Dynamic Data Driven
Sensor Network Selection and Tracking

Ioannis D. Schizas and Vasileios Maroulas
Dept. of EE, Univ. of Texas at Arlington
Dept. of Math, Univ. of Tennessee at Knoxville
http://www.uta.edu/faculty/schizas

Acknowledgment: AFOSR9550-15-1-0103
Motivation

- Ad hoc sensor network employed for tracking multiple targets using nonlinear sensor observations.

- Only a few sensors are informative
Problem Setting

➢ Network with $m$ sensors

➢ Targets spatially scattered

➢ Sensor $j$ acquires

$$x_j(t) = \sum_{\rho=1}^{R} a_\rho(t)d_{j,\rho}^{-2}(t) + w_j(t), \quad j = 1, \ldots, m$$

$a_\rho(t)$: the intensity of a signal emitted by the $\rho$th target;

$d_{j,\rho}(t)$: the distance between the $\rho$th target and sensor $j$ at time $t$;

$R$: total number of targets, $w_j(t)$: zero-mean noise with variance $\sigma_w^2$
Target’s State Model

➢ For target $\rho$ the state (position+velocity) evolves according to:

**Constant Velocity State Model:**

$$s_{\rho}(t + 1) = As_{\rho}(t) + u_{\rho}(t),$$

A: state transition matrix

$$A = \begin{bmatrix}
1 & 0 & T & 0 \\
0 & 1 & 0 & T \\
0 & 0 & 1 & 0 \\
0 & 0 & 0 & 1
\end{bmatrix}, \Sigma_u = \sigma_u^2 \begin{bmatrix}
\frac{(T)^3}{3} & \frac{(T)^2}{2} \\
\frac{(T)^2}{2} & T
\end{bmatrix} \otimes I_2$$

$\rightarrow T$ is the sampling period.

➢ Data covariance matrix contains sparse factors that indicate which groups of sensors acquire measurements about the same target

➢ Canonical correlation analysis is enhanced here with norm-one regularization to identify target-informative sensors and control the sensing process
Canonical Correlation Analysis

➢ Given data sequences \( \{x(t), y(t)\}_{t=0}^{N-1} \in \mathbb{R}^{P \times 1} \) CCA linearly extracts common features

➢ Find \( q \times p \) matrices \( E \) and \( D \) such that

\[
(\tilde{D}, \tilde{E}) = \arg \min (N^{-1}) \sum_{t=0}^{N-1} \|Ey(t) - Dx(t)\|_2^2
\]

s.to \( D\hat{\Sigma}_x D^T = I \) and \( E\hat{\Sigma}_y E^T = I \),

➢ Sample-average covariance matrices \( \hat{\Sigma}_x \) and \( \hat{\Sigma}_y \)

➢ Uncovering common targets present in both \( x(t) \) and \( y(t) \)
CCA in Clustering Sensor Data

➢ In our setting form two sequences:

\[ x(t) = \chi(t - 1) \quad \text{'past' of sensor measurements} \]

\[ y(t) = \chi(t) \quad \text{'present+future' of sensor measurements} \]

➢ Controlling memory length is possible

➢ Common components in \( x(t) \) and \( y(t) \): State vectors \( s_\rho(t) \)

➢ CCA based clustering:
Regularized CCA Formulation

- Enhance CCA with norm-one regularization mechanisms

\[
(\hat{D}, \hat{E}) = \arg \min_{D,E} \frac{1}{T} \sum_{\tau=1}^{T} \|E(y(\tau) - \hat{m}_y) - D(x(\tau) - \hat{m}_x)\|^2_2 + \sum_{\rho=1}^{M} \lambda_{E,\rho} \|E_{\rho}\|_1 + \sum_{\rho=1}^{M} \lambda_{D,\rho} \|D_{\rho}\|_1 \\
+ \nu \|E \Sigma_y E^T - I\|_F^2 + \varepsilon \|D \Sigma_x D^T - I\|_F^2
\]

- Sample-average expectation vectors \( \hat{m}_x \) and \( \hat{m}_y \)

- \( \lambda_{D,\rho}, \lambda_{E,\rho} \) are positive sparsity-controlling coefficients

- \( \varepsilon, \nu \) are positive penalty coefficients taking care of whiteness

- Employ coordinate descent mechanisms to minimize entry-by-entry

- Consensus-averaging to obtain distributed implementation
Prior Art

- Related sparse CCA formulations [Hardoon et al.'08, Witten et al.'09, Chen et al.'12, Wiesel'08]
  - Maximize correlation between two data sets and perform variable selection
  - Not decentralized approaches

- Sensor selection and clustering

  - Data model parameters should be known/available; Find sleeping intervals [Gupta et al.'06, Krishnamurthy et al.'08, Fuemmeler et al.'10, Joshi et al.'09]
  - Linear data models and memoryless sources [Schizas'13]
Algorithmic Matters

- Sparse CCA formulation

\[
(\hat{D}, \hat{E}) = \arg \min_{D,E} \frac{1}{T} \sum_{t=1}^{T} ||E(y(t) - \hat{m}_y) - D(x(t) - \hat{m}_x)||^2_2 + \sum_{\rho=1}^{M} \lambda_{E,\rho} ||E_{\rho}||_1 + \sum_{\rho=1}^{M} \lambda_{D,\rho} ||D_{\rho}||_1 \\
+ \psi ||E \Sigma_y E^T - I||^2_F + \varepsilon ||D \Sigma_x D^T - I||^2_F
\]

- Utilize 'average-like' quantities

\[
Dx(t) = \sum_{i=1}^{p} d_i x_i(t), \quad Ey(t) = \sum_{i=1}^{p} e_i y_i(t)
\]

- \(d_j\) and \(e_j\) denote the jth column of \(D\) and \(E\) respectively; Updated at sensor \(j\)

- Replacing at coordinate cycle \(k-1\) \(D\) and \(E\) in last two black terms with \(\hat{D}^{k-1}, \hat{E}^{k-1}\)
Distributed Sparse CCA

➢ Each sensor applied $K$ ADMM iterations during coordinate cycle $k$ to find estimates

$$\{\hat{\mu}^{k}_{j,\tau} \to \hat{D}^{k-1}x(\tau)\}_{\tau=0}^{t}$$

$$\{\hat{\eta}^{k}_{j,\tau} \to \hat{E}^{k-1}y(\tau)\}_{\tau=0}^{t}$$

➢ for $k=1,2,3,...$ (coordinate cycle)

➢ Sensor $j$ forms estimates $\{\hat{\mu}^{k}_{j,\tau}\}_{\tau=0}^{t}$ and $\{\hat{\eta}^{k}_{j,\tau}\}_{\tau=0}^{t}$ via K ADMM updating recursions

➢ for $j=1,...,p$

Entry-wise update of $\hat{D}^{k}(\alpha,j), \alpha = 1,\ldots,M$

Entry-wise update of $\hat{E}^{k}(\alpha,j), \alpha = 1,\ldots,M$

end for

end for

❑ Termination when e.g., $\max_{j=1,...,p}\|\hat{d}^{k}_{j} - \hat{d}^{k-1}_{j}\|_{F} + \|\hat{e}^{k}_{j} - \hat{e}^{k-1}_{j}\|_{F} < \epsilon$

❑ Convergence to a stationary point as iterations $K$ and $k$ go to infinity

➢ Communication cost proportional to $t, |N_{j}|$ and $M$
Drift Homotopy Particle Filtering

- Particle filters need a lot of particles when tracking multiple targets in order to approximate accurately the targets’ state distribution.

- Introduce one extra step to move samples in statistically significant regions.

- A drift homotopy/relaxation algorithm.
Drift Homotopy Process

Consider the signal process: \( dX_t = a(X_t)dt + \sigma(X_t)dB_t \)

Consider an SDE system with modified drift

\[
dZ_t = b(Z_t)dt + \sigma(Z_t)dB_t,
\]

\( b(Z_t) \) is suitably chosen.

Consider a collection of \( L + 1 \) modified SDE systems

\[
dZ^\ell_t = (1 - \delta_\ell)b(Z^\ell_t)dt + \delta_\ell a(Z^\ell_t)dt + \sigma(Z^\ell_t)dB_t,
\]

\( \ell = 0, \ldots, L, \) with \( \delta_\ell < \delta_{\ell+1}, \delta_0 = 0 \) and \( \delta_L = 1. \)
Drift Homotopy Algorithm

Algorithm (M. and Stinis, J. of Comp. Phys., 2012)

Instead of sampling directly from the (proposal) density

\[ p(X_{tk} | X_{tk-1}) \]  \hspace{1cm} (10)

Sample from the density

\[ p(Z^0_{tk} | X_{tk-1}) \]

and gradually morph the sample into a sample of (10) by sampling from the \( \ell \) levels:

\[ p(Z^\ell_{tk} | X_{tk-1}) \]
Remarks

- The levels from 0 to $L - 1$ are auxiliary and only serve the purpose of providing the sampler at level $L$ with a better initial condition. The final sampling is performed at the $L$th level which corresponds to the original SDE.

- The idea behind drift homotopy resembles the main idea behind *Homotopy Methods* used in deterministic optimization problems and *Simulated Annealing* used in equilibrium statistical mechanics.
Joint Sensor Selection and Tracking

→ **Start-up stage** \((t = 0)/Reconfiguration** \((t \neq 0)\): Sensors collect \(T_s\) measurements \(x_t\), S-T association scheme is applied to determine the target-informative groups \(T_{\rho_\ell,t}\) and \(\ell = 1, \ldots, \hat{r}(t)\)**

**FOR** \(\tau = t, \ldots, \)

→ Determine the leading sensor \(C_{\rho_\ell,\tau}\) in each \(T_{\rho_\ell,\tau}\)

→ Each leading sensor \(C_{\rho_\ell,\tau}\) receives particles and weights from \(C_{\rho_\ell,\tau-1}\), and \(x_j(\tau)\) from \(j \in T_{\rho_\ell,\tau}\) to perform tracking for \(\rho_\ell = 1, \ldots, \hat{r}(t)\) target via the PF recursions and obtain \(\hat{s}_{\rho_\ell}(\tau)\)

→ A set of ’candidate’ informative sensors are selected by the leading node, e.g. \(J_{\rho_\ell,t+1}\), for target \(\rho_\ell\).

→ S-T association scheme is applied in each connected set of sensors \(J_{\rho_\ell,\tau+1}\) to obtain the target-informative sets \(T_{\rho_\ell,\tau+1}\).

→ If any target-informative set \(T_{\rho_\ell,t+1}\), is empty, do **Reconfiguration** step, otherwise continue.

**END FOR**
Tracking Multiple Targets

Total number of sensors: \( m = 120 \)

→ Targets \( \rho = 1, 2, 3 \): in time interval \([1, 15]\)s.
→ Targets \( \rho = 4, 5 \): in time interval \([17, 30]\)s.
→ Targets \( \rho = 6, 7 \): in time interval \([32, 45]\)s.
→ Targets \( \rho = 8, 9, 10 \): in time interval \([47, 60]\)s.
→ Targets \( \rho = 11, 12 \): in time interval \([62, 72]\)s.
Our novel approach achieves the smallest root mean-square error
Concluding Remarks

➢ Sensor selection clustering via CCA and norm-one regularization

➢ Determining groups of correlated data; Multiple tracking processes

➢ Improved drift homotopy algorithms for tracking

Thank You!