Proceedings of Australasian Conference on Robotics and Automation, 3-5 Dec 2012, Victoria University of Wellington, New Zealand.

Sensor Selection Based Routing for Monitoring Gaussian Processes Modeled Spatial Phenomena

Linh Van Nguyen, Ravindra Ranasinghe, Sarath Kodagoda, Gamini Dissanayake
Center for Autonomous Systems
Faculty of Engineering and IT
University of Technology, Sydney, Australia
van.l.nguyen-1@student.uts.edu.au
{ravindra.ranasinghe, sarath.kodagoda, gamini.dissanayake}@uts.edu.au

Abstract

This paper addresses the trade-off between the sensing quality and the energy consumption in the wireless sensor network associated with monitoring spatial phenomena. We use a non-parametric Gaussian Process to model the spatial phenomena to be monitored and simulated annealing based approximately heuristic algorithm for sensor selection. Our novel Sensor Selection based Routing (SSR) algorithm uses this model to identify the most informative nodes, which gives the root mean square prediction error less than a specified threshold, to construct the minimal energy-expended routing tree rooted at the sink. Our experiments have verified that the proposed computationally efficient SSR algorithm has significant advantages over conventional techniques.

1 Introduction

Recent advances in micro-electro-mechanical systems and wireless communications empower wireless sensor networks (WSNs) to play a key role in a wide range of modern smart technology applications [Akyildiz *et al.*, 2002]. There is a growing momentum in the use of WSNs to monitor spatial phenomena such as temperature, humidity or rainfall particularly as a result of the substantial developments in reducing the size and the cost of wireless sensor nodes (aka motes). Nevertheless, dense deployment of WSNs is still practically prohibitive in many applications because of several key contraints. Among the most pertinent constraints, energy scarcity of motes is considered as the prominent constraint that affects the lifetime of a WSN. Furthermore, as a direct result of multiple motes co-located within the vicinity of a phenomenon in a dense WSN may generate similar data samples, there is potential to exist a sizable redundancy in sensed data traffic under utilizing the limited network bandwidth [Akkaya and Younis, 2005]. This not only adversely affects on the overall WSN lifetime but also on "goodput" (i.e. effective usage to provide the user with better intelligence about the phenomenon) of the WSN as the redundant samples do not contribute to gain any additional information about the phenomenon. Therefore, in a dense WSN, selection of the most important subset of nodes for sensing the spatial phenomenon together with the intermediate relay nodes at any given point of time is an important fundamental problem to obtain a high quality prediction of the phenomena over a longer period of time. The key objective of this work is to select the best k out of n possible potential sensor nodes and use these k sensor observations to predict about a spatial phenomena, while concurrently creating an optimal routing structure to minimize the total energy expended on the network.

A considerable amount of research has already been carried out focusing only on the WSN sensor selection problem. Cressie [Cressie, 1991] proposed a Gaussian Process (GP) model, a non-parametric generalization of linear regression, for spatial applications. It is feasible to exhaustively learn a GP model from the sensors readings and subsequently utilize this model to predict about the physical phenomenon at unobserved locations. Based on information-theoretic models, i.e. entropy and mutual information, [Caselton and Zidek, 1984], [Ko *et al.* , 1995], and [Guestrin *et al.*, 2005] developed greedy heuristic algorithms that guarantee near-optimal solutions for GP model based spatial applications. The above mentioned methods estimate the prediction error indirectly. In [Nguyen *et al.*, 2012], a computationally efficient simulated annealing based approximately heuristic algorithm was proposed to solve the sensor selection problem.

However, most of the previous works for sensor selection problem proposed solutions without taking the multi-hop routing characteristics that directly impact on the lifespan of the WSN into consideration. In previous works, there is usually a single constraint such as a constraint on the available bandwidth or a number of

active sensor or limit on the total number of message transmissions [Krause *et al.* , 2011]. Although limiting the number of selected sensors based on different constraints leads to lower the energy consumption, there is no guarantee that the selected subset is the best possible locations to ensure optimal sensor network lifetime.

It can be clearly seen that transmitting measurements directly to a center node (sink) may not be power efficient and practical. Some algorithms based on shortest path tree were proposed to create optimized multi-hop routing paths on the network to send sensor readings from sensor nodes to the sink [Akkaya and Younis, 2005]. These solutions associate with two issues: (a) the less important (less informative) nodes send their observations to the sink, (b) the more important nodes may consume more energy on relaying the data of the less important nodes.

The authors in [Shah and Beferull-Lozano, 2012] and [Portu-Repolles and Beferull-Lozano, 2009] proposed to join sensor selection and routing problems for distibuted solution. The premise behind these approaches is that every node can fuse all measurements that it receives from its children and its reading to compute an estimate for the physical phenomenon in a distributed manner. After that nodes transmit their estimated results to their parents. Nonetheless, they only utilize the linear models within certain limitation on estimation of the parameters of these models. This method is unsuitable for GP model based prediction tasks where all the measurements are necessary before performing the estimation. The incorporation of routing and sensor selection to address the prediction tasks in spatial phenomena modeled by GP, to the best of our knowledge, has not been resolved yet in the literature.

In this paper, we propose a sensor selection based routing algorithm (SSR) to address the problem stated above. This novel algorithm starts by using the algorithm proposed in our previous work [Nguyen *et al.*, 2012] to select the best informative subset of sensors. After that, sensor nodes measurements are routed sequentially to the sink from the most informative node to the least informative node using the optimal routing tree. Note that the optimal routing tree is built to minimize the overall energy consumption. Experimental results illustrates a considerable performance improvements of this algorithm compared to the others we studied in this work.

2 Problem Statement

Consider a field of n wireless sensors measuring some physical phenomenon. One special sink node is deployed deterministically at the center of the field to receive sensor measurement updates from other nodes. We further assume that the sink node has more resources to fuse the sensor measurements it receives from other nodes.

In this study we use a network of wireless sensors to monitor an indoor temperature field. We choose a finite subset of cardinality k_I sensors out of n possible sensors to develop a model for the temperature at unobserved locations. Furthermore we assume that we may need k_R number of additional nodes to create a suitable routing structure to reliably transport sensor measurements to the sink. While the energy consumption decrease with decreasing number of selected nodes (i.e. $k_I + k_R$), it has a significant impact on the sensing quality which is defined as the accuracy of prediction. This conflicting situation motivates us to carry out the multi-objective optimization problem to compromise these two conflicting metrics.

2.1 Network Formulation

Let us model the WSN comprising of n sensors as an undirected network connectivity graph $G = (W, E, A)$, where $W = \{w_1, w_2, ..., w_n\}$, is a set of vertices corresponding to the locations of sensors, E is a set of edges corresponding to one to one orthogonal communication links, and A is an adjacency matrix with following property: for each $i, j \in \{1, 2, ..., n\}$, the entry $a_{i,j} > 0$ if (w_i, w_j) is an edge of G, and $a_{i,j} = 0$ otherwise. In our network model, we suppose that each sensor has a communication range. That is a sensor can communicate with its neighbors only if its neighbors are inside its communication range. We also assume that the interference between nodes is insignificant and there exists a Medium Access Control (MAC) protocol to address the collisions on the network.

In order to send the sensor measurements from nodes to the sink, a routing structure mus be established. Every node will transmit its measurements through a multi-hop path specified by the routing structure. The amount of energy consumed at the physical layer to transmit a measurement sample from sensor i to sensor j is formulated as follows [Melodia *et al.*, 2004]:

$$E_{ij} = E^{tran} + E^{rec} + \beta.d_{ij}^{\alpha}, \tag{1}$$

where

- $E^{tran}(E^{rec})$ is the energy expended by transmitter (receiver),
- d_{ij} is the Euclidean distance between source node (i) and destination node (j),
- α is the path loss exponent, between 2 and 4 [Rappaport, 1999]. In this paper, we choose $\alpha = 2$ for free space propagation,
- β is a constant $[J/(bits.m^{\alpha})]$, and we choose $\beta = 100$ $[J/(bits.m^{\alpha})]$, [Melodia *et al.*, 2004] in this paper.

We refer E_{ij} (1) as the link metric.

The authors in [Akkaya and Younis, 2005] showed that $E^{tran}(E^{rec})$ is negligible as compared with the energy utilized by communication. Hence, we simplify the link metric to

$$E_{ij} = \beta.d_{ij}^{\alpha}. \tag{2}$$

This is a non-decreasing function of the distance d_{ij} between two adjacent communicating nodes. This results is experimentally supported by [Krishnamachari, 2005] in his work.

2.2 Sensor Selection Formulation

Consider a two dimensional sensor network that has a $\mathcal{V}$, $|\mathcal{V}| = n$, set of possible locations, which provide point measurements of some physical quantities. Sensor selection addresses the problem of choosing a subset, S, which can still represent the distribution of the physical quantity in the whole space. This can be effectively predicted using Gaussian Process (GP) [Rasmussen and Williams, 2006]. Lets denote $\mathcal{V} = [v_1, v_2, \ldots, v_n]$ as a set of locations, and $\mathcal{Z}_{\mathcal{V}} = [z_1, z_2, \ldots, z_n]$ as corresponding random variables at these locations. A joint probability distribution is given by

$$P(Z_{\mathcal{V}} = \mathbf{z}_{\mathcal{V}}) = \frac{1}{(2\pi)^{n/2}\Sigma_{\mathcal{V}\mathcal{V}}} e^{-\frac{1}{2}(\mathbf{z}_{\mathcal{V}}-\mathbf{m}_{\mathcal{V}})^T\Sigma_{\mathcal{V}\mathcal{V}}^{-1}(\mathbf{z}_{\mathcal{V}}-\mathbf{m}_{\mathcal{V}})}, \tag{3}$$

where $\mathbf{m}_{\mathcal{V}}$ is the mean vector and $\Sigma_{\mathcal{V}\mathcal{V}}$ is the covariance matrix of random variables, $\mathcal{Z}_{\mathcal{V}}$.

Rasmussen et al. [Rasmussen and Williams, 2006] demonstrated that GP has a marginalization property, which implies for any subset, $\mathcal{S}$, of $\mathcal{V}$, the joint distribution on random variables of its locations is Gaussian. Moreover, GP is comprehensively specified by its mean function $\mathcal{M}(v) = \mathbb{E}[z(v)]$, and a symmetric positive definite covariance function $\mathcal{C}(v_1, v_2)$, often called kernel function. One of frequently used kernel function is squared exponential, i.e.,

$$\mathcal{C}(v_1, v_2) = \sigma_f^2 e^{-\frac{\|v_1-v_2\|^2}{2l^2}}, \tag{4}$$

where $v_1, v_2 \in \mathcal{V}$, σ_f^2 is the maximum allowable covariance, l is the bandwidth of the kernel.

We define $\mathcal{S} \subset \mathcal{V}$ as a subset that includes all the selected sensors locations and its cardinality is predefined as k_I. For any subset $\mathcal{S}$, let $\mathcal{Z}_{\mathcal{S}}$ denote the collections of observations at locations in $\mathcal{S}$. In addition, take into account $\mathcal{U} = \mathcal{V} \setminus \mathcal{S}$ as the set of all elements in $\mathcal{V}$ but not in $\mathcal{S}$, and $\mathcal{Z}_{\mathcal{U}}$ is a vector of random variables over these unobserved locations. We will assume that sensor measurement has an additive independent identically distributed zero-mean Gaussian noise with variance σ_n^2.

It can be clearly shown that $\mathcal{Z}_{\mathcal{S}}$ and $\mathcal{Z}_{\mathcal{U}}$ are jointly Gaussian distributed as

$$\begin{bmatrix}\mathcal{Z}_{\mathcal{S}} \\ \mathcal{Z}_{\mathcal{U}}\end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix}\mathbf{m}_{\mathcal{S}} \\ \mathbf{m}_{\mathcal{U}}\end{bmatrix}, \begin{bmatrix}\Sigma_{\mathcal{S}\mathcal{S}} + \sigma_n^2 I & \Sigma_{\mathcal{S}\mathcal{U}} \\ \Sigma_{\mathcal{U}\mathcal{S}} & \Sigma_{\mathcal{U}\mathcal{U}}\end{bmatrix}\right), \tag{5}$$

where $\mathbf{m}_{\mathcal{S}}$ and $\mathbf{m}_{\mathcal{U}}$ ($\Sigma_{\mathcal{S}\mathcal{S}}$ and $\Sigma_{\mathcal{U}\mathcal{U}}$) are mean vectors (covariance matrices) of $\mathcal{Z}_{\mathcal{S}}$ and $\mathcal{Z}_{\mathcal{U}}$, respectively. $\Sigma_{\mathcal{U}\mathcal{S}} (= \Sigma_{\mathcal{S}\mathcal{U}}^T)$, are cross-covariance matrices between $\mathcal{Z}_{\mathcal{S}}$ and $\mathcal{Z}_{\mathcal{U}}$; and I is the $|\mathcal{S}| \times |\mathcal{S}|$ identity matrix.

In probabilistic terms, we derive the conditional distribution at predicted positions of $\mathcal{U}$, given $\mathcal{Z}_{\mathcal{S}}$ as follows:

$$\mathbf{m}_{\mathcal{U}|\mathcal{S}} = \mathbf{m}_{\mathcal{U}} + \Sigma_{\mathcal{U}\mathcal{S}}(\Sigma_{\mathcal{S}\mathcal{S}} + \sigma_n^2 I)^{-1}(\mathcal{Z}_{\mathcal{S}} - \mathbf{m}_{\mathcal{S}}), \tag{6}$$

$$\Sigma_{\mathcal{U}|\mathcal{S}} = \Sigma_{\mathcal{U}\mathcal{U}} - \Sigma_{\mathcal{U}\mathcal{S}}(\Sigma_{\mathcal{S}\mathcal{S}} + \sigma_n^2 I)^{-1}\Sigma_{\mathcal{S}\mathcal{U}}, \tag{7}$$

where $\mathbf{m}_{\mathcal{U}|\mathcal{S}}$ and $\Sigma_{\mathcal{U}|\mathcal{S}}$ are mean vector and covariance matrix of $\mathcal{Z}_{\mathcal{U}}$, given $\mathcal{Z}_{\mathcal{S}}$. As a consequence, using observations at locations in set $\mathcal{S}$, we can predict quantities at unobserved locations, $\mathcal{U}$. This process is described as the Gaussian regression approach [Bishop, 2006].

The quality of prediction is generally measured by calculating the errors at unobserved locations. Therefore, the goal is to select a subset $\mathcal{S} \subset \mathcal{V}$ so that it will minimize a certain measure of prediction error at unobserved locations, given observations $\mathcal{Z}_{\mathcal{S}}$. A typically used function of estimation error is the mean square error (MSE). Specifically, it is given by

$$\text{MSE}(\mathcal{U}|\mathcal{S}) = \mathbb{E}[(\widehat{\mathcal{Z}}_{\mathcal{U}|\mathcal{S}} - \mathcal{Z}_{\mathcal{U}|\mathcal{S}})^2]. \tag{8}$$

Assuming that to be an unbiased estimator (which is the case in GP regression), the mean square error can be estimated to be the variance of $\widehat{\mathcal{Z}}_{\mathcal{U}|\mathcal{S}}$ [Riley *et al.* , 2006]. Therefore, $\text{MSE}(\mathcal{U}|\mathcal{S})$ can be calculated by

$$\text{MSE}(\mathcal{U}|\mathcal{S}) = \text{trace}(\Sigma_{\mathcal{U}\mathcal{U}} - \Sigma_{\mathcal{U}\mathcal{S}}(\Sigma_{\mathcal{S}\mathcal{S}} + \sigma_n^2 I)^{-1}\Sigma_{\mathcal{S}\mathcal{U}}). \tag{9}$$

This equation can be reformulated to reflect the root mean square error (RMSE) for each predicted location. Further, we define an Average RMSE (ARMSE), by normalizing it with $|\mathcal{U}|$ i.e.,

$$\text{ARMSE}(\mathcal{U}|\mathcal{S}) = \\ = \sqrt{\frac{1}{|\mathcal{U}|}\text{trace}(\Sigma_{\mathcal{U}\mathcal{U}} - \Sigma_{\mathcal{U}\mathcal{S}}(\Sigma_{\mathcal{S}\mathcal{S}} + \sigma_n^2 I)^{-1}\Sigma_{\mathcal{S}\mathcal{U}})}. \tag{10}$$

2.3 Sensor Selection Based Routing Optimization Problem

It is clear from the above discussion that the quality of sensing grows when $\text{ARMSE}(\mathcal{U}|\mathcal{S})$ at unobserved locations reduce. Therefore, the actual problem is how to choose the best subset of sensors that corresponds to

the minimum consumption of energy and ARMSE($\mathcal{U}|\mathcal{S}$) values less than a specified threshold. The main idea behind this is to keep the best sensors alive longer.

The communication cost to transmit the data from node v to the sink can be formulated as follows:

$$\text{com}_v = \sum_{u \in path_v} \beta.(d_{u,p_u})^{\alpha} \tag{11}$$

where

$path_v$ is the multi-hop path in the routing structure, that is used to forward the measured sensor data from node v to the sink, $path_v \subset T \subset G$, T is non-spanning tree of G,

u is one of the nodes on the $path_v$,

p_u is the identity of the next hop sensor node of the sensor node u.

Therefore, we can now compute the total communication cost to transmit the data from the selected subset $\mathcal{S}$ to sink as:

$$\text{com}^{total}(\mathcal{S}) = \sum_{v \in \mathcal{S}} \text{com}_v = \sum_{v \in \mathcal{S}} \sum_{u \in path_v} \beta.(d_{u,p_u})^{\alpha} \tag{12}$$

Now, our objective is to select a subset $\mathcal{S}$ of k sensors from among a set $\mathcal{V}$ of n potential sensors, which minimizes the overall network communication cost, subject to the constraint of ARMSE($\mathcal{U}|\mathcal{S}$) is less than a specified threshold. Therefore, we define the objective function as:

$$\textbf{minimize} \quad \text{com}^{total}(\mathcal{S}) = \sum_{v \in \mathcal{S}} \sum_{u \in path_v} \beta.(d_{u,p_u})^{\alpha} \tag{13}$$

$$\begin{aligned} \textbf{subject to} \ & \text{ARMSE}(\mathcal{U}|\mathcal{S}) \leqslant \epsilon \\ & \mathcal{S} \subset \mathcal{V}, \\ & path_v \subset T, \\ & T \subset G, \end{aligned}$$

where ϵ is the threshold (for example, ϵ can be chosen to ensure an acceptable indoor thermal comfort level)

The constraint ensures that the sensing quality is to be controlled as per the application requirement. During this optimization, some sensors might not be chosen to generate sensor measurements, but they may be still utilized as relaying nodes. The following section presents an efficient approximation algorithm to solve this minimization problem.

3 Sensor Selection Based Routing Algorithm

Before presenting the algorithm, let us define an important concept called the "transmission round" that is required to fairly assess the algorithm. Intuitively this resembles with the cyclic nature of the sensor measurement delivery process in real world applications for monitoring spatial phenomena [Akkaya and Younis, 2005].

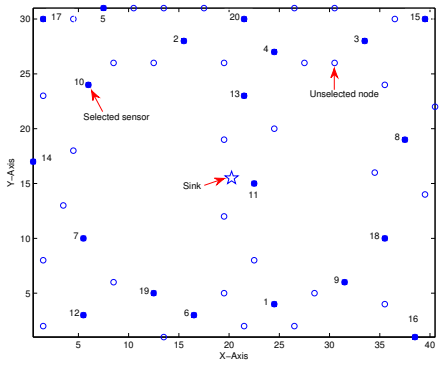

Figure 1: 20 out of 52 sensors chosen by simulated annealing based algorithm and ordered through information, deployed in Intel Berkeley Lab: Square indicates the locations selected; circle indicated the locations unselected

*Definition*1 : We define a transmission round as the time taken to transmit a plurality of sensor measurement data messages back to back from sensor nodes to the sink until ARMSE($\mathcal{U}|\mathcal{S}$) is less than the specified value ϵ. Note that the sink computes the value of ARMSE($\mathcal{U}|\mathcal{S}$) on arrival of each sensor measurement message.

Following Definition 1, in each transmission round, we attempt (a) to solve the optimization problem (13) (b) to derive an optimal routing structure.

The key idea of the algorithm is as follows.

First of all, we attempt to identify the best subset of k_I out of n potential sensor locations, which minimizes the ARMSE($\mathcal{U}|\mathcal{S}$). To resolve this problem, great deal of approaches has been proposed during last two decades. Two well-known greedy techniques based on information-theoretic models were proposed in [Caselton and Zidek, 1984], [Ko *et al.* , 1995], and [Guestrin *et al.*, 2005] to solve this problem. In [Nguyen *et al.*, 2012], we proposed a computationally efficient algorithm based on simulated annealing to address this sensor selection problem.

After choosing the subset of sensors, we order the sensor locations, from the most informative node to the least informative node. For example, as illustrated in Figure 1, from 52 sensors deployed in Intel Bekerley Lab (an indoor temperature distribution dataset from the Intel Berkeley Research Lab [Bodik et al. , 2004]), using our simulated annealing based algorithm, we selected a sensor subset of cardinality 20. Then, these 20 sensor locations were arranged from the most informative node to the least informative node.

We then generate a connectivity graph, G, rooted at

the sink. Using this graph G we have the complete picture of connectivity between adjacent sensor nodes. In the connectivity graph G, each edge is described by the link metric that is computed by (2). We use these link metric values to form the adjacency matrix A.

At each round, the active nodes transmit their sensor measurement data in the order from the most informative to the least informative. These sensor measurement data messages traverse to the sink through the minimal energy-expended multi-hop paths of the graph G. In this work we used Dijkstra algorithm [Bullo *et al.* , 2009] to build shortest paths tree of routing structures.

When each of these sensor measurement messages arrive at the sink, ARMSE($\mathcal{U}|\mathcal{S}$) is computed and compared with the threshold ϵ. This algorithm recursively processes messages and computes the ARMSE($\mathcal{U}|\mathcal{S}$) until ARMSE($\mathcal{U}|\mathcal{S}$) is less than the specified threshold. Key steps of this procedure are described in Algorithm 1. Figure 2 illustrates the optimal routing tree on the WSN using the Intel Berkeley Lad dataset, prior to any sensor becomes inactive. There are 14 sensors chosen (blue square nodes) to make ARMSE($\mathcal{U}|\mathcal{S}$) less than the specified threshold. It can be clearly seen in Figure 2 that some other nodes are selected for the optimal routing tree as relay nodes.

Algorithm 1 Pseudocode for Sensor Selection based Routing (SSR)

1: Initialize: ARMSE($\mathcal{U}|\mathcal{S}$)=infinite, ϵ, $\mathcal{S} = \oslash$,$T = \oslash$,$i = 0$
2: Choose the best sensor subset that minimizes ARMSE($\mathcal{U}|\mathcal{S}$), given cardinality
3: Order sensor locations from the most informative node to the least informative node $\{s_1, s_2, ..., s_n\}$
4: Generate the connectivity graph rooted at the sink
5: **while** ARMSE($\mathcal{U}|\mathcal{S}$) $> \epsilon$ **do**
6: $i = i + 1$
7: $\mathcal{S} \leftarrow s_i$
8: $\mathcal{U} \leftarrow \mathcal{V} \setminus \mathcal{S}$
9: Calculate ARMSE($\mathcal{U}|\mathcal{S}$)
10: Find the shortest path $path_i$ from s_i to sink using Dijkstra function
11: $T \leftarrow path_i$
12: **end while**
13: return $\mathcal{S}$ and T
14: **if** o **then**ne node dies
15: Remove the died node from the network
16: Go to step 4
17: **end if**

At the end of each round, residual energy of each node is summarized. If this energy becomes smaller than the link metric between this node and any of its neighbors, this node will be considered to be dead and hence removed from the network. Once again the algorithm starts from step 4 to develop a new connectivity graph G before optimizing a new routing tree on the network. Figure 3 illustrates a scenario when one sensor was out of energy. There are still 14 sensors chosen to make ARMSE($\mathcal{U}|\mathcal{S}$) less than the specified threshold, with one sensor (upper-right corner) was added to the subset. Eventually, a new optimal routing tree was established for the network.

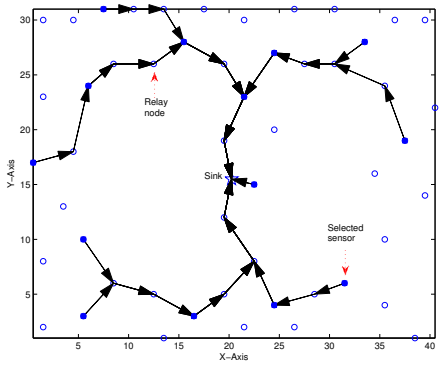

Figure 2: Sensors chosen and deployed in Intel Berkeley Lab with all sensor nodes, and optimal routing structure on the network: Square indicates the locations selected; circle indicates the locations unselected

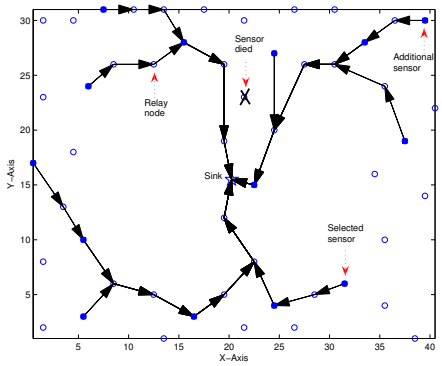

Figure 3: Sensors chosen and deployed in Intel Berkeley Lab when one sensor died, and optimal routing structure on the network: Square indicates the locations selected; circle indicates the locations unselected; ×-mark shows the sensor died

4 Results

In order to examine the behavior of this SSR algorithm, we carried out the following simulation experiments using the temperature dataset from the Intel Berkeley Research Lab Data Set [Bodik et al. , 2004]. We compared

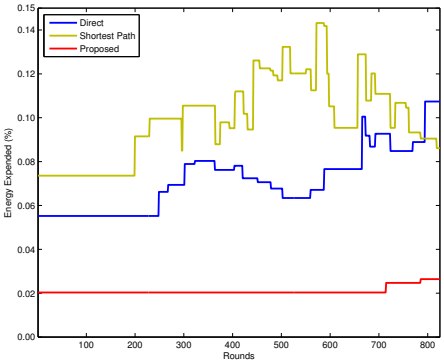

Figure 4: The percentage of total energy consumed on the network at every round

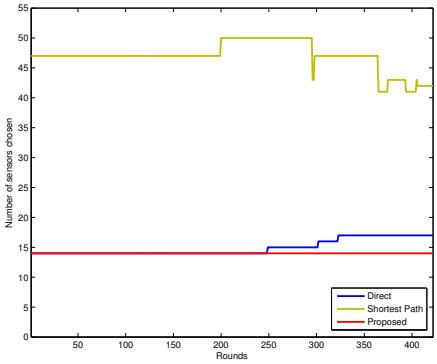

Figure 5: The number of sensors chosen to keep ARMSE($\mathcal{U}|\mathcal{S}$) under threshold in every round

our proposed approach with two other techniques: 1. Most informative subset of nodes transmitting directly to the sink (direct method), 2. Randomly selected subset transmitting to the sink using the shortest multi-hop paths (shortest path method).

At the end of every transmission round, we computed the percentage of energy expended, the number of sensor chosen, and the number of nodes died under Direct, Shortest Path and the proposed approaches.

Figure 4 shows the percentage of total energy expended by the nodes on the network through the round. It can be clearly seen from Figure 4 that the proposed method outperforms the other two methods with a fairly low and stable level of energy consumption throughout its operations. Note that the Direct and Shortest Path methods exhibit considerable fluctuations in the consumed energy. This is due to the fact that the selected nodes depletes their power at a faster rate under the Direct and Shortest Path methods causing to change the nodes of the selected set. Figure 5 shows the number of sensors chosen in every transmission round to keep the ARMSE($\mathcal{U}|\mathcal{S}$) value within the desired range. While the shortest path method queries almost all the sensors to transmit their data readings to the sink, the other two techniques utilize approximately one third of the potential measurements to obtain the sufficiently accurate prediction.

Figure 6 illustrates the total number of nodes died after every transmission round. After over 1000 transmission rounds, there are barely five nodes depleted under the proposed algorithm while there are approximately 49 and 36 depleted nodes under the Shortest Path and the Direct methods, respectively. This clearly demonstrates the sensor nodes have a significantly low lifetime under the Direct and Shortest Path methods.

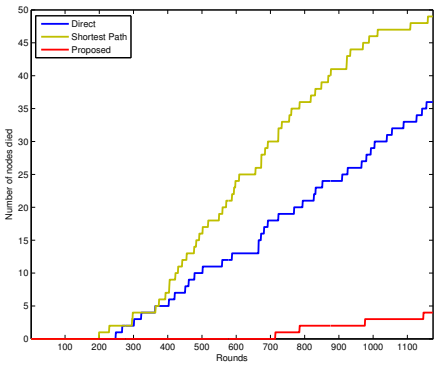

Figure 6: The number of nodes died after every round

5 Conclusions

This paper has presented an efficient novel algorithm to address the trade-off between sensing quality and energy consumption in wireless sensor networks associated with monitoring spatial phenomena by developing the routing structure based on selected most informative sensor subset. The criterion to optimize the overall communication cost is derived. The constraints to solve this optimization are based on the prediction accuracy that is linked to real world applications. It has been shown that using the limited number of more informative sensors only leads to considerable power savings.

References

[Akkaya and Younis, 2005] Kemal Akkaya and Mohamed Younis. A Survey on Routing Protocols for Wireless Sensor Networks. *Ad Hoc Networks*, 3:325–349, 2005.

[Akyildiz *et al.*, 2002] I. F. Akyildiz, W. Su, Y. Sankarasubramaiam, and E. Cayirci. Wireless Sensor Net-

works: A Survey. *Computer Networks*, 38:393–422, 2002.

[Bishop, 2006] Christopher M. Bishop. *Pattern Recognition and Machine Learning*. Springer, New York, USA, 2006.

[Bodik et al. , 2004] Peter Bodik, Carlos Guestrin, Wei Hong, Sam Madden, Mark Paskin, and Romain Thibaux. *http://db.csail.mit.edu/labdata/labdata.html*.

[Bullo *et al.* , 2009] Francesco Bullo, Jorge Cortes, and Sonia Martinez. *Distributed Control of Robotic Networks*. Princeton University Press, 2009.

[Caselton and Zidek, 1984] W. F. Caselton and J. V. Zidek. Optimal Monitoring Network Design. *Statistics and Probability Letters*, 2:223–227, 1984.

[Cressie, 1991] N. A. Cressie. *Statistics for Spatial Data*. Wiley, 1991.

[Guestrin *et al.*, 2005] C. Guestrin, A. Krause, and A. P. Singe. Near-optimal Sensor Placement in Gaussian Processes. In *Proceedings of Twenty-Second International Conference on Machine Learning* pages 265-272, Bonn, Germany, 2005. International Conference on Machine Learning.

[Ko *et al.* , 1995] C. Ko, J. Lee, and M. Queyranne. An Exact Algorithm for Maximum Entropy Sampling. *Operation Research*, 43:684–691, 1995.

[Krause *et al.* , 2011] Andreas Krause, Carlos Guestrin, Anupam Gupta, and Jon Kleinberg. Robust Sensor Placement at Informative and Communication-Efficient Locations. *ACM Transactions on Sensor Networks*, 7(31):1–33, 2011.

[Krishnamachari, 2005] Bhaskar Krishnamachari. *Networking Wireless Sensors*. Cambridge University Press, New York, USA, 2005.

[Melodia *et al.*, 2004] Tommaso Melodia, Dario Pompili, and Ian F. Akyildiz. Optimal local topology knowledge for energy efficient geographical routing in sensor networks. In *Proceedings of Twenty-third Annual Joint Conference of the IEEE Computer and Communications Societies*, pages 1705-1716, Hong Kong, November 2004. Annual Joint Conference of the IEEE Computer and Communications Societies.

[Nguyen *et al.*, 2012] Linh Nguyen, Sarath Kodagoda, Ravindra Ranasinghe, and Gamini Dissanayake. Simulated annealing based approach for near-optimal sensor selection in Gaussian Processes. *IEEE International Conference on Control, Automation and Information Science (ICCAIS 2012)* (accepted), Ho Chi Minh city, Vietnam, November 2012. International Conference on Control, Automation and Information Science.

[Portu-Repolles and Beferull-Lozano, 2009] Eduardo Portu-Repolles and Baltasar Beferull-Lozano. Joint Sensor Selection and Routing for Distributed Estimation in Wireless Sensor Networks. In *Proceedings of Seventeenth European Signal Processing Conference*, pages 774–778, Glasgow, Scotland, August 2009. European Signal Processing Conference.

[Rappaport, 1999] Theodore S. Rappaport. *Wireless Communication: Principles and Practice*. Prentice Hall, 1999.

[Rasmussen and Williams, 2006] Carl Edward Rasmussen and Christopher K. I. Williams. *Gaussian Processes for Machine Learning*. The MIT Press, 2006.

[Riley *et al.* , 2006] K. F. Riley, M. P. Hobson, and S. J. Bence. *Mathematical Methods for Physics and Engineering*. Cambridge University Press, 2006.

[Shah and Beferull-Lozano, 2012] Santosh Shah and Baltasar Beferull-Lozano. Power-aware Joint Sensor Selection and Routing for Distributed Estimation: A Convex Optimizing Approach. In *Proceedings of Eighth IEEE International Conference on Distributed Computing in Sensor Systems*, pages 230–238, Zhejiang, China, May 2012. International Conference on Distributed Computing in Sensor Systems.