Implementação de Funcionalidade Preventiva em um Sistema de Detecção de Intrusão Inteligente

Rodrigo E. Bachinski[1], Victor M. Alves[1], Eduardo Silva[1], Carla L. O. Castanho[1], Sedinei Lima[1], Pablo Espindola[1]

[1]Ciência da Computação – Universidade Regional Integrada do Alto Uruguai e das Missões (URI) Caixa Postal – 97.700-000 – Santiago – RS – Brasil

{bachinski12,victor.ccomp,eduardo.ferreira1983}@gmail.com, carla.castanho@urisantiago.br, sedinei_junior.lima@hotmail.com, pablo.espindola@yahoo.com.br

Abstract. *The intrusion prevention to computing environments is a recurring concern for network administrators. This article presents an automation solution for IIDS (Intelligent Intrusion Detection System), or proposes the transformation of an IIDS in a IIDPS (Intelligent Intrusion Detection Prevention System). The IIDPS besides to detect and report the existence of an attack, it acts proactively to prevent the attack. To actualize this transformation has developed a function that performs the integration of the IISD and the iptables firewall. This function receives as parameter the detection performed by the IISD and then executes a firewall command.*

Resumo. *A prevenção de intrusões à ambientes computacionais é uma recorrente preocupação dos administradores de redes. Este artigo apresenta uma solução de automatização para IIDS (Sistema de Identificação de Intrusão Inteligente), ou seja, propõe a transformação de um IIDS em um IIDPS (Sistema de Detecção e Prevenção de Intrusão Inteligente). O IIDPS além de detectar e notificar a existência de um ataque, ele age proativamente na prevenção do ataque. Para possibilitar esta transformação foi desenvolvida uma função que realiza a integração entre o IIDS e o firewall IPTABLES. Esta função recebe como parâmetro a detecção realizada pelo IIDS e posteriormente executa um comando de firewall.*

1. Introdução

Redes de computadores são, na maioria das vezes, ambientes propensos a ataques e tentativas de invasões. Dalmazo et. al (2009) afirma que com a popularização e custo relativamente baixo de implementação dos recursos computacionais teve-se um aumento expressivo das interconexão entre redes o que tornou, de acordo com Ferreira (2011), a segurança da informação um desafio, ocasionando por consequência um incremento na complexidade da rede e um convite aos usuários mal intencionados a obter dados empresariais ou pessoais de forma ilícita através da invasão destes ambientes.

Sendo assim, os administradores de redes empresariais ou domésticas estão, de um modo geral, continuamente procurando métodos para evitar estes tipos de anomalias. Segundo Lopez (2014) usualmente utiliza-se *firewalls* no que tange a tentativa de bloqueio de possíveis invasões, porém eventualmente, as regras do *firewall* podem ser violadas e a rede ser invadida. Diante dessa situação pode-se utilizar um IDS (Sistema de Detecção de Intrusão) ou um IPS (Sistema de Prevenção de Intrusão) como um recurso adicional de segurança ao sistema. IDS e IPS são sistemas com

funcionamento similar, o primeiro trabalha monitorando a rede para detectar possíveis intrusos, porém, ao contrário dos IPS, ele somente alerta sobre tal evento sem tomar quaisquer medidas para bloquear o ataque. Já o IPS trabalha, a exemplo do IDS, monitorando a rede, a diferença se dá quando o IPS além de detectar o intruso analisa o risco e bloqueia a ação a fim de manter a disponibilidade do serviço sem a necessidade de interferência humana.

Em contrapartida estes sistemas citados (IDS e IPS), tem como principal limitação a ocorrência de alarmes falsos. Segundo Souza (2009), isto ocorre porque nem toda atividade não usual é ilegítima ou representa um ataque, o que ocasiona uma demanda de algo com uma maior eficiência e confiabilidade, uma das formas de satisfazer esta demanda é a utilização de Inteligência Artificial no desenvolvimento do IDS/IPS.

Portanto, nesse cenário existe a necessidade de um sistema de segurança robusto, confiável e com custo-benefício satisfatório. Para satisfazer estes requisitos, este trabalho propõe a transformação do IDS apresentado por Mafra et al. (2008), intitulado de Polvo-IIDS (Sistema de Detecção de Intrusão Inteligente) em um IIDPS (Sistema de Identificação e Prevenção de Intrusão Inteligente), de forma que seja possível aliar a alta confiabilidade do sistema supracitado com a praticidade e autonomia de um IPS, bloqueando eventuais ataques antes mesmo de ocasionar grandes danos ao sistema protegido.

Este artigo busca contribuir na detecção e prevenção de ataques em ambientes suscetíveis a ataques, identificando intrusões e bloqueando-as imediatamente após sua detecção, trazendo maior segurança aos dados de corporações e usuários normais. O artigo está organizado da seguinte maneira: trabalhos relacionados sobre o assunto podem-se ser encontrados na Seção 2; na Seção 3 apresenta-se a arquitetura do Polvo-IIDS. Já na Seção 4 encontra-se a metodologia utilizada para a transformação do Sistema e proposta da nova arquitetura. Os resultados ficam a cargo da Seção 5. Por fim, na Seção 6, são apontadas as considerações finais sobre o modelo e possíveis trabalhos futuros.

2. Trabalhos Relacionados

Nesta Seção do artigo serão apesentados alguns trabalhos relacionados sobre o assunto em questão, descrevendo sucintamente o método utilizado por alguns pesquisadores, principalmente no que tange a automatização de procedimentos em IDS.

Huang et al. (2010) propõem uma integração entre um IDS e um sistema de *firewall*. Os autores justificam esta integração quando deixam evidentes as diferenças entre os dois sistemas, seus propósitos e limitações, mostrando que estes sistemas, de fato podem trabalhar integrados complementando-se um ao outro.

Outra característica ressaltada por Huang et al. (2010) é que um sistema de *firewall* opera geralmente nas camadas de enlace e rede do modelo OSI (O*pen System Interconection*) e não tem a funcionalidade de reconhecer eventuais conteúdos maliciosos que possam trafegar pela rede e ou detectar atividades maliciosas em outras camadas. Já o IDS, tradicionalmente, monitora o tráfego da rede buscando por comportamentos suspeitos ou por ataques previamente identificados por assinaturas.

Em Lei-Jun et al. (2010) encontra-se uma proposta de integração de IDS e sistemas de *firewall*, os autores apresentam suas vantagens, desvantagens e desafios. Neste trabalho, Lei-Jun et al. (2010) sugere que os dois sistemas se adaptem ao ambiente de rede. No entanto, os autores afirmam que para obter este nível de

integração e adaptação será necessário muita investigação e estudo sobre o assunto e não mencionam nenhum sistema em específico e concluem afirmando que este é ainda um modelo teórico.

Carlos et al. (2010) desenvolveu um IDPS distribuído (sensor, servidor, simulador de ataque, simulador de serviço e um executor de comandos de SO e *firewall*). Este sistema trabalha analisando o tráfego da rede e cada vez que o sensor comunica o servidor sobre a incidência de um ataque, o servidor retorna um resposta ativa (comando de *firewall* ou SO), no intuito de inibir o ataque.

Lopez et al. (2014) propõem um IDPS (Sistema de Identificação e Prevenção de Intrusão) denominado *BroFlow* para trabalhar em redes definidas por software *OpenFlow*. As contramedidas que podem ser tomadas por este IDPS ao detectar um ataque são: bloqueio de um fluxo específico ou desvio de um fluxo para uma outra estação. Todas as medidas são temporárias, ou seja, quando uma contramedida é ativada em paralelo é ativo um temporizador, e ao zerar este temporizador é feito novamente a análise do tráfego da rede e se a anomalia deixar de existir, ele desativa a regra criada anteriormente.

SnortFlow é um IDPS baseado em assinaturas proposto por Xing et al. (2013) que tem como funcionalidades a prevenção de intrusão em ambientes de nuvem baseadas em redes *OpenFlow*. Este sistema utiliza-se dos recursos disponíveis no *Snort*, e a cada nova detecção é ativado um controlador que realiza as contramedidas de reconfiguração da rede para evitar a intrusão (redirecionamento de tráfego, isolamento de tráfego, *Deep Packet Inspection*, mudança de endereço MAC, mudança de endereço IP, Bloquear a porta e Quarentena).

Neste trabalho propomos uma arquitetura diferenciada de IDPS, tal proposta utiliza um IIDS *open source* (Polvo-IIDS) trabalhando em paralelo com um *firewall open source* (iptables). Esta ideia segue alinhada com os trabalho de [Xing et al. (2013)], [Lopez et al. (2014)] e [Carlos et al. (2010)]. Estes trabalhos propõem a transformação de um IDS em um IDPS, porém nenhum deles contempla as características do sistema proposto, que é integrar um IIDS com *firewall* iptables.

3. Polvo-IIDS

Nesta Seção do trabalho serão apresentadas as principais características do projeto Polvo-IIDS descrevendo seu comportamento, sua arquitetura e o seu método de funcionamento.

3.1. Descrição do Projeto

O Polvo-IIDS é um sistema de Detecção de Intrusão Inteligente multicamadas baseado em anomalias que tem como principais características a utilização de técnicas de Inteligência Artificial e de aprendizado de máquina em seus classificadores, visando obter uma alta taxa de detecção de verdadeiros positivos e baixa taxa de falsos positivos.

Mafra et al. (2008) relataram que os Sistemas de Detecção de Intrusão Inteligentes encontrados na literatura, normalmente aplicam somente uma Rede Neural na análise dos dados de entrada e nem sempre obtêm taxas de detecção satisfatórias. Dessa forma apresentam um modelo de IDS que utiliza uma rede SOM (Self Organizing Maps) como classificador de ataques e verificadores SVMs (Support Vector Machines) para determinar se os exemplos classificados como ataque pela rede SOM são de fato ataques, ou se são apenas tráfego normal.

3.2. Arquitetura Utilizada

A partir da observação das característica da rede SOM e das SVMs, foi desenvolvido um IDS multicamadas chamado de Polvo-IIDS, um sistema de detecção de intrusão inteligente que realiza a coleta de dados na rede (NIDS) e que aplica técnicas de inteligência artificial e de aprendizado estatístico em sua construção.

Este sistema utiliza uma rede SOM como classificador, e quatro SVMs para determinar se instância observada no tráfego de rede é ou não um ataque. Na Figura 1 pode-se observar a disposição da rede SOM e das SVMs bem como as duas camadas constantes no sistema.

O classificador é composto por uma rede neural de *Kohonen* [Kohonen 1988]. Mafra et al. (2008) afirma que a escolha deste tipo de rede foi motivada pela característica das redes de *kohonen* em aprender padrões de forma automática (sem supervisão), pela facilidade em separar padrões conhecidos (treinados) e pela generalização o na detecção de padrões (detecta variações de padrões conhecidos).

A rede SOM é treinada para classificar os pacotes analisados em quatro categorias: DoS, U2L, Probe ou R2L. Onde a categoria de DoS (Denial of service), aponta os ataques de negação de serviço. A categoria de U2R (User to root), inclui tentativas de acesso a privilégios que somente usuários administradores possuem. Já na categoria de Probe (Probing) estão os ataques que caracterizam tentativas de fraudes. E em R2L (Remote to local) encontra-se os ataques que incluem acessos remotos a recursos locais.

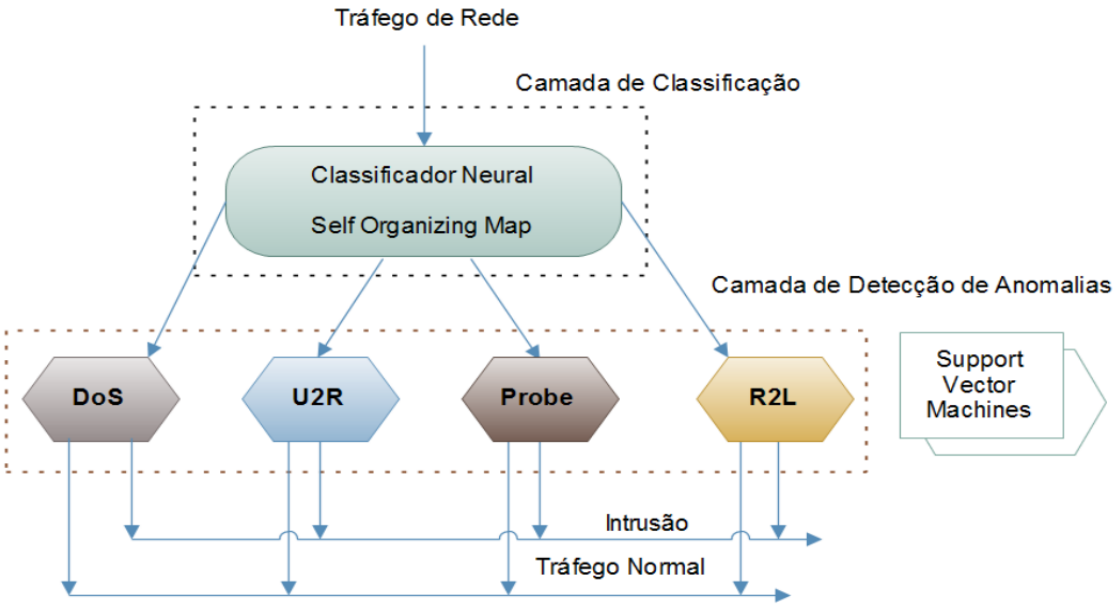

Figura 1: Arquitetura do Polvo-IIDS [Mafra et al. (2008)]

As SVMs são treinadas em específico para cada tipo de ataque, ficando cada SVM responsável somente em identificar se o pacote recebido da primeira camada é um ataque no qual ela é especialista ou somente tráfego normal.

Realizada a classificação pela primeira camada, a saída da Rede SOM é enviada para a camada de Detecção de Anomalias especialista no potencial ataque detectado e esta SVM, que é responsável pela detecção precisa das anomalias, determina se o pacote detectado realmente corresponde a um ataque ou é tráfego normal. Observando a Figura

1 pode-se perceber este fluxo descrito anteriormente.

Em seu funcionamento de acordo com o que foi documentado por Mafra et al. (2008), o Polvo-IIDS analisa os padrões de comportamento da rede, gerando mapeamentos de difusão de pacotes entre *hosts* e realiza a detecção de intrusão baseada nos mapeamentos gerados anteriormente.

4. Proposta de Funcionalidade Preventiva

A proposta de implementação de uma funcionalidade preventiva ao Polvo-IIDS é viabilizada através do desenvolvimento de uma função IPS, função esta que é responsável por interligar o IIDS com *firewall* IPTABLES, provendo assim a imediata intervenção aos ataques detectados. Nesta Seção do trabalho serão detalhadas as características, arquitetura proposta e os tipos de intervenções realizadas pela função IPS.

4.1. Características relevantes

A fim de possibilitar a transformação do Polvo-IIDS em um IIDPS, foi realizada uma integração entre o IIDS e o *firewall IPTABLES* dos Sistemas Operacionais *Linux*, de modo que cada nova detecção gere uma nova regra no *firewall* da rede, possibilitando a imediata intervenção do ataque sem a necessidade da ação humana sobre o sistema.

Para realizar a prevenção dos ataques, desenvolveu-se, utilizando a linguagem de programação Java, uma nova função para o Polvo-IIDS. Quando acionada, a Função IPS recebe como parâmetro o tipo de ataque que foi detectado, o protocolo e a porta atacada e posteriormente compila e aplica uma nova regra no *firewall IPTABLES*.

Após a inclusão da nova regra no *firewall* o sistema exibe uma janela (Figura 4) indicando que a intervenção ocorreu e mostra ao administrador da rede a nova regra inserida em seu *firewall*. Outra funcionalidade é que após a intervenção da função IPS o administrador da rede tem a opção de remover ou até mesmo inserir uma nova regra personalizada no *firewall* da rede. Nesta versão do protótipo desenvolvido faz-se necessária a intervenção do administrador da rede para remover as regras criadas no momento das detecções.

4.2. Arquitetura Proposta

A arquitetura deste projeto ficou distribuída em três etapas: detecção, interpretação e aplicação das contramedidas. Observando a Figura 2 pode-se observar o fluxo dos processos listados.

Figura 2: Fluxo dos processos do sistema proposto

A etapa de Detecção é de responsabilidade do Polvo-IIDS, no qual através de seu mecanismo de leitura analisa o tráfego com a finalidade de detectar ataques. Juntamente com a Função IPS, se fez necessária a inclusão de um novo módulo de

leitura de arquivos ao Polvo-IIDS. Durante a etapa de Detecção quando detectado um ataque, este módulo tem a responsabilidade de extrair as características do ataque (porta e protocolo) do arquivo de testes possibilitando a passagem dos mesmos como parâmetro para a Função IPS, a fim de viabilizar as duas etapas subsequentes da Função IPS.

A função IPS quando ativada pelo IIDS, recebe como parâmetros os dados disponibilizados pelo modulo de leitura implementado, e os interpreta compilando a regra que será aplicada no *firewall* da rede. Finalizada a etapa de interpretação inicia-se imediatamente a última etapa, que é a efetivação da regra compilada pela função IPS através da função *exec(String Command)* disponibilizada pela classe *Runtime.*

4.3. Tipos de Intervenções

A Função IPS proposta trabalha interpretando os dados recebidos do IIDS, quanto mais completos forem os dados transmitidos pelo IIDS mais detalhada e restritiva será a regra. Nesta versão do Polvo-IIDS utilizada, considera-se somente o tipo de ataque, o protocolo e a porta atacada, e a partir destas informações a Função IPS interpreta as informações e aplica-as diretamente no *firewall.* Um exemplo de comando compilado pela função é o seguinte: *iptables -A INPUT -p tcp --dport 23 -j DROP*

Quando o IIDS detectar mais de uma vez o mesmo tipo e ataque, a função IPS compila uma nova regra de bloqueio geral, pois entende-se que regra detalhada criada anteriormente não deu conta de bloquear o incidente notificado pelo IIDS. Um exemplo de regra de bloqueio total é o seguinte: iptables -P INPUT DROP

5. Experimentos e Discussão dos Resultados

Para avaliar o funcionamento foram realizados diversos testes na função implementada. Estes experimentos e os resultados dos mesmos estão dispostos nesta Seção do trabalho.

5.1. Testes e Resultados

Os testes com o protótipo foram executados em uma máquina com 4GB de memória RAM e processador Intel Core i3 M370 de 2.4GHz. Para a realização de testes com o modelo apresentado, foi usado o tráfego KDD Cup 1999 Data disponível na Internet [Stolfo et al. 1999]. Os experimentos seguiram os testes já aplicados no trabalho base [Mafra et al. (2008)] com a inclusão do funcionamento do módulo para análise do comportamento em relação a capacidade de criação/compilação das regras e devida aplicação no SO, sem modificação ou nova análise do POLVO em relação a detecção de intrusões.

```
Chain INPUT (policy ACCEPT)
target     prot opt source               destination
DROP       tcp  --  anywhere             anywhere            tcp dpt:telnet
```

Figura 3: Regra inserida no firewall da rede após a detecção de um ataque U2R

Primeiramente foi iniciado o monitoramento do Polvo-IIDS, com suas Redes Neurais Artificiais já treinadas com os dados do tráfego KDD Cup 1999. Como entrada de análise foi utilizado um arquivo contendo dez entradas com características do ataque ***Probe***. No momento em que o IIDS realizou a detecção do ataque ativou a Função IPS passando como parâmetro o protocolo, tipo de ataque e a porta atacada. A Função IPS por sua vez, interpretou os dados e compilou e inseriu a nova regra no *firewall*.

O segundo cenário simulado foi de um ataque ***U2R,*** a exemplo do cenário

anterior, o IIDS detectou a anomalia e executou a função IPS que realizou o trabalho de interpretação e aplicação da regra. Na Figura 3 é possível verificar a regra inserida no *firewall*. Observando a Figura 4 pode-se observar a tela exibida pela Função IPS após a prevenção de um ataque do tipo ***U2R***.

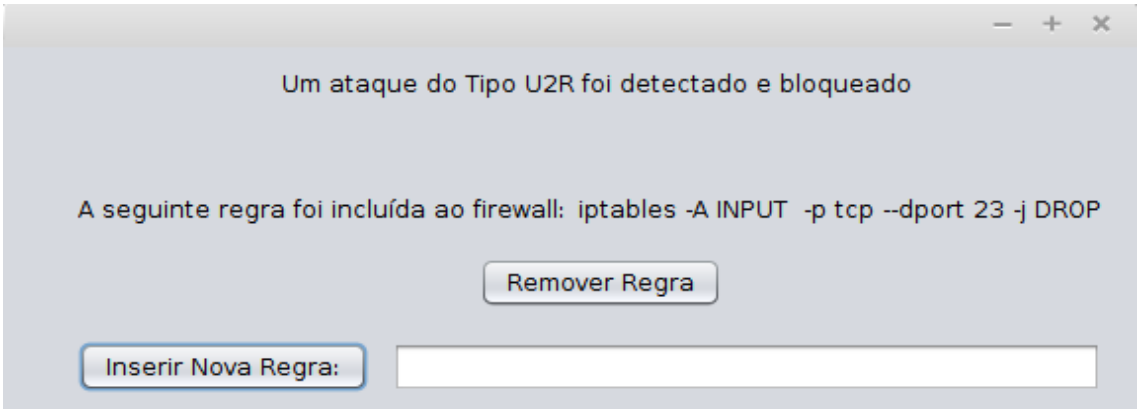

Figura 4: Intervenção gerada após a detecção de um ataque U2L

5.2. Análise dos Resultados

Após executar os testes, observou-se que a Função IPS responde de forma natural aos chamados dos IIDS, conseguindo em todos os testes realizados interpretar e aplicar as regras no *firewall*. Porém, o Polvo-IIDS classifica os ataques de forma genérica, ou seja, não leva em consideração IP de origem e destino ataque por exemplo, o que implica em inclusão de regras genéricas no *firewall*. Fato este, muitas vezes, causa inconvenientes aos usuários autênticos da rede, tendo em vista que as regras aplicadas no *firewall* pela Função IPS visam restringir o acesso a determinados serviços.

Outro aspecto levantado, é que nesta versão do protótipo fica inviável a utilização da Função IPS no bloqueio de ataques do tipo ***DoS,*** pois quando a Função IPS interpreta e compila a regra de bloqueio, ela acaba bloqueando todo o acesso ao serviço atacado, portanto esta prática não seria viável neste tipo de ataque.

6. Conclusão e trabalhos futuros

Este artigo apresentou uma solução de automatização de procedimentos com a integração de um IIDS com *firewall* IPTABES e analisando os resultados obtidos nos testes pode-se perceber que o sistema proposto oferece um elevado grau de confiabilidade a rede, haja vista que eventuais ataques são bloqueados assim que de detectados pelo IIDS. O inconveniente gerado por esta versão da função IPS é o seu modo operação, pois ela cria regras genéricas de bloqueio, fato este pode eventualmente impedir usuários legítimos de acessar a rede.

Para dissolver este inconveniente e como trabalho futuro, propõem-se a modificação da função IPS de modo que este receba como parâmetro o endereço de IP do atacante, a porta atacada e o tipo de ataque, possibilitando a compilação de uma regra de *firewall* mais detalhada, restringindo o acesso somente do atacante sem interferir nos demais usuários autênticos. Outra possibilidade que estará disponível nesta nova versão do sistema proposto, é a utilização desta função IPS em qualquer IDS, desde que o mesmo atenda as especificações dos parâmetros necessários para o funcionamento da função IPS, fato este torna o sistema genérico e aplicável a diversos IDS constantes na literatura.

Referências

Carlos, R. S.; Eduardo, J. P. Sistemas de Detecção e Prevenção de Intrusão. In: I Workshop de Trabalhos de Graduação e Pós-Graduação – DCC/UFJF, 2010.

Dalmazo, L. B.; Perlin, T.; Nunes C. R.; Kozakevicius, J. A. Filtros de alarmes de anomalias através de Wavelets. IX Simpósio Brasileiro em Segurança da Informação e de Sistemas Computacionais, Anais, p. 85-98, 2009.

Ferreira, E. T.; Carrijo, G. A.; Oliveira, R.; Araújo, N. V. S. Intrusion Detection System with Wavelet and Neural Artifical Network Approach for Networks Computers. IEEE Latin America Transactions, vol. 9, no. 5, p. 832-837, 2011.

Huang, X.; Wang, X.; Zhu, S. Study on Intelligent Firewall System Combining Intrusion Detection and Egress Access Control. In: Intelligent System Sesign and Engineering Application (ISDEA), 2010 International Conference on, 2010. v.2, p.456–459.

Kohonen, T. (1988). Self-organized formation of topologically correct feature maps. Journal of the American Society for Information Science and Technology, pages 509–521.

Lei-Jun, L.; Hong, P. A Defense Model Study Based on IDS and Firewall Linkage. In: Iinformation Science and Management Engineering (ISME) , International Conference of, 2010. v.2, p.91–94.

Lopez, M. A.; Figueiredo, U. R.; Lobato, A. G. P.; Duarte, O. C. M. B. BroFlow: Um Sistema Eficiente de Deteccao e Prevencao de Intrusao em Redes Definidas por Software. Wperfomance - XIII Workshop em Desempenho de Sistemas Computacionais e de Comunicação, p. 1919-1932, 2014.

Mafra, Paulo M.; Fraga, Joni d. S.; Moll, Vinícius; Santin, Altair O. Polvo- IIDS: Um Sistema de Detecção de Intrusão Inteligente Baseado em Anomalias. In: Simpósio Brasileiro em Segurança da Informação e de Sistemas Computacionais, (SBSeg 2008), VIII. Anais, p. 61-72, 2008.

Neto, Raimundo P. C.; Martins, Florista; Carneiro, Mateus B. Aplicações de Redes Neurais em Sistema de Detecção de Intrusos para identificar ataques de Botnet em Redes Sem Fio. IV Jornada de Informática do Maranhão, Teresina-Piauí, ISSN: 2358-8861, 2012.

Souza, E. P.; Monteiro, J. A. S. Estudo sobre sistema de Detecção de Intrusão por Anomalias: Uma abordagem utilizando Redes Neurais. In: 14o Workshop de Gerência e Operação de Redes e Serviços, 2009.

Stolfo, J. S., Wei, F., Lee, W., Prodromidis, A., and Chan, P. K. (1999). Kdd cup data - knowledge discovery and data mining competition (1999).

Xing, T.; Huang, D.; Xu, L.; Chung, C.; Khatkar, P. SnortFlow: A OpenFlow-based Intrusion Prevention System in Cloud Environment. Second GENI Research and Educational Experiment Workshop, páginas 89–92, 2013.