Detecting Failure in Jet Engines Using Uncertainty-based Changepoint Anomaly Detection

Ahmad Kamal Bin Mohd Nor^a, Srinivasa Rao Pedapati ^b, Masdi Muhammad^c

^{a,,b,c} Mechanical Department, Universiti Teknologi Petronas, Malaysia
 ^aahmad_18002773@utp.edu.my

Article History: Received: 5 April 2021; Accepted: 14 May 2021; Published online: 22 June 2021

Abstract: Anomaly detection in Prognostic and Health Management (PHM) domain by exploiting the information from deep learning uncertainty is presented in this paper. The behaviour of the uncertainty is monitored by cumulative sum (CUSUM) anomaly detection to detect abrupt changes in uncertainty, which translates the transition from healthy state to deterioration state. A probabilistic Long Short-Term Memory (LSTM) neural network is employed to predict the Remaining Useful Life (RUL) sequence distributions of engineered system. A case study of turbofan engines prognostic is presented to demonstrate the ability of this method. The proposed technique shows excellent result in term of Root Mean Square Error (RMSE) measure between ground truth anomaly and predicted anomaly and good result in scoring metric that evaluates the combination of early and accuracy of anomaly detection compared to the ground truth.

Keywords: Changepoint Detection, Anomaly Detection, Deep Learning Uncertainty, PHM, CMAPSS

1. Introduction

Engineered systems are vital for industrial operation. They ensure industrial process to be carried out as it should, adding value to products or services that is finally translated into financial gain. Downtime of these systems could result to issues related to safety, security as well as monetary to the organization. Efforts to improve reliability and availability of these machines have been undertaken by both industrial players and researchers, giving birth to various dedicated reliability engineering areas such as Multi State System (MSS) reliability and Human Reliability Engineering (HRA).

In recent decades, Prognostic and Health Management (PHM) has provided organizations with supports to manage the health condition of engineered system. PHM enables the improvement of reliability, safety, security as well as reducing the cost of maintenance in industrial context [1,2]. PHM activities consist of failure diagnostic, prognostic, and anomaly detection. Failure diagnostic is the activity of searching the root cause of failure while failure prognostic is the task of calculating the remaining useful life, i.e., the operational time of assets before failure. Anomaly detection, on the other hand, is a process to identify outliers in data points or events which deviate from a dataset's normal behaviour [3].

1.1. Anomaly Detection Techniques

Data anomaly can be divided into several categories: point anomaly, contextual anomaly, and collective anomalies. Point anomaly refers to individual point that is different from the rest of the data. Contextual anomaly points to data that is considered anomalous in specific context. Collective anomaly relates to manifestation of anomaly caused by a collection of data rather than individual data [3].

[4] presents an overview of anomaly detection techniques where 7 classes of method are described: classification-based (neural network, Bayes network, Support Vector Machine), Nearest Neighbour-based (K-Nearest Neighbour, relative density), clustering-based, statistical-based, information-theoretic-based, spectral-based and graph-based techniques.

1.2. Uncertainty-Based Anomaly Detection

Most research in deep learning applications are based on point estimates prediction. However, this is only experimental and cannot be employed in real world situation due to the absence of uncertainty quantification, which is vital for users to evaluate and trust the prediction. Uncertainty related to the quality of input data is called Aleatoric uncertainty. This kind of uncertainty occurs when data is contaminated with noise, stochasticity as well as error of acquisition.

In prognostic works, particularly when dealing with healthy-degradation-failure states, this uncertainty is expected to be relatively stable in healthy state, becomes suddenly unstable (increasing or dipping) at the degradation start point, before becoming stable again at the failure state. The uncertainty is supposed to be quite stable in the healthy and failure states as the variability of the predicted RUL sequence is low in these states. In

the degradation state however, the uncertainty becomes abruptly unstable due to the transition between healthydegradation and degradation-failure states. By monitoring the uncertainty behaviour between states, this instability can be exploited as an indicator for anomaly. Using deep learning, this instability can be amplified if the model is only trained with healthy part of the training data and tested with normal testing data. This is due to the model not recognizing the deterioration state of testing data, resulting to increase of prediction uncertainty.

The change of uncertainty level between states can be categorized as a change point detection (CPD) problem. Specifically, a time series CPD problem in this case. In CPD, the objective is to locate abrupt changes in data when a property of the time series changes [5]. [5] summarizes 2 classes of techniques available for time series CPD: supervised methods that maps input to target data [6] and unsupervised ones which comprise of likelihood ratio methods [7,8], subspace model methods [9,10], probabilistic methods [11,12], kernel-based methods [13,14], graph-based methods [15,16], and clustering methods [17,18].

In this paper, an unsupervised and statistical-based CUSUM anomaly detection method is applied to identify anomaly by exploiting the change in behaviour in prediction uncertainty. This paper claims to be the first work in anomaly detection utilizing the information of deep learning uncertainty behaviour. A case study of turbofan engines prognostic using probabilistic LSTM is employed to demonstrate the ability of this method.

2. Related Literature

Clustering technique is particularly popular in anomaly detection. In [19], anomaly detection for power electronic converters based on Principal Component Analysis (PCA) and K-Means clustering where healthy data and anomalous ones are identified from their clusters. PCA is used for feature extraction while k-means clustering with singular-value-weighted Euclidean distance is employed to define the healthy clusters. The same clustering technique is employed in [20] where anomaly in pressurized water reactor is detected using Coil current data clusters, that is classified as normal, mid-normal and off normal clusters. A semi-supervised Support Vector Data Description (SVDD) with negative samples (NSVDD) fault detection technique for rolling bearing element is proposed in [21]. The NSVDD model isolates the healthy and faulty data by using the Cyclostationary (CS) indicators to build the feature space and fits a hyper-sphere to calculate the Euclidean distances. In [22], K-Means clustering, and fuzzy modelling are used to detect anomaly of auxiliary marine diesel engine by identifying outliers. Additionally, the Event Score, that differentiates between real anomaly or false alarm is calculated from Local Outlier Factor (LOF) and Fuzzy membership function.

Supervised anomaly detection is another active research domain. Anomaly detection based on the difference between the measured and the predicted values of the AC power production for photovoltaic system is proposed in [23]. A threshold based on Hourly Lower Limit (HLL) and Hourly Upper Limit (HUL) are defined from the normal operating condition's standard deviation. In [24], Thus, the monitoring methodology by artificial neural network (ANN) to detect anomalies in the energy consumption performance of a compressed air generation system is described. The use of ANNs allows an accurate characterization of the system in a healthy state. Then, by comparing the model prediction and the actual energy consumption, residuals are calculated and plotted in a control chart. Anomaly detection for railway propulsion control systems is proposed in [25] by using Decision Tree algorithm for unsupervised learning and Naive Bayes Kernel and Ensemble Subspace KNN for supervised learning.

Changepoint detection works are equally important. An anomaly triggered-RUL estimation based on Cumulative Sum (CUSUM) control chart is proposed in [26] to detect anomaly in turbofan engines using streaming sensor data as input. In [27], a discrete wavelet transforms (DWT) modulus maximum for online change point detection by sliding dislocation window is applied for electric locomotive and forging machine anomaly identification. The current from these assets are exploited as input for the method. A stacking model consisting of Random Forest (RF), Gradient Boosting Decision Tree (GBDT), and Extreme Gradient Boosting (XGBOOST) is trained using healthy data to produce the normal or healthy model of wind turbine gearbox in [28]. Mahalanobis distance (MD) is incorporated to the model that serves as the change point detection measurement. Finally, in [29], Log-likelihood ratio estimated by Particle Filter for Lithium-Ion battery anomaly detection is applied. The method is equipped with an adaptive detection threshold for detecting the cycle at which the degradation behaviour undergoes changes in the dynamics, possibly due to faults and anomalies.

3. Methodology

3.1. Probabilistic LSTM

A probabilistic LSTM is employed to predict the RUL sequence distributions. The RUL distributions represent the health state of the studied assets from healthy state to failure. A probabilistic layer that output gaussian distribution with mean and variable standard deviation, characterizing the RUL distribution, is used in this output layer. The model is trained by only using the healthy part of the training data. The degradation and failure parts

are discarded. By doing so, the uncertainty of the prediction will show instability when the model is fed with the degradation part of the testing data, indicating anomaly.

3.2. Uncertainty Quantification & CUSUM Anomaly Detection

Uncertainty quantification will be evaluated via the rolling standard deviation plot of the Health State (HI) distribution sequence, which is the normalized RUL measurement of the system. An increasing trend indicates a growing uncertainty of the prediction while the contrary signifies that the model is more and more confident with the estimation. This change in trend especially in the transition point between health and degradation state can be an indicator of anomaly. CUSUM anomaly detection is employed to identify this change point phenomena.

The CUSUM calculates the positive (upper CUSUM) and negative (lower CUSUM) deviations of a normalised, random variable from a reference value and compares them to a threshold. For each time series instance, the lower and upper CUSUMs, initially equals to 0, are updated according to the formulas below:

$$C^{L}[n] = max(0, \frac{x[n] - \mu}{\sigma} - k + C^{L}[n-1])$$
(1)

$$C^{U}[n] = max(0, \frac{x[n] - \mu}{\sigma} - k + C^{U}[n - 1])$$
⁽²⁾

where μ and σ are the approximate mean and standard deviation of time series (x[n]) in the healthy period, and k is a predetermined reference value. Anomaly is flagged when $C^{L}[n] > h$ or $C^{U}[n] > h$ with h as the threshold.

3.3. Performance Evaluation

Root Mean Squared Error (RMSE) is used to calculate the performance of the model in anomaly detection. A score function, s, is also employed, destined to evaluate the combination of *earliness* and accuracy of the detection compared to ground truth anomaly. An early and accurate detection produces better score [30,31].

$$RMSE = \sqrt{\frac{1}{M} \sum_{i=1}^{M} (Anomaly_i^{truth} - Anomaly_i^{predicted})^2}$$
(3)

$$s = \sum_{i=1}^{M} s_i \tag{4}$$

$$s_{i} = \begin{cases} e^{\frac{-d_{i}}{13}-1}, d_{i} < 0\\ e^{\frac{di}{10}-1}, d_{i} > 0 \end{cases}$$
(5)

$$d_{i} = \left(Anomaly_{i}^{predicted} - Anomaly_{i}^{truth}\right)$$
(6)

With Anomaly_i^{truth} as the ground truth RUL for turbofan *i*, Anomaly_i^{predicted} the predicted RUL for turbofan *i*, and *M* as the total number of turbofans.

4. Result and Discussion

4.1. Case Study: Turbofan Engine Prognostic

The CMAPPS (Commercial Modular Aero Propulsion System Simulation) Turbofan run-to-failure datasets are comprised of four full sets of preparation, testing, and ground truth RUL for a variety of turbofan engines, published by Nasa Prognostic Centre (PCoE) of Ames Research Centre and designated as FD001, FD002, FD003, and FD004 [32]. Using CMAPSS software, this data was produced by modifying the operating conditions and injecting faults of varying degrees of deterioration into a simulated turbofan system [33].

In this analysis, the FD002 data was chosen. As shown in Table, this data consists of confirmed turbofan degradations whose health condition deteriorates after a certain cycle. Each turbofan has a time series sequence that includes Time (Cycle), 3 Operating Conditions (OC), and 21 sensor measurements that correspond to the system's temperature, pressure, various ratios, and bleed enthalpy. Different operating regimes (O-42K ft.), throttle resolver angle (20-100), and Mach number are referred to as OC (0-0.84). The effect of different operating conditions hides the faults found, and high levels of noise are introduced [33].

Table 1	. FD002	Dataset	Characteristics
---------	---------	---------	-----------------

Data	Fault Modes	Operating Conditions	Train Units	Test Units
FD002	2	6	260	259

4.2. Healthy Data for Anomaly Detection

The RUL target for the model's training data is obtained by using a piece-wise linear degradation model [34,35]. The health of each fleet is therefore initially considered stable until the failure start point, after which a linear deterioration occurs until failure.

The total operating period of a turbofan is represented by each time series sequence, with the last cycle representing the last instance before failure. Figure 1 shows how a turbofan's RUL is supposed to be equivalent to the previous cycle's value at first, then degrades linearly until it reaches nil (a). The turbofan 1 training data was collected over 192 cycles in this example.



Figure 1(a). Initial RUL Targets

Figure 1(b). Final RUL Targets

Using CUSUM, the first index of the upper or lower cumulative sums of each sensor's measurement that have drifted beyond 5σ threshold from the target mean indicates the initiating point of deterioration. The failure start point is determined by taking the average of all of these indexes. The transformed RUL series is shown in Figure 1 by combining the linear degradation obtained earlier with the failure start point (b).

Finally, the degradation part of the training data is removed to keep only the healthy part for anomaly detection, marked as *Stable* area in Figure 1(b).

4.3. RMSE and Score Result

As can be seen in **Table 2**, the method shows excellent result in RMSE measure and good result in Score *s* metric. The latter result indicates that even though all the detection is early, the accuracy of some detection is not very good.

Table 2.	Proposed	Method's	Performance
----------	----------	----------	-------------

Results		
RMSE	14.9	
Score, s	818.5	

4.4. Uncertainty-Based Anomaly Detection

In the following, prediction results of turbofan 2 in **Figure 2** and turbofan 4 in **Figure 4** are used as illustrations. As can be seen from **Figure 3** and **Figure 5**, the prediction uncertainty became increasingly unstable before the ground truth anomaly. CUSUM detected this behaviour early on when the standard deviation of turbofan 2 and 4 predictions exceeds the 5σ thresholds.









Figure 3. Turbofan 2 RUL Distribution Standard Deviation



Figure 4. Turbofan 4 RUL Distributions Prediction



Figure 5. Turbofan 2 RUL Distribution Standard Deviation

5. Conclusion

A changepoint detection based on deep learning's prediction uncertainty is proposed in this work. A single input, multi outputs probabilistic LSTM is employed, producing the RUL distributions estimation, characterizing the health state of engineered systems. The behaviour of the standard deviations of these distributions indicates the uncertainty state of the model's prediction. This uncertainty state can be exploited as an indicator of deterioration or anomaly particularly when the prediction transitions between heathy and deterioration state. By training the model using only healthy data and testing it with normal data, the uncertainty will show abrupt change when tested with anomalous data. This changepoint is detected using CUSUM anomaly detection with a predetermined threshold. Excellent results in term of RMSE measure and good result in score metric evaluating early, and accuracy of anomaly detection are obtained from the experiment with CMAPPS run to failure turbofan dataset.

References

- Biggio L, Kastanis I. Prognostics and Health Management of Industrial Assets: Current Progress and Road Ahead. Front Artif Intell. 2020 Nov 9;3:578613. doi: 10.3389/frai.2020.578613. PMID: 33733218; PMCID: PMC7861342.
- Bailey, Chris & Sutharssan, Thamo & Yin, Chunyan & Stoyanov, Stoyan. (2015). Prognostic and health management for engineering systems: a review of the data-driven approach and algorithms. The Journal of Engineering. 2015. 10.1049/joe.2014.0303.
- Haytham Assem, Lei Xu, Teodora S. Buda, Declan O'Sullivan, Chapter 8 Cognitive Applications and Their Supporting Architecture for Smart Cities, Editor(s): Hui-Huang Hsu, Chuan-Yu Chang, Ching-Hsien Hsu, In Intelligent Data-Centric Systems, Big Data Analytics for Sensor-Network Collected Intelligence, Academic Press, 2017, Pages 167-185, ISBN 9780128093931, https://doi.org/10.1016/B978-0-12-809393-1.00008-8.
- 4. Toshniwal, K. Mahesh and R. Jayashree, "Overview of Anomaly Detection techniques in Machine Learning," 2020 Fourth International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC), Palladam, India, 2020, pp. 808-815, doi: 10.1109/I-SMAC49090.2020.9243329.
- Aminikhanghahi, S., Cook, D.J. A survey of methods for time series change point detection. Knowl Inf Syst 51, 339–367 (2017). https://doi.org/10.1007/s10115-016-0987-z
- Ma, Jiangang & Sun, Le & Wang, Hua & Zhang, Yanchun & Aickelin, Uwe. (2016). Supervised Anomaly Detection in Uncertain Pseudoperiodic Data Streams. ACM Transactions on Internet Technology. 16. 1-20. 10.1145/2806890.
- Grana, Justin & Wolpert, David & Neil, Joshua & Xie, Dongping & Bhattacharya, Tanmoy & Bent, Russell. (2016). A Likelihood Ratio Anomaly Detector for Identifying Within-Perimeter Computer Network Attacks. Journal of Network and Computer Applications. 66. 10.1016/j.jnca.2016.03.008.
- Coluccia, Angelo & D'Alconzo, Alessandro & Ricciato, Fabio. (2013). Distribution-based anomaly detection via generalized likelihood ratio test: A general Maximum Entropy approach. Computer Networks. 57. 3446–3462. 10.1016/j.comnet.2013.07.028.
- Candy JV, Franco SN, Ruggiero EL, Emmons MC, Lopez IM, Stoops LM. Anomaly detection for a vibrating structure: A subspace identification/tracking approach. J Acoust Soc Am. 2017 Aug;142(2):680. doi: 10.1121/1.4996142. PMID: 28863574.

- Liangwei Zhang, Jing Lin, Ramin Karim, An angle-based subspace anomaly detection approach to highdimensional data: With an application to industrial fault detection, Reliability Engineering & System Safety, Volume 142, 2015, Pages 482-497, ISSN 0951-8320, https://doi.org/10.1016/j.ress.2015.05.025.
- 11. Ruggieri, Eric & Antonellis, Marcus. (2015). An exact approach to Bayesian sequential change point detection. Computational Statistics & Data Analysis. 97. 10.1016/j.csda.2015.11.010.
- 12. Akouemo, Hermine & Povinelli, Richard. (2015). Probabilistic anomaly detection in natural gas time series data. International Journal of Forecasting. 32. 10.1016/j.ijforecast.2015.06.001.
- 13. Zhang, Liangwei & Lin, Janet (Jing) & Karim, Ramin. (2017). Adaptive Kernel Density-based Anomaly Detection for Nonlinear Systems. Knowledge-Based Systems. 139. 10.1016/j.knosys.2017.10.009.
- Lindstrom MR, Jung H, Larocque D. Functional Kernel Density Estimation: Point and Fourier Approaches to Time Series Anomaly Detection. Entropy. 2020; 22(12):1363. https://doi.org/10.3390/e22121363
- Song S, Yang Y, Zhou H, Chan JC-W. Hyperspectral Anomaly Detection via Graph Dictionary-Based Low Rank Decomposition with Texture Feature Extraction. Remote Sensing. 2020; 12(23):3966. https://doi.org/10.3390/rs12233966
- 16. Bhatia, Siddharth & Hooi, Bryan & Yoon, Minji & Shin, Kijung & Faloutsos, Christos. (2019). MIDAS: Microcluster-Based Detector of Anomalies in Edge Streams.
- 17. Faroughi, Azadeh & Javidan, Reza. (2018). CANF: Clustering and anomaly detection method using nearest and farthest neighbor. Future Generation Computer Systems. 89. 10.1016/j.future.2018.06.031.
- 18. Candelieri A. Clustering and Support Vector Regression for Water Demand Forecasting and Anomaly Detection. Water. 2017; 9(3):224. https://doi.org/10.3390/w9030224
- Yizhou Lu, Aris Christou, Prognostics of IGBT modules based on the approach of particle filtering, Microelectronics Reliability, Volume 92, 2019, Pages 96-105, ISSN 0026-2714, https://doi.org/10.1016/j.microrel.2018.11.012.
- 20. Jung, Jae-Cheon & Oluwasegun, Adebena. (2020). The application of machine learning for the Prognostics and Health Management of control element drive system. Nuclear Engineering and Technology. 52. 10.1016/j.net.2020.03.028.
- "Liu, Chenyu & Gryllias, Konstantinos. (2020). A Semi-supervised Support Vector Data Descriptionbased Fault Detection Method for Rolling Element Bearings based on Cyclic Spectral Analysis. Mechanical Systems and Signal Processing. 140. 10.1016/j.ymssp.2020.106682. "
- 22. Diez-Olivan, Alberto & Pagan Rubio, Jose & Sanz, Ricardo & Sierra, Basilio. (2017). Data-driven prognostics using a combination of constrained K-means clustering, fuzzy modeling and LOF-based score. Neurocomputing. 241. 10.1016/j.neucom.2017.02.024.
- 23. Benedetti, Massimiliano & Leonardi, Fabio & Messina, Fabrizio & Santoro, Corrado & Vasilakos, Athanasios. (2018). Anomaly Detection and Predictive Maintenance for photovoltaic Systems. Neurocomputing. 310. 10.1016/j.neucom.2018.05.017.
- 24. Santolamazza, Annalisa & Cesarotti, V. & Introna, Vito. (2018). Anomaly detection in energy consumption for Condition-Based maintenance of Compressed Air Generation systems: an approach based on artificial neural networks. IFAC-PapersOnLine. 51. 1131-1136. 10.1016/j.ifacol.2018.08.439.
- Hodžić, D. Škulj and A. Čaušević, "Data-driven Anomaly Detection for Railway Propulsion Control Systems," IECON 2020 The 46th Annual Conference of the IEEE Industrial Electronics Society, Singapore, Singapore, 2020, pp. 4351-4356, doi: 10.1109/IECON43393.2020.9255026.
- Gurkan Aydemir, Burak Acar, Anomaly monitoring improves remaining useful life estimation of industrial machinery, Journal of Manufacturing Systems, Volume 56, 2020, Pages 463-469, ISSN 0278-6125, https://doi.org/10.1016/j.jmsy.2020.06.014.
- Haiyan Wang, Feng Zhou, Chunyang Jiang, Ling Qin, Huaiqing Zhang, "Change Point Detection for Piecewise Envelope Current Signal Based on Wavelet Transform", Journal of Electrical and Computer Engineering, vol. 2018, Article ID 9529870, 10 pages, 2018. https://doi.org/10.1155/2018/9529870
- Yuan, Tongke; Sun, Zhifeng; Ma, Shihao. 2019. "Gearbox Fault Prediction of Wind Turbines Based on a Stacking Model and Change-Point Detection" Energies 12, no. 22: 4224. https://doi.org/10.3390/en12224224
- Cadini, Francesco & Sbarufatti, Claudio & Cancelliere, Francesco & Giglio, Marco. (2019). State-of-life prognosis and diagnosis of lithium-ion batteries by data-driven particle filters. Applied Energy. 235. 661-672. 10.1016/j.apenergy.2018.10.095.
- X. Zhang, Y. Dong, L. Wen, F. Lu and W. Li, "Remaining Useful Life Estimation Based on a New Convolutional and Recurrent Neural Network," 2019 IEEE 15th International Conference on Automation Science and Engineering (CASE), Vancouver, BC, Canada, 2019, pp. 317-322, doi: 10.1109/COASE.2019.8843078.
- 31. K. Deng et al., "A Remaining Useful Life Prediction Method with Automatic Feature Extraction for Aircraft Engines," 2019 18th IEEE International Conference On Trust, Security And Privacy In

Computing And Communications/13th IEEE International Conference On Big Data Science And Engineering (TrustCom/BigDataSE), Rotorua, New Zealand, 2019, pp. 686-692, doi: 10.1109/TrustCom/BigDataSE.2019.00097.

- 32. Ramasso, Emmanuel & Saxena, Abhinav. (2014). Performance Benchmarking and Analysis of Prognostic Methods for CMAPSS Datasets. International Journal of Prognostics and Health Management. 5. 1-15.
- Saxena, K. Goebel, D. Simon and N. Eklund, "Damage propagation modeling for aircraft engine run-tofailure simulation," 2008 International Conference on Prognostics and Health Management, Denver, CO, USA, 2008, pp. 1-9, doi: 10.1109/PHM.2008.4711414.
- H. Miao, B. Li, C. Sun and J. Liu, "Joint Learning of Degradation Assessment and RUL Prediction for Aeroengines via Dual-Task Deep LSTM Networks," in IEEE Transactions on Industrial Informatics, vol. 15, no. 9, pp. 5023-5032, Sept. 2019, doi: 10.1109/TII.2019.2900295.
- W. Zhang, F. Jin, G. Zhang, B. Zhao and Y. Hou, "Aero-Engine Remaining Useful Life Estimation based on 1-Dimensional FCN-LSTM Neural Networks," 2019 Chinese Control Conference (CCC), Guangzhou, China, 2019, pp. 4913-4918, doi: 10.23919/ChiCC.2019.8866118.