Distributed Compressed Estimation Based on Compressive Sensing for Wireless Sensor Networks

Joint work with Songcen Xu and Vincent Poor

Rodrigo C. de Lamare
CETUC, PUC-Rio, Brazil
Communications Research Group,
Department of Electronics, University of York, U.K.

delamare@cetuc.puc-rio.br
Outline

• Review of Compressive Sensing

• Introduction

• System model

• Proposed distributed compressed estimation (DCE) scheme

• Measurement matrix optimization

• Proposed adaptive algorithm

• Simulations

• Conclusions
Review of Compressive Sensing

• Compressive sensing has gained substantial interest amongst researchers and was applied to areas such as image processing, radar, sonar and wireless communications.

• Compressive sensing theory states that an S-sparse signal $\mathbf{w}_o$ of length M can be recovered with high probability from $O(S \log_2 M)$ measurements.

• Mathematically, the vector $\bar{\mathbf{w}}_o$ with length D that carries sufficient information about $\mathbf{w}_o$ can be obtained by the linear model

$$
\bar{\mathbf{w}}_o = \Phi \mathbf{w}_o,
$$

where $\Phi$ is the D x M measurement matrix.

• A reconstruction algorithm is then applied to obtain an estimate of $\mathbf{w}_o$.

• Prior work on distributed compressive sensing is limited and has not considered compressed transmit strategies and estimation techniques.

• Distributed signal processing algorithms are of great importance for statistical inference in wireless networks.

• Such techniques enable a network to perform statistical inference from data collected from nodes that are geographically distributed.

• In this context, for each node a set of neighbour nodes collect and process their local information, and transmit their estimates to a specific node.

• Then, each specific node combines the collected information together with its local estimate to generate improved estimates.

Key problems:
- In many scenarios, the unknown parameter vector to be estimated can be very large, sparse and contain only a few nonzero coefficients.
- Most distributed algorithms to date do not exploit the structure of the unknown parameter vector and process the full-dimensional data.

Our contributions:
- Inspired by Compressive Sensing, we present a scheme that incorporates compression and decompression modules into the distributed processing.
- We present a design procedure and an adaptive algorithm to optimize the measurement matrices, resulting in improved performance.
We consider a wireless network with N nodes:
- We assume that the network is partially connected.
- Communication protocols: incremental, consensus and diffusion.

Main tasks:
- To collect and process data locally at the node and then transmit estimates.
- To perform distributed estimation by exchanging information about local estimates.

Application scenario:
- Wireless sensor networks

System model (2/2)

• System model:

\[ d_k(i) = \omega_0^H x_k(i) + n_k(i), \quad i = 1, 2, \ldots, I, \]

• Problem:

\[
\min J_\omega(\omega) = \sum_{k=1}^{N} \mathbb{E}\{|d_k(i) - \omega^H x_k(i)|^2\}
\]

• Adapt-then-combine (ATC) diffusion LMS algorithm:

For each time instant \( i = 1, 2, \ldots, n \)
For each node \( k = 1, 2, \ldots, N \)
\[
\psi_k(i) = \omega_k(i) + \mu_k x_k(i) \left[ d_k(i) - \omega_k^H(i) x_k(i) \right]^* 
\]
end
For each node \( k = 1, 2, \ldots, N \)
\[
\omega_k(i + 1) = \sum_{l \in \mathcal{N}_k} c_{kl} \psi_l(i)
\]

Proposed DCE scheme (1/3)

- Measurement:
  \[ d_k(i) = \bar{\omega}_0^H \bar{x}_k(i) + n_k(i), \quad i = 1, 2, \ldots, I, \]
  where \( \bar{\omega}_0 = \Phi_k \omega_0 \) and \( \bar{x}_k(i) \) is the D x 1 input signal vector.

- Adaptation:
  \[ \bar{\psi}_k(i) = \bar{\omega}_k(i) + \mu_k(i)e_k^*(i)\bar{x}_k(i) \]
  where \( e_k(i) = d_k(i) - \bar{\omega}_k^H(i)\bar{x}_k(i) \)

- Information exchange:
  \( \bar{\psi}_k(i) \) with D parameters

- Combination:
  \[ \bar{\omega}_k(i + 1) = \sum_{l \in \mathcal{N}_k} c_{kl} \bar{\psi}_l(i) \]

- Decompression:
  - Computation of \( \omega_o \) with the OMP algorithm.
Proposed DCE scheme (2/3)

- Each node observes the $M \times 1$ vector $x_k(i)$

- The $D \times M$ measurement matrix obtains the $D \times 1$ compressed version $\bar{x}_k(i)$, which is the input to the compression module.

- The estimation of $\omega_o$ is performed in the compressed domain.

- A decompression module employs a measurement matrix and a reconstruction algorithm to compute an estimate of $\omega_o$ at each node.
Proposed DCE scheme (3/3)

\[
\begin{align*}
\text{Initialize: } & \overline{\omega}_k(1)=0 \\
\text{For each time instant } i=1,2, \ldots, I-1 & \\
\text{For each node } k=1,2, \ldots, N & \\
\overline{\psi}_k(i) &= \overline{\omega}_k(i) + \mu(i)e_k^*(i)\overline{\omega}_k(i) \\
\text{where } e_k(i) &= d_k(i) - \overline{\omega}^H_k(i)\overline{\omega}_k(i), \\
d_k(i) &= \overline{\omega}^H_0\overline{\omega}_k(i) + n_k(i) = (\Phi_k\omega_0)^H\overline{\omega}_k(i) + n_k(i) \\
\text{and } \Phi_k &\text{ is the } D \times M \text{ random measurement matrix} \\
\text{end} & \\
\text{For each node } k=1,2, \ldots, N & \\
\overline{\omega}_k(i+1) &= \sum_{l \in N_k} c_{kl}\overline{\psi}_l(i) \\
\text{end} & \\
\text{end} & \\
\text{After the final iteration } I & \\
\text{For each node } k=1,2, \ldots, N & \\
\omega_k(I) &= f_{\text{OMP}}\{\omega_k(I)\} \\
\text{where } \omega_k(I) &\text{ is the final decompressed estimator.} \\
\text{end} & \\
\end{align*}
\]
• Computational complexity:

  – Proposed DCE scheme: $O(NDI + ND^3)$

  – Distributed NLMS algorithm: $O(NMI)$


  – Distributed sparsity-aware NLMS algorithm: $O(3NMI)$


  – Distributed compressive sensing: $O(NMI + ND^3I)$

    C. Wei, M. Rodrigues, and I. J. Wassell, “Distributed compressive sensing reconstruction via common support discovery,” in IEEE International Conference on Communications, Kyoto, Japan, June 2011
Measurement matrix optimization (1/3)

- Design of the measurement matrix:
  - MSE metric
  - Adaptive and can track changes in the environment
  - Distributed and robust against link failures

- Minimization of MSE cost function:

\[
J = \mathbb{E}\{|e_k(i)|^2\} = \mathbb{E}\{|d_k(i) - y_k(i)|^2\}
\]

where \(y_k(i) = \bar{\omega}_k^H(i) \bar{x}_k(i)\) is the estimate by sensor \(k\) and the measured signal is \(d_k(i) = \bar{\omega}_o^H(i) \bar{x}_k(i) + n_k(i) = \omega_o^H(i) \Psi_k^H \bar{x}_k(i) + n_k(i)\).
Measurement matrix optimization (2/3)

- Expanding the cost function and taking the first 3 terms, we obtain

\[
\mathbb{E}\{|d_k(i)|^2\} - \mathbb{E}\{d_k^*(i)y_k(i)\} - \mathbb{E}\{d_k(i)y_k^*(i)\} \\
= \mathbb{E}\{\omega_0^H \Phi_k^H(i)\bar{x}_k(i)|^2\} + \mathbb{E}\{(\omega_0^H \Phi_k^H(i)\bar{x}_k(i))^*n_k(i)\} \\
+ \mathbb{E}\{(\omega_0^H \Phi_k^H(i)\bar{x}_k(i))n_k^*(i)\} + \mathbb{E}\{|n_k(i)|^2\} \\
- \mathbb{E}\{(\omega_0^H \Phi_k^H(i)\bar{x}_k(i))^*y_k(i)\} - \mathbb{E}\{n_k^*(i)y_k(i)\} \\
- \mathbb{E}\{(\omega_0^H \Phi_k^H(i)\bar{x}_k(i))y_k^*(i)\} - \mathbb{E}\{n_k(i)y_k^*(i)\}.
\]

- Exploiting the fact that the random variable \(n_k(i)\) is statistically independent from the other terms and has zero mean, we have

\[
\mathbb{E}\{|d_k(i)|^2\} - \mathbb{E}\{d_k^*(i)y_k(i)\} - \mathbb{E}\{d_k(i)y_k^*(i)\} \\
= \mathbb{E}\{\omega_0^H \Phi_k^H(i)\bar{x}_k(i)|^2\} + \sigma_{n,k}^2 - \mathbb{E}\{(\omega_0^H \Phi_k^H(i)\bar{x}_k(i))^*y_k(i)\} \\
- \mathbb{E}\{(\omega_0^H \Phi_k^H(i)\bar{x}_k(i))y_k^*(i)\}.
\]
Measurement matrix optimization (3/3)

- We compute the gradient of the terms of the cost function and equate them to a zero matrix:

\[ \nabla J_{\Phi_k^*(i)} = R_k(i)\Phi_k(i)R_{\omega_0} - P_k(i) = 0 \]

where \( R_k(i) = \mathbb{E}\{\bar{x}_k(i)\bar{x}_k^H(i)\} \), \( R_{\omega_0} = \mathbb{E}\{\omega_0\omega_0^H\} \) and \( P_k(i) = \mathbb{E}\{y_k^*(i)\bar{x}_k(i)\omega_0^H\} \)

- The resulting system of equations is given by

\[ \Phi_k(i) = R_k^{-1}(i)P_k(i)R_{\omega_0}^{-1} \]

- Note that the above expression cannot be solved in closed-form because \( \omega_0 \) is unknown.

- Solution: adaptive algorithms
Proposed adaptive algorithm (1/2)

- Development of proposed adaptive algorithm:
  - We employ instantaneous values for the statistical quantities:
    \[
    \hat{R}_k(i) = \bar{x}_k(i)\bar{x}_k(i)
    \]
    \[
    \hat{R}_{\omega_0} = \omega_0\omega_0^H
    \]
    \[
    \hat{P}_k(i) = y_k^*(i)\bar{x}_k(i)\omega_0^H
    \]
  - Steepest descent rule with instantaneous values:
    \[
    \Phi_k(i+1) = \Phi_k(i) + \eta[-\nabla J_{\Phi_k^*}(i)]
    \]
    \[
    = \Phi_k(i) + \eta[\hat{P}_k(i) - \hat{R}_k(i)\Phi_k(i)\hat{R}_{\omega_0}]
    \]
    \[
    = \Phi_k(i) + \eta[y_k^*(i)\bar{x}_k(i)\omega_0^H - \bar{x}_k(i)\bar{x}_k(i)\Phi_k(i)\omega_0\omega_0^H]
    \]

  where \(\eta\) is the step size and \(\omega_0\) is the \(M \times 1\) unknown parameter vector.
Proposed adaptive algorithm (2/2)

• The $D \times 1$ parameter vector $\overline{\omega}_o$ is used in the reconstruction of the $M \times 1$ parameter vector $\omega_o$:

$$\omega_{\text{re}k}(i) = f_{\text{OMP}}\{\overline{\omega}_k(i)\}$$

where $f_{\text{OMP}}\{ . \}$ denotes the orthogonal matching pursuit (OMP) algorithm.

• Replacing $\omega_o(i)$ with $\omega_{\text{re}k}(i)$, we obtain the proposed measurement matrix optimization algorithm:

$$\Phi_k(i + 1) = \Phi_k(i) + \eta \left[ y_k(i) \bar{x}_k(i) \omega_{\text{re}k}^H(i) - \bar{x}_k(i) \bar{x}_k^H(i) \Phi_k(i) \omega_{\text{re}k}(i) \omega_{\text{re}k}^H(i) \right]$$

• Computational complexity: $O(NDI + ND^3I)$
Simulations (1/5)

- The proposed DCE scheme and algorithms are assessed in a wireless sensor network application.

- We consider a network with $N = 20$ nodes, parameter vectors with $M = 50$ coefficients, $D = 5$, and sparsity of $S = 3$ non-zero coefficients.

- The input signal is generated as $x_k(i) = [x_k(i) \ x_k(1) \ \ldots \ x_k(i-M+1)]^T$ and $x_k(i) = u_k(i) + \alpha_k x_k(i-1)$ where $\alpha_k = 0.5$ is a correlation coefficient.

- The compressed input signal is obtained by $\bar{x}_k(i) = \Phi_k x_k(i)$.

- The measurement matrix is an i.i.d. Gaussian random matrix.

- The noise is modelled as complex Gaussian noise with variance 0.001.
Simulations (2/5)

The proposed DCE scheme and algorithms are compared with:

- **Distributed NLMS algorithm: O(NMI)**
  

- **Distributed sparsity-aware NLMS algorithm: O(3NMI)**
  

- **Distributed compressive sensing: O(NMI +ND^3I)**
  
  C. Wei, M. Rodrigues, and I. J. Wassell, “Distributed compressive sensing reconstruction via common support discovery,” in IEEE International Conference on Communications, Kyoto, Japan, June 2011
Simulations (3/5)

MSE performance x time:

- The DCE scheme has a significantly faster convergence and a better MSE performance than other algorithms.

- This is because of the reduced dimension and the fact that compressive sensing is implemented in the estimation layer.
MSE performance x time:

- The DCE scheme can achieve an even faster convergence when compared with the DCE scheme without the measurement matrix optimization and other algorithms.
MSE performance x D:

• We compare the DCE scheme with the distributed NLMS algorithm with different levels of resolution in bits per coefficient, reduced dimension D and sparsity level S.

• With the increase of the sparsity level S, the MSE performance degrades.

• The MSE performance improves with the increase of the number of bits.
Conclusions

- We have proposed a novel DCE scheme and algorithms for sparse signals and systems based on CS techniques.

- In the DCE scheme, the estimation procedure is performed in a compressed dimension which reduces the required bandwidth.

- We have also proposed a measurement matrix optimization strategy and an algorithm to design the measurement matrix online.

- The results for a WSN application show that the DCE scheme outperforms prior work in terms of convergence rate, bandwidth and performance.
Thanks!

- For further details please contact me at delamare@cetuc.puc-rio.br

- The paper outlining the mains ideas here appeared in IEEE Signal Processing Letters:

References (1/4)


