Filtros de alarmes de anomalias através de *Wavelets*

Bruno Lopes Dalmazo[1], **Tiago Perlin**[1],
Raul Ceretta Nunes[1], **Alice de Jesus Kozakevicius**[1]

[1]Depto de Eletrônica e Computação – Universidade Federal de Santa Maria (UFSM)
Av. Roraima, n 1000 – Bairro Camobi – 97.105-900 – Santa Maria (RS) – Brazil

{dalmazo,ceretta}@inf.ufsm.br,{perlin81,alice.kozakevicius}@gmail.com

Abstract. *Different anomaly detection methodologies are used in Network Intrusion Detection Systems and recently signal processing techniques have been widely applied for detecting anomalies. But, this methods produce high rate of false positives. In this paper, we propose a method for reduce de number of false positives in theses systems using wavelets. To detect an attack, we used an intrusion detection system based on Time Series and we used wavelet filtering in alarms generated. We selected some characteristics from raw traffic in network to analyse in proposed anomaly detection algorithm. We then evaluate our approach with the 1999 DARPA intrusion detection dataset and the results show an increase of 22% on detection rate and a decrease of 87% on false positives.*

Resumo. *Diferentes metodologias baseadas em anomalias são utilizadas em Sistemas Detectores de Intrusão de Rede, sendo que recentemente técnicas baseadas na análise de sinais têm sido amplamente utilizadas com bons resultados. O problema, no entanto, destas técnicas ainda é o grande número de falsos positivos. Neste trabalho, é proposta uma abordagem utilizando-se a transformada wavelet para redução de falsos alarmes gerados por um detector de intrusão. Para a detecção de ataques, utilizou-se um detector de intrusão baseado em séries temporais e para correção dos falsos alarmes utilizou-se a filtragem de alarmes por wavelets. Foram selecionados alguns descritores de rede a partir do tráfego bruto para a formação de sinais de rede. A abordagem proposta foi testada usando a base de dados do DARPA 99 e resultou taxa de detecção de ataques 22% melhor, reduzindo 87% o número de falsos positivos, tornando assim, os resultados do detector mais confiáveis.*

1. Introdução

O grande crescimento da Internet e a proliferação de serviços nela disponibilizados desperta preocupações quanto a vulnerabilidade dos serviços, os quais potencialmente podem ser alvos de usos inadequados ou indevidos. O aumento da complexidade das redes, causado pela sua expansão, dificulta a detecção automática de comportamentos anômalos que podem interferir no funcionamento normal dos serviços e da própria rede. Um Sistema de Detecção de Intrusões de Rede (NIDS - *Network Intrusion Detection System*) [Kemmerer and Vigna 2002] usa informações coletadas em uma rede para identificar possíveis ataques ou comportamentos inadequados na mesma. No contexto dos NIDS, considera-se como uma intrusão um ataque que tenha sido bem sucedido.

Por meio da análise dos dados coletados da rede, um NIDS pode identificar um ataque em andamento ou que já tenha ocorrido e gerar uma resposta adequada ao evento detectado. Sua arquitetura normalmente é constituída por três componentes principais: coletor, analisador e atuador. O coletor é responsável pela aquisição dos dados de tráfego de rede de uma fonte, que pode ser uma base de dados ou, mais normalmente, uma sonda ligada diretamente a uma rede ou segmento de rede. No analisador os dados do tráfego de rede são processados por um algoritmo de análise com o objetivo de identificar um ataque ou comportamento anômalo na rede. O atuador é o que realiza uma ação quando um ataque é detectado, podendo a ação ser uma intervenção automatizada no sistema ou a geração de um alerta para a interpretação humana. O principal desafio no desenvolvimento de um NIDS é encontrar um método que identifique ataques corretamente sem gerar um grande número de falsas detecções (falsos positivos).

Os NIDS diferem uns dos outros principalmente quanto à análise dos dados de rede e são tradicionalmente divididos em dois grupos: os baseados em conhecimento, ou assinaturas, e aqueles baseados em comportamento, ou anomalias. Sistemas detectores de intrusão baseados em assinaturas [SNORT 2008] [BRO 2008] utilizam uma base de dados de assinaturas de ataques conhecidos para detectar, nos dados do tráfego de rede, ataques ocorridos. O principal problema de um NIDS baseado em assinaturas é a dificuldade em reconhecer ataques novos ou mutantes. Por outro lado, sistemas de detecção de intrusão baseados em anomalias [Thottan and Ji 2003] [Kruegel and Vigna 2003] constroem um perfil do comportamento padrão da rede, usando dados históricos, e interpretam um desvio, ou anomalia, em relação ao perfil como um possível ataque. Uma anomalia de rede pode ocorrer devido a um ataque, falha em algum equipamento, má configuração ou uso indevido de algum recurso de rede. Por identificar comportamentos anômalos, um NIDS baseado em anomalias é capaz de detectar ataques desconhecidos, porém, um problema desta abordagem é o grande número de falsos positivos gerados pelo sistema [Kemmerer and Vigna 2002]. Devido à necessidade de detecção de todo tipo de ataque, incluindo novos e desconhecidos, a pesquisa na área da detecção de anomalias de rede têm recebido atenção nos últimos anos [Thottan and Ji 2003] [Kruegel and Vigna 2003] [Guangmin 2008] [Samaan and Karmouch 2008] [Lu et al. 2008].

Na abordagem por anomalias em um NIDS, dada uma seqüência de tráfego de rede relacionada com variáveis de dados de um intervalo fixo, gera-se uma função que descreve o comportamento desta rede. Esta função é então utilizada para realizar análises de comportamento e gerar alarmes correspondentes a eventos anômalos na rede. Técnicas derivadas da Análise de Sinais foram propostas para modelagem do tráfego de rede em NIDSs baseados em anomalias [Barford et al. 2002] [Thottan and Ji 2003], como técnicas baseadas em modelos de séries temporais [Wheelwright and Makridakis 1985] e *wavelets* [Mallat 1989]. Em [Thottan and Ji 2003] foi utilizado um modelo de série temporal autoregressivo (AR) para detecção de mudanças abruptas de tráfego e em [Lu et al. 2008] foram usadas *wavelets* para modelagem do tráfego de rede. Embora haja uma vasta bibliografia na área da detecção de anomalias de rede, a incidência de falsos positivos em NIDS baseados em anomalias ainda é um fator limitante para adoção desta abordagem.

Na detecção de anomalias de rede podem ser utilizadas diversas técnicas, como: técnicas de análise estatísticas [Samaan and Karmouch 2008] e estatística *bayesiana* [Liu et al. 2008]; técnicas de garimpagem de dados, como algoritmos de

agrupamento [Li and Fang 2007] e lógica *fuzzy* [Yao et al. 2006]; técnicas de inteligência artificial, como sistemas imunológicos artificiais [Guangmin 2008] e algoritmos genéticos [Selvakani and R.S.Rajesh 2007]. Métodos derivados da análise de sinais, também, têm sido propostos para a detecção de anomalias de rede [Barford et al. 2002] [Thottan and Ji 2003], como séries temporais [Wu and Shao 2005] e a transformada *wavelet* tem sido utilizada em vários trabalhos [Barford et al. 2002] [Soule et al. 2005] [Huang et al. 2006] [Gao et al. 2006] [Lu et al. 2008] [Kim and Reddy 2008]. Porém, ainda há o problema do grande número de falsos positivos.

Alarmes falsos podem gerar interpretações incorretas e reconfigurações indevidas, causando prejuízos ao sistema e desacreditando o NIDS. Este trabalho visa melhorar os resultados apresentados em trabalhos anteriores [Lunardi et al. 2008] [Dalmazo et al. 2008] na detecção de intrusões de rede baseada em anomalias pela adição de um módulo para filtragem dos falsos alarmes gerados pelo sistema. Nos trabalhos prévios [Lunardi et al. 2008] [Dalmazo et al. 2008] foram utilizadas séries temporais para análise do tráfego de rede e detecção de intrusões. Esta abordagem mostrou-se adequada para detecção de vários ataques, porém muitos falsos positivos foram observados. Este artigo propõe a utilização de *wavelets* para filtragem dos alarmes gerados pelo Detector de Intrusão Baseado em Séries Temporais (DIBSeT) [Lunardi et al. 2008] com o objetivo de diminuir o número de falsos positivos, que ocorrem sempre que o mesmo gera incorretamente um alarme de ataque, quando na verdade trata-se de variações normais no padrão de tráfego da rede. Como resultado da inclusão do filtro de alarmes com *wavelets* os ataques detectados aumentaram de 55% para 77% e o número de alarmes falsos gerados diminuiu 87%.

Nos trabalhos relacionados baseados em *wavelets*, presentes na literatura, utilizou-se a transformada *wavelets* sozinha [Kruegel and Vigna 2003] [Gao et al. 2006] [Kim and Reddy 2008] ou em conjunto com algum modelo de séries temporais [Lu et al. 2008] basicamente para a modelagem do tráfego de rede e posterior detecção de anomalias através de alguma medida estatística. A utilização de *wavelets* para filtragem de alarmes difere das abordagens utilizadas nos trabalhos relacionados, por tratar-se de uma segunda etapa em um detector baseado em séries temporais, ou seja, analisa-se alarmes gerados e não tráfego de rede para reduzir o número de falsos positivos.

Este artigo está organizado como segue. A seção 2 representa a base teórica dos modelos de séries temporais e *wavelets* na qual se fundamenta este trabalho. Na seção 3 é proposto o NIDS baseado em séries temporais e filtragem de alarmes através de *wavelets*, além de apresentar a sua arquitetura em alto nível. A descrição da base de dados de tráfego de rede utilizada nos testes e a discussão dos resultados estão na seção 4. Por fim, na seção 5 são feitas as conclusões e a discussão de trabalhos futuros.

2. Séries Temporais e *Wavelets*

Nesta seção será apresentada a base teórica matemática na qual se fundamenta o trabalho de detecção de intrusões baseada em séries temporais com filtragem dos alarmes através de *wavelets*.

2.1. Séries Temporais

Uma Série Temporal pode ser definida como uma seqüência de dados capturados no tempo que possuem uma forte relação com o seu passado

[Wheelwright and Makridakis 1985]. Desde muito tempo, análises de séries temporais são utilizadas em diversas áreas, como eletrocardiogramas, bolsa de valores, previsões do tempo [Kim et al. 2006], dentre outros. Esta análise permite obter previsões e elaborar cenários úteis na tomada de decisões. A idéia de utilização para detecção de intrusões é uma abordagem recente, sendo assim requer cuidados na escolha dos modelos a serem utilizados e na sua implementação. A seleção de uma técnica adequada muitas vezes é dificultada devido à necessidade de ser efetuada uma análise do que se deseja como resultado e de como os dados capturados se comportam. Existem diversas técnicas, porém para que se obtenha o resultado esperado é fundamental um estudo prévio de cada caso [Ehlers 2005].

Wheelwright e Makridakis [Wheelwright and Makridakis 1985] especificaram o modelo misto Autoregressivo e de Médias Móveis (ARMA) através da equação (2.1), como sendo a combinação dos modelos Autoregressivo (AR) e o modelo de Médias Móveis (MA).

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

Analisando a equação (2.1) é possível verificar que os modelos ARMA relacionam os valores futuros com as observações passadas, assim como também com os erros passados apurados entre os valores reais e os previstos. Deste modo, o modelo escolhido para ser usado neste trabalho de detecção de intrusões foi o modelo Auto-Regressivo de Médias Móveis Integrado - *Autoregressive Integrated Moving Average* (ARIMA) - para produzir a série temporal dos dados capturados [Lunardi et al. 2008]. Este tipo de modelo é usado para trabalhar com dados aleatórios de uma série estacionária ou não estacionária [Ehlers 2005]. Séries estacionárias representam processos com a variância e covariância que ficam em torno de uma média, isto é, os dados se comportam de forma mais equilibrada. Já séries não estacionárias representam dados que não possuem uma aproximação de valores entre as amostras, podendo variar abruptamente [Tran and Reed 2001]. A capacidade do modelo ARIMA em descrever de maneira satisfatória tanto séries estacionárias como não estacionárias permite construir um método de previsão adaptativo, interessante neste trabalho, por adaptar-se ao tráfego de rede variável.

Com o uso de Séries Temporais, mais especificamente com o modelo ARIMA, é estabelecido um padrão de comportamento do tráfego de rede, e este padrão é avaliado a cada nova amostra para verificar se a série está se comportando como esperado [Nunes 2003]. Caso a série temporal se comporte de forma diferente do esperado, pode indicar o início do acontecimento de um ataque. Esta abordagem foi implementada e apresentou problemas de falsos positivos [Dalmazo et al. 2008]. Como uma solução a este problema, este trabalho implementou uma análise dos alarmes gerados pela série temporal através de *wavelets*, desta forma pode-se diminuir os falsos positivos e gerar resultados mais confiáveis.

2.2. *Wavelets*

Wavelet é uma técnica matemática capaz de decompor uma função no domínio do tempo em diferentes escalas, de modo que seja possível uma análise da função nos domínios da freqüência e tempo [Strang 1993]. A *Transformada de Wavelet* é a representação de uma função por meio de *Wavelets*. A análise através desta técnica é feita pela aplicação sucessiva da *Transformada de Wavelet*, representando a decomposição do sinal original em

diversos componentes localizados no tempo e na freqüência. Cada *wavelet* possui melhor ou pior localização nos domínios da freqüência e do tempo, por isso a análise pode ser feita com diferentes *wavelets*, de acordo com o resultado desejado. Algumas aplicações de *wavelets* são: a identificação de picos em sinais cardíacos [Kozakevicius et al. 2005], compressão de imagens [Usevitch 2001], reconhecimento da fala [Xu et al. 2007] e filtragem de sinal [Wang 2008].

A *Transformada Wavelet* decompõe um sinal dado em coeficientes *wavelets* e uma série de dilatações e translações de uma *wavelet* mãe Ψ [Mallat 1989].

$$Y = \sum_{i=-\infty}^{+\infty} C_{J,i}\varphi_{j,i} + \sum_{j=J}^{1} \sum_{i=-\infty}^{+\infty} d_{j,i}\Psi_{j,i} \text{ (2.2)}$$

Onde $C_{J,i}$ é a representação do sinal no nível mais grosseiro, $d_{j,i}$ são os coeficientes dos detalhes, φ é a função escala e Ψ a função *wavelet*.

A decomposição de um sinal em aproximação e um conjunto de coeficientes de detalhes pode ser feita pela convolução do sinal dado na resolução anterior com filtros G (passa alta) e H (passa baixa) [Mallat 1989].

$$C_{2,j} = \sum_{k=-\infty}^{+\infty} h(2n-k)C_{2,j+1}Y \text{ (2.3)}$$

$$D_{2,j} = \sum_{k=-\infty}^{+\infty} g(2n-k)C_{2,j+1}Y \text{ (2.4)}$$

A próxima seção apresentará a descrição da arquitetura proposta para o detector de intrusões de rede com análise dos dados com séries temporais e filtragem dos falsos alarmes com *wavelets*.

3. Proposta do algoritmo baseado em filtros *wavelets*

Primeiramente, alguns conceitos devem ser salientados para um melhor entendimento do funcionamento do algoritmo de filtragem de alarmes baseado em *wavelets*. Como visto anteriormente, uma Série Temporal é definida como um conjunto de amostragens de uma determinada variável, ordenadas no tempo, normalmente em intervalos eqüidistantes [Wheelwright and Makridakis 1985]. Matematicamente pode ser representada como uma função discreta de Z:

$$Z(t) = \{Z_1, Z_2, Z_3, ...\} \text{ (3.1)}$$

Na Análise de Sinais uma Série Temporal pode ser representada como a soma de uma função determinística dependente do tempo $f(t)$ e um processo estocástico, r_t, ou ruído branco (gaussiano).

$$Z_t = f(t) + r_t \text{ (3.2)}$$

Onde Z_t representa o valor observado da variável Z no instante t. Assume-se que $f(t)$ independe de r_t e r_t independe de t.

A análise baseada em Séries Temporais busca encontrar o modelo, ou o polinômio, que melhor represente a função $f(t)$ e estime os seus parâmetros componentes. Supondo que a função $f(t)$ seja difícil de ser estimada ou não viável do ponto de vista computacional, tem-se um erro na análise causado pelo erro na estimativa do modelo e pelo ruído que restou no modelo.

O preditor ARIMA do DIBSeT utiliza uma janela de análise finita e deslizante, na sua versão original ela foi definida com o tamanho de 200 elementos [Nunes 2003]. Como a janela de análise é truncada, pode-se dizer que dependências de longa duração no tráfego, maiores que a janela de análise, não possam ser modeladas corretamente pelo sistema e que conseqüentemente o ruído gaussiano também não seja devidamente considerado. Considerando a existência de alguns erros na estimativa do modelo ARIMA pelo DIBSeT pode-se assumir a geração de alguns falsos alarmes.

Para a análise de alarmes baseada em *wavelets* pode-se adotar a seguinte definição:

$$A[t] = I[t] + r_t \quad (3.3)$$

Onde A é o sinal amostrado, neste caso os alarmes gerados pelo módulo baseado em Séries Temporais do DIBSeT, $I[t]$ é o valor correto do alarme que esta sendo procurado e r_t um ruído branco gaussiano residual.

A filtragem de um dado sinal usando *wavelet* consiste no cálculo da *Transformada Wavelet Direta* seguido do corte dos valores nos detalhes considerados ruído e cálculo da *Transformada Wavelet Inversa* [Donoho and Johnstone 1995].

$$\widetilde{l} = W^{-1}\overline{W}Y \quad (3.4)$$

Onde $\widetilde{l}$ é a estimativa da função I, W^{-1} é a *Transformada Wavelet Inversa* e $\overline{W}$ é a *Transformada Wavelet Direta* com corte do ruído.

Para a função de corte dos coeficientes pode-se utilizar o *Hard Thresholding* [Donoho and Johnstone 1995] que consiste no corte dos valores menores que um determinado valor de corte t.

$$H(c) = \begin{cases} c, & |c| > t \\ 0, & |c| \leq t \end{cases} \quad (3.5)$$

Onde t representa o valor do *threshold* de corte. Este valor pode ser encontrado utilizando-se a estratégia de *threshold universal* por nível [Donoho and Johnstone 1995] definada como:

$$t = \sqrt{2log n\delta^2} \quad (3.6)$$

Onde δ^2 é estimativa da variância do ruído.

O analisador de alarmes de tráfego de redes através de *wavelets* realiza a filtragem dos alarmes gerados pelo preditor ARIMA do DIBSeT, utilizando a filtragem baseada em *wavelet* como definido por [Johnstone and Silverman 1997], usando *hard thresholding*. O valor do *threshold* é encontrado pela equação do *threshold universal* (3.6). A função *wavelet* escolhida foi a *Haar* [Mallat 1989], por ser computacionalmente mais simples. Foi considerada uma janela de análise deslizante de 256 valores. Realiza-se, também, um corte dos valores dos coeficientes *wavelet* no nível mais grosseiro como forma de evidenciar os detalhes, onde estão as variações, antes da *Transformada Wavelet Inversa*. O algoritmo completo pode ser descrito da seguinte forma:

```
1   INÍCIO
2      Enquanto existir novo_valor faça
3      {
4         Adiciona novo_valor na janela deslizante;
5         Atualiza a janela de análise deslizante;
6         Transformada Wavelet Direta do vetor de análise;
7         Cálculo do valor do threshold;
8         Filtragem do ruído por hard thresholding;
9         Corte do nível mais grosseiro do sinal;
10        Transformada Wavelet Inversa do vetor de análise;
11        Saída do novo nível de alarme correspondente;
12     }
13  FIM
```

Com a adição de uma camada de *Análise de Alarmes* por meio de *wavelts* ao DIBSeT, nomeou-se o novo NIDS como: Detector de Intrusão Baseado em Séries Temporais e filtragem Wavelet (DIBSeT-W). A Figura 1 apresenta a arquitetura final do DIBSeT-W, ao mesmo tempo em que compara-se com a arquitetura do DIBSeT. Qualquer fonte de dados que disponibilize os contadores necessários para a predição feita pelas Séries Temporais podem ser usados na *Entrada* do analizador. A filtragem dos alarmes é feita na camada de *Análise dos Alarmes*, por meio de filtragem com *wavelets*. O sistema utiliza algoritmos para análise em tempo real do tráfego de rede, porém uma análise *offline* também pode ser realizada, com a utilização de uma base de dados de tráfego de rede.

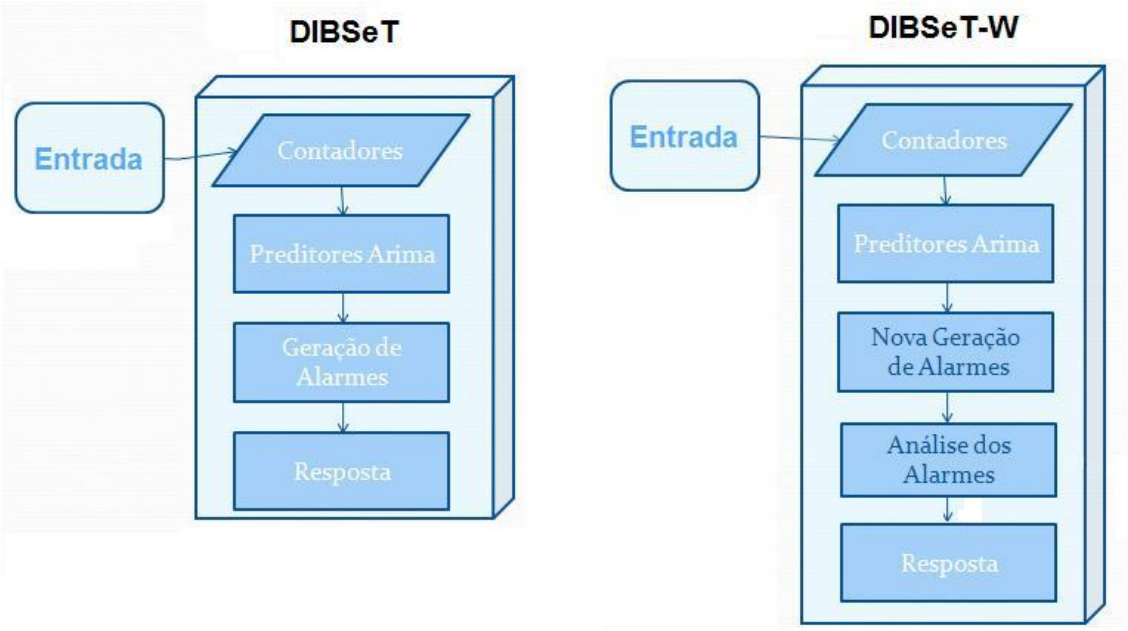

Figura 1. Comparação da arquitetura do DIBSeT com o novo NIDS com analisador de alarmes com *Wavelets*

Após a descrição da arquitetura do NIDS proposto, é preciso uma base de dados de rede conhecida e documentada na literatura para validação da arquitetura por meio de testes. A próxima seção tratará da base de dados utilizada, a descrição, discussão e apresentação dos resultados alcançados.

4. Resultados e discussões

Para possibilitar os testes com o NIDS proposto foi necessário utilizar uma base de dados de tráfego real. A base de dados de tráfego de rede do DARPA [DARPA 1999] foi usada porque possui ataques documentados, condição necessária para a contagem de erros e acertos do sistema.

Os dados originais da DARPA estão no formato bruto, desta forma, foi necessária a extração dos descritores de tráfego e geração dos contadores no formato de *logs*. Foram utilizados descritores do tráfego TCP, ICMP e tráfego total, extraídos em intervalos de 60 segundos, conforme o seguinte esquema:

tcpstat -r outside.tcpdump -o "\%T\n" 60 > w1-out60-tcp.data

O tempo médio de duração de cada ataque, encontrado na base de dados do DARPA, é de aproximadamente 10 segundos, a Figura 2 mostra um ataque de *Denial of Service* [CERT 2008] ocorrido no terceiro dia da segunda semana de treinamento. Este ataque possui cerca de 8 segundos de duração, iniciando no segundo 14536 e sendo finalizado no segundo 14544.

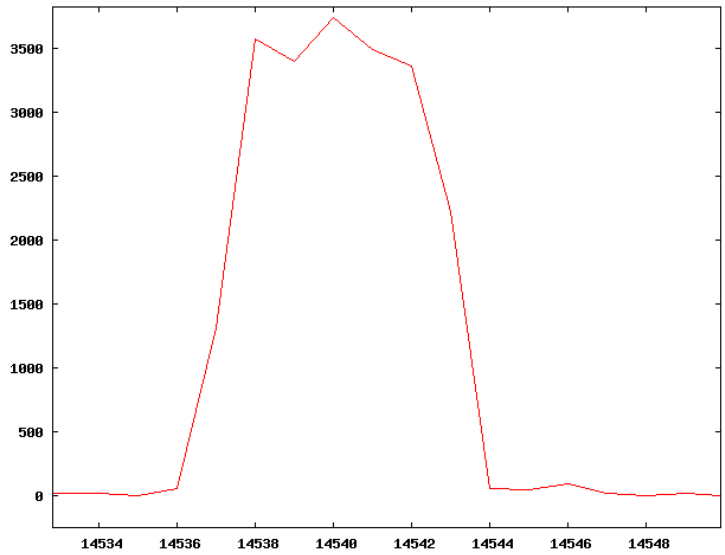

Figura 2. Duração de um ataque em segundos

Nos testes, com o tráfego total de três semanas de treinamento da base de dados do DARPA foi submetido ao analisador de alarmes de tráfego de redes através de *wavelets*. A segunda semana de treinamento possui 36 ataques documentados [DARPA 1999], porém nem todos os ataques geram alterações percepetíveis no padrão da rede. Este trabalho pode identificar ataques que geram picos no tráfego padrão da rede, os ataques enquadrados nesta categoria podem ser detectados. A seguir apresenta-se a descrição de alguns dos ataques documentados no DARPA que podem ser identificados pelo analisador de alarmes de tráfego de redes através de *wavelets*. Foram selecionados 9 ataques da base de dados que se enquadram em alguma das seguintes categorias:

- *Back*: ataque de negação de serviço contra o servidor Apache quando um cliente solicita uma URL que contenha muitas barras invertidas.

- *Dict*: adivinhar senhas de um usuário válido com variantes do nome da conta, ao longo de uma conexão por SSH ou TELNET.
- *Land*: negação de serviço, quando um *host* remoto envia vários pacotes UDP com a mesma origem e destino.
- *MailBomb*: ataque de negação de serviço onde temos um grande envio de mensagens para entregar, com o intuito de travar ou limitar o funcionamento normal de um servidor.
- *Neptune*: ataque *SYN Flood* para negação de um serviço em uma ou mais portas.
- *Smurf*: negação de serviço através de *flood* em resposta a uma requisição ICMP.

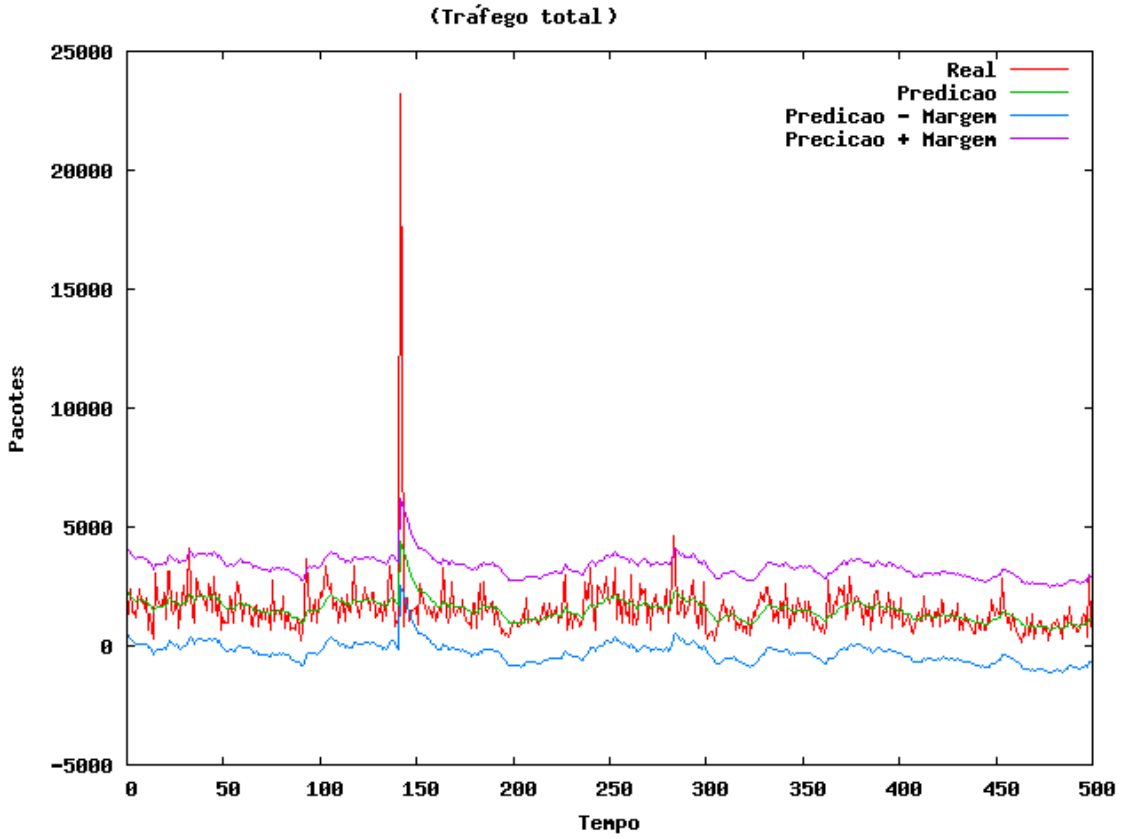

Figura 3. Tráfego total da rede com predição e margens

Primeiramente é analisado o NIDS sem o uso de filtros baseados em *wavelets*, os gráficos detalhados sobre o tráfego real Figura 4 (A) e sobre os alarmes gerados no DIBSeT sem o uso de *wavelets*, conforme mostra a Figura 4 (B). A Figura 3 mostra o tráfego normal da rede com a predição por séries temporais e as margens inferior e superior estipuladas dinamicamente. A escala de tempo contém 19216 amostras, cada amostra de tráfego possui 60 segundos de duração, deste modo, a janela de teste está compreendida em um período de aproximadamente 15 dias ou 3 semanas com 5 dias cada. Os ataques documentados pela base de dados estão na segunda semana, área destacada na Figura 4, assume-se que na primeira e na terceira semanas de tráfego não existam ataques.

Os resultados da geração dos alarmes quando submetidos ao NIDS e filtrados com *wavelets* tornaram-se mais satisfatórios, conforme apresentado na Figura 4 (C) o número de alarmes falsos diminuíram consideravelmente. Considera-se que todos alarmes

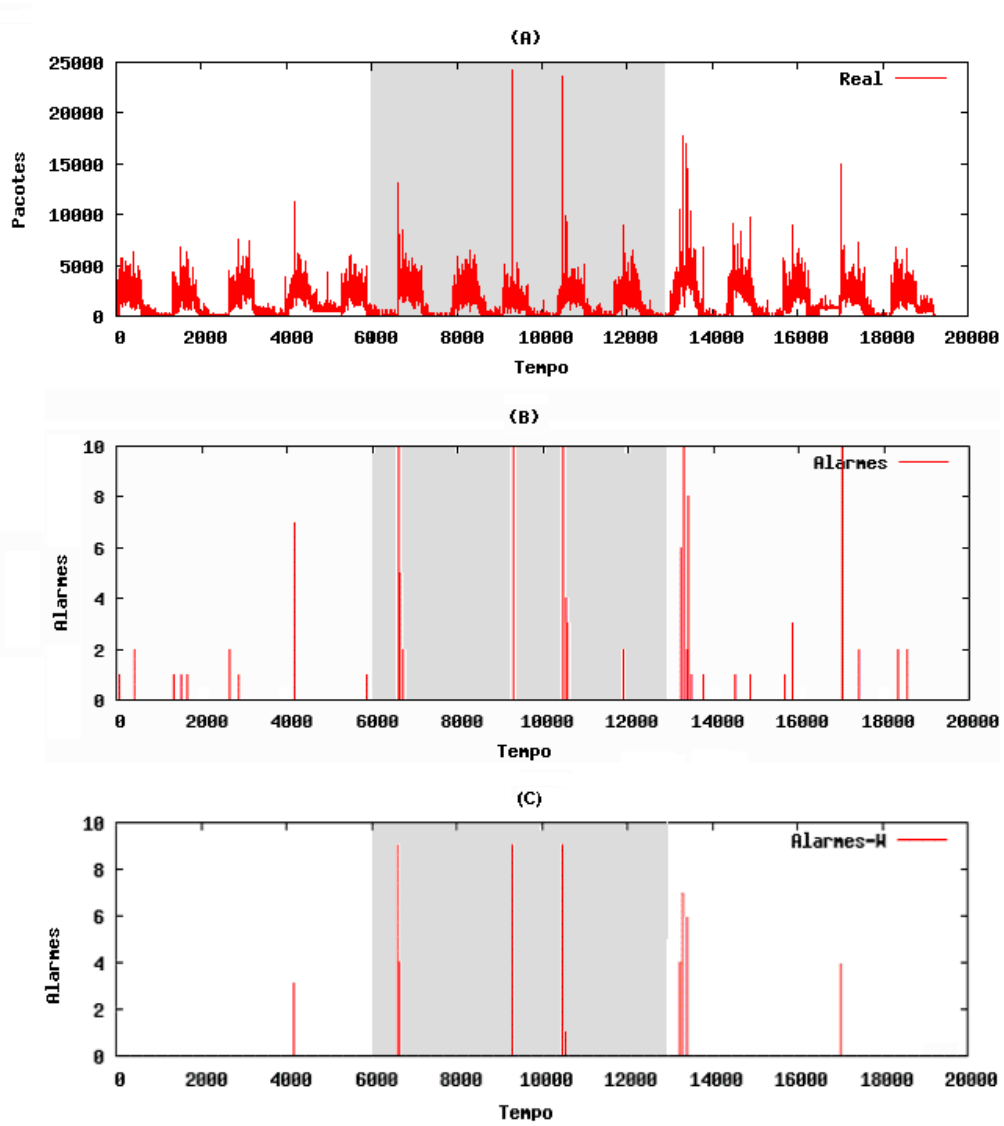

Figura 4. Tráfego real(A), Alarmes(B) sem filtros *wavelets* e Alarmes-W(C) com uso de *wavelets*

observados na primeira e terceira semanas de tráfego, fora da área destacada da Figura 4, tratam-se de falsos positivos, enquanto que os alarmes gerados na segunda semana foram verificados separadamente observando a documentação da base de dados. Na Tabela 1 compara-se o número de detecções e falsas detecções entre o DIBSeT e o DIBSeT-W, foram considerados 9 ataques da base do DARPA.

Tabela 1. Lista resumo dos resultados das 3 semanas de dados do DARPA

	Verdadeiros positivos	Falsos positivos	Falsos negativos
DIBSeT	5	39	4
DIBSeT-W	7	5	2

Na documentação do DARPA foram relatados dois ataques do tipo *Smurf* na base de dados, porém eles não puderam ser detectados quando analisado o tráfego total dos

dados, pois a grandeza das anomalias geradas por este ataque nos pacotes ICMP, se comparadas com a totalidade do tráfego da rede era muito pequena, passando assim despercebida. Entretanto quando analisado apenas o tráfego ICMP, o ataque foi detectado sem dificuldade.

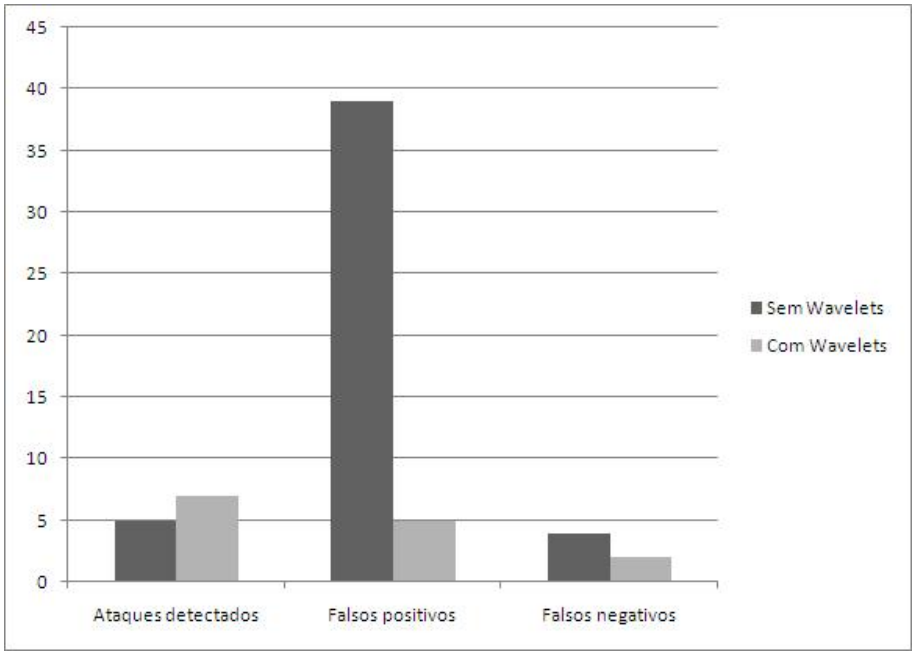

Figura 5. Comparação dos resultados com uso de *wavelets*

Através da análise gráfica da Figura 5, pode-se notar que a partir da implantação do filtro de alarmes com *wavelets* os resultados finais ficaram melhores. Os ataques detectados aumentaram de 55% para 77%. Porém, o grande destaque vai para a queda no número de alarmes falsos gerados, eles diminuíram 87%.

5. Conclusões

O crescimento das redes e a oferta crescente de serviços na web ampliaram a necessidade por sistemas de detecção de intrusão. Um dos principais objetivos dos sistemas de detecção de intrusão de rede é a rápida detecção de intrusões com uma baixa taxa de falsos positivos.

Este trabalho explorou uma técnica de detecção de intrusão por análise de anomalias, na qual foi utilizado a modelagem do tráfego de redes via séries temporais, onde os alarmes gerados são filtrados por um algoritmo baseado em *wavelets*. A experimentação realizada com a base de dados do DARPA possibilitou concluir que com o uso dos filtros *wavelets* obtém-se melhores resultados em termos de falsos positivos e verdadeiros positivos. Com a filtragem por *wavelets* obteve-se melhora significativa (22%) na taxa de detecção de ataques (verdadeiros positivos) e redução significativa no número de falsos alarmes (87%) - falsos positivos. A principal contribuição deste trabalho é apresentar um método para redução dos falsos positivos em um detector de intrusão, por meio da filtragem com *wavelets* dos alarmes gerados.

Com o aperfeiçoamento no método apresentado aqui, estamos pesquisando formas alternativas para aquisição e agregação das variáveis descritivas do tráfego de rede, bem como a correlação dos alarmes gerados em vários fluxos de dados.

Referências

Barford, P., Kline, J., Plonka, D., and Ron, A. (2002). A signal analysis of network traffic anomalies. In *IMW '02: Proceedings of the 2nd ACM SIGCOMM Workshop on Internet measurment*, pages 71–82, New York, NY, USA. ACM.

BRO (2008). Bro Intrusion Detection System. Disponível em: http://www.bro-ids.org/, último acesso em dezembro de 2008.

CERT (2008). Denial-of-Service Attack via ping. Disponível em: http://www.cert.org/advisories/CA-1996-26.html, último acesso em outubro de 2008.

Dalmazo, B. L., Vogt, F., Perlin, T., and Nunes, R. C. (2008). Detecção de intrusão baseado em séries temporais. Gramado, RS, Brasil.

DARPA (1999). Defense Advanced Research Projects Agency. disponível em: http://www.ll.mit.edu/IST/ideval/index.html. Último acesso em outubro de 2008.

Donoho, D. L. and Johnstone, I. M. (1995). De-noising by soft-thresholding.

Ehlers, R. S. (2005). Análise de séries temporais. 2005, 3rd Edição. Departamento de Estatística, Universidade Federal do Paraná, PR.

Gao, J., Hu, G., Yao, X., and Chang, R. (2006). Anomaly detection of network traffic based on wavelet packet. *Asia-Pacific Conference on Communications.*

Guangmin, L. (2008). Modeling unknown web attacks in network anomaly detection. volume 2, pages 112–116.

Huang, C. T., Thareja, S., and Shin, Y. J. (2006). *Wavelet based Real Time Detection of Network Traffic Anomalies*. Securecomm and Workshops, Columbia, SC. Departamant of Computer Sci. e Eng.

Johnstone, I. M. and Silverman, B. W. (1997). Wavelet threshold estimators for data with correlated noise. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 59:319–351.

Kemmerer, R. and Vigna, G. (2002). Intrusion detection: a brief history and overview. *Computer*, 35(4):27–30.

Kim, D. M., Cho, J. M., Lee, H. S., Jung, H. S., and Kim, J. O. (2006). Prediction of dynamic line rating based on assessment risk by time series weather model. *PMAPS 2006 International Conference on Probabilistic Methods Applied to Power Systems*.

Kim, S. S. and Reddy, A. (2008). Statistical techniques for detecting traffic anomalies through packet header data. *Networking, IEEE/ACM Transactions on*, 16(3):562–575.

Kozakevicius, A., Nunes, R. C., Rodrigues, C. R., and Filho, R. G. (2005). Adaptive ecg filtering and qrs detection using orthogonal wavelet transform. *IASTED International Conference on BioMedical Engineering (BioMed 2005)*.

Kruegel, C. and Vigna, G. (2003). Anomaly detection of web-based attacks. In *CCS '03: Proceedings of the 10th ACM conference on Computer and communications security*, pages 251–261, New York, NY, USA. ACM.

Li, Y. and Fang, B.-X. (2007). A lightweight online network anomaly detection scheme based on data mining methods. pages 340–341.

Liu, T., Qi, A., Hou, Y., and Chang, X. (2008). Method for network anomaly detection based on bayesian statistical model with time slicing. pages 3359–3362.

Lu, W., Tavallaee, M., and Ghorbani, A. (2008). Detecting network anomalies using different wavelet basis functions. *Communication Networks and Services Research Conference, 2008. CNSR 2008. 6th Annual*, pages 149–156.

Lunardi, R., Dalmazo, B. L., Amaral, E., and Nunes, R. C. (2008). Dibset: um detector de intrusão por anomalias baseado em séries temporais. Gramado, RS, Brasil.

Mallat, S. G. (1989). A theory for multiresolution signal decomposition: the wavelet representation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 11:674–693.

Nunes, R. C. (2003). Adaptação dinâmica do timeout de detectores de defeitos através do uso de séries temporais.

Samaan, N. and Karmouch, A. (2008). Network anomaly diagnosis via statistical analysis and evidential reasoning. *Network and Service Management, IEEE Transactions on*, 5(2):65–77.

Selvakani, S. and R.S.Rajesh (2007). Genetic algorithm for framing rules for intrusion detection. *IJCSNS International Journal of Computer Science and Network Security*, 7(11).

SNORT (2008). Snort. Disponível em: http://www.snort.org/, último acesso em dezembro de 2008.

Soule, A., Salamatian, K., and Taft, N. (2005). Combining filtering and statistical methods for anomaly detection. In *IMC '05: Proceedings of the 5th ACM SIGCOMM confe-*

rence on Internet Measurement, pages 31–31, Berkeley, CA, USA. USENIX Association.

Strang, G. (1993). Wavelet transforms versus fourier transforms. *Bulletin of the American Mathematical Society*.

Thottan, M. and Ji, C. (2003). Anomaly detection in ip networks. *IEEE Transactions on Signal Processing*, 51(8).

Tran, N. and Reed, D. A. (2001). Arima time series modeling and forecasting for adaptive i/o prefetching. *ACM 15th International Conference on Supercomputing*.

Usevitch, B. E. (2001). A tutorial on modern lossy wavelet image compression: Foundations of jpeg 2000. *IEEE Signal Processing Magazine*.

Wang, X. (2008). Research on effect of frequency band energy leakage to wavelet denoising. *7th World Congress on Intelligent Control and Automation*.

Wheelwright, S. C. and Makridakis, S. (1985). *Forecasting Methods for Management.* John Wiley & Sons Inc, New York.

Wu, Q. and Shao, Z. (2005). Network anomaly detection using time series analysis. pages 42–42.

Xu, Y., Wang, G., Gu, Y., and Liu, H. (2007). A novel wavelet packet speech enhancement algorithm based on time-frequency threshold. *Second International Conference on Innovative Computing, Information and Control*.

Yao, L., ZhiTang, L., and Shuyu, L. (2006). A fuzzy anomaly detection algorithm for ipv6. pages 67–67.