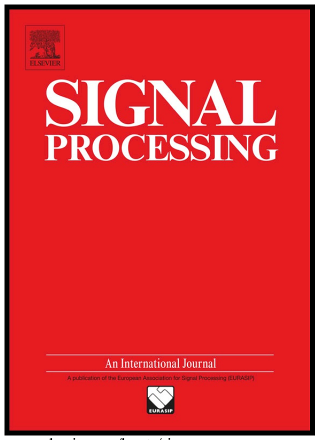
Using Graph Clustering to Locate Sources within a Dense Sensor Array

Nima Riahi, Peter Gerstoft



www.elsevier.com/locate/sigpro

PII: S0165-1684(16)30267-5
DOI: http://dx.doi.org/10.1016/j.sigpro.2016.10.001
Reference: SIGPRO6276

To appear in: *Signal Processing*

Received date: 19 February 2016
Revised date: 2 October 2016
Accepted date: 3 October 2016

Using Graph Clustering to Locate Sources within a Dense Sensor Array

Nima Riahi (corresponding author)[a], Peter Gerstoft[a]

[a]*Scripps Institution of Oceanography, University of California San Diego, La Jolla, CA 92093*

Abstract

We develop a model-free technique to identify weak sources within dense sensor arrays using graph clustering. No knowledge about the propagation medium is needed except that signal strengths decay to insignificant levels within a scale that is shorter than the aperture. We then reinterpret the spatial coherence matrix of a wave field as a matrix whose support is a connectivity matrix of a graph with sensors as vertices. In a dense network, well-separated sources induce clusters in this graph. The geographic spread of these clusters can serve to localize the sources. The support of the covariance matrix is estimated from limited-time data using a hypothesis test with a robust phase-only coherence test statistic combined with a physical distance criterion. The latter criterion ensures graph sparsity and thus prevents clusters from forming by chance. We verify the approach and quantify its reliability on a simulated dataset. The method is then applied to data from a dense 5200 element geophone array that blanketed 7 km ×10 km of the city of Long Beach (CA). The analysis exposes a helicopter traversing the array and oil production facilities.

Keywords: Seismic arrays; Source localization; Graph clustering; Non-parametric estimation

1. Introduction

Large and dense sensor arrays are becoming more common as the cost for sensor and communications hardware decreases. Examples of such arrays are the USArray initiative in seismology with 500 stations covering large parts of the continental US [1] or the seismic exploration array with 5200 sensors as presented here. As array sizes increase the occurrence of within-aperture source events that cause coherent signals over only a small fraction of all sensors becomes more common.

This work addresses the problem of localizing such weak sources in a complex and unknown environment. For known or well characterized media this problem has been addressed using frameworks such as matched field processing (MFP), maximum likelihood methods, or migration techniques in, e.g., acoustics [2, 3, 4, 5], seismology [6, 7, 8, 9, 10], infrasound acoustics [11, 12], and electromagnetics [13, 14]. The eigen-structure of the array covariance matrix or its inverse plays an important role in these approaches, in particular for data-adaptive implementations using, e.g., MVDR [15] or MUSIC [16]. A possible solution to locate sources in dense networks without relying on medium information is the spatiotemporal analysis of signal intensity [17, 18]. However, using power alone cannot detect sources near or below the noise floor.

We present a model-free analysis approach that can work also for weak signals through the use of coherent averaging. The only assumption made is that source signals enter the noise floor within a distance that is much smaller than the array aperture. That requirement is realistic for large arrays based on wave propagation in moderately attenuating media such as the earth. We follow a graph-based analysis paradigm [19, 20]: The sensors are arranged as vertices in a graph with edges between vertices existing if the corresponding sensors share a common coherent signal. Thus the graph connectivity matrix is defined as the support of the array coherence matrix.

For long observation times, we demonstrate that finding weak within-aperture sources is tantamount to identifying connected components in this graph (here referred to as clusters). Such clusters can be found through an eigenvalue decomposition of a matrix (the graph Laplacian) that is derived from the connectivity matrix [21]. In the limited-data case stochastic fluctuations in the coherence can create spurious connections in the graph, thus leading to large connected components, a well known phenomenon studied in random

Email addresses: nriahi@ucsd.edu (Nima Riahi (corresponding author)), gerstoft@ucsd.edu (Peter Gerstoft)