Improving machine learning based side-channel analysis

Jim Kok

Delft University of Technology

SIDE-CHANNEL ANALYSIS

What is Side-Channel Analysis (SCA)? And how can machine learning be applied to SCA?

- Cryptographic devices unintentionally leak physical information (e.g. power consumption).
- These leakages expose information about the secret key.
- Machine learning algorithms are used to classify the leakages (i.e., the traces) with their corresponding key (i.e., the label).

RESEARCH QUESTION

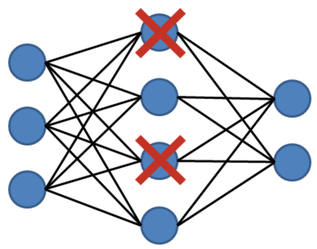
Can machine learning based SCA be improved by applying dropout?

- Dropout is a known regularization technique.
- Machine learning has been improved in various areas (e.g. speech recognition) by dropout.
- To the best of our knowledge no extensive research has been done in a SCA based context.

HYPOTHESIS

Dropout improves machine learning based SCA.

- Overfitting is one major issue regarding machine learning.
- Overfitting occurs when a model starts to memorize the inputs, instead of learning from them.
- Dropout drops out nodes, at random, from a neural network in a certain layer to prevent overfitting.



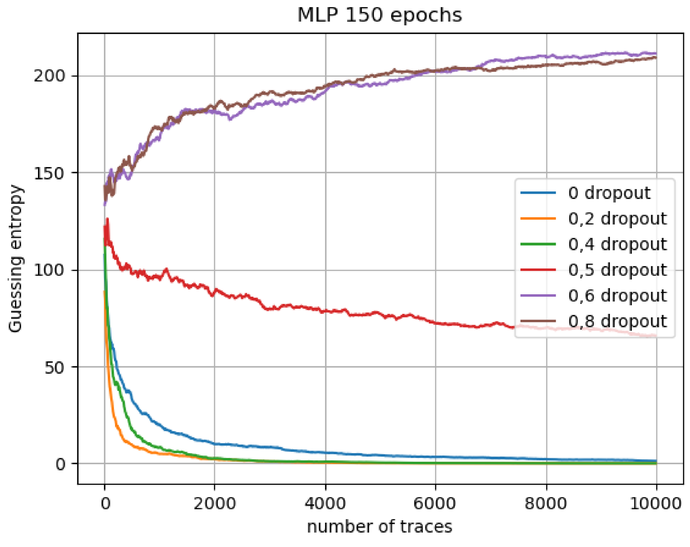
Example where the dropout rate is set to 0,5.

EXPERIMENT

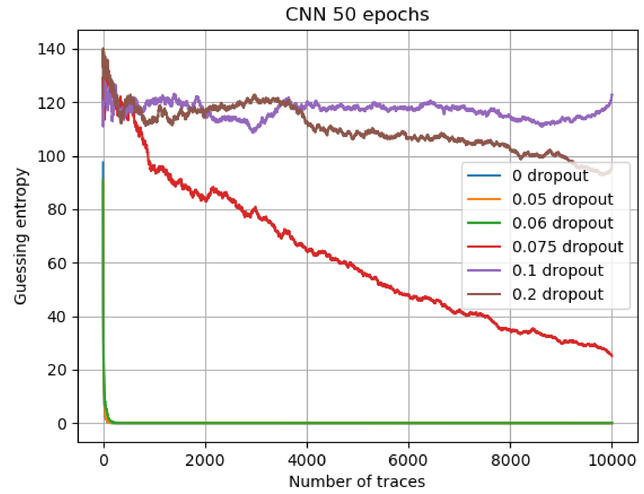
Investigation of dropout for three publicly available machine learning architectures.

- 2 Convolutional Neural Networks (CNN) architectures tested.
- CNN architectures vary in complexity and efficiency.
- 1 Multilayer Perceptron (MLP) architecture tested.
- ASCAD database (benchmark for SCA community) used for experiments.

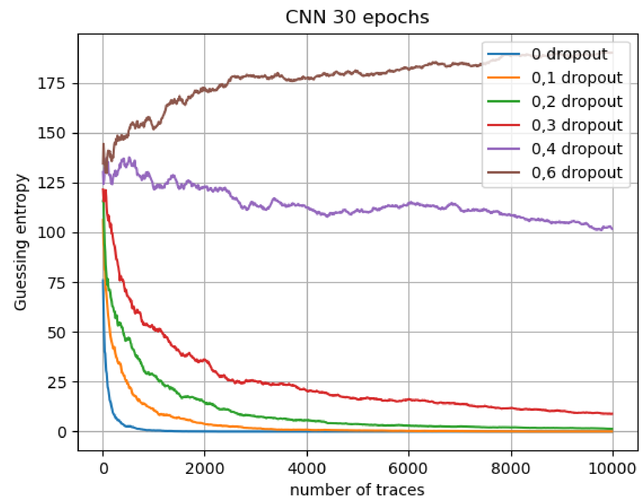
RESULTS



MLP architecture showing enhanced performance.



Efficient and uncomplicated CNN architecture showing negligible improvements.



Second CNN architecture, which is more complicated. Likewise, showing no improvements.

CONCLUSION

- Dropout showed promising results for the MLP architecture.
- Dropout showed minor to no improvements for the 2 CNN architectures.
- 2 certain consecutive dropout values drastically decrease performance.
- This threshold seems inherent to the architecture.
- The more complex CNN and MLP architecture showed a higher threshold.