Ordinal Pattern Analysis for Early Bearing Fault Detection and Classification in Rotating Machinery – First Results

Rasika Dilhani[1], and Alejandro C. Frery[2]

School of Mathematics and Statistics, Victoria University of Wellington, New Zealand

rasidilhani@gmail.com[1] and alejandro.frery@vuw.ac.nz[2]

Introduction

Ordinal Patterns are transformations that encode the sorting characteristics of values in $\mathbb{R}^D$ into $D!$ symbols.

They were proposed by Bandt & Pompe in 2002, and have proven their adequacy at extracting valuable information about the system that produces the data.

One of the possible encodings is the set of indexes that sort the D values in non-decreasing order, for example:

$$(4.2, 5.1, 7.1, 3.9, 8.6) \longmapsto (2, 3, 4, 1, 5) \longmapsto \beth \text{ (one of } 5! = 120 \text{ symbols)}.$$

A time series $\boldsymbol{x} = (x_1, x_2, \ldots, x_{D+n-1})$ can be transformed into a sequence of symbols $\boldsymbol{\pi} = (\pi_1, \pi_2, \ldots, \pi_n)$ where $\pi_i = \pi_i(x_i, x_{i+1}, \ldots, x_{i+D-1})$; D is called the "embedding dimension," and usually ranges between 3 and 6. Then, we compute $\boldsymbol{h} = (\widehat{p}_1, \widehat{p}_2, \ldots, \widehat{p}_{D!})$ the histogram of $\boldsymbol{\pi}$, and obtain two descriptors. The first is the Shannon entropy, a measure of the system's disorder:

$$H[\boldsymbol{h}] = -\frac{1}{\log D!} \sum_{i=1}^{D!} \widehat{p}_i \log \widehat{p}_i \tag{1}$$

with the convention $0 \log 0 = 0$. This quantity is bounded in the unit interval. It is zero when $p_i = 1$ for some i (and, thus, all other bins are zero), and one when $p_i = 1/D!$ for every i (the uniform probability function); it is also called "permutation entropy."

The Permutation Entropy is a powerful descriptor, but the Statistical Complexity (Lopez-Ruiz et al. 1995) enhances the analysis. First, compute the Jensen-Shannon distance between $\widehat{p}$ and $\boldsymbol{u}$, the discrete uniform distribution in $D!$ bins:

$$JS(\widehat{\boldsymbol{p}}, \boldsymbol{u}) = H_S\Big(\frac{\widehat{\boldsymbol{p}} + \boldsymbol{u}}{2}\Big) - \frac{1}{2} H_S(\widehat{\boldsymbol{p}}) - \frac{1}{2}\ln(D!).$$

Then, multiply it by the Permutation Entropy and scale it to 1:

$$C(\widehat{\boldsymbol{p}}, \boldsymbol{u}) = \frac{1}{Q_0'} JS(\widehat{\boldsymbol{p}}, \boldsymbol{u}) H(\widehat{\boldsymbol{p}}), \tag{2}$$

where Q_0' is a normalising constant.

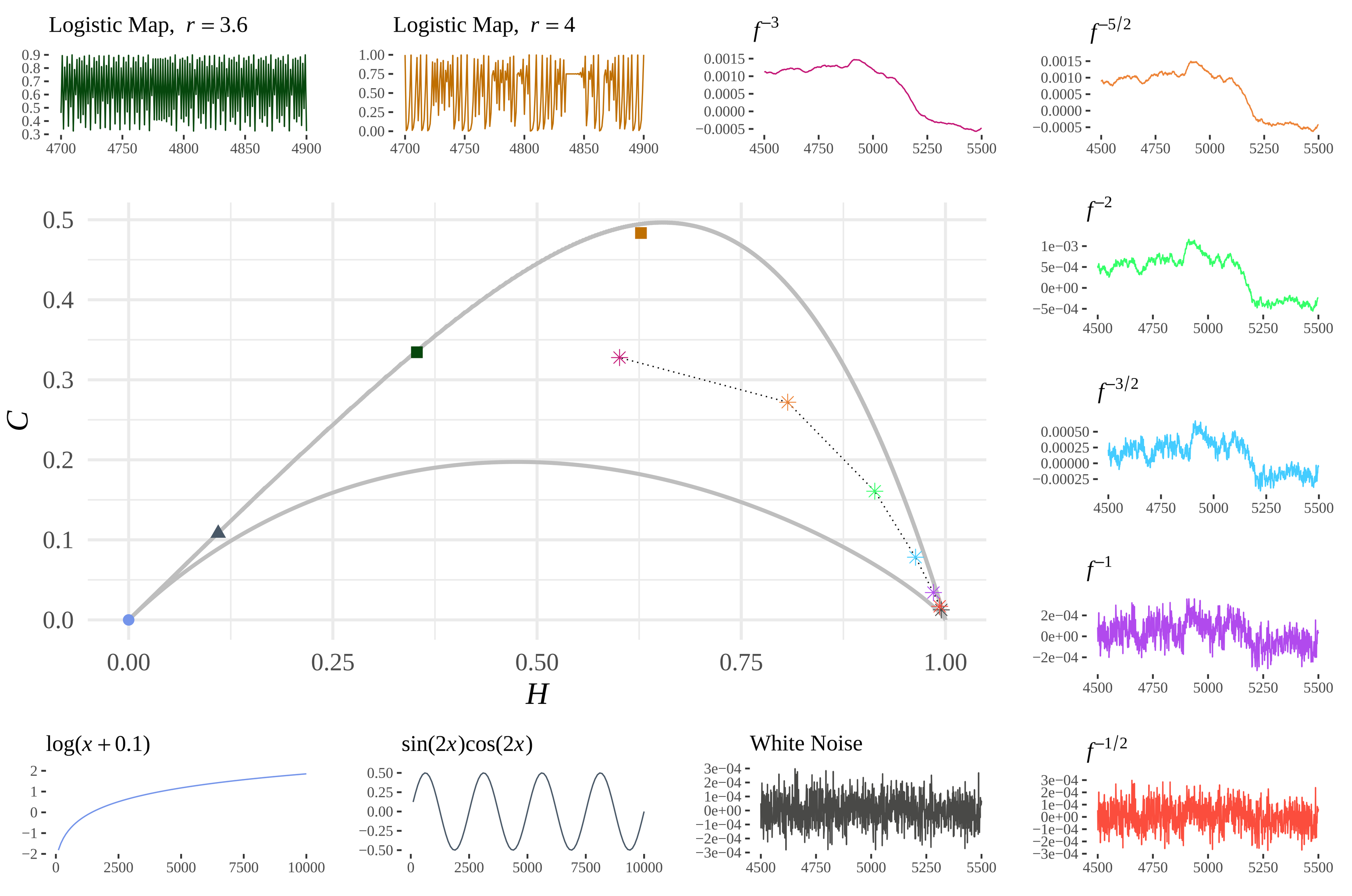

Fig:1 The $H \times C$ plane for $D = 6$ with examples of time series and their points. FROM: Chagas et al. (2022)

Rey et al. (2023), using the Generalised Delta Theorem, obtained the asymptotic distribution of the Shannon entropy, with which we have an approximation for n sufficiently large:

$$H_S(\widehat{\boldsymbol{p}}) = -\sum_{\ell=1}^{k} \widehat{p}_\ell \log \widehat{p}_\ell \overset{\mathcal{D}}{\approx} \mathcal{N}(\widehat{\mu}, \widehat{\sigma}^2), \text{ with } \widehat{\mu} = -\sum_{\ell=1}^{k} \widehat{p}_\ell \log \widehat{p}_\ell, \tag{3a}$$

and

$$\widehat{\sigma}^2 = \frac{1}{n} \sum_{\ell=1}^{k} \widehat{p}_\ell (1 - \widehat{p}_\ell)(\log \widehat{p}_\ell + 1)^2 - \frac{2}{n} \sum_{\jmath=1}^{k-1} \sum_{\ell=\jmath+1}^{k} \widehat{p}_\ell \widehat{p}_\jmath (\log \widehat{p}_\ell + 1)(\log \widehat{p}_\jmath + 1). \tag{3b}$$

Tests statistics follow this result, but it assumes that the patterns are independent.

A central result for our work stems from incorporating the temporal dependence. For $\ell = 1, 2, \ldots, D-1$, define the ordinal pattern transition matrix of order ℓ as $\mathbf{Q}^{(\ell)} \in \mathbb{R}^{D! \times D!}$, whose elements are

$$q_{ij}^{(\ell)} = \Pr\left(\pi_t = \pi^{(i)} \wedge \pi_{t+\ell} = \pi^{(j)}\right) = q_i \Pr(\pi_{t+\ell} = \pi^{(j)} \mid \pi_t = \pi^{(i)}), \tag{4}$$

for $i, j = 1, 2, \ldots, D!$.

Denote $\mathbf{D}_{\boldsymbol{q}} = \operatorname{Diag}(q_1, q_2, \ldots, q_{m!})$ the diagonal matrix obtained by $\boldsymbol{q}$, then

$$\sqrt{n}(\widehat{\boldsymbol{q}}_n - \boldsymbol{q}) \xrightarrow[n\to\infty]{\mathcal{D}} \mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma}), \tag{5}$$

with

$$\boldsymbol{\Sigma} = \mathbf{D}_{\boldsymbol{q}} - (2D-1)\boldsymbol{q}\boldsymbol{q}^{\mathrm{T}} + \sum_{\ell=1}^{D-1} \left(\mathbf{Q}^{(\ell)} + \mathbf{Q}^{(\ell)^{\mathrm{T}}}\right), \tag{6}$$

Direct computations show that, for $i, j = 1, \ldots, D!$

$$\Sigma_{ij} = \begin{cases} q_i - (2D-1)q_i^2 + 2\sum_{\ell=1}^{D-1} q_{ii}^{(\ell)} & \text{if } i = j, \\ -(2D-1)q_i q_j + \sum_{\ell=1}^{D-1} \left(q_{ij}^{(\ell)} + q_{ji}^{(\ell)}\right) & \text{if } i \neq j. \end{cases} \tag{7}$$

Then, we used the Delta Method. Let $h_1, h_2, \ldots, h_{D!}$ be continuously differentiable real functions defined in a neighbourhood of the parameter point $\boldsymbol{q}$. Consider $\boldsymbol{J}$ the matrix of partial derivatives; i.e. $J_{ij} = \partial h_i / \partial q_j$ for $i, j = 1, 2, \ldots, D!$. Assuming that $\boldsymbol{J}$ is non-singular in this neighbourhood, holds that

$$\sqrt{n}\big[h_1(\widehat{\boldsymbol{q}}_n) - h_1(\boldsymbol{q}), h_2(\widehat{\boldsymbol{q}}_n) - h_2(\boldsymbol{q}), \cdots, h_{D!}(\widehat{\boldsymbol{q}}_n) - h_{D!}(\boldsymbol{q})\big] \xrightarrow[n\to\infty]{\mathcal{D}} \mathcal{N}\big(\mathbf{0}, \boldsymbol{J}\boldsymbol{\Sigma}\boldsymbol{J}^{\mathrm{T}}\big). \tag{8}$$

Let $\boldsymbol{\alpha} = (\alpha_1, \alpha_2, \ldots, \alpha_{D!})$ be real numbers. The following convergence for the linear combination of the components holds:

$$\sqrt{n}\left[\sum_{i=1}^{D!} \alpha_i h_i(\widehat{\boldsymbol{q}}_n) - \sum_{i=1}^{D!} \alpha_i h_i(\boldsymbol{q})\right] \xrightarrow[n\to\infty]{\mathcal{D}} \mathcal{N}\left(0, \sum_{i=1}^{D!} \alpha_i^2 \big(\boldsymbol{J}\boldsymbol{\Sigma}\boldsymbol{J}^{\mathrm{T}}\big)_{ii} + 2\sum_{i=1}^{D!-1} \sum_{j=i+1}^{D!} \alpha_i \alpha_j \big(\boldsymbol{J}\boldsymbol{\Sigma}\boldsymbol{J}^{\mathrm{T}}\big)_{ij}\right). \tag{9}$$

Different choices of $\boldsymbol{\alpha}$ yield the Shannon and Tsallis entropies, and the Fisher Information measure. Another choice gives us the ingredient to obtain the Rényi entropy with temporal dependence.

Data

The data were obtained from the Bearing Data Center and the seeded fault test data at the Case Western Reserve University, School of Engineering. The dataset includes ball bearing test data for normal bearings, as well as single-point defects on the drive end and fan end. Data were collected at rates of 12,000 (12k Drive-end:DE) and 48,000 (48k Drive-end) data points per second for the drive-end bearing tests and at 12,000 (12k Fan-end:FE) data points per second for the fan-end bearing tests. Each file includes motor rotational speed, drive-end vibration data, and fan-end vibration data.

The normal baseline data include four motor load levels: 0, 1, 2, and 3, with approximate motor speeds provided in RPM (1797, 1772, 1750, and 1730). The 12k Drive End, 48k Drive End, and 12k Fan End bearing data follow the same motor load levels and speeds. This research aims to identify faulty machines using the time series of their measured features.

Methodology

The goal of this study is to identify malfunctioning machinery. We use ordinal patterns to analyze two time series in each data file. Based on the ordinal structure of the segments, we introduce distance as a metric of similarity. This metric can be used to identify malfunctioning machines with certain embedding dimensions ranging from 3 to 6. The results are analyzed using the Entropy-Complexity plane. Time series data were first examined, and the StatOrdPatt package was then used to analyze the data using the methodology outlined by Rey et al. (2024).

Results

The results of embedding dimension 4 with 12k Drive and 12k Fan end time series data are presented in this article. White noise is found close to the upper and lower limits. With the graph below, it is evident.Figure 2 shows the DE, and FE time series for engine loads. The graph is clear and it can be seen that it is failure time series data is about time series. However, it is difficult to separate the exact machine type and other data characteristics. The main idea of this research work is to identify perturbation machines using the various time series data given above.

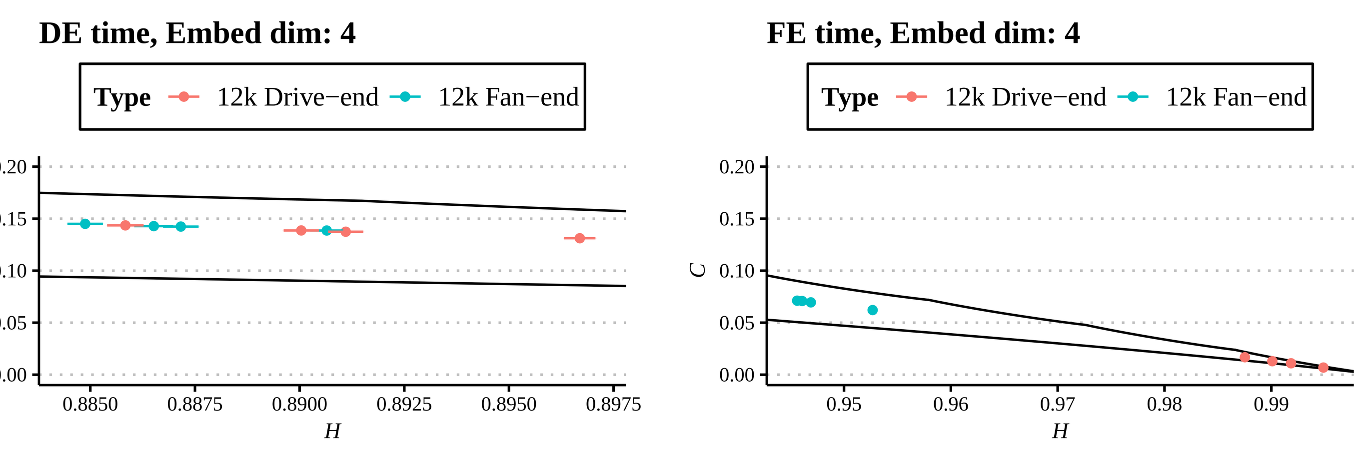

Fig:2 Entropies with their confidence intervals in the $H \times C$ plane.

Conclusion and Future Work

Permutation entropy with $D = 4$ guarantees the separation of faulty machines. The machines form clusters while maintaining individual signatures.

Confidence intervals quantify the uncertainty associated with the estimates, adding a layer of statistical rigour to multivariate analysis. Our findings establish a strong foundation for future researchers to adopt and refine similar frameworks, setting a new benchmark for multivariate time series studies.

References

Chagas, E. T., Queiroz-Oliveira, M., Rosso, O. A., Ramos, H. S., Freitas, C. G. & Frery, A. C. (2022), 'White noise test from ordinal patterns in the entropy–complexity plane', *International Statistical Review* **90**(2), 374–396.

Lopez-Ruiz, R., Mancini, H. L. & Calbet, X. (1995), 'A statistical measure of complexity', *Physics letters A* **209**(5-6), 321–326.

Rey, A. A., Frery, A. C., Lucini, M., Gambini, J., Chagas, E. T. C. & Ramos, H. S. (2023), 'Asymptotic distribution of certain types of entropy under the multinomial law', *Entropy* **25**(5), 734.

Rey, A., Frery, A. C., Gambini, J. & Lucini, M. (2024), 'Asymptotic distribution of entropies and fisher information measure of ordinal patterns with applications', *Chaos, Solitons & Fractals* **188**, 115481. **URL:** *https://www.sciencedirect.com/science/article/pii/S0960077924010336*