Um Modelo de Sistema de Detecção de Anomalias em Redes de Computadores Baseado na Extração de Características Dinâmicas

Marcelo Antonio Righi[1], Raul Ceretta Nunes[1]

[1]Programa de Pós-Graduação em Informática – CT – UFSM
Av. Roraima, 1000, B. Camobi – Santa Maria (RS) – Brasil

marcelo.righi@bol.com.br, ceretta@inf.ufsm.br

Resumo. *A detecção de anomalias baseada em redes vem sendo muito explorada atualmente devido aos inúmeros e persistentes ataques de negação de serviço. Um ponto chave nesta exploração é a técnica para extração de características utilizada para melhorar a eficiência da detecção de ataques. Esse artigo propõe um novo modelo de detecção de anomalia baseado numa técnica de extração de características dinâmicas que utiliza quatro algoritmos em conjunto: a Transformada Wavelet, a Recorrência, o K-Means e o J48. O novo modelo delimita uma Zona Crítica com base num limiar (threshold) derivado do algoritmo K-Means, permitindo uma melhor condição de detecção de anomalias.*
Palavras-Chave: *Detecção de anomalias. Recorrência. K-Means. Árvore de decisão.*
Abstract. *Due to the increasing number of denial of service attacks, the network anomaly detection has been current widely explored. A key point in this exploration is the technique for traffic feature extraction that could to improve the efficiency of detecting attacks. This paper proposes a new model of anomaly detection based on a dynamic feature extraction technique that uses a combination of four algorithms: Wavelet Transform, Recurrence, K-Means and J48. The new model defines a Critical Zone based on a threshold from K-Means algorithm, allowing a better quality on detecting anomalies.*
Keywords: *Intrusion Anomaly detection. Recurrence. K-Means. Decision Tree.*

1. Introdução

Tradicionalmente, detectores de intrusão procuram por comportamentos maliciosos utilizando técnicas baseadas em assinaturas ou anomalias [Mirkovic 2004]. A detecção por assinatura compara o tráfego com uma base de dados de ataques previamente conhecidos (assinaturas), enquanto a detecção por anomalias compara os dados coletados com registros de atividades consideradas normais no sistema.

Em detectores baseados em anomalias, foco deste artigo, as características de rede extraídas do tráfego podem ser estacionárias ou não, ou seja, não variam expressivamente em um determinado período de tempo ou podem oscilar bastante em outro. As características não estacionárias indicam que a observação de tráfego mostra características dinâmicas não lineares, se consideradas a frequência e a recorrência [Grossglauser 1999].

A Construção de novos modelos de extração e alerta, com precisão de detecção e baixa taxa de falsos alarmes, necessita de um sistema de defesa em profundidade, considerando várias camadas de segurança [Northcutt 2003].

Este artigo propõe um novo modelo chamado de Wavelet-Recorrência-Cluster-Árvore da Decisão (WRCA), para extração de características dinâmicas e detecção de

anomalias de rede. As principais contribuições deste trabalho são: (1) um novo modelo para calcular as características dinâmicas multi-escalares do tráfego da rede, utilizando a transformada wavelet e análise de recorrência; (2) um modelo de pré-detecção de anomalias (seleção de tráfego suspeito), com base nessas características dinâmicas e utilizando a clusterização; e (3) um modelo de confirmação de anomalias, com base na árvore da decisão (algoritmo J48).

O restante deste artigo está organizado da seguinte forma. Seção 2 apresenta conceitos fundamentais para o entendimento do artigo. A Seção 3 apresenta a abordagem proposta em detalhes para a implementação do modelo de WRCA. A Seção 4 apresenta detalhes de implementação para o sistema proposto. Na Seção 5 os trabalhos relacionados e na Seção 6 apresenta-se a conclusão do artigo.

2. Conceitos fundamentais

2.1. Detecção de Intrusões de Rede Baseada em Anomalias

Detectores de intrusão baseados em anomalias procuram identificar comportamentos anômalos no tráfego de rede, comparando-o com características de tráfego considerado normal (sem ataque). A principal limitação é a ocorrência de alarmes falsos, dado que nem toda atividade "não usual" (anormal) representa um ataque [Northcutt 2003].

2.2 Algumas Técnicas Utilizadas na Detecção de Anomalias de Rede

A. Transformada Wavelet Discreta
A Transformada Wavelet Discreta (TWD) é um método matemático de análise multi-escalar usada para verificar um sinal em diferentes níveis de resolução. A TWD pode ser implementada através do algoritmo de Mallat [Burrus 1997], que utiliza um banco de filtros digitais com blocos dizimadores acoplados em suas saídas filtradas para decompor o sinal original. São utilizados dois tipos de filtros: um passa-baixa (L) e um passa - alta (H). Os sinais provenientes da filtragem passa-baixa recebem o nome de coeficientes de aproximação (cA), enquanto os sinais provenientes da filtragem passa - alta recebem o nome de coeficientes de detalhes (cD).

B. Clusterização (K-Means)
O algoritmo K-Means [MacQueen 1967], também chamado de K-Médias, realiza o agrupamento (clusterização) de informações de acordo com os próprios dados para gerar as classes (Clusters) e classificar as ocorrências com base nos valores comparados com seus limiares (*threshold*) e no cálculo da distância euclidiana. O algoritmo identifica um centróide para cada classe.

C. Recorrência
A análise de recorrência [Graham 1995] é uma técnica matemática usada para definir sequências, conjuntos, operações ou algoritmos, que generalizam situações a partir de situações particulares (anteriores). A Recorrência [Eckmann 1987] tem surgido como uma técnica de análise não-linear de sistemas dinâmicos. A análise de quantificação de recorrência surgiu como forma de potencializar as avaliações, a partir do desenvolvimento das medidas de quantificação de recorrência [Webber 1994].

D. Árvore de Decisão
Uma árvore de decisão é um instrumento de apoio à tomada de decisão que consiste numa representação gráfica das alternativas disponíveis geradas a partir de uma decisão inicial. Uma das grandes vantagens de uma árvore de decisão é a possibilidade de transformação/decomposição de um problema complexo em diversos subproblemas mais simples [Breiman 1984].

E. Threshold (limiar)
A técnica de limiarização (*threshold*) define os valores de limiares que permitem rotular o tráfego de rede como padrão (normal) ou anômalo, sendo definidos valores de *threshold*, que podem ser fixos [Gao 2006] ou dinâmicos [Kim 2008].

3. Modelo Wavelet-Recorrência-Cluster-Árvore da Decisão (WRCA)

Nesta seção é apresentado o modelo de detecção de anomalias baseado na extração de características dinâmicas denominado **Wavelet-Recorrência-Cluster-Árvore da Decisão** (**WRCA**). A Figura 1 apresenta a arquitetura do modelo e as seções 3.1, 3.2 e 3.3 detalham seus módulos internos.

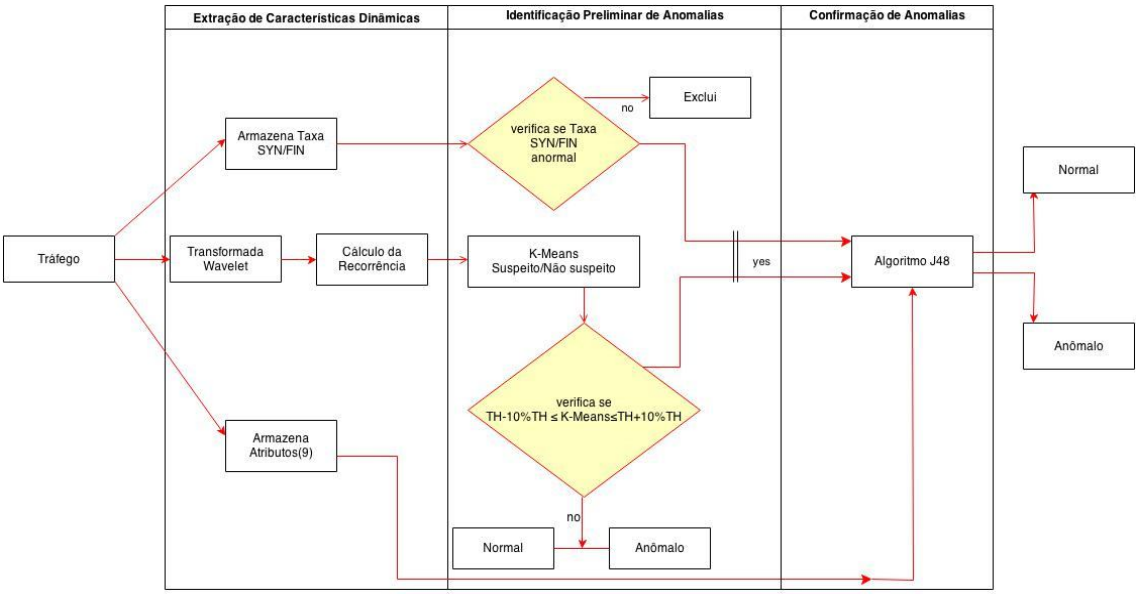

Figura 1. Arquitetura do Modelo WRCA

3.1. Módulo de Extração de Características Dinâmicas

O módulo de extração de características dinâmicas do WRCA é subdividido em três fases. A primeira extrai as características pré-selecionadas: (*i*) Taxa de SYN, Taxa de FIN/RST, Fluxo por minuto (S1), Média de Pacotes por minuto (S2), Média de Bytes por fluxo (S3), Média de Bytes por pacote (S4) e a Média S1/S4 (S5), que serão usadas no Módulo de identificação preliminar de anomalias; (*ii*) nove atributos para o algoritmo J48 (`psizeCL`, `psizeSV`, `pnumCL`, `pnumSV`, `smallpkt`, `dataDIR`, `brecvCL`, `brecvSV`, `Duration`), que só serão utilizadas no módulo de confirmação de anomalias, caso necessário. A segunda fase aplica a transformada wavelet discreta para selecionar o tráfego em diferentes frequências (L e H) (vide seção 4 Passo 1 e 2). A terceira fase computa o cálculo da recorrência e executa a extração das características dinâmicas (vide seção 3.1.1), tanto para o período de treinamento, com tráfego normal, como para o período de análise (detecção).

3.1.1. Cálculo da Equação da Recorrência e das Características Dinâmicas

A. Equação da Recorrência (ER)

Segundo [Yuan 2014] e [Maia 2011], com base em uma série de tráfego x = {xi}, i = 1, 2, ..., n, o estado do tráfego é expresso conforme Equação (1), sendo *m* a dimensão de imersão, *t* o tempo de atraso e $N = n - (m\text{-}1)\,\tau$.

$$Xj = [xj, xj + \tau, xj + (m+1)\tau],\ j = 1,2,\ldots,N \quad (1)$$

Depois de calcular os estados de tráfego, utiliza-se a Equação (2) da Recorrência para analisar os fenômenos de recorrência de cada um deles.

$$Rij = \theta(\varepsilon - \| Xi - Xj \|), \; j = 1,2..., N \quad (2)$$

Na Equação (2), Rij é um elemento da matriz de recorrência, ε é o limiar, Xi é um estado do sistema no espaço de fase m-dimensional, ‖ · ‖ norma, N é o número de estados e Θ(·) é a função definida pela Equação (3).

$$\theta(y) = \begin{cases} 0 & y \le 0 \\ 1 & y \ge 0 \end{cases} \quad (3)$$

Se a distância entre os estados *Xi* e *Xj* é menor do que o limiar (ε), então o valor de *Rij* é 1 e existe um ponto preto em (*i, j*) na equação de recorrência **ER**; caso contrário, o valor de *Rij* é 0 e existe um ponto branco em (*i, j*).

B. Cálculo da Razão da Recorrência, Determinismo e Entropia

Para poder avaliar qualquer série de tráfego após a fase de treinamento as texturas da estrutura **ER** são quantificadas através do cálculo da Razão de Recorrência (RR), Determinismo (DET) e Entropia (ENT), como segue.

1) Razão de Recorrência (RR) - mede a densidade dos pontos de recorrência em **ER**.

$$\mathbf{RR} = \frac{1}{N^2} \sum_{i,j=1}^{N} R_{i,j}$$

2) Determinismo (DET) – mede a relação entre os pontos de recorrência que formam as estruturas de linhas diagonais em ER e todos os pontos de recorrência.

$$\mathbf{DET} = \frac{\sum_{l=l\min}^{N} lP(l)}{\sum_{i,j=1}^{N} R_{i,j}}$$

3) Entropia de Shannon (ENT) - mede a distribuição de frequência dos comprimentos das linhas diagonais.

$$\mathbf{ENT} = \sum_{l=l\min}^{N} p(l)\log_2 p(l) \qquad \mathrm{p(l)} = \frac{P(l)}{\sum_{l=l\min}^{N} P(l)}$$

3.2. Módulo de Identificação Preliminar de Anomalias

O módulo de identificação preliminar de anomalias procura identificar anormalidades no tráfego e indicar tráfegos suspeitos. Para tal, aplica-se o algoritmo K-Means [Yuan 2014] para realizar uma classificação e avalia-se se a maioria dos valores de K-Means estão dentro do intervalo considerado como Zona Crítica (*threshold* = (ε) +/- 10%ε). Em paralelo, verifica-se as taxas de SYN/FIN.

O uso de uma margem de segurança (*threshold* (ε) +/- 10%ε) delimita a Zona Crítica e permite obter melhor eficiência na detecção, através de uma avaliação mais acurada do tráfego limítrofe ao (ε). Em resumo, este módulo executa três atividades:

1) Classificação do tráfego pelo Algoritmo K-Means (suspeito ou não suspeito);
2) Verificação da Taxa SYN/FIN;
3) Verificação se os valores de K-Means ficam no intervalo:

$$\varepsilon - 10\%\varepsilon \le MaioriadeValores(K - Means) \le \varepsilon + 10\%\varepsilon$$

Após análise e classificação do tráfego pelo algoritmo K-Means, é verificado o comportamento da Taxa SYN/FIN e dos valores da classificação K-Means frente a Zona Crítica delimitada. Havendo suspeitas, o tráfego será dirigido ao Módulo de

Confirmação de Anomalias para avaliação pelo algoritmo J48 (confirmação ou não de suspeito ou não suspeito) no Módulo de Confirmação de Anomalias. Caso contrário, se não ocorrerem suspeitas nas atividades 2 ou 3 o tráfego classificado suspeito é denominado ANÔMALO e o tráfego classificado não suspeito é denominado NORMAL.

3.3. Módulo de Confirmação de Anomalias

Esse módulo faz a verificação do tráfego que apresentou características de suspeito ou não suspeito no Módulo de Identificação Preliminar de Anomalias (vide seção 3.2). Ele tem a finalidade de gerar uma árvore de decisão baseada em um conjunto de dados de treinamento, sendo usado para classificar as instâncias no conjunto de teste.

O Módulo adota o algoritmo J48, uma implementação em java do algoritmo C 4.5 [Quinlan 1993], que, segundo [Librelotto 2013], se mostra adequado para os procedimentos envolvendo as variáveis (dados) qualitativas contínuas e discretas presentes nas bases de dados e é é considerado o que apresenta o melhor resultado na montagem de árvores de decisão a partir de um conjunto de dados de treinamento. Para a montagem da árvore, o algoritmo J48 utiliza a abordagem de dividir-para-conquistar, onde um problema complexo é decomposto em subproblemas mais simples.

A aplicação do algoritmo J48 foi realizada considerando o banco de dados de atributos construído por [Dos Santos 2011] (vide Tabela 1), sendo os atributos extraídos no Módulo de Extração de Características Dinâmicas. Esta etapa do modelo WRCA só é aplicada para confirmação, ou não, da suspeita no tráfego sob análise.

4. Implementação do Modelo de Detecção WRCA

A implementação do modelo ainda está em curso e utilizará dados de bases de tráfego, tal como a da base DARPA 1999, bem como dados de coleta realizada na rede da instituição.

Considerando os dados disponíveis, os experimentos estão planejados para comparar, no Módulo de Identificação Preliminar de Anomalias, cinco estatísticas: Fluxo por minuto (S1), Média de Pacotes por minuto (S2), Média de Bytes por fluxo (S3), Média de Bytes por pacote (S4) e a Média S1/S4 (S5). Os dados de treinamento serão confrontados com o fluxo corrente de tráfego.

Tabela 1. Atributos utilizados pelo classificador J48. Adaptado de [Dos Santos 2011]

Atributo	Descrição
`psizeCL (bytes)`	Tamanho médio dos pacotes recebidos pelo cliente.
`psizeSV (bytes)`	Tamanho médio dos pacotes recebidos pelo servidor
`pnumCL`	Número de pacotes recebidos pelo cliente
`pnumSV`	Número de pacotes recebidos pelo servidor
`Smallpkt`	Porcentagem de pacotes pequenos
`dataDIR`	Direção do tráfego
`brecvCL (bytes)`	Total de dados recebidos pelo cliente
`brecvSV (bytes)`	Total de dados recebidos pelo servidor
`Duration`	Diferença em segundos - último pacote e o primeiro

Caso o tráfego seja processado pelo Módulo de Confirmação de Anomalias, os atributos indicados na Tabela 1 serão comparados pelo algoritmo J48, que deverá emitir o resultado final de confirmação ou não do que foi pré-determinado pelo Algoritmo K-Means.

A seguir o detalhamento dos passos do algoritmo resultante:

Entrada: séries temporais de tráfego (cinco estatísticas) $x = \{xi\}, i = 1,2,...,n$, Taxa SYN/FIN e Atributos segundo Tabela 1.
Saída: **tráfego normal ou anômalo**
Passo 1: para uma série de tempo de tráfego *x*, empregar a transformada wavelet para reconstruir a baixa freqüência L e a alta frequência H;
Passo 2: com base na janela deslizante, utilizar o método da Recorrência para calcular as características de recorrência de diferentes séries de tráfego L e H, respectivamente;

$$L = \{FLr\} = \{[\, f_{r;RR}, f_{r;DET}, f_{r;ENT}\,]\}, r = 1,2,...,N_w$$

$$H = \{FHr\} = \{[\, f_{r;RR}, f_{r;DET}, f_{r;ENT}\,]\}, r = 1,2,...,N_w$$

onde r é a r-enésima subsérie, N_w é o número de subsérie, $N_w = \frac{n-W}{W_s} + 1$. Onde:

$$x = \{F_r\} = \{[FLr, FHr]\}$$

Passo 3: para cada série de tráfego (cinco estatísticas), repetir as etapas 1 e 2 e, em seguida, combinar as características dinâmicas da série de cinco tráfegos juntos para descrever os padrões de comportamento do tráfego, de acordo com a seguinte expressão:

$$X = \{Xr\} = \{[F_r^1, F_r^2, F_r^3, F_r^4, F_r^5]\}$$

Passo 4: usar o algoritmo K-Means para classificar cada X_r em diferentes grupos e identificar o tráfego em suspeito ou não suspeito com base na regra de Limite (*Threshold*).
Passo 5: verificar se a Taxa de SYN/FIN está alterada e se o tráfego está compreendido na Zona Crítica.
Passo 6: Caso ocorra uma e outra condição do Passo 5, o tráfego é analisado pelo algoritmo J48, no Módulo de Confirmação de Anomalias, com base nos nove atributos (Tab. 1), extraídos no Módulo de extração de características dinâmicas e comparados com [Dos Santos 2011].
Passo 7: Caso confirme tráfego suspeito, a saída será **ANÔMALO**, do contrário a saída será **NORMAL**; caso confirme tráfego não suspeito, a saída será **NORMAL,** do contrário a saída será **ANÔMALO**. Se não ocorrer ao menos uma condição do Passo 5, o Módulo de Confirmação de Anomalias não será utilizado, os nove atributos serão descartados e a saída será a mesma determinada pelo algoritmo K-Means (**suspeito-ANOMALO, Não suspeito-NORMAL**).

5. Trabalhos Relacionados

Em [Wang 2002] é proposta uma detecção usando a razão entre o número de pacotes TCP SYN e o número de pacotes TCP FIN e RST, mostrando que o normal seria uma relação perto de 1 em um período suficientemente longo, uma vez que a maioria das sessões TCP começa com um SYN e termina com um FIN.

Em [Grossglauser 1999] é sugerido que o tráfego de rede se expõe a propriedades onipresentes de auto-similaridade e dependência de longa duração, ou seja, de correlações em uma ampla gama de escalas de tempo, demonstrando a Recorrência como técnica para detecção de anomalias.

A extração de características dinâmicas é primeiramente descrita em [Yuan 2014], que contribuiu de maneira fundamental para a detecção de anomalias, pois pode

ser utilizada independentemente do fluxo de rede estar elevado ou não no momento do ataque.

A literatura sugere que a combinação de múltiplos classificadores pode melhorar a acurácia da detecção, como demonstra [Chou 2009].

Uma contribuição importante deste trabalho é a delimitação de uma Zona Crítica do *threshold*, que após a fase de testes pode produzir uma confiabilidade maior dos limiares, construindo uma pré-identificação de anomalias e uma análise mais "refinada" caso necessário, sem sobrecarregar o sistema e diminuindo os falsos alarmes. E, também, a combinação de quatro algoritmos, de características dinâmicas e estacionárias em diferentes profundidades ou níveis, melhorando a eficácia do sistema.

6. Considerações finais

Este artigo relata uma abordagem nova na busca da redução do número de falsos alarmes na detecção de intrusão em redes baseadas em anomalias, utilizando diversas técnicas existentes de maneira híbrida, com a extração de características dinâmicas combinadas de forma efetiva e qualitativa durante o tráfego em um determinado espaço de tempo.

O Modelo WRCA mostra-se promissor na detecção de anomalias, pois se caracteriza pela análise do tráfego em níveis de profundidade, que combinados podem melhorar o desempenho dos sistemas atuais, fazendo uma verificação "grosseira" e outra mais "refinada" sem sobrecarregar a memória, pois só utiliza a árvore da decisão (J48) em caso de necessidade. Isto faz com que a maioria das requisições seja determinada pelo algoritmo K-Means (verificação grosseira), sem a necessidade de uma verificação mais profunda (árvore de decisão), só realizada se a combinação de fatores a exigirem.

Referências

Breiman, L., Friedman, J. H., Osshen, R. A. (1984). Classification and regression tress. Belmont: Cahpman&Hall.

Burrus, S.C.; Gopinath, R.A. and Guo (1997). H..Introduction to Wavelets and Wavelet Transforms: A Primer. PrenticeHall.

Chou, T., Fan, J., Fan, S. and Makki, K. (2009). Ensemble of machine learning algorithms for intrusion detection. In Systems, Man and Cybernetic, pages 3976-3980.

Dos Santos, Adriana (2011). Uma Metodologia para Caracterização do Tráfego de Redes de Computadores: Uma Aplicação em Detecção de Anomalias. Disponível em: sid.inpe.br/mtc-m19/2011/02.15.17.55-TDI.

Eckmann, J. P.; Kamphorst S. O.; Ruelle, D.(1987). Recurrence plots of dynamical systems.Europhys. Lett., 56(5):973–977.

Gao, J.(2006) et al. Anomaly Detection of Network Traffic Based on Wavelet Packet. In: Asia-pacific Conference on Communications. *Proceedings.*

Graham, Ronald J., Knuth, Donald E., Patashnik, Oren (1995). Matemática Concreta: Funda-mentos para a Ciência da Computação. Rio de Janeiro. Livros Técnicos e Científicos Editora.

Grossglauser, M.; Bolot, J. C. (1999). On the relevance of long-range dependence in network traffic, *IEEE/ACM Transactions on Networking*, 7(5):629-640.

Kim, S. S.; Reddy, A. L. N.(2008). Statistical techniques for detecting traffic anomalies through packet header data. *IEEE/ACM Transaction on Networking*, Piscataway, NJ, USA: IEEE Press, v. 16, n. 3, p. 562-575.

Librelotto, S. R.; Mozzaquatro, P. M. (2013). Análise dos Algoritmos de Mineração J48 e Apriori Aplicados na Detecção de Indicadores da Qualidade de Vida e Saúde. *Revista Interdisciplinar de Ensino, Pesquisa e Extensão* (RevInt), v.1, n.1, pp.26-37.

MacQueen, J. B. (1967). "Some Methods for Classification and Analysis of Multivariate Observations", Em *Proceedings of the Fifth Symposium on Math, Statistics, and Probability*, pp. 281–297.

Maia, Leonardo P.; Souza, Iberê O. K.(2011). Gráficos de Recorrência de Sistemas Dinâmicos. Disponível em: https://uspdigital.usp.br/siicusp/cdOnlineTrabalho VisualizarResumo?numeroInscricaoTrabalho=60&numeroEdicao=19.

Mirkovic J., P. Reiher (2004). A taxonomy of DDoS attack and DDoS defense mechan- isms, ACM SIGCOMM Computer Communications Review 34 (2) 39.

Northcutt, S. (2003). Novak, J. *Network Intrusion Detection* – Ed. New Riders Publishing.

Quinlan, J. R. (1993). C4.5: Programs for machine learning. Morgan Kaufmann PublishersInc., San Francisco, CA, USA.

Wang H., D. Zhang, K.G. Shin (2002) Detecting SYN flooding attacks, in: Proceedings of IEEE INFOCOM'2002, New York City, NY, pp. 1530–1539.

Webber, C. L.Recurrence Quantification Analysis, v. 13.1. June 2009. Software Package disponível em: < http://homepages.luc.edu/~cwebber/> Acesso em 29 ago 2011.

Yuan J., R. Yuan, X. Chen. (2014). Network Anomaly Detection based on Multi-scale Dynamic Characteristics of Traffic. INT J COMPUT COMMUN, ISSN 1841-9836, 9(1):101-112, February.