Learning-based Adaptive Sensor Selection Framework for Multi-sensing WSN

Sushmita Ghosh, Swades De, Shouri Chatterjee, and Marius Portmann

***Abstract*— Wireless sensor nodes equipped with multiple sensors often have limited energy availability. To optimize the energy sustainability of such sensor hubs, in this paper a novel adaptive sensor selection framework is proposed. Multiple sensors monitoring different parameters in the same environment often possess cross-correlation, which makes the system predictive. To this end, a learning-based optimization strategy is developed using Upper Confidence Bound algorithm to select an optimum active sensor set in a measurement cycle based on the cross-correlations among the parameters, energy consumed by the sensors, and the energy available at the node. Further, a Gaussian process regressor-based prediction model is used to predict the parameter values of inactive sensors from the cross-correlated parameters of active sensors. To evaluate the performance of the proposed framework in real-life applications, an air pollution monitoring sensor node consisting of seven sensors is deployed in the campus that collects data at a default high sampling rate. Simulation results validate the efficiency and efficacy of the proposed framework. Compared to the current state-of-the-art the proposed algorithm is 54% more energy efficient, with complexity $\mathcal{O}(2^{\mathsf{P}})$ for P sensors in the node, while maintaining an acceptable range of sensing error.**

***Index Terms*— adaptive sensor selection framework, energy saving, Gaussian process regression model, Multi-parameter sensor hub, parameter cross-correlation, reinforcement learning, upper confidence bound algorithm.**

I. Introduction

Wireless Sensor Networks (WSNs) consist of a large number of sensor nodes. Most of them are powered by batteries which limit the lifetime of such networks. Two main energy consuming parts of WSNs are sensing and transmission. A sensor node consists of a set of sensors to monitor various parameters of the particular environment. As an example, an air pollution monitoring sensor node consists of sensors for monitoring several parameters, namely, temperature, humidity, particulate matter (PM), and various hazardous gases. The energy consumption of many good quality sensors is more than that of the transmission module. To solve the problem of limited battery capacity, sensor nodes are equipped with an energy harvesting mechanism for replenishing the batteries. For reduced energy consumption in sensing, an optimal sampling interval can be decided for each parameter without compromising on the quality of sensing [1]. Though such mechanisms increase the lifetime of WSNs, designing a sustainable wireless sensor node monitoring multiple parameters with less energy consumption and good sensing quality is still an open challenge.

In many WSNs field data is sent to the local collection point routinely, which can be analyzed to create a feedback mechanism to adapt the system parameters [2]. In an environment with multiple parameters of sensing interest, although the sensing elements may be very different based on their composition, these parameters often manifest cross-correlation among them, which can be exploited to estimate one parameter value from the other one or more parameters. Thus, by choosing an optimal number of active sensors an energy-efficient sensing process at a field node can be designed.

Sushmita Ghosh, Swades De, and Shouri Chatterjee are with the UQ-IITD Academy and Department of Electrical Engineering, Indian Institute of Technology Delhi, New Delhi, India (e-mail: qiz198439@uqidar.iitd.ac.in, swadesd@ee.iitd.ac.in, shouri@ee.iitd.ac.in)

Marius Portmann is with the School of IT and Electrical Engineering, University of Queensland, Brisbane, Australia (e-mail: marius@itee.uq.edu.au).

A. Related Work

Numerous works have been presented in the literature on energy-efficient sensing. These can be divided into two categories: a) adaptive sampling algorithm applied on each parameter at a sensor node independently [1], [3], [4], b) selecting an active set of sensors based on the spatio-temporal correlation of data among multiple sensor nodes, monitoring one parameter in a densely-deployed WSN [5]–[7]. Compared to traditional fixed-rate sampling, both the processes can reduce data volume and energy consumption in sensing and transmission.

An adaptive sampling algorithm was proposed for snow monitoring application in [4] that estimates the optimal sampling frequencies online for sensors by applying fast Fourier transform (FFT) on a sufficiently large sample set. As per Nyquist criteria, the minimum sampling rate can be calculated from the maximum detectable frequency of the sensed signal. When the frequency changes, a new sampling rate is calculated after detecting a user-defined W number of samples. Three different data collection mechanisms were proposed for temperature and humidity sensing independently in [1] to adapt the sampling rate with the variation of environment. In the first approach, the T-statistic value (ratio of the variance calculated based on the collected measurements) was computed using

one-way Anova model and Bartlett test, which is used in behavior function to find a new sampling rate. The second and third models used the Jaccard similarity function and Euclidean distance function respectively to find dissimilarity between two consecutive data-sets. Imposing new parameter values (obtained from these methods) on behavior function, new sampling rates were calculated for the two individual models. A decentralized approach of adaptive sampling was proposed in [3], that employs Kalman Filtering (KF) based estimation technique to autonomously adjust the sampling interval within a given range based on KF estimation error.

In a densely-deployed WSN, spatio-temporal correlations among the sensing signals are used to select a fraction of sensors to collect samples of a particular sensing parameter [6]. Compressive sensing was applied in [5], [8] for selecting fewer sensors from a large set of sensor nodes in a heterogeneous sensing environment. An optimization function was defined to select a particular set of sensors for the next measurement cycle considering sensing quality, process dynamics, and energy availability at the nodes.

In [9], a sequential Bayesian approach for spatio-temporal Gaussian process regression was formulated to develop an adaptive sampling algorithm for mobile sensor network. In [10], Gaussian process regression (GPR) was used to predict the mean and covariance functions of spatio-temporally varying signals, as in predicting environmental monitoring parameters with localization uncertainty in mobile wireless sensor networks. GPR has been widely used to draw statistical inference from environmental data [11]. GPR is used in [12], [13] to predict Particulate Matter (PM) from its historical values for low-cost PM monitoring.

Various regression and neural network-based techniques were proposed in [14]–[17] to predict the values of air pollution parameters, $PM_{2.5}$, PM_{10}, from their spatio-temporal correlations with the meteorological parameters: wind speed, temperature, and humidity.

An adaptive sensing framework, proposed in [18], applied reinforcement learning in mobile sensing environment to improve the energy efficiency. However, *the work does not deal with multi-parameter sensor hubs.*

B. Research Gap

Although the aforementioned approaches substantially reduce energy consumption in the WSNs, the limited battery capacity of the miniature sensor hub in a large-scale deployment restricts the lifetime of such networks. Moreover, with the development of good quality sensors, their cost and energy consumption are increasing, and the number of sensing elements attached to a single node is increasing as well. This in turn increases data volume and energy consumption of the sensor nodes.

The node-level adaptations in [1], [3], [4] are on individual sensing parameters; *they do not exploit multi-parameter cross-correlation.* The studies in [2], [5], [6], [9] consider densely-deployed WSNs for monitoring the same parameter. Because of spatio-temporal correlations, the data is sparse in nature. In contrast, *the process dynamics in multi-parameter sensing on a sensor hub is expected to be different.*

The studies in [14]–[17] target weather forecasting, rather than energy efficiency.

The use of GPR models on spatio-temporally varying signals in mobile sensing networks [10]–[13] with high accuracy prediction performance motivates its use in predicting multiple parameter values in a sensor node.

As noted above, *energy-efficient sensor data collection by exploiting the process dynamics of a multi-parameter sensor hub is yet to be studied in the literature.* Motivated by good prediction performance of GPR in [10]–[13], an energy-aware sensing mechanism is proposed in this work that activates online a set of sensors that minimizes prediction error while ensuring an upper bound on energy consumption. The upper bound is decided by the energy available at the node.

C. Motivation

The need for machine learning-based prediction models to optimize the energy efficiency in WSNs has been discussed in Section I-A. In recent studies, sensor nodes in a network are proposed to be selected by solving the trade-off between sensing/estimation error and energy consumption [5], [19]. However, these methods were developed for network-level adaptation, where the data collected at the present cycle is exploited to decide the optimal set of nodes to be activated in the next cycle based on the energy available at the nodes. In these formulations it is assumed that the signal is slowly varying, therefore, the next state can be predicted from the present state.

Apart from the network-level sensing, if a single node consists of multiple sensors to capture the randomly varying parameter values in the environment, the above formulation may not be accurate in choosing optimal set of sensors at a node. Therefore, in deciding the optimal sensor set in a measurement cycle, besides the data from the immediate past measurement cycle, history of the past cycles would be useful. To this end, a reinforcement learning model can be created where the system learns with experience. Based on the dynamics of the system, the sensor selection algorithm can be modeled as a multi-armed bandit (MAB) problem. Finding an optimal sensor set for the next cycle is analogous to finding an optimal arm for the next time slot. MAB is a reinforcement learning problem, used widely in many real-time systems for online decision making purpose. Various algorithms, such as Upper Confidence Bound (UCB) algorithm, Thompson sampling, epsilon-greedy algorithm, etc. are used to solve MAB problems. Among them, UCB or UCB1 algorithm is adapted in the proposed framework, as the regret or the loss suffered by the system for not choosing an optimal set is minimum in UCB. It is a deterministic algorithm that works based on the principle of optimism in the place of uncertainty. It is computationally less expensive and hence easy to implement.

D. Contributions

In this work, a new method of sensor data collection mechanism is proposed for a miniature sensor hub consisting of multiple sensors monitoring various parameters in the environment. The proposed approach is broadly studied for node-level analysis that can be used in a distributed network.

The key features and contributions of this work are as follows:

1) The proposed multi-sensing framework is modelled as a MAB problem that exploits cross-correlation among the various sensing parameters at a node to activate an optimal set of sensors while accounting for the non-stationarity of the environment.
2) A modified upper confidence bound (UCB) algorithm is developed for activation of an optimal set of sensors; its reward is calculated in every measurement cycle based on the cross-correlation among the parameters, sensing energy requirement, and the node-level energy availability.
3) The length of measurement cycle is adaptive to the process dynamics, which inherently exploits the variation of cross-correlation of the sensed processes.
4) A new method is proposed to adapt the length of measurement cycle based on a suitably developed parametric mapping function that maps the change in cross-correlation among the parameters to the length of next measurement cycle.
5) A GPR model is introduced to predict the missing parameters at the central entity (CE) from the sampled parameters based on cross-correlation among them.
6) The proposed algorithm is tested on real data-set from a pollution monitoring node deployed on the campus. Compared to the adaptive sampling algorithm proposed in [1], up to 54% energy saving can be achieved using the proposed data collection method while maintaining an acceptable range of sensing error.

Organization: The system model is presented in Section II, followed by a brief description on the GPR based prediction model in sensor data prediction in Section III. Section IV contains the proposed adaptive sensor selection algorithm. The experimental setup is explained Section V, followed by results and conclusion in Sections VI and VII, respectively.

Notations: $\mathcal{A}$ and $\underline{\mathbf{A}}$ respectively denote a set and a matrix. $\underline{\mathbf{Z}} \in \mathbb{R}^{P\times I}$ represents a real valued matrix of size $P \times I$, $\mathbf{z} \in \mathbb{R}^{P\times 1}$ denotes a vector having P elements and $|\mathcal{A}| = A$ denotes the cardinality of set $\mathcal{A}$.

II. System Model

Consider a sensor node having P sensors, each monitoring a parameter in the environment. Let $\mathcal{P} = \{P_p; 1 \leq p \leq P\}$ be the set containing all the parameters. In the x^{th} measurement cycle the sensed data at a sampling instant is denoted as [20]:

$$\mathbf{z}^x = \mathbf{y}^x + \boldsymbol{\eta}^x \tag{1}$$

where $\mathbf{z}^x \in \mathbb{R}^{P\times 1}$ is the measurements vector containing the observations of all the parameters. $\mathbf{y}^x \in \mathbb{R}^{P\times 1}$ is the true signal vector and $\boldsymbol{\eta}^x \in \mathbb{R}^{P\times 1}$ is the measurement noise vector associated with the P sensors. η for all the parameters are considered additive Gaussian, identically and independently distributed with zero mean and σ^2 variance. If I^x is the number of instances/samples collected for each parameter in x^{th} measurement cycle, then $\underline{\mathbf{Z}}^x \in \mathbb{R}^{P\times I^x}$ is the matrix containing the temporal measurements of all the parameters.

Since the sensors in a node monitor different spatio-temporally varying signals in the same environment, they often exhibit cross-correlation, which makes the system feasible to predict one parameter from its cross-correlated parameters.

The total number of possible sensor sets that can be constructed from P sensors, excluding the null set and the set containing all parameters, is $N = (2^P - 2)$. In a measurement cycle x if $\mathcal{A}_i^x$ denotes the active sensor set, the corresponding inactive/sleeping sensor set is $\mathcal{B}_i^x = \mathcal{P} - \mathcal{A}_i^x; 1 \leq i \leq N$. Therefore, $\mathcal{A}_i^x = \{P_{\mathcal{A}_i^x,m}; 1 \leq m \leq A_i^x\}$ and $\mathcal{B}_i^x = \{P_{\mathcal{B}_i^x,k}; 1 \leq k \leq B_i^x\}$. Let $\mathcal{S}$ be the set containing all active and sleeping subset of nodes, i.e., $\mathcal{S}^x = \{(\mathcal{A}_i^x, \mathcal{B}_i^x); 1 \leq i \leq N\}$. $|\mathcal{A}_i^x| = A_i^x$, $|\mathcal{B}_i^x| = B_i^x$, $|\mathcal{S}^x| = N$, and $A_i^x + B_i^x = P$.

This sensing model is used in the proposed framework in Section IV. The sensors belonging to the active set collect samples for a duration, called measurement cycle, and store them in the memory. At the end of the measurement cycle, the node transmits the gathered data to a central controller or local data collection center which does the computations based on the algorithm proposed in Section IV-C and sends some feedback to the node for subsequent adaptation on sensing quality and energy efficiency. In a given environmental setting, the average number of communication exchanges and the corresponding energy requirement is reasonably fixed. Moreover, sensing energy of a wireless node in environmental monitoring is significantly high, which can dominate over communication energy cost [4]. Since the sensing energy and quality optimization by exploiting cross-correlation of parameter values is the main focus, the communication cost is not accounted in this study.

III. GPR based Prediction Model

As stated in Section I-A, inspired by its good signal recovery scheme, GPR is used in this work to predict the parameter values for the inactive set in a multi-parameter sensor hub. *The literature survey reveals that, so far GPR has not been used for signal recovery in a multi-sensing node*.

A modified UCB algorithm, discussed in Section IV, returns an optimal sensor set. According to the selected set some sensors are to be activated and rest of the parameter values are predicted using the GPR model, described in this section.

A Gaussian process can be defined as a distribution over functions, where the inference takes place directly at the function space view [11]; in contrast, in Bayesian Linear Regression the inference takes place in the white space view. GPR exploits all possible functions drawn from a Gaussian distribution passing through the points with a standard deviation that would correspond to some uncertainty about the function.

Let $\underline{\mathbf{Z}}$ be the design matrix consisting of n feature vectors, $\mathbf{y}$ be the target vector having n target values from the training set, and $\mathbf{w}$ be the weight vector. For a test input $\mathbf{z}_*$:

$$\text{Prior: } Pr(f(\mathbf{z}_*)) = \int_{\mathbf{w}} Pr(f|\mathbf{w}, \mathbf{z}_*) Pr(\mathbf{w}) d\mathbf{w}. \tag{2}$$

$Pr(\mathbf{w})$ is assumed Gaussian and $Pr(f|\mathbf{w}, \mathbf{z}_*)$ is deterministic, hence $Pr(f(\mathbf{z}_*))$ is Gaussian.

$$\text{Posterior: } Pr(f(\mathbf{z}_*)|\underline{\mathbf{Z}}, \mathbf{y}) = \int_{\mathbf{w}} Pr(f|\mathbf{w}, \mathbf{z}_*) Pr(\mathbf{w}|\underline{\mathbf{Z}}, \mathbf{y}) d\mathbf{w}. \tag{3}$$

$Pr(f(\mathbf{z}_*)|\underline{\mathbf{Z}}, \mathbf{y})$ is also Gaussian in nature [21, Ch. 2].

According to the function view, there is a Gaussian at $f(\mathbf{z}_*)$ for every $\mathbf{z}_*$, and those Gaussians are correlated through some hyper-parameters, called weights. The correlation among the Gaussians is drawn from a user-defined covariance function, such as exponential, linear, etc. Such a large number of functions with univariate or multivariate Gaussian distributions constitute a Gaussian process [21, Ch. 2]. A Gaussian process $f(\mathbf{z})$ is expressed as: $f(\mathbf{z}) \sim GP(m(\mathbf{z}), k(\mathbf{z}, \mathbf{z}')) \ \forall \ \mathbf{z}, \mathbf{z}'$, where $m(\mathbf{z}) = E[f(\mathbf{z})]$ is the mean function and $k(\mathbf{z}, \mathbf{z}') = E[(f(\mathbf{z}) - m(\mathbf{z}))(f(\mathbf{z}') - m(\mathbf{z}'))]$ is the covariance function.

In the proposed approach, the prediction model consists of N sub-models for N number of sensor sets created from P parameters. Based on the selected active sensor set $\mathcal{A}_i^x \subset \mathcal{P}$ in the x^{th} measurement cycle, its corresponding sub-model is selected to predict the parameter values of $\mathcal{B}_i^x$. Intuitively, the correlated parameters have more significance in prediction.

During training the model, the active (sleep) set is denoted as $\mathcal{A}_i$ ($\mathcal{B}_i$) and the corresponding measurement vector is $\mathbf{z}_{\mathcal{A}_i}$ ($\mathbf{z}_{\mathcal{B}_i}$). During prediction in x^{th} cycle, the measurement vectors of the active set $\mathcal{A}_i^x$ and the inactive set $\mathcal{B}_i^x$ are respectively $\mathbf{z}^x_{\mathcal{A}_i} = \{z^x_{\mathcal{A}_i}(1), z^x_{\mathcal{A}_i}(2), \cdots, z^x_{\mathcal{A}_i}(A_i^x)\}$ and $\mathbf{z}^x_{\mathcal{B}_i} = \{z^x_{\mathcal{B}_i}(1), z^x_{\mathcal{B}_i}(2), \cdots, z^x_{\mathcal{B}_i}(B_i^x)\}$. Thus, $\forall i \in \mathcal{S}^x$, (1) is rewritten as: $\mathbf{z}^x_{\mathcal{A}_i} = \mathbf{y}^x_{\mathcal{A}_i} + \boldsymbol{\eta}$ for $\mathcal{A}_i^x$, and $\mathbf{z}^x_{\mathcal{B}_i} = \mathbf{y}^x_{\mathcal{B}_i} + \boldsymbol{\eta}$ for $\mathcal{B}_i^x$.

The sub-models are initially trained using n training samples. The i^{th} sub-model having B_i^x number of regressors to predict each parameters of $\mathcal{B}_i^x$. k^{th} regressor of the i^{th} sub-model is trained using n input/feature vectors belonging to the design/training matrix $\underline{\mathbf{Z}}_{\mathcal{A}_i} \in \mathbb{R}^{A_i \times n}$ and n target values belonging to the target vector $\mathbf{z}_{\mathcal{B}_i}(k) \in \mathbb{R}^{n \times 1}$. If $\mathcal{A}_i^x$ is the optimal sensor set selected to activate in the x^{th} measurement cycle, for a point observation $\mathbf{z}^x_{\mathcal{A}_i,*} \in \mathbb{R}^{A_i^x \times 1}$, GPR predicts $y^x_{\mathcal{B}_i,*}(k) \in \mathbb{R}, \forall k \in \mathbf{y}^x_{\mathcal{B}_i}$. Measurement noise η is assumed to have zero mean and variance σ^2 for all the parameters.

$$z^x_{\mathcal{B}_i,*}(k) = y^x_{\mathcal{B}_i,*}(k) + \eta = f_k(\mathbf{z}^x_{\mathcal{A}_i,*}) + \eta = f^x_{i,*}(k) + \eta \quad (4)$$

where $f \sim N(0, K_{n\times n})$. $K_{n\times n} = [k(\mathbf{z}_{\mathcal{A}_i,r}, \mathbf{z}_{\mathcal{A}_i,s})]_{n\times n}$, where $k(\mathbf{z}_{\mathcal{A}_i,r}, \mathbf{z}_{\mathcal{A}_i,s})$ is the covariance function between the r^{th} and s^{th} instant of design/training matrix $\underline{\mathbf{Z}}_{\mathcal{A}_i} \in \mathbb{R}^{A_i \times n}$.

In a WSN, measurement of sensing parameters generates time-series data, which implies that they can be temporally correlated. However, in many cases, such as in air pollution monitoring, data has low temporal correlation because of large sampling intervals. In such cases squared exponential covariance function fits better than the other functions [21, Ch. 2]. In a noisy environment, the covariance function is:

$$k(\mathbf{z}_{\mathcal{A}_i,r}, \mathbf{z}_{\mathcal{A}_i,s}) = e^{-\frac{1}{2l^2}\sum_{m=1}^{A_i} |z_{\mathcal{A}_i,r}(m) - z_{\mathcal{A}_i,s}(m)|^2} + \sigma^2\delta_{rs}. \quad (5)$$

Prediction model for GPR in a noisy environment is: $f^x_{i,*}(k)|\underline{\mathbf{Z}}_{\mathcal{A}_i}, \mathbf{z}_{\mathcal{B}_i}(k), \mathbf{z}^x_{\mathcal{A}_i,*} \sim N(\overline{f^x_{i,*}(k)}, \mathrm{Cov}(f^x_{i,*}(k)))$, where $\overline{f^x_{i,*}(k)}$ is the mean function and $\mathrm{Cov}(f^x_{i,*}(k))$ is the covariance function that can be derived from (2) and (3) [21, Ch. 2]:

$$\overline{f^x_{i,*}(k)} = K(\underline{\mathbf{Z}}_{\mathcal{A}_i}, \mathbf{z}^x_{\mathcal{A}_i,*})[K(\underline{\mathbf{Z}}_{\mathcal{A}_i}, \underline{\mathbf{Z}}_{\mathcal{A}_i}) + \sigma^2 I]^{-1}\mathbf{z}_{\mathcal{B}_i}(k) \quad (6)$$

$$\mathrm{Cov}(f^x_{i,*}(k))) = k(\mathbf{z}^x_{\mathcal{A}_i,*}, \mathbf{z}^x_{\mathcal{A}_i,*}) + K(\mathbf{z}^x_{\mathcal{A}_i,*}, \underline{\mathbf{Z}}_{\mathcal{A}_i})[K(\underline{\mathbf{Z}}_{\mathcal{A}_i}, \underline{\mathbf{Z}}_{\mathcal{A}_i}) + \sigma^2 I]^{-1}K(\underline{\mathbf{Z}}_{\mathcal{A}_i}, \mathbf{z}^x_{\mathcal{A}_i,*}). \quad (7)$$

The complexity of the GPR depends on $(K+\sigma^2 I)^{-1}$ which is cubic in the number of training points.

Although there are a few signal recovery schemes available in the literature, these are suitable to estimate a sparse signal. As presented in [19], Sparse Bayesian Learning (SBL) method is used to estimate a spatially varying sparse signal in a densely deployed sensor network. In contrast, the proposed work considers a single sensor hub fixed at a point in space which is unaware of the spatial variations of the parameters. At a multi-parameter sensing hub, multiple cross-correlated parameter values do not exhibit strong sparsity, which makes the sparse signal recovery schemes such as SBL [19], Orthogonal Matching Pursuit (OMP) [22], inefficient in estimating the signals. To this end, GPR based prediction model is used in this work, as it assumes Gaussian prior on the data which fits well with the sensing signals at the node.

IV. Proposed Sensor Selection Framework

In contrast to the existing sensor selection frameworks, which deal with the network-level sensor node adaptation, the proposed framework deals with the node-level sensor selection analysis of a multi-parameter sensor hub.

This section describes the proposed optimal sensor set selection framework using a modified UCB algorithm. With the sample values from the selected set, using the corresponding GPR sub-model, parameter values of the inactive set are predicted following the procedure in Section III.

Let c_{th} be the correlation threshold. If a and b have correlation greater than c_{th}, they are declared as correlated, else they are uncorrelated. For each subset $\mathcal{A}_i^x$ in x^{th} cycle,

$$c^x_{i,k} = \sum_{m=1}^{A_i^x} \mathbb{1}_{[|c^x(m,k)| \geq c_{th}]}; \forall k \in \mathcal{B}_i^x \text{ and } m \in \mathcal{A}_i^x. \quad (8)$$

If $c^x_{i,k} \geq 1; \forall k \in \mathcal{B}_i^x$, then $\mathcal{A}_i^x$ is correlated with $\mathcal{B}_i^x$. Therefore, all the parameters of $\mathcal{B}_i^x$ can be predicted from the parameters of $\mathcal{A}_i^x$. Let $\tilde{\mathcal{C}}^x_{i,k}$ be the set containing the parameters of $\mathcal{A}_i^x$ correlated with k^{th} parameter of $\mathcal{B}_i^x$. For each correlated pair of subset $(\mathcal{A}_i^x, \mathcal{B}_i^x) \in \mathcal{S}^x$, the average cross-correlation between the parameters of $\mathcal{A}_i^x$ and $\mathcal{B}_i^x$ is:

$$C_i^x = \frac{1}{B_i^x}\sum_{k=1}^{B_i^x} \frac{1}{c^x_{i,k}} \sum_{q=1}^{c^x_{i,k}} |c^x(q,k)|; \forall k \in \mathcal{B}_i^x \text{ and } q \in \tilde{\mathcal{C}}^x_{i,k}. \quad (9)$$

$C_i^x = 0$ for each uncorrelated pair of subset $(\mathcal{A}_i^x, \mathcal{B}_i^x) \in \mathcal{S}^x$. C_i^x can be defined as cross-correlation factor indicating the correlation between the parameters of $\mathcal{A}_i^x$ and $\mathcal{B}_i^x$.

Energy consumption for the subset $\mathcal{A}_i^x$ is:

$$E_i^x = \sum_{m=1}^{A_i^x} En_m^x; \forall m \in \mathcal{A}_i^x. \quad (10)$$

En_m^x is the sensing energy consumed by the m^{th} sensor.

Prediction error for the subset $\mathcal{B}_i^x$ is:

$$PE_i^x = \frac{1}{B_i^x}\sum_{k=1}^{B_i^x} Pe_k^x; \forall k \in \mathcal{B}_i^x. \quad (11)$$

Pe_k^x is the prediction error for k^{th} parameter of $\mathcal{B}_i^x$.

C_i^x and PE_i^x are dynamic, but E_i^x does not change as the sensors are assumed to consume fixed energy in sensing.

A. Modified UCB Algorithm for Sensor Selection

As discussed in motivation, UCB is a popular value-based method used to solve MAB problems, where the decision of choosing an optimal arm is taken online based on previous experiences. The goal is to minimize the regret by maximizing the reward achieved based on the action [23].

In this work, the UCB algorithm is modified in order to use in the proposed sensor selection framework. As mentioned in Section III, N is the total possible number of sensor sets created from P sensors. The UCB algorithm is used to find an optimal set in every measurement cycle. Considering the sensing environment with a sensor hub collecting samples from the environment as a random process, a reinforcement learning model is created where the learner can learn the system with time, based on the previous experience. Since the samples were drawn from some non-stationary distribution exhibiting correlations, the learner chooses an optimal sensor set having $M < P$ sensors, which gives the knowledge of all the P parameters in the environment that is sufficient to reconstruct all the time-varying signals.

Choosing an optimal sensor set at some measurement cycle depends on the strength of cross-correlations among the parameters of the active set and the inactive/sleep set. Therefore, the cross-correlation factor defined in (9) acts as the performance parameter to decide the reward for a sensor set. Since, the main objective of this study is to optimize the energy sustainability of the node, the energy consumed by the sensors upon selecting a sensor set defined in (10) is termed as another performance parameter to calculate reward.

An optimal set $\mathcal{A}_i^x \in S^x$ needs to be selected that minimizes both prediction error PE_i^x in (11) and energy consumption E_i^x in (10) in a way that maximizes the cross-correlation factor C_i^x in (9) and minimizes energy consumption E_i^x.

To incorporate energy awareness, the residual energy is considered as another performance parameter for finding the optimal sensor set. Let E_{batt} be the battery capacity of the node and E_0^x be its available energy in x^{th} measurement cycle. $\lambda^x \triangleq \frac{E_0^x}{E_{batt}}$ is the normalized energy available at the node.

Let R_i^x be the reward obtained from some distribution $P_{R_i}^x$ for selecting a set $\mathcal{A}_i^x \in \mathcal{S}^x$ at the x^{th} measurement cycle, then R_i^x can be defined as:

$$R_i^x = \frac{\lambda^x (C_i^x)^\gamma}{\nu^x (E_i^x)^\beta}, \tag{12}$$

where $\nu^x = \max_{i \in \mathcal{S}^x} \frac{(C_i^x)^\gamma}{(E_i^x)^\beta}$.

The reward is bounded in $[0, 1]$. A higher value of C_i^x means a better sensing quality. If $C_i^x > c_{th}$ for some set $i \in \mathcal{S}$, the corresponding active-sleep sets are correlated, which ensures faithful reconstruction of the predicted signals. β and γ in (13) are imposed to control the weight on C_i^x and E_i^x in the reward. Optimum values of β and γ would minimize the regret.

In contrast to the traditional UCB algorithm where the reward is assigned to the selected set only, in the modified UCB algorithm the reward is calculated for every set at every measurement cycle based on the collected and predicted samples. Since correlation among the slowly-varying signals does not change rapidly, in deciding the next state the present state carries significant information compared to the previous states.

Assuming the rewards for i^{th} sensor set up to x^{th} cycle $R_i^1, R_i^2, \cdots, R_i^x$ as a sequence of independent Gaussian random variables with true mean μ and variance 1, the empirical mean of the distribution of reward $P_{R_i}^x$ is estimated as:

$$\widehat{\mu}_i^x = \frac{1}{x} \sum_{t=1}^{x} R_i^t = \frac{1}{x} \sum_{t=1}^{x} \frac{\lambda^t (C_i^t)^\gamma}{\nu^t (E_i^t)^\beta}. \tag{13}$$

The objective function for selecting an optimal sensor set at the $(x+1)^{th}$ measurement cycle is:

$$(\mathcal{P}\mathbf{1}) : \mathcal{A}_\mathbf{i}^{\mathbf{x+1}} = \underset{i \in \mathcal{S}^x}{\text{maximize}} \frac{1}{x} \sum_{t=1}^{x} \frac{\lambda^t (C_i^t)^\gamma}{\nu^t (E_i^t)^\beta} + \sqrt{\frac{2 \ln \frac{1}{\delta}}{T_i^x}} \tag{14}$$
$$\text{s. t. } C_i^x > c_{th} \text{ and } I^{x+1} E_i^{x+1} < E_0^x.$$

For each sensor set, C_i^x is updated after every measurement cycle whereas E_i^x remains unchanged because of fixed sensing energy consumption. I^{x+1} denotes the number of samples collected in $(x+1)^{th}$ cycle for each parameter, T_i^x is the number of times the i^{th} sensor set is activated, and δ is defined as a confidence bound. The parameter λ controls the weight on energy efficiency in the reward. If E_0 increases (in energy harvesting nodes) C_i^x gets relatively more weight, whereas for decreasing E_0 (in battery-operated or energy harvesting node) the energy parameter in (14) gets more weight. The impact of λ on the algorithm performance is explained in Section VI.

In the x^{th} measurement cycle, if $\mu_{i^*}^x$ is the true mean of the optimal sensor set i^* and $\widehat{\mu}_i^x$ is the empirical mean of the selected sensor set i, from [23] for any $\delta \in (0, 1)$,

$$P(\widehat{\mu}_i^x \geq \mu_{i^*}^x + \sqrt{\frac{2 \ln \frac{1}{\delta}}{T_i^x}}) \leq \delta \tag{15}$$

$$P(\widehat{\mu}_i^x \leq \mu_{i^*}^x - \sqrt{\frac{2 \ln \frac{1}{\delta}}{T_i^x}}) \leq \delta. \tag{16}$$

In (14), the first term is the empirical mean of the distribution of reward while the second term provides the confidence bound. T_i^x is the number of times the set $\mathcal{A}_i^x$ is selected up to x^{th} measurement cycles. A higher value of T_i^x makes the learner more confident about the distribution of reward. At the end of every measurement cycle, the objective function in (14) explores all the sets and selects an optimal set that satisfies the constraints in (14). The first constraint is $C_i^x > c_{th}$; if a set is not correlated in x^{th} measurement cycle, the probability that it will be optimal for $x+1^{th}$ measurement cycle is negligible, although it might have obtained a higher reward in the previous measurement cycles. The second constraint is related to the node's energy availability. The residual energy of the node has to be more than the total energy consumed in sensing. Initially, $T_i = 1; \forall i \in \mathcal{S}$, as the reward for all the sets can be calculated from the training samples. $\delta = \frac{1}{(x+2)^2}$ at x^{th} measurement cycle is found to be appropriate [23].

The algorithm always returns a sensor set to be activated at the next measurement cycle based on the past experience unless the two extreme conditions occur: a) parameters are totally uncorrelated, b) residual energy at the node is not enough to continue sensing operation. The traditional UCB algorithm works on the assumption of the stationary distribution of reward and selects the optimal set having the maximum

probability of getting the highest reward. In contrast, the distribution of reward in the considered system is non-stationary. To incorporate this non-stationarity, the reward is calculated at every measurement cycle for every set, which helps the learner to learn the distributions with the recently drawn samples. Therefore, the function does not stick to choose some sub-optimal set subsequently; rather it tends to select different sets at different points of time. The UCB algorithm is re-initiated while retraining the GPR model given in Section III.

The main objective of the learner is to minimize the regret which is the price paid by the learner for not choosing the optimal set. Regret is the expected deviation between the mean of the optimal set and the selected set. If the selected set is not optimal then the regret is greater than zero. Therefore, the underlying cost function is a minimization problem. The expression of upper bound on regret up to some measurement cycle for the proposed method is given in (17) and the derivation is given in the appendix.

$$Reg^x \leq \sum_{i \in \mathcal{S}^x; \Delta_i > 0} (1 + 2x^2\delta)\Delta_i + \sum_{i \in \mathcal{S}^x; \Delta_i > 0} \frac{8 \ln \frac{1}{\delta}}{\Delta_i}. \quad (17)$$

UCB algorithm is effective as it provides an upper bound on regret. Although the regret grows with the number of measurement cycles, UCB offers minimal growth. Δ_i is the difference between the mean reward of the optimal sensor set and the i^{th} sub-optimal sensor set selected in that measurement cycle.

The literature survey reveals that, this is the first reinforcement learning-based sensor selection framework for multi-parameter sensor hub.

B. Adaptive Data Collection Interval

In WSN, data transmission consumes a large percentage of energy. In a slowly varying process, the correlation among the sensing parameters changes slowly. Hence, data need not to be transmitted after every sampling instant, rather the sampled data can be stored in memory and transmitted after a certain period which can be called the length of the measurement cycle or data collection interval. Let, the length of x^{th} measurement cycle is $\mathcal{T}^x$.

The communication module consumes significant energy during its ON period. By optimally activating the communication module, the energy consumption can be reduced. On the other hand, the data correlation changes with time due to the non-stationary nature of the environment at parameter values. A smaller length of measurement cycle $\mathcal{T}^x$ maintains a similar cross-correlation among the parameters over the whole cycle, however at the cost of a higher communication energy. On the other hand, a larger $\mathcal{T}^x$ reduces the communication energy but the data correlation may change abruptly, thereby affecting the optimality of sensor selection within the measurement cycle. The fixed sample-based approach, proposed in [1], works on individual sensors. Therefore, it is not suitable for the proposed multi-sensing framework. Hence, *the data collection interval in the proposed work is adaptive to the stochasticity of the environment.*

The data collection interval $\mathcal{T}^x$ in the proposed framework is adapted using a parametric mapping function given in (19) based on the change in cross-correlations among the parameters given in (18) [24].

$$d = \frac{1}{N} \sum_{i=1}^{N} |C_i^x - C_i^{x-1}| \quad (18)$$

$$y(d) = a_o + a_1(d)^\alpha \text{ (cycle/hour)}, \quad (19)$$

where d is the average difference between present and past cross-correlation factors over set $\mathcal{S}$ and $y(d)$ is the frequency of measurement cycle in cycle/hour. The length of x^{th} measurement cycle is $\mathcal{T}^x = \frac{1}{y(d)}$. a_o and a_1 can be computed using the boundary conditions: $d \in [0, 1]$ for $y(d) \in [b_o, b_1]$ where b_o and b_1 are user-defined. α is the shape parameter which decides the rate of adaptation. The optimum value of α is determined numerically by minimizing the regret in (17), which is convex with respect to Δ_i. Regret varies with the variation of α, as the number of samples collected at a particular measurement cycle varies with α. Hence there exists an optimum value of α which minimizes the regret function.

The sampling rate of the selected sensors is decided based on Nyquist criteria. The maximum frequency of each parameter is computed after predicting the missing samples. The largest one among all frequencies is declared as the maximum frequency of the sensor hub and the Nyquist rate is calculated accordingly. In a dynamic stochastic environment, this sampling frequency may change in every measurement cycle, hence it is recomputed.

C. Feedback Mechanism for Adaptive Sensor Selection

The adaptive sensor selection algorithm is given in Algorithm 1 which runs at the CE and gives feedback to the sensor node. Initially, the prediction model is trained at the CE using a sufficient number of samples n collected at the node. The prediction model consists of N sub-models for N subsets in S. To retrain the model, the temporal correlation of the individual parameters is exploited. As these are slowly varying temporal signals, it is expected to have a strong temporal correlation among the signals from cycle to cycle. Let, ct_p^x is the temporal correlation of the p^{th} parameter in x^{th} measurement cycle and $ct_{p,th}$ is the temporal correlation threshold of p^{th} parameter, defined by the user to maintain an acceptable rage of prediction error. ct_p^x is computed by drawing an equal number of samples from the reconstructed signals at the present and past cycles at a fixed interval. Since prediction errors are random, the increase in prediction error reduces the temporal correlations among the signals. Hence, all the sensors are to be activated in the $(x+1)^{th}$ measurement cycle if $ct_p^x \ngeq ct_{p,th} \forall p \in \mathcal{P}$ in x^{th} measurement cycle. Samples collected in $(x+1)^{th}$ measurement cycle is used as a test data-set to find prediction error (test error) for all the N sub-models. If the error for any sub-model differs from the recorded error during training (PE_i) by some value (ϵ_i) then the complete GPR model is retrained by collecting n training samples from all the sensors. In Algorithm 1, PE_i and PE_i^x are the prediction errors of the i^{th} sensor set during training/retraining and during x^{th} measurement cycle respectively. The CE collects input from the node, runs the algorithm, and sends feedback to the sensor node. The inputs to the sensor hub are the optimal sensor

Algorithm 1: Active sensor set selection at the CE

Input: Sampled data and E_0^x from the sensor node
if *e=1* **then**
  Retrain and test the model with recently collected samples
  Find $PE_i; \forall i \in \mathcal{S}$ using (11)
  Set $x = 0$, $e = 0$, and $\mathcal{T}^x = \mathcal{T}'$
else
  if *e'=1* **then**
    Find $PE_i^x; \forall i \in \mathcal{S}$ using (11)
    if $|PE_i^x - PE_i| \leq \epsilon_i; \forall i \in \mathcal{S}^x$ **then**
      Calculate $y(d)$ from (18) and (19)
      Set $\mathcal{T}^x = \frac{1}{y(d)}$
      Set $e' = 0$
    else
      Set $e' = 0$, and $e = 1$
      Find maximum frequency $f_m^p; \forall p \in \mathcal{P}$ using FFT
      Set $f_s^x = 2 \times max\{f_m^1, f_m^2, f_m^3 .... f_m^P\}$
      Set $\mathcal{T}^x = \frac{n}{f_s^x}$
    end
  else
    Predict the missing samples using the appropriate sub-model of GPR
    Calculate $y(d)$ from (18) and (19)
    Set $\mathcal{T}^x = \frac{1}{y(d)}$
    Find $ct_p^x; \forall p \in \mathcal{P}$
    while $ct_p^x \not\geq ct_{p,th}; \forall p \in \mathcal{P}$ **do**
      Set $e' = 1$
    end
  end
end
while *e=0* **do**
  Construct $\mathcal{S}^x$ and calculate C_i^x and $E_i^x; \forall i \in \mathcal{S}^x$ using (9) and (10) respectively
  Find optimal active sensor set $\mathcal{A}_i^{(x+1)}$ by solving (14)
  Find maximum frequency $f_m^p; \forall p \in \mathcal{P}$ using FFT
  Set $f_s^x = 2 \times max\{f_m^1, f_m^2, f_m^3 .... f_m^p\}$
end
Set $x = x + 1$
Output: Transmit updated $\mathcal{T}^x, f_s^x, \mathcal{A}_i^{(x+1)}, e, e'$ to the node

set $\mathcal{A}_i^x$ that contains the sensors to be activated in the x^{th} cycle, the length of x^{th} measurement cycle, sampling rate f_s^x in the x^{th} cycle, and the status of the two flags e, e'. The outputs of the sensor hub are the data collected by the sensors belonging to $\mathcal{A}_i^x$ and the remaining energy of the node, which are transmitted to the CE. Based on the status of the two flags e, e' and the sensors belonging to the active set $\mathcal{A}_i^x$, the sensor node activates the sensors. $\mathcal{T}'$ is chosen by the user for some $y(d) \in [b_o, b_1]$ as the length of first measurement cycle.

D. Complexity of the Proposed Framework

The computational complexity in training/retraining and prediction with feedback mechanism are different. During training, the parameters are computed where the complexity is the $\mathcal{O}(N)$. Complexity of GPR model is $\mathcal{O}(n^3)$, where n is the number of training samples. If t, u, and v are the numbers of α, β, and γ respectively comprising the search space, then the total computational complexity in training N sub-models with k-fold cross-validation is $\approx (t + u + v)\mathcal{O}(N) + k(\sum_{i=1}^{N} B_i)\mathcal{O}(n^3)$. Complexity in prediction in the x^{th} measurement cycle is $\approx B^x I^x \mathcal{O}(n'^3)$, where B^x is the number of parameters in the inactive set selected to turn off and I^x is the number of samples collected for each parameter in x^{th} measurement cycle. $n' = 1$ is the number of predictions made at an instant of time. Complexity in finding optimal sensor set using (14) is $\mathcal{O}(N)$. Therefore, complexity in prediction and feedback mechanism in x^{th} measurement cycle is given by $\approx B^x I^x \mathcal{O}(n'^3) + \mathcal{O}(N)$, where $N = 2^P - 2$.

In contrast, the complexity of the algorithm in [1] is $\mathcal{O}(P)$. Although the complexity of the proposed framework is higher than the existing competitive methods, the whole computation is done by the CE and the sensor node has the only task of sensing based on the active set input from the CE.

TABLE I: Specifications of sensors

Sensors	Parameters	Energy consumption (J)
DHT11	Temperature, Humidity	12m
MQ-137	NH_3	1.1
AFE-A4 alphasense	NO_2, Ozone, CO, SO_2	54m
Alphasense OPC N3	PM_{10}, $PM_{2.5}$	29.55

V. Experimental Setup

The performance of the proposed algorithm explained in Section IV is evaluated using experimental and simulation results. To demonstrate the efficiency of the algorithm in real-life applications, an air pollution monitoring sensor hub is deployed on the campus to collect data. The sensor node consists of seven sensors monitoring nine parameters in the air. The various sensors used to collect data are listed in Table I along with their energy consumption. The energy consumption of sensors includes the energy to turn on the sensor, to heat-up (heat-up time is the time required to prepare for sensing), and sensing (to collect the first two good data) from the environment. It is assumed that the energy consumption of the sensor does not change with time or place. Four different gas concentration monitoring sensors (NO_2, Ozone, CO, SO_2), embedded in the 3^{rd} sensor board, powered by a single supply pin, makes it difficult to measure the energy consumption of individual sensors. Therefore the sensing energy of the four sensors is considered as 1.57J, 0.054J, 0.642J, and 1.35J respectively [6].

Correlation exhibited among the parameters has been studied and it has been observed that most of the parameters have a cross-correlation coefficient higher than 0.5. In Table I, it can be observed that the OPC N3 used as a particulate matter sensor consumes huge energy. Therefore it is required to smartly handle this sensor. The main objective of the proposed framework is to make such nodes energy efficient by optimally operating the sensors.

The experimental setup gave a set of raw data based on certain sampling periodicity, which gives a parameter variation profile. The algorithm described in Section IV is applied on that data set, which returns an active sensor set. Correspondingly, some data has been chosen from that same data set while the rest of the values are predicted. The simulation results given in the next section validate that the predicted values

combined with the actual values are sufficiently good with respect to the actual values to reconstruct the signals.

VI. Results and Discussion

This section demonstrates the significance of results with the proposed framework. As discussed in Section V, the dataset collected by deploying an air-pollution monitoring node with nine sensors mounted on it is used to study the proposed algorithm. The performance of the learning-based algorithm is compared with the most recent and nearest competitive sensor data collection algorithm in [1], based on the Anova model and Bartlett test. The simulations have been performed in MATLAB along with training/retraining the prediction models.

To find a suitable correlation threshold, the average of $max(\frac{C_i}{E_i}); \forall i \in \mathcal{S}$ is calculated up to some measurement cycle for $c_{th} \in [0, 1]$, shown in Fig. 1. It has been observed that the maximum value lies on $c_{th} \in (0.5, 0.7)$. According to the literature, if correlation coefficient between two parameters is higher than 0.5, they can be declared as correlated[25]. Form the Fig. 1 up to $c_{th} < 0.5$, C_i is increases with increasing c_{th} but after $c_{th} > 0.7$, E_i increases, as the number of sensor set satisfying the criteria given in (8) is reduces. after $c_{th} > 0.9$ it is almost 0 because none of the parameters are having cross-correlation higher than 0.9. Choosing $c_{th} = 0.5$ make the system more energy-efficient with relatively larger prediction error, while choosing $c_{th} = 0.7$ increases the predicted signal quality by reducing the energy efficiency.

To find optimum training length for the models average value of prediction error ($\frac{1}{N}\sum_{i=i}^{N} PE_i; \forall i \in \mathcal{S}$) has been studied, where PE_i is the prediction error of i^{th} inactive sensor set computed from (11). The prediction error of each parameter is measured in terms of mean relative error (MRE), given in (20). Form Fig. 2, it can be observed that increasing the length of the training sequence reduces test error until the model starts over-fitting the data. An optimum training length can be chosen between $1500 - 1600$ assuming that all N sub-models perform well with minimum train and test error.

MRE for a parameter $p \in \mathcal{P}$ is defined as:

$$\text{MRE}_p = \frac{1}{I}\sum_{i=1}^{I} \frac{|z_p(i) - \widehat{z}_p(i)|}{|z_p(i)|}. \tag{20}$$

$z_p(i)$ and $\widehat{z}_p(i)$ are the i^{th} samples of the actual and predicted data sequence respectively. I is the total number of samples.

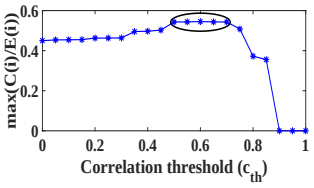


Fig. 1: Optimum correlation threshold 0.5.

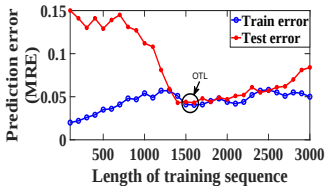


Fig. 2: Optimum training length 1500.

A. Finding Optimum values Parameters

The optimum values of parameters (α, β, and γ) that minimize the regret function defined in (17) have been determined numerically based on the training samples and these values are recomputed at the time of retraining the models.

$\beta = \gamma = 1$, $b_o = 1 cycle/hr$, and $b_1 = 2 cycle/hr$ is set for the objective function given in (14) to find the optimum value of α for the function given in (19) that minimizes the regret given in (17). Fig. 3(a) shows that the regret is minimum for some α and slightly varies with the variation of measurement cycles. An optimum value of $\alpha = 0.7$ is chosen to simulate the proposed algorithm. The $\Delta_i; \forall i \in \mathcal{S}$ is changing with time, as the reward distribution is non-stationary in practical. Since the true mean is unknown, the averaged empirical mean is used to find Δ_i for the i^{th} sensor set.

Similarly, β is chosen by setting $\alpha = 0.7$ and $\gamma = 1$, shows in Fig. 3(b). An optimum value of $\beta = 1.8$ gives minimal regret. From Fig. 3(c), $\gamma = 1$ can be picked up as it minimizes the regret for $\alpha = 0.7$ and $\beta = 1.8$.

B. Simulation Results of the Proposed Algorithm

Simulation results on the proposed algorithm for $c_{th} = 0.5$ using (14) are shown in Fig. 4. Fig. 4(a) presents the various sensor sets activated at different measurement cycles. Since optimal sensor set changes with time due to the non-stationary distribution of reward, the learner selects different active sensor sets at different measurement cycles. The cross-correlation factor of the active sensor set at different measurement cycles is shown in Fig. 4(b). Although the c_{th} is set to 0.5, the cross-correlation factor of the active sensor set is much higher than the threshold, c_{th}, which ensures a good prediction performance. Fig. 4(c) shows that the data collection interval is adaptive to the system dynamics.

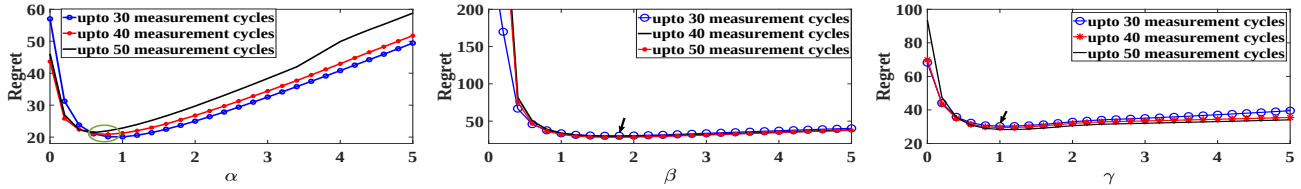
Average energy consumed by the active sensors and average prediction error of the parameters belongs to the inactive set are considered up to some measurement cycles to show the trade-off between energy consumption and prediction error, which can be clearly observed in Fig. 5(a). Increasing the number of active sensors with the order given in Table-1 increases sensing energy consumption while reducing the prediction error. The variation of prediction error with c_{th} is presented in Fig. 5(b), which shows that increasing c_{th} reduces the prediction error while increasing the energy consumption, as more number of sensors are to be activated.

Fig. 6 shows that, using E_0 increases the network lifetime for lower values of c_{th}. A higher value of c_{th} reduces the energy efficiency; the effect of E_0 is negligible at $c_{th} = 0.8$.

Incorporating energy awareness in the reward makes the system adaptive to energy availability at the node.

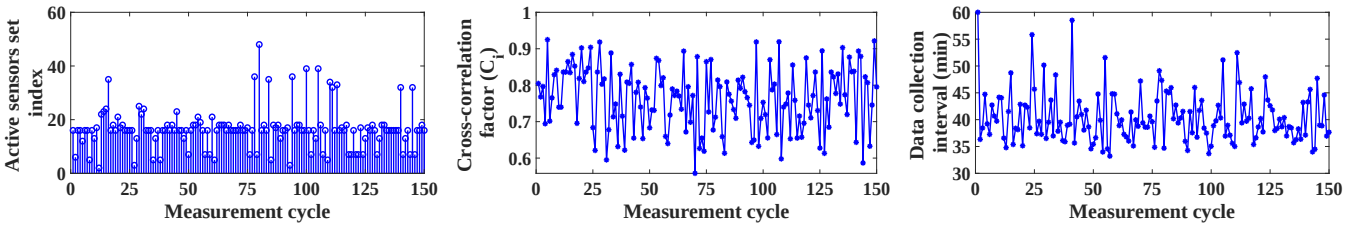
C. Performance Comparison with State-of-the-Art:

As discussed in Section I-B, although many works have been reported on sensor selection algorithm, most of the recent works are dedicated to network-level adaptation. On the contrary, the proposed sensor selection algorithm is dedicated to node-level adaptation. The performance of the proposed algorithm is compared with the most recent competitive data collection algorithm proposed in [1]. The algorithm proposed in [1] is based on a statistical approach, which uses one-way Anova model and Bartlett test to find the new sampling rate.



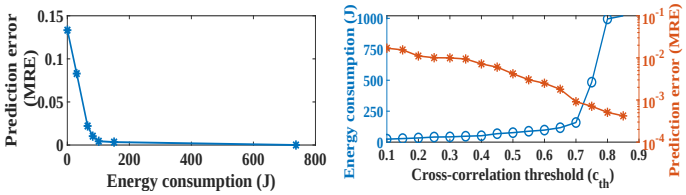
(a) Regret versus α; $\beta = \gamma = 1$ (b) Regret versus β; $\alpha = 0.7$ and $\gamma = 1$ (c) Regret versus γ; $\alpha = 0.7$ and $\beta = 1.8$

Fig. 3: Finding optimum value of the hyper-parameters.



(a) Active sensor set index (b) Correlation factor of the active sensor set (c) Data collection interval

Fig. 4: Simulation results of the proposed algorithm, for $c_{th} = 0.5$, $l = 1$, $\sigma = 0.0005$.



(a) (b)

Fig. 5: a)Trade-off between sensing energy consumption and prediction error, b) Average energy consumption and prediction error versus cross-correlation threshold.

(a) $c_{th} = 0.5$ (b) $c_{th} = 0.8$

Fig. 6: Network life time with/without using E_0 in the reward.

On the other hand, the proposed algorithm is based on learning technique that selects a set of sensors in a measurement cycle from the total number of sensors available in the node, and these sensors collect samples based on the Nyquist rate. Unlike the work in [1], where the adaptive sampling algorithm is applied in the node itself, the proposed sensor selection algorithm is computed at the CE, which reduces the processing energy at the node.

Samples collected by the active sensors are received at the central entity, where the missing samples of the parameters belonging to the sleeping set are predicted. For a faithful reconstruction of the signal, MRE given in (20) is used as a performance index. $z_P(i)$ and $\widehat{z}_p(i)$ in (20) are respectively the i^{th} samples of the actual and reconstructed data sequence for the p^{th} parameter. Following [4], MRE threshold is set as 1%. Reconstruction errors for the nine parameters using Anova-based adaptive sampling [1] and proposed framework for $c_{th} = 0.5$, 0.7, and 0.8 are listed in Table II. It has been observed that the reconstruction error in the proposed framework for $c_{th} > 0.5$ is below 1% for all the parameters.

TABLE II: Reconstruction error (MRE)

Parameters	Anova model	Proposed framework ($c_{th} = 0.5$)	Proposed framework ($c_{th} = 0.7$)	Proposed framework ($c_{th} = 0.8$)
Temperature	0.00091	0.0007	0.00065	0.00051
Humidity	0.00072	0.00082	0.00061	0.0005
PM_{10}	0.00093	0.0092	0.0019	0.0011
$PM_{2.5}$	0.00086	0.0091	0.0015	0.0012
CO	0.0028	0.0098	0.0039	0.00107
NH_3	0.0012	0.0035	0.0004	0.0003
NO_2	0.00082	0.0033	0.00083	0.00081
Ozone	0.0011	0.0064	0.0018	0.001
SO_2	0.00098	0.0048	0.001	0.0006

The parameters having lower energy consumption (Temperature, Humidity) are activated more frequently. Hence, for those parameters sensing error is lower in the proposed model compared to the Anova-based model.

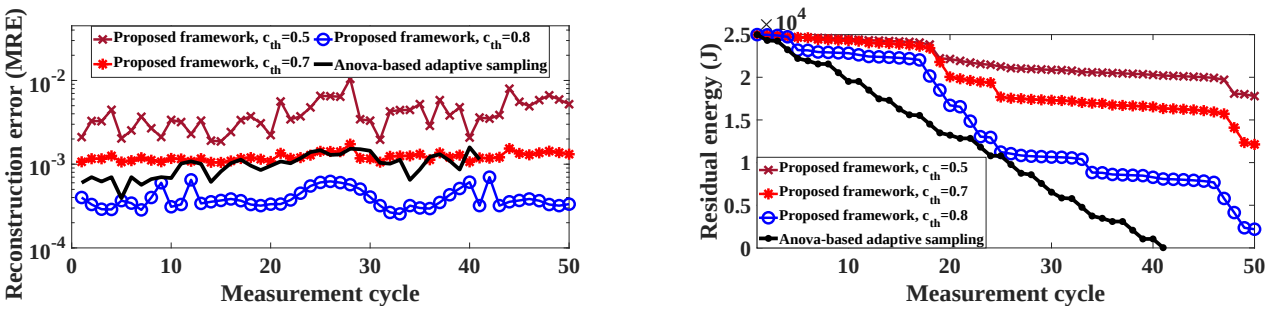
The average reconstruction error for P parameters at the x^{th} measurement cycle is given by,

$$\text{RE}^x = \frac{1}{P}\sum_{p=1}^{P} \text{MRE}_p^x, \tag{21}$$

where MRE_p^x is the reconstruction error for the p_{th} parameter at the x^{th} measurement cycle.

Fig. 7(a) shows a comparison of average reconstruction error, computed using (21), for the Anova-based model and the proposed framework for three different values of c_{th}. It is analytically observed that, with $c_{th} = 0.7$, the error for the proposed learning-based model is similar to the Anova-based model. Considering $c_{th} = 0.5$, the error is higher than the Anova-based model but it is still lower than 10^{-2}. Following [4], MRE $< 10^{-2}$ is acceptable. The error goes below 10^{-3} for $c_{th} = 0.8$. Thus, in the proposed model a suitable value c_{th} can be set to achieve the required error performance.

To obtain the energy efficiency of the proposed model over the existing methods, the sensing energy consumption at the



(a) (b)

Fig. 7: Comparison of (a) energy consumption and (b) error of the proposed framework with Anova-based model [1].

TABLE III: Performance comparison of adaptive data collection interval with fixed sample based approach

Parameters	Fixed ($\mathcal{W} = 30$ samples)	Fixed ($\mathcal{W} = 50$ samples)	Fixed ($\mathcal{W} = 100$ samples)	Adaptive ($\mathcal{T}^x \in [30, 60]$ min)
ΔC^x_{avg}	0.083	0.129	0.153	0.0922
Communication energy (J)	166.3277	117.8177	83.1677	110.8877

node is compared assuming the node is battery operated. If E^l_p and E^l_a are respectively the sensing energy consumed up to l cycles in the proposed framework and the Anova-based adaptive sampling algorithm, then the percentage of energy saved is given by,

$$\text{Energy saved} = \left[\frac{1}{L}\sum_{l=1}^{L}\frac{E^l_a - E^l_p}{E^l_a}\right] \times 100\%. \quad (22)$$

From Fig. 7(b), it can be observed that the residual energy of the node at any measurement cycle in the proposed model with $c_{th} = 0.8, 0.7,$ and 0.5 are relatively much higher compared to the Anova-based model. The energy efficiency of the proposed model over the Anova-based model is computed using (22). For $c_{th} = 0.7$, the proposed framework is 46% more energy-efficient compared to the Anova-based adaptive sampling algorithm while maintaining a similar error performance. Considering MRE in the acceptable range (1%), the proposed framework with $c_{th} = 0.5$ saves 54% energy compared to the Anova-based model. For $c_{th} = 0.8$, the proposed model can save 27% energy with 30% improvement in accuracy compared to the Anova-based model.

As discussed in Section IV-B, the data collection interval $\mathcal{T}^x$ in the proposed model is adaptive to the process dynamics. This is in contrast with the Anova-based model [1], where the data collection window (number of samples, $\mathcal{W}$) is fixed. If the i^{th} sensor set is selected at the x^{th} measurement cycle, then the change in cross-correlation factor $\Delta C^x = |C^x_i - C^{x-1}_i|$ should be minimum for optimal operation. The average of ΔC^x (ΔC^x_{avg}) over a large number of measurement cycles is computed to compare the performance of the proposed method with that in [1]. Although ΔC^x_{avg} is lower for lower values of $\mathcal{T}^x$, it increases the communication energy consumption. Let $E_c = E_{c_1} + E_{c_2}$ is the total communication energy consumed by the radio module over a day, where E_{c_1} is the ON state energy required for the communication module and E_{c_2} is the transmission energy. According to [26], E_{c_2} is much higher than E_{c_1}. Thus, E_c is mainly decided by the number of times data is sent to the CE. Considering that 32 bits are required to represent a sample packet, E_{c_1} and E_{c_2} are calculated form the theoretical measurements given in [26]. As shown in Table III, ΔC^x_{avg} increases and E_c reduces with increase in $\mathcal{W}$. The optimal condition is achieved for the adaptive length of measurement cycle, where the ΔC^x_{avg} is slightly higher than that with the minimum fixed window size $\mathcal{W} = 30$, but and E_c is much lower than that with $\mathcal{W} = 30$ and $\mathcal{W} = 50$.

D. Sensitivity of the Proposed Framework to Different Reinforcement Learning Algorithms

The proposed framework uses the UCB algorithm to define the objective function and solve the trade-off between the performance parameters. There are various algorithms available in the literature, such as Thompson sampling, Q-learning, etc., to solve reinforcement learning problems. Q-learning uses the Bellman equation to find the Q-values of the states by estimating future rewards. The proposed framework is considered as stateless, which is modeled as a MAB problem. The MAB problems can be solved using ϵ-greedy, Thompson sampling, UCB algorithm, etc. However, the stateless variant of Q-learning also can be used to solve such problems [27]. The performance of the UCB algorithm in the proposed work is compared with Q-learning and Thompson sampling with respect to sensing energy consumption, cumulative reward, and computation complexity. The results are listed in Table IV. The ϵ-greedy method is used in Q-learning to choose an optimal sensor set, where ϵ is decaying with time. In this simulation initial value of ϵ, decay, and the learning rate are chosen as 1, 0.95, and 0.9, respectively [27]. It has been observed that the total sensing energy consumed at the node and the total reward achieved over a long number of measurement cycles are similar in the case of UCB and Q-learning. However, Thompson sampling performance is poorer. The complexity of UCB, Q-learning, and Thompson sampling are $\mathcal{O}(N)$, $\mathcal{O}(NI_{itr})$, and $\mathcal{O}(N)$, respectively, where I_{itr} is the number of iterations required to find the Q-value of each sensor set. Although in terms of cumulative reward and sensing energy consumption Q-learning performs marginally better, its computational complexity is higher than UCB. Hence, UCB is considered to be appropriate in the proposed framework.

TABLE IV: Performance comparison of different reinforcement learning algorithms in the proposed framework

Algorithms	UCB	Q-learning	Thompson sampling
Sensing energy consumption (J)	2504.2	2502.3	2602
Cumulative reward	27.04	27.08	25.42
Computational complexity	$\mathcal{O}(N)$	$\mathcal{O}(NI_{itr})$	$\mathcal{O}(N)$

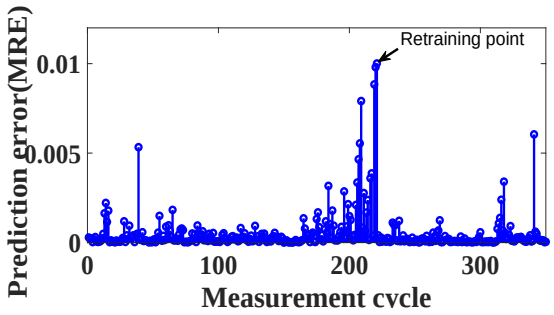


Fig. 8: Variation of prediction error with measurement cycle

E. Signal Recovery

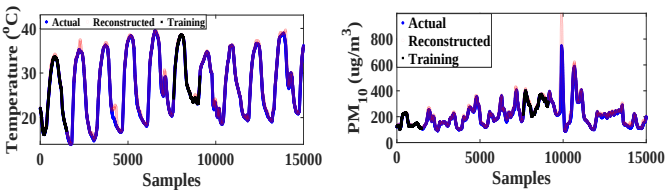
GPR (presented in Section III) directly predicts the time-varying function with a certain confidence interval. Although many sparse signal recovery techniques are discussed in the literature, those are efficient if the data has a large feature dimension and they exhibit sparsity. Unlike that, in the proposed framework the sensor node consists of a limited number of sensors, which makes the feature dimension very small. Moreover, signals with multiple parameters exhibit less sparsity compared to the signals with a single parameter. Therefore GPR is appropriate to use in the proposed framework.

From Fig. 8, it can be observed that the prediction error PE^x given in (11) increases with measurement cycle. Hence, the GPR model is retrained when prediction error exceeds an error threshold. For N sub-models different values of $\epsilon_i; \forall i \in \mathcal{S}$ are chosen such that the MRE remains within the 1% i.e, $\epsilon_i = 0.01 - PE_i$. By exploring the data, $ct_{p,th} = 0.88$ is chosen for all the parameters such that the prediction error remains within the error threshold. The sub-models having MRE higher than 1% are discarded from $\mathcal{S}$, hence the sets corresponding to those sub-models are not considered while choosing an optimal sensor set.

A comparison of actual and reconstructed signals for temperature and PM_{10} are shown in Fig. 9. The actual signal corresponds to the data-set collected from the experimental setup and the reconstructed signal combines the actual data (when the sensor set is active) with the predicted data (when the sensor set is inactive). It can be observed that the reconstructed signals follow the actual signals, hence these are overlapping, which validates the performance of the GPR model.

VII. Concluding Remarks

In this paper, a novel adaptive sensor selection strategy is presented based on the GPR and UCB algorithm. In a non-stationary environment, learning the system with all previous experience gives better insight about the dynamics



(a) Temperature (^oC) (b) PM_{10} in $\mu g/m^3$.

Fig. 9: Actual and reconstructed signals of the air pollution monitoring parameters.

of the system rather than focusing on the immediate previous experience only. The correlation factor developed in this paper acts as a good performance index to indicate the sensing quality of the predicted signals due to the infeasibility of finding actual prediction error. The proposed method finds an optimal sensor set based on the process dynamics with efficient energy management. Extensive studies on the air pollution monitoring data validate that the proposed algorithm is significantly energy-efficient compared to the current state-of-the-art. In future, temporal correlations for each parameter will be exploited to find optimal sampling instants instead of doing at Nyquist rate, which will further reduce the number of samples that need to be collected.

Appendix

Let $\widehat{\mu}_i^x$ be the empirical mean of set i after x measurement cycles; set i is selected for $s \leq x$ times. From [23, Ch. 7], with δ as the confidence bound, we have:

$$Pr(\exists\ s \leq x : \widehat{\mu}_i^x \geq \mu_i^x + \sqrt{\frac{2\ln\frac{1}{\delta}}{T_i^x}}) \leq x\delta. \tag{A.1}$$

Let i be any sub-optimal set and i^* be the optimal. Assume for all x the estimated mean is within the confidence bound.

$$\text{From (15): } \mu_i^{x-1} + \sqrt{\frac{2\ln\frac{1}{\delta}}{T_i^{x-1}}} \geq \widehat{\mu}_i^{x-1}$$

$$\text{From (16): } \widehat{\mu}_{i^*}^{x-1} + \sqrt{\frac{2\ln\frac{1}{\delta}}{T_i^{x-1}}} \geq \mu_{i^*}^{x-1}.$$

Suppose in x^{th} measurement cycle set $\mathcal{A}_i^x$ is activated. If $\mathcal{A}_i^x$ is a sub-optimal set then the loss suffered by the learner is $\Delta_i^x = \mu_{i^*}^x - \mu_i^x$. Thus, we can write:

$$\mu_i^{x-1} + 2\sqrt{\frac{2\ln\frac{1}{\delta}}{T_i^{x-1}}} \geq \widehat{\mu}_i^{x-1} + \sqrt{\frac{2\ln\frac{1}{\delta}}{T_i^{x-1}}} \geq \widehat{\mu}_{i^*}^{x-1} + \sqrt{\frac{2\ln\frac{1}{\delta}}{T_i^{x-1}}}$$
$$\geq \mu_{i^*}^{x-1} = \mu_i^{x-1} + \Delta_i^{x-1}.$$

$$\therefore T_i^{x-1} \leq \frac{8\ln\frac{1}{\delta}}{\left(\Delta_i^{x-1}\right)^2} \text{ and } T_i^x \leq 1 + \frac{8\ln\frac{1}{\delta}}{\left(\Delta_i^x\right)^2}. \tag{A.2}$$

(A.2) infers that the sub-optimal set is not selected too often. Regret decomposition over the sets in x measurement cycles:

$$
\begin{aligned}
Reg^x &= \mathbb{E}[\sum_{t=1}^{x} \mu_{i^*}^t - \mu_i^t] = \mathbb{E}[\sum_{t=1}^{x} \Delta_i^t] \\
&= \mathbb{E} \sum_{t=1}^{x} [\sum_{i \in \mathcal{S}^x} \mathbb{1}_{(A^t = A_i)} \Delta_i] = \sum_{i \in \mathcal{S}^x} \Delta_i \mathbb{E}[T_i^x] \\
&\leq \sum_{i \in \mathcal{S}^x; \Delta_i > 0} \Delta_i [1 + 2x^2\delta + \frac{8 \ln \frac{1}{\delta}}{\Delta_i^2}] \\
&\leq \sum_{i \in \mathcal{S}^x; \Delta_i > 0} (1 + 2x^2\delta)\Delta_i + \sum_{i \in \mathcal{S}^x; \Delta_i > 0} \frac{8 \ln \frac{1}{\delta}}{\Delta_i}.
\end{aligned} \tag{A.3}
$$

The upper bound on regret can be achieved using (A.3). In this decomposition it is assumed that the loss suffered by selecting the i^{th} sub-optimal set at any cycle is the same.

REFERENCES

[1] H. Harb and A. Makhoul, "Energy-efficient sensor data collection approach for industrial process monitoring," *IEEE Trans. Ind. Informat.*, vol. 14, no. 2, pp. 661–672, 2017.

[2] J. Wang, S. Tang, B. Yin, and X.-Y. Li, "Data gathering in wireless sensor networks through intelligent compressive sensing," in *proc. IEEE INFOCOM, Orlando, FL, USA*. IEEE, 2012, pp. 603–611.

[3] A. Jain and E. Y. Chang, "Adaptive sampling for sensor networks," in *proc. of the 1st international workshop on Data management for sensor networks: in conjunction with VLDB 2004, Toronto, Canada*, 2004, pp. 10–16.

[4] C. Alippi, G. Anastasi, M. Di Francesco, and M. Roveri, "An adaptive sampling algorithm for effective energy management in wireless sensor networks with energy-hungry sensors," *IEEE Trans. Instrum. Meas.*, vol. 59, no. 2, pp. 335–344, 2009.

[5] S. Hwang, R. Ran, J. Yang, and D. K. Kim, "Multivariated Bayesian compressive sensing in wireless sensor networks," *IEEE Sensors J.*, vol. 16, no. 7, pp. 2196–2206, 2015.

[6] V. Gupta and S. De, "Collaborative multi-sensing in energy harvesting wireless sensor networks," *IEEE Trans. Signal Inf. Process. Netw.*, vol. 6, pp. 426–441, 2020.

[7] R. Prabha, M. V. Ramesh, V. P. Rangan, P. V. Ushakumari, and T. Hemalatha, "Energy efficient data acquisition techniques using context aware sensing for landslide monitoring systems," *IEEE Sensors J.*, vol. 17, no. 18, pp. 6006–6018, 2017.

[8] J. Liu and C. Ling, "Adaptive compressed sensing using intra-scale variable density sampling," *IEEE Sensors J.*, vol. 18, no. 2, pp. 547–558, 2018.

[9] Y. Xu, J. Choi, S. Dass, and T. Maiti, "Sequential Bayesian prediction and adaptive sampling algorithms for mobile sensor networks," *IEEE Trans. Autom. Control*, vol. 57, no. 8, pp. 2078–2084, 2011.

[10] M. Jadaliha, Y. Xu, J. Choi, N. S. Johnson, and W. Li, "Gaussian process regression for sensor networks under localization uncertainty," *IEEE Trans. Signal Process.*, vol. 61, no. 2, pp. 223–237, 2012.

[11] D. Gu and H. Hu, "Spatial Gaussian process regression with mobile sensor networks," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 23, no. 8, pp. 1279–1290, 2012.

[12] T. Zheng, M. H. Bergin, R. Sutaria, S. N. Tripathi, R. Caldow, and D. E. Carlson, "Gaussian process regression model for dynamically calibrating and surveilling a wireless low-cost particulate matter sensor network in delhi." *Atmospheric Measurement Techniques*, vol. 12, no. 9, 2019.

[13] V. C. Guizilini and F. T. Ramos, "A nonparametric online model for air quality prediction," in *Twenty-Ninth AAAI Conference on Artificial Intelligence, Austin, Texas, USA*, 2015.

[14] Y. Zhang, Y. Wang, M. Gao, Q. Ma, J. Zhao, R. Zhang, Q. Wang, and L. Huang, "A predictive data feature exploration-based air quality prediction approach," *IEEE Access*, vol. 7, pp. 30 732–30 743, 2019.

[15] W. Qiao, W. Tian, Y. Tian, Q. Yang, Y. Wang, and J. Zhang, "The forecasting of PM2. 5 using a hybrid model based on wavelet transform and an improved deep learning algorithm," *IEEE Access*, vol. 7, pp. 142 814–142 825, 2019.

[16] J. Ma, Y. Ding, V. J. Gan, C. Lin, and Z. Wan, "Spatiotemporal prediction of PM2. 5 concentrations at different time granularities using IDW-BLSTM," *IEEE Access*, vol. 7, pp. 107 897–107 907, 2019.

[17] B. Zou, Q. Pu, M. Bilal, Q. Weng, L. Zhai, and J. E. Nichol, "High-resolution satellite mapping of fine particulates based on geographically weighted regression," *IEEE Geosci. Remote Sens. Lett.*, vol. 13, no. 4, pp. 495–499, 2016.

[18] L. Cai, M. Boukhechba, N. Kaur, C. Wu, L. E. Barnes, and M. S. Gerber, "Adaptive passive mobile sensing using reinforcement learning," in *proc. IEEE WoWMoM, Washington, DC, USA, USA*, 2019, pp. 1–6.

[19] V. Gupta and S. De, "SBL-based adaptive sensing framework for WSN-assisted IoT applications," *IEEE Internet Things J.*, vol. 5, no. 6, pp. 4598–4612, 2018.

[20] R. Tan, G. Xing, X. Liu, J. Yao, and Z. Yuan, "Adaptive calibration for fusion-based wireless sensor networks," in *proc. IEEE INFOCOM, San Diego, CA, USA*. IEEE, 2010, pp. 1–9.

[21] C. Rasmussen and C. Williams, "Gaussian processes for machine learning.,(MIT press: Cambridge, MA)," 2006.

[22] S. G. Mallat and Z. Zhang, "Matching pursuits with time-frequency dictionaries," *IEEE Trans. Signal Process.*, vol. 41, no. 12, pp. 3397–3415, 1993.

[23] T. Lattimore and C. Szepesvári, *Bandit algorithms*. Cambridge University Press, 2020.

[24] K. Egoh and S. De, "A multi-criteria receiver-side relay election approach in wireless ad hoc networks," in *proc. IEEE MILCOM, Washington, DC, USA*. IEEE, 2006, pp. 1–7.

[25] J. F. Hemphill, "Interpreting the magnitudes of correlation coefficients." *American Psychological Association*, 2003.

[26] "Espressif inc, esp 8266." [Online]. Available: https://www.espressif.com/sites/default/files/documentation/0a-esp8266ex_datasheet_en.pdf

[27] T. B. de Oliveira, A. L. Bazzan, B. C. da Silva, and R. Grunitzki, "Comparing multi-armed bandit algorithms and Q-learning for multi-agent action selection: a case study in route choice," in *proc. IJCNN, Rio de Janeiro, Brazil*. IEEE, 2018, pp. 1–8.