REVIEWER'S OPINION OF FINAL THESIS

I. IDENTIFICATION DATA

Thesis name:	**Identifying Malicious Hosts by Aggregation of Partial Detections**
Author's name:	**Ondrej Lukás**
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.

I consider this thesis to be challenging for the following reasons. First, it is based on a software that was not made by the student. Second, it was focused on a quite unusual field of research: the decrease of false positive alarms in a *working* algorithm and *independently* of this algorithm. Third, because the data available is scarce, making the analysis more difficult.

Satisfaction of assignment **fulfilled with minor objections**

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.

The thesis was fulfilled with satisfaction in most of its parts. The analysis of the problem was correct, the development of the software solution was satisfactory, and the work in the algorithms was good. However, there were some parts that can be improved. In particular the dataset can be improved to include better normal data and a better assortment of malware captures. Regarding the description of the technique in the thesis, it could be improved to better reflect the work of the student.

Method of conception **correct**

Assess that student has chosen correct approach or solution methods.

The method of conception of the thesis is correct. The path was difficult but the student managed to understand and improve a previous program, to find new ways of dealing with errors and to explore the different algorithms that can be used to solve the problem.

Technical level **B - Very Good.**

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 very good. The student make good use of the knowledge gained and by the sources and data. In particular he did very good use of the datasets and the experiments. The area of false positives is very hard because it highly depends on the dataset used.

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

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

The thesis is highly technical and has two main parts. The first part was the development of an extended version of the original Slips tools, improving its use in real networks. The second part was the improvement of the tool in regard of its false positive detections using machine learning methods. The first part was correctly developed, while the second part could use a more polished description. Finally, the major aspect of the thesis that should be improved is the use of the English language.

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.

The analysis of previous work was one of the major weaknesses of the thesis. The student had a basic idea of the previous work in the area, but a more deep study could have improved the analysis. However, the student did a good work by finding a novel technique for dealing with errors in the detection.

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.

The student did a good bachelor work for fulfilling the goals of the thesis. Including the goal of learning how to do a basic research and how to write a thesis. The level of results are good, and they show just how difficult was to solve the problem completely. Moreover, the student did a good work of adapting to a previous software.

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.

REVIEWER'S OPINION OF FINAL THESIS

In this thesis Ondrej Lukas focused on two important part of the improvement of detections in the network. First, he developed a useful and real software that implemented his proposed solution. The software was made available for the community and it is being widely used nowadays. Second, he experimented and research on how to improve the detection results by using a basic machine learning algorithm.

The first part of the thesis deals with the implementation of the solution in real code. This is a very hard task since the code should be good enough to run in large networks. This means that the student spent quite some time dealing with implementation issues. This work was very important for showing that his results are applicable to real networks.

In the second part of his thesis, the student dealt with a more difficult problem. How to compare different machine learning algorithms in order to find which solution was the best. The results are promising.

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

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