FACULTY OF INFORMATION TECHNOLOGY CTU IN PRAGUE

Review report of a final thesis

Student:	Bc. Tomáš Stefan
Reviewer:	Ing. Josef Kokeš
Thesis title:	Security assessment of web application penetration testing tool
Branch of the study:	Computer Security

Date: 10. 1. 2021

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 assignment was fulfilled, although some parts would benefit from a more in-depth approach.

Evaluation criterion:	*The evaluation scale: 0 to 100 points (grade A to F).*
2. Main written part	*65 (D)*

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 written part follows the requirements of the assignment and describes all the necessary parts. I did notice a few minor factual errors: in particular, the claim that the application violates HTTP/1.1 specifications is not relevant because the request was not made in HTTP/1.1 - as far as I can tell, it was a valid, albeit unusual, HTTP/0.9. What is far more important, the work tends to skip over relevant preliminaries such as the theoretical background into security assessment methodologies (at least the one used) or the meaning of penetration testing; this also occurs with some technical descriptions. The assessment was done but the results are only informally described; I would expect a formal vulnerability assessment methodology to be used to evaluate the found weaknesses objectively. I very much miss explanations of motivation for various choices made throughout the analysis; in particular, the work doesn't even mention the possible alternative approaches such as reverse engineering and their possible effect on the results, much less explain why they were not used. Instead of addressing these issues, the thesis spends a lot of time on copying the source code of the implementation without any clear benefit to the reader.

Regarding the formal side of the thesis: It is written in very good English. In places I found the language a bit cumbersome, but it never prevented me from understanding what the student wanted to explain. The typography is fine, except for page 38 where the footnote overflows to page 39 and combines with another footnote in a rather nasty way. Speaking of footnotes, their number should be significantly reduced, 31 is way too much. The most significant formal flaw is the structure of the chapters: Chapter 1 has one subchapter which has one subsubchapter which then contains a number of lesser sections. Chapter 2 has two subchapters, but the first one is extremely short. Chapter 3 is long and I don't think all text in it can actually be described as "Practical part" - many explanations which should have been made earlier are also present here; overall, the structuring feels like the thesis was written in chronological order, without the necessary restructuring of the content when it became apparent that a new area needs to be opened. Often, a subsection follows a section immediately, without any introductory text in between (e.g. section 2.1, 2.2.2.1 and others).

The bibliography is large but redundant - there are no less than 9 references to RFC6455, 7 references to RFC2616, and others; if different pages of the same source are to be referenced, which in tself is a good thing, it should be done through an attribute of that reference rather than by creating a new reference.

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

3. Non-written part, attachments

80 (B)

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:

The non-written part is very short but adequate for the student's purpose - that is, it does what the author set to do. I do have some complaints about the chosen approach, but they mostly go back to the written part and its rather weak theoretical foundation - for example, I think that it is not enough to only generate valid inputs and verify that they get handled properly by the application; it is perhaps even more important to generate *invalid* inputs and see how the application handles them.

I do not understand why the gathered data such as the TCP capture files were not stored on the attached medium and I am very much at a loss as to why the ZIP archive described in section 3.5.5.4 was shown in hex in Appendix B and not stored on the CD in its binary form.

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

4. Evaluation of results, publication outputs and awards

75 (C)

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:

The thesis describes the student's own original research into the security of Burps Suite which makes it valuable. The rather lacking formal foundations tend to reduce the value by casting doubt on the reliability of the results. The lack of discussion of alternate approaches magnifies this issue. Despite that I find the results useful, even if they are not as useful as they could be.

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:

1) How could reverse engineering enhance or support your findings, if at all?
2) What was the reaction of the developers of Burps Suite to your report?

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

6. The overall evaluation

70 (C)

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 presented thesis researches a very interesting topic and provides some quite interesting new results. For example, I find the undocumented REST endpoints to be an important and rather disturbing discovery. Unfortunately, the thesis seems to lack the solid theoretical foundations which makes it difficult to trust it fully - the missing discussion of alternatives in particular detracts from the reliability of the results. The structure of the written part also needs improvement. Overall, the work is an acceptable master thesis, but unfortunately fails to reach the level it could have attained. Despite that, I recommend it for the defense and grade it C-good.

Signature of the reviewer: