Detecção de Intrusão Utilizando Análise de Séries Temporais com Modelos ARMAX/GARCH

Igor Forain[1], Adilson E. Guelfi[1,2], Elvis Pontes[2], Anderson Silva[2]

[1]Instituto de Pesquisas Tecnológicas (IPT)
Caixa Postal 05508-901– São Paulo – SP – Brasil

[2]Escola Politécnica
Universidade de São Paulo (USP) – São Paulo, SP – Brasil

iforain79@gmail.com, guelfi@lsi.usp.br, elvis.pontes@usp.br,
anderson.silva@pad.lsi.usp.br

Abstract. *In this paper is proposed a method of intrusion detection based on network anomaly identification using Autoregressive Moving Average Exogenous (ARMAX) and Generalized Autoregressive Conditional Heteroskedasticity (GARCH) statistical models. Experiments with DARPA (1999) intrusion dataset showed during SYN flood denial of service and TCP SYN scanning attacks that the proposed method achieved a detection rate of 100% and false positive rate below 5%.*

Resumo. *O objetivo deste trabalho é propor um método de detecção de intrusão por anomalia no tráfego de pacotes de rede aplicando modelos autoregressivos de média móvel com entradas exógenas (Autoregressive Moving Average Exogenous - ARMAX) e autorregressivos com heteroscedasticidade condicional (Generalized Autoregressive Conditional Heteroskedasticity – GARCH). Em termos experimentais, utilizando as bases de tráfego (dataset) disponibilizadas pela DARPA (1999), durante a análise de ataques de negação de serviço synflood e comportamentos de varredura de redes executados por meio de pacotes TCP SYN, o método proposto neste trabalho apresentou probabilidade de detecção de intrusão próxima a 100% e índice de falsos positivos abaixo de 5%.*

1. Introdução

O crescimento da Internet possibilitou a interconexão de diferentes topologias de redes e o acesso de milhões de usuários de computadores a sistemas geograficamente distribuídos. Acompanhando o crescimento deste meio de comunicação, ocorreu o desenvolvimento das técnicas de ataques aos sistemas responsáveis pelo armazenamento e transmissão das informações que trafegam na Internet, podendo ser citados principalmente os ataques de negação de serviço (Denial of Service – DoS) [Moore et al. 2006]. Para proteger as redes, são utilizadas ferramentas que permitem detectar, identificar e mitigar ameaças. Uma destas ferramentas é o Sistema de Detecção de Intrusão (Intrusion Detection System - IDS) cuja função é emitir alertas para o administrador de rede quando detectada a presença de um ataque.

Conforme detalhado por Axelsson (2000), existem duas abordagens para o IDS: assinatura de comportamento malicioso e identificação de tráfego de rede considerado anormal. A primeira baseia-se na criação de um banco de dados de assinaturas de

ataques a partir do tráfego de rede previamente identificado como originário de uma ação maliciosa, acusando um possível ataque em fluxos de pacotes que se encaixem no comportamento já modelado. A segunda abordagem fundamenta-se na modelagem do tráfego de dados considerado normal, apontando um possível ataque em fluxos de pacotes que indiquem um comportamento diferente do esperado. Quando comparadas, a primeira possui as vantagens da precisão na identificação de ataques conhecidos e menor incidência de falsos positivos, porém é frágil na detecção de novos tipos de ataques. A segunda é capaz de identificar novos ataques, mas possui maior incidência de falsos positivos.

Entre as técnicas utilizadas pelos IDSs baseados na identificação de tráfego de rede anormal, podem ser citadas a inteligência artificial e a inferência estatística [Kabiri and Ghorbani 2005]. A análise estatística de séries temporais provenientes do tráfego de rede objetiva estimar o comportamento normal esperado de uma rede ao longo do tempo, de modo a obter uma base de comparação para detectar possíveis anormalidades futuras [Thottan, Liu and Ji 2010]. A eficiência destas técnicas está diretamente relacionada à característica do tráfego de dados de uma rede no que diz respeito ao tipo de protocolo utilizado e ao comportamento previsto ao longo do tempo. No artigo de Zhou, He e Sun (2006) é realizada uma análise estatística do tráfego de dados de uma rede, concluindo que ele tende a manter algum nível de autocorrelação em um horizonte maior de tempo e autossimilaridade em diferentes janelas temporais, apresentando simultaneamente um comportamento quase estacionário em intervalos mais curtos. O trabalho de Kai, Zhengwei e Bo (2009) corrobora o artigo de Zhou, He e Sun (2006) sobre a não estacionariedade do tráfego de rede ao longo de todo o domínio temporal, e acrescenta conclusões sobre a inexistência de uma única função de distribuição de probabilidade, capaz de modelar o fluxo de pacotes resultante dos principais protocolos de rede empregados na atualidade.

Os ataques de DoS do tipo *synflood* e os comportamentos de varredura pré-ataque por meio de pacote TCP SYN ainda figuram entre os principais ataques executados na Internet [Anstee, Bussiere and Sockride 2012]. Ambos são uma ameaça para a plena disponibilidade de serviços computacionais de empresas privadas e órgãos estatais, sendo constantemente empregados por motivações políticas [Nazario 2009]. Os ataques de *synflood* foram recentemente utilizados em ações de guerra cibernética para indisponibilizar os serviços de redes do País atacado antes do envio de forças convencionais [Saleem and Hassan 2009].

Neste contexto, o objetivo deste trabalho é propor um método de detecção de intrusão por anomalia no tráfego de pacotes de rede que utiliza análise de séries temporais. O método proposto é baseado na aplicação de modelos estatísticos para estimar o comportamento normal do tráfego. Este trabalho restringe-se à aplicação de modelos estatísticos autoregressivos de média móvel com entradas exógenas (Autoregressive Moving Average Exogenous - ARMAX) e de modelos autorregressivos com heteroscedasticidade condicional (Generalized Autoregressive Conditional Heteroskedasticity – GARCH) sobre séries temporais geradas a partir de variáveis resultantes da agregação de pacotes de rede. O método proposto é aplicado para ataques de negação de serviço do tipo *synflooding* e para comportamentos de varredura utilizando pacotes TCP SYN.

Este artigo está organizado da seguinte forma: a seção 2 introduz alguns conceitos teóricos utilizados neste artigo. A seção 3 apresenta alguns trabalhos relacionados à detecção de intrusão por meio de análise de séries temporais e modelos estatísticos. A seção 4 detalha o método de detecção de intrusão proposto neste artigo. A seção 5 apresenta e analisa os resultados obtidos por meio da validação do método proposto. A seção 6 conclui este artigo e apresenta possibilidades de trabalhos futuros.

2. Análise do Synflood e Séries Temporais

Na disponibilidade de um serviço que utilize o protocolo TCP, ao receber a solicitação de uma conexão por um pacote SYN, um servidor aloca memória para esta conexão e responde com um pacote SYN+ACK. Sob um ataque de *synflood*, alguns cenários podem ser observados:

- O pacote de ACK final nunca será recebido pelo servidor, fazendo com que este permaneça enviando pacotes de SYN+ACK até o valor máximo de tempo de retransmissão (Retransmission Timeout - RTO) ou até o valor máximo de tentativas de retransmissão ser atingido;
- Na indisponibilidade do serviço, haverá uma falta de resposta por parte do servidor.

Em condições normais, a diferença entre a quantidade de pacotes SYN e SYN+ACK ao longo do tempo é muito pequena (próxima de zero). Por este motivo, a diferença entre a quantidade desses dois tipos de pacotes é um bom parâmetro para detecção de ataques de *synflood* [Divakaran, Murthy and Gonsalves 2006].

Séries temporais são sequências de observações de uma variável em intervalos regulares de tempo. Dentre os modelos existentes para análise dessas séries, alguns são voltados para o domínio da frequência e outros para o domínio temporal. Entre estes últimos, parte deles, como o modelo ARMA, objetiva estimar o valor médio condicional da série ao longo do tempo e outros, como o GARCH, estimam a variância condicional [Kirchgässner e Wolters 2008].

A palavra heteroscedástico da denominação do modelo GARCH refere-se à variação do desvio padrão condicional de uma série ao longo do tempo. Já o modelo ARMA tem como hipótese, a independência temporal do desvio padrão e variação do valor médio ao longo do tempo. As palavras condicional e autorregressivo indicam dependência e mecanismos de retroalimentação por meio dos quais observações passadas da série temporal influem nos valores futuros do valor médio e da variância da própria série [Enders 2009]. O modelo ARMA é responsável por estimar fenômenos temporais de prazos mais longos, ao contrário do GARCH, capaz de estimar características de prazos mais curtos, como por exemplo, comportamentos de tráfego em rajada [Zhou, He e Sun 2006].

3. Trabalhos Relacionados

No artigo de Divakaran, Murthy e Gonsalves (2006), o processo de estabelecimento de uma conexão TCP é utilizado para detectar ataques de *synflood*. Uma conexão TCP é estabelecida pela troca de mensagens na seguinte ordem: SYN, SYN+ACK e ACK. Esse trabalho utiliza um esquema de predição linear baseado em series temporais elaboradas a partir da diferença na quantidade de pacotes SYN e SYN+ACK que trafegam em uma conexão. Já em Ranjan, Murthy e Gonsalves (2010) são aplicados

modelos GARCH para detecção daqueles mesmos ataques de *synflood*, e foi concluído que o tempo de resposta e a incidência de falsos positivos são inferiores com o uso de modelos heteroscedásticos quando comparados com os valores obtidos pelos modelos de predição linear. Já em James e Murthy (2011) são utilizadas transformações matemáticas sobre séries temporais para satisfazer critérios de estabilidade e estacionariedade, permitindo o uso de modelos autorregressivos. Nesses três trabalhos é estudado somente o ataque *synflood* com o emprego do mesmo tipo de série temporal baseada na captura de todos os pacotes TCP SYN para calcular a diferença com a quantidade de pacotes SYN+ACK ao longo do tempo. Ranjan, Murthy e Gonsalves (2010) concluíram que o modelo GARCH de primeira ordem é capaz de reagir a mudanças no padrão do tráfego de rede mais rápida e precisamente, sendo eficaz na detecção de ataques de *synflood*.

Em James e Murthy (2011) é realizado um estudo sobre a relevância e aplicabilidade dos modelos autoregressivos (AutoRegressive – AR) e de média móvel (Moving Average - MA) para detecção de ataques de *synflood*. Nesse trabalho, novamente é utilizada uma série temporal baseada no número de conexões TCP semiabertas, por meio da contagem da diferença entre a quantidade de pacotes SYN e SYN+ACK que trafegam em um enlace de rede a cada 10s. Esse estudo baseia-se na necessidade de uma série temporal estacionária para permitir a aplicação de um único modelo de análise ao longo de todo o domínio temporal, evitando assim, os processos de extração de janelas da série e de recálculo dos coeficientes do modelo estatístico.

O trabalho de He, Zhou e Sun (2005) utiliza modelos ARMAX/GARCH para analisar e sintetizar tráfego IP proveniente de redes LAN e WAN. A série temporal prevista por esses modelos (quantidade de bytes trafegados pela rede a cada 10ms) apresentou características de autossimilaridade e variância que correspondiam às do tráfego real. Zhou, He e Sun (2006) propõem um método de predição de tráfego de rede baseado na utilização de modelos estatísticos ARIMA/GARCH para estimar o comportamento do tráfego TCP de uma rede. Partindo da hipótese de que o comportamento do tráfego de rede apresenta dependência temporal de curto e longo prazo, Zhou, He e Sun (2006) usam o modelo ARIMA em conjunto com o modelo GARCH para prever comportamentos de rede que se apresentam em um horizonte maior de tempo.

Sperotto et al. (2009) propõem um Modelo de Estados de Markov para síntese de séries temporais de fluxos de pacotes de rede, proporcionando massa de dados de teste para validação de outros esquemas de detecção de intrusão. Sperotto, Sadre e Pras (2008) analisam o comportamento de séries temporais resultantes da agregação de fluxos de pacotes na presença de ataques de força bruta contra o serviço de terminal seguro (Secure Shell - SSH) e de DoS contra o sistema de nomes de domínios (Domain Name System - DNS). Esses dois ataques podem ser detectados por meio de identificação de anormalidades em variáveis resultantes da agregação de tráfego de rede [Sperotto et al. 2010].

Siris e Papagalou (2006) aplicam algoritmos de limiar adaptativo e de somas cumulativas (Cumulative Sum – CUSUM) para detecção de ataques *synflood*. Nesse trabalho, a série temporal sobre a qual são aplicados os algoritmos é baseada na quantidade de pacotes TCP SYN por intervalo de amostragem de 10 s. A massa de testes utilizada é baseada no tráfego não malicioso disponibilizado por DARPA (1999) e

na amostragem de 14.5 horas de tráfego obtido no enlace de rede que liga a Universidade de Creta ao Greek Research and Technology Network (GRNET). Siris e Papagalou (2006) concluem que o algoritmo de CUSUM apresenta melhores resultados do que o algoritmo de limiar adaptativo, apesar de este último ser mais rápido na detecção dos ataques.

O processamento digital de sinais já é utilizado como ferramenta de análise de séries temporais no domínio da frequência para detecção de ataques de intrusão. Entre os trabalhos, pode ser citado o artigo de Zhou e Leang (2003) cujo objetivo é encontrar padrões nos espectros de freqüência, por meio da aplicação da transformada rápida de Fourier (Fast Fourier Transform - FFT) em sinais temporais, que indiquem a presença de ataques de força bruta, DoS, varredura de portas e ataques baseados em dicionários.

Erickson (2008) desenvolveu um trabalho cujo objetivo é encontrar características no domínio da frequência que indiquem a presença de tráfego na rede proveniente dos seguintes ataques: Tabela de Processos, Dicionário e Teardrop. Primeiramente, este estudo baseia-se na criação de dois sinais: um a partir do intervalo de tempo entre a chegada dos pacotes e o outro originado do tamanho do conteúdo de cada pacote. Para superar as deficiências do uso da FFT, Erickson (2008) utiliza a Transformada de Fourier de Lomb-Scargle, que permite a geração de espectros para séries temporais incompletas. Em Lu e Ghorbani (2009), é proposto o uso de funções wavelets para detectar tráfego de rede que caracterize comportamento malicioso. Nesse trabalho, são gerados quinze sinais temporais a partir da conversão do tráfego de pacotes para o formato de fluxos de dados.

Tabela 1. Comparação dos Trabalhos

Trabalho	Acurácia	Mod. Matemático	Tempo de Resposta	Localização	Proposta	Técnica de Captura	Tipo de Experimento
Divakaran, Murthy, Gonsalves (2006)	FN[1]: 0% FP[2]: 7%	Predição linear adaptativa	15.291 s	Borda do roteador de acesso	Método de detecção (*synflood*)	Todo tráfego TCP	Produção / simulação
Ranjan, Murthy e Gonsalves (2010)	FN: 0% FP: 5%	GARCH adaptativo	13.291 s	Borda do roteador de acesso	Método de detecção (*synflood*)	Todo tráfego TCP	Produção / simulação
James e Murthy (2011)	FN: 0% FP: 11%	ARIMA estático	5 s	Borda do roteador de acesso	Método de detecção (*synflood*)	Todo tráfego TCP	Produção / simulação
Siris e Papagalou (2006)	PD[3]: 100% FP: 32%	Limiar adaptativo e CUSUM	100 s	Não informada	Método de detecção (*synflood*)	Todo tráfego TCP	Darpa / Simulação
Zhou, He e Sun (2006)	Não se aplica	ARIMA / GARCH estático	Não se aplica	Borda do roteador de acesso	Método de predição de tráfego	Todo tráfego TCP	Produção / Público
He, Zhou e Sun (2005)	Não se aplica	ARMAX / GARCH estático	Não se aplica	Não se aplica	Método de predição de tráfego	Todo tráfego IP	Produção

OBS: (1) falso negativo, (2) falso positivo, (3) probabilidade de detecção.

Comparando este trabalho com os apresentados na tabela 1, ele também faz uso dos modelos estatísticos ARMAX/GARCH, assim como os trabalhos de Zhou, He e Sun (2006) e He, Zhou e Sun (2005), com a diferença de que esses modelos são usados aqui com fins de detecção de ataques e não somente previsão de tráfego, além de serem aplicados sobre outro tipo de série temporal. Ao contrário dos trabalhos de Divakaran, Murthy, Gonsalves (2006) e Ranjan, Murthy e Gonsalves (2010), os modelos utilizados neste trabalho não são constantemente recalculados, sendo computacionalmente menos intensivos. Este trabalho também usa técnicas de detecção de anormalidade em séries temporais, mas diferentes das apresentadas por Siris e Papagalou (2006). E, diferentemente de James e Murthy (2011), este trabalho utiliza uma variação do modelo ARIMA em conjunto com o GARCH, sendo este último estático. Além disso, este

trabalho realiza experimentos sobre uma base de dados de tráfego de domínio público, disponibilizada por DARPA (1999).

4. Proposta de Detecção de Intrusão com Modelos ARMAX / GARCH

O método proposto neste trabalho atua primeiramente na caracterização do comportamento normal do tráfego de rede, possibilitando identificar ações maliciosas a partir de anormalidades em parâmetros do tráfego. Portanto, o método é constituído de duas partes: caracterização do tráfego normal de rede e detecção do evento de intrusão propriamente dito.

O comportamento do tráfego normal é obtido por meio de duas abordagens: a primeira utilizando um modelo estatístico ARMAX, e a segunda abordagem utilizando modelos ARMA /GARCH. O objetivo é estimar a média e variância condicionais ao longo do tempo das séries geradas a partir do tráfego de rede. Logo, o modelo estatístico empregado possui como hipótese inicial para sua síntese, a existência de tráfego de rede sem quaisquer eventos maliciosos, de forma a representar o comportamento normal da rede na qual o método proposto é utilizado. Além disso, para caracterizar o tráfego normal, este trabalho analisa as séries temporais que representam a quantidade de pacotes TCP SYN e TCP SYN+ACK, em intervalos constantes de tempo T_s.

A detecção do evento de intrusão em si compõe a segunda parte do método, quando o modelo estatístico normal é usado em conjunto com o ferramental de decisão para identificar a presença de ataques. Essa parte decisória do método é executada por meio do emprego de algoritmos de identificação de anormalidades em séries temporais. Este trabalho utiliza o erro absoluto e normalizado para identificar ataques.

4.1. Síntese do Modelo Normal ARMAX

Na primeira abordagem, que consiste na utilização somente do modelo ARMAX, considera-se o sistema em que o método de detecção vai ser utilizado como uma "caixa preta" cuja função de transferência discreta deve ser determinada. Dessa forma, a interface de rede é vista como um sistema que tem como entrada um sinal constituído de pacotes TCP SYN e como saída, um sinal constituído de pacotes TCP SYN+ACK, após um atraso δ, conforme apresentado na figura 01.

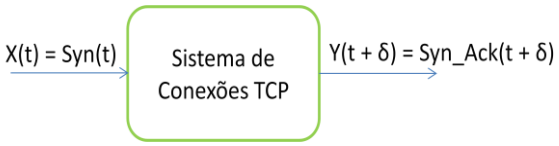

Figura 1. Sistema representado pelo modelo ARMAX.

4.2. Síntese do Modelo Normal ARMA / GARCH

Na segunda abordagem, emprega-se um modelo ARMA, que nada mais é do que um modelo ARMAX sem a componente exógena, em conjunto com o modelo GARCH. O processo de síntese dos modelos estatísticos utilizados pode ser segmentado nos seguintes passos conforme sistematizado na figura 02: captura de tráfego normal, geração de séries temporais, estacionarização, síntese do ARMA, teste de heteroscedasticidade e síntese do ARMA/GARCH. Cada um dos passos é analisado com mais detalhes após a figura 02.

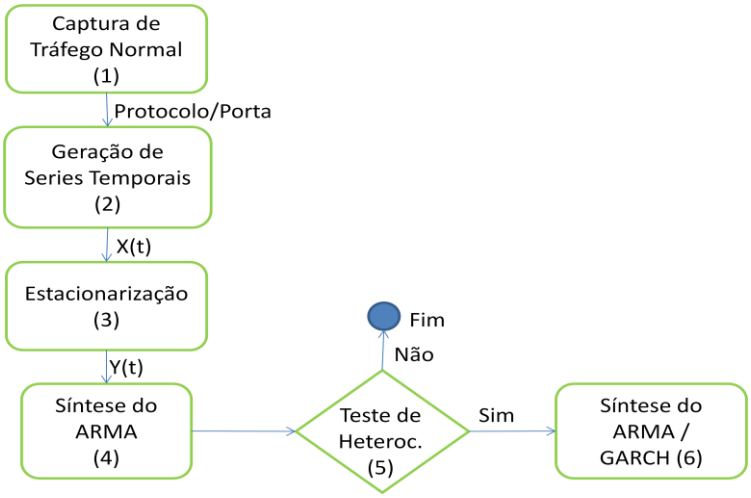

Figura 2. Esquema de síntese do modelo normal ARMA / GARCH.

PASSO 1 – Captura de Tráfego Normal: A captura de tráfego normal consiste na obtenção de tráfego de rede que utiliza o protocolo TCP. Uma vez obtida essa massa de dados e tendo por hipótese a inexistência de tráfego malicioso dentro dela, pode-se utilizá-la para iniciar o processo de geração de séries temporais;

PASSO 2 – Geração de Séries Temporais: Depois de obtida a massa de tráfego de pacotes (passo 01 da fig. 2) isenta de ações maliciosas, é realizada uma amostragem em intervalos constantes de tempo T_s, da quantidade de pacotes com o campo SYN marcado;

PASSO 3 – Estacionarização: Cada uma das séries temporais geradas no passo 2 é testada quanto a sua estacionariedade para identificar a necessidade da aplicação de transformações matemáticas na série original que a levem a uma condição de estabilidade. Como a estacionariedade é uma hipótese para aplicação de modelos ARMA / GARCH na descrição do comportamento temporal, é necessário aplicar testes estatísticos para determinar a presença de componentes não estacionários.

O teste aplicado neste trabalho consiste no cálculo da função de autocorrelação (Autocorrelation Function – ACF), descrita na equação 4.2.1, das séries temporais geradas no passo 2.

$$r_s = \frac{\sum_{t=s+1}^{T}(y_t-\bar{y})(y_{t-s}-\bar{y})}{\sum_{t=1}^{T}(y_t-\bar{y})^2} \qquad (4.2.1)$$

em que r_s é o valor de autocovariância entre termos defasados de s instantes de amostragem, y_t é o valor da série temporal no instante t e $\bar{y}$ é o valor médio. Uma vez aplicada a equação 4.2.1 sobre a série temporal e normalizando os resultados em relação a variância dos dados sob teste, analisa-se o gráfico dos valores de r_s em função do atraso s, gráfico este denominado de correlograma. Para uma série temporal estacionária, o correlograma correspondente converge rapidamente para próximo de zero, apresentando em seguida um comportamento oscilante de centralidade nula. Caso a série temporal estudada não apresente este comportamento, mas sim um valor decrescente e assintótico em relação ao valor nulo, é indicativo de um comportamento não estacionário. Para estas situações é sugerida a aplicação de uma transformação de diferenciação para estabilizar a série temporal [Enders 2009].

PASSO 4 – Síntese do Modelo ARMA: A elaboração de um modelo ARMA (passo 4 da figura 02) que representa o comportamento estatístico da série temporal estudada é dividida em três fases distintas e cronologicamente dependentes: identificação, estimação e diagnóstico. Este processo, constituído destas três fases, deve ser iterado novamente caso os resíduos numéricos calculados na fase de diagnóstico demonstrem a incapacidade do modelo proposto de estimar os valores da série temporal estudada [Morettin e Toloi 2006].

PASSO 5 – Teste de Heteroscedasticidade: Todo o modelo ARMA parte do pressuposto de que a série temporal modelada é estacionária e possui variância condicional constante ao longo do tempo. Porém, as séries sintetizadas no passo 3 podem não apresentar esse comportamento. Além disso, os próprios valores de variância condicional podem seguir um modelo ARMA. Esta forma de descrever a série poderia representar o comportamento do tráfego de pacotes mais próximo da realidade, uma vez que poderia reproduzir as rupturas de variância características do tráfego em rajada ou as mudanças repentinas nos valores das séries de fluxos de dados apresentadas no passo 3 [Enders 2009].

PASSO 6 – Síntese do Modelo ARMA/GARCH: Dependendo dos resultados do teste anterior, apontando a presença de componentes heteroscedásticas na série, é necessário o uso de modelos GARCH para estimar o comportamento esperado da variância condicional das séries temporais elaboradas no Passo 3. A variância condicional do modelo GARCH é definida por meio das seguintes equações:

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^{P} \alpha_i X_{t-i}^2 + \sum_{j=1}^{Q} \beta_j \sigma_{t-j}^2 \quad (4.2.2)$$

$$X_t = \sqrt{h_t}\varepsilon_t \quad (4.2.3)$$

Caso o modelo seja de primeira ordem, tem-se:

$$h_t = \alpha_0 + \alpha_1 X_{t-1}^2 + \beta_1 h_{t-1} \quad (4.2.4)$$

em que h_t é a variância condicional e ε_t é uma sequência de variáveis aleatórias independentes e identicamente distribuídas com média zero e variância um, $\alpha_0 \geq 0$, $\alpha_1 \geq 0$, $\beta_1 \geq 0$, $\alpha_1 + \beta_1 < 1$.

Depois de confirmada a presença de componentes heteroscedásticos na série temporal de pacotes analisada, a estimação dos coeficientes descritos nas equações 4.2.2, 4.2.3 e 4.2.4 é realizada com o uso de funções de máxima verossimilhança condicional [Morettin e Toloi 2006].

4.3. Identificação do Ataque

A identificação do ataque é realizada por meio do cálculo do erro entre os valores e momentos estatísticos da série temporal prevista com a realizada. Na primeira abordagem é usado o erro absoluto por meio do cálculo dos resíduos entre a quantidade de pacotes TCP SYN+ACK estimada pelo modelo ARMAX e a efetivamente realizada. Os erros absolutos são usados com o modelo ARMAX, pois conforme experimentos mostrados na Seção 5, os resíduos numéricos claramente identificam a ocorrência dos ataques. Na segunda abordagem, usam-se erros calculados por meio dos resíduos entre a quantidade de pacotes TCP SYN estimada pelo modelo ARMA/GARCH e a efetivamente realizada, normalizados pela variância condicional estimada. Nessa última

abordagem, optou-se por utilizar o erro normalizado, pois o ARMA/GARCH estima tanto o valor médio quanto a variância condicional.

A medida de erro calculada é comparada com um valor limite a partir do qual seria indicativo de um ataque de intrusão conforme descrito na figura 03. Nesta figura, p é a ordem do modelo estatístico utilizado, X(t) a série temporal de entrada utilizada pelo modelo ARMAX, Y(t) a série realizada, $\hat{Y}(t)$ a série prevista e k é a janela de previsão, indicando quantas amostras futuras serão estimadas pelo modelo.

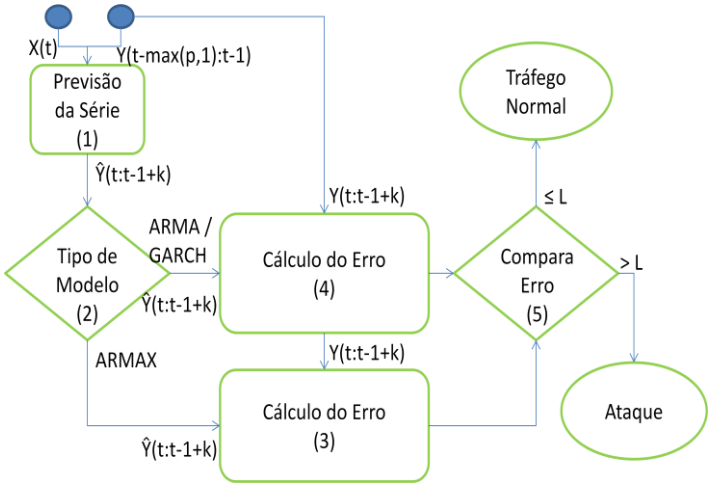

Figura 3. Esquema de identificação do ataque.

5. Experimentos e Validação com Modelos ARMAX / GARCH

Para validar o método de detecção de intrusão proposto neste trabalho optou-se por realizar dois experimentos, usando o tráfego disponibilizado por DARPA (1999). A base de dados distribuída por DARPA (1999) inclui cinco semanas de tráfego de pacotes armazenados no formato *tcpdump*. A primeira e terceira semanas possuem somente tráfego normal e a segunda, quarta e quinta semanas contêm ataques em meio ao tráfego normal. Em DARPA (1999) existem 190 ocorrências de 57 tipos de ataques, dentre os quais 37 são de varredura e 63 de DoS. Entre essas ocorrências, 9 baseiam-se no uso de pacotes TCP SYN [Thomas, Sharma and Balakrishnan 2008].

Os experimentos foram realizados no Matlab R2008a, em uma máquina com processador Intel Core I5-3210m, 4Gb de memória RAM e sistema operacional Windows 7. Os valores calculados nesses experimentos independem da máquina empregada, visto que o Matlab é um ambiente de simulação numérica.

5.1. Experimento I – Uso do modelo ARMAX

Nesse experimento, o primeiro dia da primeira semana de DARPA (1999) foi usado para sintetizar o modelo normal, que em seguida foi utilizado para prever a quantidade de pacotes TCP SYN+ACK nos outros dias em que não ocorreram ataques. Os resultados de previsão são apresentados na Tabela 2. A precisão do modelo na previsão de tráfego é calculada por meio da equação 5.1.1 de normalização:

$$Previsão(\%) = 100 * (1 - norm(\hat{Y}(t) - Y(t))/norm(Y(t) - E(Y(t)) \qquad (5.1.1)$$

na qual $\hat{Y}(t)$ é a série prevista, Y(t) a realizada e E(Y(t)) o valor médio da série realizada.

Tabela 2. Precisão do modelo ARMAX

	S1D1	S1D2	S1D3	S1D4	S1D5	S3D1	S3D1E	S3D2	S3D2E	S3D3	S3D3E	S3D4	S3D5
Previsão(%)	100	99.57	100	100	100	94.90	95.07	95.06	84.52	98.40	94.24	91.04	98.02

OBS: S = semana, D = dia e E = extra.

Durante o teste de detecção, observaram-se os seguintes indicadores numéricos a fim de medir a acurácia e eficiência do método: quantidade de verdadeiros positivos, falsos positivos, falsos negativos e tempo de resposta. Sendo que este último é o intervalo de tempo necessário para identificação do ataque a partir do seu instante inicial. Nesse teste, os erros foram calculados e apresentados graficamente na figura 04 para os cinco dias do tráfego da DARPA (1999) que contêm ataques de *synflood* ou comportamentos de varredura com pacotes TCP SYN.

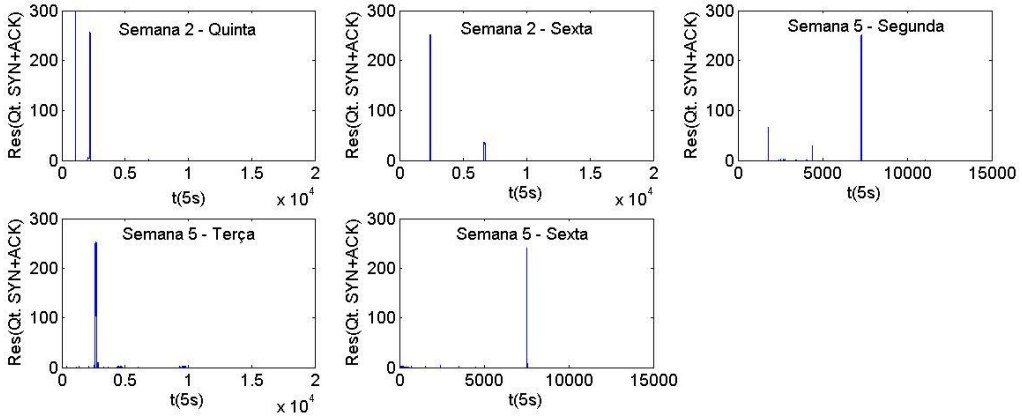

Figura 4. Resíduos do modelo ARMAX.

Observando-se as barras na figura 04, que representam as incidências de erro ao longo do dia analisado, conclui-se a respeito da ocorrência de 2 ataques na semana 2 – quinta, 2 ataques na semana 2 – sexta, 3 ataques na semana 5 – segunda, 1 ataque na semana 5 – terça e 1 ataque na semana 5 - sexta, identificando todos os ataques com pacotes TCP SYN presentes na base de dados da DARPA (1999) com um tempo de resposta inferior aos 5s de amostragem. Estes indicadores representaram 100% de verdadeiros positivos, 0% de falsos negativos e 0% de falsos positivos.

5.2. Experimento II – Uso do modelo ARMA/GARCH

Nesse experimento, o primeiro dia da primeira semana de DARPA (1999) também é utilizado para sintetizar o modelo normal, mas não são usados os critérios de previsão calculados no primeiro experimento visto que não há componentes exógenas. Analisando-se os correlogramas A e C na figura 05, correspondentes a série temporal e a sua forma diferenciada, respectivamente, conclui-se a respeito de sua não estacionariedade, resolvida por meio de uma transformação de diferenciação. Na figura 05, os correlogramas C e D demonstram um forte caráter de média móvel de primeira ordem e o correlograma B indica possíveis componentes heteroscedásticas na série diferenciada, sugerindo, portanto, a utilização de um modelo ARMA(1, 1) / GARCH(1, 1).

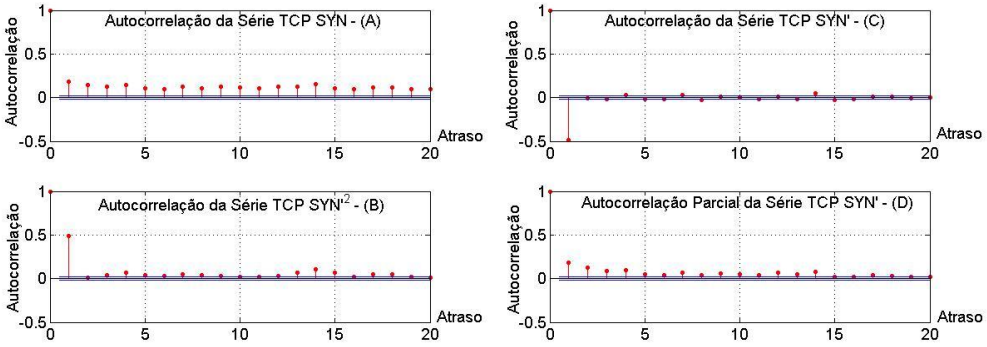

Figura 5. Correlogramas da série TCP SYN, TCP SYN', TCP SYN'^{2}.

Sintetizando-se o modelo normal conforme descrito na Seção 4.2, e estimando o comportamento da série temporal, obtêm-se as previsões de tráfego TCP SYN apresentadas na figura 06, que são comparadas com suas realizações para os dois dias úteis seguintes da primeira semana. Por meio da figura 06, percebe-se que a série temporal de pacotes TCP SYN estimada reproduz os mesmos comportamentos de variância da série real.

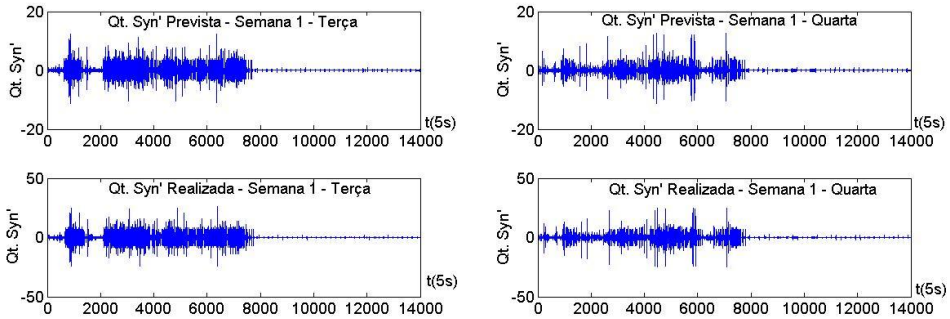

Figura 6. Séries realizadas e previstas pelo ARMA / GARCH.

Durante o teste de detecção, observaram-se os mesmos indicadores numéricos utilizados no experimento I. Nesse teste, foram obtidos os quatro gráficos de erro da figura 07. Esses erros foram calculados e apresentados graficamente na figura 07 para os mesmos cinco dias do tráfego da DARPA (1999) do experimento I. Nos gráficos da figura 07, são notadas as ocorrências de detecção por meio da observação das barras que representam os valores de erro ao longo do tempo.

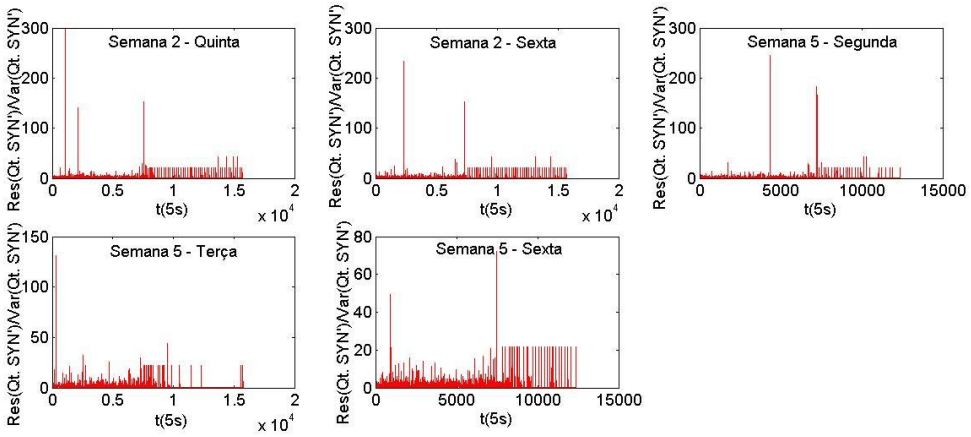

Figura 7. Resíduos das séries normalizados pela variância condicional.

Contabilizando-se os eventos de erro, obtiveram-se os seguintes valores de verdadeiros positivos, falsos positivos, falsos negativos e tempo de resposta, respectivamente: 9, 85, 0 e 5s. Esses valores absolutos representam um índice de falsos positivos inferior a 5%, visto que foram 85 falsas detecções em mais de 20000 amostras corretamente classificadas como não maliciosas e um índice de verdadeiros positivos de 100%, já que os 9 ataques foram identificados. Quanto ao tempo de resposta obtido, 5s, demonstra que o método consegue identificar o ataque com um atraso igual ao tempo de amostragem, em termos de eficiência na identificação do ataque. A tabela 03 apresenta os resultados deste trabalho comparados com dois IDSs baseados em assinatura (Snort e Cisco IDS) e dois baseados em anormalidade (PHAD e ALAD) segundo Thomas, Sharma and Balakrishnan (2008) e também com o método proposto por Siris e Papagalou (2006).

Tabela 3. Comparação de IDSs

IDS / Indicador	Este Trabalho	Snort	PHAD	ALAD	Cisco IDS415	Siris e Papagalou (2006)
PD	100%	100%	100%	< 50%	< 50%	100%
FP	ARMAX: 0% GARCH: < 5%	N.I	N.I	N.I	N.I	0% a 32%
TP	ARMAX: < 5s GARCH: 5s	10 s	N.I	N.I	N.I	10 a 15s

OBS: PD = probabilidade de detecção, FP = falso positivo, TP = tempo de resposta e N.I = não informado.

De acordo com a tabela 03, o método proposto possui probabilidade de detecção igual a do Snort e a do PHAD e superior ao do ALAD e do Cisco IDS415 para os ataques abordados. Apesar de o tempo de resposta deste trabalho ser inferior ao do Snort, devem ser levadas em consideração diferenças no *hardware* utilizado e também o fato de que o Snort estava aplicando um conjunto completo de regras objetivando detectar todos os ataques presentes em DARPA (1999). Quando comparado ao método de Siris e Papagalou (2006), este trabalho apresenta na média, índice de falsos positivos e tempo de detecção inferiores.

5. Conclusões e Trabalhos Futuros

Neste artigo, foi apresentado mais um método de detecção de ataques por meio do emprego de modelos estatísticos ARMAX / GARCH sobre séries temporais geradas a partir da quantidade de pacotes TCP SYN e SYN+ACK que trafega em uma interface de rede a cada 5s. Por meio da utilização do modelo ARMAX, foi possível observar sua capacidade de prever a relação esperada entre a quantidade de pacotes SYN+ACK e SYN com uma precisão próxima de 100%. Como resultado desta capacidade preditiva, o modelo ARMAX é capaz de detectar os ataques baseados no pacote SYN com uma probabilidade de 100%, incorrendo em 0% de falsos positivos. Isso resulta também da própria base de dados disponibilizada pela DARPA (1999), que não emprega uma grande variabilidade de tipos de ataque de *synflood* [Mchugh 2000]. A desvantagem da utilização do modelo ARMAX está na necessidade do emprego de duas séries temporais simultaneamente. Já o modelo GARCH, foi capaz de detectar 100% dos ataques e com uma taxa de falsos positivos próxima a 0%, visto que foram poucas falsas indicações em um universo diário de cerca de 14000 amostras de tráfego. Apesar da aparente desvantagem do modelo GARCH, ele alcança esses resultados somente utilizando a série de pacotes TCP SYN. Como o intervalo de amostragem de 5s é bem maior do que o tempo de processamento requerido pelos equipamentos disponíveis na atualidade, dificilmente execuções do método proposto em dispositivos dedicados alcançariam tempo de resposta diferente do obtido neste trabalho.

Entre trabalhos em andamento e futuros, está o emprego do método proposto em outras bases de dados de ataques, como a apresentada por Shiravi et al. (2012), que reproduz condições mais atuais e diversificadas de ataques de DDoS. O método também será empregado para detecção de outros tipos de ataques, como os de força bruta sobre conexões de SSH. Além disso, o modelo GARCH ainda precisa ser mais explorado em sua capacidade de estimar tráfego, como por exemplo, utilizando outras distribuições estatísticas leptocúrticas (t-Student, por exemplo) para gerar os valores de variância condicional autoregressiva.

Referências

Anstee, D., Bussiere, D. and Sockride, G. (2012) "Worldwide Infrastructure Security Report", Volume VII, Arbor Networks, 2012.

Axelsson, S. (2000) "Intrusion Detection Systems: A Survey and Taxonomy". In: DEPT. OF COMPUTER ENGINEERING, CHALMERS UNIVERSITY OF TECHNOLOGY, 2000, Sweden, 27p.

DARPA (1999) "1999 darpa intrusion detection evaluation data set". In: MIT LINCOLN LABORATORIES, 1999.

Divakaran, D., Murthy, H. and Gonsalves, T. (2006) "Detection of SYN flooding attacks using linear prediction analysis". In: IEEE INTERNATIONAL CONFERENCE ON NETWORKS, 14., 2006, TBD, Singapore, p. 1-6.

Enders, W. (2009) "Applied Econometric Time Series". 3. ed. [S.I.]: Wiley, 2009. 544p.

Erickson, T. J. (2008) "Digital Signal Processing Leveraged for Intrusion Detection". Ohio, USA, 2008. 91p. Dissertação (Mestrado) - Air Force Institute of Technology, Air University, Ohio, USA, 2008.

He D., Sun Z. and Zhou B. (2005) "An ARMAX/GARCH time series model for IP Traffic Trace", in ITC19, Beijing, China, Aug 2005.

James, C. and Murthy, H. (2011) "Time Series Models and its Relevance to Modeling TCP SYN Based DoS Attacks". In: CONFERENCE ON NEXT GENERATION INTERNET, 7., 2011, Kaiserslautern, Germany, p. 1-8.

Kabiri, P. and Ghorbani, A. A. (2005) "Research on Intrusion Detection and Response: A Survey". International Journal of Network Security, [S.I.], [S.I.], V.1, N.2, p. 84-102, 2005.

Kai, H., Zhengwei, Q. and BO, L. (2009) "Network Anomaly Detection Based on Statistical Approach and Time Series Analysis". In: IEEE International Conference on Advanced Information Networking and Applications Workshops, 23., 2009, Bradford, United Kingdom, p. 205-211.

Kirchgässner, G. and Wolters, J. (2008) "Introduction to Modern Time Series Analysis". [S.I.]: Springer, 2008. 284p.

Lu, W. and Ghorbani, A. A. (2009) "Network Anomaly Detection Based on Wavelet Analysis". EURASIP Journal on Advances in Signal Processing, NY, USA, v.1, n.4, p. 1-16, 2009.

Mchugh, J. (2000) "Testing Intrusion Detection Systems: A Critique of the 1998 and 1999 DARPA Intrusion Detection System Evaluations as Performed by Lincoln

Laboratory". ACM Transactions on Information and System Security, v. 3, n. 4, p. 262–294, November 2000.

Moore, D., Shannon, C., Brown, D. J., Voelker, G. M. and Savage, S. (2006) "Inferring Internet Denial-of-Service Activity". ACM Transaction on Computer Systems, [S.I.], [S.I.], V.24, N.2, p. 115-139, 2006.

Morettin, P. A. and Toloi, C. M. C. (2006) Análise de Séries Temporais. 2. ed. [S.I.]: Blucher, 2006. 544p.

Nazario, J. (2009). "Politically motivated denial of service attacks". The Virtual Battlefield: Perspectives on Cyber Warfare, p. 163-181.

Ranjan, N., Murthy, H. A. and Gonsalves, T. A. (2010) "Detection of syn flooding attacks using generalized autoregressive conditional heteroskedasticity (GARCH) modeling technique". IN: NATIONAL CONFERENCE ON COMMUNICATIONS, 16., 2010, I.I.T Madras, India, p. 1-5.

Saleem, M. and Hassan, J. (2009) "Cyber warfare, the truth in a real case". In: Project Report for Information Security Course, Linköping Universitetet, 2009, Sweden, 7p.

Shiravi A., Shiravi H., Tavallaee M. and Ghorbani A. A. (2012) "Toward developing a systematic approach to generate benchmark datasets for intrusion detection", Computers & Security, Volume 31, Issue 3, May 2012, Pages 357 - 374, ISSN 01674048, 10.1016/j.cose.2011.12.012.

Siris, V. and Papagalou, F. (2006) "Application of anomaly detection algorithms for detecting SYN flooding attacks". Computer Communications, Amsterdan, v.29, n.6, p.1433-1442, may. 2006.

Sperotto, A., Sadre, R., Boer, P-T. and Pras, A. (2009) "Hidden Markov Model Modeling of SSH Brute-force Attacks". In: International Workshop on Distributed Systems: Operation and Management, 9., Venice, Italy, volume 5841, p. 164-176.

Sperotto, A., Sadre, R. and Pras, A. (2008) "Anomaly Characterization in Flow-Based Traffic Time Series". In: IEEE International Workshop on IP Operations and Management, 8., 2008, Samos, Greece, p. 15-27.

Sperotto, A., Schaffrath, G., Sadre, R., Morariu, C., Pras, A. and Stiller, B. (2010) "An Overview of IP Flow-based Intrusion Detection". IEEE Communications Surveys & Tutorials, [S.I.], v. 12, n. 3, p. 343-356, 2010.

Thomas, C., Sharma, V. and Balakrishnan, N. (2008) "Usefulness of DARPA dataset for intrusion detection system evaluation". In SPIE Defense and Security Symposium (pp. 69730G-69730G). International Society for Optics and Photonics.

Thottan, M., Liu, G. and Ji, C. (2010) "Anomaly Detection Approaches for Communication Networks". Algorithms for Next Generation Networks. [S.I.]: Springer, 2010. p. 239-261.

Zhou, B., He, D. and Sun, Z. (2006) "Traffic modeling and prediction using ARIMA/GARCH model". Modeling and Simulation Tools for Emerging Telecommunication Networks, [S.I.]: Springer, 2006. p. 101-121.

Zhou, M. and Lang, S-D. (2003) "A Frequency-Based Approach to Intrusion Detection". Systemics, Cybernetics and Informatics, [S.I.], v. 2, n. 3, p. 52-56, 2003.

