**Feature-Based Clustering of Simulated Time Series Using Ordinal Pattern Analysis**

Ordinal pattern analysis has emerged as a promising alternative to analysing time series data, providing a robust and computationally efficient approach. Ordinal pattern analysis is used in this research to analyse simulated time series under Autoregressive (AR), Moving Average (MA), and Autoregressive Moving Average (ARMA) models. Ordinal pattern analysis involves converting time series data into a sequence of symbols that represent ordering relationships among data points within specific time windows. Then, the Shannon entropy and the Statistical Complexity are computed from the histogram of symbols. The method detects small differences between models and sample sizes.

This study presents a simulation-based experiment into time series clustering using ordinal pattern analysis. The primary research focus was on identifying the distinguishing characteristics of AR, MA, and ARMA models through ordinal patterns features. Synthetic time series were generated with the time window called as an embedding dimension of 3 and varying lengths (n=500 and n=1000), ensuring models satisfy stringent parameter constraints for stationarity and invertibility. ~~The experimental design incorporates four models structural variations (M1–M4), which depend on the parameter variation within each AR, MA, and ARMA type of the model, with coefficients systematically varied as specified in the experimental table (Attached in Annex 1).~~  The experimental design considers several variants for each AR, MA, and ARMA model type, labelled in the form Model\_Type\_M1 to Model\_Type\_M4, with each model representing a distinct parameter setting (see Annex 1). For robustness, 100 independent replications were performed for each model configuration.

~~Ordinal patterns were extracted for each time series to compute permutation entropy and statistical complexity. Clustering structures were then analyzed in the entropy–complexity plane, serving as the main feature space for model discrimination. Distinct clusters corresponding to the type of the model classes (AR, MA, and ARMA) were observed, with clear separation the highest coefficients with -0.8 and 0.8, particularly evident for series of length 1000 (Figures 1 and 2). For small sample series (n = 500), observed patterns were more dispersed, indicating greater variability in entropy–complexity estimates for smaller samples. As the series length increased, cluster centers became more stable and intra-group scatter decreased, demonstrating the stabilizing effect of larger sample sizes on ordinal-based features.~~

Ordinal patterns were extracted to compute entropy and complexity. Features were analyzed in the entropy–complexity plane for model discrimination. ARMA(1,1) shows two distinct clusters. AR coefficients -0.8 and 0.8 produce lower entropy and higher complexity. Coefficients 0.1 and -0.1 yield higher entropy and lower complexity. ARMA(2,2) shows overlapping groups across all four models. Clear separation appears by sample size, with larger samples forming more stable clusters. (Figures 1 and 2).

The findings highlight the effectiveness of feature-based time series clustering, leveraging permutation entropy and complexity to capture intrinsic model differences. The simulation framework further validates the method’s reliability under different structural and stochastic configurations. This research provides a foundation for applying ordinal pattern analysis in unsupervised time series grouping, with potential applications in diverse data-rich domains such as finance, engineering, and bioinformatics.

**Entropy–Complexity Based Clustering of ARMA Simulated Time Series Models**

Ordinal pattern analysis is used to cluster time series in this research work. This method is fast and handles noise well. It changes each time series into symbols that show the order of data points in a small window. This helps capture the key patterns of each model. The technique gives clear features and shows small differences between series with different structures, embedding dimensions, and sample sizes. ~~Ordinal pattern analysis helps group time series data quickly and reliably. It works well with noisy signals. The method turns data points into symbols based on their order in a set window. This reveals important signal patterns. It shows small differences between various model types and sample sizes.~~ Ordinal pattern analysis helps to group time series data quickly and reliably, and it is particularly effective when working with noisy signals. The method converts data points into symbols based on their relative order within a given window, which reveals important patterns in the signal. In doing so, it highlights subtle differences between model types and across different sample sizes.

In this study, a simulation-driven study was conducted to evaluate clustering behaviours of synthetic time series generated from a range of Autoregressive (AR), Moving Average (MA), and Autoregressive Moving Average (ARMA) models. Each model was parameterized rigorously to meet stability and invertibility criteria, and the experimental design included series lengths of n=500 and n=1000 as well as four structural variations per model class. A total of 100 replications were performed for each configuration to ensure statistical reliability.

The core analytical procedure consisted of extracting ordinal patterns from each series, followed by the computation of permutation entropy and statistical complexity. These features were then used to characterize clustering in the entropy–complexity plane. The analysis revealed distinct groupings corresponding to the underlying model class, particularly when the time series length was 1000. For shorter series (n = 500), the clusters appeared less distinct and more dispersed, indicating increased variability of entropy–complexity estimates at smaller sample sizes. Increasing the series length led to more stable and compact clusters, supporting the reliability of ordinal pattern features for clustering.

Overall, this work demonstrates the utility of ordinal features for unsupervised time series grouping. The experimental findings point to broad applicability in fields where data are heterogeneous and prone to noise, such as finance, engineering, and biomedical monitoring. The simulation framework further validates ordinal pattern-based clustering as a foundational methodology for future research in time series discrimination.

Annex 1

Table for parameter setting of Research Work

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  |  | **Coefficients** | | | |
| **Type of model** | **Model** | **ar1** | **ar2** | **ma1** | **ma2** |
| AR(1) | AR1\_M1 | 0.8 |  |  |  |
| AR1\_M2 | 0.1 |  |  |  |
| AR1\_M3 | -0.8 |  |  |  |
| AR1\_M4 | -0.1 |  |  |  |
| AR(2) | AR2\_M1 | 0.1 | 0.8 |  |  |
| AR2\_M2 | -0.8 | 0.1 |  |  |
| AR2\_M3 | 0.1 | -0.8 |  |  |
| AR2\_M4 | -0.8 | -0.1 |  |  |
| MA(1) | MA1\_M1 |  |  | 0.8 |  |
| MA1\_M2 |  |  | 0.1 |  |
| MA1\_M3 |  |  | -0.8 |  |
| MA1\_M4 |  |  | -0.1 |  |
| MA(2) | MA2\_M1 |  |  | 0.1 | 0.8 |
| MA2\_M2 |  |  | -0.8 | 0.1 |
| MA2\_M3 |  |  | 0.1 | -0.8 |
| MA2\_M4 |  |  | -0.8 | -0.1 |
| ARMA(1,1) | ARMA11\_M1 | 0.8 |  | 0.8 |  |
| ARMA11\_M2 | 0.1 |  | 0.1 |  |
| ARMA11\_M3 | -0.8 |  | -0.8 |  |
| ARMA11\_M4 | -0.1 |  | -0.1 |  |
| ARMA(2,2) | ARMA22\_M1 | 0.1 | 0.8 | 0.1 | 0.8 |
| ARMA22\_M2 | -0.8 | 0.1 | -0.8 | 0.1 |
| ARMA22\_M3 | 0.1 | -0.8 | 0.1 | -0.8 |
| ARMA22\_M4 | -0.8 | -0.1 | -0.8 | -0.1 |

Results

