Martin Rehak
Department of Computers, FEE
Czech Technical University in Prague

Master Thesis Assessment

Student name: **Michal Stanke**
Thesis Title: **Identifying Similarities in malicious network behaviour**

This is the supervisor's assessment of Michal Stanke's Diploma thesis entitled "Identifying Similarities in malicious network behaviour".

Thesis Description

The goal of the thesis is to identify similarities between communication activities of a single malware samples and grouping them together according to this similarity. Typical activities may include command & control, malware monetization, secondary attacks and other behaviours. These behaviours can be noisy, unstable and our understanding of thesis goal does frequently evolve over time. Therefore, a single activity can be classified with several labels and split into several seemingly independent activities. Their grouping is needed in order to reduce the number of activities our users have to inspect and also in order to make the analysis more robust w.r.t classification changes. Robustness would come from the fact that a single, yet long-running activity can be classified into several classes due to the above-mentioned dynamic behaviour. it does also evolve as our understanding of the events that constitute the activity evolves. Grouping the activities together would allow us to create consistent classification and apply them retroactively, using all the knowledge gained since the first classification.

Thesis Assessment

Michal has been working within my team since his bachelor's thesis and has been a very active and well appraised team member and software engineer. Therefore, when the time came to select his diploma thesis topic, he was offered an opportunity to work independently on a research topic that became his thesis. With regret, I have to admit that the result of the work has not fully met my expectations. The thesis did formally achieve its goals, but the quality is lacking. For starters, the thesis lacks the title page. Then, comparatively large part of the work is spent on introduction, but provides little clarity on the goals and the specific context of the work itself. The (very brief) description of the related work would be more appropriate for a bachelor's thesis. But these are only secondary observations.

The main issue I take with the thesis is the lack of clarity in description of the method in Chapter 3. Even knowing the subject, it would be really difficult to understand the algorithm and its limitations. The same criticism applies to Chapter 5. The description of experiments is severely lacking and would not allow reproduction of the work. But my biggest issue is the lack of argumentation (or

empirical arguments) why is the presented solution optimal. Michal did design a solution and measured its properties. We know that some improvement can be achieved on the problem. But the work has not increased our understanding of the problem as much as it could due to methodological limitations.

As a supervisor, I must accept partial responsibility for this failure. But, during the last months, I did not have many opportunities to encounter Michal and to provide him with feedback. Also, besides an early draft, I have not seen the work before it was handed over to the university, which is unusual.

Despite the above considerations, I still consider this work as formally **fulfilling the objectives and acceptable as a diploma thesis. I propose the grade "D"** and wish Michal luck in his future career.

In Prague, on January 27, 2018

Martin Rehak