

ProgCon: um Sistema para Prognóstico de Congestionamento de Tráfego de Redes

Jorge L. C. e Silva Marcilia A. Campos Paulo R. F. Cunha

Universidade Estadual do Ceará (UECE)
Centro de Informática - Universidade Federal de Pernambuco (UFPE)
Caixa Postal 7851 - 50740-540 - Recife - PE - Brasil

{jlcs,mac,prfc}@cin.ufpe.br

***Abstract** - The behavior of the current networks traffic is chaotic, possessing statistical properties which include complex temporal correlations. In this scenario, a problem not yet solved is related to traffic congestion. A non stationary characteristic, short and long-range temporal dependences and the present noises in the traffic give opportunity to use new techniques that allow the use of simple models for its analysis. This work proposes a mechanism for forecasting possible congestions through the identification of the tendency component in time series that represent the network real traffic, using discrete wavelet transform. The proposed mechanism is efficient since can detect congestion in the traffic network.*

Keywords: wavelet, traffic congestion, prediction, self-similar .

***Resumo** – O comportamento do tráfego das redes atuais é caótico, possuindo propriedades estatísticas que incluem correlações temporais complexas. Neste cenário, um problema ainda não resolvido é o do congestionamento do tráfego. A natureza não estacionária, as dependências temporais de curto e de longo alcance e os ruídos presentes no tráfego ensejam o uso de novas técnicas que permite a utilização de modelos simples para a sua análise. Este trabalho propõe um mecanismo para prever possíveis congestionamentos, através da identificação da componente de tendência em séries de tempo que representam o tráfego real de redes, utilizando a transformada de wavelet discreta. O mecanismo proposto mostrou ser eficiente pois detectou o possível congestionamento da rede.*

Palavras-chave: wavelet, congestionamento, prognóstico, auto-similar.

1. Introdução

O objetivo deste trabalho é apresentar o sistema ProgCon, o qual identifica situações que podem causar congestionamentos em redes de computadores. A idéia é modelar as propriedades estatísticas do tráfego no domínio *wavelet* (Daubechies, 1992) e, em seguida, utilizar técnicas de estimação de tendências para inferir possíveis casos de congestionamento.

Os benefícios de se prever possíveis congestionamentos em um tempo futuro próximo (na ordem de milissegundos) são relevantes para os mecanismos de controle de congestionamento das atuais redes, pois evitam o desperdício de recursos através de medidas de controle que vão desde o descarte de pacotes até o provisionamento de

largura de banda adicional nos canais, passando pela limitação da taxa de entrada de pacotes na rede.

Para avaliar os resultados, o prognóstico efetuado pelo ProgCon foi comparado com o modelo de previsão de séries temporais FARIMA (*Fractional Autoregressive Integrated Moving Average*), proposto por Hosking (1981).

As principais contribuições deste trabalho estão na proposta do mecanismo e na sua aplicabilidade prática e, também, na investigação da transformada de *wavelet* para fins de prognóstico de congestionamento. O ProgCon constitui, portanto, uma nova abordagem para prognósticos de congestionamentos de tráfego de redes.

Este trabalho está organizado da seguinte forma. A primeira seção introduz a terminologia e trabalhos relacionados para a leitura do artigo. Na seção 2 tem-se o cenário do mecanismo e do tráfego e a descrição dos procedimentos do mecanismo proposto e dos parâmetros da modelagem matemática. A seção 3 mostra os resultados obtidos e a validação do ProgCon, enquanto que a seção 4 apresenta as conclusões.

1.1 Fundamentos Teóricos

Wavelets (Daubechies, 1992) podem ser consideradas como sendo funções da forma $\omega_a^b(t)$, geradas a partir de translações (ou deslocamentos) e dilatações (ou contrações) de uma única função $\omega(t)$, que permitem a representação, tanto no domínio da frequência quanto no tempo, dos elementos de determinados subespaços do conjunto de todas as funções. A variável independente t representa o tempo, e seu domínio é referido com domínio do tempo. A forma geral de uma “família” de *wavelet* é

$$\omega_a^b(t) = \frac{1}{\sqrt{a}} \omega\left(\frac{t-b}{a}\right), a > 0, -\infty < b < 0 \quad (1.1)$$

A função original, $\omega(t)$, corresponde ao caso em que $a=1$ e $b=0$. O procedimento de dilatações binárias e translações diádicas consiste em representar operações aplicadas à função original de acordo com

$$\omega_j^k(t) = 2^{\frac{j}{2}} \omega(2^j t - k), j, k \in Z,$$

fazendo em (1.1) $a = 2^{-j}$, $b = k2^j$, $j, k \in Z$.

Seja $x = (x_0, x_1, \dots, x_T)$ um vetor de dados de tamanho 2^n , isto é, $T = 2^n - 1$, $n \in Z^+ = \{1, 2, \dots\}$. Define-se a transformada de *wavelet* discreta de x , com respeito à *wavelet* mãe ω , como

$$d_j^k = \sum_{t=0}^T x_t \omega_j^k(t/T), j = 0, 1, \dots, n-1 \text{ e } k = 0, 1, \dots, 2^j - 1,$$

onde os d_j^k são chamados de coeficientes.

Redução de ruído (*de-noising*) é uma aplicação de *wavelet* (Gomes; Velho; Goldstein, 1999), cujo objetivo é reduzir ou mesmo remover o ruído presente num sinal ou série de tempo, através da diminuição da magnitude dos coeficientes.

Este procedimento aplicado a uma dada série de tempo $\{x_t\}_{t=0}^{n-1}$ atua na forma de uma eventual redução ou encolhimento de cada um dos coeficientes

$\{d_j^k\}$, $0 \leq j \leq n-1, 0 \leq k \leq 2^j - 1$, de acordo com um determinado conjunto de critérios. Os coeficientes reduzidos são obtidos através da aplicação de limiares aos coeficientes originais. O procedimento de redução dos coeficientes de *wavelet* de um sinal consiste de: (i) aplicação da transformada de *wavelet* discreta aos dados x_0, x_1, \dots, x_{n-1} , obtendo-se n (potência de 2) coeficientes d_j^k , todos com ruído; (ii) utilização de limiares (*thresholds*) para diminuir os coeficientes, tornando nulos aqueles abaixo de um valor especificado, obtendo-se nesta etapa os coeficientes sem ruído, \tilde{d}_j^k ; (iii) aplicação da transformada inversa de *wavelet* sobre os coeficientes da etapa (ii) para obter as estimativas do sinal.

Com relação aos limiares, as questões referentes a escolha da política de limiares e dos parâmetros que governam esta política assumem fundamental relevância. Duas políticas bastante difundidas de redução dos coeficientes são denominadas, respectivamente, de *dura* (*hard threshold*) e *suave* (*soft threshold*). Uma vez definido um dado limiar λ , as políticas dura e suave podem ser expressas, respectivamente, por

$$\tilde{d}_j^k = \begin{cases} 0, & \text{se } |d_j^k| \leq \lambda \\ d_j^k, & \text{se } |d_j^k| > \lambda. \end{cases}$$

e

$$\tilde{d}_j^k = \begin{cases} 0, & \text{se } |d_j^k| \leq \lambda \\ \text{sign}(d_j^k)(|d_j^k| - \lambda), & \text{se } |d_j^k| > \lambda, \end{cases}$$

onde a função sinal ($\text{sign}(d_j^k)$) representa o sinal do coeficiente de *wavelet*.

Com respeito ao estabelecimento dos parâmetros (limiares) que governam as políticas, existem duas propostas denominadas, respectivamente, de *quantil* e *universal*. A proposta quantil consiste em determinar um parâmetro λ -quantil do conjunto de todos coeficientes de *wavelet* (Vidakovic; Müller, 1998). Por exemplo, deseja-se substituir os 30% menores coeficientes (em valores absolutos) por zero. Então o limiar λ -quantil corresponde ao maior valor dos coeficientes que estão dentro da percentagem de 30%. Donoho, Johnstone, Kerkycharian e Picard (1995) propuseram um estimador, chamado universal, para o parâmetro λ que consiste no estabelecimento de um único λ , relacionado com o desvio padrão dos coeficientes. Para estimar o parâmetro de limiaridade é usada a fórmula

$$\lambda = \sigma \sqrt{2 \log(n)},$$

onde n é o número de coeficientes e σ o desvio padrão.

O comportamento tráfego das redes atuais é caótico, podendo ser representado por processos auto-similares dependentes de longa duração. O modelo FARIMA (Hosking 1981) tem sido muito usado em modelagem de tráfego de rede (Shu et al. 2000). O modelo FARIMA(p, d, q) é uma extensão natural do processo ARIMA(p, d, q) (Beran, 1994) quando se utilizam valores do conjunto dos reais para o parâmetro d que representa o grau ou parâmetro de diferenciação. Uma característica importante desse processo é a propriedade de exibir longa dependência quando $d \in (0.0 ; 0.5)$, curta

dependência quando $d = 0$ e memória intermediária quando $d \in (-0.5 ; 0)$. X_t é um processo FARIMA(p,d,q) se:

$$\phi(B)\Delta^d X_t = \theta(B)\varepsilon_t,$$

onde $\phi(B)$ e $\theta(B)$ são, respectivamente, processos AR (AutoRegressive) e MA (Moving Average) estacionários (Beran,1994), Δ é o operador de diferença, isto é, $\Delta^d X_t = X_t - X_{t-d}$ e ε_t é o ruído branco.

1.2 Trabalhos relacionados

Uma das principais motivações no uso de *wavelet* é a capacidade que a transformada tem em reduzir a correlação temporal de modo que a auto-correlação dos coeficientes seja menor que a do correspondente processo temporal. Kaplan e Kuo (1993) investigaram a estrutura de correlação de processos FGN (*Fractional Gaussian Noise*), dependentes de longo alcance, com parâmetro Hurst (Hurst, 1955), H ($0.5 < H < 1$), e mostraram que estes processos, depois de transformados por uma *wavelet* de Haar (Daubechies, 1992) apresentam uma série de coeficientes onde a soma das correlações converge para uma constante. Estudos teóricos (Kaplan; Kuo, 1993; Flandrin, 1993) da estrutura de correlação dos coeficientes destes processos indicam que a transformada de *wavelet* reduz a dependência temporal, implicando que a correlação dos coeficientes decaia exponencialmente. As propriedades estatísticas dos coeficientes de *wavelet* podem ser usadas para caracterizar o processo estocástico original.

Roughan; Veitch; Abry (1998) desenvolveram um sistema, baseado em *wavelet*, para estimar o parâmetro de Hurst de modo *online* em um roteador, com o objetivo de fazer modificações adaptativas no controle de congestionamento e prover informações sobre a dinâmica do tráfego. A introdução de técnicas baseadas em *wavelets* para analisar *traces* de tráfego (Abry; Veitch, 1998) representa um passo na direção do desenvolvimento de alternativas de modelagem, em especial, em tráfego não estacionário. Riedi et al. (1999) mostraram simulações de tráfego de rede usando *wavelets*, com resultados que assemelham-se aos *traces* reais de tráfego *Ethernet*. Qiao; Skicewicz; Dinda (2003) estudaram empiricamente a previsibilidade de coeficientes *wavelet*, gerados a partir de tráfego real, para uso na determinação de tempos de transferência de mensagens. O sistema WIND (Huang; Feldmann; Willinger, 2001) usa análises de escala baseada em *wavelet* para detectar problemas de desempenho de rede. Uma classe de processos, chamada de multifractais, foi proposta por Peltier; Lévy-Véhel (1997) e relacionada com *wavelet* por Feldmann; Gilbert; Willinger (1998) que, na prática, investigaram a natureza multifractal do tráfego Internet e utilizaram *wavelets* para detectar, identificar e descrever comportamentos caóticos. As causas físicas da auto-similaridade no tráfego de rede e o impacto da propriedade multifractal em mecanismos de controle de congestionamento do protocolo TCP foram investigadas em Feldman et al. (1999). Ma; Ji (2001) modelaram o tráfego auto-similar gerado por essas causas no domínio da *wavelet*.

2. Metodologia e Cenário

Os atuais mecanismos de controle de congestionamento, como o RED (Braden et al., 1998), estão localizados nos nodos intermediários das redes (roteadores, *switches*, etc.), onde acontece a maioria dos congestionamentos. Da mesma forma que o mecanismo de

controle, o ProgCon deve estar localizado nas interfaces dos roteadores. A Figura 1 ilustra a possível localização do ProgCon. Nesta figura, os componentes funcionais de um NIC (*Network Interface Card*) podem processar tanto o tráfego de entrada quanto o de saída. O ProgCon atua sobre os fluxos enviados para as portas de saída do roteador antes destes serem armazenados nas filas (*buffers*).

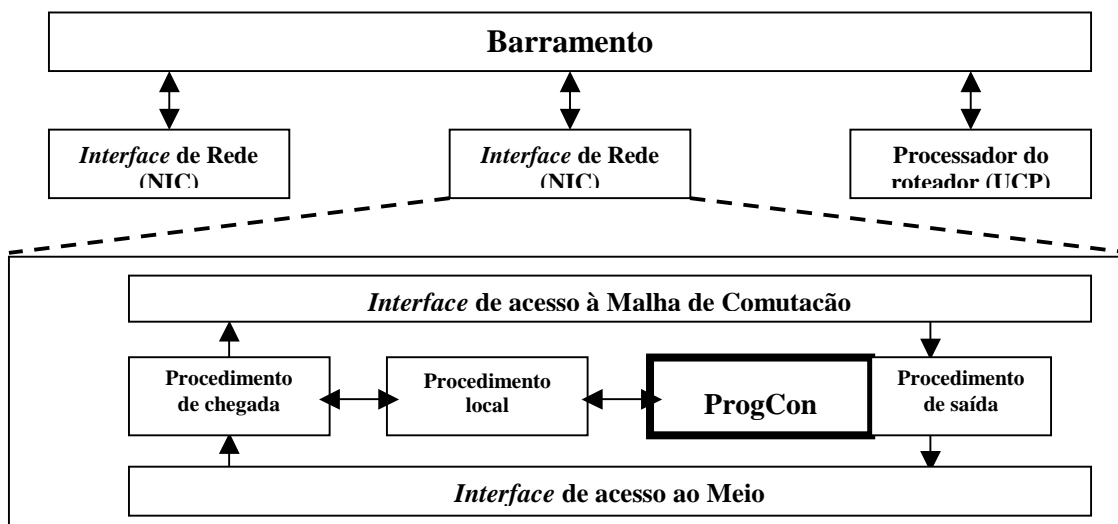


Figura 1. Diagrama Funcional de um Roteador

ProgCon é um sistema desenvolvido em C/C++ que utiliza o processo de contagem do número de bits ou de pacotes que chegam no componente ou na porta de saída de um dado roteador, no intervalo de tempo $[0, t]$. Dentro de uma janela de tempo T (janela de observação), onde $T = 2^n$, $n = 1, 2, \dots$, e $T > t$, a série formada pelo processo de contagem é dinâmica porque à medida que a próxima observação (número de pacotes no intervalo $[0, t]$) se torna disponível, a observação mais antiga é desprezada e a mais nova incorporada à série. Assim, todos os componentes funcionais do mecanismo estão continuamente trabalhando com séries atualizadas que representam o comportamento do tráfego em tempo real.

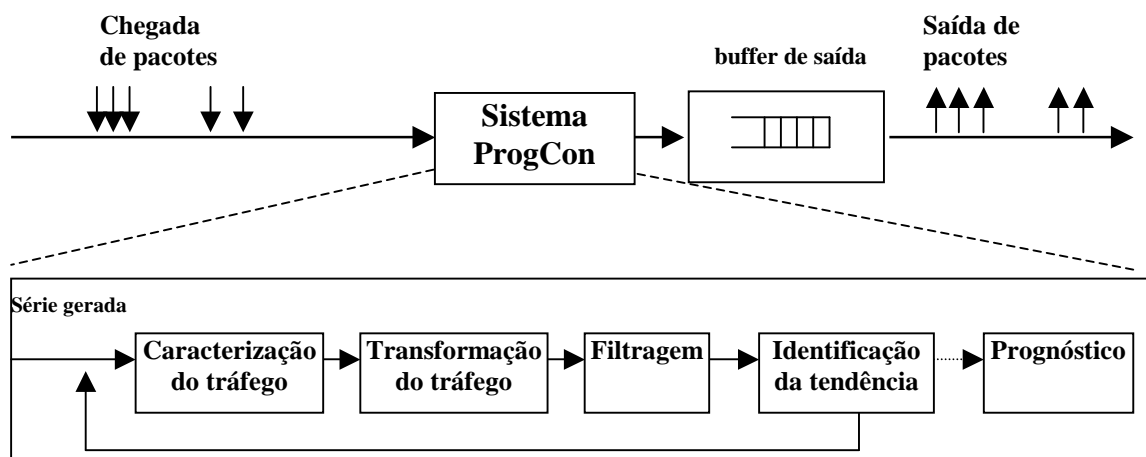


Figura 2. Componentes Funcionais do Sistema

A Figura 2 ilustra os procedimentos funcionais do sistema: (i) Caracterização do tráfego, (ii) Transformação do tráfego, (iii) Filtragem dos coeficientes, (iv) Identificação da tendência e (v) Prognóstico de congestionamento.

O primeiro procedimento tem o objetivo de identificar o grau de auto-similaridade do tráfego, através da utilização da função de autocorrelação, análise da estatística R/S e do parâmetro de Hurst. Durante essa fase o sistema calcula as estatísticas média aritmética, variância, estatística R/S e o parâmetro de Hurst. Estas mesmas estatísticas são também calculadas sobre a série transformada, após o procedimento de identificação de tendência.

O segundo procedimento consiste em aplicar à série coletada, que corresponde ao tráfego real em um intervalo de tempo T , uma função de *wavelet* discreta, isto é, efetuar a transformação direta da série. Para realizar este procedimento, o sistema estima, inicialmente, a base de *wavelet* a ser empregada durante o seu funcionamento, usando para isso a primeira janela de observação e 20 bases de *wavelet* mais conhecidas. A estimação consiste em: (i) calcular a transformada direta de *wavelet*, (ii) efetuar a filtragem dos coeficientes, (iii) calcular a transformada inversa e (iv) comparar o valor observado (original) com o correspondente valor calculado (reconstruído), computando-se o erro quadrático médio, EQM_k , $k=1, \dots, 20$. Seja $EQM_{(k)}$ a ordenação de EQM_k , isto é, $EQM_{(1)} \leq \dots \leq EQM_{(20)}$. Portanto, a base escolhida é $EQM_{(1)}$.

O passo seguinte consiste em aplicar filtros aos coeficientes de *wavelet* para encolher ou até mesmo zerar aqueles que estão a partir de um determinado limiar. O limiar deve ser calibrado de maneira que a série reconstruída não apresente valores próximos da série original, mas também não apresente valores muito distantes. Realizada a filtragem, o procedimento efetua a transformada inversa usando a mesma base utilizada na transformação direta.

O quarto procedimento identifica a tendência na série transformada, através dos testes de Cox-Stuart e de Spearman (Morettin; Toloi, 1987). Para isto define-se um parâmetro de indicação de congestionamento, IC, o qual pode ser a média amostral do tráfego ou definido pelo usuário do sistema. O procedimento consiste em comparar os maiores valores da série de tráfego com o IC.

O quinto procedimento consiste de prognosticar possíveis congestionamentos, com base nos cálculos efetuados no procedimento anterior. A indicação de congestionamento em portas do equipamento monitorado é inferida através da comparação do número de pacotes previsto, n_p , com os valores limites que balizam o fluxo de tráfego, mínimo, m , e máximo, M , de pacotes por unidade de tempo. Quando $n_p < m$, o sistema funciona sem indicação de congestionamento. Se $m \leq n_p < M$, o mecanismo proposto informa ao controle de congestionamento do equipamento a possibilidade de ocorrência de engarrafamento do tráfego. A caracterização de congestionamento ocorre quando $n_p \geq M$. Neste caso, o sistema alerta ao controle sobre o acontecimento.

Para validar os resultados do ProgCon foram coletadas trinta séries com 1024 e 2048 observações, representando tráfegos reais e simulados com dependência de longa duração.

O sistema utiliza os seguintes parâmetros definidos na chamada do programa: (i) a base de *wavelet*; (ii) a política de redução de ruído; (iii) o método de escolha do limiar dos coeficientes (iv) o tamanho da janela de observação e (v) o indicador de congestionamento (IC).

Os parâmetros selecionados para o sistema influem na modelagem do tráfego. Para otimizar a escolha dos parâmetros do ProgCon, foi realizado um desenho experimental (Jain, 1991), onde a variável de interesse foi o percentual de acertos do sistema, do tipo 2^k , onde k é a quantidade de fatores. Com base no desenho experimental os parâmetros escolhidos e considerados padrões para o sistema ProgCon foram:

- (i) política de filtragem *hard*;
- (ii) limiar de redução universal-2;
- (iii) janela de tamanho 128;
- (iv) base de *wavelet* com menor erro quadrático médio;

3. Resultados

Os resultados apresentados a seguir mostram o comportamento do sistema proposto usando os parâmetros escolhidos na seção anterior. As bases de *wavelet* usadas para as transformações foram estimadas pelo sistema, como descrito na Seção 2 e estão apresentadas no Quadro 1.

Quadro 1. Bases escolhidas pelo Sistema

Tamanho da janela de observação	Tráfego com 1024 observações	Tráfego com 2048 observações
	Bases de wavelet	
128	Coiflet_4	Daubechies_8
512	Daubechies_10	Spline_2_2

A Tabela 1 mostra os parâmetros de Hurst, tanto dos coeficientes quanto das séries (tráfego) simuladas e transformadas. As duas primeiras colunas desta tabela apresentam estes parâmetros para o tráfego com 1024 observações e as duas últimas para o tráfego com 2048. Observa-se que os parâmetros de Hurst das séries simuladas estão acima de 0.84 e das transformadas acima de 0.90, enquanto seus respectivos coeficientes apresentam valores menores que 0.56, indicando que a transformada de *wavelet* reduz a dependência temporal dos coeficientes. Portanto, os coeficientes obtidos pelas transformadas apresentam dependência de curta duração, o que favorece modificações nos seus valores sem alterar suas autocorrelações.

Tabela 1. Parâmetro de Hurst dos coeficientes e das séries simuladas e transformadas

Séries	Tráfego com 1024 observações		Tráfego com 2048 observações	
	Janela de 128	Janela de 512	Janela de 128	Janela de 512
Coefficientes	0.54	0.51	0.53	0.55
Simuladas	0.87	0.87	0.85	0.85
Transformadas	0.93	0.93	0.91	0.91

A Tabela 2 apresenta os estatísticos dos coeficientes obtidos pela transformação das séries simuladas com 1024 e 2048 observações. Os dados desta tabela mostram que: (i) os desvios padrões dos coeficientes são maiores que os correspondentes desvios das séries simuladas e transformadas (vide Tabela 3), sendo até três vezes maior que os desvios padrões da série com 2048 observações; (ii) a variabilidade dos coeficientes é grande, como pode ser observado através do valor da amplitude, comparada com a das séries simuladas e transformadas (vide Tabelas 3 e 4); (iii) as médias aritméticas dos coeficientes são menores do que 5.20, sobressaindo uma próxima de zero (0.17), enquanto as médias das séries simuladas e transformadas estão acima de 36 (vide Tabela 3 e 4).

Tabela 2. Estatísticos dos coeficientes de wavelets

Estatísticos	Tráfego com 1024 observações		Tráfego com 2048 observações	
	Janela de 128 Coiflet_4	Janela de 512 Daubechies_10	Janela de 128 Daubechies_8	Janela de 512 Spline_2_2
Média	3.56	0.17	5.19	3.48
Variância	2375.49	2388.12	5387.16	6118.79
Desvio padrão	48.74	48.87	73.40	78.22
Coef. Variação	13.70	294.22	14.14	22.50
Mínimo	-197.78	-296.28	-252.76	-435.24
Máximo	698.98	1085.90	1159.66	1785.58
Amplitude	896.76	1382.18	1412.42	2220.82

A Tabela 3 apresenta os estatísticos das séries simuladas e transformadas (reconstruídas). Os dados desta tabela mostram que: (i) a média aritmética da série simulada é menor que as médias das séries transformadas; (ii) os desvios padrões das séries reconstruídas são menores que a correspondente medida da série simulada, indicando uma menor variabilidade das observações transformadas; (iii) a série transformada pela base de Daubechies_10 apresentou menor dispersão; (iv) todos os parâmetros de Hurst estão acima de 0,86, indicando auto-similaridade em todas as séries. Observa-se porém, que as transformações produziram séries com parâmetro de Hurst acima do parâmetro da série original.

Tabela 3. Estatísticos da série simulada com 1024 observações

Estatísticos	Série Simulada	Série transformada (Coiflet_4)	Série transformada (Daubechies_10)
Média	36.77	37.02	37.15
Variância	1036.21	440.38	384.51
Desvio padrão	32.19	20.99	19.61
Coef. de variação	0.88	0.57	0.53
Mínimo	0.00	0.00	0.00
Máximo	130.00	87.00	74.00
Amplitude	130.00	87.00	74.00
Parâmetro de Hurst	0.87	0.93	0.93

A Figura 3 mostra as autocorrelações das séries transformada e simulada e dos coeficientes gerados a partir da base Coiflet_4. Observa-se que as autocorrelações das séries simulada e transformada decaem lentamente se comparadas com a autocorrelação da série dos coeficientes a qual está próxima de zero, confirmando-se que a transformada de *wavelet* obtém coeficientes com dependência de curta duração. As autocorrelações das demais transformadas apresentam o mesmo comportamento e por isso os gráficos são omitidos.

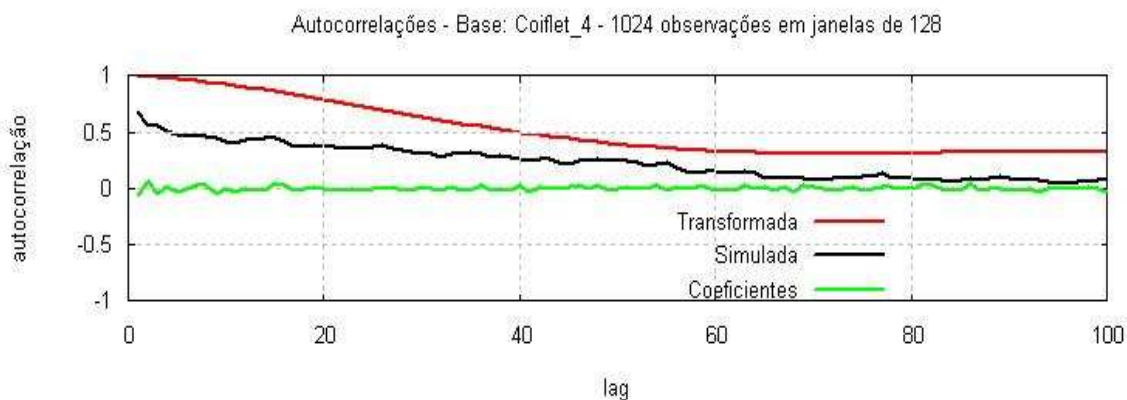


Figura 3. Autocorrelações das séries transformada, simulada e dos coeficientes

A execução da transformada inversa, a partir dos coeficientes gerados pela transformada discreta de *wavelet*, constitui um dos procedimentos do sistema ProCon para prognosticar congestionamento. A série obtida pela transformada inversa, após o procedimento de filtragem, apresenta um comportamento suave, como pode ser visto no tráfego da Figura 4 com 2048 observações; O IC, neste caso, foi a média amostral do tráfego.

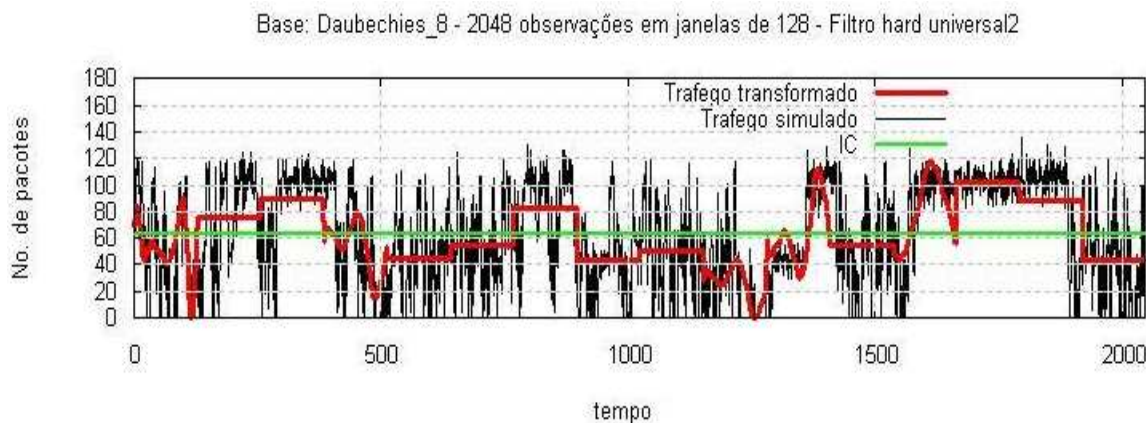


Figura 4. Série original e série transformada pela base de Daubechies_8

A Figura 5 contém as mesmas séries da Figura 4, porém com mais detalhes, isto é, o gráfico mostra as primeiras 512 observações. Verifica-se na Figura 4, e com mais detalhes na Figura 5 (vide setas), que a série transformada acompanha as inflexões da série original, indicando as tendências.

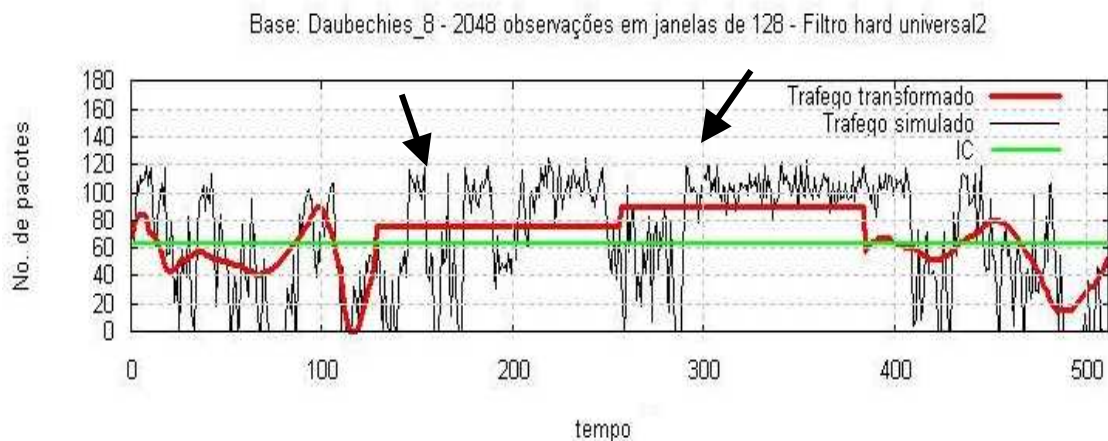


Figura 5. Séries original e transformada com as primeiras 512 observações

O sistema ProCon infere um possível congestionamento verificando a tendência da série gerada pela transformada inversa em cada janela de observação. Por exemplo, na série simulada com 2048 observações e janelas de tamanho 128, o mecanismo executou 16 vezes o teste de tendência, uma para cada janela de observação.

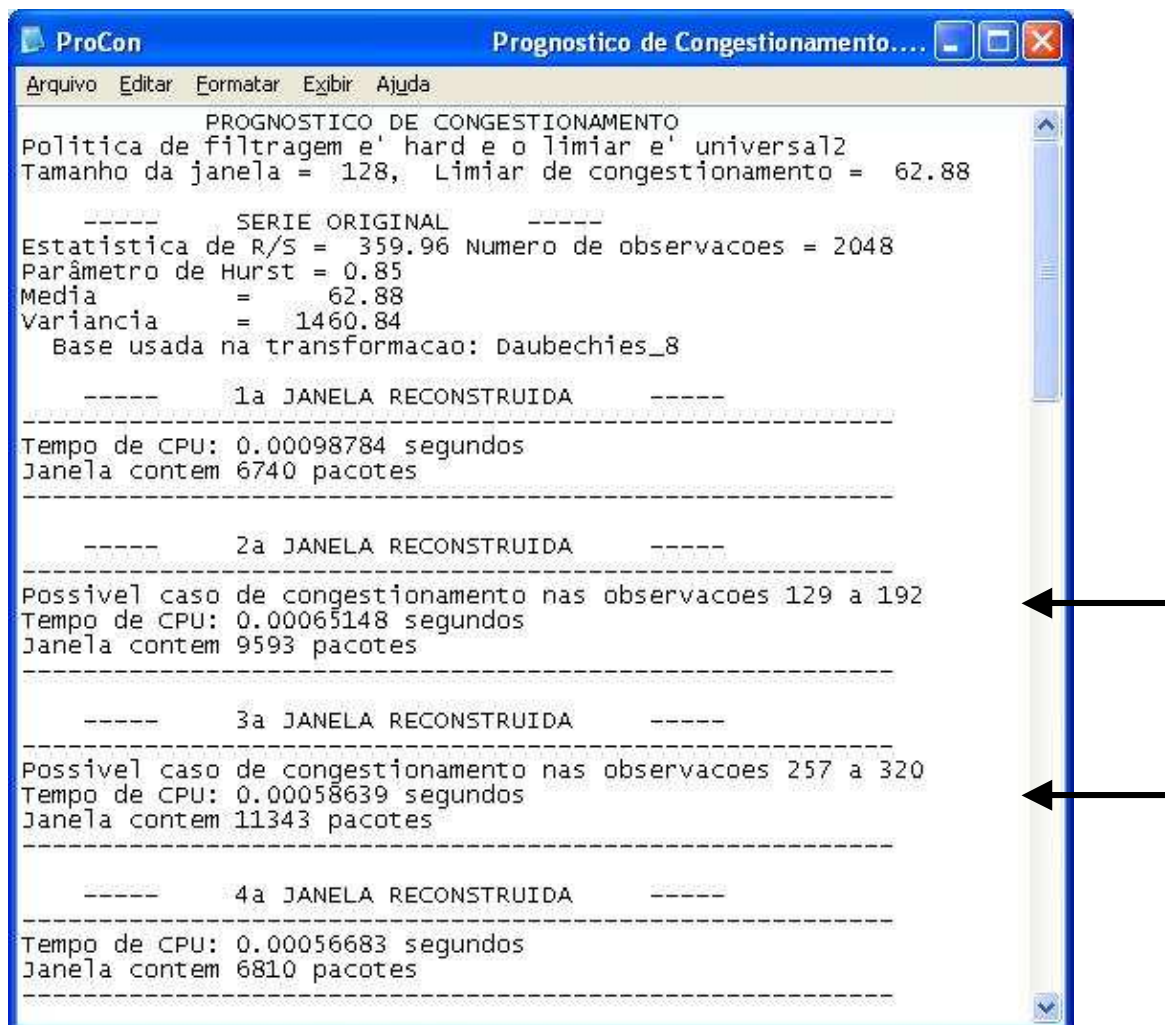


Figura 6. Exemplo de resultado do sistema ProCon – da 1ª à 4ª janela

A Figura 6 apresenta o resultado inferido pelo sistema após a aplicação do teste de tendência nas quatro primeiras janelas de observação da série citada como exemplo. Este resultado mostra possíveis casos de congestionamento na 2ª e na 3ª janela em função da tendência crescente da série transformada, nas observações 129 a 192 e 257 a 320. Confrontando esses resultados com o gráfico da Figura 5, observa-se que nesses intervalos existe realmente tendência crescente, indicado pelas setas. Além desses intervalos, observa-se também a existência de tendência crescente nas observações próximas de 100 e no início da série, que não foram detectadas pelo sistema. Nota-se pelo gráfico, que o comportamento no início da série e na observação 100, é oscilante, estando acima do patamar de congestionamento, mas nos instantes seguintes apresenta valor zero. Este tipo de comportamento é interpretado pelo sistema como rajadas curtas e súbitas que podem ser absorvidas pelo *buffer* do controle de congestionamento do roteador, sem contudo caracterizar congestionamento.

3.1 Viabilidade do Sistema

O ProgCon foi projetado para trabalhar em tempo real e utiliza o relógio da UCP do equipamento, no qual está instalado, para executar seus procedimentos no momento certo e de forma constante, mesmo que o processamento da janela anterior tenha terminado. O sistema ProgCon, durante a sua execução, consulta o relógio da UCP antes do cálculo da transformada discreta *wavelet* e após o teste de tendência. A diferença entre a 2ª consulta ao relógio e a 1ª corresponde ao tempo demandado pelo mecanismo para realizar um prognóstico, isto é, para executar os procedimentos de transformação da série, de filtragem dos coeficientes, de transformação inversa e de teste da tendência. Para estimar o tempo médio de execução do mecanismo, foram simuladas 30 (trinta) séries de tráfego com 2048 observações cada.

O Quadro 2 apresenta os tempos de execução do mecanismo em três diferentes computadores. Os testes realizados com o ProgCon mostram que, para janelas de tamanho 128, os tempos médios de execução medidos, usando os processadores Celerone Pentium IV, estão na casa de microssegundos (608 e 350 microssegundos respectivamente) e com o processador Pentium MMX o tempo médio é de 5.436 milissegundos. Para janelas de tamanho 512, os tempos medidos, usando o processador Pentium IV, estão abaixo de 2 milissegundos, com o processador Celeron próximos de 5 milissegundos e no Pentium MMX acima de 20 milissegundos. É importante ressaltar que o mecanismo demora menos de 1 milissegundo para processar uma janela de tamanho 128 usando processadores de clock 933 MHz e 2,2 GHz e na ordem de 20 milissegundos, para uma janela de tamanho 512, em um computador de 233 MHz.

Considerando-se: (i) que testes realizados com o sistema, utilizando processadores mais rápidos e janelas de tamanho 128, mostraram que os tempos demandados para um prognóstico são inferiores a 1 milissegundo (em média 479 microssegundos), (ii) que a programação do mecanismo não está otimizada, (iii) que nos processador de 933 MHz e 2.2 GHz, o ProgCon gastou menos de 1 milissegundo para cada janela de tamanho 128, (iv) que o mecanismo proposto deve ser “embarcado ou embutido” em um equipamento, como um roteador, conclui-se que o sistema é viável.

Quadro 2. Tempos de execução

Equipamento	Tamanho da janela	Tempo médio (segundos)
Processador Pentium MMX, 233 MHz	128 observações	0.005436
	512 observações	0.021252
Processador Celeron, 995 MHz	128 observações	0.000608
	512 observações	0.005563
Processador Pentium IV, 2.2 GHz	128 observações	0.000350
	512 observações	0.001760

3.2 Validação do Sistema

A validação dos prognósticos efetuados pelo ProgCon foi realizada através da comparação dos acertos obtidos pelo sistema com resultados obtidos pelo modelo FARIMA. Optou-se por este modelo, primeiro devido à natureza auto-similar e a não-estacionaridade das séries de tráfego. Segundo, em função do parâmetro de diferenciação fracionária d desse modelo poder ser estimado a partir do parâmetro Hurst calculado pelo próprio sistema. O processo de estimação e previsão usando FARIMA foi feito através de uma biblioteca suportada pela ferramenta Ox (Doornick; Ooms, 2001).

Para aplicar os procedimentos de comparação de prognósticos, foram simuladas 30 (trinta) séries de tráfego com 2048 observações cada. Para cada uma dessas séries, o ProgCon utilizou, baseado no desenho experimental citado na Seção 2, os seguintes parâmetros: (i) janela de tamanho 128, (ii) base de *wavelet* estimada pelo próprio sistema, (iii) política *hard* de redução de coeficientes, e (iv) limiar universal-2. No total, foram prognosticadas 480 séries de tamanho 128, isto é, cada série de tráfego com 2048 observações contém 16 séries de 128 observações.

A comparação do ProgCon com o modelo FARIMA consistiu na contagem do número de casos de congestionamento detectados e não detectados, onde se sabia previamente as janelas que continham esses casos. Também foi computado o número de casos detectados por ambos, onde não havia congestionamento. Os Quadros 3 e 4 mostram, respectivamente, o número de casos de congestionamento detectados pelo ProgCon e pelo modelo FARIMA.

Quadro 3. Número de acertos do ProCon

ProCon	Detectado		Não detectado		Total
Houve congestionamento	120	73%	45	27%	165
Não houve congestionamento	28	9%	287	91%	315
Total	148		332		480

Quadro 4. Número de acertos do modelo FARIMA

FARIMA	Detectado		Não detectado		Total
Houve congestionamento	130	79%	35	21%	165
Não houve congestionamento	38	12%	277	88%	315
Total	168		312		480

Observa-se nos quadros citados, que o modelo FARIMA apresenta um percentual de acertos maior que o mecanismo ProgCon quando realmente há congestionamento. Contudo, nos casos em que não há congestionamento, o modelo FARIMA detectou mais falsos congestionamentos do que o ProgCon.

Apesar do mecanismo ProgCon ter acertado menos que o modelo FARIMA, o seu processamento foi extremamente rápido quando comparado com o FARIMA. Enquanto o ProgCon processou uma série de 2048 observações em menos de 15 milissegundos, a ferramenta Ox demandou mais de 15 minutos para processar a mesma série em um processador de 933 MHz.

4. Conclusões

Este sistema é uma nova abordagem para prognóstico de congestionamento de tráfego de redes. O ProgCon infere somente prognóstico de tendência, não prevendo valores futuros.

As principais conclusões deste trabalho são:

- (i) a transformada de *wavelet* discreta, a redução dos coeficientes e a transformada inversa geram uma série suavizada;
- (ii) testes de tendência do comportamento da série, que antes não podiam ser aplicados em decorrência da natureza auto-similar do tráfego de redes, podem ser usados em função da suavização da série reconstruída pela transformada de *wavelet*;
- (iii) a escolha correta dos parâmetros do mecanismo proposto, tais como tamanho da janela de observação, política de filtragem, método de escolha do limiar de redução dos coeficientes e patamar de congestionamento, favorece a corretude do prognóstico;
- (iv) os tempos de execução obtidos, da ordem de unidades de microssegundos, e a possibilidade de otimização do sistema viabilizam sua utilização.

Agradecimentos: Este trabalho contou com o apoio da UECE, do CIn / UFPE e do CNPq.

Referências

Abry, P.; Veitch, D. “Wavelet Analysis of Long-Range Dependent Traffic”. *IEEE Transaction Information Theory*, v.44, n.1, p.2-15, 1998.

Beran, J., “Statistics for Long-Memory Processes”. Chapman & Hall, 1994.

Braden, B.; Clark, D.; Crowcroft, J.; Davie, B.; Deering, S.; Estrin, D.; Floyd, S.; Jacobson, V.; Minshall, G.; Partridge, C.; Peterson, L.; Ramakrishnan, K.; Shenker, S.; Wroclawski, J.; Zhang, L., “Recommendations on Queue Management and Congestion Avoidance in the Internet”. Request for Comments 2309, IETF, Apr. 1998.

Daubechies, I. “Ten Lectures on Wavelets”. SIAM, Philadelphia, 1992.

Donoho, D. L.; Johnstone, I. M.; Kerkyacharian, G.; Picard, D. “Universal near Minimality of Wavelet Shrinkage”. Technical Report, Stanford University, 1995.

Doornik, J. A.; Ooms, M., “Introduction to Ox”, Timberlake Consultants Press, 2001.

- Feldman A.; Gilbert, A. C.; Huang P.; Willinger W. “Dynamics of IP Traffic; a Study of the Role of Variability and the Impact of Control”. In: ACM SIGCOMM’99, **Proceedings ...**, p.301-313, 1999.
- Feldman A.; Gilbert, A. C.; Willinger W. “Data Networks as Cascades: Investigating the Multifractal Nature of Internet WAN Traffic”. In: ACM SIGCOMM’98, **Proceedings ...**, p.42-55, 1998.
- Flandrin, P. “Fractional Brownian Motion and Wavelets”. In M. Farge and J. C. Vassilicos, Eds., *Wavelets, Fractals and Fourier Transforms*, p.109-122, Clarendon Press, Oxford, 1993.
- Gomes, J.; Velho, L.; Goldstein, E. “From Fourier Analysis to Wavelets”. In: SIGGRAPH’99. **Anais...**, 1999, 196p.
- Hosking, J. R. M. “Fractional Differencing”. *Biometrika*, V.83, N.1, p.165-176, Apr. 1981.
- Huang, P.; Feldmann, A.; Willinger, W., “A Non-intrusive, Wavelet based Approach to Detecting Network Performance Problems”. In: ACM SIGCOMM Internet Measurement Workshop, **Proceedings ...**, p.213-227, Nov. 2001.
- Hurst, H. E. “Methods of Long-Term Storage in Reservoirs”. *Institution Civil Engineers. Proceedings...*, Part I, p.519-577, 1955.
- Jain, R., “The Art of Computer Systems Performance Analysis”. Jhon Wiley & Sons, Inc., 1991.
- Kaplan, L. M.; Kuo, C. J. “Fractal Estimation from Noisy Data via Discrete Fractional Gaussian Noise (DFGN) and the Haar Basis”. *IEEE Transactions on Information Theory*, v.41, n.3, p.3554-3562, 1993.
- Ma, S.; Ji, C. “Modeling Heterogeneous Network Traffic in Wavelet Domain”. *IEEE/ACM Transactions on Networking*, v.9, n.5, 2001.
- Morettin, P. A.; Toloi, C. M. C., “Séries temporais”. Editora Atual, 136 p., 1987.
- Peltier R.; Lévy-Véhel J. “Multifractal Brownian Motion: Definition and Preliminary Results”. Research Report, INRIA, 39p, Aug. 1997.
- Qiao, Y.; Skicewicz, J.; Dinda, P., “Multiscale Predictability of Network Traffic”. Technical Report TR NWU-CS-02-13, Northwestern University, Evanston, IL, 2003.
- Riedi R.; Crouse M.; Ribeiro V. J.; Baraniuk R. G. “A Multifractal Wavelet Model with Application to Network Traffic”. *IEEE Transaction Information Theory*, Special issue on Multiscale Statistical Signal Analysis and Its Applications, v.45, n.3, p.992-1018, 1999.
- Roughan, M.; Veitch, D.; Abry, P., “On-Line Estimation of the Parameters of Long-Range Dependence”. In: Globecom’98, **Proceedings ...**, v. 6, p. 3716-3721, Nov. 1998.
- Shu, Y; Jin, Z; Wang, J.; Yang, O. W., “Prediction-based Admission Control Using FARIMA Models”. In: IEEE ICC’00, **Proceedings ...**, v.3, p.1325-1329, New Orleans, USA, June 2000.
- Vidakovic, B.; Müller, P. “Wavelets for Kids: A Tutorial Introduction”. Technical Report, Duke University, 1998.