35th Annual International Conference of the IEEE EMBS
Osaka, Japan, 3 - 7 July, 2013

A Bio-Inspired Cooperative Algorithm for Distributed Source Localization with Mobile Nodes

Azam Khalili[1], Amir Rastegarnia[1], Md Kafiul Islam[2] and Zhi Yang[2]

Abstract— **In this paper we propose an algorithm for distributed optimization in mobile nodes. Compared with many published works, an important consideration here is that the nodes do not know the cost function beforehand. Instead of decision-making based on linear combination of the neighbor estimates, the proposed algorithm relies on information-rich nodes that are iteratively identified. To quickly find these nodes, the algorithm adopts a larger step size during the initial iterations. The proposed algorithm can be used in many different applications, such as distributed odor source localization and mobile robots. Comparative simulation results are presented to support the proposed algorithm.**

I. INTRODUCTION

Consider a set of N nodes as $\mathcal{N} = \{1, 2, \cdots, N\}$, where the objective of each node is to estimate the $M \times 1$ vector $\boldsymbol{w}^o$ that maximizes a cost function $J(\boldsymbol{w})$. The cost function in many estimation criteria possess in an important form as a sum of N local functions $\{f_k(\boldsymbol{w}\}_{k=1}^N$

$$\arg\min_{\boldsymbol{w}} J(\boldsymbol{w}) = \sum_{k=1}^{N} f_k(\boldsymbol{w}) \tag{1}$$

where $f_k(\boldsymbol{w})$ only depends on data available at node k. This optimization problem arises in a variety of applications, ranging from sensor networks to precision agriculture, environment monitoring, disaster relief management, smart spaces, target localization, as well as medical applications [1], [2]. Although each node can solve the optimization problem via a non-cooperative algorithm [3], in many applications, it is more desirable to have a fully decentralized solution where the statistical information for the underlying processes of interest is not available. This motivates the development of distributed adaptive estimation schemes (also known as adaptive networks [4]). More details on adaptive networks including different implementation schemes and performance comparisons can be found in [5-10]

To implement adaptive networks, one has to compute the gradient vector of cost function and make instantaneous approximations for it. Reported adaptive network implementations rely on a basic assumption that the form of the cost function, $J(\boldsymbol{w})$, is known beforehand by all nodes in the network. Consequently, its gradient vector and instantaneous approximations can be computed. However, there are applications that the nodes do not know the form of the cost function beforehand [11-16]. In such applications, nodes can only sense variations in the values of the objective function as they diffuse through the space. An example is foraging model for bacteria where the bacterial foraging for food by means of moving towards the direction of increasing nutrients in response to chemical signaling [11, 12]. In this scenario, one can interpret the cost function $J(\boldsymbol{w})$ as the concentration of nutrients and local estimate $\boldsymbol{w}_{k,i} \in \mathbb{R}^2$ as the position in the plane of the kth bacterium at time i. Thus, the nodes (here the bacteria), must sense the cost function at their locations and move toward the location of the peak of the concentration. This example reveals that it is necessary to find a distributed solution to (1) when the cost function is not available and nodes are required to converge to the peak of the objective function through an adaptive diffusive process

In [11] a diffusion and cooperation model has been introduced to understand the role of collaboration in bacteria foraging. The model addresses four factors: motion, diffusion, observation, and decision. In [12] an iterative algorithm for optimization over networks with mobile nodes is proposed. It is assumed in [12] that the nodes have limited abilities and they are allowed to cooperate with their neighbors to optimize a common objective function. Moreover, in developing the algorithm, it is assumed that the nodes do not know the form of the cost function beforehand. However, the algorithm in [12] uses a fixed step-size iterative algorithm and picks the search vector as a linear combination of the neighbors' last steps. In this paper we consider the problem of form (1) with an important assumption that the nodes do not know the form of the cost function beforehand. We interpret the successive $\boldsymbol{w}_{k,i} \in \mathbb{R}^2$ as location vectors and propose an algorithm for adaptation over networks with mobile nodes. We show that, the performance of algorithm in [12] can be improved by employing the following ideas

- Using the data related to information-rich node[1] instead of linear combination of the neighbors' estimates. More precisely, when the measured signals are very noisy, a linear combination of adjacent nodes does not provide a good estimate to calculate gradient vector, which is necessary in developing optimization algorithms with unknown objective function.
- Using variable step-size in the iterative optimization algorithm. We choose to use larger step sizes in the initial iterations to increase the probability of finding the information-rich nodes. This scheme not only improves the convergence rate, but also improves the cost function

[1]A. Khaliil and A. Rastegarnia are with Department of Electrical Engineering, Malayer Uuniversity, Malayer, 65719-95863, Iran, Email: {a.khalili,a_rastegar}@ieee.org

[2]Kafiul Islam and Zhi Yang are with the Department of Electrical and Computer Engineering, National University of Singapore, Singapore. Email: {kafiul_islam, eleyangz }@nus.edu.sg

[1]The information-rich node is defined in the Section III.

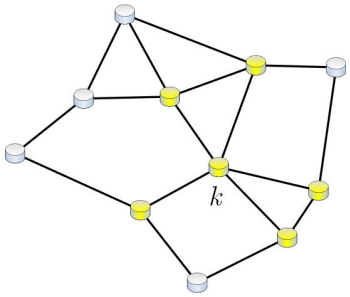

Fig. 1. A network with N nodes: the neighborhood of node k are distinguished.

averaged per node defined as (in this paper, the average locations of nodes)

$$\eta_{av} = \frac{1}{N}\sum_{k=1}^{N} J(\boldsymbol{w}_{k,i}) \tag{2}$$

The proposed algorithm can be applied to different applications, such as distributed odor source localization [16], cooperative prey herding modeling [17] and mobile robots [18, 19].

The rest of this paper is organized as follows. Section II gives the problem description. Section III focuses on formulation and analysis. In section IV, comparative simulation results are presented. Section V gives concluding remarks.

II. PROBLEM STATEMENT

Consider a network with N nodes as shown in Fig. 1. We denote $\mathcal{N}_k(i)$ as the neighborhood of node k at time i which is defined as the set of nodes that are connected to node k at time i including itself. Moreover, define $x_k(i-1)$ as the noisy measurement of the cost function at node k and time index $i-1$ which is given by[2]

$$x_k(i-1) = J(\boldsymbol{w}_{k,i-1}) + v_k(i-1) \tag{3}$$

where v_k is measurement noise with variance σ_v^2. Assume that at time i, each node k has access to noisy measurements of the cost function at times $i-1$ and $i-2$. Using these measurements, the local error signal is

$$z_k(i) = x_k(i-1) - x_k(i-2) \tag{4}$$

On the other hand, using a first-order Taylor series expansion, we have

$$J(\boldsymbol{w}_{k,i-1}) \approx J(\boldsymbol{w}_{k,i-2}) + [\nabla J(\boldsymbol{w}_{k,i-2})]^T \boldsymbol{u}_{k,i} \tag{5}$$

where $\boldsymbol{u}_{k,i}$ is defined as

$$\boldsymbol{u}_{k,i} \triangleq \boldsymbol{w}_{k,i-1} - \boldsymbol{w}_{k,i-2} \tag{6}$$

It must be noted that the vector $\boldsymbol{u}_{k,i}$ denotes the direction of motion from $\boldsymbol{w}_{k,i-1}$ to $\boldsymbol{w}_{k,i-2}$. According to (3),(4) and (5), we can relate the gradient vector to local error signal via

$$e_k(i) \approx [\nabla J(\boldsymbol{w}_{k,i-2})]^T (\boldsymbol{w}_{k,i-1} - \boldsymbol{w}_{k,i-2}) \tag{7}$$

[2]Throughout the paper, we use boldface letters for vectors and small letter for scalars.

where

$$e_k(i) = z_k(i) + v'_k(i) \tag{8}$$

where $v'_k(i) = v_k(i-1) - v_k(i-2)$. When the gradient vector at $\boldsymbol{w}_{k,i-1}$ is available, we can use it to update from $\boldsymbol{w}_{k,i-1}$ to $\boldsymbol{w}_{k,i}$ as

$$\boldsymbol{w}_{k,i} = \boldsymbol{w}_{k,i-1} + \mu \nabla J(\boldsymbol{w}_{k,i-1}) \tag{9}$$

where μ is the step size parameter. However, we consider the fact that the nodes do not know the form of the cost function beforehand. So we can not use (9) to update the local estimates. Thus, the objective for each node becomes that of determining a good estimate for this gradient vector.

III. PROPOSED COOPERATIVE OPTIMIZATION SCHEME

A. Motivation

In a non-cooperative scheme, (i.e. when $\mathcal{N}_k(i) = \{k\}$) each node can estimate the gradient vector at $\boldsymbol{w}_{k,i-1}$ as $\frac{\boldsymbol{u}_{k,i}}{\|\boldsymbol{u}_{k,i}\|}$. Therefore, for non-cooperative scheme, Eq (9) changes to

$$\boldsymbol{w}_{k,i} = \boldsymbol{w}_{k,i-1} + \mu \frac{\boldsymbol{u}_{k,i}}{\|\boldsymbol{u}_{k,i}\|} I(e_k(i)) \tag{10}$$

where $I(x)$ is the indicator function: it is equal to one when $x > 0$ and zero otherwise. It is important to note that we can get better result if we allow cooperation among nodes. The algorithm in [12] is a cooperation based one that has the following update equation

$$\boldsymbol{w}_{k,i} = \boldsymbol{w}_{k,i-1} + \mu \sum_{\ell=1}^{|\mathcal{N}_{k,i-1}|} a_{\ell k} \boldsymbol{u}_{\ell,i}^T \tag{11}$$

where $|\cdot|$ denotes the cardinality of a set at time $i-1$ and $a_{\ell k}$ are the combination weights. In this work, we propose a new algorithm to enhance the performance of (11) under the conditions of the signals are noisy, thus a linear combination of the neighboring nodes could not provide accurate estimate of gradient vector. To improve the estimation performance, we need to use the data related to information-rich nodes more than other nodes. We can also use bigger step sizes in the initial iterations to increase the probability of finding the information-rich nodes. In the sequel, we propose a new algorithm that considers the mentioned issues.

B. Method

As described in the introduction, the successive $\boldsymbol{w}_{k,i}$ are defined as location vectors. Thus, the neighborhood of node k at time $i-1$ is given by

$$\mathcal{N}_{k,i-1} = \{\ell \in \mathcal{N} : \|\boldsymbol{w}_{\ell,i-1} - \boldsymbol{w}_{k,i-1}\| \leq r_0\} \tag{12}$$

where $r_0 > 0$ is some radius value. As we mentioned before, if we want to improve the estimation performance then we need to use the data related to information-rich nodes more than other nodes. An information-rich node in the the neighborhood of node k is a node that has the following property

$$\{j \in \mathcal{N}_k(i) \mid J(\boldsymbol{w}_{j,i-1}) > J(\boldsymbol{w}_{\ell,i-1}),\ \forall \ell \in \mathcal{N}_k(i)\} \tag{13}$$

Since we do not have access to $J(\boldsymbol{w}_{\ell,i-1})$ the noisy measurements are used to find the information-rich node at node k and time index $i-1$ as

$$j = \arg\max_{\ell}\{x_\ell(i-1) \mid \ell \in \mathcal{N}_{k,i-1}\} \tag{14}$$

Then, in (6) we replace $\boldsymbol{w}_{j,i-1}$ into $\boldsymbol{w}_{k,i-1}$ to get

$$\boldsymbol{u}_{k,i}^m = \boldsymbol{w}_{j,i-1} - \boldsymbol{w}_{k,i-2} \tag{15}$$

where $\boldsymbol{u}_{k,i}^m$ is the modified direction vector. To further improve the performance, we replace the fixed step-size with a variable step-size as

$$\mu_k(i) = \mu_k(i-1) + \beta_k z_k(i) \tag{16}$$

where β_k is a positive constant. Using (15) and (16), we can modify the update equation in (10) and (11) as follows

$$\begin{cases} \boldsymbol{w}_{k,i} = \boldsymbol{w}_{k,i-1} + \mu_k(i)\frac{\boldsymbol{u}_{k,i}^m}{\|\boldsymbol{u}_{k,i}^m\|} I(e_k(i)) \\ \mu_k(i) = \mu_k(i-1) + \beta_k z_k(i) \end{cases} \tag{17}$$

The pseudo code of the proposed algorithm is shown in sequel.

The pseudo code of the proposed algorithm

Initialize $\boldsymbol{w}_{k,i} = 0$
for $i = 1, 2, \cdots$
  for $k = 1 : N$
    find the information-rich node at node k using (14)
    compute modified direction vector $\boldsymbol{u}_{k,i}^m$ using (15)
    update the step size $\mu_k(i)$ using (16)
    update local estimate $\boldsymbol{w}_{k,i} = \boldsymbol{w}_{k,i-1} + \mu_k(i)\frac{\boldsymbol{u}_{k,i}^m}{\|\boldsymbol{u}_{k,i}^m\|} I(e_k(i))$
  end for
end for

IV. SIMULATION RESULTS

In this section we present the simulation results to evaluate the performance of the proposed scheme. As a first example, we consider a cost function as shown in Fig. 2. The cost function is a two-dimensional Gaussian distribution with two peaks placed at locations $(-15,-12)$ and $(15,12)$. Accordingly, the value of cost function at $\boldsymbol{w} = (x,y)$ is given by

$$\begin{aligned} J(x,y) &= b_{\max}\exp\left(-\frac{(x-x_1)^2+(y-y_1)^2}{2\sigma_b^2}\right) \\ &+ b_{\max}\exp\left(-\frac{(x-x_2)^2+(y-y_2)^2}{2\sigma_b^2}\right) \end{aligned} \tag{18}$$

where $(x_1,y_1) = (-15,-12)$, $(x_2,y_2) = (15,12)$, $b_{\max} = 10$, and $\sigma_b = 4$. It is important to note that the shape of cost function may change in time. For example in foraging model for bacteria, the bacteria consume food during the foraging process, we assume that the density of nutrition (cost function) is not affected appreciably. This can be achieved by, for example, continually replenishing the nutrition level [11 , 12]. At time index $i = 0$, nodes are randomly and uniformly distributed over a 40×40 rectangular region centered at $(0,0)$ as shown in Fig. 3. The step size is $\mu = 0.8$ and $\beta = 0.015$ in the proposed algorithm. Moreover, we model the measurement $v_k^2(i)$ in (2) as i.i.d. Gaussian random variable with zero mean and unit variance.

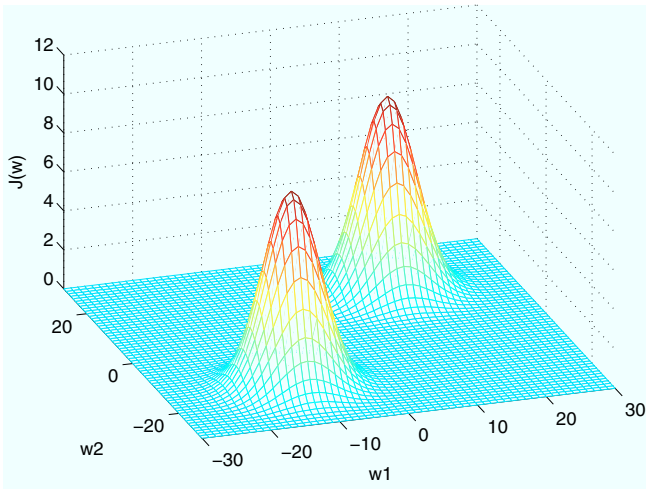

Fig. 2. The distribution of the cost function.

In Fig. 4 we have plotted the final locations of different nodes for non-cooperative scheme in (10), the given algorithm in [12] and the proposed algorithm. Obviously, in the non-cooperative scheme does not work well and only a small fraction of the nodes can find the optimum location. It is clear from Fig. 4 that in the proposed algorithm, the most of nodes have moved from their initial locations to the locations of peak values. So, we can conclude that without cooperation, only the bacteria close to the food source are able to reach the peak values. Fig. 5 shows the average location of nodes $(\eta_k(i))$ per iteration i. We can see again that the proposed algorithm provides better performance in comparison with the given algorithm in [12].

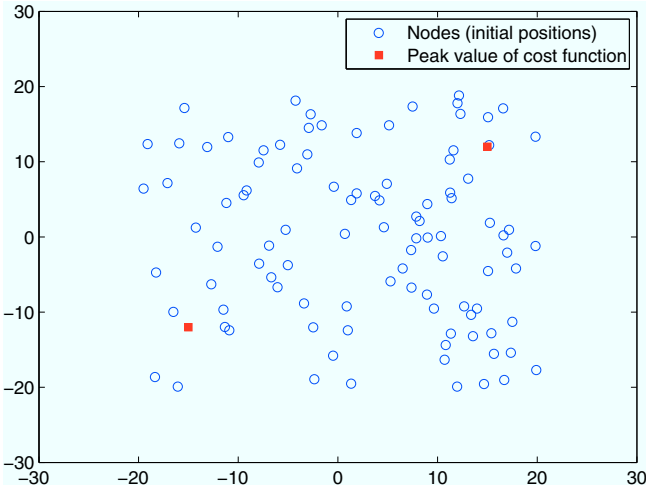

Fig. 3. The initial position of nodes and the location of peak values of the cost function at the $(x_1,y_1) = (-15,-12)$, $(x_2,y_2) = (15,12)$.

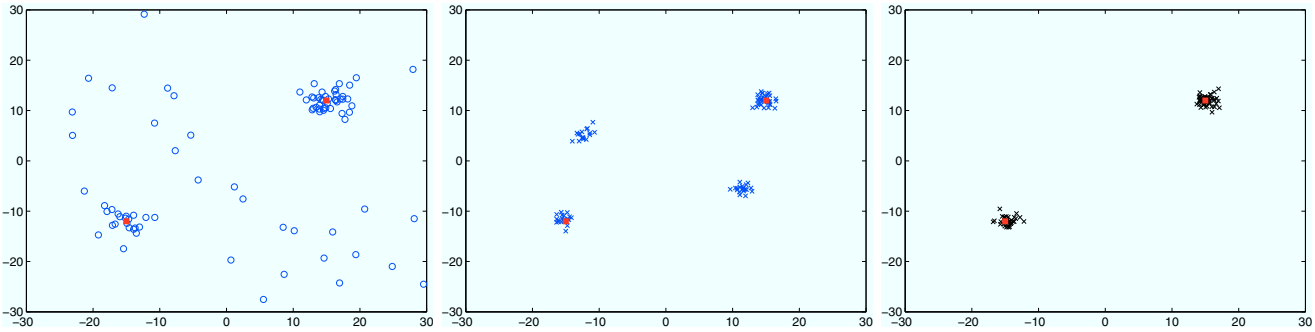

Fig. 4. The final locations of different nodes for non-cooperative scheme in (10) (left), the algorithm given in [12] (middle) and the proposed algorithm (right).

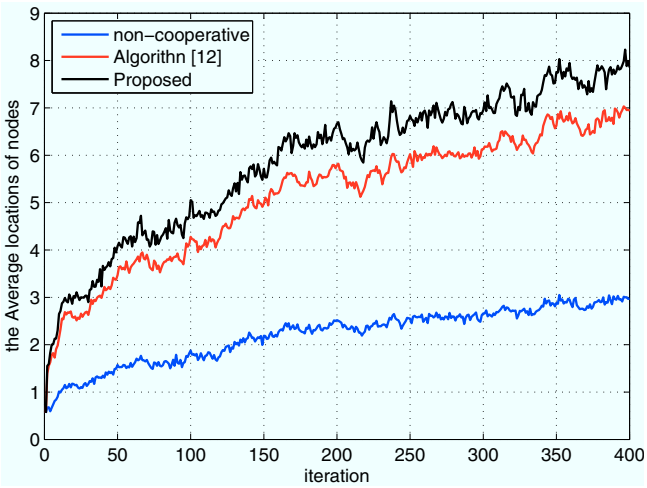

Fig. 5. The average location of nodes ($\eta_k(i)$) per iteration i.

V. CONCLUSIONS

The distributed estimation problems appear in many applications. Although adaptive networks are excellent solution for such problems, however, in most available implementations it is assumed that the shape of the cost function is known beforehand by all nodes in the network. To address this problem, in this paper we proposed a new algorithm that uses the data related to information-rich node, at every node at every iteration, instead of linear combination of the neighbors' estimates. Moreover, in the proposed algorithm bigger step sizes in the initial iterations are used to increase the probability of finding the information-rich nodes. Simulation results indicate the suitable performance of the proposed algorithm. In our future work we will provide mathematical substantiation, additional controlled experiment and algorithm implementation into a real system to support the claims.

REFERENCES

[1] D. Estrin, G. Pottie, and M. Srivastava, "Intrumenting the world with wireless sensor setworks," *in Proc. IEEE Int. Conf. Acoustics, Speech, Signal Processing (ICASSP)*, Salt Lake City, UT, May 2001, pp. 2033-2036.

[2] C. G. Lopes and A. H. Sayed, "Distributed processing over adaptive networks," *in Proc. Adaptive Sensor Array Processing Workshop*, MIT Lincoln Lab., Lexington, MA, Jun. 2006.

[3] A. H. Sayed, *Fundamentals of Adaptive Filtering*, John Wiley and Sons, New York, NJ, USA, 2003.

[4] A. H. Sayed, "Diffusion adaptation over networks," *to appear in E-Reference Signal Processing*, R. Chellapa and S. Theodoridis, editors, Elsevier, 2013.

[5] C. G. Lopes and A. H. Sayed, "Incremental adaptive strategies over distributed networks," *IEEE Trans. Signal Processing*, vol. 55, no. 8, pp. 4064-4077, Aug. 2007.

[6] C. G. Lopes and A. H. Sayed, "Diffusion least-mean squares over adaptive networks: Formulation and performance analysis," *IEEE Trans. on Signal Processing*, vol. 56, no. 7, pp.31223136, July 2008.

[7] F. S. Cattivelli, C. G. Lopes, and A. H. Sayed, "Diffusion recursive least-squares for distributed estimation over adaptive networks," *IEEE Trans. on Signal Processing*, vol. 56, no. 5, pp. 18651877, May 2008.

[8] F. S. Cattivelli and A. H. Sayed, "Multilevel diffusion adaptive networks," *in Proc. IEEE Int. Conf. Acoustics, Speech, Signal Processing (ICASSP)*, Taipei, Taiwan, April 2009.

[9] N. Takahashi, I. Yamada, and A. H. Sayed, "Diffusion least-mean-squares with adaptive combiners," *in Proc. IEEE Int. Conf. Acoustics, Speech, Signal Processing (ICASSP)*, pp. 2845-2848, Taipei, Taiwan, April 2009.

[10] F. Cattivelli and A. H. Sayed, "Self-organization in bird flight formations using diffusion adaptation," *in Proc. 3rd International Workshop on Computational Advances in Multi-Sensor Adaptive Processing*, Aruba, Dutch Antilles, Dec. 2009, pp. 4952.

[11] J. Chen, X. Zhao and A. H. Sayed, "Bacterial motility via diffusion adaptation," *in Proc. 44th Asilomar Conference on Signals, Systems and Computers*, Pacific Grove, CA, Nov. 2010.

[12] J. Chen and A. H. Sayed, "Bio-inspired cooperative optimization with application to bacteria motility," *in Proc. IEEE Int. Conf. Acoustics, Speech, Signal Processing (ICASSP)*, Prague, Czech Republic, pp. 5788-5791, May 2011.

[13] S-Y. Tu and A. H. Sayed, "Foraging behavior of fish schools via diffusion adaptation," *International Workshop on Cognitive Information Processing*, pp. 63-68, Elba Island, Italy, June 2010.

[14] F. Cattivelli and A. H. Sayed, "Self-organization in bird formations using diffusion adaptation," *Proc. IEEE CAMSAP*, Aruba, Dec. 2009.

[15] K. Passino, "Biomimicry of bacterial foraging for distributed optimization and control," *IEEE Control Systems Magazine*, vol. 22, no. 6, pp. 5267, 2002.

[16] Hayes A. T., Martinoli A., and Goodman R. M., "Distributed Odor Source Localization," *Special Issue on Artificial Olfaction, IEEE Sensors*, Vol. 2, No. 3, pp. 260-271, 2002.

[17] S-Y. Tu and A. H. Sayed, "Cooperative prey herding based on diffusion adaptation," *in Proc. IEEE Int. Conf. Acoustics, Speech, Signal Processing (ICASSP)*, Prague, Czech Republic, pp. 3752-3755, May 2011.

[18] H. Ishida, T. Nakamoto, T. Moriizumi, T. Kikas, and J. Janata, "Plumetracking robots: A new application of chemical sensors," *Biol. Bull.*, vol. 200, pp. 222-226, April 2001.

[19] C. R. Kube and E. Bonabeau, "Cooperative transport by ants and robots," *Robot. Auton. Syst.*, vol. 30, pp. 85101, 2000.