} \right] \\ &= E(R^2)^{N-p-1} \left[\frac{1}{2} - E(R^2) \right] \left(\frac{1}{2} \right)^{2p} \end{aligned} \quad (54)$$

Inserindo a equação (54) em (52) e utilizando a média (equação (17)) e a variância (equação (18)) do processo $X(\Delta t_0)$, obtemos a função de autocorrelação desejada:

$$E[X(\Delta t_0)_n, X(\Delta t_0)_{n+k}] = e^{2\rho+\gamma^2} \left(\frac{\alpha(\alpha+1)^{N-1}}{(\alpha+1/2)^N} k^{-\log_2\left(\frac{\alpha+1}{\alpha+1/2}\right)} \right) \quad (55)$$

Apêndice B

Prova da Proposição 1

Seja $X(t)$ um processo cuja função de autocorrelação é representada por $r_x(k) = E[X(t+k)X(t)]$. No método de Levinson-Durbon, calcula-se inicialmente o valor do coeficiente de reflexão Γ_{j+1} ($j = 1, \dots, n$) para encontrar um modelo AR de ordem n [31]. O coeficiente de reflexão é usado no cálculo dos pólos do modelo AR (Auto-Regressivo). Neste método, os valores dos coeficientes $a_{n+1}(j)$ do modelo AR são calculados a partir de $a_n(j)$ de tal forma que:

$$r_x(k) + \sum_{l=1}^p a_p(l)r_x(k-l) = 0; \quad k = 1, 2, \dots, p \quad (56)$$

Assim, o pólo de um modelo com ($j = 1$) é dado por [31]:

$$\Gamma_{j+1} = -\frac{\gamma_j}{\varepsilon_j} \quad (57)$$

onde ε_{j+1} é o erro de modelagem.

Inserindo a equação de autocorrelação (19) em (57), e sabendo que $p = \Gamma_1$, temos a seguinte equação para o pólo p :

$$p = -\frac{1}{2 \log_2\left(\frac{\alpha+1}{\alpha+1/2}\right)} \quad (58)$$

7 Referências Bibliográficas

- [1] J. C. Bezdek. "Fuzzy models-what are they and why?", *IEEE Trans. Fuzzy Syst.*, pp. 1-6, Fevereiro 1993.
- [2] B.-S. Chen; S.-C. Peng e K.-C. Wang. "Traffic modeling, prediction, and congestion control for high-speed networks: A fuzzy AR approach". *IEEE Transactions on Fuzzy Systems*, vol.8, no. 5, pp. 491 - 508, Outubro 2000.
- [3] E. Kim, M. Park, S. Ji e M. Park. "A new approach to fuzzy modelling". *IEEE Trans. Fuzzy Syst*, vol.5, pp.328-337, Agosto 1997.
- [4] T. Takagi e M. Sugeno. "Fuzzy identification of systems and its applications to modeling and control". *IEEE Trans. Syst., Man, Cybern.*, vol. 15, pp.116-132, Jan. 1985.
- [5] M. Sugeno and T. Yasukawa. "A fuzzy-logic-based approach to qualitative modelling". *IEEE Trans. Fuzzy Syst*, vol.1, pp.7-31, Fevereiro 1993.
- [6] H.-H Liu and P.-L. Hsu. "Design and simulation of adaptive fuzzy control on the traffic network". In *International Joint Conference SICE-ICASE*, pp. 4961-4966, Outubro 2006.
- [7] S. Molnar and G. Terdik. "A general fractal model of internet traffic". In *IEEE LCN 2001*, Tampa, Florida, Novembro 2001.
- [8] M. S. Crouse e R. G. Baraniuk V. J. Ribeiro, R. H. Riedi. "Multiscale queueing analysis of long-range dependent traffic". *Proc. IEEE INFOCOM*, vol.2, pp. 1026-1035, Março 2000.
- [9] I. W. C. Lee and A. O. Fapojuwo. "Stochastic processes for computer network traffic modelling". *Computer Communications*, vol. 29, pp.1-23, Março 2005.
- [10] M. Grossglauser e J.-C. Bolot. "On the relevance of long-range dependence in network traffic". *IEEE/ACM Transactions on Networking*, vol. 7, no.5, pp.629-640, Outubro 1999.
- [11] Y. C. Ouyang, C.-W. Yang e W. S. Lian. "Neural networks based variable bit rate traffic prediction for traffic control using multiple leaky bucket". *Journal of High Speed Networks*. vol. 15, no.2, pp.11-122, 2006.
- [12] V. A. Aquino e J. A. Barria. "Multiresolution FIR neural-network-based learning algorithm applied to network traffic prediction". *IEEE Transactions on Systems, Man and Cybernetics-C*, vol. 36, no.2, pp.208-220, Março 2006.
- [13] R. R. Yager e D. P. Filev. *Essentials of Fuzzy Modeling*. John Wiley and Sons, 1994.
- [14] C.-S. Ouyang, W.-J. Lee, and S.-J. Lee. "A TSK-type neurofuzzy network approach to system modeling problems". *IEEE Transactions on Systems, Man and Cybernetics, Part B*, vol. 35, no.(4), pp.751-767, Agosto 2005.
- [15] G. H. C. Oliveira, R. J. G. B. Campello e W.C. Amaral. "Fuzzy models within orthonormal basis function framework". In *IEEE International Fuzzy Systems Conference Proceedings*, Seoul, Korea, 22-25 Agosto 1999.
- [16] B. Ninness e F. Gustafsson. Orthonormal bases for system identification. In *Proc. of 3rd European Control Conference*, pp. 13-18, Roma, Itália, Setembro 1995.
- [17] B. Wahlberg e L. Ljung. "Hard frequency-domain model error bounds from least-squares like identification techniques". *IEEE Trans. on Automatic Control*, vol. 37, no.7, pp.900-912, Julho 1992.

- [18] G. A. Dumont e Y. Fu. "Non-linear adaptive control via Laguerre expansion of Volterra kernels". *Int. J. Adaptive Control and Signal Processing*, vol. 7, pp.367-382, 1993.
- [19] K.Park e W. Willinger. *Self-similar Network Traffic and Performance Evaluation*. John Wiley and Sons, New York, 2000.
- [20] A. Feldmann, A. C. Gilbert e W. Willinger. "Data networks as cascades: Investigating the multifractal nature of Internet WAN traffic". pp. 25-38. ACM/SIGCOMM'98, Vancouver, 1998.
- [21] R. H. Riedi, M. S. Crouse, V. J. Ribeiro e R. G. Baraniuk. "A multifractal wavelet model with application to network traffic". *IEEE Trans. on Information Theory*, vol. 45, no.3, pp. 992-1018, Abril 1999.
- [22] T. D. Dang., S. Molnar e I. Maricza. "Capturing the complete characteristics of multifractal network traffic". In *Globecom 2002*, Taipei, Taiwan, Novembro 2002.
- [23] T. D. Dang; S. Molnar e I. Maricza. "Queuing performance estimation for general multifractal traffic". *Int. J. Commun. Syst.*, vol. 16, no.2, pp.117-136, 2003.
- [24] M. H. Hayes. *Statistical Digital Signal Processing and Modeling*. John Wiley of Sons, 1996.
- [25] M. S. Bazaraa, H. D. Sherali e C. M. Shetty. *Nonlinear Programming: Theory and Algorithms*. John Wiley and Sons, New York, segunda edição, 1993.
- [26] P. Young. *Recursive Estimation and Time Series Analysis: An Introduction*. Communications and Control Engineering Series. Springer-Verlag, New York, 1984.
- [27] B. K. Lee, B. S. Chen, Y. S. Yang e T. H. Lee. "Fuzzy adaptive predictive flow control of ATM network traffic". *IEEE Trans. on Fuzzy Syst.*, vol. 11, no.4, pp.568-581, Agosto 2003.
- [28] C. C. Chuang, C. C. Hsiao e J. T. Jeng. "Adaptive fuzzy regression clustering algorithm for TSK fuzzy modelling". In *IEEE International Symposium on Computational Intelligence in Robotics and Automation*, Julho 2003.
- [29] Lee Luan Ling et al. A computational tool and optimization methods for multimedia traffic characterization and effective bandwidth estimation on modern communication networks. Relatório Técnico. Projeto Ericson UNI-20, Laboratório de Reconhecimento de Padrões e Redes de Comunicações (LRPRC) -Unicamp, Março 2002.
- [30] B. U. Toreyin, M. Trocan e A. E. Cetin. "LMS based adaptive prediction for scalable video coding". In ICASSP, Toulouse, França, 14 a 19 de Maio 2006.
- [31] S. S. Haykin. *Modern Filters*. Macmillan Publishing Company, New York, 1989.
- [32] F. H. T. Vieira e L.L. Lee. "Fuzzy modeling and prediction with confidence bound estimation for traffic rate allocation in high-speed networks". *Simpósio Brasileiro de Redes Neurais*, São Luís, 29 de Setembro a 01 de outubro 2004.
- [33] A. Papoulis. *Probability, Random Variables, and Stochastic Processes*. McGraw-Hill, New York, 3rd edition, 1991.
- [34] D. A. Rolls, G. Michailidis e F. Hernandez-Campos. "Queueing analysis of network traffic: methodology and visualization tools". *Computer Networks*, vol. 48, pp.447-473, Janeiro 2005.
- [35] B. Qiu. "Analysis of fuzzy logic and autoregressive video source predictors using T-tests". In *International Symposium on Signal Processing and its Applications (ISSPA)*. Brisbane, Austrália, Agosto 1999.
- [36] H. T. Tran e T. Ziegler. "Adaptive bandwidth provisioning with explicit respect to QoS requirements". *Computer Communications*, vol. 28, pp.1862-1876, Janeiro 2005.
- [37] A. M. Adas. "Using adaptive linear prediction to support real-time VBR video under RCBR network service model". *IEEE/ACM Trans. Net.*, vol. 6, no.5, pp.635-645, Outubro 1998.
- [38] S. Chong, S.-Q.Li e J. Ghosh. "Predictive dynamic bandwidth allocation for efficient transport of real-time VBR video over ATM". *IEEE JSAC*, vol. 13, no.1, pp.12-23, Jan. 1995.