DISTRIBUTED DECISION-MAKING OVER MOBILE ADAPTIVE NETWORKS

Sahar Khawatmi, Xinxin Huang, Abdelhak M. Zoubir

Signal Processing Group
Technische Universität Darmstadt
Merckstr. 25, 64283 Darmstadt, Germany
{khawatmi,xhuang,zoubir}@spg.tu-darmstadt.de

ABSTRACT

In this paper, we study distributed decision-making over mobile adaptive networks where nodes in the network collect data generated by two different models. The nodes need to decide which model to estimate and track. However, they do not know beforehand which model they observe. Therefore, an effective clustering technique is needed. We apply a clustering technique that reduces the clustering error. Furthermore, introduce an additional term to the motion model to ensure that the nodes move coherently without fragmentation in the network during the decision-making process. Once the network reaches agreement on the desired model, the cooperation among nodes enhances the performance of the estimation task by relaying data throughout the network.

Index Terms— Decentralized processing, multi-task networks, learning, adaptive networks, decision-making.

1. INTRODUCTION AND RELATED WORK

Inspired by distinctive biological phenomena, several algorithms are designed to mimic the behavior of animal groups that move together in an amazing coherence, such as bee swarms, birds flying in formation, and schools of fish [1–9]. Diffusion strategies can be used to solve estimation tasks in cooperative networks, which consist of a collection of nodes with adaptation and learning abilities [10, 11]. In some situations, the nodes in the network need to decide between multiple options, for example, to track only one of two food sources [12]. In the presence of multiple targets another situation is considered in [13] where the nodes switch the target they are tracking and form distinct clusters. The resulting clusters split up while moving and pursue their distinct target over time.

We consider a distributed mean-square-error estimation problem over an N-node network. The connectivity of the nodes is described by a graph (see Figure 1). Data sensed by any particular node can arise from one of two different models. The objective is to reach an agreement among all nodes in the network on one model to estimate and track. Two definitions are introduced: the observed model, which refers to the one, from which a node collects data, and the desired model, which refers to the one the node decides to move towards. The nodes do not know which model generated the data they collect; they also do not know which other nodes in their neighborhood sense data arising from the same model. Therefore, each node needs to determine the subset of its neighbors that observes the same model.

The proposed classification scheme in [12], which determines the subset neighbors that observes the same model, has some demerits. The performance of this scheme depends on the initial location of the network and the location of the models. Since the decision-making objective depends on the classification output, errors made in the classification process have an impact on the global decision. Several clustering algorithms have been proposed in [14–16]. A fast clustering technique that lets the nodes distinguish the neighbors in real-time is needed in mobile networks because the topology changes quickly due to the movement of the nodes. In this paper, we replace the proposed classification scheme in [12] by the clustering algorithm in [16] to ensure fast and accurate clustering. Changes in topology over time imply that the network may be separated into two groups before reaching the agreement on one model. Now, while groups are moving far away from each other towards their different desired models, they will lose the connections between each other. This means that the decision-making process fails to ensure that the network converges to only one desired model. We add a new term to the velocity control, this term helps to keep the network moving in a cohesive manner, even if nodes move and do not make a decision yet.

The paper is organized as follows: the network and data model are described in Section 2. We illustrate the decision-making algorithm and the motion mechanism technique in Sections 3 and 4, respectively. Simulation results and discussion are presented in Section 5.

2. NETWORK AND DATA MODEL

Consider a collection of N nodes distributed in space. Figure 1 shows the network structure where nodes with the same color observe the same model. The unknown models are denoted by $\{z_1^\circ, z_2^\circ\}$ each of size $M \times 1$. We denote the set of neighbors of node k by $\mathcal{N}_{k,i}$ of size $n_{k,i}$ (i.e, the number of neighbors of node k). While the set of neighbors of node k excluding k itself is denoted by $\mathcal{N}_{k,i}^{-}$. We represent the network topology at time instant i by means of the $N \times N$ adjacency matrix E_i whose entries $e_{\ell k}(i)$ are defined as follows:

$$e_{\ell k}(i) = \begin{cases} 1, & \ell \in \mathcal{N}_{k,i} \\ 0, & \text{otherwise} \end{cases} \tag{1}$$

We consider an $N \times N$ combination matrix A_i where its $(\ell, k)^{\text{th}}$ entry contains the combination weight $a_{\ell k}(i)$ reflects the weight that node k assigns to data received from node ℓ. The entries of the combination matrix A_i are non-negative real-valued, satisfying

$$a_{\ell k}(i) = 0 \text{ for } \ell \notin \mathcal{N}_{k,i}, \quad \sum_{\ell=1}^{N} a_{\ell k}(i) = 1. \tag{2}$$

Furthermore, we define the nodes observed model vector by

$$w^\circ \triangleq \text{col}\{w_1^\circ, w_2^\circ, \cdots, w_N^\circ\}, \quad w^\circ \in \mathbb{R}^{MN \times 1}. \tag{3}$$

Figure 1 shows that node k collects data from model z_1°, i.e., $w_k^\circ = z_1^\circ$, while node ℓ collects data from model z_2° which implies $w_\ell^\circ = z_2^\circ$. We denote the estimate vector of the desired model at time instant i of node k by $w_{k,i}$. We define $w_i \triangleq \text{col}\{w_{1,i}, w_{2,i}, \cdots, w_{N,i}\}$. The objective of the network is to have all $w_{k,i}$ converge to only one model, either $\{z_1^\circ\}$ or $\{z_2^\circ\}$. We can write that for each node $k \in \{1, 2, \cdots, N\}$

$$w_{k,i} \to z_j^\circ \quad \text{as} \quad i \to \infty \tag{4}$$

where j is either 1 or 2. The nodes seek to estimate the vector parameter z_j°, which leads to the fact that nodes with $w_k^\circ = z_j^\circ$ track their own observed model, but others with $w_k^\circ \neq z_j^\circ$ do not track their own observed model, but track z_j° instead, although they do not have any streaming data from z_j°. Then, the aggregate cost function $J^{\text{glob}}(w)$ is defined as:

$$J^{\text{glob}}(w) = \sum_{k=1}^{N} ||w_{k,i} - z_j^\circ||^2. \tag{5}$$

The location and velocity vectors of node k at time instant i are denoted by $x_{k,i}$ and $v_{k,i}$, respectively. The modified diffusion strategy in [12] is given by the following steps:

$$\psi_{k,i} = w_{k,i-1} + \mu(q_{k,i} - w_{k,i-1}) \tag{6}$$

$$w_{k,i} = \sum_{\ell \in \mathcal{N}_{k,i}} (\dot{a}_{\ell k}(i)\psi_{\ell,i} + \ddot{a}_{\ell k}(i) w_{\ell,i-1}) \tag{7}$$

where μ is a positive step-size parameter and $q_{k,i}$ is the noisy location of the model that node k observes and is given by,

$$q_{k,i} = w_k^\circ + \eta_{k,i} \tag{8}$$

where $\eta_{k,i}$ is a zero-mean white random process with variance $\sigma_k^2(i) = \kappa ||w_k^\circ - x_{k,i}||^2$, for $\kappa > 0$. The combination coefficients $\dot{a}_{\ell k}(i)$ and $\ddot{a}_{\ell k}(i)$ are two sets of non-negative entries in the combination matrices $\dot{A}_i$ and $\ddot{A}_i$ which satisfy:

$$\dot{A}_i + \ddot{A}_i = A_i. \tag{9}$$

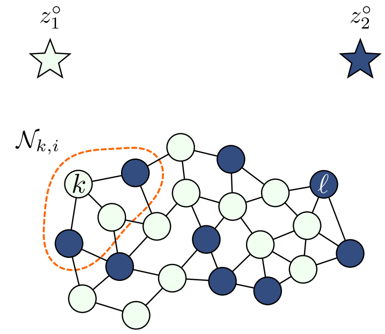

Fig. 1: Illustration of a network model with two observed models represented by two colors.

The design method of the matrices $\dot{A}_i$ and $\ddot{A}_i$ is illustrated in Section 3. We define the central location and velocity of the network at time instant i by averaging location and velocity of all nodes over the network, respectively, as

$$x_i^\circ \triangleq \frac{1}{N}\sum_{k=1}^{N} x_{k,i}, \quad v_i^\circ \triangleq \frac{1}{N}\sum_{k=1}^{N} v_{k,i}. \tag{10}$$

The central network location x_i° and velocity v_i° are estimated using the diffusion strategy in a distributed manner by considering the following global cost functions:

$$J^x(x^g) = \sum_{k=1}^{N} ||x_{k,i}^g - x_i^\circ||^2, \quad J^v(v^g) = \sum_{k=1}^{N} ||v_{k,i}^g - v_i^\circ||^2. \tag{11}$$

Applying the diffusion strategy structure to estimate $x_{k,i}^g$ and $v_{k,i}^g$, respectively, we obtain

$$\theta_{k,i} = x_{k,i-1}^g + \mu(x_{k,i} - x_{k,i-1}^g) \tag{12}$$

$$x_{k,i}^g = \sum_{\ell \in \mathcal{N}_{k,i}} a_{\ell k}(i)\theta_{\ell,i} \tag{13}$$

$$\phi_{k,i} = v_{k,i-1}^g + \mu(v_{k,i} - v_{k,i-1}^g) \tag{14}$$

$$v_{k,i}^g = \sum_{\ell \in \mathcal{N}_{k,i}} a_{\ell k}(i)\phi_{\ell,i}. \tag{15}$$

Note that we estimate the central network location and velocity using matrix A_i rather than $\dot{A}_i$ or $\ddot{A}_i$ due to the fact that the nodes are required to share velocity and location information with all neighbors, regardless of their observed models.

Let the true clustering matrix, which gives information about the observed model and is not known beforehand, be denoted by F_i°. The assignment of the $(\ell, k)^{\text{th}}$ entry to one means

$$f_{\ell k}^\circ(i) = 1 \Rightarrow \{\ell \in \mathcal{N}_{k,i} \text{ and } w_k^\circ = w_\ell^\circ\}. \tag{16}$$

We apply the clustering technique proposed in [16] to create the estimated clustering matrix F_i of size $N \times N$. Each node

k runs the following steps for the clustering process:

$$\psi^c_{k,i} = \psi^c_{k,i-1} + \mu_k(q_{k,i} - \psi^c_{k,i-1}) \tag{17}$$

$$w^c_{k,i} = \sum_{\ell \in \mathcal{N}'_{k,i}} a'_{\ell k}(i)\psi^c_{\ell,i} \tag{18}$$

$\psi^c_{k,i}$ and $w^c_{k,i}$ both converge to the observed model w°_k without being affected by the decision process. Initializing $\psi^c_{k,-1} = 0$ and $B_{-1} = S_{-1} = F_{-1} = I_N$, where I denotes the identity matrix of appropriate size. The combination matrix A'_i is designed using the following steps [16]:

$$b_{\ell k}(i) = \begin{cases} 1, \text{ if } ||\psi^c_{\ell,i} - w^c_{k,i-1}|| \leq \epsilon \\ 0, \text{ otherwise} \end{cases} \tag{19}$$

$$s_{\ell k}(i) = \xi \times s_{\ell k}(i-1) + (1-\xi) \times b_{\ell k}(i) \tag{20}$$

$$f_{\ell k}(i) = \lfloor s_{\ell k}(i) \rceil \tag{21}$$

for $\epsilon > 0$, $0 \leq \xi \leq 1$, and $0 < \zeta \leq 1$. The notation $\lfloor \cdot \rceil$ denotes rounding to the nearest integer. The combination coefficients $a'_{\ell k}(i)$ satisfy:

$$a'_{\ell k}(i) = 0, \text{ if } \ell \notin \mathcal{N}'_{k,i}, \quad \sum_{\ell=1}^{N} a'_{\ell k}(i) = 1 \tag{22}$$

where $\mathcal{N}'_{k,i}$ consists of neighbors believing that they belong to the same cluster, i.e., $f_{\ell k}(i) = 1$ implies $\ell \in \mathcal{N}'_{k,i}$.

3. SELECTION OF COMBINATION MATRICES

Applying the strategy in [12], we use matrix G_i of size $N \times N$ to estimate the desired model of each node. Node k assigns the value of the $(\ell, k)^{\text{th}}$ entry for each node $\ell \in \mathcal{N}^-_{k,i}$ using (24). The meaning of the value is as follows:

$$\begin{cases} g_{\ell k}(i) = 1 : w_{\ell,i} \to w^\circ_k \\ g_{\ell k}(i) = 0 : w_{\ell,i} \not\to w^\circ_k \end{cases} \tag{23}$$

Since node k has access to the desired models of its neighbors $g_{\ell\ell}(i)$, it adjusts the desired model of node ℓ from its perspective $g_{\ell k}(i)$ according to the following rule:

$$g_{\ell k}(i) = \begin{cases} g_{\ell\ell}(i-1), & \text{if } f_{\ell k}(i) = 1 \\ 1 - g_{\ell\ell}(i-1), & \text{otherwise.} \end{cases} \tag{24}$$

Each diagonal entry $g_{kk}(i)$ indicates whether node k wishes to track its own observed model or not,

$$\begin{cases} g_{kk}(i) = 1 : w_{k,i} \to w^\circ_k \\ g_{kk}(i) = 0 : w_{k,i} \not\to w^\circ_k, \end{cases} \tag{25}$$

where node k updates its desired model $g_{kk}(i)$ according to:

$$g_{kk}(i) = \begin{cases} g_{kk}(i-1), & \text{with probability } p_k(i) \\ 1 - g_{kk}(i-1), & \text{with probability } 1 - p_k(i) \end{cases} \tag{26}$$

and $p_k(i)$ is given by,

$$p_k(i) = \frac{[n^g_k(i)]^K}{[n^g_k(i)]^K + [n_k(i) - n^g_k(i)]^K} \tag{27}$$

for a positive constant K, with $n^g_k(i)$ being the size of the set $\mathcal{N}^g_{k,i}$ that contains the subset of nodes that are in the neighborhood of node k and have the same desired model as node k at time instant $i-1$. Herein, $\mathcal{N}^g_{k,i}$ is constructed as follows:

$$\mathcal{N}^g_{k,i} = \{\ell | \ell \in \mathcal{N}_{k,i},\ g_{\ell k}(i) = g_{kk}(i-1)\}. \tag{28}$$

The entries of $\dot{A}_i$ and $\ddot{A}_i$ are set according to the following rules:

$$\dot{a}_{\ell k}(i) = \begin{cases} a_{\ell k}(i), \text{ if } \ell \in \mathcal{N}_{k,i} \text{ and } f_{\ell k}(i) = g_{kk}(i) \\ 0, \text{ otherwise} \end{cases} \tag{29}$$

$$\ddot{a}_{\ell k}(i) = \begin{cases} a_{\ell k}(i), \text{ if } \ell \in \mathcal{N}_{k,i} \text{ and } f_{\ell k}(i) \neq g_{kk}(i) \\ 0, \text{ otherwise} \end{cases} \tag{30}$$

As a result, in Eq. (7) node k combines $\{\psi_{\ell,i}\}$ if it wishes to estimate w°_ℓ, otherwise it combines $\{w_{\ell,i-1}\}$ instead, where $\ell \in \mathcal{N}_{k,i}$.

4. MOTION MODEL

Every node k updates its location vector according to the rule

$$x_{k,i+1} = x_{k,i} + \triangle t \cdot v_{k,i+1} \tag{31}$$

where $\triangle t$ is a positive time step and $v_{k,i+1}$ is the updated velocity vector of node k. Several factors influence the determination of the velocity $v_{k,i+1}$ of node k, such as (i) the desire to move towards the desired model z°_j, (ii) the desire to move in coordination with other nodes, and (iii) the desire to avoid collision. The velocity vector $v^a_{k,i+1}$, which allows node k to move towards the desired model, is given by

$$v^a_{k,i+1} = \begin{cases} w_{k,i} - x_{k,i}\ , & \text{if } ||w_{k,i} - x_{k,i}|| \leq \delta \\ \delta \cdot \dfrac{w_{k,i} - x_{k,i}}{||w_{k,i} - x_{k,i}||}\ , & \text{otherwise,} \end{cases} \tag{32}$$

where δ is a positive scaling factor used to bound the node speed. To move in a harmonious manner, the velocity vector $v^b_{k,i+1}$ of node k is updated as, $v^b_{k,i+1} = v^g_{k,i}$. Nodes should keep a safe distance r from their neighbors to avoid collision during the movement. The velocity vector $v^c_{k,i+1}$ of node k is given by

$$v^c_{k,i+1} = \frac{1}{n_k(i) - 1} \sum_{\ell \in \mathcal{N}^-_{k,i}} (1 - \frac{r}{||x_{\ell,i} - x_{k,i}||})(x_{\ell,i} - x_{k,i}). \tag{33}$$

Before reaching the agreement on one desired model, the network might become separated into two groups where each group moves towards its desired model. If these two groups move away from each other and lose the connections, the decision-making process fails. To resolve this issue, we define an $N \times 1$ vector ι_i, each entry $\iota_k(i)$ is given by,

$$\iota_k(i) = \begin{cases} 1, \text{ if } p_k(i) > 0.5 \text{ AND } n_k^g(i) < n_k(i) \\ 0, \text{ otherwise.} \end{cases} \tag{34}$$

The condition in Eq. (34) implies that $\mathcal{N}_{k,i}$ does not agree yet on one desired model and $p_k(i) > 0.5$ (i.e., node k will keep its previous decision). If $\iota_k(i) = 1$, an additional action has to be taken by node k, where $\iota_k(i)$ controls the term $v_{k,i+1}^d$ that enforces node k to move towards the center of the network in order to keep the network cohesive. Herein, $v_{k,i+1}^d$ is given by,

$$v_{k,i+1}^d = \frac{x_{k,i}^g - x_{k,i}}{||x_{k,i}^g - x_{k,i}||}. \tag{35}$$

Finally, for the non-negative weighting factors λ and β satisfying: $\lambda + \beta = 1$, each node adjusts its velocity according to the following rule:

$$\begin{aligned} v_{k,i+1} =& \big[1 - \iota_k(i)\big] \cdot \big[\lambda \cdot v_{k,i+1}^a + \beta v_{k,i+1}^b\big] \\ &+ \iota_k(i) v_{k,i+1}^d + v_{k,i+1}^c. \end{aligned} \tag{36}$$

5. SIMULATION RESULTS AND DISCUSSION

We consider a fully connected network with 40 randomly distributed nodes. The maximum number of neighbors of node k is $n_{k,i} = 7$, as long as they are within radius $R = 15$. Nodes observe data originating from two different models: $z_1^\circ = [-10; 10]$ and $z_2^\circ = [10; 10]$. The assignment of nodes to the models is random. We use a uniform combination policy to generate the coefficients $a_{\ell k}(i)$ and $a'_{\ell k}(i)$. The clustering parameters are set as follows: $\{\epsilon, \xi\} = \{5, 0.6\}$. The velocity parameters are set as follows: $\{\lambda, \beta, r, \Delta t, \delta\} = \{0.2, 0.8, 3, 0.1, 1\}$ and $\{\mu, \kappa, K\} = \{0.05, 0.02, 20\}$. The simulation results are obtained by averaging over 1000 independent experiments with different setup of the initial network location. Table 1 displays the success rate R_r of the decision-making to agree on one model, and the average required time to achieve this agreement T_r. Obviously, the proposed strategy provides better performance with almost 100% success rate of decision-making in mobile networks while needing less iterations to achieve this agreement.

Table 1 Decision-making success rate R_r and the average required time to achieve the agreement T_r .

	T_r(sec)	R_r(%)
Strategy [12]	72	64.2%
Proposed strategy	58	99.3%

Figure 2 depicts the maneuver of fish schools with two food sources over time where the nodes agree on the model z_1°. The transient network mean-square deviation (MSD) at each time instant i is defined by, $\text{MSD}_d(i) \triangleq \frac{1}{N}\sum_{k=1}^N ||z_d^\circ - w_{k,i}||^2$, where z_d° is the desired model. By substituting the undesired model $z_{\bar{d}}^\circ$ in the previous equation, we obtain the second curve $\text{MSD}_{\bar{d}}$. Figure 3(a) represents both curves and shows how the network converges to z_d°. Figure 3(b) shows the transient network mean-square error of estimating the central velocity v_i° which is given by: $\text{MSE}_v(i) \triangleq \frac{1}{N}\sum_{k=1}^N ||v_i^\circ - v_{k,i}^g||^2$. The learning curve indicates that the network moves coherently.

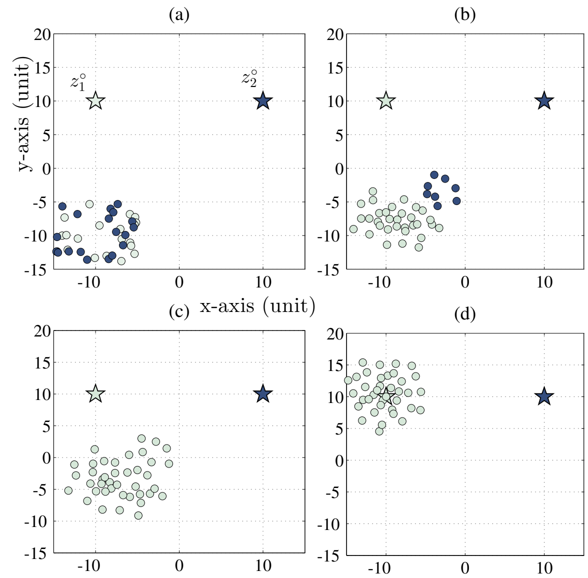

Fig. 2: Maneuver of fish schools with two food sources over time (a) $i = 10$, (b) $i = 30$, (c) $i = 100$, and (d) $i = 500$. The length unit is the body length of the node.

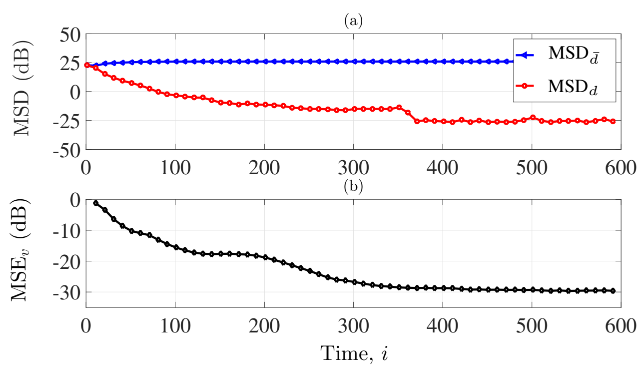

Fig. 3: The transient network mean-square deviation MSD (a). The transient network mean-square error MSE_v (b).

We studied distributed decision-making over mobile adaptive networks. We have shown that our proposed clustering technique reduces the error that affects the decision-making process. We have added a new term to the motion equation to ensure that the nodes move coherently without fragmentation. Simulation results show that the proposed strategy is insensitive to the initial network location and has high success rate of agreement on one model.

6. REFERENCES

[1] F. S. Cattivelli and A. H. Sayed, "Modeling bird flight formations using diffusion adaptation," *IEEE Trans. Signal Process.*, vol. 59, no. 5, pp. 2038–2051, May 2011.

[2] S. Y. Tu and A. H. Sayed, "Tracking behavior of mobile adaptive networks," in *Proc. Asilomar Conferences on Signals, Systems and Computers*, Pacific Grove, CA, Nov. 2010, pp. 698–702.

[3] S. Y. Tu and A. H. Sayed, "Mobile adaptive networks," *IEEE J. Sel. Topics Signal Process.*, vol. 5, no. 4, pp. 649–664, Aug. 2011.

[4] J. Chen, X. Zhao, and A. H. Sayed, "Bacterial motility via diffusion adaptation," in *Proc. Asilomar Conference on Signals, Systems and Computers*, 2010, pp. 1930–1934.

[5] A. Avitabile, R. A. Morse, and R. Boch, "Swarming honey bees guided by pheromones," *Ann. Entonomol. Soc. Am.*, vol. 68, pp. 1079–1082, 1975.

[6] H. Berg, "Motile behavior of bacteria," *Physics Today*, vol. 53, no. 1, pp. 24–29, 2000.

[7] A. Rastegarnia, A. Khalili, and M. K. Islam, "A self-organizing diffusion mobile adaptive network for pursuing a target," *American Journal of Signal Processing*, vol. 6, no. 2, pp. 25–31, 2016.

[8] F. Dressler and O. B. Akan, "Bio-inspired networking: from theory to practice," *IEEE Communications Magazine*, pp. 176–183, Nov. 2010.

[9] S. Camazine, J. L. Deneubourg, N. R. Franks, G. Theraulaz J. Sneyd, and E. Bonabeau, *Self-Organization in Biological Systems*, Princeton, NJ: Princeton Univ. Press, 2003.

[10] A. H. Sayed, S.Y. Tu, J. Chen, X. Zhao, and Z. Towfic, "Diffusion strategies for adaptation and learning over networks," *IEEE Signal Process. Mag.*, vol. 30, pp. 155–171, May 2013.

[11] A. H. Sayed, "Adaptation, learning, and optimization over networks," *Found. Trends in Mach. Learn.*, vol. 7, no. 4–5, pp. 311–801, Jul. 2014.

[12] S. Y. Tu and A. H. Sayed, "Distributed decision-making over adaptive networks," *IEEE Trans. Signal Process.*, vol. 62, no. 5, pp. 1054–1069, Mar. 2014.

[13] M. Z. Lin, M. N. Murthi, and K. Premaratne, "Mobile adaptive networks for pursuing multiple targets," in *Proc. IEEE International Conference on Acoust., Speech, and Signal Process. (ICASSP)*, South Brisbane, QLD, Apr. 2015, pp. 3217 – 3221.

[14] X. Zhao and A. H. Sayed, "Distributed clustering and learning over networks," *IEEE Trans. Signal Process.*, vol. 63, no. 13, pp. 3285–3300, Jul. 2015.

[15] J. Chen, C. Richard, and A. H. Sayed, "Adaptive clustering for multitask diffusion networks," in *Proc. 23rd European Signal Processing Conference (EUSIPCO)*, Nice, France, Sep. 2015, pp. 200–204.

[16] S. Khawatmi, A. M. Zoubir, and A. H. Sayed, "Decentralized clustering over adaptive networks," in *Proc. 23rd European Signal Processing Conference (EUSIPCO)*, Nice, France, Sep. 2015, pp. 2745–2749.