Reflection Paper for "Tiresias: Predicting Security Events Through Deep Learning"

Yash Arun Patade

Department of Computer Science, New York Institute of Technology

DTSC 615: Optimization Methods for Data Science

Dr. Taoufik Ennoure

May 07, 2021

Reflection Paper for "Tiresias: Predicting Security Events Through Deep Learning"

This paper will consist summary and personal thoughts of the paper in consideration. This paper will also contain details about the testing of the Tiresias prediction model and other related prediction models which use baseline models of lower complexity.

History of the name 'Tiresias' in Odyssey and Greek Mythology.

The Odyssey or the Odyssey of Homer is a Greek poem which describes the quest of return of Odysseus. The journey was the aftermath to the war of Troy, which lasted for 10 years. Odysseus's journey was to the island of Ithaca.

In Greek Mythology, Tiresias was a blind prophet. He was the son of the Shepherd Everes and the nymph Chariclo. When Tiresias was blinded by Athena, he was also granted the power of foresight. He gained the ability of 'seeing the future' and prediction of events.

Tiresias was an important character in many Greek Mythologies. He foretold a dark future for Narcissus, the infamous character who was obsessed with his own beauty. And where the term 'narcissist' comes from. Tiresias predicted that Narcissus would live a long life, as long as he does not see his own reflection. Ironically, Narcissus drowned in the pool of water while gazing at his own reflection.

Recurrent Neural Networks.

Tiresias uses the RNN for predicting events of the future. The past data is to be taken into consideration for any prediction to be done. The RNNs are powerful neural networks which have an internal memory. The basic neural network lacks an internal memory. The basic neural network takes in a vector of fixed size as its input. Whereas, the an RNN is designed to take in a series of inputs with no predefined ranges.

Data Collection for Testing Tiresias.

The authors of the said paper collected the data from 740,000 nodes over a period of 27 days. The total collected data, which was 3.4 billion security events, was then used for testing the Tiresias. The total data was divided into two separate datasets namely; D1 and D2. D1 was the data collected from 17 days (Nov 1 – Nov 17). D2 dataset has 1.2 billion security events of every 8th and 23rd day of all the months between Nov 2017 and Feb 2018, and the first three days of Jan 2018.

Comparison of baseline models and Tiresias.

The authors of the paper perform a comparative analysis of whether a high complexity RNN like Tiresias is required for the prediction of the security events, or just simple, baseline models with lower complexity would suffice. The baseline models used here were the Spectral, Markov Chain and 3-gram. In the study, the Tiresias performs better than all the baseline models. Further into the comparative study, it is found out that the 3-gram performs better than the Markov Chain. And Markov Chain performs better than Spectral. The authors then conclude that, this order of performance illustrates the significance of

sequence memory. The computational time for training and running the models could not be compared fairly. The reason was that the Tiresias uses GPU to train the RNN models, while the other baseline systems rely on traditional CPUs. The 3-gram training and running required then days whereas the Tiresias (using GPU) required 10 minutes per epoch.

Testing Tiresias in real-world scenarios.

A multi-step, co-ordinated network attack is an attack where network exploits are carried out concurrently and very closely to one another. This increases the degree of harm the attacker can inflict on the network. The difficulty of identifying such attacks is that it may pass as non-harmful or non-threatful when the IPS engine views them one-by-one.

The multi-step attack was successfully memorized by the prediction tool. The Tiresias did not always perform correct predictions. In the test conducted by the authors, Tiresias failed to perform correct event prediction twice. But, it could adapt itself by using the actual observed events to correct its own errors.

Conclusion.

The Tiresias is a robust system which implements the Recurrent Neural Network. Security attacks were predicted using the same system. Other baseline models with lower complexity like 3–gram, Spectral and Markov Chain were compared with Tiresias, with Tiresias outperforming all the three. The results from the test concluded how important it is to implement the use of sequential memory.

References

[1] Shen, Y., Mariconti, E., Vervier, P. A., & Stringhini, G. (2018, October). *Tiresias: Predicting Security Events Through Deep Learning*. CCS '18. https://doi.org/10.1145/3243734.3243811

[2] Donges, N. (2020, September 3). *A Guide to RNN: Understanding Recurrent Neural Networks and LSTM*. Built In. https://builtin.com/data-science/recurrent-neural-networks-and-lstm

[3] Venkatachalam, M. (2019, August 20). *Recurrent Neural Networks - Towards Data Science*. Medium. https://towardsdatascience.com/recurrent-neural-networks-d4642c9bc7ce