Distributed conjugate gradient strategies for distributed estimation over sensor networks

Songcen Xu and Rodrigo C. de Lamare
Department of Electronics, University of York,
York, United Kingdom
Emails: sx520@york.ac.uk, rcdl500@york.ac.uk

THE UNIVERSITY of York

Outline

- Introduction
- System model and problem statement
- Proposed incremental distributed CG – based algorithms
- Proposed diffusion distributed CG – based algorithms
- Analysis of the algorithms
- Simulations
- Conclusions

Introduction

- Distributed processing has become popular in wireless communication networks. It can collect data at each node in a given area and convey the information to the whole network in a distributed way.
- When compared with the centralized solution, the distributed solution has significant advantages.
- This kind of strategy can significantly reduce the amount of processing and the communications bandwidth requirements.
- The problem we are interested in solving: distributed estimation over sensor networks.
- We propose distributed CG algorithms with both incremental and diffusion adaptive strategies for distributed estimation over sensor networks.

System model and problem statement

- We focus on the processing of an adaptive filter for adjusting the weight vector $\boldsymbol{\omega}_o$ with coefficients $\boldsymbol{\omega}_k$.

- The desired signal of each node at time instant i:

$$d^{(i)} = \boldsymbol{\omega}_0^H \boldsymbol{x}^{(i)} + n^{(i)}, \quad i = 1, 2, \ldots, N,$$

- The output of the adaptive filter for each node:

$$y^{(i)} = \boldsymbol{\omega}^{(i)H} \boldsymbol{x}^{(i)}, \quad i = 1, 2, \ldots, N,$$

- The problem: minimize the cost function J(ω)

$$J(\boldsymbol{\omega}) = \boldsymbol{E}[||\boldsymbol{d} - \boldsymbol{X}\boldsymbol{\omega}||^2],$$

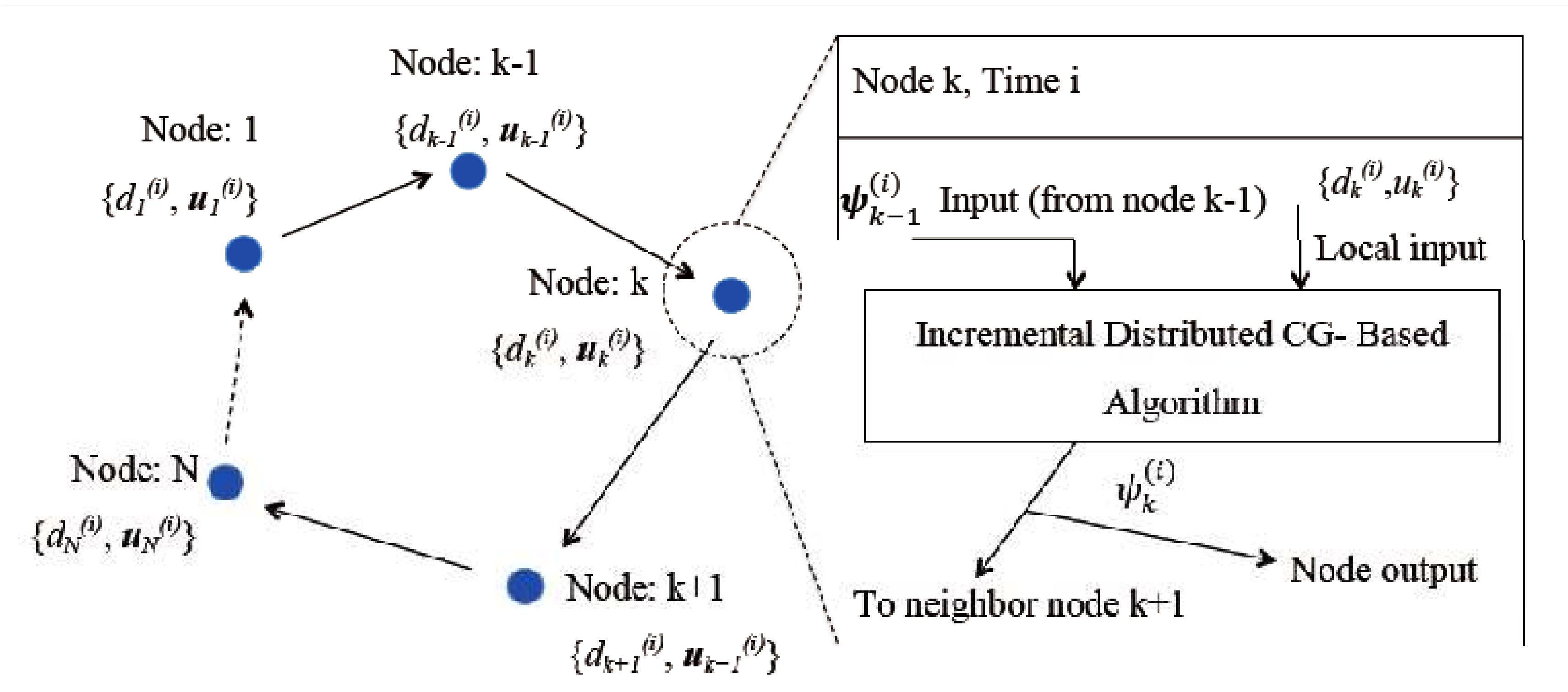
where

$$\boldsymbol{X} = [\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots \boldsymbol{x}_N], \ (N \times M)$$

$$\boldsymbol{d} = [d_1, d_2, \ldots d_N]^T, \ (N \times 1)$$

Proposed incremental distributed CG – based algorithms(1/3)

- Incremental distributed CG- based network processing:



Proposed incremental distributed CG – based algorithms(2/3)

Incremental distributed conventional CG

- It solves the following equation:

$$\boldsymbol{R}_k^{(i)}\boldsymbol{\omega}_k^{(i)} = \boldsymbol{b}_k^{(i)}$$

- In the CG- based algorithm, the iteration procedure is introduced. For the jth iteration, we choose the negative direction as:

$$\boldsymbol{g}_k^{(i)}(j) = \boldsymbol{b}_k^{(i)} - \boldsymbol{R}_k^{(i)}\boldsymbol{\omega}$$

- The CG-based weight vector is defined as:

$$\boldsymbol{\omega}_k^{(i)}(j) = \boldsymbol{\omega}_k^{(i)}(j-1) + \alpha_k^{(i)}(j)\boldsymbol{p}_k^{(i)}(j)$$

- The direction vector is defined as:

$$\boldsymbol{p}_{k+1}^{(i)} = \boldsymbol{g}_k^{(i)} + \beta_k^{(i)}\boldsymbol{p}_k^{(i)}$$

where $\beta_k^{(i)}(j)$ is calculated by the Gram Schmidt orthogonalization procedure for the conjugacy

- ’Exponentially decaying data window’ is introduced to define the correlation and cross-correlation matrices

$$\boldsymbol{R}_k^{(i)} = \lambda_f \boldsymbol{R}_{k-1}^{(i)} + \boldsymbol{x}_k^{(i)}[\boldsymbol{x}_k^{(i)}]^H$$

$$\boldsymbol{b}_k^{(i)} = \lambda_f \boldsymbol{b}_{k-1}^{(i)} + d_k^{(i)*}\boldsymbol{x}_k^{(i)}$$

Proposed incremental distributed CG – based algorithms(3/3)

Incremental distributed modified CG:

- We redefine the negative gradient vector with a recursive expression:

$$g_k^{(i)} = b_k^{(i)} - R_k^{(i)}\omega_k^{(i)} = \lambda_f g_{k-1}^{(i)} - \alpha_k^{(i)} R_k^{(i)} p_k^{(i)} + x_k^{(i)}[d_k^{(i)} - x_k^{(i)H}\omega_{k-1}^{(i)}]$$

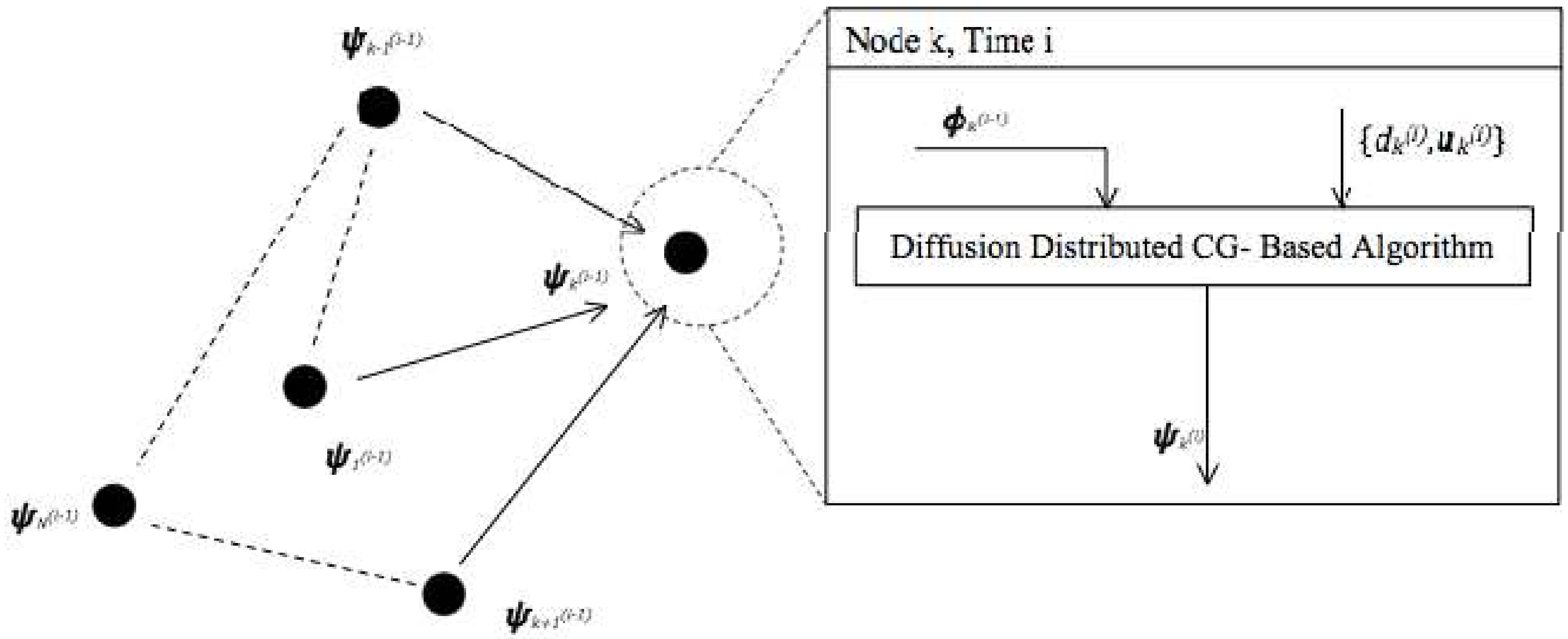
- The direction vector is defined as:

$$p_{k+1}^{(i)} = g_k^{(i)} + \beta_k^{(i)} p_k^{(i)}$$

where β_k is computed to avoid the residue produced by using the Polak-Ribiere approach

Proposed diffusion distributed CG – based algorithms(1/2)

- Diffusion distributed CG- based network processing



Proposed diffusion distributed CG – based algorithms(2/2)

- Local estimates are combined at node k as:

$$\phi_k^{(i-1)} = \sum_{l \in N_{k,i-1}} c_{kl} \psi_l^{(i-1)}$$

where C_{kl} should be satisfied:

$$\sum_l c_{kl} = 1, l \in N_{k,i-1} \forall k$$

- The combiner C is defined through the Metropolis rule:

$$\begin{cases} c_{kl} = \frac{1}{max(n_k, n_l)}, & \text{if } k \neq l \text{ are linked} \\ c_{kl} = 0, & \text{for } k \text{ and } l \text{ not linked} \\ c_{kk} = 1 - \sum_{l \in N_k / k} c_{kl}, & \text{for } k = l \end{cases}$$

- The unbiased estimates for node k are calculated as:

$$\psi_k^{(i)}(j) = \phi_k^{(i-1)}(j) + \alpha_k^{(i)}(j) p_k^{(i)}(j)$$

Analysis of the algorithms: incremental distributed CG

- The computational complexity is used to analyse the proposed incremental distributed CG algorithms

Algorithm	Additions	Multiplications
IDCCG	m^2+m $+J(m^2+6m-4)$	$2m^2+2m$ $J(m^2+7m+3)$
IDMCG	$2m^2+10m-4$	$3m^2+12m+3$
IDLMS	$4m-1$	$3m+1$
IDRLS[3]	$4m^2+12m+1$	$4m^2+12m-1$

Analysis of the algorithms: diffusion distributed CG

- The computational complexity is used to analyse the proposed diffusion distributed CG algorithms

Algorithm	Additions	Multiplications
DDCCG	$m^2 + m$ $+J(m^2 + 6m$ $+Lm - 4)$	$2m^2 + 2m$ $+J(m^2 + 7m$ $+Lm + 3)$
DDMCG	$2m^2 + 10m - 4$ $+Lm$	$3m^2 + 12m + 3$ $+Lm$
DDLMS	$4m - 1 + Lm$	$3m + 1 + Lm$
DDRLS	$4m^2 + 16m + 1 + Lm$	$4m^2 + 12m - 1 + Lm$

Simulations (1/4)

- There are 20 nodes in the network
- The number of taps of the adaptive filter is 10
- The number of repetitions is 1000
- The variance for the input signal and the noise are 1 and 0.001, respectively.
- The noise samples are modeled as complex Gaussian noise.

Simulations (2/4)

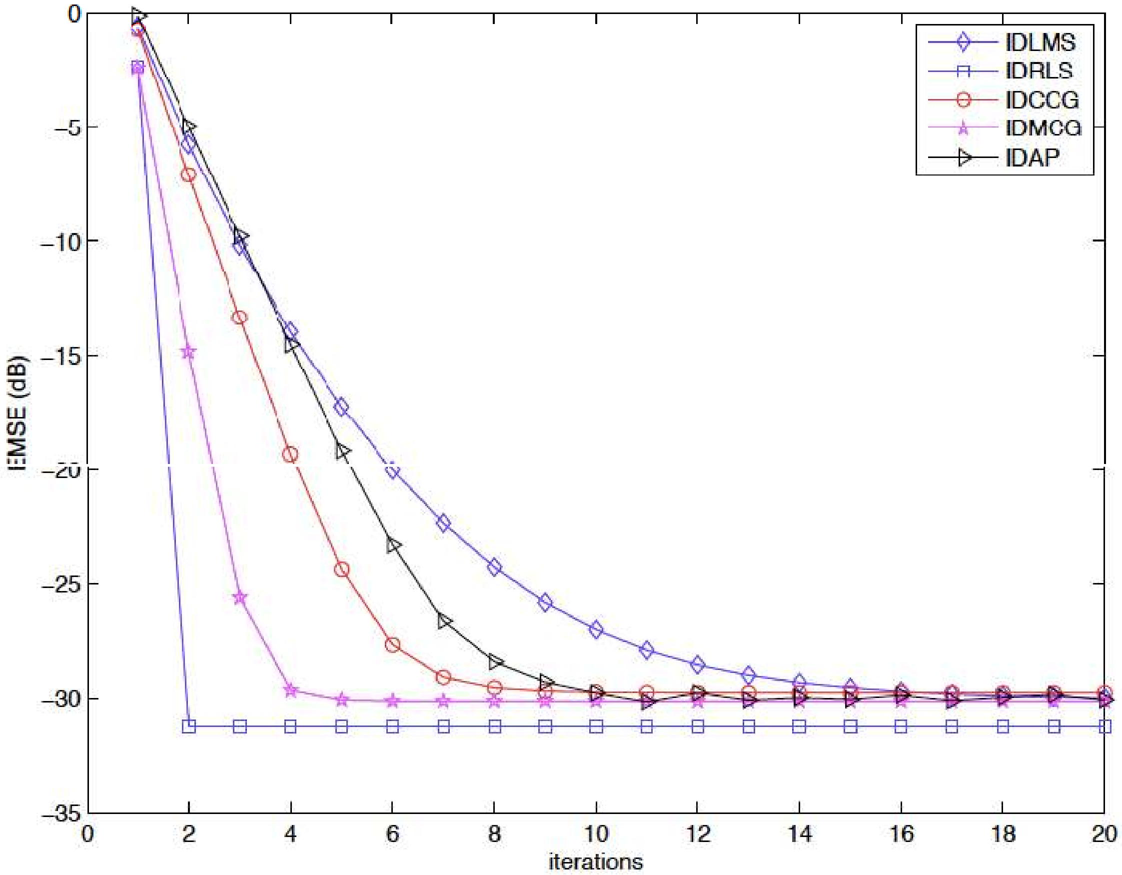


Fig. 3. Output EMSE against the number of iterations for Incremental Strategy with α_{IDLMS}=0.005, λ=0.2, $\lambda_{f-IDCCG}$=0.3,$\lambda_{f-IDMCG}$=0.25, $\eta_{f-IDCCG}$=$\eta_{f-IDMCG}$=0.15, j=5, α_{IDAP}=0.06, K=2

Simulations (3/4)

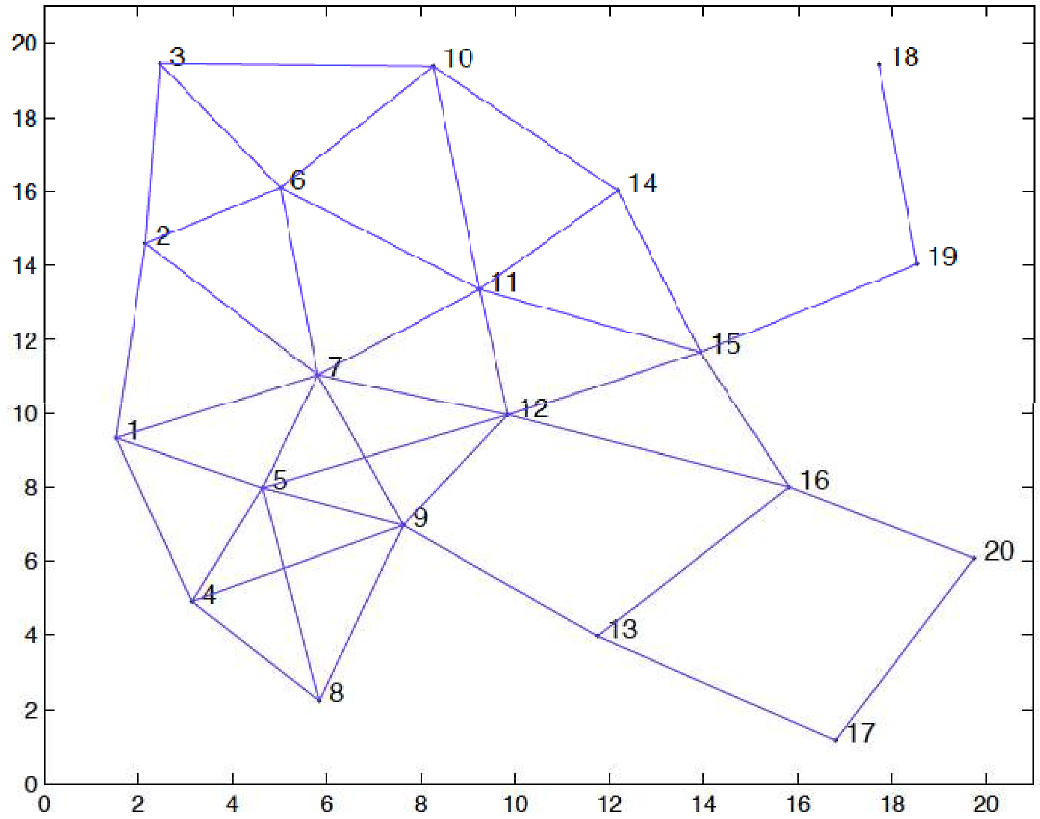


Fig. 4. Network Structure

Simulations (4/4)

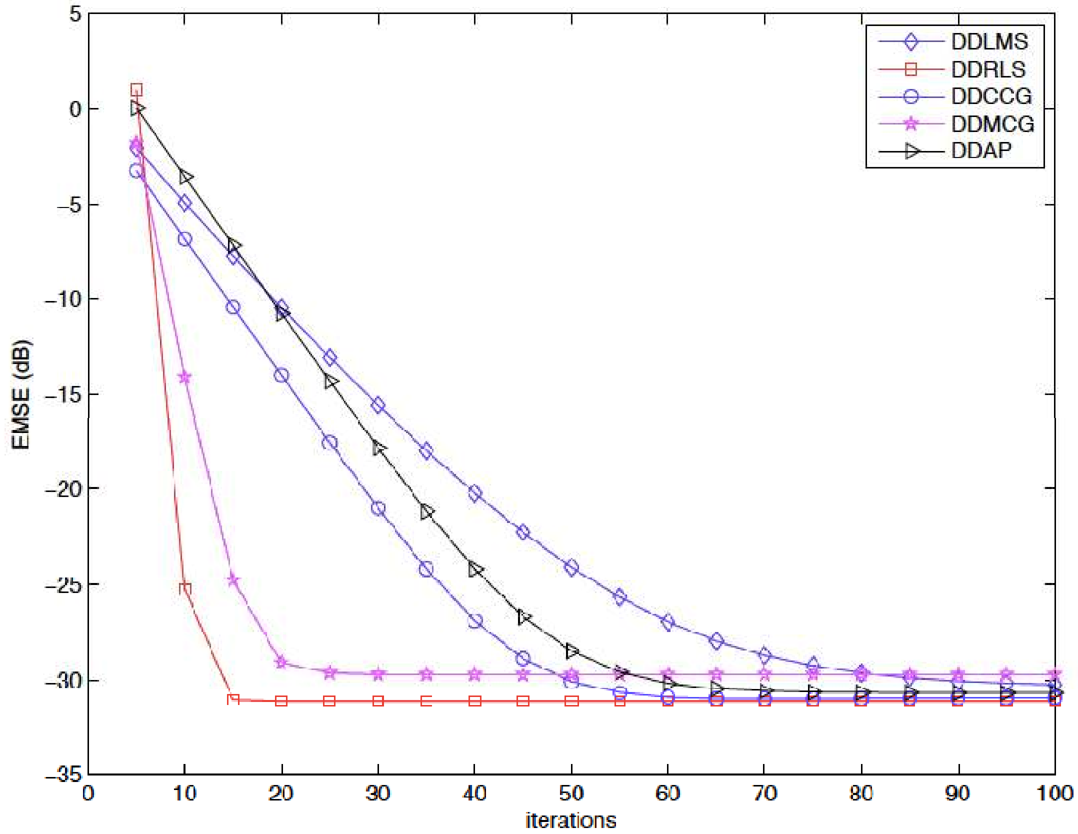


Fig. 5. Output EMSE against the number of iterations for Diffusion Strategy with α_{DDLMS}=0.0075, λ=0.998, $\lambda_{f-DDCCG}$=0.25, $\eta_{f-DDCCG}$=0.25, j=5, $\lambda_{f-DDMCG}$=0.46, $\eta_{f-DDMCG}$=0.45, α_{DDAP}=0.075, K=2

THE UNIVERSITY *of York*

Conclusions

- We have developed distributed CG algorithms for incremental and diffusion type distributed estimation over sensor networks.

- The CG- based strategies avoid the matrix inversion and numerical instability of RLS algorithms and have a faster convergence than LMS and AP algorithms.

- Simulation results have shown that the proposed IDMCG and DDMCG algorithms have a better performance than the LMS and AP algorithm, and a close performance to the RLS algorithm.

References

[1] C. G. Lopes and A. H. Sayed, ”Incremental adaptive strategies over distributed networks”, IEEE Trans. Sig. Proc., vol. 55, no. 8, pp. 4064- 4077, August 2007.

[2] C. G. Lopes and A. H. Sayed, “Diffusion least-mean squares over adaptive networks: Formulation and performance analysis”, IEEE Trans. Sig. Proc., vol. 56, no. 7, pp. 3122-3136, July 2008.

[3] L.L. Li, J.A. Chambers, C.G. Lopes,and A.H. Sayed, “Distributed Estimation Over an Adaptive Incremental Network Based on the Affine Projection Algorithm”, IEEE Trans. Sig. Proc., vol. 58, issue. 1, pp. 151-164, Jan. 2010.

[4] F. Cattivelli, C. G. Lopes, and A. H. Sayed, ”Diffusion recursive leastsquares for distributed estimation over adaptive networks,” IEEE Trans. Sig. Proc., vol. 56, no. 5, pp. 1865-1877, May 2008.

[5] G. Mateos, I. D. Schizas, and G. B. Giannakis, ”Distributed Recursive Least-Squares for Consensus-Based In-Network Adaptive Estimation,” IEEE Trans. Sig. Proc., vol. 57, no. 11, pp. 4583-4588,November 2009.

[6] O. Axelsson, Iterative Solution Methods, New York: Cambridge Univ. Press, 1994.

[7] G. H. Golub and C. F. Van Loan, Matrix Computations, 2nd Ed.. Baltimore, MD: Johns Hopkins Univ. Press, 1989.

[8] P. S. Chang and A. N. Willson, Jr, “Analysis of Conjugate Gradient Algorithms for Adaptive Filtering”, IEEE Transactions on Signal Processing, vol. 48, no. 2, pp. 409-418, Febrary 2000.

[9] L. Wang, and R.C.de Lamare , “Constrained adaptive filtering algorithms based on conjugate gradient techniques for beamforming ”, IET Signal Processing, vol. 4, issue. 6, pp. 686-697, Feb. 2010.

[10] R. Fa, R. C. de Lamare, and L. Wang, “Reduced-Rank STAP Schemes for Airborne Radar Based on Switched Joint Interpolation, Decimation and Filtering Algorithm”, IEEE Trans. Sig. Proc., vol. 58, no. 8, pp. 4182-4194, Aug. 2010.

[11] R. C. de Lamare and R. Sampaio-Neto, “Adaptive Reduced-Rank Processing Based on Joint and Iterative Interpolation, Decimation, and Filtering”, IEEE Trans. Sig. Proc., vol. 57, no. 7, pp. 2503 - 2514, Juy. 2009.