



Indian Institute of
Technology Delhi



THE UNIVERSITY
OF QUEENSLAND
AUSTRALIA

LEARNING BASED ADAPTIVE SENSOR SELECTION FRAMEWORK FOR MULTI-SENSING WSN

SUSHMITA GHOSH

UQ-IITD ACADEMY OF RESEARCH

INDIAN INSTITUTE OF TECHNOLOGY DELHI

&

THE UNIVERSITY OF QUEENSLAND

JANUARY 2024

© Indian Institute of Technology Delhi (IITD), New Delhi, 2024



Indian Institute of
Technology Delhi



THE UNIVERSITY
OF QUEENSLAND
AUSTRALIA

LEARNING BASED ADAPTIVE SENSOR
SELECTION FRAMEWORK FOR
MULTI-SENSING WSN

by

SUSHMITA GHOSH

UQ-IITD Academy of Research

Submitted

in fulfillment of the requirements for the joint degree of

DOCTOR OF PHILOSOPHY

to the

INDIAN INSTITUTE OF TECHNOLOGY DELHI

&

THE UNIVERSITY OF QUEENSLAND

JANUARY 2024

Dedicated to my parents

Certificate

This is to certify that the thesis entitled “**Learning based Adaptive Sensor Selection Framework for Multi-Sensing WSN**” being submitted by **Ms. Sushmita Ghosh** to the Indian Institute of Technology Delhi and The University of Queensland for the award of the degree of **Doctor of Philosophy** is a record of bonafide research work carried out by her. **Ms. Sushmita Ghosh** has worked under our guidance and supervision and has fulfilled the requirements for the submission of this thesis, which to our knowledge has reached the requisite standard. The results contained in this thesis are original and have not been submitted, in part or full, to any other University or Institute for the award of any other degree or diploma.

Prof. Swades De

IITD supervisor
Department of EE
IIT Delhi
New Delhi, India, 110016

Prof. Shouri Chatterjee

IITD supervisor
Department of EE
IIT Delhi
New Delhi, India, 110016

Prof. Marius Portmann

UQ supervisor
School of ITEE
University of Queensland
St. Lucia, Australia - 4076



Abstract

Wireless sensor networks are gaining enormous attention for monitoring physical conditions in various applications. The network consists of many densely/sparsely deployed wireless sensor nodes. The field-deployed battery-powered sensor nodes equipped with power-hungry sensors often suffer from limited energy availability due to the limited battery capacity. The energy consumption is higher for the sensor hubs equipped with multiple sensors monitoring multiple parameters in the environment. To this end, this dissertation aims at developing efficient smart sensing approaches to enhance the energy sustainability of such WSNs. A sensor node equipped with a sensor monitoring the variation of a particular parameter in time often exhibits a high temporal correlation that has been exploited in this study to smartly sense the parameter. In the case of a multi-sensing node, the signal variation of multiple parameters sensed in the same environment possesses a good cross-correlation among them, which has been exploited to optimize energy consumption in a multi-sensing node.

As a first step, a learning-based optimization strategy is developed for a multi-sensing node 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, the algorithm has been implemented in an air pollution monitoring sensor node consisting of seven sensors. The node has been deployed on the campus of IIT Delhi. It has been observed that the proposed algorithm significantly reduces the energy consumption of the node.

Further, to optimize the sensing energy consumption of each sensor in the node, a learning-based adaptive sampling framework is developed that explores the sparsity in the time series data and finds optimal sampling instants for the next measurement cycle. The Principal component analysis is used to sparsify the time domain signal and the sparse signal is reconstructed from its low-dimensional signal using the sparse Bayesian learning method. An optimization function is formed that solves the trade-off between accuracy and energy consumption and finds the optimal sampling instants for the next measurement cycle. The simulation results validate that the proposed smart sensing method significantly reduces the sensing energy consumption of the individual sensors in the node.

In the next study, a novel edge intelligence-based data-driven dynamic priority sensing and transmission framework is developed to optimize energy efficiency as well as sensing accuracy. The proposed framework jointly exploits the cross-correlation among the sensing parameters and temporal correlation of the individual sensing signals to find an optimal active sensor set and optimal sampling instants of the sensors in the next measurement cycle. The length of measurement cycle is dynamically decided based on the change in cross-correlation among the parameters and the system state. A discounted upper confidence bound algorithm-based optimization function is formulated to find the optimal active sensor set by solving the trade-off among cross-correlation, energy consumption, and length of measurement cycle. The proposed framework uses Gaussian process regressor-based prediction models to estimate the temporal and cross-correlated parameters of the active and inactive sensor set, respectively. The sampling interval of each active sensor is dynamically adapted based on the temporal prediction error. Extensive simulations are performed on air pollution monitoring dataset to validate the efficacy of the proposed framework in both real-time and non-real-time applications.

The real-life deployments of air pollution monitoring systems are sparse, due to large size, high cost, and high-power consumption. Such sparsely deployed sensing stations are unable to provide a fine granular pollution mapping of a given geographical area. By deploying low cost, low power, miniature air pollution monitoring sensor nodes, the air pollution of the whole area can be accurately measured. However, accuracy of the sensed data of the low cost miniature sensing nodes needs to be addressed. Thus, an autocalibration method of low cost MSNs, with the help of sparsely deployed high cost sensing stations is proposed. The datasets from the HCSSs are collected and used to calibrate the MSN using a suitable learning-based regressor model at the nearby edge node. Further, a cross-correlation based method of determining the optimum time to recalibrate the low cost sensors in a multi-sensing node is also proposed. This method eliminates the requirement of taking the MSNs offline to calibrate/re-calibrate them. To apply the proposed autocalibration method, this paper additionally presents the design of a low cost, low power particulate matter sensor. To validate the performance of the low cost PM sensor, the calibrated PM data are compared with the data collected from a colocated commercially available PM sensor, which is considered as a reference. The low cost PM sensor is 91% more cost efficient and 57% more energy efficient compared to the commercial high cost PM sensor, while maintaining the sensing error within a given threshold.

सार

(Abstract in Hindi)

विभिन्न अनुप्रयोगों में भौतिक स्थितियों की निगरानी के लिए वायरलेस सेंसर नेटवर्क पर अत्यधिक ध्यान दिया जा रहा है। नेटवर्क में कई सघन / कम तैनात वायरलेस सेंसर नोड होते हैं। बिजली की खपत वाले सेंसर से लैस क्षेत्र-तैनात बैटरी-संचालित सेंसर नोड्स अक्सर सीमित बैटरी क्षमता के कारण सीमित ऊर्जा उपलब्धता से पीड़ित होते हैं। पर्यावरण में कई मापदंडों की निगरानी करने वाले कई सेंसर से लैस सेंसर हब के लिए ऊर्जा की खपत अधिक होती है। इसके लिए, इस शोध प्रबंध का उद्देश्य ऐसे डब्ल्यूएसएन की ऊर्जा स्थिरता को बढ़ाने के लिए कुशल स्मार्ट सेंसिंग दृष्टिकोण विकसित करना है। समय में एक विशेष पैरामीटर की भिन्नता की निगरानी करने वाले सेंसर से लैस एक सेंसर नोड अक्सर एक उच्च अस्थायी सहसंबंध प्रदर्शित करता है जिसका उपयोग इस अध्ययन में स्मार्ट तरीके से पैरामीटर को समझने के लिए किया गया है। मल्टी-सेंसिंग नोड के मामले में, एक ही वातावरण में महसूस किए गए कई मापदंडों की सिग्नल भिन्नता उनके बीच एक अच्छा क्रॉस-सहसंबंध रखती है, जिसका बहु-सेंसिंग नोड में ऊर्जा खपत को अनुकूलित करने के लिए शोषण किया गया है।

पहले कदम के रूप में, अपर कॉन्फिडेंस बाउंड एल्गोरिद्म का उपयोग करते हुए मल्टी-सेंसिंग नोड के लिए एक लर्निंग-बेस्ट ऑप्टिमाइजेशन रणनीति विकसित की गई है, जो पैरामीटर्स के बीच क्रॉस-सहसंबंधों, सेंसर द्वारा खपत ऊर्जा के आधार पर माप चक्र में एक इष्टतम सक्रिय सेंसर सेट का चयन करने के लिए है। , और नोड पर उपलब्ध ऊर्जा। इसके अलावा, सक्रिय सेंसर के क्रॉस-सहसंबद्ध पैरामीटर से निष्क्रिय सेंसर के पैरामीटर मानों की भविष्यवाणी करने के लिए एक गॉसियन प्रक्रिया रेजिस्टर-आधारित भविष्यवाणी मॉडल का उपयोग किया जाता है। वास्तविक जीवन के अनुप्रयोगों में प्रस्तावित ढांचे के प्रदर्शन का मूल्यांकन करने के लिए, एल्गोरिद्म को वायु प्रदूषण निगरानी सेंसर नोड में लागू किया गया है जिसमें सात सेंसर शामिल हैं। नोड को IIT दिल्ली के परिसर में तैनात किया गया है। यह देखा गया है कि प्रस्तावित एल्गोरिद्म नोड की ऊर्जा खपत को काफी कम कर देता है।

इसके अलावा, नोड में प्रत्येक सेंसर की ऊर्जा खपत को अनुकूलित करने के लिए, एक सीखने-आधारित अनुकूली नमूनाकरण ढांचा विकसित किया गया है जो समय श्रृंखला डेटा में विरलता की पड़ताल करता है और अगले माप चक्र के लिए इष्टतम नमूनाकरण का पता लगाता है। प्रिंसिपल कंपोनेंट एनालिसिस का उपयोग टाइम डोमेन सिग्नल को कम करने के लिए किया जाता है और स्पास बायेसियन लर्निंग मेथड का उपयोग करके इसके लो-डायमेंशनल सिग्नल से स्पास सिग्नल को फिर से बनाया जाता है। एक ऑप्टिमाइजेशन फ़ंक्शन बनता है जो सटीकता और ऊर्जा की खपत के बीच व्यापार-बंद को हल करता है और अगले माप चक्र के लिए इष्टतम नमूनाकरण का पता लगाता है। सिमुलेशन के परिणाम इस बात की पुष्टि करते हैं कि प्रस्तावित स्मार्ट सेंसिंग विधि नोड में अलग-अलग सेंसर की सेंसिंग ऊर्जा खपत को काफी कम कर देती है।

अगले अध्ययन में, ऊर्जा दक्षता के साथ-साथ संवेदन सटीकता को अनुकूलित करने के लिए एक उपन्यास एज इंटेलिजेंस-आधारित डेटा-संचालित गतिशील प्राथमिकता संवेदन और संचरण ढांचा विकसित किया गया है। प्रस्तावित रूपरेखा संयुक्त

रूप से अगले माप चक्र में एक इष्टतम सक्रिय सेंसर सेट और सेंसर के इष्टतम नमूने के उदाहरणों को खोजने के लिए संवेदन मापदंडों और व्यक्तिगत संवेदी संकेतों के अस्थायी सहसंबंध के बीच क्रॉस-सहसंबंध का शोषण करती है। माप चक्र की लंबाई गतिशील रूप से पैरामीटर और सिस्टम स्थिति के बीच क्रॉस-सहसंबंध में परिवर्तन के आधार पर तय की जाती है। क्रॉस-सहसंबंध, ऊर्जा खपत और माप चक्र की लंबाई के बीच व्यापार-बंद को हल करके इष्टतम सक्रिय सेंसर सेट खोजने के लिए एक रियायती ऊपरी आत्मविश्वास बाध्य एल्गोरिदम-आधारित अनुकूलन फ़ंक्शन तैयार किया गया है। प्रस्तावित रूपरेखा क्रमशः सक्रिय और निष्क्रिय सेंसर सेट के अस्थायी और क्रॉस-सहसंबंधित मापदंडों का अनुमान लगाने के लिए गॉसियन प्रक्रिया प्रतिगामी-आधारित भविष्यवाणी मॉडल का उपयोग करती है। प्रत्येक सक्रिय सेंसर का नमूना अंतराल अस्थायी भविष्यवाणी त्रुटि के आधार पर गतिशील रूप से अनुकूलित होता है। वास्तविक समय और गैर-वास्तविक समय दोनों अनुप्रयोगों में प्रस्तावित ढांचे की प्रभावकारिता को मान्य करने के लिए वायु प्रदूषण निगरानी डेटासेट पर व्यापक सिमुलेशन किए जाते हैं।

बड़े आकार, उच्च लागत और उच्च बिजली की खपत के कारण वायु प्रदूषण निगरानी प्रणालियों की वास्तविक जीवन तैनाती विरल है। इस तरह के कम तैनात संवेदन स्टेशन किसी दिए गए भौगोलिक क्षेत्र के सूक्ष्म दानेदार प्रदूषण मानचित्रण प्रदान करने में असमर्थ हैं। कम लागत, कम बिजली, लघु वायु प्रदूषण निगरानी सेंसर नोड्स तैनात करके, पूरे क्षेत्र के वायु प्रदूषण मानचित्र को सटीक रूप से मापा जा सकता है। हालांकि, कम लागत वाले मिनिएचर सेंसिंग नोड्स के सेंसिबिलिटी डेटा की सटीकता पर ध्यान देने की जरूरत है। इस प्रकार, कम तैनात उच्च लागत संवेदन स्टेशनों की मदद से कम लागत वाले एमएसएन की एक ऑटोकैलिब्रेशन विधि प्रस्तावित है। एचसीएसएस से डेटासेट एकत्र किए जाते हैं और एमएसएन को कैलिब्रेट करने के लिए पास के किनारे नोड पर एक उपयुक्त शिक्षण-आधारित प्रतिगामी मॉडल का उपयोग किया जाता है। इसके अलावा, मल्टी-सेंसिंग नोड में कम लागत वाले सेंसर को फिर से कैलिब्रेट करने के लिए इष्टतम समय का निर्धारण करने के लिए एक क्रॉस-सहसंबंध आधारित पद्धति भी प्रस्तावित है। यह विधि एमएसएन को कैलिब्रेट/री-कैलिब्रेट करने के लिए ऑफ़लाइन लेने की आवश्यकता को समाप्त करती है। प्रस्तावित ऑटोकैलिब्रेशन पद्धति को लागू करने के लिए, यह पेपर अतिरिक्त रूप से कम लागत, कम पावर पार्टिकुलेट मैटर सेंसर का डिज़ाइन प्रस्तुत करता है। कम लागत वाले पीएम सेंसर के प्रदर्शन को मान्य करने के लिए, कैलिब्रेटेड पीएम डेटा की तुलना व्यावसायिक रूप से उपलब्ध पीएम सेंसर से एकत्रित डेटा के साथ की जाती है, जिसे एक संदर्भ माना जाता है। कम लागत वाला पीएम सेंसर वाणिज्यिक उच्च लागत वाले पीएम सेंसर की तुलना में 91% अधिक लागत कुशल और 57% अधिक ऊर्जा कुशल है, जबकि एक निश्चित सीमा के भीतर संवेदन त्रुटि को बनाए रखता है।

Declaration by author

This thesis is composed of my original work, and contains no material previously published or written by another person except where due reference has been made in the text. I have clearly stated the contribution by others to jointly authored works that I have included in my thesis.

I have clearly stated the contribution of others to my thesis as a whole, including statistical assistance, survey design, data analysis, significant technical procedures, professional editorial advice, financial support and any other original research work used or reported in my thesis. The content of my thesis is the result of work I have carried out since the commencement of my higher degree by research candidature and does not include a substantial part of work that has been submitted to qualify for the award of any other degree or diploma in any university or other tertiary institution. I have clearly stated which parts of my thesis, if any, have been submitted to qualify for another award.

I acknowledge that an electronic copy of my thesis must be lodged with the University Library and, subject to the policy and procedures of The University of Queensland, the thesis be made available for research and study in accordance with the Copyright Act 1968 unless a period of embargo has been approved by the Dean of the Graduate School.

I acknowledge that an electronic copy and hard copy of my thesis must be lodged with the Institute Library and, subject to the policy and procedures of The Indian Institute of Technology Delhi, the thesis be made available for research and study in accordance with the Copyright Policy of IIT Delhi.

I acknowledge that copyright of all material contained in my thesis resides with the copyright holder(s) of that material. Where appropriate I have obtained copyright permission from the copyright holder to reproduce material in this thesis and have sought permission from co-authors for any jointly authored works included in the thesis.

Publications

Refereed Journals

1. **S. Ghosh**, S. De, S. Chatterjee, and M. Portmann, "Learning-Based Adaptive Sensor Selection Framework for Multi-Sensing WSN," in *IEEE Sensors Journal*, vol. 21, no. 12, pp. 13551-13563, 15 June, 2021. (Chapter 2)
2. **S. Ghosh**, S. De, S. Chatterjee, and M. Portmann, "Edge Intelligence Framework for Data-Driven Dynamic Priority Sensing and Transmission", *IEEE Transactions on Green Communications and Networking*, vol. 6, no. 1, pp. 376-390, March 2022. (Chapter 4)
3. **S. Ghosh**, P. Das, S. De, S. Chatterjee and M. Portmann, "Local Reference-Free In-Field Calibration of Low-Cost Air Pollution Monitoring Sensors", *IEEE Transactions on Instrumentation and Measurement*, vol. 71, pp. 1-13, 2022, Art no. 2517613, Sep. 2022. (Chapter 5)
4. **S. Ghosh**, A. K. Mandal, S. De, S. Chatterjee, and M. Portmann, "Light-weight ML aided autonomous IoT networks", *IEEE Communications Magazine*, Jan. 2023.

Refereed Conferences

1. **S. Ghosh**, S. De, S. Chatterjee, and M. Portmann, "Learning-based Smart Sensing for Energy-Sustainable WSN," in *Proc. IEEE ICC 2021*, Montreal, Canada, June 2021. (Chapter 3)
2. **S. Ghosh**, P. Das, S. Muruges, S. De, S. Chatterjee, and M. Portmann, "Energy Aware Smart Sensing and Implementation in Green Air Pollution Monitoring System," in *Proc. IEEE ICC 2023*, Rome, Italy, May 2023.

Other Publications During PhD

Refereed Patents

1. S. Chatterjee, **S. Ghosh**, and S. De, "A system for optimizing energy transmission and a method thereof," applied for Indian Patent, ref. no. 202111043815, Sep. 2021.
2. S. Chatterjee, P. Das, **S. Ghosh**, and S. De, "Pollution monitoring system and method thereof," applied for Indian Patent, ref. no. 202111043780, Sep. 2021.

Refereed Journals

1. P. Das, **S. Ghosh**, S. Chatterjee, and S. De, "A low cost outdoor air pollution monitoring device with power controlled built-in PM sensor," IEEE Sensors Journal, May 2022.

Refereed Conferences

1. W. A. Khan, P. Das, **S. Ghosh**, M. Roy Chowdhury, S. Tripathi, S. Kaur, S. Chatterjee, and S. De, "Smart IoT Communication: Circuits and Systems," Research exhibit, in Proc. COMSNETS, pp. 1-2, Bangalore, India, Jan. 2020. (The runner-up research demo award)
2. P. Das, **S. Ghosh**, S. Chatterjee, and S. De, "Energy harvesting-enabled 5G advanced air pollution monitoring device," in Proc. IEEE 5F World Forum, Bangalore, India, September, 2020.
3. P. Das, **S. Ghosh**, W. A. Khan, S. Chatterjee, and S. De, "Prototype implementation of dynamic data pruning in smart energy meter," in Proc. IEEE GLOBECOM Wksp., pp. 1-6, Madrid, Spain, Dec. 2021.
4. S. Biswas, Himanshu, **S. Ghosh**, P. Das, K. Saha, and S. De, "Efficient data transfer mechanism for DLMS/COSEM enabled smart energy metering platform," in Proc. ACM SIGMETRICS/ IFIP Performance, Poster paper, pp. 1-3, Mumbai, India, Jun. 2022.

Contributions by others to the thesis

Editorial suggestions for all the chapters of my dissertation were provided by the supervisors, Prof. Swades De and Prof. Shouri Chatterjee from the Indian Institute of Technology Delhi, India, and Prof. Marius Portmann from the University of Queensland, Australia, and the Ph.D. committee, Prof. Seshan Srirangarajan (Chairperson from IIT Delhi), Prof. Ankesh Jain (Internal expert from IIT Delhi), and Prof. Dan Kim (External expert from UQ).

Statement of parts of the thesis submitted to qualify for the award of another degree

No works submitted towards another degree have been included in this thesis.

Research involving human or animal subjects

No animal or human subjects were involved in this research.

Acknowledgements

I would like to express my deep and sincere gratitude to my supervisors, Prof. Swades De and Prof. Shouri Chatterjee from the Indian Institute of Technology Delhi, India, and Prof. Marius Portmann from the University of Queensland, Australia for their enormous support, endless help, and invaluable guidance during my Ph.D., which led to the fruition of this research work. Their insightful advice and continuous encouragement during the regular discussions helped me to accumulate skills, knowledge, and professional experiences. Over the years, I have learned a lot from them and it is a great honor to work under their supervision.

I would then like to thank my committee members, Prof. Seshan Srirangarajan (Chairperson from IIT Delhi), Prof. Ankesh Jain (Internal expert from IIT Delhi), and Prof. Dan Kim (External expert from UQ). Their valuable suggestions have significantly improved the quality of my dissertation. I would like to extend my heartfelt gratitude to other professors with whom I have interacted during different courses including Prof. Prathosh A P, Prof. Shantanu Pal, and Prof. Deepak U. Patil. I consider myself fortunate enough to be a part of great institutions like IIT Delhi and UQ, where I have gained knowledge from the excellent courses under the Department of Electrical Engineering.

I am greatly indebted to Prof. Bidyut K. Bhattacharyya from the National Institute of Technology, Agartala (later at Georgia Institute of Technology, USA), who encouraged me to pursue my career in the field of research and academia. I also express my sincere gratitude to him for encouraging me to join the department of UQ-IITD joint Ph.D. program and the research group of Prof. Swades De at IIT Delhi.

Finally, I would like to thank my mother Smt. Doli Ghosh, my father Sri. Swapan Ghosh, and the other family members for their consistent support, and encouragement during my Ph.D. study. Without their constant support and inspiration, this acknowledgment and PhD dissertation would have never appeared.



Sushmita Ghosh

Financial support

This research was supported in parts by an Indian Institute of Technology Delhi Research Scholarship, by the Prime Minister's Doctoral Research Fellowship, by a Research grant from the UQ-IITD Academy of Research (UQIDAR), by the Science and Engineering Board, Department of Science and Technology (DST), Government of India, under Grant CRG/2019/002293, and by the Indian National Academy of Engineering (INAE) through the Abdul Kalam Technology Innovation National Fellowship.

Keywords

Temporal correlation, cross-correlation, edge intelligence, optimal sensor set, optimal sampling instants, adaptive sensor selection framework, energy saving, Gaussian process regression model, multi-parameter sensor hub, upper confidence bound algorithm.

Australian and New Zealand standard research classifications (ANZSRC)

ANZSRC code: 090609 Signal Processing, 60%

ANZSRC code: 080101 Adaptive Agents and Intelligent Robotics, 20%

ANZSRC code: 090699 Electrical and Electronic Engineering not elsewhere classified, 20%

Fields of research classification

FoR code: 0906 Electrical and Electronic Engineering, 70%

FoR code: 0801 Artificial Intelligence and Image Processing, 30%

Contents

Abstract	i
Abstract in Hindi	iii
Declaration by author	v
Publications	vii
Contributions by others to the thesis	ix
Acknowledgements	xi
Financial support	xiii
Table of Contents	xv
List of Figures	xix
List of Tables	xxiii
List of Symbols	xxv
List of Abbreviations	xxvii
1 Introduction	1
1.1 Background	1
1.2 Energy Efficient Sensor Data Collection Mechanisms	2
1.3 Need for Sustainable IoT Solutions	4
1.4 Need for Calibration of Low Cost Sensors	7

1.5	Organization	9
2	Cross-Correlation based Adaptive Sensing Framework	11
2.1	Introduction	11
2.2	Chapter Organization	12
2.3	System Model	14
2.4	GPR based Prediction Model	15
2.5	Proposed Sensor Selection Framework	17
2.5.1	Modified UCB Algorithm for Sensor Selection	18
2.5.2	Adaptive Data Collection Interval	22
2.5.3	Feedback Mechanism for Adaptive Sensor Selection	23
2.5.4	Complexity of the Proposed Framework	24
2.6	Experimental Setup	25
2.7	Results and Discussion	27
2.7.1	Finding Optimum Parameter Values	28
2.7.2	Simulation Results of the Proposed Algorithm	28
2.7.3	Performance Comparison with State-of-the-Art	30
2.7.4	Sensitivity of the Proposed Framework to Different Reinforce- ment Learning Algorithms	33
2.7.5	Examples of Individual Signal Recovery	34
2.8	Summary	35
3	Temporal Correlation based Adaptive Sensing Framework	37
3.1	Introduction	37
3.2	Chapter Organization	38
3.3	Preliminaries	38
3.3.1	Data Acquisition Model	38
3.3.2	Sparse Signal Representation and Signal Recovery	40
3.4	Adaptive Sampling Framework	42
3.4.1	Complexity of the Proposed Framework	44
3.5	Experimental Setup and Results	45
3.5.1	Experimental Setup for Data Collection	45
3.5.2	Results and Discussions	46
3.6	Summary	49

CONTENTS

4	Edge Intelligence Framework for Data-Driven Dynamic Priority Sensing and Transmission	51
4.1	Introduction	51
4.2	Chapter Organization	53
4.3	System Model and Problem Definition	53
4.3.1	System Model	53
4.3.2	Problem Definition	55
4.4	Characterization of Sensing Signal Dynamics	56
4.5	GPR-based Temporal and Cross-Correlated Parameter Prediction Model	58
4.5.1	GPR Model for Time Series Prediction	59
4.5.2	GPR Model for Cross-Correlated Parameter Prediction	59
4.6	Proposed Framework	60
4.6.1	Optimal Length of Measurement Cycle	61
4.6.2	Optimal Lead Values of the Sensing Parameters	63
4.6.3	Discounted UCB-based Optimization Function for Sensor Selection	64
4.6.3.1	Concavity of the Proposed Optimization Function	66
4.6.4	Data-Driven Dynamic Priority Sensing and Transmission Mechanism	67
4.6.5	Computational Complexity of the Proposed Framework	70
4.7	Simulation Results and Discussion	71
4.7.1	Finding Optimum Values of Hyper-Parameters	72
4.7.2	Validation of the Proposed Algorithm	74
4.7.3	‘bad’ States Detection Accuracy	75
4.7.4	Reconstruction Performance Evaluation	76
4.7.5	Performance Comparison with State-of-the-Art	77
4.7.6	Sensitivity of the Proposed Framework to RL-based Optimization Functions	79
4.8	Experimental Demonstration of the Proposed Framework	79
4.8.1	Experimental Setup	80
4.8.2	Experimental Results and Discussion	81
4.9	Summary	83
5	Autocalibration of Low Cost Sensors	85
5.1	Introduction	85

CONTENTS

5.2	System Model	86
5.3	Proposed Local Reference Free Calibration	87
5.3.1	Local Reference Free Calibration Method	87
5.3.2	Machine Learning-based Calibration Model	90
5.3.3	Automatic Recalibration of Low Cost Air Pollution Monitoring Sensors in MSN	92
5.4	Experimental Setup for Data Collection	94
5.4.1	Design of Low Cost PM sensor	94
5.4.2	Deployment of Low Cost PM Sensor	95
5.5	Results and Discussions	96
5.5.1	Estimation of Actual Data \bar{Z}_p^{act} from Long Term Averaged Esti- mated Data \bar{Z}_p^{est}	96
5.5.2	Variation of Prediction Error with Distance of MSN from HCSSs	99
5.5.3	GPR based Calibration Model	100
5.5.4	Automatic Recalibration Performance	102
5.5.5	Recalibration Overhead	103
5.5.6	Energy Saving Performance	105
5.6	Summary	106
6	Concluding Remarks and Future Works	109
6.1	Concluding Remarks	109
6.2	Future Works	111
	Bibliography	112
	Biodata of the Author	121

List of Figures

1.1	Taxonomy of state-of-the-art adaptive sensing frameworks.	3
1.2	Taxonomy of state-of-the-art calibration methods.	8
2.1	(a) Optimum correlation threshold 0.5 and (b) Optimum training length 1500.	28
2.2	Finding optimum value of the hyper-parameters.	28
2.3	Simulation results of the proposed algorithm, for $c_{th} = 0.5$, $l = 1$, $\sigma =$ 0.0005	29
2.4	a)Trade-off between sensing energy consumption and prediction error, b) Average energy consumption and prediction error versus cross-correlation threshold.	29
2.5	Network life time with/without using E_0 in the reward.	30
2.6	Comparison of (a) energy consumption and (b) error of the proposed framework with Anova-based model [1].	30
2.7	Variation of prediction error with measurement cycle	35
2.8	Actual and reconstructed signals of the air pollution monitoring parameters.	35
3.1	Prototype of air pollution monitoring sensor board.	45
3.2	(a) Power spectral density of the sensing signals, (b) measurement cycle versus sensing and communication energy, and (c) Number of samples collected for temperature and PM _{2.5} parameters at various measurement cycles.	46

LIST OF FIGURES

3.3	Comparison of (a) number of samples, (b) reconstruction error of temperature, (c) reconstruction error of PM _{2.5} , and (d) Total energy consumed by the node at different measurement cycles in the proposed framework, Anova-based adaptive sampling [1], and Nyquist-based adaptive sampling [2].	47
3.4	Reconstruction performance of the proposed algorithm.	47
4.1	Smart sensing system model.	53
4.2	(a) Two-state Markov model of sensing signal, (b) distribution of 1 st order time derivative of PM _{2.5} , and (c) distribution of 1 st order time derivative of CO.	56
4.3	Variation of temporal prediction error with l_p	71
4.4	Variation of temporal prediction error and temporal correlation with lead value ζ^t	72
4.5	(a) Average energy consumption and reconstruction error versus cross-correlation threshold and (b) Variation of cross-correlation coefficient with ζ^c	74
4.6	(a) Optimal active sensor set, (b) adaptive length of measurement cycle, and (c) deviation of average cross-correlation coefficient of the active set.	75
4.7	Comparison of (a) number of ‘bad’ state detected, (b) residual energy at the sensor node, and (c) reconstruction error of the proposed framework with Nyquist-based framework [2].	75
4.8	Comparison of actual and reconstructed sensing signals.	77
4.9	Reconstruction error of the parameters at the edge node.	77
4.10	Performance comparison of the proposed framework with SBL-based approach [3].	78
4.11	Flow chart of the smart sensing algorithm programmed in the edge node.	80
4.12	The designed air pollution monitoring board, (a) the front side of the board with off-the-shelf gas sensors and DHT sensor, (b) the back side of the board with on-board PM sensor circuitry, and (c) Deployed IP67 rated prototype with solar panel in IIT Delhi campus.	81
4.13	Solar power harvested in a day.	82
4.14	Prediction error and temporal correlation versus ζ	82
4.15	Performance comparison of the algorithm implemented in the air pollution monitoring system.	83

LIST OF FIGURES

4.16	Actual and reconstructed signals at the edge node.	83
5.1	System model	87
5.2	Estimate the sensing parameter values at point P_0 using section formula. .	89
5.3	Cross-correlation among the sensing parameters collected from high cost reference sensors.	93
5.4	Prototype implementation of low cost PM sensor.	94
5.5	(a) Experimental setup for data collection; (b) high cost air pollution monitoring nodes deployed by CPCB [4].	95
5.6	Short term variation of (a) Estimated $PM_{2.5}$ and actual $PM_{2.5}$; (b) Estimated PM_{10} and actual PM_{10}	97
5.7	Variation of mean absolute error of (a) $PM_{2.5}$, and (b) PM_{10} with averaging period.	97
5.8	Variation of long term averaged (a) $PM_{2.5}$ and (c) PM_{10} with time, (b) Estimated $PM_{2.5}$ versus actual $PM_{2.5}$, and (d) Estimated PM_{10} versus actual PM_{10}	98
5.9	Comparison of the estimated, actual, and predicted long term average values of (a) $PM_{2.5}$, and (b) PM_{10}	99
5.10	Long term average of the particulate matter (PM) data collected from the uncalibrated MSN and the high cost reference sensing station installed in IIT Delhi campus.	101
5.11	Comparison of uncalibrated, actual, and calibrated sensor data: (a) Long term average values of $PM_{2.5}$; (b) Instantaneous values of $PM_{2.5}$; (c) Variation of calibration error of $PM_{2.5}$	101
5.12	Comparison of uncalibrated, actual, and calibrated sensor data: (a) Long term average values of PM_{10} ; (b) Instantaneous values of PM_{10} ; (c) Variation of calibration error of PM_{10}	101
5.13	Variation of training and cross validation error of (a) $PM_{2.5}$, and (b) PM_{10} with number of training examples.	102
5.14	Comparison of cross-correlation among the sensing parameters collected from calibrated and uncalibrated low cost sensors.	103
5.15	GPR based calibration of the uncalibrated sensing parameters.	104
5.16	(a) Variation of cross-correlation between temperature and $PM_{2.5}$, calibration error of $PM_{2.5}$ with time, and (b) Cross-correlation based adaptive retaining of the calibration models.	104

List of Tables

1.1	Comparison of state-of-the-art	5
2.1	List of Symbols	13
2.2	Specifications of sensors	26
2.3	Reconstruction error (MRE)	31
2.4	Performance comparison of adaptive data collection interval with fixed sample based approach	32
2.5	Performance comparison among different reinforcement learning algo- rithms	34
3.1	List of Symbols	39
3.2	Performance comparison	48
4.1	List of Symbols	54
4.2	Specification of sensors	71
4.3	Comparison with state-of-the-art	75
4.4	Performance comparison of the proposed framework with and without RL	79
4.5	Performance comparison among different cases	84
5.1	List of Symbols	88
5.2	Comparison of error between predicted and actual data for different re- gressors	99
5.3	Variation of prediction error with the area of triangle	99
5.4	Comparison of sensing error between calibrated MSNs data and actual data	101
5.5	Performance comparison	106

List of Symbols

P	Number of parameters
N'	Number of sensor sets
X	Number of measurement cycles
\mathcal{P}	Parameter set
$z_p^x(i)$	i^{th} sample of the p^{th} parameter at x^{th} measurement cycle
\mathbf{z}_p^x	x^{th} measurement vector of the p^{th} parameter
\mathbf{y}_p^x	x^{th} true signal vector of the p^{th} parameter
η_p^x	x^{th} noise vector of the p^{th} parameter
N_p^x	Number of samples collected for p^{th} parameter at x^{th} cycle
τ^x	Length of x^{th} measurement cycle
t_p^x	Sampling interval of p^{th} parameter at x^{th} cycle
\mathcal{A}_i^x	i^{th} active sensor
\mathcal{B}_i^x	i^{th} inactive sensor set
A_i^x	Cardinality of \mathcal{A}_i^x
M_p^x	Number of optimal sampling instants of the p^{th} parameter at x^{th} cycle
$c^x(m, k)$	cross-correlation between m and k at the x^{th} measurement cycle
c_{th}	correlation threshold
C_i^x	Average cross-correlation between the parameters of \mathcal{A}_i^x and \mathcal{B}_i^x
Es_i^x	Sensing energy consumption of \mathcal{A}_i^x
Et_i^x	Transmission energy consumption of \mathcal{A}_i^x
E_c	Communication energy consumption
PE_i^x	Prediction error of \mathcal{B}_i^x
E_{batt}	Battery capacity of the node
E_0^x	Energy available at node in x^{th} measurement cycle
$\mathbf{Z} \in \mathbb{R}^{m \times n}$	Data matrix of size $m \times n$
I^x	Number of samples collected in x^{th} cycle for each parameter

LIST OF SYMBOLS

T_i^x	Number of times the i^{th} sensor set is activated
Pg_p^x	good state probability of p^{th} parameter at x^{th} cycle
Pb_p^x	bad state probability of p^{th} parameter at x^{th} cycle
R_i^x	Reward obtained for selecting i^{th} active set at the x^{th} measurement cycle
δ	Confidence bound
n	Number of training/retraining samples
MRE_p	MRE for a parameter $p \in \mathcal{P}$
$nMSE_p^x$	Normalized mean squared error of the p^{th} parameter in x^{th} cycle
ct_p^x	Temporal correlation of p^{th} parameter at x^{th} cycle
$et_p(n)$	Temporal prediction error of the p^{th} parameter for the n^{th} sample
$Z_p^n(i)$	i^{th} instantaneous value of the p^{th} parameter at the n^{th} node
\bar{Z}_p^n	Long term average of p^{th} parameter at n^{th} node

List of Abbreviations

APMD	Air Pollution Monitoring Device
BCRB	Bayesian Cremer-Rao Bound
CE	Central Entity
FFT	Fast Fourier Transform
GPR	Gaussian Process Regressor
HCSS	High Cost Sensing Station
IoT	Internet of Things
KF	Kalman Filtering
MAB	Multi-Armed bandit
MRE	Mean Relative Error
MSE	Mean Square Error
MSN	Miniature Sensing Node
nMSE	Normalized Mean Square Error
OMP	Orthogonal Matching Pursuit
PCA	Principal Component Analysis
PE	Prediction Error
PM	Particulate Matter
RE	Reconstruction Error
RMSE	Root Mean Square Error
SBL	Sparse Bayesian Learning
UCB	Upper Confidence Bound
WSN	Wireless Sensor Network