Change Point Detection Algorithms for Identifying Bearing Damage Onset

Havikja van As – Bachelor student at The Hague University of Applied Sciences

In this project, we explored the use of Change Point Detection (CPD) algorithms for identifying early onsets of bearing damage, a crucial step in predictive maintenance. We evaluated two CPD algorithms—Relative Unconstrained Least-Squares Importance Fitting (RuLSIF) and Model-Fitting Algorithm—using vibration and load data both from the Bearing Test System (BETSY) developed in the PrimaVera project and from publicly available datasets. The results indicated that the Model-Fitting algorithm performs best, and although there were some challenges in obtaining real-world failure data and further testing is needed, our findings suggest that CPD algorithms can effectively detect bearing faults.

Introduction

Bearings are critical components in many machines, including ship engines, and their failure can cause costly downtime. To prevent this, various techniques are being investigated to identify the early onset damage in bearings and predict their remaining useful life (RUL). In this student project, we explored the use of **Change Point Detection (CPD) algorithms** for the early identification of bearing damage.

Data for this study were collected with the Bearing Test System (BETSY), built by The Hague University of Applied Sciences (THUAS). With BETSY, bearing operation can be simulated under controlled conditions while vibration and load sensor data are collected, which provide clues about the bearing's condition. However, due to technical challenges related to BETSY, publicly available datasets [1] were also used, which contain similar vibration data from damaged bearings.

The goal of this study was to find a CPD algorithm capable of **detecting the onset of bearing deterioration** with minimal error, irrespective of the type of bearing damage or the specific system it occurs in. CPD algorithms work by identifying moments when the statistical properties of sensor data change, signaling potential damage. However, these algorithms must also distinguish between damage-related changes and routine variations, such as adjustments to operational settings.

The central question of this study was:

"Which Change Point Detection algorithm can identify, with minimal error, the initiation of bearing damage, regardless of the damage type or system?"

To answer this question, we focused on:

1. Identifying which sensor measurements best indicate bearing damage;
2. Reviewing existing CPD algorithms to find candidates suitable for real-time monitoring;
3. Testing the performance of selected algorithms using vibration and load sensor data;
4. Evaluating algorithm performance; and
5. Assessing whether the algorithms can differentiate between damage-related and operational changes.

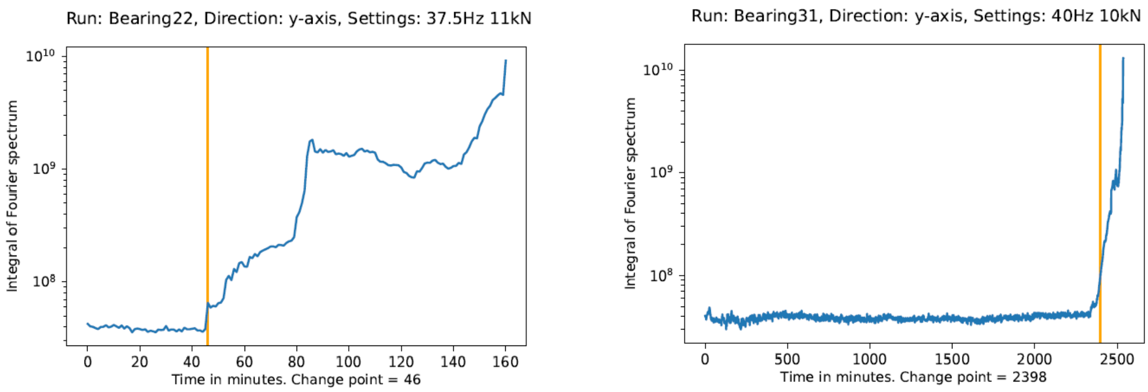

Figure 1: Bearing vibration energy time series from XJTU-SY data [1] and manually added change points (orange lines). Left: bearing outer race fault after 2h 41 min under 37.5 Hz rotation and 11 kN axial load. Right: bearing outer race fault after 42 h 18 min under 40 Hz rotation and 10 kN axial load.

Methods and approach

Data acquisition

Data from BETSY. The sensors on the THUAS Bearing Test System measure parameters such as vibration in three directions (*x*, *y*, and *z*), the axial force, bearing house temperature, and rotational speed (in rotations per minute, or RPM). With BETSY, various operational conditions can be replicated by adjusting the loads on the bearing and the speed of the motor. However, due to technical constraints, BETSY could not generate run-to-failure data during the study period.

Data from XJTU-SY. [1] To compensate for BETSY's limitations, the study utilized 15 run-to-failure datasets provided by the Institute of Design Science and Basic Component at Xi'an Jiaotong University (XJTU), Shaanxi, P.R. China and the Changxing Sumyoung Technology Co., Ltd. (SY), Zhejiang, P.R. China. [1]. These datasets were collected using a similar test system and included vibration measurements in two directions (*x* and *y*). Bearings in these datasets were also exposed to different operating conditions and exhibited different damage types, providing a diverse set of scenarios for algorithm training and evaluation. Due to the availability of only vibration data, we decided to focus on this measurement type only.

For both sets of vibration data, the measurement frequency f_m was 1 sample per minute (1/60 Hz) and the sampling frequency f_s was 25 kHz. The XJTU-SY datasets contain 32,768 data points per measurement (measurement duration of 1.28 s) and the BETSY dataset 250,000 (10 s).

Each vibration measurement was Fourier transformed and then the integral of the Fourier spectrum was calculated as to obtain one (energy) value per measurement. All further analysis was performed on these vibration energy time series (Figure 1).

Change point detection algorithms

A literature review identified two CPD algorithms as the most suitable for this application [2]:

1. **Relative Unconstrained Least-Squares Importance Fitting (RuLSIF)**. A non-parametric algorithm designed for detecting shifts in data distributions [3].
2. **Model-Fitting Algorithm**. A clustering-based approach that identifies change points when new data deviates from established patterns [4].

The algorithms were trained and tested on the XJTU-SY datasets [1] – which were split into training and validation sets – and tested on a single time series from BETSY where controlled changes were introduced in axial load and

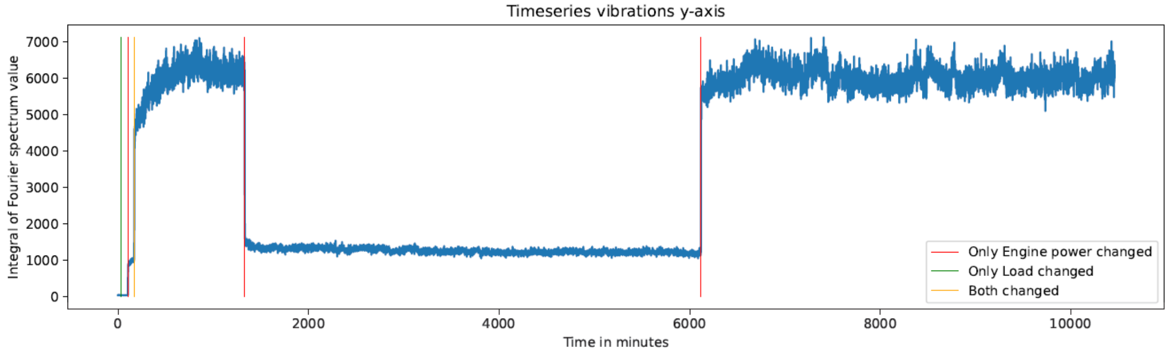

Figure 2: Preprocessed vibration data time series collected with the THUAS Bearing Test System (BETSY) with system (engine power and load) changes.

motor speed (Figure 2). This additional experiment was designed to determine whether the algorithms could distinguish between actual bearing deterioration and normal system fluctuations.

During training, the hyperparameters of the CPD algorithms, such as window size, kernel bandwidth, and significance thresholds, were fine-tuned using a threefold cross-validation.

Results and Discussion

Performance of CPD Algorithms

To measure performance of the CPD algorithms, we used **Mean Absolute Error** (MAE), a widely used metric in predictive modeling. Here, MAE quantifies the time difference between the actual and predicted change points, providing a clear measure of an algorithm's accuracy.

We found that the Model-Fitting algorithm is the most effective method, achieving an MAE of 40.4 min. This result suggests that the algorithm can reasonably estimate the initiation of bearing damage within a short timeframe, allowing for timely maintenance interventions. However, its ability to handle diverse damage types remains uncertain. While it performed well for most cases, it struggled with complex combinations of bearing failures in a single dataset.

RuLSIF, on the other hand, also showed potential but had a higher variability in its predictions. One possible limitation is its sensitivity to the assumption that the data are stationary, which may not always hold in real-world applications [5].

Robustness and False Positives

An important aspect of CPD algorithms is their ability to ignore irrelevant in-system variations. With BETSY, machine settings, in this case engine power / motor speed and axial load, can be manually and intentionally changed (Figure 1). Both algorithms successfully identified these system changes as non-damage-related, with a maximum error of only 1 minute. This indicates that CPD models can be trained to disregard routine adjustments, making them suitable for industrial applications.

Conclusion

We demonstrated in this study that Change Point Detection (CPD) algorithms are a viable tool for early detection of bearing damage, with the Model-Fitting algorithm showing the most promise. However, additional research is needed to confirm its ability to handle a wider variety of bearing failures and operate effectively across different mechanical systems.

Future work includes (1) expanding BETSY's capabilities to **collect more diverse run-to-failure datasets** with which we will be able to further validate the viability of CPD algorithms; (2) evaluating additional, more **complex CPD models**, such as Gaussian Process-based methods; and (3) optimizing **real-time implementation** for practical deployment.

References

[1] Wang, B., Lei, Y., Li, N., & Li, N. (2020). A Hybrid Prognostics Approach for Estimating Remaining Useful Life of Rolling Element Bearings. IEEE Transactions on Reliability, 69 (1), 401–412.

[2] Aminikhanghahi, S., & Cook, D. J. (2016). A survey of methods for time series change point detection. Knowledge and Information Systems, 51 (2), 339–367.

[3] Liu, S., Yamada, M., Collier, N., & Sugiyama, M. (2013). Change-point detection in time-series data by relative density-ratio estimation. Neural Networks, 43, 72–83.

[4] Tran, D. (2013). Clustering-based change point detection in time-series analysis. Journal of Machine Learning Research, 14, 3459–3492

[5] Mazzoleni, M., Antoniadis, P., Petrillo, A., & Travaglini, N. (2020). Machine learning-based approaches for predictive maintenance in mechanical systems. IEEE Transactions on Industrial Informatics, 16(4), 2376–2384.

Facts

Researchers	Havikja van As Derek Land Sam Aerts* John Bolte
Academic Partners	The Hague University of Applied Sciences (THUAS)

THE HAGUE UNIVERSITY OF APPLIED SCIENCES

* s.h.aerts@hhs.nl

PrimaVera – powered by

NWO