Supervisor's opinion
Ing. Tomáš Vintr

Martin Rektoris
Katedra řídící techniky - K13135
Anomaly detection in periodical stochastic phenomena

During the research part, student works very independently. He read quite a considerable amount of papers connected to the topic. Then he started to attend the consultations regularly, asking for complex and detailed information about applying the methods. During the last stage, he was extraordinarily active.
The student expected to get many different datasets for applying studied methods, but unfortunately, we could not gather the data for various reasons not connected to this thesis. Afterwards, when he found out there is no possibility to have real datasets, he decided to gather his own training and test dataset, which took him whole three weeks and heavily influenced the writing stage part of his thesis, during which he was forced to work more than 14 hours a day. I did not expect such an attitude, and therefore I have to grade him with A - excellent.

Classification: **A - excellent**

25.5.2021 in Prague

ČVUT
ČESKÉ VYSOKÉ UČENÍ TECHNICKÉ V PRAZE

REVIEWER'S OPINION OF FINAL THESIS

I. IDENTIFICATION DATA

Thesis name:	**Anomaly detection in periodical stochastic phenomena**
Author's name:	**Martin Rektoris**
Type of thesis :	bachelor
Faculty/Institute:	Faculty of Electrical Engineering (FEE)
Department:	Department of Control Engineering
Thesis reviewer:	Dr. Ransalu Senanayake
Reviewer's department:	Dept. of Aeronautics & Astronautics, Stanford University

II. EVALUATION OF INDIVIDUAL CRITERIA

Assignment — **challenging**

Evaluation of thesis difficulty of assignment.

Forecasting future and outlier detection are considered challenging in the fields of Statistics and Machine Learning. When applying these techniques to real-world problems such as social robotics and epidemiological modeling, it becomes even harder. Since the assignment of this thesis is an analysis of such outlier techniques in forecasting in robotic settings, I would consider this a challenging task.

Satisfaction of assignment — **fulfilled**

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.

The thesis adequately addresses the given assignment. It discusses various outlier detection algorithms and experimentally analyze them.

Method of conception — **outstanding**

Assess that student has chosen correct approach or solution methods.

The thesis considers various combinations of techniques to analyze and determine the most appropriate method.

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.

An adequate literature review is provided. Proposing a novel technique to address the specific problems in outlier detection in spatiotemporal models can further improve the technical quality of this work.

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

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

The thesis is easy to read. However, there are numerous language issues, especially in the abstract and introduction.

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.

I do not see any breach of ethics.

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.

Experimental results can be useful in future publications.

ČVUT
ČESKÉ VYSOKÉ UČENÍ TECHNICKÉ V PRAZE

REVIEWER'S OPINION OF FINAL THESIS

III. OVERALL EVALUATION, QUESTIONS FOR DEFENSE, CLASSIFICATION SUGGESTION

Summarize thesis aspects that swayed your final evaluation. Please present apt questions which student should answer during defense.

I evaluate handed thesis with classification grade **A - excellent.**

Forecasting spatiotemporal phenomena, as well as outlier detection, are extremely challenging and important problems, especially in the context of robots operating in dynamic environments and epidemiological modeling. This thesis provides an analysis of outlier detection for such spatiotemporal processes. Having performed a thorough literature review on forecasting methods and outlier detection methods in chapters 2 and 3, respectively, the author gathers real and synthetic datasets as described in chapter 4. An adequate rationale for using particular techniques is also provided. The thesis proposes using various combinations of forecasting and outlier detection techniques in chapter 5. A thorough analysis of all results is provided in chapter 6. In addition to measuring the accuracy, the thesis provides sensitivity analysis in terms of area under ROC curve. The thesis also provides a method for calibrating thresholds. My main criticism is about the numerous language issues which can be easily fixed before the final submission of the thesis. Although the technical content is very easy to follow, the use of language in the abstract, motivation, and literature review sections can be improved. Overall, excellent work!

Questions:

1. In chapter 4 of the thesis, six categories are provided. Would the model performance depend on the initial categorization? How can we perform this categorization for a new spatiotemporal phenomenon? What is the relationship between the number of categories (e.g. six) and the accuracy? Would it be worth learning the categorization within the model itself?
2. What are the chances of extending the proposed method to phenomena that exhibit aperiodic patterns or highly volatile patterns?
3. It is stated in the thesis that the STARIMA models cannot be used because of the irregular measurements. However, on contrary, in the datasets chapter, it is mentioned that "Every place was measured every thirty minutes with a small deviation as possible." Please elaborate.
4. Future predictions are always subjected to uncertainty. And, outlier detection is an integral part of uncertainty quantification. What are your thoughts on quantifying uncertainty in your predictions? How do we know what future predictions to believe in what degree of trustworthiness?
5. There are lots of external factors that affect spatiotemporal processes. For instance, in epidemiology, how people interact or travel; in robotics, how the environment changes, etc. How can we incorporate these various factors into the proposed framework?

Date: **31.5.2021**

Signature: