



Review report of a final thesis

Student: Michal Hloušek
Reviewer: Mgr. Jan Kohout
Thesis title: Machine learning based guest network detection
Branch of the study: Computer Security and Information technology

Date: 11. 6. 2020

<i>Evaluation criterion:</i>	<i>The evaluation scale: 1 to 4.</i>
1. Fulfilment of the assignment	1 = assignment fulfilled, 2 = assignment fulfilled with minor objections, 3 = assignment fulfilled with major objections, 4 = assignment not fulfilled
<i>Criteria description:</i> 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.	
<i>Comments:</i> The thesis describes clearly the problem and goals which the author wants to achieve. The objectives stated in the assignment were fulfilled.	
<i>Evaluation criterion:</i>	<i>The evaluation scale: 0 to 100 points (grade A to F).</i>
2. Main written part	90 (A)
<i>Criteria description:</i> 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.	
<i>Comments:</i> The work describes well the target domain and the problem being solved. The machine learning algorithms are discussed sufficiently as well, although I would prefer different ordering of the chapters - first to describe the problems being solved and challenges and then discuss the possible ML algorithms that might help to solve them (and why).	
<i>Evaluation criterion:</i>	<i>The evaluation scale: 0 to 100 points (grade A to F).</i>
3. Non-written part, attachments	95 (A)
<i>Criteria description:</i> 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.	
<i>Comments:</i> The experiments are well designed and performed in a very rigorous way. I appreciate the careful preparation of training and testing datasets to estimate the classification performance properly.	
<i>Evaluation criterion:</i>	<i>The evaluation scale: 0 to 100 points (grade A to F).</i>
4. Evaluation of results, publication outputs and awards	95 (A)
<i>Criteria description:</i> 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.	
<i>Comments:</i> The results bring interesting insights into the domain and could be possibly used as a basis for developing a production system which automatically identifies guest networks.	
<i>Evaluation criterion:</i>	<i>No evaluation scale.</i>

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:

page 24 - "we clip the client vectors to the interval $[0, 1]$ " - if zero means that a client was not active (had zero flows) and 1 means that the client issued at least one flow, what does a value between 0 and 1 mean?

More general discussion: Given that each subnet is represented by a vector which aggregates information about its behavior in multiple time windows, can we think about models/algorithms that could work with representations of individual windows directly, e.g., algorithms that would work with such sequences of observations/feature vectors?

Evaluation criterion:

The evaluation scale: 0 to 100 points (grade A to F).

6. The overall evaluation

93 (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:

Overall, I consider this work successful. I have minor objections regarding ordering of the chapters and introducing many different variables in the overview of ML techniques (which might be a little bit messy), but I like the experimental part, which was done above the standards of bachelor theses.

Signature of the reviewer: