REVIEWER'S OPINION OF FINAL THESIS

I. IDENTIFICATION DATA

Thesis name:	**Graph-Based Analysis of Malware Network Behaviors**
Author's name:	**Daniel Smolik**
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.

This thesis was a challenging assignment. The student was faced with the problem of analyzing the behaviors of malware traffic with a graph representation. This means that there were several different topics to be studied, together with the implementation of the solution.

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.

The assignment was satisfactorily fulfilled. The student analyzed the problem, obtained the data, developed the solution and verify it with machine learning algorithms.

Method of conception **correct**

Assess that student has chosen correct approach or solution methods.

The student has chosen the correct path for solving his thesis. In particular the implementation of the graph analysis is very good and the results are sound.

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 first part of the thesis deals with an analysis of how to represent the behavior of malware in a new graph form. This idea is novel and took some time to be created. In this first part the student also implemented the graph. In the second part the student obtained a dataset of behaviors and used machine learning tools to identify the malware behaviors from the normal ones. It can be seen that the student learned from the literature and gained experience.

Formal and language level, scope of thesis **D - Satisfactory.**

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

The student managed to represent the formal idea of the graph accurately. However, the thesis lacks a good structure and deep explanations, making it very difficult to grasp the concepts. Some topics are explained in a vague way and this makes some descriptions confusing. Moreover, the English language needs a lot of improvement, further difficulting the analysis. The major weakness of the thesis is that most of the work of the student was not adequately represented in the thesis. There is a lot of 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.

The explanation of the previous work is barely enough for the thesis. This is a weak part of the work and it could be improved. It is true that the area of graph analysis is large, but a deeper understanding of the previous attempts to solve the problem could have improved the thesis. However, the student managed to create a technique that is novel for analysing the malware traffic.

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 primary goals of the thesis were fulfilled: a new technique to create the graph representations of malware traffic, the implementation of the solution in software, the usage of a dataset and the experiments to find the best solution. However, it is clear that the student failed to completely add to the thesis all the work he did. From the thesis it is very hard to follow and understand important parts of his work.

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

The work done by Daniel Smolik in his thesis can be separated in two parts. First, he created a new representation of the behavior of malware in the network and implemented it in software. Second, he developed new features for this graph based on the knowledge of how malware and normal computers behave in the network. In the second part, he extracted the features to train a machine learning algorithm that found the best combination of thresholds to identify the malware traffic.

This thesis is novel in the way the traffic is represented in the network, in the selection of features and in the results obtained. The implementation of the concept was very good and it was also important to show how the ideas can be put to the test in real networks. Finally, the results obtained are good and promising for future applications.

Given the technical implementation of this thesis and the novel idea presented, I evaluate handed thesis with classification grade **C - Good.**

Date: June 11th **2017, Prague.**

Signature: Sebastian Garcia