HUMS2023 Data Challenge Result Submission

Team Name: LM-RMS-DAI

Team Members: Nathaniel Rigoni, Daniel Wade

Institutions: Lockheed Martin Rotary and Missions Systems

Publishable: Yes

1. Summary of Findings

A neural network tuned for multiplexed time series Virtual Sensor (VS) anomaly detection is trained and evaluated against a faulted epicyclic transmission in this paper. The VS is shown to provide good estimation of fault presence and progression of an artificially induced planetary gear crack. The team used three varieties of the same VS architecture: leave-one-out (VS-LOO), two-class (VS-2C), and unsupervised operations clustering (VS-UOC). This paper focuses on the VS-LOO and VS-UOC.

Each aspect of the performance versus the fault detection and growth is tied to a combination of the VS capabilities (Table 1). We claim detection (row 1) when the VS-LOO exceeds the threshold across multiple readings. Due to the nature of our method, we do not claim a difference between multi and single sensor detection (row 2, 3). We find that VS-LOO error increases as the fault progresses from the Day 22 minima (row 4). We claim exponential growth detection and state change based on VS-UOC and VS-LOO (row 5).

Table 1 Summary of Analysis Results

#	Detection & Trending	Data file name/number	Comments
1	Consistent detection on at least one signal channel; i.e. the fault indicators remain consistently above the threshold.	Day025 20220111_100623	Figure 5, reconstruction error consistently greater than 40
2	Confirmed detection on at least two signal channels; i.e. the fault indicators remain consistently above the threshold.		
3	Clear multi-channel indication of the characteristic fault features; i.e. faulty planet gear meshing with both the ring and sun gears.		
4	Confirmed trend of fault progression; i.e. a consistent increasing trend started from which file number/name.	Day022 20211209_125436	Figure 5, Error minima
5	Confirmed trend of accelerated fault progression; i.e. a consistent exponential increasing trend started from which file number/name	Day025 20220111_153659	Figure 5, Cluster Change, Error Change

2. Analysis Methods

Following review of recent investigations (1, 2) into epicyclic transmission fault detection capabilities the team decided to use our multiplexed time series VS anomaly detection python module to evaluate this new test rig data. Our VS module has multiple flavors of modeling, of which three are considered applicable to this problem: VS-LOO, VS-2C, and VS-UOC. These flavors are deployable into modern HUMS data analysis pipelines and could be combined with traditional level alert and physics of failure methods.

VSs are typically trained from a wide variety of nominal operations and result in a physics-informed-data-driven algorithm for anomaly detection and unsupervised operations-clustering. Typically we suggest a hybrid approach (HUMS plus data driven methods) to guide customers in diagnosing and prognosing incipient faults which can help avoid the pitfalls of under-performing HUMS algorithms, and non-ideal sensor choices/placements. This paper shows only the data-driven approaches.

For the VS-LOO, the team used reshaped, 99-point planetary gear sequences from the provided data. The sequences were tagged with time stamps and ordered for input into our VS module. The neural network uses an encoder-decoder multi-headed attention architecture. For this data challenge the neural network receives input from three randomly selected sensors and the output predicts the fourth held back sensor. The randomly selected reconstruction sensor is held constant for all remaining reconstructions.

The VS-LOO sensor (input pinion flange, IP) was chosen at random from among the four total accelerometers: IP, upper housing flange left and right (RF, RL), and rear of gearbox (RR). The training was taken from the first 250 000 (99 point) sequences. Early stopping validation was performed on sequences 250 000 through 300 000. This train-validate duration was selected after a basic design of experiments was completed, reviewing results from multiple durations based on overall in-sample error performance. It is important to note that the duration was kept at the absolute minimum so as to learn from the *nominal* vibration relationships. The remainder sequences are treated as out-of-sample testing (1 853 970 total). The performance of the trained VS is evaluated by computing the signal reconstruction error.

$$\sqrt{\sum (actual - predicted)^2}$$

From a HUMS perspective, the VS-LOO can be considered a Condition Indicator (CI) computed from the time synchronous average, based on a 99 point sequence. Figure 1 shows an example reconstruction.

It is helpful to improve the model outlier performance through the application of a median filter, a common method used in real-time CI computation on currently deployed HUMS (e.g. MSPU). The median filter also improves visual distinction when trending the CI as a time-series. The median filtered output shown in this paper results in a total of 2 152 995 points. Median error per sequence is also aggregated as overall *median error per data file* as our final CI, shown in Figure 5.

While the team did develop a VS-2C model, it is omitted from this paper due to space constraints. The model performed well, but provides less value to the discussion.

The third method explored uses the entire dataset along with unsupervised clustering to create our VS-UOC implementation. The team used the Fourier transform of sequences from each sensor concatenated together to discover clusters of operation. The VS-UOC uses the output from the neural network hidden state of the encoding layer to create a ***planetary operating mode encoding***. These multidimensional encodings result in *modes* of the synchronous time average spectra separated by distance in the n-dimensional space (Figure 2). We explored several settings for the VS-UOC, and finally settled on the simplest, 2 separable clusters (0 and 1) and one non-separable group (denoted as the -1 *cluster*). The -1 *cluster* is not a cluster, rather data that are part of this group cannot be clustered; we refer to this group as the ***noise cluster***.

Regardless of the number of clusters found by the algorithm (which is related to cluster sensitivity settings), they are related to each other through hierarchical clustering, i.e. certain clusters are sub clusters of higher level clusters. The algorithm seeks to continuously distinguish new clusters until it reaches a minimum error state at which point it stops and returns the learned cluster labels.

The latent space representation clustering is shown in a two-dimensional space for each sequence in Figure 2. The embedding was trained using reshaped 99 point planet gear sequences per all four sensors. The first group, cluster 1 in green, is the latent space representation of the normal operating behavioral modes of the test article. The centroid space, cluster 0 in red, is a smaller operational mode of the dataset, appearing to be related to startup sequences (see Figure 5 for temporal representation). The third group, cluster negative 1 in blue, is a transient state of behaviors that are unique or do not occur in a great enough frequency to be clustered. These cluster groups identify groups of behavior of the gearbox as measured across all sensors. It is assumed that internal events occur to transition these groups away from the nominal behavioral group, e.g. crack propagation, test article start up, or fault state events.

3. Illustrating Figures

The team presents two figures in this section to illustrate how the VS-LOO and VS-UOC algorithms function. Figure 1 shows three example reconstructions from 'Day021_Hunting_SSA_20211208_110248'. The sequence lengths shown are 750 time steps in duration and the actual and predicted IP sensor signals are shown together. Each of these files are rolled up to a final reconstruction error CI.

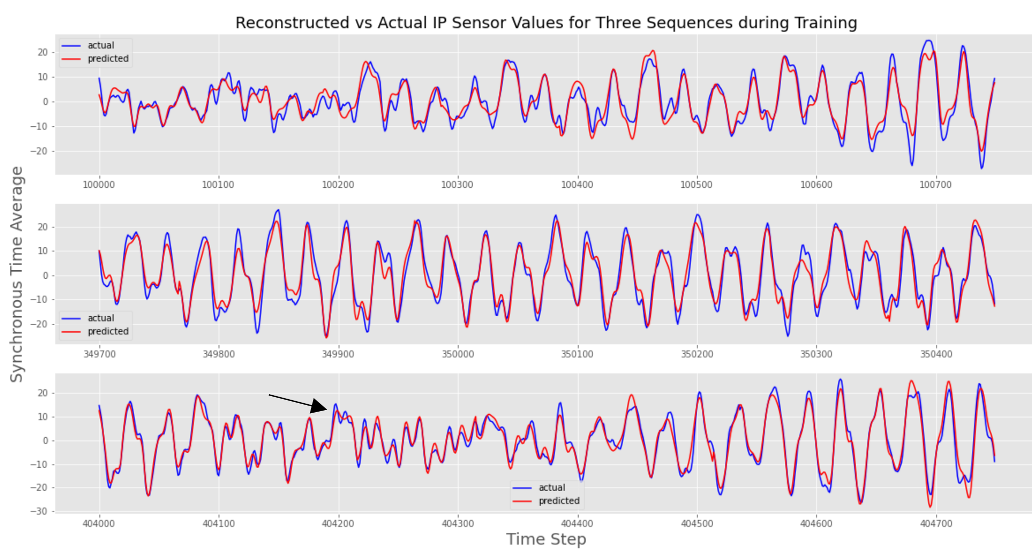

Figure 1. Example VS-LOO reconstruction for three sequences on the *assumed nominal* Day 21 of testing. The details of *good* high frequency reconstruction are pointed out in the third sequence.

Figure 2 shows the VS-UOC algorithm results in two dimensional space. These groups indicate behaviors of the gearbox as measured by the vibrational response. Groups occupying different areas in this latent space are observed to be independent and separate behaviors as observed by the model. Temporal associations can be seen in Figure 5.

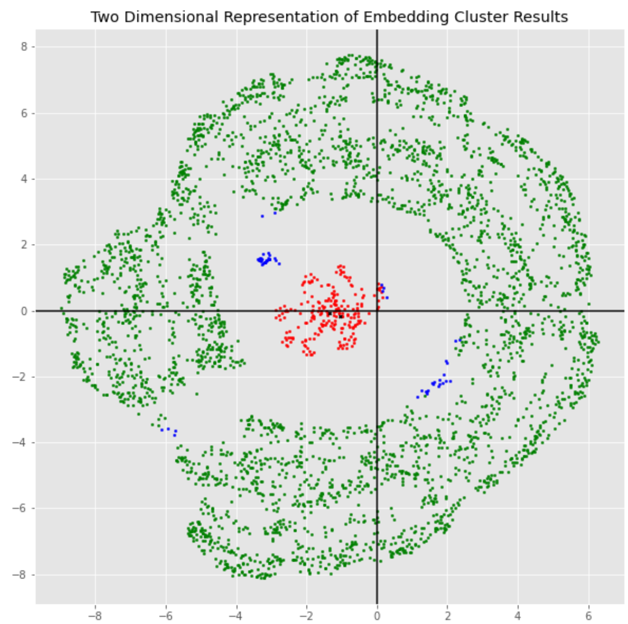

Figure 2. Latent space clustering of vibrational behaviors of the planet gear with three clusters.

4. Characteristic Fault Signatures of Early Detection

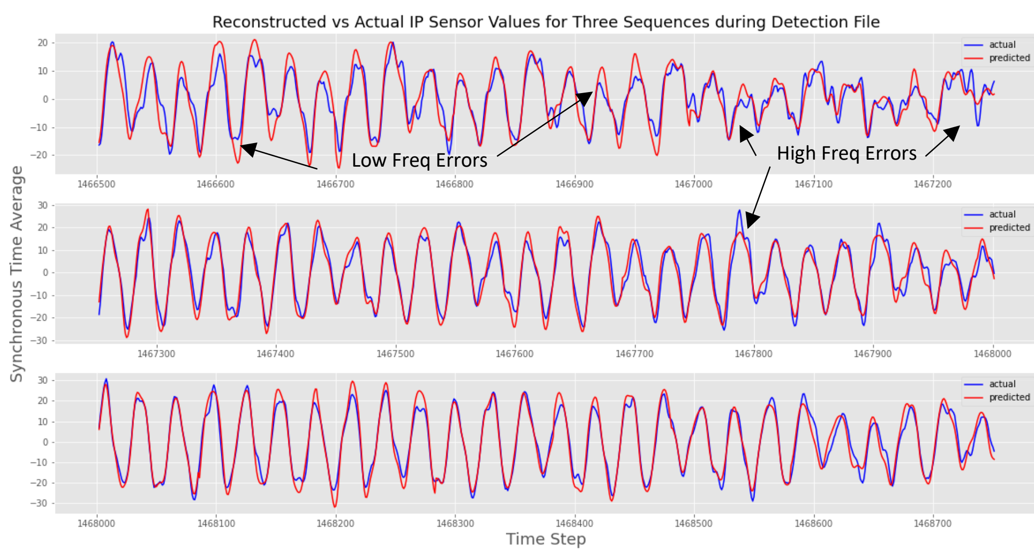

Figure 3. Example reconstruction at the detection point for the VS-LOO algorithm.

The VS method chosen for this data challenge uses all four of the installed sensors. After completion of results based on the IP sensor, the team did review one additional sensor (RF) to verify that the results were repeatable. Similar behaviors were observed between the two methods (Figure 5, cyan line). Figures 3 and 4 show the reconstruction created by the VS-LOO at the first detection point and the acceleration point respectively (Figure 5, Table 1). Notably, the difference between Figures 3 and 4 versus Figure 1 show that the reconstruction degraded as a function of fault progression. Figure 4 shows that the reconstruction struggled with high frequency content near the acceleration point.

Figure 4. Example reconstruction at the acceleration point for the VS-LOO algorithm showing significant reconstruction amplitude degradation at low and high frequencies with consistent under-prediction.

5. Fault Progression Trending Curve

The VS-LOO is the primary method for discovery of the fault trending curve. The VS-LOO median output is shown as a function of data file during fault progression in Figure 5 (green, red, and blue data points). The trending curve is computed as a median filter of the reconstruction error per file (black line). The VS-UOC cluster memberships are represented as green (cluster zero), red (cluster 1), and blue (*noise cluster*). Annotations are shown for each element of Table 1 as well as the membership of data used for training and validation early stopping. The grid lines indicate the first file per test day of the File indices.

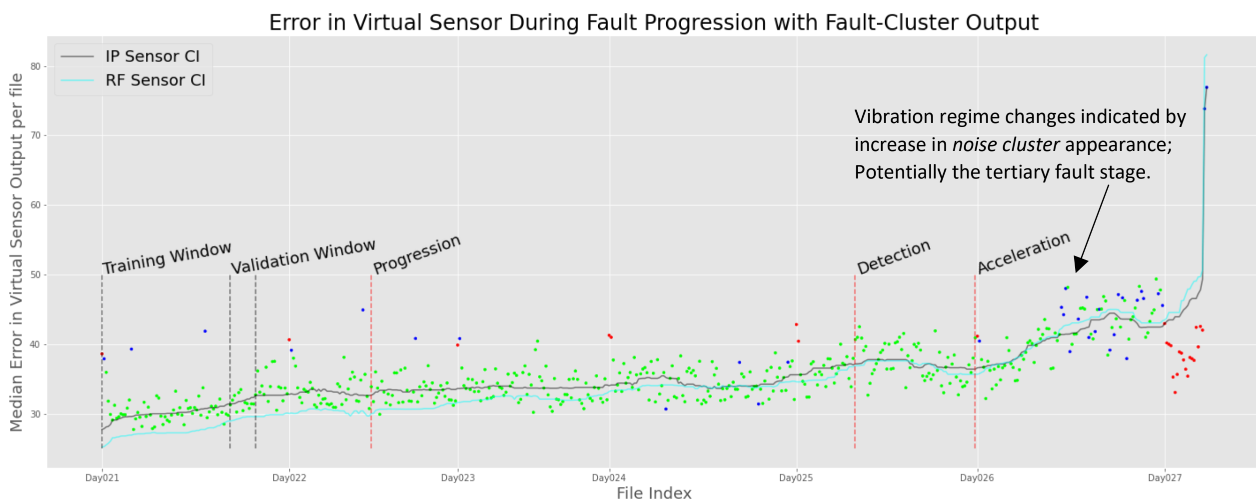

Figure 5. The summary plot of VS-LOO and VS-UOC outputs shown as a function of reconstruction error per data file versus data file index during fault progression. Data points are colored by VS-UOC cluster membership (1 is green, 0 is red, and -1 is blue).

6. Description of Analysis Methods

Description of fault detection method

The best reasonable detection threshold for our VS ensemble techniques to this specific test rig was to determine detection as *consistent exceedance above our threshold*. The threshold was determined by review of the box and whisker plots (3) of data from days 21 and 22 (Figure 6). The upper whisker for these test days is just below 40 units of error and therefore that was chosen as our fault detection threshold. The first consistent (back-to-back) exceedance of the threshold occurs at Table 1, row 1.

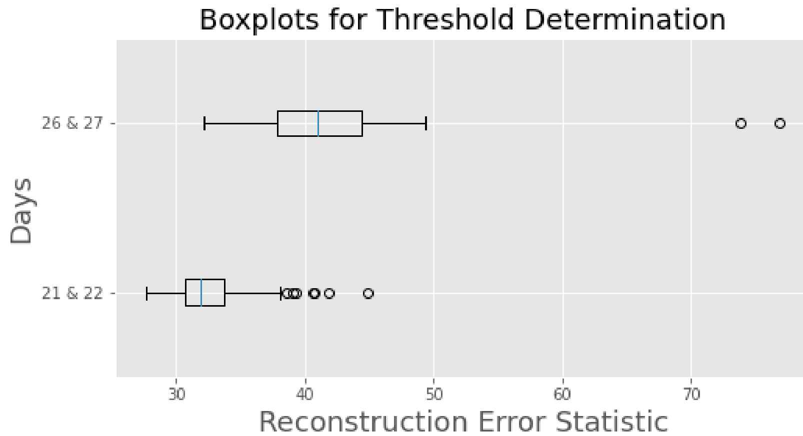

Figure 6. VS-LOO reconstruction error box and whisker plots for Days 21 and 22 versus Days 26 and 27.

The VS-UOC algorithm assists the VS-LOO identify a shift in behavior and confirm that the reconstruction has drifted too far from the original sequences. It also incorporates the frequency domain techniques often used by traditional HUMS analysis.

Description of fault trending method

The fault trending method chosen for our VS is a median filter of the output per data file. The median filter allows the team to show the general trend in the VS-LOO reconstruction error. Median filters are frequently employed in HUMS to provide improved detection criteria and explainability to maintainers.

Conclusions

Care has been taken to review the results without over-fitting, in short though there are clear cluster groups displaying relevance to the problem: data collected at test-rig startup is unique; noise cluster operations may indicate significant fault propagation; fault propagation was steady for the majority of data collections (cluster 1 operations); startup on Day 26 propelled the test rig into a tertiary fault stage; on Day 27, the test rig never achieved cluster 1 vibration operations; a deployed VS-LOO and VS-UOC requires engineering inputs to assemble a true ensemble model.

7. Supplement Information

References

1. Accident Investigation Board Norway. Report SL 2018/04 Summary Report on the Air Accident Near Turoy, Oygarden Municipality, Hordaland County, Norway 29 April 206 with Airbus Helicopters EC 225 LP, LN-OJF, Operated by CHC Helikopter Service. , July 2018
2. Wade, Daniel, Brian Tucker, Mark Davis, Doug Knapp, Sophie Hasbroucq, Moreno Saporiti, Malcom Garrington, and Alex Rudy. Joint Military and Commercial Rotorcraft Mechanical Diagnostics Gap Analysis. Proceedings of the American Helicopter Society 73rd Annual Forum. Fort Worth, TX. May 2017.
3. https://matplotlib.org/stable/api/_as_gen/matplotlib.pyplot.boxplot.html