FACULTY OF INFORMATION TECHNOLOGY CTU IN PRAGUE

Review report of a final thesis

Student:	Bc. Daniel Nemčík
Reviewer:	Ing. Lukáš Machlica, Ph.D.
Thesis title:	Behavioural graph-based classification of infected network hosts
Branch of the study:	Computer Security

Date: 22. 1. 2020

Evaluation criterion:	*The evaluation scale: 1 to 4.*
1. Fulfilment of the assignment	***1 = assignment fulfilled,*** *2 = assignment fulfilled with minor objections,* *3 = assignment fulfilled with major objections,* *4 = assignment not fulfilled*

Criteria description:
Assess whether the submitted FT defines the objectives sufficiently and in line with the assignment; whether the objectives are formulated correctly and fulfilled sufficiently. In the comment, specify the points of the assignment that have not been met, assess the severity, impact, and, if appropriate, also the cause of the deficiencies. If the assignment differs substantially from the standards for the FT or if the student has developed the FT beyond the assignment, describe the way it got reflected on the quality of the assignment's fulfilment and the way it affected your final evaluation.

Comments:

The problem is well outlined, aims and focus of the thesis are stated clearly. The work is coherent, experiments are designed properly, and derived conclusions are logical and legitimate.

Evaluation criterion:	*The evaluation scale: 0 to 100 points (grade A to F).*
2. Main written part	*85 (B)*

Criteria description:
Evaluate whether the extent of the FT is adequate to its content and scope: are all the parts of the FT contentful and necessary? Next, consider whether the submitted FT is actually correct – are there factual errors or inaccuracies? Evaluate the logical structure of the FT, the thematic flow between chapters and whether the text is comprehensible to the reader. Assess whether the formal notations in the FT are used correctly. Assess the typographic and language aspects of the FT, follow the Dean's Directive No. 26/2017, Art. 3. Evaluate whether the relevant sources are properly used, quoted and cited. Verify that all quotes are properly distinguished from the results achieved in the FT, thus, that the citation ethics has not been violated and that the citations are complete and in accordance with citation practices and standards. Finally, evaluate whether the software and other copyrighted works have been used in accordance with their license terms.

Comments:

The thesis is well structured, individual chapters are logically ordered and contain the right amount of information in order to follow the experiments described in the last chapter of the thesis. List of tables and abbreviations is also provided, however it would be more beneficial to place the glossary of abbreviations at the beginning rather than at the end of the work. Citations in the work are used correctly. Chapters related to individual techniques are easy to read, but the reader found harder to read parts focusing on the description of the training steps in which the reference graph is build. I would appreciate more detailed discussion and formulas how the event relation matrix is constructed, how the reference graph is aggregated across different users and how exactly is the feature importance estimated.

Evaluation criterion:	*The evaluation scale: 0 to 100 points (grade A to F).*
3. Non-written part, attachments	*100 (A)*

Criteria description:
Depending on the nature of the FT, comment on the non-written part of the thesis. For example: SW work – the overall quality of the program. Is the technology used (from the development to deployment) suitable and adequate? HW – functional sample. Evaluate the technology and tools used. Research and experimental work – repeatability of the experiment.

Comments:

Chapter 5.3 provides detailed discussion on individual technologies used during the experimental part of the thesis. Student proved to be able to work with variety of state-of-the-art environments and services. A lot of experiments were carried out. They are well described and analysed.

Evaluation criterion:	*The evaluation scale: 0 to 100 points (grade A to F).*
4. Evaluation of results, publication outputs and awards	*95 (A)*

Criteria description:
Depending on the nature of the thesis, estimate whether the thesis results could be deployed in practice; alternatively, evaluate whether the results of the FT extend the already published/known results or whether they bring in completely new findings.

Comments:

I appreciate that even if the results are not perfect from the efficacy perspective and standard random forrest classifier outperformed the proposed solution, author did not hand picked any specific working points at which the results would shine. Instead, full evaluation curves are plotted and individual caveats are well discussed. The work was developed within CISCO research group, therefore a focus on practical application can be anticipated.

Evaluation criterion:	*No evaluation scale.*

5. Questions for the defence

Criteria description:
Formulate questions that the student should answer during the Presentation and defence of the FT in front of the SFE Committee (use a bullet list).

Questions:

Could the student elaborate in more depth on:
1) how the event relation matrix is constructed for the reference graph?
2) how exactly is the feature importance estimated?
3) how the reference graph is aggregated across different users?

Evaluation criterion:	*The evaluation scale: 0 to 100 points (grade A to F).*
6. The overall evaluation	*90 (A)*

Criteria description:
Summarize which of the aspects of the FT affected your grading process the most. The overall grade does not need to be an arithmetic mean (or other value) calculated from the evaluation in the previous criteria. Generally, a well-fulfilled assignment is assessed by grade A.

Comments:

The student demonstrated his ability to work with literature, state-of-the-art techniques and services, propose well founded experiments and derive conclusion based on the outcomes. Still, I found harder to follow the descriptions of how the reference extraction algorithm exactly works, and therefore got the feeling that this crucial section was written in a rush without proper polishing and clarification of the ideas.

Signature of the reviewer: