Federation University ResearchOnline

<https://researchonline.federation.edu.au>

Copyright Notice

This article is the version that was accepted for publication by Cambridge University Press. The final and definitive version of this article can be found at: <https://doi.org/10.1017/S026357472000079X>

Nguyen, Kodagoda, S., Ranasinghe, R., & Dissanayake, G. (2021). Mobile Robotic Sensors for Environmental Monitoring using Gaussian Markov Random Field. *Robotica* 39 (5), 1–23.

See this record in Federation ResearchOnline at:

<http://researchonline.federation.edu.au/vital/access/HandleResolver/1959.17/176959>

Mobile Robotic Sensors for Environmental Monitoring using Gaussian Markov Random Field

Linh Nguyen^{†*}, Sarath Kodagoda[‡], Ravindra Ranasinghe[‡]
and Gamini Dissanayake[‡]

[†]*School of Engineering, Information Technology and Physical Sciences, Federation University Australia, Churchill, VIC 3842, Australia*

[‡]*Centre for Autonomous Systems, University of Technology Sydney, Ultimo, NSW 2007, Australia*

)

SUMMARY

This paper addresses the issue of monitoring spatial environmental phenomena of interest utilizing information collected by a network of mobile, wireless and noisy sensors that can take discrete measurements as they navigate through the environment. It is proposed to employ Gaussian Markov random field (GMRF) represented on an irregular discrete lattice by using the stochastic partial differential equations method to model the physical spatial field. It then derives an GMRF based approach to effectively predict the field at unmeasured locations, given available observations, in both centralized and distributed manners. Furthermore, a novel but efficient optimality criterion is then proposed to design centralized and distributed adaptive sampling strategies for the mobile robotic sensors to find the most informative sampling paths in taking future measurements. By taking advantage of conditional independence property in the GMRF, the adaptive sampling optimization problem is proven to be resolved in a deterministic time. The effectiveness of the proposed approach are compared and demonstrated using pre-published data sets with appealing results.

KEYWORDS: Environmental monitoring; Mobile robotic sensors; Sensor networks; Spatial prediction; Gaussian Markov random field.

1. Introduction

In recent years, there have been increasing demands for monitoring environments such as exploring ecosystem change in ocean and on land, observing toxic pollutants and detecting forest fires [1]–[3]. These tasks if performed by humans who manually harvest measurements are tiresome and time-consuming. Therefore, in this paper we propose to utilize mobile robotic wireless sensor networks (MRWSNs) [4], [5] to estimate and predict spatial phenomena. With a set of networked mobile sensor nodes and wireless communication, the MRWSNs are competent to provide services required not only for monitoring and exploring the environment but also for exchanging information. The mobility of the robotic wireless sensors can be utilized to generate the optimal strategies for collecting measurements in the spatial environmental field of interest. Furthermore, the MRWSNs are capable of estimating and predicting a spatial phenomenon at unmeasured locations by modelling available measurements. For instance, Leonard *et al.* employed a linear model to predict the ocean ecosystem and presented a performance metric specified by uncertainty to derive a set of parameterized paths for mobile sensors

* Corresponding author. E-mail: l.nguyen@federation.edu.au

[1]. The work in [6] introduced a Kalman filter based technique to estimate the parameters of a physical spatio-temporal process model and then maximized a specified performance to navigate robotic sensors throughout an environment. Graham *et al.* described the random field models by tools from geostatistics, i.e. Kriging, and proposed to utilize either a known [7] or an unknown [8] covariance function. The sensor network includes static computing nodes and mobile sensing agents taking measurements. The static nodes compute gradient of variance and send control commands to robotic sensors. In [9], it was illustrated that the information about the model was observed to consider the adaptive sampling paths. Moreover, the author of [10] developed a distributed Kriged Kalman filter for robotic wireless sensors to predict a spatial field. A consensus algorithm was implemented on new measurements to calculate state predictions of the field. By combining with coverage control [11], Martínez [12] derived a distributed interpolation scheme based on a distributed non-parametric prediction approach for field estimation in mobile sensor networks. In [13], the authors designed a graph based strategy that maximizes information gain from measurements to find visiting locations and moving paths for a mobile robot. Xu *et al.* primarily used the Gaussian process (GP) regression [14], [15] for estimating and predicting the generally scalar field and designed optimality criteria based on the Fisher Information Matrix [16] and the average of the prediction error variances [17], [18] for the optimal sampling paths for the MRWSNs. In terms of the compressive sensing framework, Huang *et al.* [19] maximized the entropy of next measurements to find the next most informative positions for networked mobile sensors. A new utility function based on travelled distances of mobile robotic wireless sensors (MRWSs) was proposed in [20] to present the Bayesian optimization problem that choose the relevant informative locations for the MRWSs in the GP modelled field. In our previous works [21]–[24], by exploiting the model prediction in the spatial phenomena, which relies on the GP model, we developed the algorithms to drive robots through environmental fields, relied on locational optimization [21], conditional entropy [23] and posterior uncertainty [24].

In most of the existing works using mobile robotic sensors for environmental monitoring, GP has been widely demonstrated to be most efficient model for spatial field modelling. The GP delineates the continuous interpretation of the spatial field very well; nevertheless, the size of the covariance matrix is proportional to the number of sensor measurements. That is, the level of computational complexity in the GP regression as well as the sampling optimization problem increases as the MRWSs navigate throughout the environment over time. There have been some pre-published works addressing this issue in different ways. For instance, in their work [25], Williams *et al.* proposed to employ the Nyström method, which is used to numerically address the eigenproblems, to approximately compute the eigendecomposition of the Gram matrix. The proposition enables the kernel-based predictors to be carried out on a smaller system compared with the original one of all the observations. Besides, [26] introduced a simple sparse greedy approach to calculate the approximation of the maximum posteriori estimation of the GP, which diminishes the amount of computation required to find the solution scales. In an effort to overcome the computational complexity in the GP regression, the authors in [27] introduced a framework for sparse GP techniques. The method first considered randomized greedy selections based on information-theoretic principles, then alternatively represented the GP model that facilitates generalizations to regression. In [18], the authors proposed to truncate the sensor measurements in order to deal with the complexity in computing the covariance matrix. The proposition was assumed that the model parameters were known a priori. However, in the present study, it is proposed to estimate the mean parameters and hyperparameters online at each time iteration as observations are gathered, which requires all the measurements to be presented. Recently, the challenge regarding the covariance matrix in the GP has been dealt with by replacing the GP by a computationally efficient Gaussian Markov random field (GMRF) [28]–[35]. The GMRF approximately represents the spatial field on a lattice. It is to be noted that in the extreme cases of very small neighbours in the lattice, the GMRF can approximate

the GP surprisingly well [29]. The computational benefits of the GMRF come from the sparsity of the precision matrix, an inverse of a dense covariance matrix, whose zero elements relate directly to conditional independence property. Due to computational advancement and scalability, the GMRF has received much attention for resource-constrained MRWSNs as compared to the standard GP [33], [36]–[39]. In [39], Xu *et al.* proposed a new class of GP built on a GMRF for the spatial prediction problem. Nonetheless, this proposition requires the precision matrix to be given a priori [36]; and the optimal learning and sampling strategy was not proposed in the work. Although [36]–[38] provided algorithms to derive the sampling locations for the MRWSNs, their solutions were restricted to a regular lattice. Fitting the GMRF to the GP on the regular lattice, as discussed in [32], has to be precomputed for a discrete set of parameter values such as smoothness and range. For instance, in [36], the authors assumed that the support of the discrete prior to the hyperparameters is known, yet how to choose them is unknown.

On the other hand, designing optimally adaptive sampling paths that enable the MRWSs to be driven to the most informative locations to observe the physical environment, as proven in [23], [40], is NP-hard but can be approximately computed by a greedy heuristic algorithm. However, the complexity of resolving sampling strategies by the greedy approach that requires factorizing of covariance or precision matrices has not received much attention from the existing works. Particularly, the algorithm often becomes intractable due to the need of finding the optimal sampling points in a large number of unmeasured locations of interest.

Therefore, in this work, to study a spatial estimation approach in the resource-constrained MRWSNs, we propose to exploit the GMRF model presented on an irregular lattice [35] that is created from the points of interest and the mobile sensor locations in the spatial environment by the use of the Voronoi tessellation theory. It is more realistic since the points of interest and the locations of the mobile sensors are random, not on vertices of a grid. Moreover, all the parameters and hyperparameters are directly estimated from the data and online updated when more measurements are collected. The posterior mean vector and precision matrix are adaptively computed by learning the observations without prior information requirement. It is noticed that the centralized algorithm for the spatial inference was discussed in our previous work [35]. In this work, we propose the distributed algorithm to compute the spatial predictions in a distributed mobile sensor network, where each mobile sensor has a short wireless communication range.

More importantly, since the spatial field is modelled by the GMRF on an irregular lattice, which leads to different schemes to compute the posterior mean vector and precision matrix as compared with the previous literatures (e.g. [36], [39]), we propose a new adaptive centralized sampling algorithm based on the conditional entropy to adaptively optimize the sampling locations for the robotic sensors at each movement. The proposed optimality criterion is proved to be resolved in deterministic time even with higher number of measurements or a larger unmeasured locations on a very dense mesh. Likewise, for the mobile sensors in a distributed network, a distributed sampling approach is developed to find their best sampling paths. Although the proposed distributed schemes require each robotic sensor to communicate with its neighbouring sensors only, it enables the mobile sensors to approximately reach to global estimation and prediction values. Finally, we have extensively evaluated the proposed approaches using two pre-published data sets, where the obtained results are appealing.

The remainder of this paper is organized as follows. In Section 2, we introduce a model of a mobile robotic sensor network employed for environmental monitoring of a spatial field. Section 3 presents both GMRF based centralized and distributed methods for predicting a spatial environmental field given measurements gathered by the MRWSNs. To enhance quality of the spatial field prediction, centralized and distributed adaptive sampling strategies for the MRWSs are discussed in Section 4. Section 5 demonstrates efficiency of the proposed approaches by experimental results before conclusions are drawn in Section 6.

2. A Network of Mobile Robotic Sensors

2.1. Notation

Let $\mathbb{R}$ and $\mathbb{R}_{>0}$ denote the set of real and positive real numbers. $\mathbb{Z}_{>0}$ denotes the set of nonnegative integer numbers. The Euclidean distance function is defined by $\|\cdot\|$. We let $\det(\cdot)$ denote the operator of the determinant. While $\mathcal{O}(\cdot)$ denotes the operator of running time of an algorithm, $\log$ denotes the logarithm operator. $a_t[i]$ is defined as the i -th element of the vector a_t , and $b_t[i, j]$ is referred as an element at the i -th row and j -th column of the matrix b_t . Other notation will be explained as and when they occur.

2.2. Network Model

Consider N spatially distributed mobile sensing agents indexed by $i \in \{1, \dots, N\}$. Agents are equipped with identical sensors, and take measurements at times $t \in \mathbb{Z}_{>0}$. At time t , agent i takes a noisy measurement at its current location $s_{t,i} \in \mathbb{R}^d$, specified by

$$y_{t,i} = X(s_{t,i})\beta_t + z(s_{t,i}) + \varepsilon(s_{t,i}), \quad (1)$$

where $X(s_{t,i})$ denotes a vector of spatially referenced non-random variables (known as covariates) at the location $s_{t,i}$ and β_t is a mean parameter vector estimated at time t . $z(s_{t,i})$ is a latent random variable, which is modelled by a GMRF [32]. $\varepsilon(s_{t,i})$ is a noise with a zero mean and a known variance σ_ε^2 .

For notational simplicity, we denote all mobile robotic sensor locations at time t by $s_t = (s_{t,1}^T, \dots, s_{t,N}^T)^T$, and the measurements made by all robotic sensors at time t by $y_t = (y_{t,1}, \dots, y_{t,N})^T$. Moreover, we denote the collection of all mobile sensors' locations and the collective measurements from time 1 to t by $s_{1:t} = (s_1^T, \dots, s_t^T)^T$, and $y_{1:t} = (y_1, \dots, y_t)^T$ respectively.

Between measurement instants, each mobile sensor moves according to the discrete dynamics,

$$s_{t+1,i} = s_{t,i} + l_{t,i}, \quad (2)$$

where $\|l_{t,i}\| \leq l_{max} \in \mathbb{R}_{>0}$, in which l_{max} is the maximum distance an agent can move between time instances.

It is noted that a robotic sensor has a communication range R and is a neighbour of other sensors iff it can communicate to them. Furthermore, for the safe navigation in a network of mobile sensors, we employ the scheme presented in our previous work [23], where a minimum distance between any two agents is predefined.

3. Spatial Predictive Inference in MRWSNs

As discussed in our previous work [35], a spatial physical field can be presented by an efficiently computational GMRF model, which approximates a continuously indexed Gaussian process but possesses the conditional independence property. It is noticed that the GMRF model is built on an irregular lattice that is created from the points of interest and the locations of measurements by the use of the Voronoi tessellation theory. More importantly, by the use of stochastic partial differential equations (SPDE), the GMRF model projects observations in a spatial field onto an irregular triangulation mesh, which allows the model to construct its very sparse precision matrix. For more details, interested readers are referred to [35]. In this section, we present how the GMRF model can be employed to predict the spatial field in a centralized and a distributed manner.

3.1. Centralized Prediction Schemes

Let the GMRF model creates a mesh of n vertices for a spatial field and impose unmeasured locations of interest upon the vertices on the mesh. Suppose that the latent

random field at those vertices at time t is distributed as

$$z_t | \theta_t \sim \mathcal{N}(\mathbf{0}, Q_t^{-1}), \quad (3)$$

where Q_t is a sparse precision matrix. Interested readers are referred to [35] for more details of how to calculate the precision matrix. And, θ_t is the hyperparameter vector learned at time step t , $Q_t = Q(\theta_t)$.

We define A as a projector matrix, which projects the latent random field modelled at the triangulation mesh vertices to the data locations. At the time instant t , size of the projector matrix, A_t , is $tN \times n$. The projector matrix A_t is built in such a way that each row has only one non-zero element of value 1. This non-zero element is indexed by the correspondingly observed location in the vertices set of the triangulation mesh.

As presented in [35], the collective measurements of the spatial field up to time t are modelled as follows,

$$y_{1:t} | z_t, \theta_t, \beta_t, \sigma_\varepsilon^2, A_t \sim \mathcal{N}(X(s_{1:t})\beta_t + A_t z_t, \sigma_\varepsilon^2 I), \quad (4)$$

where $X(s_{1:t})$ denotes a matrix of covariates, β_t, θ_t are parameters estimated by using all the available observations from time 1 to t , and I denotes the $tN \times tN$ identity matrix. If all model parameters are learned, the joint distribution that is computed by employing the technique in [41] is specified by

$$z_t, y_{1:t} | \theta_t, \beta_t, \sigma_\varepsilon^2, A_t \sim \mathcal{N} \left(\begin{bmatrix} \mathbf{0} \\ X(s_{1:t})\beta_t \end{bmatrix}, \begin{bmatrix} Q_t^{-1} & Q_t^{-1} A_t^T \\ A_t^T Q_t^{-1} \sigma_\varepsilon^2 I + A_t Q_t^{-1} A_t^T \end{bmatrix} \right). \quad (5)$$

One can also represent the covariance matrix in (5) in terms of the corresponding partitioned precision matrix. To do this, we utilize the block-wise inversion approach [42] and the Schur complement. It is given by

$$\begin{bmatrix} Q_t^{-1} & Q_t^{-1} A_t^T \\ A_t^T Q_t^{-1} \sigma_\varepsilon^2 I + A_t Q_t^{-1} A_t^T \end{bmatrix}^{-1} = \begin{bmatrix} Q_t + A_t^T (\sigma_\varepsilon^2 I)^{-1} A_t & -A_t^T (\sigma_\varepsilon^2 I)^{-1} \\ -(\sigma_\varepsilon^2 I)^{-1} A_t & (\sigma_\varepsilon^2 I)^{-1} \end{bmatrix}. \quad (6)$$

If VT is a set of locations at the triangulation mesh vertices, the full conditional distribution of z_t at VT , given $y_{1:t}$, is also Gaussian, derived from (5) and (6) as follows

$$z_t | y_{1:t}, \theta_t, \beta_t, \sigma_\varepsilon^2, A_t \sim \mathcal{N}(\mu_{z_t | y_{1:t}}, Q_{z_t | y_{1:t}}^{-1}), \quad (7)$$

where

$$\mu_{z_t | y_{1:t}} = X(VT)\beta_t + Q_{z_t | y_{1:t}}^{-1} A_t^T (\sigma_\varepsilon^2 I)^{-1} (y_{1:t} - X(s_{1:t})\beta_t), \quad (8)$$

$$Q_{z_t | y_{1:t}} = Q_t + A_t^T (\sigma_\varepsilon^2 I)^{-1} A_t, \quad (9)$$

where $X(VT)$ is a matrix of the covariates at VT .

The primary computation of (8) is to factorize the sparse matrix $Q_{z_t | y_{1:t}}$, which can be obtained in a short time. In other words, since $Q_{z_t | y_{1:t}}$ is not dependent on collection of measurements and its non-zero elements in most case are n out of n^2 entries [43], $\mu_{z_t | y_{1:t}}$ can be computed in $\mathcal{O}(n^{\frac{3}{2}})$. In contrast, if the spatial field is represented by the standard GP model [15], the computational complexity of the prediction scheme is cubic in the dimension of the dataset [44], which leads to computational intractability with the increase in the number of observations over time.

3.2. Distributed Prediction Schemes

3.2.1. Discrete-time average consensus (DAC) approach. The discrete-time average consensus algorithm [45] is frequently utilized in a distributed mobile network in order to enable each agent to trace the average value of whole network signals. Consider a MRWSN in which every mobile sensor i has its own measurement signal g_i . It is aimed that sensor i needs to compute the arithmetic mean of elements in the vector $g = (g_1, g_2, \dots, g_N)^T \in \mathbb{R}^N$. It is assumed that each robotic sensor location is a vertex of a graph. If this graph is connected, then the arithmetic mean of g can be obtained by the following iterative procedure

$$x_i[k+1] = x_i[k] + \gamma \sum_{j \neq i} a_{ij} (x_j[k] - x_i[k]) \quad (10)$$

with initial condition $x[0] = g$, where $a_{ij} = 1$ iff i -th sensor and j -th sensor are neighbours and 0 otherwise, and γ is the step size. Olfati-Saber *et al.* [45] demonstrated that if γ holds

$$0 < \gamma < \frac{1}{\Delta},$$

the DAC algorithm converges to the solution, where $\Delta = \max_i \left(\sum_{j \neq i} a_{ij} \right)$ is the maximum degree of the graph. At the end of the algorithm, every mobile sensor obtains the average value of the vector g that is equivalent to $\sum_{i=1}^N \frac{g_i}{N}$.

3.2.2. Distributed estimators over time. Consider a MRWSN of N agents and suppose that every robotic sensor knows the mesh of the triangulation and is located at an arbitrary vertex of the mesh. However, each sensor does not know the locations of any other sensors out of its communication range. At the time step t , the environmental measurements collected at the i^{th} mobile wireless sensor are included its own observations and its neighbours' readings, accumulated up to time t , as specified by

$$y_{i,1:t} = \left(\bigcup_{k=1}^t y_{k,i} \right) \cup \left(\bigcup_{k=1}^t \left(\bigcup_{g=1}^{N_{ik}} y_{k,g} \right) \right),$$

where N_{ik} is the number of the i^{th} sensor's neighbours at the time k . Moreover, the sampling locations corresponding to $y_{i,1:t}$ are known to the i^{th} mobile sensor as follows,

$$\mathbf{s}_i = \left(\bigcup_{k=1}^t s_{k,i} \right) \cup \left(\bigcup_{k=1}^t \left(\bigcup_{g=1}^{N_{ik}} s_{k,g} \right) \right).$$

For the sake of simplicity, we denote $y_i = y_{i,1:t}$.

We suppose that by exchanging information with the neighbours, every mobile agent can only predict the field at an unobserved vertex v_j that is inside the sphere of radius $0.5R$ centered at the sensor. Hence, the global posterior prediction at any unmeasured locations can be obtained as the following proposition.

Proposition 1. Given the measurements y_i collected by the mobile robotic wireless sensor i from time 1 to t and the global parameters of the spatial field model, i.e. β_t, θ_t , every mobile agent i can approximately compute the global posterior prediction at v_j , where $\|v_i - v_j\| \leq 0.5R$, as follows.

(1) Mean:

$$m_{v_j|s_{1:t}} = X(v_j)\beta_t + \frac{1}{N_j} \sum_{i=1}^{N_j} \varphi_i[j], \quad (11)$$

where

$$\varphi_i = Q_{z_{t,i}|y_{1:t}}^{-1} A_{t,i}^T (\sigma_\varepsilon^2 I_i)^{-1} (y_i - X(\mathbf{s}_i) \beta_t),$$

and

- N_j is the number of sensors whose distances to v_j are less than or equal to $0.5R$;
- $A_{t,i}$ is the projector matrix that projects the random field modelled at the triangulation vertices to $\mathbf{s}_i$;
- I_i is a $|\mathbf{s}_i| \times |\mathbf{s}_i|$ identity matrix;
- $Q_{z_{t,i}|y_{1:t}} = Q_t + A_{t,i}^T (\sigma_\varepsilon^2 I_i)^{-1} A_{t,i}$.

(2) Variance:

$$\sigma_{v_j|s_{1:t}}^2 = \frac{1}{N_j} \sum_{i=1}^{N_j} Q_{z_{t,i}|y_{1:t}}^{-1} [j, j]. \quad (12)$$

Proof. In the centralized schemes, it can be seen that $\phi = A_t^T (\sigma_\varepsilon^2 I)^{-1} (y_{1:t} - X(s_{1:t}) \beta_t)$ is a n column vector with non-zero elements corresponding to sensor locations on the mesh. Product of the j^{th} row of $Q_{z_t|y_{1:t}}^{-1}$ and ϕ is the prediction mean of the random field at the vertex v_j on the triangulation.

In the distributed schemes, due to lack of information of all sensor locations, the mobile agent i cannot compute the projector matrix A_t . This leads to (8) not being calculated in a distributed way. However, look into (8), suppose that if sensor locations v_k and v_j are distant, then covariance $Q_{z_t|y_{1:t}}^{-1} [j, k]$ between the random variables at v_k and v_j is trivial. In other words, the predicted field at v_j would not change much if the robotic sensor at v_k does not participate in this prediction computation. Therefore, in this consideration, we propose that each mobile agent i employs its own measurements and its neighbour sensors' readings y_i but uses the global parameters (β_t, θ_t) to predict the random field φ_i at the triangulation vertices. We can see that φ_i is a n column vector but only the element $\varphi_i[j]$ that is estimated by the mobile sensor i is used to calculate the global random field at v_j .

Notice that the assumption $\|v_i - v_j\| \leq 0.5R$ is required for making communications among robotic sensors that can estimate the field at v_j , which holds that $\frac{1}{N_j} \sum_{i=1}^{N_j} \varphi_i[j]$ and the variance can be obtained by the group of these agents as average consensuses. $\square$

Each mobile agent in the v_j -estimated group predicts the spatial field at v_j in a distributed way; yet its own predicted mean and variance are exchanged with its neighbouring sensors in the group to compute the global prediction mean and variance at this predicted location. The distributed algorithm for a robotic sensor is summarised in Algorithm 1.

3.3. Parameters Estimation

The mean parameters β_t and hyperparameters θ_t can be estimated by utilizing generalized least squares technique [14] and the maximum likelihood approach [46]. In the following, a recursive algorithm for estimating the mean parameters β_t and hyperparameters θ_t at time step t is described.

3.3.1. Centralized Learning. From (5), it is straightforward to write down the marginal distribution of $y_{1:t}$ as

$$y_{1:t} | \theta_t, \beta_t, \sigma_\varepsilon^2, A_t \sim \mathcal{N}(X(s_{1:t}) \beta_t, \sigma_\varepsilon^2 I + A_t Q_t^{-1} A_t^T). \quad (13)$$

Let $\Sigma_t = \sigma_\varepsilon^2 I + A_t Q_t^{-1} A_t^T$, in accordance with the best linear unbiased estimator [14], β_t can be obtained by minimizing the function

$$f(\beta_t) = (y_{1:t} - X(s_{1:t}) \beta_t)^T \Sigma_t^{-1} (y_{1:t} - X(s_{1:t}) \beta_t).$$

Algorithm 1 Distributed prediction algorithm for each mobile agent i at time step t

Input: $y_i = \left(\bigcup_{k=1}^t y_{k,i} \right) \cup \left(\bigcup_{k=1}^t \left(\bigcup_{g=1}^{N_{ik}} y_{k,g} \right) \right)$, $\mathbf{s}_i = \left(\bigcup_{k=1}^t s_{k,i} \right) \cup \left(\bigcup_{k=1}^t \left(\bigcup_{g=1}^{N_{ik}} s_{k,g} \right) \right)$, β_t, θ_t, v_j

Output: Predicted field at unobserved location v_j , including mean $m_{v_j|s_{1:t}}$ and variance $\sigma_{v_j|s_{1:t}}^2$

At time $t \in \mathbb{Z}_{>0}$, do

- 1: Compute Q_t based on θ_t
 - 2: Compute $A_{t,i}$
 - 3: Compute $Q_{z_{t,i}|y_{1:t}} = Q_t + A_{t,i}^T (\sigma_\varepsilon^2 I_i)^{-1} A_{t,i}$
 - 4: Compute $\varphi_i = Q_{z_{t,i}|y_{1:t}}^{-1} A_{t,i}^T (\sigma_\varepsilon^2 I_i)^{-1} (y_i - X(\mathbf{s}_i) \beta_t)$
 - 5: Compute average random value $\frac{1}{N_j} \sum_{i=1}^{N_j} \varphi_i[j]$ via DAC approach, where j is the index of the location v_j on triangulation mesh
 - 6: Compute mean $m_{v_j|s_{1:t}} = X(v_j) \beta_t + \frac{1}{N_j} \sum_{i=1}^{N_j} \varphi_i[j]$
 - 7: Compute variance $\sigma_{v_j|s_{1:t}}^2 = \frac{1}{N_j} \sum_{i=1}^{N_j} Q_{z_{t,i}|y_{1:t}}^{-1}[j, j]$ via DAC approach
-

If given θ_t , i.e. $\Sigma_{1:t}$ is known, the estimated β_t can be specified by

$$\hat{\beta}_t = (X(s_{1:t})^T \Sigma_t^{-1} X(s_{1:t}))^{-1} X(s_{1:t})^T \Sigma_t^{-1} y_{1:t}. \quad (14)$$

Then, from (13), the log-likelihood function can be written as

$$f(\theta_t, \beta_t) = -\frac{1}{2} \{ (y_{1:t} - X(s_{1:t}) \beta_t)^T \Sigma_t^{-1} (y_{1:t} - X(s_{1:t}) \beta_t) + \log \det(\Sigma_t) + tN \log(2\pi) \}. \quad (15)$$

By substituting $\hat{\beta}_t$ into (15) and then numerically maximizing this log-likelihood function with respect to θ_t , we can obtain the hyperparameters $\hat{\theta}_t$. These hyperparameters are employed to compute $\hat{\beta}_t$ in the next iteration. The algorithm converges to the model parameters estimation as error of two estimated values of each parameter at two consecutive iterations is trivial.

The complexity of the evaluation of the log-likelihood function is considerably dependent on inverting the covariance matrix Σ_t . The complexity of inverting this matrix is $\mathcal{O}(n^3)$. Since n is vary large, it is worth considering how to efficiently compute the logarithm of the determinant of the covariance matrix Σ_t and its inversion, i.e. the precision matrix Σ_t^{-1} .

By the matrix determinant lemma, the logarithm of the determinant of Σ_t can be given by

$$\begin{aligned} \log \det(\Sigma_t) &= \log \det(Q_t + \frac{1}{\sigma_\varepsilon^2} A_t^T A_t) - \log \det(Q_t) \\ &\quad + 2tN \log(\sigma_\varepsilon). \end{aligned}$$

Similarly, the precision matrix Σ_t^{-1} can be obtained by employing the Woodbury maxtrix identity as follows

$$\Sigma_t^{-1} = \frac{1}{\sigma_\varepsilon^2} \left(I - A_t (Q_t + A_t^T A_t \frac{1}{\sigma_\varepsilon^2})^{-1} A_t^T \right).$$

As a result, these formulae ameliorate the calculations remarkably in the MRWSNs with limited resources.

3.3.2. Distributed Learning. Consider a distributed MRWSN where we assume that a graph with vertices of mobile sensor locations is connected over time. In a distributed way, every robotic sensor i can learn its own parameters and hyperparameters by utilizing its collective measurements y_i . Each mobile agent can employ the recursive algorithm introduced in the previous section to estimate its own model parameters. Nevertheless, in order to predict the physical field at any unobserved locations, the GMRF regression technique, as described in (11) and (12), requires the global parameters that can be obtained by the following proposition.

Proposition 2. If a MRWSN is connected, each mobile sensor i can calculate the global parameters and hyperparameters by exchanging local information with its neighbours as given

$$\xi_i[k+1] = \xi_i[k] + \gamma \sum_{j \neq i} a_{ij}(\xi_j[k] - \xi_i[k]), \quad (16)$$

where $\xi_i = (\theta_{t,i}, \beta_{t,i})^T$, and $\beta_{t,i}$ and $\theta_{t,i}$ are a mean parameter vector and a hyperparameter vector estimated at the time instant t by the mobile sensor i .

Proof. Since, it can be clearly seen that (16) is similar to (10), the results are directly derived by the DAC approach. $\square$

4. Adaptive Sampling Approach

The objective of this paper is to estimate and predict efficiently and effectively the spatial field by using a limited number of the MRWSs. Therefore, there is a need to design optimal adaptive sampling paths that enable the MRWSs to be driven to the most informative locations to observe the physical environment. It is shown that each MRWS always looks for a next location at time instant $t+1$ given all measurements up to time t . Note that the robotic sensors collect measurements of the field at the vertices of the triangulation. In the following, we consider the one-step-ahead spatial prediction technique before presenting the adaptive sampling approach.

4.1. One-step-ahead Spatial Inference

It can be clearly seen that looking for the most information locations s_{t+1} at time instant $t+1$ without y_{t+1} is intractable. Hence, in order to address this problem, the one-step-ahead forecast is utilized to predict the latent spatial values at time $t+1$, given the collective observations up to time t . The readers are referred to [47] for more detailed discussion.

The spatial field model at time t can be written as

$$y_{1:t} = X(s_{1:t})\beta_t + A_t z_t + \varepsilon(s_{1:t}), \quad (17)$$

where $\varepsilon(s_{1:t}) \sim \mathcal{N}(\mathbf{0}, \sigma_\varepsilon^2 I)$.

Given $z_t|y_{1:t}$ (11), the latent random process at time step $t+1$ given observations $y_{1:t}$ can be forecasted by

$$\hat{z}_{t+1}|y_{1:t} = z_t|y_{1:t} + \eta_{t+1}, \quad (18)$$

where $\eta_{t+1} \sim \mathcal{N}(\mathbf{0}, \sigma_{sys}^2 I_\eta)$, I_η is a $n \times n$ identity matrix, and σ_{sys}^2 is a system error that describes the changes in the elements of the parameters vector between times t and $t+1$.

Let $Q_{\eta,t+1}^{-1} = \sigma_{sys}^2 I_\eta$ and suppose that σ_{sys}^2 is to be known, the prior conditional distribution of $\hat{z}_{t+1}$ given $y_{1:t}$ can be obtained by

$$\hat{z}_{t+1}|y_{1:t} \sim \mathcal{N}(\mu_{z_t|y_{1:t}}, Q_{z_t|y_{1:t}}^{-1} + Q_{\eta,t+1}^{-1}). \quad (19)$$

Therefore, the measurements forecasted at time $t + 1$ is

$$\hat{y}_{t+1}|y_{1:t} = X(s_{t+1})\beta_t + A_t(\hat{z}_{t+1}|y_{1:t}) + \varepsilon(s_{t+1}), \quad (20)$$

where $\varepsilon(s_{t+1}) \sim \mathcal{N}(\mathbf{0}, \sigma_\varepsilon^2 I_{t+1})$ and I_{t+1} is a $N \times N$ identity matrix.

Moreover, if we define a column vector $b_{t+1} \in \mathbb{R}^n$ at time $t + 1$ as

$$b_{t+1}[i] = \begin{cases} 1, & \text{if } VT(i) = s_{t+1,j} \text{ and } j = 1, \dots, N \\ 0, & \text{otherwise.} \end{cases}$$

Then, the posterior precision matrix of the posterior distribution of $\hat{z}_{t+1}$ conditioned on $y_{1:t}$ and $\hat{y}_{t+1}$ is specified by

$$\hat{Q}_{z_{t+1}|y_{1:t}, \hat{y}_{t+1}} = (Q_{z_t|y_{1:t}}^{-1} + Q_{\eta, t+1}^{-1})^{-1} + \frac{1}{\sigma_\varepsilon^2} b_{t+1} b_{t+1}^T. \quad (21)$$

For the sake of simplicity, from here we will use $\hat{Q}_{z_{t+1}} = \hat{Q}_{z_{t+1}|y_{1:t}, \hat{y}_{t+1}}$, $Q_{z_t} = Q_{z_t|y_{1:t}}$, and $Q_\eta = Q_{\eta, t+1}$.

Note that these precision matrices are very sparse and that will be utilized for fast computation in the following sections.

4.2. Centralized Sampling Strategy

In recent related works, the approaches commonly used to find the most informative sampling locations at the next time step $t + 1$ are derived from information-theoretic criteria such as entropy [48]. These works proposed to optimize the conditional entropy on the GP models [9], [16], [23], or on the GMRF models restricted to a regular lattice [36]–[39]. In this subsection, we extend the conditional entropy based adaptive sampling method on the GMRF models that are described on an irregular lattice. Moreover, we demonstrate the complexity of these existing methods and propose a novel approach that is computationally practical for the MRWSNs.

The crucial idea of our proposed technique is to drive the MRWSs to new sampling positions in such a way that the overall uncertainty at the n spatial sites is minimized. In terms of the GP technique, let us define a very dense covariance matrix $\hat{\Sigma}_{z_{t+1}}$ of $\hat{z}_{t+1}|y_{1:t}, \hat{y}_{t+1}$. $\hat{\Sigma}_{z_{t+1}}$ can be obtained by $\hat{\Sigma}_{z_{t+1}} = \hat{Q}_{z_{t+1}}^{-1}$. The conditional entropy of $\hat{z}_{t+1}|y_{1:t}, \hat{y}_{t+1}$ is computed in terms of the closed form expression (21) as

$$H(\hat{z}_{t+1}|y_{1:t}, \hat{y}_{t+1}) = \frac{1}{2} \{ \log \det(\hat{\Sigma}_{z_{t+1}}) + n \log(2\pi) + n \}.$$

In a simplified form, we denote

$$H = \log \det(\hat{\Sigma}_{z_{t+1}}) \quad (22)$$

as a conditional entropy function. As a result, the next optimal sampling locations at time instant $t + 1$, s_{t+1} , can be obtained by minimizing the uncertainty as follows

$$s_{t+1}^{opt} = \underset{s_{k+1,i} \in \Omega_{VT}^{(t)}(i)}{\operatorname{argmin}} \log \det(\hat{\Sigma}_{z_{t+1}}), \quad (23)$$

where $\Omega_{VT}^{(t)}(i)$ is a set of the spatial sites inside VT but restricted by $\Omega_i^{(t)}$, which is the allowable movement region of the MRWS i at time t [23].

This optimization problem can be rewritten in terms of the precision matrix. We have

$$s_{t+1}^{opt} = \underset{s_{k+1,i} \in \Omega_{VT}^{(t)}(i)}{\operatorname{argmax}} \log \det(\hat{Q}_{z_{t+1}}). \quad (24)$$

It can be clearly seen that the above combinatorial optimization problems are NP-hard [49], [50]. Usually, these issues can be addressed by employing a greedy algorithm that sequentially finds a set of the next near-optimal sampling locations at time $t + 1$, s_{t+1}^* . Nonetheless, the costs of these algorithms need to be considered in the resource-constrained MRWSNs. The computational complexity of the algorithms can be described as below.

Lemma 1. Given a GP model, the greedy algorithm can resolve the optimization problem (23) in time $\mathcal{O}(Nn^4)$.

Proof. Since the size of $\hat{\Sigma}_{z_{t+1}}$ is $n \times n$, exact computations of $\log \det(\hat{\Sigma}_{z_{t+1}})$ require $\mathcal{O}(n^3)$ operations [51]. In addition, VT has n elements, so there are $\mathcal{O}(n)$ such logarithms of the determinants to be calculated.

Moreover, in the adaptive sampling strategy, we need to find N next MRWS locations at every time step. Thus, the greedy algorithm has to run N iterations in time $\mathcal{O}(N)$. So, the optimization problem (23) can be finalized in running time $\mathcal{O}(Nn^4)$. $\square$

In general, n is very large; despite being under a greedy approach, the problem (23) is still intractable in the resource-constrained MRWSNs.

Lemma 2. Given a GMRF model, the greedy algorithm can address the optimization problem (24) in time $\mathcal{O}(Nn^{\frac{5}{2}})$.

Proof. Note that the size of $\hat{Q}_{z_{t+1}}$ is also $n \times n$. As discussed in [43],

$$\log \det(\hat{Q}_{z_{t+1}}) = 2 \sum_{i=1}^n \log(L_{ii}),$$

where $\hat{Q}_{z_{t+1}} = LL^T$. This Cholesky decomposition is calculated in time $\mathcal{O}(n^{\frac{3}{2}})$ in two dimensional space.

The rest of this proof is similar and refers to the proof of the **Lemma 1**. Therefore, the computational complexity of (24) is $\mathcal{O}(Nn^{\frac{5}{2}})$. $\square$

The GMRF models enable the algorithm to remarkably reduce the computational time of finding the near-optimal sampling locations for the MRWSNs. Nevertheless, this approach is still impractical. In the following, we propose an efficient novel optimality criterion for the adaptive sampling problem in the MRWSNs.

Proposition 3. Given Q_{z_t} computed at time t by (9) and σ_{sys}^2 , the next sampling positions for the MRWSs at the time step $t + 1$ is the optimal solution of the following problem.

$$s_{t+1}^{opt} = \underset{s_{t+1,i} \in \Omega_{VT}^{(t)}(i)}{\operatorname{argmax}} b_{t+1}^T (Q_{z_t}^{-1} + Q_{\eta}^{-1}) b_{t+1}. \quad (25)$$

Proof. Let $\Psi = (Q_{z_t}^{-1} + Q_{\eta}^{-1})^{-1}$. From (24) and (21), and by utilizing the matrix determinant lemma, we have

$$\begin{aligned} H &= -\log \det\{\Psi + \frac{1}{\sigma_{\varepsilon}^2} b_{t+1} b_{t+1}^T\} \\ H &= -\log \left\{ \left(1 + \frac{1}{\sigma_{\varepsilon}^2} b_{t+1}^T \Psi^{-1} b_{t+1} \right) \det(\Psi) \right\} \\ H &= -\log \left(1 + \frac{1}{\sigma_{\varepsilon}^2} b_{t+1}^T \Psi^{-1} b_{t+1} \right) - \log \det(\Psi) \end{aligned}$$

It is stated that Ψ is constant at each time instant, Ψ^{-1} is symmetric positive definite, and the logarithm function is monotonic. For these reasons, it can be in turn defined

$$H = b_{t+1}^T \Psi^{-1} b_{t+1}$$

to complete the proof. $\square$

The optimization problem in the **Proposition 3** is still combinatorial and NP-hard. Nonetheless, the greedy algorithm can cope with this maximization issue in a reasonable computational time as follows.

Theorem 1. Given a GMRF model, the near-optimal solution of (25) can be obtained in time $\mathcal{O}(Nn)$.

Proof. Due to the sparse structure of Q_{z_t} , $Q_{z_t}^{-1}$ can be computed in a very short time at time t and constant at time $t + 1$. Moreover, the diagonal matrix Q_{η}^{-1} is constant as presented in Section 4.1. It is noted that in most case the precision matrix has only n non-zero elements out of its n^2 entries [43]. More importantly, b_{t+1} has only N non-zero elements as defined in Section 4.1. Therefore, the computation of $b_{t+1}^T (Q_{z_t}^{-1} + Q_{\eta}^{-1}) b_{t+1}$ requires $\mathcal{O}(N\sqrt{n})$ operations.

It is akin to the proof in the **Lemma 1**, computing $b_{t+1}^T (Q_{z_t}^{-1} + Q_{\eta}^{-1}) b_{t+1}$ is repeated $\mathcal{O}(n)$ times, and the adaptive sampling algorithm is required to run $\mathcal{O}(N)$ iterations to find N locations for the robotic sensors to move in the next step. Hence, the total time to resolve the optimization problem (25) is $\mathcal{O}(N^2 n^{\frac{3}{2}})$. $\square$

It is apparent that the proposed approach can be efficiently implemented in the resource-limited MRWSNs. It is also to be noted that the computational complexity remains approximately constant even with an increasing number of measurements.

4.3. Distributed Sampling Strategy

We have designed the distributed algorithms in the previous section to predict the spatial field at time t . In this subsection, an efficient navigation strategy is developed to implement the equation (25) in a distributed fashion. Under the assumption of the allowable measurement region, each mobile agent i can find the most informative location inside $\Omega_{VT}^{(t)}(i)$ as below.

Proposition 4. Given Q_t obtained at time t , the next sampling location for the robotic sensor i at the time step $t + 1$ is $s_{t+1,i} = v_k \in \Omega_{VT}^{(t)}(i)$ that satisfies

$$s_{t+1,i}^{opt} = \underset{s_{t+1,i} = v_k \in \Omega_{VT}^{(t)}(i)}{\operatorname{argmin}} Q_t[k, k]. \quad (26)$$

Proof. At the mobile sensor i , we denote a column vector $b_{t+1,i} \in \mathbb{R}^n$ as

$$b_{t+1,i}[j] = \begin{cases} 1, & \text{if } VT(j) = s_{t+1,i} \\ 0, & \text{otherwise.} \end{cases}$$

We can see that $b_{t+1,i}$ is a vector with one non-zero element. Hence, from (25), σ_ε^2 will be excluded from the result as the noise corresponds to measurements. The constant-diagonal matrix Q_η^{-1} can also be eliminated. Eventually, the result of the right hand side of (25) obtained by the sensor i is $Q_t^{-1}[k, k]$ for computing iteration of v_k . $\square$

Obviously, the computational complexity of (26) is $\mathcal{O}(N_i)$, where N_i is the number of the vertices inside $\Omega_{VT}^{(t)}(i)$. This computational time is very promising for each resource-constrained robotic sensor to find an optimal sampling path.

5. Results and Discussions

In this section, in order to illustrate the efficiency of the proposed approaches, experiments were conducted using published data sets described as follows. The aim was to compute the predicted field and the prediction error variances in the whole environment as well as the root mean square errors. More particularly, the computational time for finding the efficient sampling paths for MRWSs was also calculated. Simulated results are demonstrated in two following subsections: *centralized scheme* and *distributed scheme*. Usually, the centralized approach is utilized in a small network, where all sensors can easily transmit all measurements to the sink. Hence, the centralized algorithm was carried out on the numerically generated data set [36], where a small network of 5 mobile sensors was utilized to observe the physical field. In contrast, the distributed method is often used in a large-scale network. In this experiment, the distributed algorithm was implemented in a network of 30 mobile sensing agents that were deployed in a simulated agricultural field. Data set for this experiment was derived from a real-world application [52] and regenerated by [53]. Note that all experiments were implemented in two dimensional environments.

5.1. Centralized Navigation Scheme

In this discussion, one considers a situation where, at time step t , all mobile sensors make new observations and transmit them to the sink via a specific routing tree. Then the base station computes the centralized sampling strategy and sends control commands back to each robotic sensor. All prediction activities are conducted by the sink.

Fig. 1: The true field of the numerically generated data set. Range of the fields is shown in color bar.

Fig. 2: The results of the numerically generated field in Case 1: The predicted fields at time (a) $t = 1$, (b) $t = 10$ and (c) $t = 20$. Ranges of the fields are demonstrated in color bars.

A spatial field of the physical quantity was generated within 100 units $\times$ 50 units as shown in Fig. 1. The range of the field varies from 15.00 to 28.30. There were five ($N = 5$) MRWSs used with constrained individual displacements of a maximum radius of 5 units in every time step. The desired buffer width was set to $\omega = 0.3$ units. In this evaluation, three cases are considered in three different starting conditions as illustrated in Case 1 Fig. 3a, Case 2 Fig. 4a and Case 3 Fig. 4b. In Case 1, for comparison purposes, all

Fig. 3: The results of the numerically generated field in Case 1: The predicted variances at time (a) $t = 1$, (b) $t = 10$ and (c) $t = 20$. Ranges of the errors are demonstrated in color bars. The sampling paths for MRWSs up to time steps 1, 10 and 20, respectively, are illustrated by white circles, where current mobile sensor locations are shown in white dots.

MRWSs were started from the pre-defined locations as shown in Fig. 3a (in white dots), similar to the work in [36]; however, the robots were started from random locations in Case 2 and Case 3. It was assumed the knowledge of the measurement noise $\sigma_\varepsilon^2 = 0.2$ and the system error $\sigma_{sys}^2 = 0.25$. The spatial field was represented by approximately 5000 vertices of triangles which were considered as spatially interested sites. In this case, the

percentage of non-zero elements in the sparse precision matrix is around 18%, whereas the covariance matrix is comprehensively dense.

Figures 2a, 2b, and 2c demonstrate the predicted fields at times $t = 1, 10$ and 20 . It can be seen that the predicted means are approaching the true field as illustrated in Fig. 1 when the number of the observations increases. White circles in Fig. 3a, 3b, and 3c show the trajectories of the near-optimal paths of the MRWSs up to time instants 1, 10, and 20, respectively. Figures 3a, 3b and 3c also show the gradual reduction in the prediction error variances with the growth of time steps, corresponding to 1^{st} , 10^{th} , and 20^{th} instant respectively. The prediction errors are higher in the areas distant from the measured points. Nonetheless, the predicted variances in the whole space are trivial when 5 MRWSs collect 20 observations each.

Fig. 5 shows the root mean square errors (RMSEs) of the predicted field at the spatial sites. As expected, increased number of observations or with more exploration leads to gradual reduction of the RMSEs.

Fig. 4: The sampling paths for MRWSs up to time step 20 are illustrated by white circles, where current mobile sensor locations are shown in white dots. (a) Case 2 and (b) Case 3.

For comparison purposes, a generic GP regression [15] for estimating and predicting the spatial phenomenon was implemented in Case 1. Note that since the standard GP continuously interpret the spatial field, its prediction results should illustrate the best performance in all scenarios studied. However, computing the dense covariance matrix is a main obstacle to addressing the spatial prediction issue, which we have proposed to exploit the GMRF in this paper. Under starting conditions in Case 1, as can be seen in Fig. 5, results of the proposed approach are highly comparable to those of the standard GP based method. Furthermore, the implementation of the approach based on a regular lattice proposed by Xu *et al.* [36] was also conducted. It can be clearly seen

Fig. 5: The root mean square errors for the numerically generated field at different starting conditions.

that the proposed GMRF outperforms [36]. The rationale could be in the fact that to compute the Bayesian inferences [36] assumes that the mean values are known to be zero and the support of the discrete prior to the hyperparameters is known, which should be learned from the data. In contrast, in the proposed approach, no prior information is required, and the parameters and hyperparameters are directly estimated from the sensor readings and repeatedly updated when new measurements are collected. In order to highlight the consistency of resulting performance, the proposed approach was then evaluated in Case 2 and Case 3, respectively. Results corresponding to the prediction quality as demonstrated in Fig. 5 in Cases 1, 2 and 3 show that the proposed algorithm generates highly promising solutions for all the cases studied in this work.

It is noted that in reality the robotic sensors may be deployed to start from the same location. In that case, measurements collected the mobile sensors in the first step are similar, which leads to predicting a plain spatial field in the whole environment in which every location has a constant prediction. In other words, there is no uncertainty in predictions of the spatial field, and the locations for the robots to move in the next (second) step cannot be optimized or algorithmically found. Thus, in that case, it is suggested that all the robotic sensors move to random locations in their first movements. Once they are located at random positions, the proposed algorithm can efficiently drive them from the second movements onwards, as presented in Case 2 and Case 3.

Another important aspect to compare is the computational complexity in finding the sampling paths for the MRWSs. In section 4.2, three strategies to optimize the informative navigations in the MRWSNs have been formulated. They all aim to minimize the uncertainty at unobserved locations of interest. The first strategy uses a GP model to describe the spatial field and then to find sensing locations by minimizing the logarithm of the determinant of the covariance matrix. The strategy is specified by (23). The second strategy represents the field by the GMRF and solves the sampling problem by maximizing the logarithm of the determinant of the precision matrix, which is shown in (24). In the third scheme, the GMRF is utilized to model the spatial field, but a new criterion for the sampling path issue has been proposed, which removes the need

(a)

(b)

(c)

Fig. 6: The results of the agricultural field: The model parameters of GMRF are estimated by centralized and distributed approaches proposed. These parameters are (a) mean parameter β_0 , (b) $\log(\tau)$, and (c) $\log(\kappa)$.

for computing the logarithm of the determinant. The proposed criterion is described by (25). These strategies are all resolved by the greedy algorithm. It was noted that the run time of the algorithms did not ascend with the rise of the number of observations. More importantly, in this particular illustrative experiment, in each time instant, the algorithm took approximately 9 seconds to address the problem (25), whereas solutions

Fig. 7: The true field of the agricultural data set. Percentage of the soil organic matter is shown in color bar.

of the problems (23) and (24) required approximately 7 hours and 4 hours run time respectively implemented on R V3.0 with a PC of 3.1GHz Intel Core i5-2400 Processor.

5.2. Distributed Navigation Scheme

Consider a MRWSN in a distributed fashion in which at time instant t every individual mobile agent can transmit data to and receive information from its neighbouring sensors. There is no central station in this case; each robotic sensor itself computes prediction of the field of interest and finds the optimal sampling path by employing its own and its neighbours' measurements as well as exchanging information with the neighbours. The data set in this consideration is derived from a real application [52]. It has been known that nitrogen fertilizer applications are essential for production of crops in the agricultural field. Since nitrogen levels, supplied by mineralization of soil organic matter, can vary considerably across the field, soil sampling is crucial for the accurate description of spatial patterns in soil fertility. It was proposed utilizing the MRWSNs to construct the percentage of the soil organic matter patterns that recommend a strategy of nitrogen fertilizing.

The percentage of the soil organic matter was regenerated in [53] from the data of 1375 soil measurements. The size of this field is $400\text{m} \times 400\text{m}$. The true levels of the percentage of the soil organic matter are illustrated in Fig. 7. In this experiment, at the beginning, $N = 30$ mobile robotic wireless sensors were randomly deployed in the field as demonstrated as white dots in Fig. 9a. The maximum distance that each mobile agent can travel between time steps was set to 15m. The communication range R was set to 100m. The minimum distance between any two robotic sensors at any time was $\omega = 3\text{m}$. It was assumed that σ_ϵ^2 was known and set to 0.01.

In the proposed method, it is assumed that each mobile sensor can store all its own measurements from time step 1 to current time step t and transmit them to its neighbours when required. Therefore, at a particular time, every mobile sensing agent possesses not only its own measurements but also its neighbours' readings by only exchanging local information. Particularly, these collective measurements include the measurements gathered distantly from the sensor. This leads to the fact that the sensor can obtain measurements at locations even without navigating them. As a result, there are encouraging results demonstrated in Figures 6 and 9. For instance, in order to evaluate the distributed technique, the mean parameters and the hyperparameters of the GMRF model are estimated using both centralized and distributed methods and illustrated in

(a)

(b)

(c)

Fig. 8: The results of the agricultural field: The predicted fields at time (a) $t = 1$, (b) $t = 10$ and (c) $t = 20$. Ranges of the soil organic matter percentages are demonstrated in color bars.

(a)

(b)

(c)

Fig. 9: The results of the agricultural field: The predicted variances at time (a) $t = 1$, (b) $t = 10$ and (c) $t = 20$. Ranges of the errors are demonstrated in color bars. The sampling paths for MRWSs up to time steps 1, 5 and 10, respectively, are shown by white circles, where current robotic mobile sensor locations are shown in white dots.

Fig. 10: The root mean square errors for the agricultural field.

Fig. 6. It can be clearly seen that the global model parameters obtained by the distributed method approach those obtained by the centralized method. In the left column of Fig. 9, the predicted fields are illustrated at particular times $t = 1, 5$, and 10 , respectively. The fields represent the percentage of the organic matter in the soil. When the MRWSs navigate throughout the environment, the predicted fields shown in Figures 8a, 8b, and 8c are gradually getting close to the true field shown in Fig. 7.

In order to find the new optimal sampling paths for the MRWSs, the centralized criterion demonstrated by (25) was solved in a distributed fashion. In other words, the centralized strategy (25) is replaced by the distributed scheme (26) so that each mobile sensor can itself find the next sensing location. Consequently, the trajectories of the sampling paths of the mobile robotic wireless sensors at time steps 1, 5 and 10 are demonstrated by white circles in Figures 9a, 9b, and 9c, where the current sensor locations are shown in white dots. Although each mobile sensing agent cannot communicate to other sensors outside of its communication range, the mobile sensors adjust themselves according to the proposed strategy so that their measured locations cover the whole space as time accumulates. Moreover, the prediction error variances in the whole environment are also described in the right column of Fig. 9, where Figures 9a, 9a, and 9c correspond to 1^{st} , 5^{th} , and 10^{th} sampling instants. In accordance with the color bars, in the vicinity of the measured locations, the predicted variances are small compared with those at distant points. Particularly, it can be seen that the prediction errors in the whole space gradually reduce when more observations are collected.

In this experiment, for comparison purposes, the centralized standard GP and the centralized GMRF methods were also implemented in the reference field to compute the RMSEs. Notice that due to using all measurements to continuously model the physical field, the centralized GP algorithm should show the best performance in all cases. The results of the RMSEs obtained by the centralized and distributed GMRF approaches and the centralized standard GP method are shown in Fig. 10. Generally speaking, the centralized GMRF approach outperforms the distributed GMRF method. In addition, the RMSEs obtained by the proposed approaches are not much different from those obtained by the centralized standard GP. More importantly, in regards to computational time of addressing the sampling path problem (26) to find the next optimal sampling

location, each mobile robotic sensor can successfully solve (26) in approximately one second.

6. Conclusions

The paper has proposed to utilize the MRWSN to observe and then predict the spatial environmental field of interest using the GMRF model presented on a triangulation of an irregular discrete lattice, which not only allows the parameters and hyperparameters of the model to be updated online but also contributes to significant reduction in computational time owing to the sparse structure of the precision matrix. An optimal adaptive sampling strategy was proposed based on minimizing the conditional entropy of random variables at unobserved locations, which aims to drive the mobile robotic sensors on the most informative sampling paths. The adaptive sampling optimization problem was proven to be tractably solved in a polynomial time. The capability of the MRWSN to effectively predict the spatial environmental field and informatively navigate through the environment was investigated in both centralized and distributed fashions. The proposed algorithms based on the approximate GMRF model were demonstrated with convincing results as compared with those obtained by the counterpart GP model. Future work will consider the implementation of the proposed approach in a realistic network of mobile robotic sensors.

References

1. N. E. Leonard, D. Paley, F. Lekien, R. Sepulchre, D. M. Fratantoni, and R. Davis, "Collective motion, sensor networks and ocean sampling," *Proceedings of the IEEE*, vol. 95(1), pp. 48–74, 2007.
2. M. Mysorewala, D. Popa, and F. Lewis, "Multi-scale adaptive sampling with mobile agents for mapping of forest fires," *Journal of Intelligent and Robotic Systems*, vol. 54(4), pp. 535–565, 2009.
3. D. V. Le, H. Oh, and S. Yoon, "Environment learning-based coverage maximization with connectivity constraints in mobile sensor networks," *IEEE Sensors Journal*, vol. 16(10), pp. 3958–3971, 2016.
4. S. Susca, F. Bullo, and S. Martínez, "Monitoring environmental boundaries with a robotic sensor network," *IEEE Transactions on Control Systems Technology*, vol. 16, pp. 288–296, 2008.
5. M. Jadhaliha and J. Choi, "Environmental monitoring using autonomous aquatic robots: Sampling algorithms and experiments," *IEEE Transactions on Control Systems Technology*, vol. 21, pp. 899–905, 2013.
6. J. Choi, J. Lee, and S. Oh, "Swarm intelligence for achieving the global maximum using spatio-temporal Gaussian processes," in *Proc. American Control Conference*, Washington, USA, June 2008, pp. 135–140.
7. R. Graham and J. Cortes, "A cooperative deployment strategy for optimal sampling in spatiotemporal estimation," in *Proc. IEEE Conference on Decision and Control*, Cancun, Mexico, 2008, pp. 2432–2437.
8. —, "Distributed sampling of random fields with unknown covariance," in *Proc. American Control Conference*, MI, USA, 2009, pp. 4543–4548.
9. —, "Cooperative adaptive sampling via approximate entropy maximization," in *Proc. IEEE Conference on Decision and Control*, Shanghai, China, 2009, pp. 7055–7060.
10. J. Cortés, "Distributed Kriged Kalman filter for spatial estimation," *IEEE Transactions on Automatic Control*, vol. 54, pp. 2816–2827, 2009.
11. J. Cortes, S. Martínez, T. Karatas, and F. Bullo, "Coverage control for mobile sensing networks," *IEEE Transactions on Robotics and Automation*, vol. 20, pp. 243–254, 2004.
12. S. Martínez, "Distributed interpolation schemes for field estimation by mobile sensor networks," *IEEE Transactions on Control Systems Technology*, vol. 18, pp. 491–500, 2010.
13. J. Binney, A. Krause, and G. S. Sukhatme, "Optimizing waypoints for monitoring spatiotemporal phenomena," *The International of Robotics Research*, vol. 32(8), pp. 873–888, 2013.
14. N. A. Cressie, *Statistics for spatial data*. Wiley, 1991.
15. C. E. Rasmussen and C. K. I. Williams, *Gaussian processes for machine learning*. The MIT Press, Cambridge, Massachusetts, London, England, 2006.
16. Y. Xu and J. Choi, "Adaptive sampling for learning Gaussian processes using mobile sensor networks," *Sensors*, vol. 11, pp. 3051–3066, 2011.
17. Y. Xu, J. Choi, S. Dass, and T. Maiti, "Sequential Bayesian prediction and adaptive sampling algorithms for mobile sensor networks," *IEEE Transactions on Automatic Control*, vol. 57, pp. 2078–2084, 2012.
18. Y. Xu, J. Choi, and S. Oh, "Mobile sensor network navigation using Gaussian processes with truncated observations," *IEEE Transactions on Robotics*, vol. 27, pp. 1118–1131, 2011.

19. S. Huang and J. Tan, "Adaptive sampling using mobile sensor networks," in *Proc. IEEE International Conference on Robotics and Automation*, Minnesota, USA, May 2012, pp. 657–662.
20. R. Marchant and F. Ramos, "Bayesian optimization for intelligent environmental monitoring," in *Proc. IEEE/RSJ International Conference on Intelligent Robots and Systems*, Algarve, Portugal, October 2012, pp. 2242–2249.
21. L. V. Nguyen, S. Kodagoda, R. Ranasinghe, and G. Dissanayake, "Locational optimization based sensor placement for monitoring gaussian processes modeled spatial phenomena," in *Proc. 8th IEEE International Conference on Industrial Electronics and Applications*, Melbourne, Australia, June 2013, pp. 1706–1711.
22. —, "Spatially-distributed prediction with mobile robotic wireless sensor networks," in *Proc. 2014 13th IEEE International Conference on Control, Automation, Robotics and Vision*, Marina Bay Sands, Singapore, December 2014, pp. 1153–1158.
23. —, "Information-driven adaptive sampling strategy for mobile robotic wireless sensor network," *IEEE Transaction on Control Systems Technology*, vol. 24(1), pp. 372–379, 2016.
24. —, "Adaptive placement for mobile sensors in spatial prediction under locational errors," *IEEE Sensors Journal*, vol. 17(3), pp. 794 – 802, 2017.
25. C. K. I. Williams and M. Seeger, "Using the Nyström method to speed up kernel machines," *Advances in Neural Information Processing Systems*, vol. 13, 2001.
26. A. J. Smola and P. Bartlett, "Sparse greedy Gaussian process regression," *Advances in Neural Information Processing Systems*, vol. 13, 2001.
27. N. Lawrence, M. Seeger, and R. Herbrich, "Fast sparse Gaussian process methods: The informative vector machine," in *Proc. 16th Annual Conference on Neural Information Processing Systems*, Vancouver, BC, USA, 2003, pp. 609–616.
28. H. Rue and L. Held, *Gaussian Markov random field: Theory and applications*. Chapman and Hall - CRC Press, London, England, 2005.
29. H. Rue and H. Tjelmeland, "Fitting Gaussian Markov random fields to Gaussian fields," *Scandinavian Journal of Statistics*, vol. 29, pp. 31–49, 2002.
30. N. Cressie and N. Verzele, "Conditional-mean least-squares fitting of Gaussian Markov random fields to Gaussian fields," *Computational Statistics and Data Analysis*, vol. 52, pp. 2794–2807, 2008.
31. L. Hartman and O. Hössjer, "Fast kriging of large data sets with Gaussian Markov random fields," *Computational Statistics and Data Analysis*, vol. 52, pp. 2331–2349, 2008.
32. F. Lindgren, H. Rue, and J. Lindström, "An explicit link between Gaussian fields and Gaussian Markov random fields: The stochastic partial differential equation approach," *Journal of the Royal Statistical Society (Series B)*, vol. 73, pp. 423–498, 2011.
33. J. L. Ny and G. J. Pappas, "On trajectory optimization for active sensing in Gaussian process models," in *Proc. Joint 48th IEEE Conference on Decision and Control and 28th Chinese Control Conference*, Shanghai, China, 2009, pp. 6286–6292.
34. L. V. Nguyen, S. Kodagoda, R. Ranasinghe, and G. Dissanayake, "Mobile robotic wireless sensor networks for efficient spatial prediction," in *Proc. 2014 IEEE/RSJ International Conference on Intelligent Robots and Systems*, Chicago, IL, USA, September 2014, pp. 1176–1181.
35. L. V. Nguyen, S. Kodagoda, and R. Ranasinghe, "Spatial sensor selection via Gaussian Markov random fields," *IEEE Transaction on Systems, Man, and Cybernetics: Systems*, vol. 46(9), pp. 1226–1239, 2016.
36. Y. Xu, J. Choi, S. Dass, and R. Maiti, "Efficient Bayesian spatial prediction with mobile sensor networks using Gaussian Markov random fields," *Automatica*, vol. 49, pp. 3520–3530, 2013.
37. M. Jadaliha and J. Choi, "Fully Bayesian simultaneous localization and spatial prediction using Gaussian Markov random fields," in *Proc. American Control Conference*, Washington DC, USA, June 2013, pp. 4599–4604.
38. M. Jadaliha, Y. Xu, and J. Choi, "Efficient spatial prediction using Gaussian Markov random fields under uncertain localization," in *Proc. ASME Dynamic Systems and Control Conference*, Florida, USA, October 2012.
39. Y. Xu and J. Choi, "Spatial prediction with mobile sensor networks using Gaussian processes with built-in Gaussian Markov random fields," *Automatica*, vol. 48, pp. 1735–1740, 2012.
40. A. Singh, A. Krause, C. Guestrin, and W. J. Kaiser, "Efficient informative sensing using multiple robots," *Journal of Artificial Intelligence Research*, vol. 34, pp. 707–755, 2009.
41. C. M. Bishop, *Pattern recognition and machine learning*. Springer, New York, USA, 2006.
42. D. S. Bernstein, *Matrix mathematics: Theory, facts, and formulas with application to linear systems theory*. Princeton University Press, 2005.
43. H. Rue, S. Martino, and N. Chopin, "Approximate Bayesian inference for latent Gaussian models by using integrated nested Laplace approximations," *Journal of the Royal Statistical Society (Series B)*, vol. 71, pp. 319–392, 2009.
44. S. Banerjee, B. Carlin, and A. Gelfand, *Hierarchical modeling and analysis for spatial data*. Chapman and Hall, New York, 2004.
45. R. Olfati-Saber, J. A. Fax, and R. M. Murray, "Consensus and cooperation in networked multi-agent systems," *Proceedings of the IEEE*, vol. 95, pp. 215–233, 2007.
46. P. J. Diggle and P. J. Ribeiro, *Model-based geostatistics*. Springer, New York, USA, 2007.
47. M. West and J. Harrison, *Bayesian forecasting and dynamic models*. Springer, New York, 1999.

48. T. M. Cover and J. A. Thomas, *Elements of information theory*. John Wiley and Sons, Lnc, Hoboken, New Jersey, 2006.
49. C. Ko, J. Lee, and M. Queyranne, “An exact algorithm for maximum entropy sampling,” *Operations Research*, vol. 43, pp. 684–691, 1995.
50. L. Nguyen, G. Hu, and C. J. Spanos, “Efficient sensor deployments for spatio-temporal environmental monitoring,” *IEEE Transaction on Systems, Man, and Cybernetics: Systems*, vol. Early Access, pp. 1–11, 2018.
51. M. L. Stein, Z. Chi, and L. J. Welty, “Approximating likelihood for large spatial data sets,” *Journal of the Royal Statistical Society (Series B)*, vol. 66, pp. 275–296, 2004.
52. D. Mulla, M. Beatty, and A. Sekely, “Evaluation of remote sensing and targeted soil sampling for variable rate application of nitrogen,” in *Proc. 5th International Conference on Precision Agriculture*, America, 2001.
53. P. Tokekar, J. V. Hook, D. Mulla, and V. Isler, “Sensor planning for a symbiotic uav and ugv system for precision agriculture,” *IEEE Transactions on Robotics*, vol. 32, pp. 1498–1511, 2016.