Modeling effect of time delay for large network of seismic monitor

Jordi Anguera* Leevi Annala† Stefan Dimitrijevic‡
Patricia Pauli§ Liisa-Ida Sorsa¶ Dimitar Trendafilov‖
Christophe Picard**

17.7.2018

Abstract

Seismic monitoring is used to study the behavior and composition of the underground floor. For earthquake prediction and underground works precise timing and positioning information is needed. Drilling companies use equipments that are linked in a network and are generally connected to a global positioning system for synchronization. However, instruments are not continuously synchronized and may deviate in time. Hence, the periods to which the vibration of the underground floor are caught might lack accuracy. Consequently, the precise localisation of the events becomes impossible. We have delay measurements and distance data available for an example seismic network and use it to correct the timing of the data in periods without GPS reception.

1 Introduction

In operation of seismic networks high quality of data is required for accurate prediction of seismic events. Precise timing is crucial but continuous GPS

*Autonomous University of Barcelona, Spain
†University of Jyväskylä, Finland
‡University of Novi Sad, Serbia
§Technical University of Darmstadt, Germany
¶Tampere University of Technology, Finland
‖University of Sofia "St. Kliment Ohridski", Bulgaria
**University of Grenoble Alpes and Grenoble INP, France

synchronization of the stations' internal clocks is not possible due to high energy consumption.

The purpose of this work is to analyze noise in a large network of seismic monitors and to extract clock drift from the delay. This work suggests an alternative way to determine clock drift based on the data only.

2 Problem Statement

2.1 Assumptions

All the stations have the same orientation. The medium the wave is propagating in is homogenous.

2.2 Definitions

Δ_i	clock drift at station i
δ_{ij}	Measured delay between station i and station j
$\hat{\delta}_{ij}$	Actual delay between station i and station j
ε_{ij}	Noise between two stations i and j, including clock drift and ambient noise
r_{ij}	Distance

2.3 Modelling time delay between seismic monitors

The network of monitors can be modelled as a complete graph. The signal of each monitor is cross-correlated with other monitor signals, finally yielding time-delay data $\hat{\delta}_{ij}$ between each monitor. The time delay between the monitors comprise information on the real time the signal travels between the monitors δ_{ij}, the inaccuracies of the clock (time drift), and other noise affecting the measurement. The position of each seismic monitor in the network is available that can be used to compute the distances d_{ij} between all stations.

3 Data

The data is collected from a network of 73 seismic monitor stations recording ground vibrations in the Southeast region of France. Their average distance is 166km. The GPS coordinate position of each station in the network is known (Figure 1) and they each contain a clock which is synchronized to the GPS once a month. The seismic monitor sensor records compression and decompression as discrete values -1 and 1, respectively, every second. The sensor responds to any event in the area, be it an earthquake, tremor from road traffic, airplane or any other pressure wave which travels in the ground.

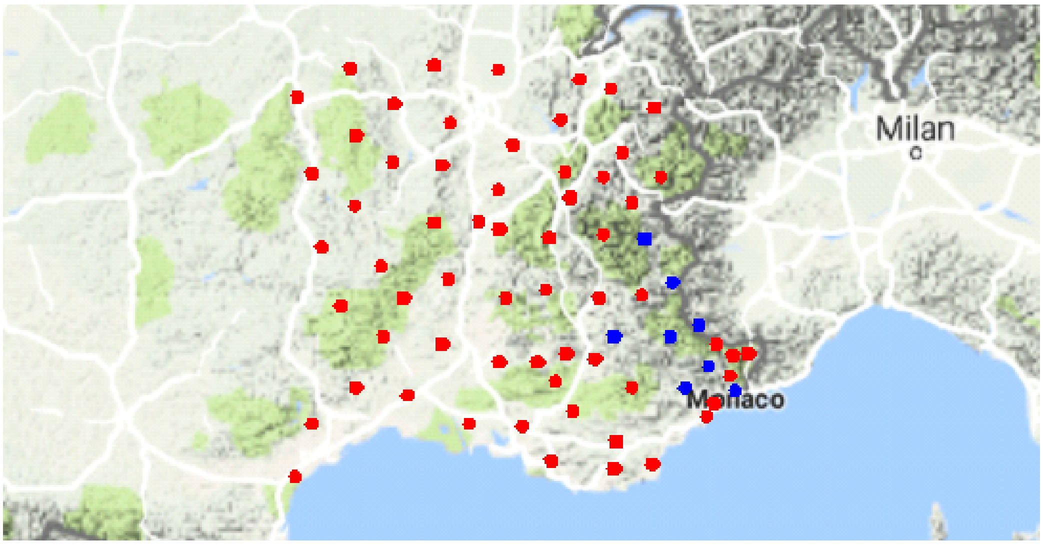


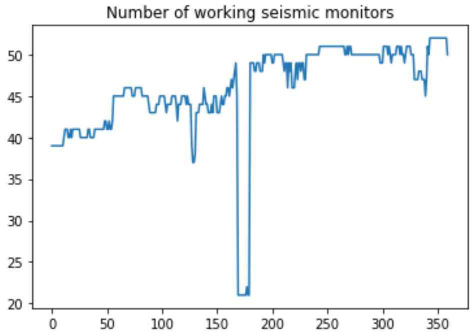
Figure 1: The network of seismic monitors are located in the Southeast France, mostly in the regions of Provence-Alpes-Côte d'Azur and Rhône-Alpes. The eight stations chosen as the small test set is shown as green in the southeast part of the area.

The stations work independent of each other and are occasionally shut down for some time frame for maintenance, repair or just random events. Just as occasionally they are brought up to measure again. Therefore, the number of active stations varies over time. The Figure 2(a) shows the number of working stations, and the Figure 2(b) the number of working connections between the monitor stations over one year of measurements.

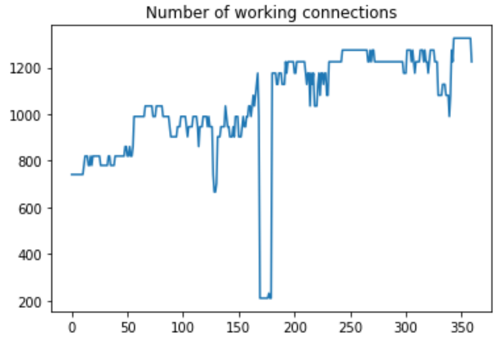
The compression and decompression data recorded by the stations is retrieved and run through initial data cleaning and filtering procedures. The data is then cross-correlated in one-hour time windows to yield time-delay data of signals between all monitors (Figure 3).

The time delays between each monitor are assumed to depend on the distance between the monitor. The further away the monitors are from each other, the larger the time delay caused by the same event recorded by the monitor. Figure 4 shows the time signal of the station pairs that are closest and furthest apart over a time frame of 10 days. One can see that the variance of the delays is much higher for the stations that are far apart.

The Figure 5 shows an example of time-delay evolution of all monitors



(a) The number of working seismic monitor stations.



(b) The number of working connections between the stations.

Figure 2: Working stations and connections over one year time period.

over 24 hours. The distributions are symmetric with zero mean which implies that the delays at an instant cancel out over the whole area the 73 stations are mounted in. The interesting feature of the histograms is how the shape of the distribution changes slowly over time. However, the mean remains zero.

3.1 The most influential stations

Jordi to describe how he found the most influential stations (SVD) and figures

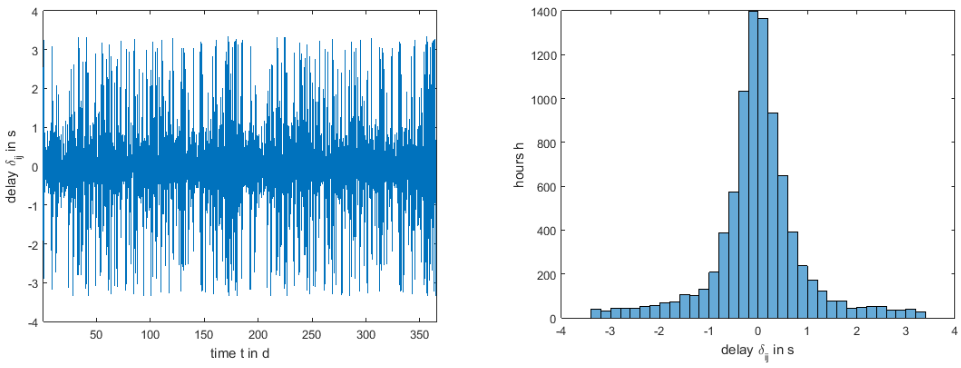
4 Methods

Let $\hat{\boldsymbol{\delta}} \in \mathbb{R}^{n\times n}$ be the measured pairwise time delays of the system, $\boldsymbol{\delta}$ the actual pairwise time delays, and $\boldsymbol{\varepsilon}$ an error term including clock drift of the station's clock and other errors

$$\hat{\boldsymbol{\delta}} = \boldsymbol{\delta} + \boldsymbol{\varepsilon}. \tag{1}$$

Each of the monitors are equipped with a clock which runs independent of others. It is synchronized via GPS system once a month and then runs independently. The drift, ε_i in clock i is caused by variation in the clocks oscillator.

The position of each seismic monitor in the network is known. As the monitors are spread across a large area, it is assumed that local tremors are detected by stations that are close by, and therefore the correlations found in the pairwise cross-correlations and time delay data between them have higher likelihood to be linked.



(a) Time Signal. (b) Histogram.

Figure 3: The time-delay signal recorded between one station pair over a year. The time delay is computed in one-hour windows, yielding 24 time-delay values over a day.

4.1 Denoising model

4.1.1 Definitions

The computational denoising model involves weighted network estimation by the use of topological graph metrics, described in detail in [2].

We have a weighted graph $\mathcal{G} = (\mathcal{V}, \mathcal{E}, \hat{\delta}_{\mathbf{t}})$ defined by a finite set of nodes $\mathcal{V}$ with $|\mathcal{V}| = n$, a set of edges $\mathcal{E} = \{(v_i, v_j) \in \mathcal{E}\}$, with $|\mathcal{E}| = n^2 - n$ and the weighted adjacency matrix $\hat{\boldsymbol{\delta}}$ with $\hat{\delta}_{ii} = 0$ for all i. The matrix $\hat{\boldsymbol{\delta}}$ is symmetric and describes the time delays between events in the graph, and is normalized, i.e. $\hat{\delta}_{ij} \in [0, 1]$. The weighted adjacency matrix indicates the strength of connection between nodes.

4.1.2 Graph metrics

Graph metrics are scalar functions of the weight matrix, $f_i(\hat{\boldsymbol{\delta}}) = K_i : \mathbb{R}^{n \times n} \to \mathbb{R}$. They quantify a property of a network. We deisigned a metric, the connection strength of a node i that is the row sum of the normalized weighted adjency matrix

$$f_i(\hat{\boldsymbol{\delta}}) = \sum_j \hat{\delta}_{ij}. \tag{2}$$

The connection strength can be interpreted as the sum of all normalized delays of all of the links to node i.

Figure 6 shows two nodes. Node A has four connections yielding a connection strength $0.1 + 0.5 + 0.7 + 0.6 = 1.9$ whereas node B only has three

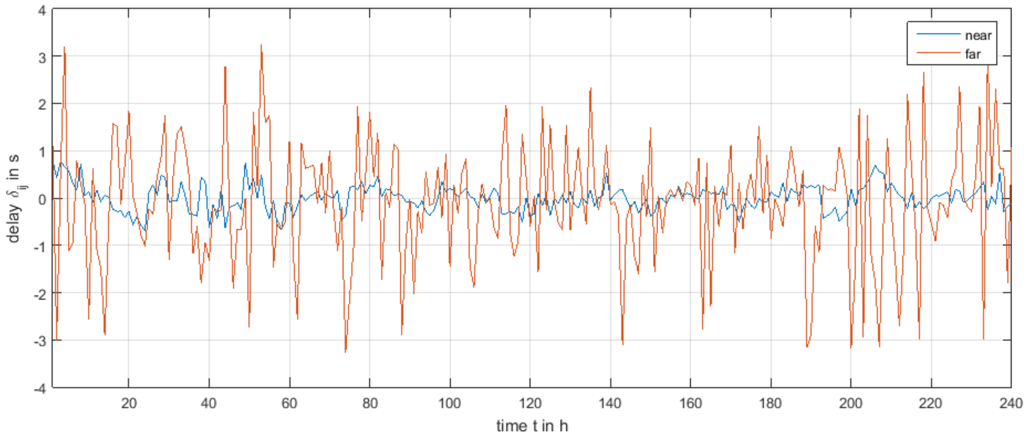


Figure 4: Comparison of the delay of the furthest and closest station pairs over 10 days.

connections. Thus, the connection strength is $0.3 + 0.6 + 0.1 = 1$. Next, we look at the K_i. Since we have the distance measurements available and expect a correlation between the delays and distances, we want to use this information to denoise the delay measurements and define

$$K_i = f_i(\boldsymbol{W}_d). \tag{3}$$

It is the connection strength as well. But here we plug in the distance matrix $\boldsymbol{W_d}$. The distances are constants and yield constants K_i. The weight matrix $\boldsymbol{W_d}$ is a function of the distance between node i and node j.

1. First, we choose the function to be the reciprocal of the distances of the node to all other nodes, $w_{ij} = 1/r_{ij}$, normalized between $[0, 1]$. Higher values correspond to nodes with short distances to all other nodes.

2. Second, we try a metric with weights proportional to the distances ($w_{ij} \propto r_{ij}$), normailzed between $[0, 1]$ accordingly. Here, high values correspond to long distances between two stations.

Other possible metrics are average neighbor degrees (resilience), transitivity, or a clustering coefficient. The analysis based on these metrics is beyond scope of this work.

4.1.3 Cost function

We assume to have estimates of M differentiable graph metrics, one for each node i, of the original matrix $\hat{\boldsymbol{\delta}}$, i.e. $f_i(\boldsymbol{\delta}) = K_i$, where $i \in \{1, \ldots, M\}$, then

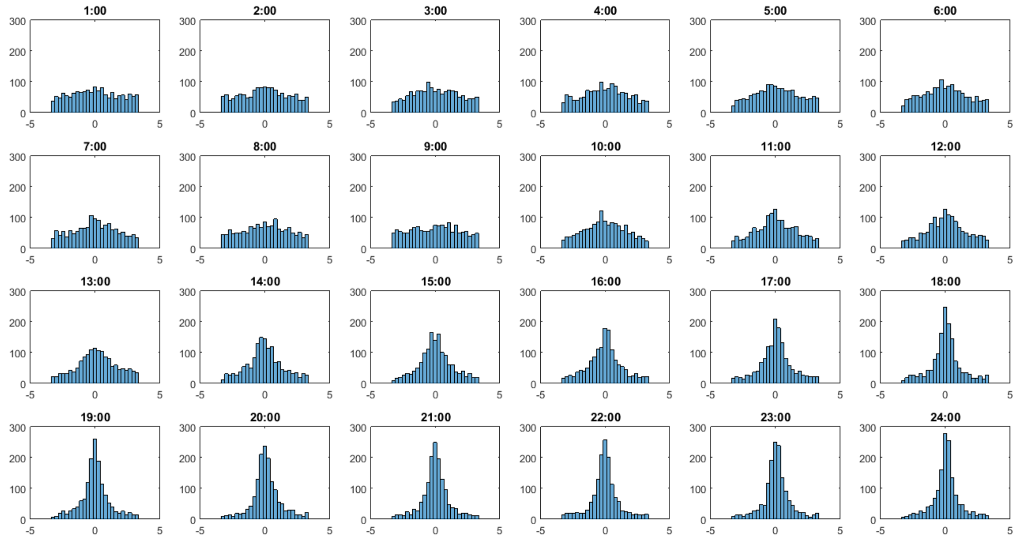


Figure 5: Histograms of time delays of all active stations recorded over 24 hours. The histogram shows how many connections have a particular delay at the instant.

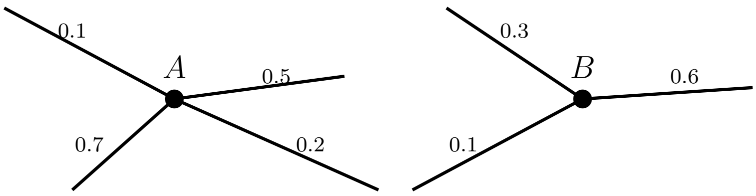


Figure 6: Distance metric.

we can formulate a cost function that measures the deviation of the observed weight matrix's metrics $f_i(\hat{\boldsymbol{\delta}})$ to the estimates K_i as:

$$c(\hat{\boldsymbol{\delta}}) = \sum_i E_i^2(\hat{\boldsymbol{\delta}}) = \sum_i (f_i(\hat{\boldsymbol{\delta}}) - K_i)^2 \tag{4}$$

The error is minimized with gradient descent updates on $\hat{\boldsymbol{\delta}}$:

$$\hat{\boldsymbol{\delta}}^{(t+1)} = \hat{\boldsymbol{\delta}}^t - \mu \sum_i E_i(\hat{\boldsymbol{\delta}}^t) \frac{df_i(\hat{\boldsymbol{\delta}}^t)}{d\hat{\boldsymbol{\delta}}^t}, \tag{5}$$

where t is the iteration index and μ the learning rate. We find the derivative for the gradient descent update step. For $i = 2$ it is

$$\frac{df_2}{d\boldsymbol{\delta}} = \frac{\sum_j \hat{\delta}_{2j}}{d\hat{\boldsymbol{\delta}}} = \begin{bmatrix} 0 & 1 & 0 & \dots & 0 \\ 1 & 0 & 1 & \dots & 1 \\ 0 & 1 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 1 & 0 & \dots & 0 \end{bmatrix}. \tag{6}$$

It is a symmetric matrix with ones in the ith row and ith column. The diagonal entries remain zero.

4.2 Clock drift estimate

The denoising model in section 4.1 computes the denoised estimate $\boldsymbol{\delta}$ of the measurement data $\hat{\boldsymbol{\delta}}$. The algorithm is run over a series of time points and an estimate for the clock drift between each station pair is obtained.

A polynomial fitting algorithm !!!TODO: specify is run to find an estimate for the trend in noise. The error term in the fitting procedure will account for all the other errors in the system. The trend is considered to be caused by the clock drifts of the two stations.

The noise, ε_{ij}, between each pair between two stations in the system is modelled with

$$\varepsilon_{ij} = \Delta_i - \Delta_j + e_{ij}, \tag{7}$$

in which $\Delta_i - \Delta_j$ is the clock drift between two stations i and j, and e_{ij} accounts for all other ambient noise in the system in addition to the clock drift.

The clock errors of individual stations can be written in matrix form

$$\begin{bmatrix} 1 & -1 & 0 & 0 & \dots & 0 \\ 1 & 0 & -1 & 0 & \dots & 0 \\ 1 & 0 & 0 & -1 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 0 & 0 & 0 & \dots & -1 \\ 0 & 1 & -1 & 0 & \dots & 0 \\ 0 & 1 & 0 & -1 & \dots & 0 \\ 0 & 1 & 0 & 0 & \ddots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 0 & 1 & -1 \end{bmatrix} \begin{bmatrix} \Delta_1 \\ \Delta_2 \\ \Delta_3 \\ \Delta_4 \\ \vdots \\ \Delta_k \\ \vdots \\ \Delta_{n-1} \\ \Delta_n \end{bmatrix} = \begin{bmatrix} \Delta_1 - \Delta_2 \\ \Delta_1 - \Delta_3 \\ \Delta_1 - \Delta_4 \\ \vdots \\ \Delta_1 - \Delta_n \\ \Delta_2 - \Delta_3 \\ \Delta_2 - \Delta_4 \\ \vdots \\ \Delta_2 - \Delta_k \\ \vdots \\ \Delta_{n-1} - \Delta_n \end{bmatrix} = \begin{bmatrix} \Delta_{12} \\ \Delta_{13} \\ \Delta_{14} \\ \vdots \\ \Delta_{1n} \\ \Delta_{23} \\ \Delta_{24} \\ \vdots \\ \Delta_{2k} \\ \vdots \\ \Delta_{(n-1)n} \end{bmatrix},$$

which represents a classical overdetermined inversion problem of the form $\mathbf{Gm} = \mathbf{s}$, with m being the unknown model vector and s being the known data [1]. The matrix $\mathbf{G}$ has rank $n-1$ because a constant offset can be added to all stations simultaneously without changing the time differences. Hence it is lacking full rank, therefore Tikhonov reqularization is employed to overcome this issue. The ordinary least squares seeks to minimize the sum of squared residuals

$$\|\mathbf{Gm} - \mathbf{s}\|_2^2. \tag{8}$$

When regularisation is added to the Equation 8, we obtain

$$\|\mathbf{Gm} - \mathbf{s}\|_2^2 + \|\mathbf{\Gamma m}\|, \tag{9}$$

for some suitable Tikhonov matrix $\mathbf{\Gamma}$. We choose this matrix as a multiple of the identity matrix $\mathbf{\Gamma} = \alpha \mathbf{I}$, giving preference to smaller norms.

The explicit solution is hence given by

$$\hat{\mathbf{s}} = (\mathbf{G}^T\mathbf{G} + \mathbf{\Gamma}^T\mathbf{\Gamma})^{-1}\mathbf{G}^T\mathbf{m}. \tag{10}$$

4.3 Algorithm

Leevi to plug in pseudo code of the implementation of the method

5 Results

5.1 Signal denoising

5.2 Clock deviations

6 Conclusions and Outlook

By analyzing the delay data we clearly showed that it changes over time showing patterns other than only white noise. This is due to clock drift. We used a method of weighted network estimation by the use of topological graph metrics exploiting the relation between delay and distances to denoise the data. The method is robust to temporarily disconnected stations that are only included as they are only included in the calculations when they are active. Hence, the algorithm can be run for on larger networks. The algorithm produced denoised delays we could deduce clock drifts from. Due to missing validation data we cannot make conclusions on the quality of our estimates. The algorithm could either be run on synthetic data where the results are known or on data sets that include continuous GPS synchronized data. To get more compelling results other metrics could be used, or even a combination of multiple metrics.

References

[1] C. Sens-Schonfelder. Synchronizing seismic networks with ambient noise. *Geophysical Journal International*, 174(3):966–970, 2008.

[2] L. Spyrou and J. Escudero. Weighted network estimation by the use of topological graph metrics. *CoRR*, abs/1705.00892, 2017.