Signal Processing 107 (2015) 109–122

Contents lists available at ScienceDirect

Signal Processing

journal homepage: www.elsevier.com/locate/sigpro

Distributed estimation of the inverse of the correlation matrix for privacy preserving beamforming

Yuan Zeng *, Richard C. Hendriks

Delft University of Technology, Mekelweg 4, 2628 CD Delft, The Netherlands

ARTICLE INFO

Article history:
Received 3 February 2014
Received in revised form
9 July 2014
Accepted 9 July 2014
Available online 24 July 2014

Keywords:
Distributed estimation
Gossip processing
Correlation matrix
Wireless acoustic sensor networks

ABSTRACT

In this paper, we consider a privacy preserving scenario where users in the network want to perform distributed target source estimation with a wireless acoustic sensor network (WASN), without revealing the actual source of interest to other entities in the network. This implies that users do not share the steering vector of the beamformer with any other party. For distributed multi-channel noise reduction in WASNs, distributed estimation of the inverse noise or noise+target correlation matrix is an important aspect and in general a challenging problem.

To make both privacy preservation and distributed multi-channel noise reduction possible, we make use of the fact that recursive estimation of the inverse correlation matrix can be structured as a consensus problem and can be realized in a distributed manner via the randomized gossip algorithm. This makes it possible to compute the MVDR in distributed manner without revealing the steering vector to any of the other entities in the network, and providing privacy about the actual source of interest. We provide theoretical analysis and numerical simulations to investigate the convergence error between the gossip-based estimated correlation matrix and the centralized estimated correlation matrix. It is shown that the convergence error accumulates across time without using a sufficient number of transmissions in the gossip-based algorithm. To eliminate this convergence error, we propose in addition a clique-based algorithm for distributed estimation of the inverse correlation matrix (CbDECM). Theoretical analysis shows that the CbDECM algorithm converges to the centralized estimate of the matrix inverse.

We investigate the performance of the presented clique-based distributed framework in combination with a distributed privacy preserving MVDR beamformer, where information about the actual source of interest is kept private. Simulations show that the proposed algorithm converges to the centralized MVDR beamformer.

1. Introduction

To improve the quality and intelligibility of speech processing applications under noisy environments, it is customary to equip them with a single- or multi-microphone noise reduction algorithm (for an overview see e.g., [1–3]). As multi-microphone noise reduction algorithms can take advantage of spatial diversity, they usually lead to better speech quality and intelligibility than their single-microphone counterparts. In particular it is the number of microphones and their placement that determine the potential performance of a multi-microphone noise reduction algorithm. However, as most mobile speech processing

* Corresponding author.
E-mail addresses: Y.Zeng@tudelft.nl (Y. Zeng), R.C.Hendriks@tudelft.nl (R.C. Hendriks).

http://dx.doi.org/10.1016/j.sigpro.2014.07.006
0165-1684/

devices have relatively small dimensions, the number of microphones as well as their placement is rather restricted.

Using the so-called wireless acoustic sensor networks (WASNs), it is possible to use a much larger number of microphones that are distributed in the environment and where their placement is not restricted by the device itself. This allows a further increase in noise reduction performance. However, the conventional multi-microphone noise reduction algorithms (e.g. [2,3]) are characterized by having one processor where all data is processed centrally. Such centralized algorithms are less suitable for a WASN, as they may require higher energy consumption or transmission bandwidth than necessary. The fact that the sensors in a WASN are all equipped with a (simple) processor potentially owned by different users allows us to perform intermediate processing of data without the need to first send all data to a single point in the network. This has recently led to an increased research interest to distributed speech enhancement algorithms, see e.g., [4–7].

As the processors and sensors in the WASN context are not necessarily anymore owned by a single user, distributed processing might come with serious privacy risks. These could range from an increased risk of being eavesdropped to an increased risk that private data or information becomes public. Within the speech enhancement context, such privacy issues were first addressed in [8,9] for two scenarios. The scenario in [8] considered the case where a user keeps the exact source of interest private for other users, while [9] considered the scenario where eavesdropping by untrusted third parties is overcome. Both contributions employed homomorphic encryption [10] to provide the necessary privacy. However, homomorphic encryption is computationally very complex, and requires very high bit rates for data transmission. In the current paper, we consider a different approach and develop a framework for distributed signal estimation employing a WASN, while providing the user a certain level of privacy with respect to the source of interest. We consider the case where the users of the network do not want to share to which specific source in the environment they are listening, while they do want to make use of the WASN to estimate their signal of interest.

More specifically, the application scenario that we consider in this paper is the one where multiple users make use of a WASN that consists of many processors (including their own) and where each processor is equipped with multiple microphones. The users can use the additional sensors in the WASN to obtain an improved estimate of their signal of interest, which can be different for each user and is usually determined by the steering vector of the beamformer. However, due to privacy reasons, the users want to keep their source of interest private (i.e., the steering vector). Although the microphone signals might be public, hiding the steering vector will overcome that the exact combination of microphone signals required for specific target signal estimation is publicly known. Moreover, hiding the steering vector makes sure that none of the entities with access to the network is able to reveal which conversation or source is apparently of interest for a particular user. This will guarantee a certain amount of privacy to the users of the network. One way to guarantee privacy preservation on the source of interest would be to send all data to all nodes and compute a conventional beamformer in every node. In this way, users do not need to make the steering vector public. However, this requires a lot of data transmission. Performing calculations in a distributed way will reduce the number of data transmissions, due to the fact that local nodes perform intermediate calculations. This leads to a data compression depending on the number microphones per node. We investigate thus the possibility that each user in the WASN estimates his signal of interest by performing distributed computations on the WASN data, while keeping the particular source of interest private. To do so, we concentrate on distributed estimation of one of the most well-known beamformers, the minimum variance distortionless response (MVDR) beamformer.

The MVDR beamformer depends on the inverse of the noise or noise+target spectral correlation matrix. Computing this inverse in a distributed manner is not trivial, as the data in a WASN is not centrally present and each element of the inverse of the correlation matrix is a function of the statistics of the noise or the noise+target at multiple microphones. In [7], a distributed MVDR was presented based on a randomized gossip algorithm [11] where the inversion of the noise correlation matrix was overcome by assuming the correlation matrix to be diagonal. This simplifies distributed computation of the MVDR, but it also compromises the performance as the noise is assumed to be uncorrelated across microphones. In [12] the MVDR was computed using a message passing algorithm [13]. However, this requires the network topology to be consistent with the noise correlation matrix, where two nodes are neighbors if their noise cross correlation is unequal to zero. This would require to adjust the transmission range of the nodes in the network to the correlation matrix and consequently, increase the energy usage for transmission or decrease the connectivity in the network. In [5] and [6], computation of the inverse of the correlation matrix was overcome by employing the generalized sidelobe canceller structure. However, this algorithm also constrains the topology of the network to be fully connected.

The contribution of this paper is twofold. First, we present a method where each user can estimate a different signal of interest from a mix of many different signals by means of a distributed MVDR beamformer without the need to reveal the source of interest to other entities in the network. In order to do this, we develop an algorithm that enables distributed estimation of the inverse of a correlation matrix, which is the second contribution of this paper. This algorithm for distributed estimation of the matrix inverse is based on the observation that in practice, correlation matrices are usually estimated recursively by exponential smoothing. Using the Sherman–Morrison formula [14], estimation of the inverse of the correlation matrix can be seen as a consensus problem and can be realized using gossip algorithms. Although the convergence error per time frame is decreased with increasing number of iterations when using gossip algorithms, estimation errors might accumulate. This is caused by the fact that the correlation matrix is recursively estimated across time. These convergence errors can be eliminated using