ČESKÉ VYSOKÉ UČENÍ TECHNICKÉ V PRAZE

SUPERVISOR'S OPINION OF FINAL THESIS

I. IDENTIFICATION DATA

Thesis name:	**Industrial Control System Security Analytics**
Author's name:	**Marcel Német**
Type of thesis :	master
Faculty/Institute:	Faculty of Electrical Engineering (FEE)
Department:	Department of Computer Science
Thesis supervisor:	Andreas Wespi
Supervisor's department:	External – IBM Research Zurich

II. EVALUATION OF INDIVIDUAL CRITERIA

Assignment **challenging**

Evaluation of thesis difficulty of assignment.

The thesis is about behavior-based security analytics, also referred to as anomaly detection, in Industrial Control Systems (ICS) networks. Behavior-based security analytics systems are powerful technologies to detect new, so far not known attacks or system misbehaviors. However, such behavior-based systems cannot be deployed "out of the box". Some tuning and adaptation to a given environment are needed. Current tuning approaches are either ad-hoc and manual, or they are automated without giving the security professional much insight into why a certain parameter setting is better than another.

As experience shows, expert knowledge is required to configure and tune behavior-based systems. A challenge in ICS is that there is a multitude of sensors and therefore also a multitude of sensor data streams. A purely human-based approach does not scale. There are simply too many data streams to be configured.

This is where Marcel Német's thesis start. The objective of Marcel's thesis is to build a solution that assists the security professional to easily select the right analytics algorithm and to tune it for the data stream in scope.

The thesis assignment is challenging because there are various concepts and technologies involved. One has to

- understand the specifics of ICS and the data produced by ICS devices
- understand the principles of machine learning in general and the specifics of behavior-based security analytics systems in particular
- have very good software engineering skills to build a usable solution that assists in selecting and tuning behavior-based security analytics systems

Furthermore, the thesis was written in an industrial research environment where the students are not just given well defined tasks but where room is left for exploring new ideas and approaches.

Satisfaction of assignment **fulfilled with minor objections**

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.

Initially five tasks were defined:

1. *Familiarize with the problem of analyzing ICS protocol data.*
   Marcel has a deep analytics background. He used his skills to analyze the ICS network data streams that we collected in a real ICS environment. Marcel is familiar with the tools that can help to analyze large volumes of data and visualize specific aspects of the data. It was very interesting to discuss with him the results he had obtained. In that sense he has clearly met the goal of this task.

2. *Identify existing analytics modules that can be used for the analysis of ICS protocols. Develop own analytics modules.*
   It is very demanding to develop a new anomaly detection module. I was aware of this but Marcel was confident

that he could develop his own analytics solution. Therefore, we included this task. However, during the course of the thesis it became obvious that a lot of work would be needed to achieve an excellent solution. Therefore we decided to skip this task and spend more time on the main thesis goal, i.e., task 3.

3. *Design and implement a platform that allows a user to interactively select the best analytics module or combination of modules for a given task.*
   The platform that Marcel has built is an extension of an already existing ICS data exploration and analytics platform that was developed at IBM Research - Zurich. Marcel familiarized himself very quickly with the IBM platform and its features as well as with the underlying code base. Although many of the technologies that were used to build the platform were new to him, within short time he got to a level where he could make his own contributions.
   Surprisingly fast he came up with a first version of the assistant platform that security professionals can use to tune behavior-based security analytics systems. He further developed the assistant platform in an agile style. The final result is a very good solution that IBM plans to use in follow-up projects. I am very pleased with this part of the work.
   I can also mention that during the assessment of the solution by some test persons the newly developed software has always worked.

4. *Test the platform with ICS network data.*
   Marcel has continuously used the ICS network data to test and validate his software solution. He repeatedly presented the newest enhancements of the assistant platform based on our ICS network data. This task I consider fulfilled although in the thesis report little information can be found about this task.

5. Assess usability of the analytics platform and the quality of the analytics solution.
   The usability of the platform has been assessed. Several people working at IBM Research – Zurich tested and assessed Marcel's solution. Ideally the test persons would have been security professionals working in an ICS environment. However, the ICS security solution developed by IBM Research – Zurich has only been used in pilot projects so far. Therefore, there were no security professionals available who could have compared the manual tuning with the functionality provided by the newly developed assistant platform.

Activity and independence when creating final thesis **D - satisfactory.**

Assess that student had positive approach, time limits were met, conception was regularly consulted and was well prepared for consultations. Assess student's ability to work independently.

I liked the interactions with Marcel. He used to ask critical questions and contributed interesting ideas in our meetings. Marcel has a lot of interest in many areas. He likes to understand technologies in detail and to try them out.

There were phases where he was very efficient and produced excellent results. For example, he very quickly got acquainted with IBMS's ICS security analytics platform and within short time he had a first version of his solution running.

However, there were also phases where there was little progress. This led to the situation that eventually there was little time left for writing the thesis report. Aside from jointly discussing the overall structure of the thesis report*, Marcel wrote the report on his own. There was no time left for doing a review and providing feedback before Marcel had to submit the thesis report.

* In my evaluation I use the terms "thesis" and "thesis report". Thesis covers the overall work that comprises both the software solution Marcel has built and the "thesis report".

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.

From a technical perspective, Marcel has done a very good job. He has become an expert in various technical areas, and he demonstrated that he can find solutions to technical problems on his own. He was able to identify related literature and adapt it to his own work.

While initially he did not have much experience in doing larger software projects, he started to like software engineering and became more and more fluent.

Formal and language level, scope of thesis **C - good.**

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

The structure of the thesis report is sound. Marcel has a very good command of English. There are various grammatical mistakes that are repeatedly made. For example, quite often there is a missing article. There are several typos but this is a minor issue. The overall impression in terms of language is very good. Most parts of the thesis report are easy to read.

In terms of arrangement of the thesis report, there are several deficiencies. The problem to be solved is not well introduced, and it is not clearly stated what the contribution of the thesis is. Most of the sections are rather short and the content often deviates from what is expected based on the section title. The "assessment and evaluation" section is much too long compared with the other sections.

Providing additional material in form of appendices is a good approach. However, there has to be some valuable content in an appendix. In appendix C, providing screenshots without much description does not invite the reader to have a closer look at the screenshots. Appendix D shows that Marcel did a thorough job in recording and analyzing the assessment sessions. However, this is detailed information that does not add much to the understanding of the thesis.

Marcel tried to introduce some formalism by giving definitions for key terms being used. In particular in Section 4 many definitions are given. However, definitions are not always precise, see for example the definition of an anomaly (Def. 4.1) , or definitions are introduced but not used later on. Also, there are terms such as Precision and Recall that are introduced but not in the form of a definition. This means that from a formal notation point of view there is certainly room for improvement.

Selection of sources, citation correctness **A - excellent.**

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 thesis report references sources in the areas of ICS technology in general, ICS security analytics, automated parameter optimization in behavior-based analytics systems, and software engineering. All the relevant areas the thesis touches upon are covered. Citation ethics has not been breached, and the bibliographic citations are complete and in accordance with citation convention and standards.

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 thesis deliverable consists of two parts: the thesis report and the software platform that supports security professionals in tuning anomaly detection systems. While I am very pleased with the software solution, in particular considering that Marcel has continuously improved his software engineering skills, I am less happy about the quality of the report. The report has a lot of deficiencies due to the fact that Marcel had to finish it under time pressure. Although we had defined a realistic time plan, Marcel underestimated the effort of writing the thesis report. With some little assistance from my side I am sure that Marcel could have turned in a nicely written report that reflects all the good work he has done during his thesis time.

SUPERVISOR'S OPINION OF FINAL THESIS

III. OVERALL EVALUATION, QUESTIONS FOR DEFENSE, CLASSIFICATION SUGGESTION

Summarize thesis aspects that swayed your final evaluation.
I am very pleased with the software solution Marcel has developed. I rate this work with a "B". I am less happy about the quality of the thesis report. For this part I give a "D". I consider both parts equally important and therefore give an overall grade "C".

I evaluate handed thesis with classification grade **C - good.**

Date: **21.1.2017** Signature: