ETH zürich

Master Thesis Proposal

Dimensionality Reduction and Source Separation of Graph-Temporal Data

Advisor: Bastian Seifert
Professor: Markus Püschel

October 13, 2021

A classical problem in signal processing is to recover original component signals from a mixture signal. Examples include the cocktail party problem, where several people are talking in room and one wants to recover the individual speech of each from measurements in the room, or EEG signals, where one tries to find the sources of certain electrical activities in the brain. Typically, there are also sources of noise in the mixture signal.

The recently proposed method Dynamical Component Analysis (DyCA) [1, 2, 3] assumes that the multivariate time-series $\mathbf{q}(t) \in \mathbb{R}^N$ is the mixture of time-dependent amplitudes $\mathbf{x}_i(t)$

$$\mathbf{q}(t) = \sum_{i=1}^{n} \mathbf{x}_i(t)\mathbf{w}_i, \tag{1}$$

with $\mathbf{w}_i \in \mathbb{R}^N$ being the mixing vectors, and the amplitudes are described by a set of differential equations

$$\begin{aligned} \frac{\partial \mathbf{x}_i}{\partial t} &= \sum_{k=1}^{N} a_{i,k}\mathbf{x}_k, \text{ for } i = 1, \dots, m \\ \frac{\partial \mathbf{x}_i}{\partial t} &= f_i(\mathbf{x}_1, \dots, \mathbf{x}_n), \text{ for } i = m+1, \dots, n, \end{aligned} \tag{2}$$

with f_i non-linear smooth functions. DyCA then performs simultaneous dimensionality reduction and source separation by finding an optimal projection onto a subspace of $\mathbb{R}^N$ using time-averaged correlation matrices of $\mathbf{q}$ and its time derivative.

However, the use of correlation matrices does not make use of all available information, as the sensors measuring the multivariate time-series are typically connected by a network structure, modelled by a graph. Graph signal processing (GSP) [4, 5] on the other hand makes explicit use of the network structure, by defining a graph Fourier transform for signals supported on a network, but is not taking into account the temporal dynamics. For example, if as adjacency matrix the correlation matrix of a data set is taken, the graph Fourier transform specializes to Principal Component Analysis (PCA) [6].

Your contribution The goal of this project is to implement a graph prior into DyCA and to evaluate the perfomance of the obtained method.

In particular your tasks are:

a) Derive a variant of DyCA, which uses a graph prior in form of adjacency matrices instead of correlation matrices.

b) Evaluate the obtained method on epileptic EEG data against DyCA and temporal-agnostic graph Fourier transform.

c) Find additional application areas, e.g., from opinion dynamics, and evalute the method on data sets from these areas.

Deliverables

Final report: The final report may be written in English or German. It must contain an abstract written in both English and German. It should include an introduction, an analysis of related work, and a complete documentation of all used software tools and mathematical derivations. Three copies of the final report must be delivered to the supervisor.

Reproducible experimental setup: Implementations, configuration scripts and instructions to reproduce the results reported in the thesis must be delivered in electronic form.

Presentation: The results of the thesis must be presented during a software-group seminar. The presentation is capped to 30 minutes and should give an overview as well as the most important details of the work.

Contact If you are interested in pursuing this master thesis, please contact bastian.seifert@inf.ethz.ch or pueschel@ethz.ch.

References

[1] B. Seifert, K. Korn, S. Hartmann, and C. Uhl. Dynamical Component Analysis (DyCA): Dimensionality reduction for high-dimensional deterministic time-series. In *Proc. Int. Workshop Mach. Learn. Signal Process. (MLSP)*, pages 1–6, 2018.

[2] K. Korn, B. Seifert, and C. Uhl. Dynamical Component Analysis (DyCA) and its application on epileptic EEG. In *Proc. Int. Conf. Acoust., Speech, and Signal Process. (ICASSP)*, pages 1100–1104, 2019.

[3] C. Uhl, M. Kern, M. Warmuth, and B. Seifert. Subspace Detection and Blind Source Separation of Multivariate Signals by Dynamical Component Analysis (DyCA). *IEEE Open J. Signal Process.*, 1:230–241, 2020.

[4] D. Shuman, S. K. Narang, P. Frossard, A. Ortega, and P. Vandergheynst. The emerging field of signal processing on graphs: Extending high-dimensional data analysis to networks and other irregular domains. *IEEE Signal Process. Mag.*, 30(3):83–98, 2013.

[5] A. Sandryhaila and J. M. F. Moura. Discrete Signal Processing on Graphs. *IEEE Trans. Signal Process.*, 61(7):1644–1656, 2013.

[6] G. Mateos, S. Segarra, A. G. Marques, and A. Ribeiro. Connecting the dots: Identifying network structure via graph signal processing. *IEEE Signal Process. Mag.*, 36(3):16–43, 2019.