Assignment 3 (ML for TS) - MVA

Richard Goudelin rgoudelin@gmail.com
Thomas Gravier thomas.gravier@gadz.org

December 30, 2024

1 Introduction

Objective. The goal is to implement (i) a signal processing pipeline with a change-point detection method and (ii) wavelets for graph signals.

Warning and advice.

- Use code from the tutorials as well as from other sources. Do not code yourself well-known procedures (e.g. cross validation or k-means), use an existing implementation.
- The associated notebook contains some hints and several helper functions.
- Be concise. Answers are not expected to be longer than a few sentences (omitting calculations).

Instructions.

- Fill in your names and emails at the top of the document.
- Hand in one report per pair of students.
- Rename your report and notebook as follows:
`FirstnameLastname1_FirstnameLastname1.pdf` and
`FirstnameLastname2_FirstnameLastname2.ipynb`.
For instance, `LaurentOudre_CharlesTruong.pdf`.
- Upload your report (PDF file) and notebook (IPYNB file) using the link given in the email.

2 Dual-tone multi-frequency signaling (DTMF)

Dual-tone multi-frequency signaling is a procedure to encode symbols using an audio signal. The possible symbols are 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, *, #, A, B, C, and D. A symbol is represented by a sum of cosine waves: for $t = 0, 1, \ldots, T-1$,

$$y_t = \cos(2\pi f_1 t / f_s) + \cos(2\pi f_2 t / f_s)$$

where each combination of (f_1, f_2) represents a symbols. The first frequency has four different levels (low frequencies), and the second frequency has four other levels (high frequencies); there are 16 possible combinations. In the notebook, you can find an example symbol sequence encoded with sound and corrupted by noise (white noise and a distorted sound).

Question 1

Design a procedure that takes a sound signal as input and outputs the sequence of symbols. To that end, you can use the provided training set. The signals have a varying number of symbols with a varying duration. There is a brief silence between each symbol.

Describe in 5 to 10 lines your methodology and the calibration procedure (give the hyperparameter values). Hint: use the time-frequency representation of the signals, apply a change-point detection algorithm to find the starts and ends of the symbols and silences, and then classify each segment.

Answer 1

To decode the DTMF symbols from the input sound signal, the following procedure is employed:

1. **Time-Frequency Representation**: A spectrogram is computed to obtain the time-frequency representation of the filtered signal, focusing on the frequency range pertinent to DTMF (600 Hz to 2000 Hz).

2. **Frequency Selection**: For each frequency within the selected range in the spectrogram, autocorrelation is calculated to identify stable frequency components. The top frequencies with the highest autocorrelation values are selected as candidates representing the DTMF $f1$ or $f2$.

3. **Preprocessing**: The signal is filtered to remove white noise using an FFT-based thresholding method, retaining only significant frequency components.

4. **Change-Point Detection**: Breakpoints between symbols and silences are detected by identifying intervals where the power of the selected frequencies exceeds a predefined threshold. This segmentation isolates individual symbols within the signal. We tried to use the ruptures package but it wasn't working as well as the threshold method.

5. **Symbol Classification**: Each segmented interval is analyzed to determine the active frequency pairs. These pairs are then mapped to their corresponding symbols using a frequency-to-symbol mapping dictionary that has been manually created as it is really simple to annotated the corresponding frequencies using the function for frequency selection.

6.Calibration:

- **Spectrogram Parameters**: We set the number of samples per segment (n_{perseg}) to 512. Given that the minimum sound duration is 0.03 seconds, it give a window length of 661 with our sampling rate. After tweaking, 512 window size provides a good balance between time and frequency resolution. The overlap (noverlap) was set to 256 samples, which produced one of the cleanest spectrograms during visual inspection.
- **Autocorrelation Threshold**: By plotting the autocorrelation values for several samples, we determined that a threshold of $v = 0.0001$ effectively separates frequencies with a clear fundamental component from those that only contain partial sinusoidal/noise information within the spectrogram.
- **Power Threshold**: Analysis of the energy levels revealed that when a DTMF note is played, its energy value exceeds -22 dB (in 10log scale). Therefore, we set the power threshold ($\text{threshold}_{\text{db}}$) to -22 dB to accurately detect active symbols.
- **Frequency Selection**: The maximum number of frequencies to select was set to 8, corresponding to the 4 low frequencies (f_1) and 4 high frequencies (f_2) used in DTMF signaling.

Question 2

What are the two symbolic sequences encoded in the test set?

Answer 2

Using our method, we obtained the following symbol sequences (both are correct):

- **Sequence 1**: `['7', '2', '1', 'C', '9', '9']`
- **Sequence 2**: `['1', '#', '2', '#']`

Here is an example of how the method performed on the second test case:

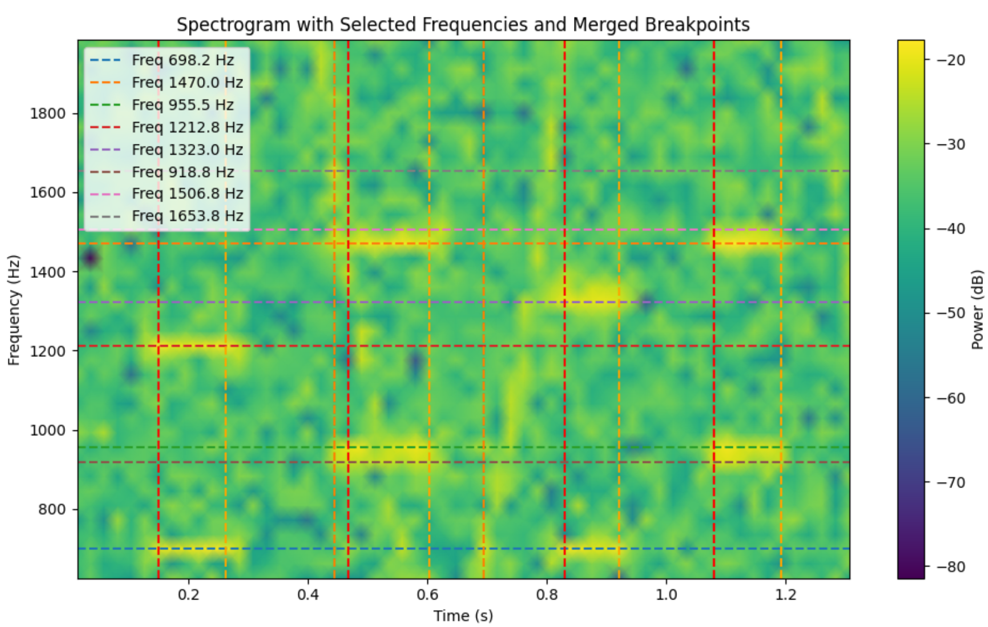


Figure 1: Segmented Part of the Second Example

3 Wavelet transform for graph signals

Let G be a graph defined a set of n nodes V and a set of edges E. A specific node is denoted by v and a specific edge, by e. The eigenvalues and eigenvectors of the graph Laplacian L are $\lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_n$ and $u_1, u_2, \ldots, u_n$ respectively.

For a signal $f \in \mathbb{R}^n$, the Graph Wavelet Transform (GWT) of f is $W_f : \{1, \ldots, M\} \times V \longrightarrow \mathbb{R}$:

$$W_f(m, v) := \sum_{l=1}^{n} \hat{g}_m(\lambda_l) \hat{f}_l u_l(v) \tag{1}$$

where $\hat{f} = [\hat{f}_1, \ldots, \hat{f}_n]$ is the Fourier transform of f and $\hat{g}_m$ are M kernel functions. The number M of scales is a user-defined parameter and is set to $M := 9$ in the following. Several designs are available for the $\hat{g}_m$; here, we use the Spectrum Adapted Graph Wavelets (SAGW). Formally, each kernel $\hat{g}_m$ is such that

$$\hat{g}_m(\lambda) := \hat{g}^U(\lambda - am) \quad (0 \leq \lambda \leq \lambda_n) \tag{2}$$

where $a := \lambda_n / (M + 1 - R)$,

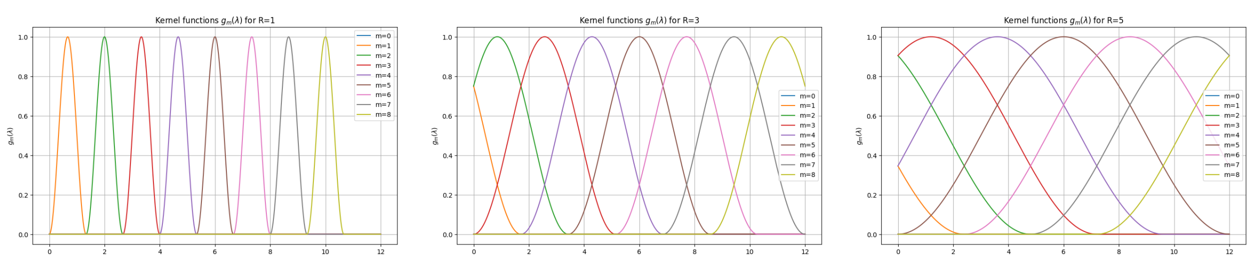
$$\hat{g}^U(\lambda) := \frac{1}{2}\left[1 + \cos\left(2\pi\left(\frac{\lambda}{aR} + \frac{1}{2}\right)\right)\right] \mathbb{1}(-Ra \leq \lambda < 0) \tag{3}$$

and $R > 0$ is defined by the user.

Question 3

Plot the kernel functions $\hat{g}_m$ for $R = 1$, $R = 3$ and $R = 5$ (take $\lambda_n = 12$) on Figure 2. What is the influence of R?

Answer 3



(a) $R = 1$ (b) $R = 3$ (c) $R = 5$

Figure 2: The SAGW kernel functions for different values of R.

L'influence de R est de réguler la largeur des noyaux $g_m(\lambda)$. Un R élevé produit des noyaux étroits, offrant une meilleure séparation spectrale et capturant des détails locaux, tandis qu'un R faible génère des noyaux larges, favorisant une analyse plus globale avec un chevauchement spectral accru.

We will study the Molene data set (the one we used in the last tutorial). The signal is the temperature.

Question 4

Construct the graph using the distance matrix and exponential smoothing (use the median heuristics for the bandwidth parameter).

- Remove all stations with missing values in the temperature.
- Choose the minimum threshold so that the network is connected and the average degree is at least 3.
- What is the time where the signal is the least smooth?
- What is the time where the signal is the smoothest?

Answer 4

The stations with missing values are: BATZ, BEG MEIL, CAMARET, PLOUGONVELIN, RIEC SUR BELON, ST NAZAIRE-MONTOIR, PLOUAY-SA, VANNES-MEUCON, LANNAERO, PLOUDALMEZEAU, LANDIVISIAU, SIZUN, QUIMPER, OUESSANT-STIFF, LANVEOC, ARZAL, BREST-GUIPAVAS, and BRIGNOGAN.

The threshold is equal to 0.83 .

The signal is the least smooth at 2014-01-21 06:00:00.

The signal is the smoothest at 2014-01-24 19:00:00.

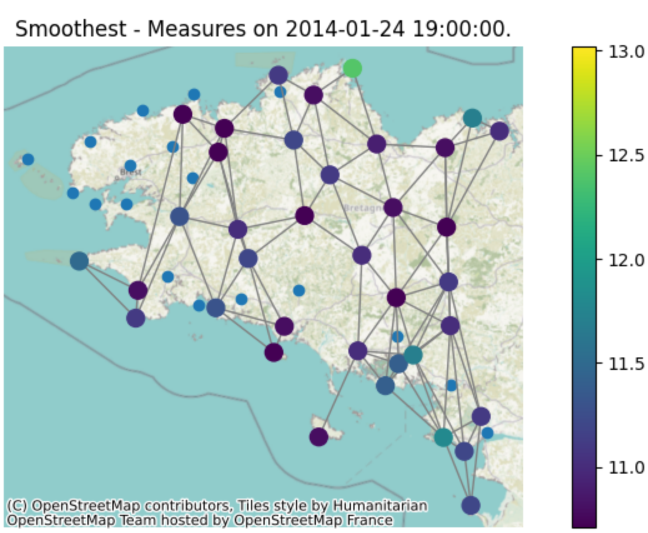


Figure 3: Smoothest Signal

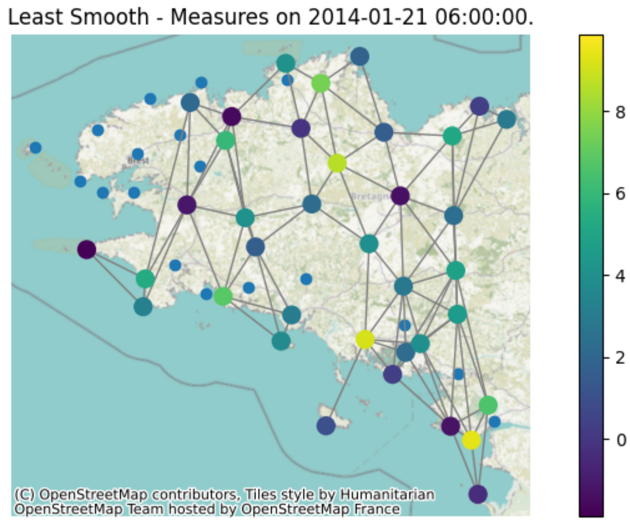


Figure 4: Least smooth Signal

Question 5

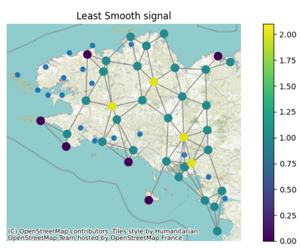
(For the remainder, set $R = 3$ for all wavelet transforms.)

For each node v, the vector $[W_f(1, v), W_f(2, v), \ldots, W_f(M, v)]$ can be used as a vector of features. We can for instance classify nodes into low/medium/high frequency:

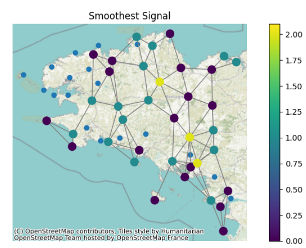
- a node is considered low frequency if the scales $m \in \{1, 2, 3\}$ contain most of the energy,
- a node is considered medium frequency if the scales $m \in \{4, 5, 6\}$ contain most of the energy,
- a node is considered high frequency if the scales $m \in \{6, 7, 9\}$ contain most of the energy.

For both signals from the previous question (smoothest and least smooth) as well as the first available timestamp, apply this procedure and display on the map the result (one colour per class).

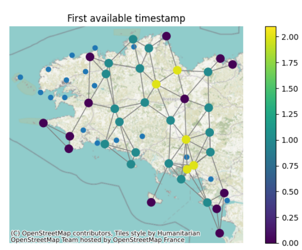
Answer 5



(a) Least smooth signal



(b) Smoothest signal



(c) First available timestamp

Figure 5: Classification of nodes into low/medium/high frequency

Question 6

Display the average temperature and for each timestamp, adapt the marker colour to the majority class present in the graph (see notebook for more details).

Answer 6

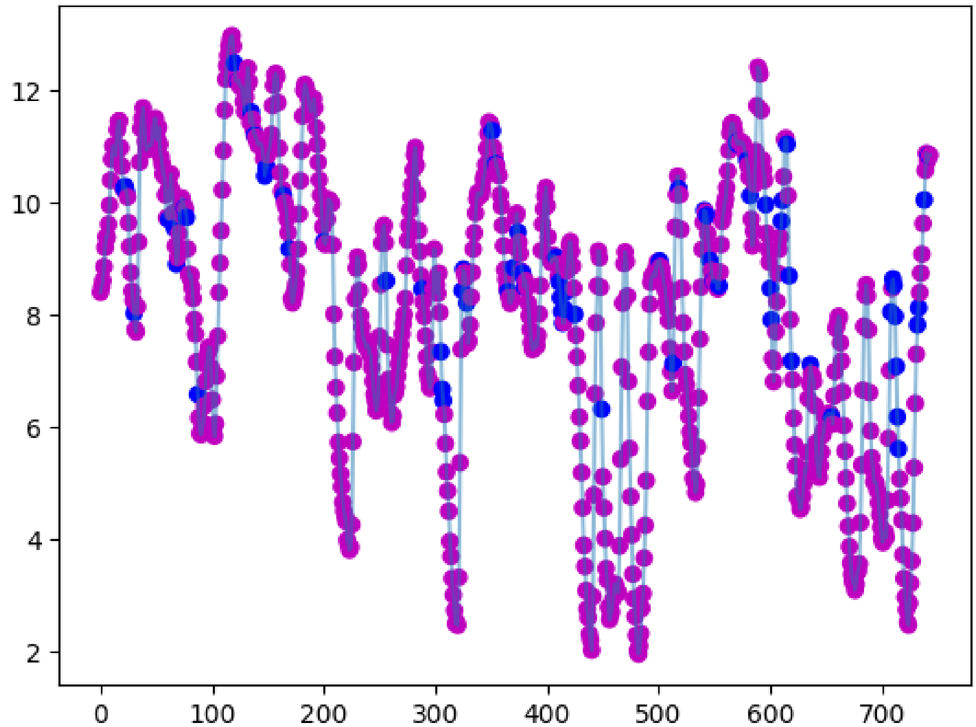


Figure 6: Average temperature. Markers' colours depend on the majority class.

Question 7

The previous graph G only uses spatial information. To take into account the temporal dynamic, we construct a larger graph H as follows: a node is now *a station at a particular time* and is connected to neighbouring stations (with respect to G) and to itself at the previous timestamp and the following timestamp. Notice that the new spatio-temporal graph H is the Cartesian product of the spatial graph G and the temporal graph G' (which is simply a line graph, without loop).

- Express the Laplacian of H using the Laplacian of G and G' (use Kronecker products).
- Express the eigenvalues and eigenvectors of the Laplacian of H using the eigenvalues and eigenvectors of the Laplacian of G and G'.
- Compute the wavelet transform of the temperature signal.
- Classify nodes into low/medium/high frequency and display the same figure as in the previous question.

Answer 7

To express the Laplacian of the spatio-temporal graph H (L_H) in terms of the Laplacians of the spatial graph G (L_G) and the temporal graph G' ($L_{G'}$), we use the Cartesian product of G and G'. The Laplacian of H is given by:

$$L_H = L_G \otimes I_m + I_n \otimes L_{G'}$$

where I_n and I_m are identity matrices of dimensions corresponding to the nodes in G and G', respectively.

From this, it follows that the eigenvalues of H are obtained as the sums of the eigenvalues of G (λ_i) and G' (μ_j):

$$\text{Eigenvalues of } H: \quad \{\lambda_i + \mu_j \mid i = 1, \ldots, n,\ j = 1, \ldots, m\}$$

The eigenvectors of H are formed by the Kronecker products of the eigenvectors of G (v_i) and G' (u_j):

$$\text{Eigenvectors of } H: \quad \{v_i \otimes u_j \mid i = 1, \ldots, n,\ j = 1, \ldots, m\}.$$

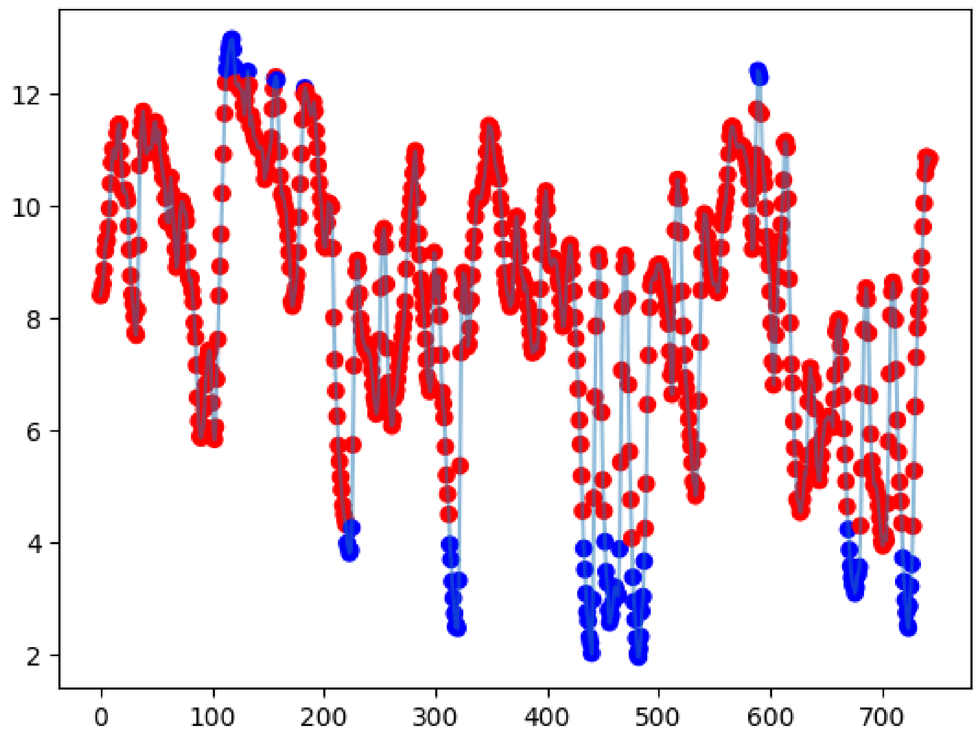


Figure 7: Average temperature. Markers' colours depend on the majority class.