Rodrigo Mendonça da Paixão
Lucas Teles Agostinho

Uma abordagem de redes neurais artificiais para sistemas de detecção e prevenção de instrusão

São Paulo – Brasil

2015

Rodrigo Mendonça da Paixão
Lucas Teles Agostinho

Uma abordagem de redes neurais artificiais para sistemas de detecção e prevenção de instrusão

Pré-monografia apresentada na disciplina Trabalho de Conclusão de Curso I, como parte dos requisitos para obtenção do título de Bacharel em Ciência da Computação.

Centro Universitário Senac

Bacharelado em Ciência da Computação

Orientador: Eduardo Heredia

São Paulo – Brasil

2015

Resumo

Palavras-chaves: IDS,Rede,Internet

Lista de abreviaturas e siglas

IDS	Intrusion Detection System
IPS	Intrusion Prevention System
NIDS	Network Intrusion Detection System
RNA	Rede Neural Artificial
SVM	Suport Vector Machine
AI	Artificial Intelligence
ML	Machine Learning

Sumário

1 Introdução

1.1 Contexto

Nos dias atuais pessoas e empresas estão cada vez mais dependentes do uso da internet para realizar suas tarefas. Com o aumento da utilização também temos um aumento de casos de incidentes sobre quebra de segurança. Segundo o CERT (CENTRO DE ESTUDOS, RESPOSTA E TRATAMENTO DE INCIDENTES DE SEGURANçA, 2015) tivemos um crescimento de 197% de incidentes no ano de 2014 relativo a 2013. A necessidade de se proteger contra estes ataques acabou despertando interesse por ferramentas automatizadas para detectar ataques e analisar formas de aprimorar as técnicas atuais para tal.

As ferramentas de detecção de intrusão são chamadas de IDS (Intrusion Detection System), seu trabalho é monitorar as atividades e analisar os eventos em uma rede em busca de anomalias que sugiram uma invasão. Estes não costumam executar qualquer ação para impedir intrusões, sua principal função é alertar os administradores de sistemas que há uma possível violação de segurança, sendo desta forma uma ferramenta passiva. Existem as ferramentas de prevenção de intrusão, que são conhecidas como IPS (Intrusion Prevention System) estas são ferramentas que assim como IDS analisam o trafego e os eventos de uma rede, porem reagem de forma a bloquear o acesso ou atividade maliciosa, senso assim uma ferramenta ativa.

Podemos classificar os IDS da seguinte forma.

Baseado em Host ou rede, onde o primeiro faz uso de arquivos de log para cada computador individualmente e o segundo captura pacotes que trafegam na rede para analisar seu conteúdo.

Online ou Offline, onde um é capaz de detectar e marcar um intruso enquanto a esta sendo realizada a intrusão, e o outro analisa registros após o evento ocorrer e indica que houve uma violação de segurança tinha ocorrido desde a última verificação, respectivamente.

Baseado em abuso ou anomalia, onde por anomalia o sistema identifica comportamento fora do padrão, e por abuso compara as atividades na rede com comportamentos de ataques já conhecidos.

A maioria dos métodos utilizados para detecção são baseados em inteligencia artificial (AI), entre as varias técnicas conhecidas de AI, a que tem tido melhores resultados e consequentemente mais usada é a de Redes Neurais Artificiais (RNA)(Jake Ryan, Meng-

Jang Lin, Risto Miikkulainen, 1998)(Miroslav Stampar, 2014).

As RNA são uma classe de algorítimos para aprendizado de maquina (AM), usada para realizar classificação de dados. A rede neural é treinada de forma a dar mais importância para as principais características de uma determinada instancia de um problema, para ajudar a classificar os dados que ainda estão por vir. RNAs tem sido utilizadas com sucesso na detecção de intrusão (C. Zhang, J. Jiang, M. Kamel, 2005) (X. Tong, Z. Wang, H. Yu, 2009) (S.-C. O. K. Y. Wonil Kims, 2004), porem elas necessitam de uma quantidade substancial de dados para realizar o treinamento, a partir desse treinamento que ela passara a ter capacidade de identificar os padrões para posteriormente receber os dados novos para classificação.

1.2 Motivação

Esforços realizados para proporcionar segurança em ambientes computacionais, tem como motivação o fato de existirem riscos que podem comprometer a integridade, confiabilidade e disponibilidade da informação. Esses riscos são avaliados de acordo com as chances do mesmo ocorrer e com os custos envolvidos para tratá-lo. Técnicas de defesa vêm sendo aprimoradas, porém ainda existem diversas limitações que as impedem de estarem efetivamente preparadas para o qualquer tipo de ataque (R. Beghdad, 2008), sendo assim necessário soluções inovadoras para tratar os níveis de ameaças atuais e futuras. Este cenário é a principal motivação deste trabalho que consiste em propor, implementar e mensurar resultados de uma solução para treinamento de RNA para detecção e prevenção de intrusão.

1.3 Justificativa

Muitas técnicas de AI tem sido utilizadas para IDS/IPS, a mais utilizada e a RNA(Miroslav Stampar, 2014), porem existem tipos de ataques que não são facilmente detectados, por ocorrerem com menor frequência, tendo poucas entradas para o treino da RNA(R. Beghdad, 2008), resultando em mais erros, por esse motivo escolhemos trabalhar de forma a aprimorar seu resultados.

1.4 Objetivos

O objetivo deste trabalho é propor uma forma de aprimorar o sistema de aprendizado de redes neurais artificiais para detecção e prevenção de intrusão. Para isso sera necessário, gerar uma base em um ambiente controlado para testes específicos, implementar uma solução de IPS/IDS que utilize RNA, implementar uma metodologia de treinamento,

realizar o treino da RNA e por fim comparar resultados com outras técnicas de treinamento para RNA.

1.5 Método de trabalho

Utilizaremos a base de dados de trafego em rede KDD Cup 99, por ser uma das bases mais completas e amplamente utilizada nos testes de IDS, sera essencial para se realizar uma comparação consistente de resultados.

Para gerar nossa base mais especifica iremos monitorar um ambiente de rede durante um período de tempo, no qual serão realizados alguns ataques controlados periodicamente, deste sera gerado um log que usaremos no nosso sistema.

Desenvolver uma solução para analise dos pacotes e eventos de uma rede, esta sera desenvolvida em Go, utilizara RNA para classificar as atividades na rede.

Realizar treinamento em ambas as bases de dados, serão formas diversificadas de treinamento, sera feito uma comparação de acertos/erros e tempo necessário para treino.

Faremos uma comparação de desempenho e efetividade de nossa solução e algumas que temos hoje.

1.6 Organização do trabalho

Este trabalho esta dividido em três partes. Na próxima seção, sera apresentado o estado da arte, onde sera revisado a literatura sobre detecção e prevenção de intrusão utilizando RNA.

Logo apos teremos a proposta de forma mais detalhada do que é pretendido realizar na próxima etapa do trabalho.

Por fim um cronograma de controle sobre como prosseguira a segunda etapa deste trabalho.

2 Revisão de Literatura

Um trabalho feito na Information Systems Security Bureau (Miroslav Stampar, 2014), fez um comparativo de técnicas publicas entre os anos de 2010 e 2014 para a detecção de intrusão.

A pesquisa indica que existe um pequeno crescimento das técnicas de aprendizado de maquina (ML) e inteligencia artificial (AI) comparadas com outras técnicas que não são informadas, mostrando as técnicas como uma sub-área importante e com uma forte tendencia, concluindo ML contribui como a principal área da AI utilizada para a detecção de intrusão.

Outro comparativo foi com o numero de publicação entre 2010 e 2014 das diferentes algoritmos de inteligencia artificial e aprendizado de maquina como Redes Neurais Artificiais, Logica Fuzzy, Algoritmo Genético, Arvore de Decisões, etc. A pesquisa mostrar que algorítimos com base em redes neurais artificiais são os algoritmos de inteligencia artificial mais popular entre as publicações, mesmo com uma pequena queda em seu uso no ultimo ano. Algoritmo Genético, K-vizinhos mais próximos e arvore de decisão ficam muito próximos, quase empatando no segundo lugar.

O artigo conclui que inteligencia artificial desempenha um papel substancial no estudo de detecção de intrusão, redes neurais artificiais são as mais populares, e que a pesquisa nessa área ainda é muito necessária, ha muitos resultados promissores nesses algorítimos, especialmente em abordagens hibridas, nos quais se utiliza a combinação entre técnicas diferentes.

No trabalho (Marley Maria Bernardes Rebuzzi Vellasco, 2007), Marley da uma visão geral sobre Redes Neurais Artificiais(RNA), explicando que a expressão "rede neural"é a tentativa de representar a capacidade que cérebro humano possui de reconhecer, associar e generalizar padrões, sendo as principais áreas de atuação para a classificação de padrões e previsão.

A modelagem de uma rede neural consiste em três etapas: (1) Treinamento e Aprendizado, obtido pelo ambiente gerador dos dados, (2) Associação, obtido pelo reconhecimento de padrões distintos e (3) Generalização, relacionado a capacidade da rede de reconhecer com sucesso o ambiente que origina os dados, não propriamente os dados utilizados no treinamento.

O conhecimento é passado por um algoritmo de treinamento e aprendizado, este é transformado e armazenado nas conexões. O aprendizado é resultado de muitas apresentações de conjuntos de exemplos de treinamento.

O treinamento pode ser dado de duas maneiras: (1)batelada ou ciclos, onde a atualização somente acontece depois da apresentação de todos os pesos e (2) padrão a padrão ou incremental,onde a atualização é feita após a apresentação de cada novo padrão. O procedimento de treinamento pode ser classificado em dois tipos: supervisionado e não supervisionado.

Existem basicamente dois tipos de arquiteturas ou topologia de redes neurais: redes não concorrentes e redes concorrentes. Redes não concorrentes são aqueles que não recebem atualização das suas saídas para suas entradas, organizando sua estrutura em uma ou múltiplas camadas. Feedfoward é um tipo de rede não concorrente, o sinal é sempre propagado para a frente, da entrada para a saída.

Redes concorrentes são redes com a atualização das saídas para as entradas, não tendo obrigatoriamente sua estrutura organizada por camadas e podendo ter interligações entre neurônios da mesma camada case esteja organizada em camadas.

No trabalho em conjunto do Departamento de Automação e Sistemas da Universidade Federal de Santa Catarina (Paulo M. Mafra,Joni da Silva Fraga, Vinícius Moll, 2008) e com a Pontifícia Universidade Católica do Paraná (Altair Olivo Santin, 2008), foi desenvolvido um sistema multi-camadas chamado de POLVO-IIDS, utilizando redes neurais de Kohonen, que classifica os dados de forma genérica, comportamentos considerados normais ou anômalos e para cada classe de ataque foi utilizada um algoritmo de inteligencia artificial do tipo Support Vector Machine(SVM) especializada na detecção da classe correspondente e tendo como saída a indicação de trafego normal ou atividade maliciosa. A ideia de utilizar outra rede neural é para minimizar o numero de falsos positivos, com apenas um tipo de ataque para classificar, aumenta a precisão para identificar apenas duas categorias, trafego normal ou anomalia.

O valor de cada neurônio pode variar de 0 a 1, normalmente em redes neurais de Kohonen algum neurônio deve estar de 0,1 a 0,9, mas no POLVO-IIDS, foi utilizado de 0.2 a 0.8 para evitar erros de margem. O trabalho indica de teve uma melhora nos resultados obtidos relacionados a outras literaturas, mostrando que o modelo POLVO-IIDS é um modelo eficiente.

Shraddha Surana propôs um modelo (Shraddha Surana, 2014) utilizando clusterização Fuzzy e Redes neurais artificiais, onde o dataset de treinamento é dividido em N subsets, que são distribuídos em N redes neurais intermediarias, com mais uma camada de RNA que agrega os resultados das redes intermediarias para a partir dai realizar a classificação de novos dados. Esta abordagem conseguiu uma taxa de detecção de 81,6%.

Um trabalho feito na Pontifícia Universidade Católica do Rio de Janeiro (Renato Maia Silva, Marco Antonio Grivet M. Maia, 2004), demonstra um teste de desempenho para algumas diferentes entradas para o treinamento de Redes neurais artificiais, utilizando

4 redes, uma utiliza as 41 categorias disponíveis, as outras 3 utilizam apenas 9 categorias básicas, usando um intervalo de -1 a 1. A primeira rede utilizando todas as 41 categorias, a segunda e terceira rede com apenas 9 categorias e uma saída, identificando como 1 sendo ataque ou -1 atividade normal, indicando como -1 normal, 0 neptune ou 1 smurf e a ultima com 4 saídas e 9 categorias, indicando (-1 1 1 1) normal,(1 -1 1 1) neptune, (1 1 -1 1) back e (1 1 1 -1) smurf.

As taxas de acertos foram acima de 90 por cento para todas as configurações testadas, a rede que teve um melhor resultado foi a terceira rede, tendo 97,5% na sua taxa de acertos.

Na Universidade Federal de Santa Maria (Bruno Lopes Dalmazo, Francisco Vogt, Tiago Perlin, Raul Ceretta Nunes, 2008), utilizando serie temporal, um modelo matemático para apresentar amostragens periódicas que apresentam dependência entre as amostras, classificando os sistemas de detecção de intrusão em três componentes fundamentais: fonte de informação, analise e resposta.

A fonte de informação representada por um coletor associado a um host, rede ou segmento de rede. Analise sendo parte do IDS que verifica eventos derivados da fonte de informação onde eh determinada a indicação que um evento eh uma intrusão que esta ocorrendo ou ja ocorreu.

O detector de intrusões baseado em series temporais teve como resultado preliminar uma demostração que a utilização de series temporais para a detecção de ataques, apresentam resultados satisfatórios para a detecção de um ataque e em pouco consumir de tempo de processamento.

No trabalho feito na universidade de Waterloo no canada(Chunlin Zhang, Ju Jiang, Mohamed Kamel, 2004),foi apresentado duas técnicas de redes neurais baseados em hierarquia para IDS, hierarquia em series e hierarquia paralela, onde o objeto para a detecção de ataques de abuso e anomalia em tempo real sem a interrupção humana.

Usando dois pré-requisitos para redes neurais hierarquias. O primeiro, cada classificação individual deve acertar seu desempenho, caso contrario,o erro é enviado para os níveis acima, acumulando e influenciados os níveis mais baixos. A taxa de detecção e de falso positivo são os principais indicadores de desempenho.

A segunda, as classificações,basicamente, podem ser dividas em grupos seguindo alguns critérios, cada grupo pode ser associado para seu própria classificação, então as classificações ou sua saída pode ser combinada em conjuntos.

O artigo conclui em seus resultados, uma ótima habilidade para detectar instruções conhecidas e desconhecidas com um curte período de tempo para o treinamento, uma alta taxa de detecção e um baixo índice de falsos positivos. O IDS de hierarquia de series conseguiu monitorar em tempo real o trafego na rede, treinando automaticamente novas

intrusos modificando sua estrutura para novas classificações. O IDS de hierarquia paralela, resolveu em partes, os problemas da hierarquia em series, sendo mais rápido para ser executado.

3 Proposta

Será implementada na próxima etapa desse trabalho um modelo para Detecção e prevenção de intrusão fazendo uso de técnicas de inteligência artificial, especificamente redes neurais artificiais, utilizar suas características de reconhecimento de padrões e generalização para realizar a classificação dos eventos ocorridos em redes de computadores, classificando-os em normais ou intrusivos.

Sera desenvolvida uma aplicação baseada no modelo, ela ira monitorar os pacotes em uma rede de computadores do tipo IPv4, este ira classificar e prevenir atividades maliciosas.

A principio sera utilizada uma RNA do tipo *feed-foward*, com três camadas apenas, uma de entrada, uma intermediaria e uma de saída. Os dados serão passados para a camada de entrada, processados pela rede e classificados como uma das cinco classes da camada de saída. Estas sendo Normal, DoS, U2R, R2L e Probe.

O modelo terá foco no treinamento da RNA, a principio para aprender os pesos da rede em multicamadas sera usado o algorítimo de *backpropagation* com a regra de atualização de peso do gradiente descendente.

Desse ponto serão mensurados alguns testes a partir da abordagem do POLVO-IIDS (Paulo M. Mafra,Joni da Silva Fraga, Vinícius Moll, 2008), no qual usaremos uma rede para classificação, e quatro SVM para detecção de anomalias.

Outra abordagem é a de clusterização (Shraddha Surana, 2014), separando o *dataset* de treino em N *subsets*, cada um sera utilizado para realizar o treinamento de uma rede neural, para no final outra rede realizar a agregação dos resultados destas N redes.

Pretende-se analisar a viabilidade de aplicação destas abordagems, bem como detalhar suas vantagens ou desvantagens em relação a métodos convencionais de detecção de intrusão.

4 Cronograma

Para modelar a proposta descrita, sera seguido um cronograma semanal, este ira descrever semana a semana as tarefas que devem ser realizadas de forma a se concluir o trabalho.

Semana	**Atividade**
1ª	Analise de trafego de pacotes em rede, criar aplicação em Go para ler e realizar log das operações na rede
2ª	Criar base de dados de trafego em rede de forma controlada. Verificar como implementações de IDS atuais reagem a esses dados. / trabalhar na monografia
3ª	Implementar na aplicação de log de trafego um sistema de redes neurais
4ª	Realizar treinamento da RNA da aplicação com os *datsets* KDD'99 e o modelado para o projeto / trabalhar na monografia
5ª	Mensurar resultados da aplicação do modelo / trabalhar na monografia
6ª	Comparar com os resultados obtidos por outros IDS, e com resultados publicados de outros modelos baseados em RNA
7ª	Analizar se for possivel como aprimorar os resultados do modelo
8ª	Implementar o modelo de POLVO-IIDS
9ª	Comparar com os resultados obtidos anteriormente / trabalhar na monografia
10ª	Implementar RNA clusterizado.
11ª	Comparar com os resultados obtidos anteriormente / trabalhar na monografia
13ª	Implementar modelo hibrido POLVO-IIDS Clusterizado
14ª	Comparar com os resultados obtidos anteriormente / trabalhar na monografia
15ª	Trabalhar na monografia - desenvolvimento
16ª	Trabalhar na monografia - resultados
17ª	Trabalhar na monografia - resultados
18ª	Apresentação dos resultados

Referências

Altair Olivo Santin. *POLVO-IIDS: Um Sistema de Detecção de Intrusão Inteligente Baseado em Anomalias*. [S.l.], 2008. Citado na página 9.

Bruno Lopes Dalmazo, Francisco Vogt, Tiago Perlin, Raul Ceretta Nunes. *Detecção de Intrusões baseada em Séries Temporais*. [S.l.], 2008. Citado na página 10.

C. Zhang, J. Jiang, M. Kamel. Intrusion detection using hierarchical neural networks. *Pattern Recognition Letters*, p. 779–791, 2005. Citado na página 6.

CENTRO DE ESTUDOS, RESPOSTA E TRATAMENTO DE INCIDENTES DE SEGURANçA. 2015. Disponível em: <http://www.cert.br/stats/incidentes/>. Acesso em: 25/05/2015. Citado na página 5.

Chunlin Zhang, Ju Jiang, Mohamed Kamel. *Intrusion detection using hierarchical neural networks*. [S.l.], 2004. Citado na página 10.

Jake Ryan, Meng-Jang Lin, Risto Miikkulainen. Intrusion detection with neural networks. *Advances in Neural Information Processing Systems 10*, 1998. Citado na página 6.

Marley Maria Bernardes Rebuzzi Vellasco. *REDES NEURAIS ARTIFICIAIS*. [S.l.], 2007. Disponível em: <http://www2.ica.ele.puc-rio.br/Downloads/33/ICA-introdu%C3%A7%C3%A3o%20RNs.pdf>. Citado na página 8.

Miroslav Stampar. *Artificial Inteligence in network intrusion detection*. [S.l.], 2014. Citado 2 vezes nas páginas 6 e 8.

Paulo M. Mafra,Joni da Silva Fraga, Vinícius Moll. *POLVO-IIDS: Um Sistema de Detecção de Intrusão Inteligente Baseado em Anomalias*. [S.l.], 2008. Citado 2 vezes nas páginas 9 e 12.

R. Beghdad. Critical study of neural networks in detecting intrusions. *Computers & Security*, p. 168–175, 2008. Citado na página 6.

Renato Maia Silva, Marco Antonio Grivet M. Maia. *REDES NEURAIS ARTIFICIAIS APLICADAS À DETECÇÃO DE INTRUSOS EM REDES TCP/IP*. [S.l.], 2004. Citado na página 9.

S.-C. O. K. Y. Wonil Kims. Intrusion detection based on feature transform using neural network. *Computational Science - ICCS 2004*, p. 212–219, 2004. Citado na página 6.

Shraddha Surana. *Intrusion Detection using Fuzzy Clustering and Artificial Neural Network*. [S.l.], 2014. Disponível em: <http://www.wseas.us/e-library/conferences/2014/Gdansk/FUNAI/FUNAI-32.pdf>. Citado 2 vezes nas páginas 9 e 12.

X. Tong, Z. Wang, H. Yu. A research using hybrid rbf / elman neural networks for intrusion detection system secure model. *Computer Physics Communication*, p. 1795–1801, 2009. Citado na página 6.