Analisador de alarmes de tráfego de redes através de *Wavelets*

Bruno Lopes Dalmazo * **Raul Ceretta Nunes**
Depto de Eletrônica e Computação - CT - Universidade Federal de Santa Maria
E-mail: dalmazo@inf.ufsm.br, ceretta@inf.ufsm.br

Alice J. Kozakevicius
Depto de Matemática - CCNE - Universidade Federal de Santa Maria
97150-900, Campus Camobi, Santa Maria, RS
E-mail: alice.kozakevicius@gmail.com.

RESUMO

1 Introdução

Um Sistema de Detecção de Intrusão (IDS - *Intrusion Detection System*) [1] usa dados coletadas em uma rede para identificar anomalias no seu comportamento, caracterizando um ataque (ou tentativa) em tempo real ou que já tenha ocorrido. Através de uma seqüência de tráfego de rede relacionada com variáveis de dados de um intervalo fixo gera-se uma função matemática que descreve o comportamento desta rede. Todo valor que diferir do padrão determinado por esta função é associado a um alarme. A proposta deste trabalho é apresentar um filtro de alarmes *Wavelets* para um IDS baseado em Séries Temporais, com o objetivo de minimizar o grande número de *falsos positivos* ainda reportados pela maioria dos trabalhos na área [2].

2 Desenvolvimento e resultados

Série Temporal é definida como uma seqüência de dados capturados no tempo que possuem relação com o seu passado [3]. O modelo misto ARIMA (*Autoregressive Integrated Moving Average*), definido pela equação (2.1), é a combinação dos modelos Autoregressivo (AR) e de Médias Móveis (MA) mais um fator de integração (I) para tornar a série estacionária. Através da Série Temporal é estabelecido um padrão de comportamento do tráfego de rede, que ao ser reavaliado, verifica a cada nova amostra se o comportamento da série está dentro do esperado [4].

$$\chi_t = \phi_1\chi_{t-1} + \phi_2\chi_{t-2} + \ldots + \phi_p\chi_{1-p} + \epsilon_t - \theta_1\epsilon_{t-1} - \theta_2\epsilon_{t-2} - \ldots - \theta_q\epsilon_{t-q}. \quad (2.1)$$

Transformada Wavelet é uma técnica capaz de decompor uma função no domínio do tempo em diferentes escalas, sendo ainda possível uma análise da função nos domínios da freqüência e tempo [5]. A *Transformada Wavelet Discreta* decompõe um sinal discreto $Y = \{c_{0,j}\}_{j=0,\ldots,N_0-1}$ com $N_0 = 2^s$ amostras, em uma componente *de escala* $\{c_{J,i}\}$, no nível J mais grosseiro e em várias componentes *de detalhe* $\{\{d_{k,j}\}_{j=0,1,\ldots,N_k-1}\}_{k=1,2,\ldots,J}$, sendo $N_k = 2^{s-k}$ o número de amostras em cada nível de decomposição.

A idéia central por trás do algoritmo rápido para a transformada é a obtenção dos coeficientes de detalhe e de escala por meio de sucessivas convoluções com filtros associados a família de

*Apoio BIC/FAPERGS (Proc.08514069)

Wavelets considerada. A equação (1) nos dá a expansão do sinal Y na base wavelet, sendo a função *Wavelet* denotada por Ψ e a função escala, φ:

$$Y = \sum_{i=0}^{N_J} c_{J,i}\varphi_{j,i} + \sum_{j=J}^{1}\sum_{i=0}^{N_j} d_{j,i}\Psi_{j,i}. \quad (1)$$

Para a análise de alarmes baseada na expansão *Wavelet* do sinal, pode-se assumir que $A[t]$ é composto por duas componentes: $A[t] = L[t] + r_t$, com $L[t]$, a função que está sendo procurada e r_t, um ruído gaussiano residual. $L[t]$ representa o valor correto do alarme no tempo t. A filtragem de um dado sinal consiste na sequência de operações sobre os coeficientes da expansão wavelet do sinal: $\widetilde{L} = W^{-1}Thr(WY)$, sendo $\widetilde{L}$ a estimativa para a componente $L[t]$. W e W^{-1} denotam as transformadas *Direta* e *Inversa*, respectivamente. Thr é o operador de truncamento dos detalhes. Consideramos a transformada wavelet de *Haar* [5] e operador de truncamento com estratégia de *Hard Thresholding* [6] que torna zero todos os valores dos coeficientes menores que um determinado valor de *Threshold*, λ. Esta operação evidencia os detalhes, que concentram as variações mais significativas. Além disso, o sinal é analisado através de um sistema de janelas deslizantes com 256 valores.

Neste trabalho compararam-se os resultados entre o uso do IDS baseado em Séries Temporais e o uso do IDS com filtros *Wavelet*, destacando-se as seguintes observações: a partir da implantação do filtro de alarmes com *Wavelets* as anomalias detectadas aumentaram de 55,56% para 77,78%. Porém, o grande destaque vai para a queda no número de falsos alarmes gerados, eles diminuíram 87,18% em comparação ao IDS baseado em Séries Temporais.

3 Conclusões

Este trabalho apresenta técnicas de detecção de anomalias através da análise de predições com Séries Temporais, combinada com um método de filtragem da expansão wavelet do sinal. Com isso a produção de alarmes se dá através dos coeficientes *Wavelets* que restaram após o truncamento. Bons resultados foram obtidos com o uso da filtragem *Wavelet* como forma de gerar alarmes, dando destaque para a grande redução dos falsos alarmes gerados pelo IDS, em comparação aos resultados obtidos utilizando somente Séries Temporais.

Palavras-chave: *Séries temporais, Transformada de Haar, alarmes*

Referências

[1] R. A. Kemmerer and G. Vigna, Intrusion detection: a brief history and overview, *Computer*, 35(4) pp. 27-30, 2002.

[2] M. Thottan and C. Ji, Anomaly Detection in IP Networks, *IEEE Transactions on Signal Processing*, 51(8), pp. 2191-2204, 2003.

[3] S. C. Wheelwright and S. Makridakis, "Forecasting Methods for Management", John Wiley & Sons Inc, New York, 1985.

[4] R. C. Nunes and I. Jansch-Pôrto, QoS of Timeout-Based Self-Tuned Failure Detectors: The Effects of the Communication Delay Predictor and the Safety Margin, In: Int. Conf. on Dependable Systems and Networks, IEEE Computer Society, 2004.

[5] S. G. Mallat, A theory for multiresolution signal decomposition: the wavelet representation, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, v.11, pp. 674-693, 1989.

[6] I. M. Johnstone and B. W. Silverman, Wavelet threshold estimators for data with correlated noise, *Journal of the Royal Statistical Society: Series B*, v.59, pp.319-351, 1997.