REVIEWER'S OPINION OF FINAL THESIS

I. IDENTIFICATION DATA

Thesis name:	**Detection of HTTPs Malware Traffic**
Author's name:	**František Střasák**
Type of thesis :	bachelor
Faculty/Institute:	Faculty of Electrical Engineering (FEE)
Department:	Department of Cybernetics
Thesis reviewer:	Sebastian Garcia
Reviewer's department:	Department of Computer Science

II. EVALUATION OF INDIVIDUAL CRITERIA

Assignment **challenging**

Evaluation of thesis difficulty of assignment.

The difficulty assignment of this thesis was challenging. The student worked on a problem that is even challenging for most security companies: the analysis of malware using encrypted connections. The topic requires several different skills to be analyzed properly: understanding of encrypted connections, understanding of network traffic and understanding of basic machine learning algorithms.

Satisfaction of assignment **fulfilled**

Assess that handed thesis meets assignment. Present points of assignment that fell short or were extended. Try to assess importance, impact or cause of each shortcoming.

This thesis completely fulfills the assignment. The student correctly worked with the data, understood the problem, generated new solutions and performed the required experiments.

Method of conception **correct**

Assess that student has chosen correct approach or solution methods.

The student has chosen the correct path for solving the thesis. He spend time analyzing the data, extracting features, evaluating the results and writing the process.

Technical level **A - excellent.**

Assess level of thesis specialty, use of knowledge gained by study and by expert literature, use of sources and data gained by experience.

The technical level of the thesis is excellent because the student applied the basic machine learning concepts correctly, never overestimated the results and understood the problem of obtaining the correct dataset. This thesis is also an excellent technical level because it can be applied directly to real network traffic.

Formal and language level, scope of thesis **C - good.**

Assess correctness of usage of formal notation. Assess typographical and language arrangement of thesis.

This thesis doesn't have the need for a very formal notation, and therefore the formal level is good enough. Where it was needed, the notation is correct. Despite the complexity of the topic the English is good but not outstanding. Therefore, there is still room for improvement.

Selection of sources, citation correctness **D - satisfactory.**

Present your opinion to student's activity when obtaining and using study materials for thesis creation.Characterize selection of sources. Assess that student used all relevant sources. Verify that all used elements are correctly distinguished from own results and thoughts. Assess that citation ethics has not been breached and that all bibliographic citations are complete and in accordance with citation convention and standards.

This is the weakest area of the Thesis. The analysis of the previous work is barely enough and there are works that were left out. However, it is true that the topic is quite unsolved yet and therefore the thesis is novel. I would have liked to see more analysis of previous works in some of the topics used in the thesis.

Additional commentary and evaluation

Present your opinion to achieved primary goals of thesis, e.g. level of theoretical results, level and functionality of technical or software conception, publication performance, experimental dexterity etc.

I'm pleased to see a bachelor student analyzing this topic in such way. The primary goals of the thesis were fulfilled, including the goal for the student for learning about research. The results obtained are good and promising, specially considering that the student knows how to continue.

III. OVERALL EVALUATION, QUESTIONS FOR DEFENSE, CLASSIFICATION SUGGESTION

Summarize thesis aspects that swayed your final evaluation. Please present apt questions which student should answer during defense.

In this thesis František Střasák tackled a very important problem of computer security: To analyze the behavior of malware using HTTPs in the network in order to improve its detection using machine learning algorithms. This is a challenging problem with important consequences for the community.

REVIEWER'S OPINION OF FINAL THESIS

The thesis starts with a first analysis of the network traffic and how to better obtain it for the goal of detecting HTTPs malware traffic. This ends up with the choose of Bro as the source of information and an analysis of the structure of HTTPs.

The most difficult and challenging part of the thesis was finding the correct features for the analysis. The student did a remarkable work in identifying meaningful and strong features. This was a large task that required smart solutions.

With the features separated, the student analyzed and compared several state-of-the-art machine learning algorithms, such as XGBoost, and obtained very good results. It should be noted that the student felt the necessity of a better dataset and he created it himself.

Given how good the thesis was solved and the importance of its results, I evaluate handed thesis with classification grade **A - excellent.**

Date: June 11th **2017, Prague.** Signature: Sebastian Garcia