ETH
Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

Prof. Laurent Vanbever
Networked Systems Group

Evaluating and Defeating Network Flow Classifiers Through Adversarial Machine Learning

Semester thesis proposal

Background Analyzing and understanding network activity can be extremely challenging in large and heterogeneous networks. The use of machine learning classifiers, complementary to classical intrusion detection systems to identify malicious activity is gaining importance. In order to allow live detection, classifiers are trained on small and efficiently computable feature sets, i.e., based on the flow's metatata, timing and volume. While such systems typically work well in settings where the adversary does not know of their existence, adversarial inputs can easily fool them by altering packets aspect and number without changing the effective payload of the flow.

Goals The high-level goal of this thesis is to analyze existing classifiers and to develop a system to circumvent them. We will start with a classifier that we developed in our group [3] and which was tested on data from the worlds largest cyber defense exercise [1]. As a final result, we envision a system that can modify network traffic in real time such that classifiers fail. For example, this could be achieved by *(i)* implementing a proxy; *(ii)* using programmable data planes [2]; or modifying the operating system's networking stack.

More precisely, the tasks of this thesis are the following:

- Analyze existing network traffic classifiers and measure their robustness against malicious inputs
- Implement a system that transparently modifies network traffic such that the classifier's outcome is wrong
- Evaluate the effectiveness of the resulting system as well as its impact on the network performance
- Propose measures to improve the classifiers robustness

If promising, the results of this work are expected to be used in a follow-up international cyber defense exercise.

Requirements

- Knowledge in communication networks
- Knowledge in machine learning
- Programming skills (e.g., Python)

Contacts This thesis is offered in collaboration with armasuisse Science and Technology within the Cyber Defence Campus initiative and can be performed either at ETH or at the research labs of armasuisse in Thun. If you are interested in the topic, please contact us:

- Roland Meier, meierrol@ethz.ch
- Dr. Luca Gambazzi, luca.gambazzi@ar.admin.ch
- Prof. Dr. Laurent Vanbever, lvanbever@ethz.ch

References

[1] Locked shields. https://ccdcoe.org/exercises/locked-shields/.

[2] P. Bosshart, D. Daly, G. Gibb, M. Izzard, N. McKeown, J. Rexford, C. Schlesinger, D. Talayco, A. Vahdat, G. Varghese, et al. P4: Programming protocol-independent packet processors. *ACM SIGCOMM Computer Communication Review*, 44(3):87–95, 2014.

[3] N. Känzig, R. Meier, L. Gambazzi, V. Lenders, and L. Vanbever. Machine learning-based detection of c&c channels with a focus on the locked shields cyber defense exercise. In *2019 11th International Conference on Cyber Conflict (CyCon)*. IEEE, 2019.