université angers

LARIS
Laboratoire angevin de recherche en ingénierie des systèmes
62 avenue Notre-Dame du Lac | 49000 ANGERS
Tél. 02 44 68 75 36 | Fax 02 44 68 75 61

POLYTECH

Angers, the 7th of September 2022

Review of Dr. Nizar CHATTI, Associate professor in the University of Angers

On the doctoral thesis of Mr. Martin RON intitled

« Pattern Discovery, Learning and Detection in Time Series »

The PhD manuscript, entitled « Pattern Discovery, Learning and Detection in Time Series», belongs to the great family of approaches dealing with time series analysis in continuous data (also known as functional data), motif discovery and detection which aim to summarize and visualize in a better way massive series databases and to mine motifs in time series. It can be used in various datamining tasks including classification, clustering...

In this research work, the main objective of analyzing time series data emanating from power consumption signals of robotic-manipulators, is to detect the production pauses (that may be either planned or unplanned) and to put the robots on a sleep mode when these pauses occur in order to decrease the energy consumption. The developed work is also generic and aims to be applied for any industrial process described by general repetitive tasks that can be observed in time series.

The overall manuscript is very well written and it clearly presents the contributions and the potential of applying them in different use cases. The state of the art of related techniques is also described, allowing the reader to easily appreciate the originality of the work.

The PhD manuscipt encompasses 8 chapters including an introduction and a conclusion. Each chapter focuses on a specific topic, provides separate definition of the goals and discusses the findings from simulations and experiments. The different chapters are complementary to each other and the overall structure of the manuscript is coherent and relevant.

Chapter 1 is an introductive chapter which starts by presenting the dataset used for the different simulations and experiments. Mr. Ron also recalls the steps leading to the obtained results and he highlights the genericity of the developed technique and the potential of applying it in diffrent use cases. Then, he briefly decribes the content of each chapter before listing the set of the achieved goals. Finally, Mr. Ron presents some notation rules that will be used throughout the manuscript.

LARIS | *Laboratoire angevin de recherche en ingénierie des systèmes* | laris.univ-angers.fr
62 avenue Notre-Dame du Lac | 49000 ANGERS
Tél. 02 44 68 75 36 | Fax 02 44 68 75 61

This synthetic chapter is a good basis for reading the manuscript and it briefly recalls the main concepts. Given the content of the overall manuscript and for sake of clarity, it would have been interesting if this first chapter had been extended to include a review of the most recently published works regarding energy time series forecasting, motifs discovery and detection, clustering of patterns as well as a brief summary of the mathematical fundamentals and definitions underlying these concepts. This would helped to better situate, from the beginning, the different contributions with existing relavant works. However, it is worth noting that the mentionned elements are clearly defined within the following chapters.

Chapter 2 presents the different steps leading to a characterization of a 6 axis KUKA KR5 arc robot dynamic model. This later will be used for defining the expected structure of the stochastic model. The different model assumptions have been clearly depicted and presented i.e. the assumption of rigid body dynamics, no estimation of external forces...
It would have been interesting to further discuss the impact of these assumptions and to provide a table with all parameters of the used KUKA KR5 industrial robot.

Chapter 3 focuses on behaviour pattern recognition by analyzing the robot manipulator operations in a welding cell based on its measured power consumption. First, a feature vector is extracted from the power data, the different sections are labelled and a model pattern is chosen after applying a median filter. Relevant segments are selected and similarity assessment is fulfilled after calculating Pearson correlation coefficient. Some of the power data have been measured from the real robot. However, due to some constraints related to the availability of robots in production environment and for being able to change the robotic paths to perform additional movements, a simulation environnement has been used i.e. Tecnomatix Process Simulate, making it possible to simulate the power consumption on a specific robot trajectory. The different steps have been clearly described and argued.

The proposed technique presents some limitations that have been also highlighted and discussed especially the fact that viable information are manually extracted from the original model pattern that need to include representative segments. Furthermore, the thresholds are tuned manually for both detection and feature extraction. It will be interesting to discuss with Mr. RON during the defense about the possibility for using other metrics for similarity detection as the RMS (Root Mean Square), the VAF (percentage Variance Accounted For) or the IoA (Index of Agreement). This can improve the accuracy of the obtained results.

Another problem deserves to be addressed namely the problem of sensor biased measurement which can decrease significantly the segment detection performances. The development of an intergrated auxiliary model to track and estimate in real-time any bias in used sensors, could be relevant to explore. This can be performed, for instance, by using a dedicated decision making-test called CR2 (Two Confidence Regions) that is able to perform robust and fast detection based on the parameters estimated by RARX (Recursive Auto Regressive with eXogenous input) approach. It is worth noting that the selected approach should satisfy technical requirements in terms of computational burden for viable implementation.

LARIS | *Laboratoire angevin de recherche en ingénierie des systèmes* | laris.univ-angers.fr
62 avenue Notre-Dame du Lac | 49000 ANGERS
Tél. 02 44 68 75 36 | Fax 02 44 68 75 61

In chapter 4, the motif discovery task is formulated. This chapter followed the structure of a submitted publication (ref [22]). The main contribution lies in creating a Motif Discovery Framework (MoDiF) that reduces in a significant way the number of motif candidates and it is based on the use of approximation methods for developing a modular approach to motif discovery. The contribution is relevant and all steps are detailed in a good way. The experimentation results show the relevance of the developed approach which has been compared with existing algorithms (i.e. CHIME, SCRIP++) using different real data sets.

The detection of on-line motifs in continuous data have been discussed in this chapter. However, some other techniques and particularly the ones that combine probabilistic models and polynomial least-squares approximations diserve to be further investigated and compared with the developed approach (for instance the works of Fuchs et al. (2009) or Mueen and Keogh (2010)).

Among the motif discovery algorithms, it would be interesting also to investigate and to compare the obtained results with the k-motif algorithm that has been used in different applications and which is also dependent on a pre-fixed pattern length. The algorithms can be compared using a divergence measure for probability densities.

Chapter 5 presents an unsupervised data preprocessing technique (based on OPTICS clustering algorithm) which segments-out repeating patterns and prepares a set of training examples for subsequent unsupervised CS-HGMM training.

Furthemore, The mode-detection CUSUM algorithm has been used to segment measured data with sufficient lag after the robot mode switches into the resting mode. The chapter is well structured and the main outcomes are highlighted. However, is would be interesting to further describe the used data presented on Fig. 5.4 and 5.5. What about simulated data presented in chapter 3 ?
Beyond the clustering algorithm, it is worth noting that a time series can be segmented with a data stream segmentation method for instance and the segments can be modeled using normal distributions with time-dependent means and constant variances. This can be investigated on future works.

Chapter 6 investigates the use of continuous-state hidden Gauss-Markov models CS-HGMM in order to identify the robot operations. This chapter is presented as a proof of concept and a use case evaluation has been considered. It would be interesting to provide further details about the HMM states and to present how The HMM-based classifiers is trained and tested. It would be also interesting to compare the obtained results with those of other approaches such as artificial neural network and naive Bayes classifier.

Chapter 7 presents an original approach of the continuity preference in training the time-varying hidden Gauss-Markov models (HGMM). The dynamic model presented in chapter 2 is used. The proposed approach has been tested using different datasets of different robot manipulators. It clearly appears that the models trained with the continuity preference are able to distinguish robotic operations and to evaluate the likelihood of an operation given a series of power data. The chapter is also well structured and the main outcomes are highlighted.

LARIS | *Laboratoire angevin de recherche en ingénierie des systèmes* | laris.univ-angers.fr
62 avenue Notre-Dame du Lac | 49000 ANGERS
Tél. 02 44 68 75 36 | Fax 02 44 68 75 61

The PhD manuscript ends with chapter 8 by providing final conclusions. The main contributions are also recalled, fulfillment of dissertation goals is discussed and future works are presented.

The entire manuscript is remarkably consistent. A single subject was discussed but from many points of view : theoretical, algorithmic, application. The obtained results defend the idea that transversality theory seems to be the right framework to discuss the time series analysis in functional data, motif discovery and detection. The scientific contributions of the thesis, which are developed in chapters 3, 4, 5, 6 and 7 are well-argued and well-constructed. The presented results are technical, solid and, to my knowledge, original. The developed work is also generic and aims to be applied not only for robot manipulators but also for any industrial process described by general repetitive tasks. For the different reasons set out in this report, I therefore recommend the thesis defense, which represents remarkable work on a both difficult and fascinating problem.

The author of the thesis proved to have an ability to perform research and to achieve scientific results. I do recommend the thesis for presentation with the aim of receiving a Ph.D. degree.

Angers, the 7th of September 2022

Signature

Dr. Nizar CHATTI

LARIS | *Laboratoire angevin de recherche en ingénierie des systèmes* | laris.univ-angers.fr
62 avenue Notre-Dame du Lac | 49000 ANGERS
Tél. 02 44 68 75 36 | Fax 02 44 68 75 61