

10^o

FEPEG FÓRUM

ENSINO • PESQUISA
EXTENSÃO • GESTÃO
RESPONSABILIDADE SOCIAL: INDISSOCIABILIDADE
ENSINO, PESQUISA E EXTENSÃO UNIVERSITÁRIA



ISSN 1806-549 X

Autor(es): GUSTAVO HENRIQUE REINALDO DE OLIVEIRA LIMA, JOHN DENNER DIAS DO ROSÁRIO, VITOR GIOVANE PEREIRA ALVES, NILTON ALVES MAIA

Classificação de protocolos de aplicação utilizando Redes Neurais Artificiais do tipo MLP através de interceptação simulada de características obtidas sobre parâmetros de tráfego de rede

Introdução

As Redes Neurais Artificiais (RNAs) são um campo de estudo da Inteligência artificial que frequentemente são utilizadas para diversos propósitos pois conseguem perceber padrões que são complexos demais para os seres humanos [1].

O Multilayer Perceptron (MLP) é uma RNA que permite que sejam solucionados problemas não-linearmente separáveis justamente por permitir a utilização de camadas escondidas e utilizar uma função de ativação não-linear [2]. No MLP há três grandes divisões: a camada de entrada, as camadas escondidas ou ocultas e as camadas de saída. A camada de entrada é a porta de acesso dos padrões de entrada à RNA. As camadas escondidas ou ocultas, são as camadas localizadas entre a camada de entrada e a camada de saída. Não há limite para a quantidade de camadas escondidas, e também não é necessário que uma RNA tenha esse tipo de camada. Por fim, a camada de saída gera a saída da rede, ou seja, a resposta da rede sobre um determinado conjunto de padrões apresentado à camada de entrada [2].

Com a crescente utilização da quantidade de usuários de redes de computadores também cresce a preocupação com a qualidade de gerenciamento e com a segurança dessas redes. A Inteligência Artificial e mais especificamente as RNAs podem ser grandes aliadas para garantir um bom funcionamento dessa volumosa troca de dados através das redes de computadores. Dentre os inúmeros benefícios pode-se citar o reconhecimento de padrões de utilização de protocolos de aplicação, nesse sentido o administrador de redes pode buscar soluções para problemas específicos, reduzindo o tempo de procura do problema.

Diante da relevância da utilização de Inteligência Artificial em redes de computadores, o presente estudo visa fornecer uma abordagem simples dessa tecnologia, realizando uma classificação de protocolos da camada de aplicação a partir de características simuladas de dados de tráfegos de rede. Para realizar essa classificação utilizou-se uma RNA do tipo MLP.

Material e métodos

A. Criação do cenário

Foi criado um cenário para propiciar a utilização da RNA do tipo MLP. O cenário imaginado é o de um setor de uma empresa que possua apenas um computador pessoal e uma servidor *web*. Esses dois dispositivos estão conectados a um roteador que propicia o acesso direto à *internet* caso o protocolo de aplicação trafegado seja HTTP, SSH e FTP. Caso o protocolo seja SMTP ele passará por um módulo *antispam* que apenas permitirá a transmissão do fluxo à *internet* caso não detecte *spam*. Esse cenário é ilustrador na Figura 1.

Nesse cenário, o computador pessoal apenas utiliza os protocolo de aplicação HTTP e SSH. Já o servidor *web* utiliza os protocolos FTP, HTTP e SMTP.

B. Aquisição do conjuntos de padrões

Foram gerados, através do desenvolvimento de um algoritmo na linguagem de programação PHP, 2000 (dois mil) padrões que refletem uma interpretação das características binárias resultantes de uma interceptação simulada de pacotes de rede. As características binárias estão organizadas: quanto a origem, PC ou servidor; quanto ao tamanho do *payload* do pacote, pequeno ou grande; quanto ao protocolo de transporte, TCP ou UDP; quanto à necessidade do pacote passar pelo módulo de *antispam*, sim ou não.

C. Rede MLP utilizada na classificação dos dados

O objetivo do presente trabalho é que a RNA do tipo MLP utilizada, a partir de determinadas características de tráfego de uma rede de computadores, classifique cada pacote de rede em FTP, HTTP, SSH ou SMTP. Para que essa

10^o

FEPEG FÓRUM

ENSINO • PESQUISA
EXTENSÃO • GESTÃO

RESPONSABILIDADE SOCIAL: INDISSOCIABILIDADE
ENSINO, PESQUISA E EXTENSÃO UNIVERSITÁRIA



ISSN 1806-549 X

RNA consiga determinar qual desses protocolos de aplicação foi utilizado nos 2000 padrões de entrada, foi desenvolvido um algoritmo no *software* Matlab R2015a. Como o MLP possui aprendizado supervisionado, onde as entradas e saídas desejadas são fornecidas por um módulo supervisor [3], nesse algoritmo foram apresentados os padrões possíveis e suas saídas desejadas. Em seguida foi definida a estrutura da rede MLP com 8 camadas escondidas e duas saídas possíveis. O número máximo de épocas definido é 7000 à taxa de aprendizado 0,01 como pode ser observado na Tabela 1.

Resultados e discussão

Após a realização dos treinamentos, observou-se que todos algoritmos alcançaram o resultado esperado, ou seja, classificaram corretamente todos os padrões. Os algoritmos de treinamento *Resilient backpropagation* (*trainrp*) e Gradiente descendente *backpropagation* com *momentum* e taxa adaptativa (*traingdx*) alcançaram erro zero, já o Gradiente descendente *backpropagation* com taxa adaptativa (*traingda*) alcançou como taxa de erro, em sua melhor performance, o valor 7.0559e-11.

Analisou-se também a quantidade de épocas utilizada por cada algoritmo de treinamento. Nesse quesito o algoritmo *Resilient backpropagation* saiu-se melhor que os demais, já que conseguiu realizar o treinamento da RNA em 18 épocas, a menor quantidade de épocas dentre os algoritmos de treinamento utilizados no presente trabalho. O algoritmo de treinamento Gradiente descendente *backpropagation* com *momentum* e taxa adaptativa treinou a RNA em 109 épocas. Por último, o algoritmo de treinamento Gradiente descendente *backpropagation* com taxa adaptativa foi o que utilizou a maior quantidade de épocas de treinamento da RNA, 281, um valor maior que o dobro da soma da quantidade de épocas dos dois outros algoritmos de treinamento. Esses resultados podem ser observados na Figura 2.

Considerações finais

O presente trabalho apresentou uma arquitetura a partir de um cenário fictício, um comparativo gráfico da performance alcançada pelos algoritmos de treinamento escolhidos, os valores resultantes de uma classificação de protocolos de aplicação sobre um conjunto de padrões com diferentes algoritmos de treinamento.

Como trabalho futuro pretende-se desenvolver com a linguagem de programação Java um sistema que seja capaz de interceptar pacotes de rede e armazenar em um banco de dados informações resultantes de processamento sobre os mesmos, como *jitter*, atraso e perda de pacotes. Sobre esse banco de dados pretende-se aplicar uma RNA do tipo não-supervisionado, ou seja, quando não se conhece os valores de saída para um conjunto de padrões de entrada.

Agradecimentos

À FAPEMIG pelo financiamento da pesquisa.

Referências bibliográficas

- [1] TISSOT, Hegler C.; CAMARGO, Luiz C.; POZO, Aurora T. R.. Treinamento de Redes Neurais Feedforward: comparativo dos algoritmos Backpropagation e Differential Evolution. In: ENCONTRO NACIONAL DE INTELIGÊNCIA ARTIFICIAL, 9., 2012, Curitiba. **BRACIS**. Curitiba: Bracis, 2012. Disponível em: <http://www.ppgia.pucpr.br/~enia/anais/enia/artigos/105243_2.pdf>. Acesso em: 6 nov. 2016.
- [2] FERRARI, Hevelin Viviane et al. Uma comparação entre Redes Neurais Wavelet, LMS, MLP e RBF para classificação de DPOC. In: CONGRESSO DE MATEMÁTICA E SUAS APLICAÇÕES, 1., 2006, Foz do Iguaçu. **Uma comparação entre Redes Neurais Wavelet, LMS, MLP e RBF para classificação de DPOC**. Foz do Iguaçu: Congresso Internacional de Matemática e Suas Aplicações, 2006.
- [3] NARDO JUNIOR, Aparecido. **APLICAÇÃO DE REDES NEURAIS UTILIZANDO O SOFTWARE MATLAB**. 2005. 75 f. TCC (Graduação) - Curso de Ciência da Computação, Centro Universitário Eurípides de Marília – Univem, Fundação de Ensino Eurípides Soares da Rocha, Marília, 2005. Disponível em: <[http://aberto.univem.edu.br/bitstream/handle/11077/412/Aplicação de Redes Neurais utilizando o software MATLAB.pdf?sequence=1](http://aberto.univem.edu.br/bitstream/handle/11077/412/Aplicação%20de%20Redes%20Neurais%20utilizando%20o%20software%20MATLAB.pdf?sequence=1)>. Acesso em: 7 nov. 2016.

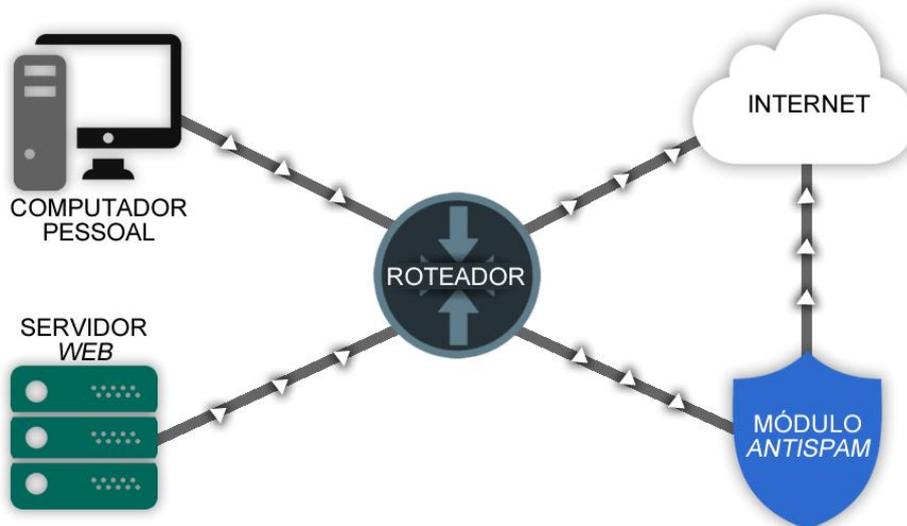


Figura 1 – Arquitetura do cenário criado

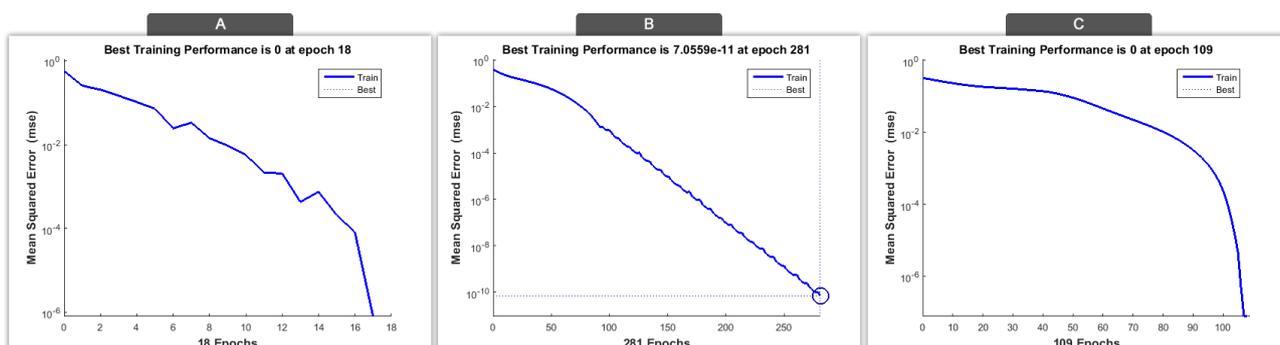


Figura 2 – Comparativo da performance apresentada pelos algoritmos de treinamento escolhidos durante a fase de treinamento da Rede Neural MLP. Figura 2A, *Resilient backpropagation* (trainrp); Figura 2B, Gradiente descendente *backpropagation* com taxa adaptativa (trainгда); Figura 2C, Gradiente descendente *backpropagation* com *momentum* e taxa adaptativa (trainгдаx).

Tabela 1 – Configurações de treinamento utilizadas

Configuração	Nº de camadas escondidas	Taxa de aprendizado	Épocas	Função de treinamento
1	8	0,01	18	trainrp
2	8	0,01	281	trainгда
3	8	0,01	109	trainгдаx

10^o

FEPEG FÓRUM

ENSINO • PESQUISA
EXTENSÃO • GESTÃO

RESPONSABILIDADE SOCIAL: INDISSOCIABILIDADE
ENSINO, PESQUISA E EXTENSÃO UNIVERSITÁRIA



ISSN 1806-549 X

Tabela 2 – Quantidade de cada protocolo de aplicação detectado

Protocolos	FTP	HTTP	SSH	SMTP
Quantidade absoluta	234	432	439	895
Quantidade percentual	11.7%	21.6%	21.95%	44.75%

S ou PC	Tam Payload	Protocolo Transp	Módulo anti-spam	Saída
S	P	TCP	0	FTP
P	G	TCP	0	HTTP
S	G	TCP	0	HTTP
P	P	TCP	0	SSH
P	P	UDP	0	SSH
S	P	TCP	1	SMTP
S	G	UDP	1	SMTP
S	G	TCP	1	SMTP
S	P	UDP	1	SMTP