ICANN19
28th International Conference on Artificial Neural Networks

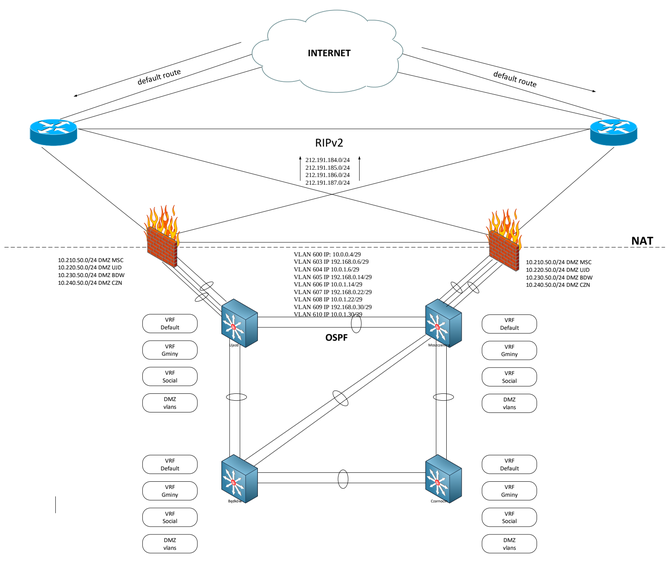
Convolutional Recurrent Neural Networks For Computer Network Analysis

Jakub Nowak [0000−0002−1572−3426], Marcin Korytkowski [0000−0002−6002−2733], Rafal Scherer [0000−0001−9592−262X]
Computer Vision and Data Mining Lab, Institute of Computational Intelligence,
Czestochowa University of Technology
Al. Armii Krajowej 36, 42-200 Czestochowa, Poland
{jakub.nowak, marcin.korytkowski, rafal.scherer}@iisi.pcz.pl http://iisi.pcz.pl

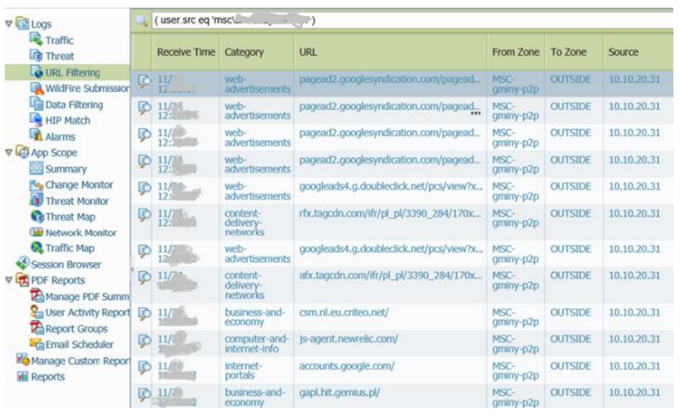
Introduction

A method of computer network user detection with recurrent neural networks. We use LSTM and gated recurrent unit neural networks. We added convolutional input layers. We transform requested URLs by one-hot character-level encoding. The system was checked on real network data collected in a local municipal network. It can classify network users; hence, it can also detect anomalies and security compromises

Local LAN to collect data



Logs from Paloalto firewall:



Neural Networks and Encoding

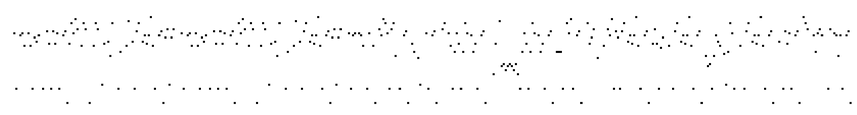
Four types of recurrent neural networks: LSTM, BiLSTM, GRU and BiGRU with convolutional input layers. The input data to the networks was a 70x45-pixel image with character-level one-hot encoding, as we have 70 possible characters and up to 45 characters for one URL. The convolutional layer in all RNNs had two versions:

Version 1:

Input 45x70x1
Embedding 32 (45x32x1)
Convolution 128 feature maps, filter 1D size 5 stride 1
Convolution 256 feature maps, filter 1D size 3 stride 1
Convolution 512 feature maps, filter 1D size 2 stride 1
MaxPooling 6
512 x 7 = 3584 (input to LSTM)

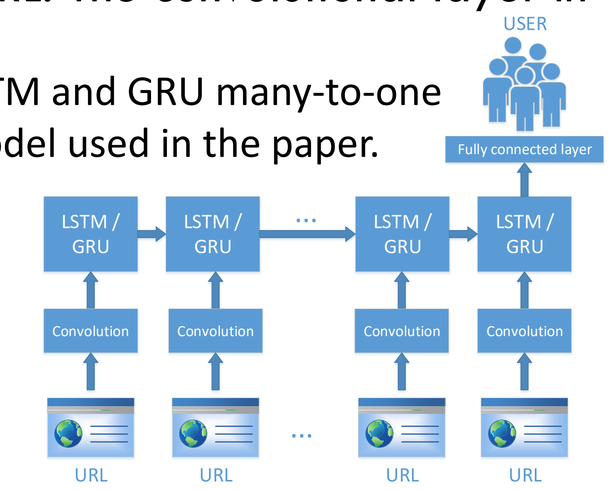
Version 2:

Input 45x70x1
Embedding 32 (45x32x1)
Convolution 64 feature maps, filter 1D size 5 stride 1
Convolution 128 feature maps, filter 1D size 3 stride 1
Convolution 256 feature maps, filter 1D size 2 stride 1
MaxPooling 6 2
56 x 7 = 1792 (input to LSTM)

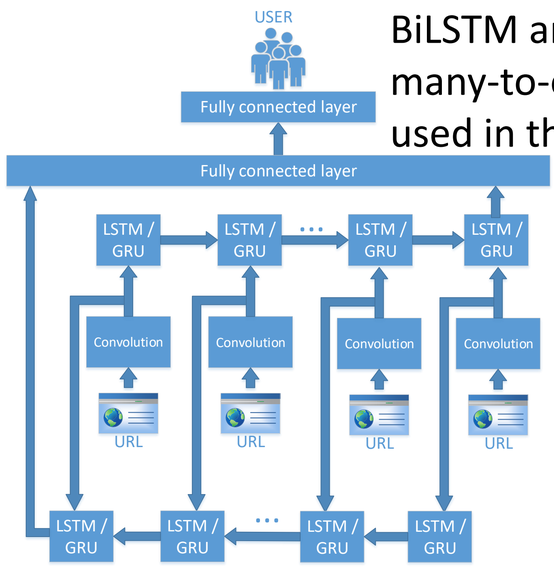


Example of encoded URL

LSTM and GRU many-to-one model used in the paper.



BiLSTM and BiGRU many-to-one model used in the paper.

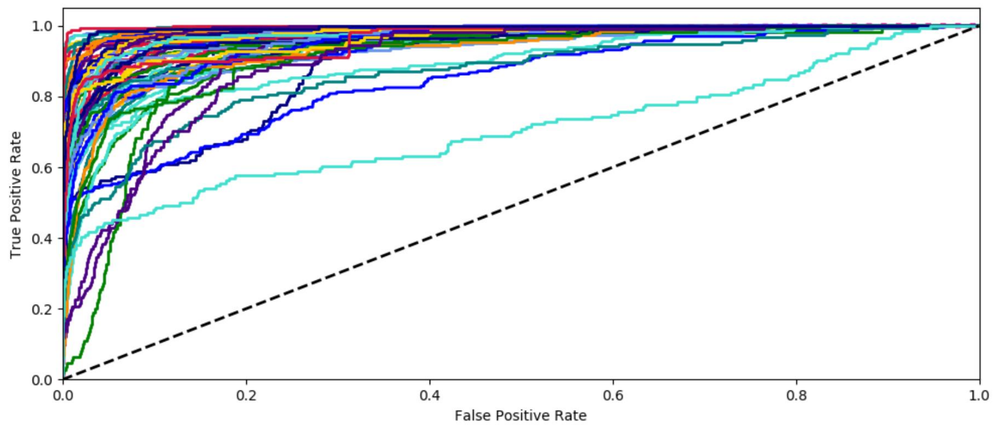


Results

Testing error for all the neural networks with two variants of input convolutional layers. The networks were tested on future data, not used during training.

Network	Error %	
	Version 1	Version 2
LSTM	28.90%	27.10%
BiLSTM	28.00%	28.20%
GRU	26.58%	26.60%
BiGRU	25.40%	26.20%

ROC curves for all the users for the BiGRU network (the best one)



Conclusions

We showed that LSTM and GRU networks with input convolutional layers are suitable for identifying network users based on URLs they requesting.

The convolutional layers and one-hot encoding on the character level we applied, entirely replace the use of a dictionary or other ways of feeding text to recurrent networks.

Such an approach is especially useful in the case of URLs which often do not use regular English (or any other language) words.

ManuscriptID: 534