

Peer Review: Traffic Classification by Hardware-accelerated Machine Learning

by Yvonne-Nadine Preda, 01404589

reviewed by Michael Mente 01634435, a01634435@unet.univie.ac.at

University Vienna, Faculty of Computer Science

This bachelor thesis aims to achieve performance improvements in the context of network traffic classification by combining neural networks and low-level programming. On the one hand, the focus is on possible optimizations by using SIMD instructions (within a neural network), on the other hand on supporting preprocessing on the data plane.

The expose clearly presents the current problem and also shows why a solution is needed. Furthermore, a brief overview of existing implementations in similar fields of application that are currently in use is given, thus ensuring a good contextual embedding of the topic. By clearly articulating the status quo, the research question of whether network traffic can be classified using a neural network is almost obvious. However, it is novel to include the benefits of hardware acceleration of SIMD instructions on the classification process.

The sub-topics on which the focus is particularly placed are clearly specified, using areas from various sub-disciplines of computer science. The combination of neural networks and hardware-oriented programming using P4 presumably holds great potential for performance improvements.

The methods, i.e., the development of prototypes with subsequent performance measurements, are well suited to achieve the stated goals. However, it is not entirely clear whether a second implementation that does not support hardware acceleration is necessary to measure the benefit of hardware accelerated SIMD instructions. The same is true for the measurements of preprocessing efficiency using P4 in the data plane. Regardless, the measurement results can provide new insights into load balancing with respect to traffic preprocessing and hardware acceleration.

In my opinion, the scope of the research questions is relatively broad. In the course of the research work, this could entail the risk that problems arise that take an unexpectedly long time to solve. I am referring to the use of a neural network for traffic analysis, which alone probably contains enough topics for a

bachelor's thesis. If everything goes as planned, I do not see any problems staying on schedule. However, to be on the safe side, I would budget a little more time for the implementation phase. I am also surprised that more time is allocated for the preparation of the interim presentation than for the final presentation.

In general, the exposé already gives a good outlook on the bachelor thesis. The topic, the goals and the methods are clearly defined and already offer a preview of the upcoming work, and there is nothing to criticize in terms of content. However, I would find it helpful to divide the expose into subsections for easier navigation.