Spatial Correlation-based Efficient Communication Protocols for Wireless Sensor Networks

by

RAJEEV K. SHAKYA

DEPARTMENT OF ELECTRICAL ENGINEERING

INDIAN INSTITUTE OF TECHNOLOGY, KANPUR, INDIA

February, 2014

Spatial Correlation-based Efficient Communication Protocols for Wireless Sensor Networks

A Thesis Submitted

in Partial Fulfillment of the Requirements

for the Degree of

DOCTOR OF PHILOSOPHY

by

RAJEEV K. SHAKYA

to the

DEPARTMENT OF ELECTRICAL ENGINEERING

INDIAN INSTITUTE OF TECHNOLOGY, KANPUR, INDIA

February, 2014

CERTIFICATE

It is certified that the work contained in the thesis titled **Spatial Correlation-based Efficient Communication Protocols for Wireless Sensor Networks**, by **Rajeev Kumar Shakya**, has been carried out under my supervision. In my opinion, the thesis has reached the standard fulfilling the requirement of regulation of the Ph.D. degree. The work embodied in this thesis has not been submitted elsewhere for the award of any degree or diploma.

Dr. Yatindra Nath Singh
Department of Electrical Engineering,
Indian Institute of Technology,
Kanpur -208016.

Dr. Nishchal K. Verma
Department of Electrical Engineering,
Indian Institute of Technology,
Kanpur -208016.

SYNOPSIS

Name of the Student: Rajeev Kumar Shakya Roll Number: Y8104073

Degree for which submitted: Ph.D. Department: Electrical Engineering

Thesis Title: **Spatial Correlation-based Efficient Communication Protocols for Wireless Sensor Networks**

Thesis Supervisors: Dr. Yatindra Nath Singh and Dr. Nishchal K. Verma

Month and year of submission: **February, 2014**

Wireless Sensor Networks (WSNs) is an emerging area of mobile adhoc networks that presents novel networking issues because of their different application requirements, limited resources capabilities and functionalities, small packet size, and the dynamic multi-hop transmissions. Applications include environmental and habitat monitoring such as forest fire detection, air pollution and rainfall observation in agriculture; they also extend to the military, education, monitoring of machine health, and medical diagnostics. It is unlikely to have a unified communication protocol which is suitable for all the sensor network applications. In this context, we present the design and implementation of communication protocols for wireless sensor networks that are tailored for event-driven systems. Since the sensor nodes are densely deployed for sufficient data resolution, the redundancy of sensor nodes results in highly correlated data transmissions from nodes due to overlapped sensing regions. The degree of spatial correlation increases with the increase in overlapped sensing areas between the nodes (i.e., decreasing inter-node distance). Therefore, every sensor node observing the event need not be active for sensing and communication. A significant amount of energy-saving is possible by taking advantage of spatial correlation. Besides, for the dense WSNs, when the number of nodes is more than the minimum required in the field, multiple nodes detect

the same event simultaneously. As a result, the correlated transmissions and collisions occur unnecessarily. The motivation is to design energy-efficient communication protocols (a medium access control and a routing protocol) for wireless sensor networks that are tailored for the event-driven systems.

In this thesis, a basic spatial correlation function has been proposed to model the correlation characteristics of the event information observed by multiple sensors with omni-directional coverage and communication capabilities. Two frameworks – mathematical and analytical – using this new alternative correlation model have been presented at both medium access control (MAC) and network layers. At the network layer, it is demonstrated that how, through proper tuning of both the sensing range and the correlation threshold, WSNs can be partitioned into the disjoint correlated clusters without any degradation in the information reliability, thus enabling significant energy-saving during data collection. At the MAC layer, extending the work of Vuran and Akyildiz [14], a theoretical framework has been developed to estimate the event source and reconstruct the distortion at the sink. The impact of correlation between nodes on achieved distortion at the sink is investigated using various parameters such as sensing range, selected node numbers and spatial node density. The performance of the proposed model in terms of observed event distortion have been compared with what one gets with the correlation models found in the existing literature.

Motivated by the benefits of the framework and methods, energy-efficient communication protocols (a MAC protocol called Event-MAC and a routing protocol called PS-NLEACH) are also developed for event-driven sensing systems. They exploit the spatial correlation to improve the system's life-span. The spatial correlation-based, localized event-oriented MAC (Event-MAC) is developed first. It aims to reduce energy consumption by scheduling event reports based on correlation and priorities, without compromising the achieved distortion constraint and event latency. Next, a correlation-

based hierarchical routing protocol (PS-NLEACH) is developed that optimizes the network structure according to correlation between nodes without degrading the data resolution or fidelity. The proposed protocols are simulated and compared with the existing protocols to examine their performance.

In this thesis, the research has been presented in six Chapters. Chapter 1 provides an introduction to sensor networks, research challenges and a brief overview of the existing solutions. A review of existing literature is presented in Section 2. In Chapter 1, the observations on correlations between sensor nodes, the advantages of considering spatial correlation and the motivation behind the work presented in the thesis, have also been highlighted.

The concept of spatial correlation using sensors with omni-directional coverage and communication capabilities, and its advantages for energy-efficiency in WSNs has been given in Chapter 3. The relationships between the spatial correlation and the number of reporting nodes in the network for a given correlation threshold and distortion constraint have been derived at the network and MAC layers in this chapter. It is found that the densely deployed WSN can be partitioned into non-overlapping correlated regions at the network layer. Thus, only a single node needs to be allowed to report data rather than every node doing so, from a given correlated region, resulting in significant energy savings during data collection. Furthermore, the results show that required data reliability can be achieved through proper tuning of both the sensing range (ϑ) and the correlation threshold (ξ).

With the signal processing perceptive, for the estimation and reconstruction of the event source at the sink, the optimum/minimum node density within an event area can be calculated to minimize the observed event distortion. This happens when the reporting nodes are chosen such that (i) they should be as close as possible to the event source location, and (ii) as far as possible from each other. At MAC layer, the performance of

the proposed model in terms of achieved distortion at the sink has been compared with the existing methods. The comparative study shows that the proposed correlation model outperforms the existing correlation models. It is observed that the minimum number of nodes can be achieved by tuning of sensing range (ϑ) for a given distortion constraint (D_{max}) according to the requirement of the application.

In Chapter 4, the developed spatial correlation-based, localized event-oriented MAC (Event-MAC) has been presented based on the framework presented in Chapter 3 for event-driven wireless sensor networks. The prioritized selection of reporting nodes according to the spatial correlation, the received energy level from the event source, residual energy, type of data packet, queue length and packet delay, has also been discussed. The MAC performance could be improved by prioritized scheduling of the nodes in the influencing region when an event is detected by the nodes. A dynamic slot allocation technique is activated in the influencing region that assigns time-slots to the reporting nodes according to the correlation and priorities. Event-MAC guarantees collision-free channel access and eliminates redundant transmissions.

Based on the framework developed in Chapter 3, a correlation based hierarchical routing protocol, called PS-NLEACH has been developed and presented in Chapter 5. As the node selection phase is inserted at the start of N-LEACH in this algorithm, it is named as Pre-Selected N-LEACH (PS-NLEACH). PS-NLEACH divides the entire sensor field into several correlated sub-regions and selects a subset of nodes as the representatives of the regions. These representative nodes later execute a dynamic clustering N-LEACH algorithm to collect data during each round. Simulation results show that PS-NLEACH reduces the energy consumption of nodes and improves the system's life-span under the given data resolution which is tuned by the correlation threshold parameter.

Finally, the major conclusions of the present work and the scope for future work have been given in Chapter 6.

Acknowledgement

It is my pleasure to thank many people who made this thesis possible. I feel immense pleasure in expressing my profound sense of gratitude to my thesis supervisor Dr. Yatindra Nath Singh. I whole heartedly thank Dr. Singh for his invaluable support and guidance during my study, for providing me freedom to work on the topic of my choice. He has provided me the tools, the power, and will to strive for success. My cordial thanks are also extended to my co-advisor, Dr. Nishchal K. Verma.

It takes a long time to write a PhD thesis and it took me more than the most. I would like to express my gratitude to the people who helped in every little way they could, ever since I enrolled as a student at IIT Kanpur. I express my heartiest thanks to Prof. P. K. Kalra and Dr. A. R. Harish. Among all who taught me, I can safely make the claim that they are the best. I am grateful to all the faculty members of the Department of Electrical Engineering at IIT Kanpur for their constant support and encouragement.

I would also like to thank my colleagues for kindly sharing their experience about their studies and work. The generous supports from IIT Kanpur and Department of Electrical Engineering is greatly appreciated.

The last but not the least, I would like to thank my family for the unselfish support during the many years of my doctoral study. The dissertation would not have happened without their understanding and sacrifice.

Rajeev K Shakya

Dedicated

to

my family, Teachers

and

my beloved Parents

Contents

List of Figures

List of Tables

List of Symbols

$\beta,\ \gamma$	Signal attenuation parameters, default values are $\beta = 1$ and $\gamma = 2$
$\hat{S}_i$	Estimated value of S_i
$\hat{S}_{avg}$	Average of all estimated values $\hat{S}_i$'s from all N nodes.
λ	Spatial node density as the number of nodes per unit area $(1/d^2)$
$\mathbb{C}ov(.)$	Mathematical Covariance
$\mathbb{E}(.)$	Mathematical expectation
$\mathbb{V}ar(.)$	Mathematical variance
$\rho_{(i,j)}$	Correlation coefficient between nodes n_i and n_j located at coordinates s_i and s_j
σ_N	Noise variance of measured information
σ_S^2	Signal variance of measured information
$\mathbf{C_{ss}}$	Correlation matrix computed by a node with its neighbors
ϑ	Correlation parameter which is twice of r sensing range say $\theta = 2r$
$d_{(i,j)}$	Euclidean between nodes n_i and n_j located at s_i and s_j
$d_{(i,s)}$	Euclidean between node n_i and event source S.
D_E	The event MSE distortion measure
D_{max}	Maximum tolerable event distortion allowed by a sensor network
$INTC(d)$	Intersection area of two circles centered at two points located at d distance.
M	A total number of selected reporting nodes $(M < N)$
m_i	Mean value of measured signal information of node n_i

N	A total number of potential reporting nodes
N_i	Noise component in signal observed by node n_i
P_s	Sending power or transmission power required by a node
$p_{max},\ p_{min}$	Maximum and minimum received energy level of a event source
R_c	Effective event radius as $R_c = R_E/\sqrt{2}$
R_{corr}	Correlation radius
R_E	Event radius of circular influencing region
S	Event source or energy-radiating signal
S_0	Initial signal strength of event source S
S_i	Observed signal information of node n_i
t_{cs}	A time-slot duration for a control packet
t_{ds}	A time-slot duration for a data packet
U_e	Energy consumption per node per unit time
X_i	Observed signal with noise component for node n_i

List of abbreviations

ACK	Acknowledgement
ARP	Address Resolution Protocol
BP	Battery Power
CCA	Carrier Channel Assessment
CH	Cluster Head
CS	Control Slot
CSMA	Carrier Sense Multiple Access
CTS	Clear to Send
DCF	Distributed Coordinate Function
DIFS	Distributed Inter-Frame Space
DS	Data Slot
EWSNs	Event Driven Wireless Sensor Networks
GPS	Global Positioning System
IEEE	Institute of Electrical and Electronics Engineers
IFq	Interface Queue
INS	Iterative Node Selection
ISO	International Standards Organization
JGRV	Jointly Gaussian Random Variable
LPL	Low Power Listening
MAC	Medium Access Control
MMSE	Minimum Mean Square Error

MSE	Mean Square Error
NPI	Net Priority Index
NS-2	Network Simulator 2
OTcl	Object-oriented Tool Command Language
PD	Packet Delay
PI	Priority Index
PT	Packet Type
QL	Queue Length
QoS	Quality of Service
RCS	Reserved Control Slot
RCET	Random Correlated Event Traffic
REG	Random Event Generation
RSS	Received Signal Strength
RTS	Request to Send
SIFS	Short Inter-Frame Space
SP	Source Packet
SYN	Synchronization
TDMA	Time Division Multiple Access
TSA	Traditional Slot Assignment
UDP	User Datagram Protocol
VQ	Vector Quantization
WSNs	Wireless Sensor Networks

Chapter 1

Introduction

The recent developments in low-cost, low-power sensor nodes which are capable of sensing, computing and transmitting sensory data over a large geographical area by co-operatively monitoring physical or environmental conditions, e.g., temperature, sound, vibration, pressure, etc., have enabled many sensor network applications. Applications include environmental and habitat monitoring such as forest fire detection, air pollution and rainfall observation in agriculture; they also extend to the military, education, monitoring of machine health, and medical diagnostics. Wireless Sensor Networks (WSNs) usually comprise of a large number of sensor nodes deployed randomly in a highly dynamic and hostile environment. As the sensory coverage (sensing range) and network connectivity (communication range) of sensor nodes are limited, in any random deployment, one has to deploy nodes with sufficiently high density. This helps in maintaining the desired level of coverage for sufficient sensory data resolution for sufficient operational lifetime [1].

WSNs are usually event-driven systems where several nodes try to transmit data when any physical phenomena of interest (PoIs) is detected. Thus the collective efforts of these sensor nodes play an important role in reliable detection of PoIs. The sensing process is usually omni-directional [2]. A node can detect an event within its limited sensing range.

Event detection depends on the sensed strength of the emitted event signal, properties of physical phenomena, the sensed events' features and the hardware of the node. An event may trigger a large number of nodes for sensing and transmission. For example, in surveillance applications, a target in a sensor field will be detected by several nodes in its proximity. These nodes become active and send reports about the target to the sink.

WSNs are densely deployed to assure fine-grain monitoring in various applications. It may be as high as 20 nodes$/m^2$ or more [1]. Since the nodes are densely deployed, transmissions of information from nodes are spatially correlated. The redundant correlated data is observed due to the common sensing area between the nodes. The degree of spatial correlation increases with the increase in common sensing area between nodes. Various solutions have been proposed in the literature to provide information-driven strategies by the observing data correlation, specially in the densely deployed WSNs. To avoid redundancy and achieve just the required resolution of data, it is desirable to limit the number of reporting nodes, so that the nodes send only the data containing useful information to the sink. However, there could be a trade-off between energy consumption and the quality of information. In this thesis, a set of protocols is proposed to improve the life-span of the network while achieving the desired quality of information.

A brief overview of WSNs is given in this chapter. After that, the existing solutions and the motivation behind the work in the present thesis are discussed. The problem formulation and assumptions are briefly summarized at the end. The performance metrics used in present thesis for analysis have also been discussed. Finally, the organization of the thesis is outlined.

1.1 Background

As stated earlier, a WSN consists of a large number of sensor nodes deployed randomly inside or close to the physical phenomenon. These sensor nodes cooperatively

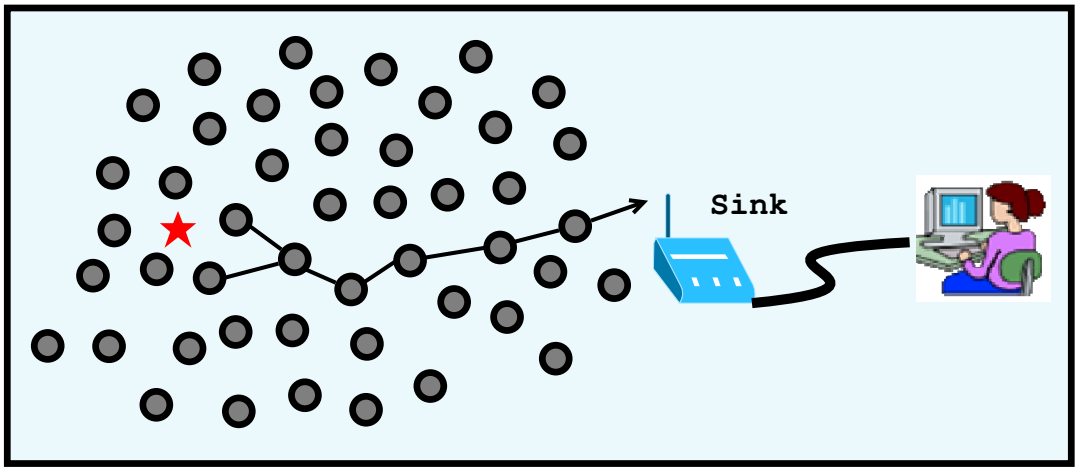


Figure 1.1: *Wireless Sensor Network Overview*

monitor physical and environmental conditions such as temperature, sound, vibration and humidity. The sensor readings collected by these sensor nodes are transmitted over a wireless channel to the distant node, known as the base station or the sink. The primary function of a sensor network is to sample sensory information from the designated area and transmit it to the sink. The sink mostly forwards all the data by a wire-line or an independent wireless network to the control center as shown in Fig. 1.1. The sensor nodes in the network operate as a collective structure which makes this type of network different from traditional ad-hoc networks. Due to dependency on sensors for generating data, the traffic is variable and correlated. Most of the time, the traffic changes little over long periods, and can be very intense for short periods of time. This correlated bursty traffic is characteristic of densely deployed sensor networks.

In general, the application domains of WSNs are classified into three main categories: periodic reporting, event-based reporting, and query-based reporting. In periodic reporting, sensor nodes collect data regularly and send it periodically to the sink. In event-based reporting, sensor nodes send the reports only when a special event of interest is detected. For example, in the case of an application to detect forest fires, when temperature above a threshold limit is detected by the sensor nodes in the region, these nodes become active

and send alarm messages to the sink. In query-based reporting, the sink node requests a subset of nodes to send their collected data. This thesis deals with the event-based reporting applications.

A high density of sensor nodes is required to provide full coverage of an area under observation [2, 3]. For dense WSNs, when the number of nodes is more than the minimum required number of nodes in the field, multiple nodes detect the same event simultaneously. As a result, the collisions and correlated transmissions of event reports become critical issues. In this thesis, the main focus is on densely deployed WSNs with the correlated data.

Each node has only limited battery power to operate, and this can deplete easily. The life of a network depends on the power available with the nodes. It would be impossible to replace/change the battery since sensor networks are usually deployed in unattended environments. In some cases, sensor nodes can harvest energy from surroundings such as solar power. However external power supply cannot be reliable solution due to possible non-deterministic behavior of external sources. Consequently, the major bottleneck in WSNs is battery life. Therefore, the primary challenge is to focus on a network's life-span by conserving the battery power, so that the network can work for months or years after deployment without any further power supply requirement [1, 4, 5]. It is important to note that communication among nodes consumes more power than the computation.

Effective deployment will decide how efficiently the communication is managed over a network that operates for years. Communication is essential (i) to provide collaboration between neighboring nodes to help in observing accurate PoIs and (ii) to report these PoIs to the sink. But, the communication is also error-prone, and consumes a large amount of energy and has a limited range requiring more nodes to relay the information over the network [1]. All these factors limit the life-span of the sensor network. In typical sensor motes like IRIS-It or MicaZ, the energy consumed to relay one bit of data is equivalent

to that consumed in the execution of several hundred instructions. Therefore, managing communication according to network traffic is the key challenge in the design of energy-efficient communication protocols for WSNs.

1.2 Existing Solutions and Motivation

As stated before, since a major resource bottleneck is energy rather than computing capability, several energy-efficient MAC protocols [6, 7] and energy-efficient routing protocols [8, 9] have been proposed in the literature. These protocols aim to turn off some nodes completely when they are not needed. As a result, a significant amount of energy saving is achieved by these schemes. MAC protocols [6, 7] suggest wake-up scheduling schemes to put the nodes in sleep mode whenever they are not needed. On the other hand, routing protocols [8, 9] take into account the sleep schedules of the nodes for their efficient routing. Although these protocols increase the life-span of the network, there can always be some redundant data gathered by multiple sensors in the network.

In densely deployed WSNs, a large volume of data may be generated and is available for transmission when an event is detected by the sensor nodes in their proximity. In such a scenario, all these nodes transmit a lot of redundant data. For example, in Fig. 1.2, nodes A, B, C and D are shown with their respective sensing ranges which are indicated by dashed circles. Here, the sensing area of node B is also fully covered by its neighbours A, C, D together. Therefore, event information observed by the sensor node B will also be sensed by its neighbours. B is found to be a redundant node in this case as, even if B is not there, the reliable event reporting will happen. The high density of redundant nodes in a WSN permits higher data resolution but at the expense of the energy used in transmitting extra bits to the sink. It is desirable to limit the number of reporting nodes to achieve just the desired resolution.

Based on the observations made above, most of the solutions [10, 3, 5] adopt coverage-

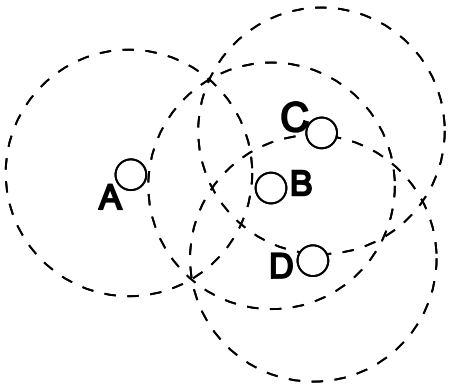


Figure 1.2: *Coverage area of nodes, indicating that node B is redundant.*

based scheduling schemes. In these schemes, redundant nodes are identified based on the sensing ranges of nodes. A coverage-based scheduling scheme first selects the nodes using pairwise geographic distance between them. The selected nodes then sample the sensor field uniformly over the space. The coverage-based schemes may not work well where sensor readings may change in some sub-regions due to a change in the local measurement. For example, in the case of temperature measurement, two nearby nodes can observe different readings if one of them is in sunshine while the other one is in shadow. Consequently, these methods rely only on sensing coverage which may not be an efficient or effective method. In addition, these methods do not consider the observed data for spatial correlation between nodes. Intuitively, correlating the data measured by the nodes may help to select the nodes more efficiently.

To capture the spatial correlation in event data and take advantage of it for energy efficiency, several theoretical methods and protocols have been proposed [11, 12, 13, 14, 15, 16, 17, 18]. Recently, the relation between the distortion, spatio-temporal correlation and the event estimation reliability for large sensor networks has been investigated [11, 19, 20]. In [19], the combined effect of spatial and temporal correlation for both point and field sources has been studied. The spatio-temporal characteristics are derived in terms

of distortion functions using models of point and field sources. Based on the aspects of correlation in WSNs, according to information theory, efficient medium access [14] and reliable event transport [20] protocols have been proposed to exploit the spatial and temporal correlations respectively. Another design criteria is to compress the sensed information at an optimum rate by using distributed source coding techniques [12, 13, 21, 22]. These include Slepian-Wolf coding and communication coding to reduce the amount of data. It has been proved that event information can be reconstructed with high confidence even from a reduced number of sources, if they are correlated.

The above research relies on the intuition that sensor nodes in spatially close proximity can observe the correlation in data. However, there is no suitable model which considers the real network conditions of WSNs. Since there are various choices in implementation and in the configuration of the networks, it is possible to design different models to fit different circumstances. A model must be chosen according to the properties of the physical phenomena and the features of the sensed events. The events' features vary significantly for energy-radiating physical phenomena originating in the field. A fitting model would also take into account the real conditions of the network such as sensing range, location of nodes and the distances between them.

1.3 Objectives

WSNs are usually event-driven systems where several nodes try to transmit data when any event of interest happens. Thus the collective efforts of these sensor nodes play an important role in reliable event detection [11]. Since the sensor nodes are densely deployed, the redundancy of sensor nodes results in highly correlated data transmissions from nodes due to overlapped sensing regions. The degree of spatial correlation increases with the increase in overlapped sensing areas between the nodes (i.e., decreasing inter-node distance). Therefore, every sensor node observing the event need not be ac-

tive for sensing and communication. A significant amount of energy-saving is possible by taking advantage of spatial correlation. The motivation is to design energy-efficient communication protocols for energy-constrained wireless sensor networks by exploiting spatial correlation.

Besides, for the dense WSNs, when the number of nodes is greater than the minimum required in the field, multiple nodes detect the same event simultaneously. As a result, correlated transmissions and collisions occur unnecessarily.

The fundamental aim of the work in the present thesis is to increase the life-span of the sensor networks. In this thesis, a model framework has been presented to discover the correlation of event information as observed by the reporting nodes. The model is applicable only to the omni-directional sensor nodes which are widely used in WSNs. The event source estimation and reconstruction of the distortion at the sink has been investigated. The impact of correlation between nodes on achieved distortion at the sink is investigated. The performance of the proposed model in terms of observed event distortion have been compared with what one gets with existing correlation models. Based on this theoretical framework, energy-efficient MAC and routing protocols have been proposed. The proposed MAC protocol, known as Event-MAC is a novel hybrid TDMA/CSMA MAC which takes into account spatial correlation and priorities on node scheduling for efficient event reporting. By activating Event-MAC, only when an event is detected, its dynamic time-slot allocation guarantees collision-free channel access with full channel utilization in the event area. This avoids the unnecessary transmissions and collision in densely deployed event-driven sensor networks. At the network layer, it is demonstrated that how, through proper tuning of both the sensing range and the correlation threshold, WSNs can be partitioned into the disjoint correlated clusters without degradation of the information reliability, thus enabling significant energy-saving during data collection. Based on the results obtained from the model framework, a hierarchical routing protocol

named PS-NLEACH has been proposed, which takes advantage of spatial correlation for energy efficiency during data collection in WSNs.

The two fundamental measures of quality of service (QoS) in WSNs are (i) sensing coverage that provides the maximum detection rate of events and (ii) network connectivity that provides fast delivery of a detected event to the sink. By considering both the measures, this research captures the spatial correlation of an underlying sensor field and use it for efficient communication protocols in WSNs. As an extension work of [14], this thesis deals with techniques of node selection which exploit the spatial correlation for reliable and efficient communication. The following objectives have been achieved in this thesis.

- Modelling of the correlation characteristics based on the Omni-directional sensing and to theoretically analyze the impact of the spatial correlation between the nodes at both the MAC and the network layers.

- Development of node selection techniques using a new alternative correlation model, at both MAC and network layers and study the impact of the correlation between nodes using various parameters such as distortion constraints, spatial node density, node selection number and sensing range.

- Development of a collision-free, localized, event-oriented MAC protocol based on the framework that exploits the spatial correlation in a distributive fashion for event-driven sensor network applications.

- Development of an energy-efficient clustering technique based on the correlation threshold without compromising the information reliability during data collection for WSNs.

1.4 Network model assumptions

The following assumptions are made to develop framework and protocols in this thesis.

- *Homogeneous sensor network:* The sensor network applications are assumed to consider the sensor nodes having equal initial energy and similar capabilities (communication, processing etc.).

- *Sensor nodes without mobility:* The sensor nodes are assumed to be static in the networks. Their positions are fixed throughout their lifetime.

- *Single Channel sensor node:* The proposed protocols are designed to work for simplex radio architecture only.

- *Sensing and communication rage:* A sensor node is characterized by sensing coverage (sensing range) and network connectivity (communication range). All sensor nodes follow the omni-directional sensing mode, in which each node can reconstruct event at every point in a r-radius disk area centered at its location. A sensor node can detect all events within such a disk whereas no event outside the disk can be detected [2, 23].

- *Densely deployed sensor network:* To promise the fine-grain monitoring, sensor nodes are assumed to be densely deployed (e.g., 20 $nodes/m^3$) [4].

- *Event-driven Systems:* WSNs are considered as event-driven systems, where several nodes get activated for reporting data when an event of interest happens in the sensor field. It is considered that the collective efforts of sensor nodes play an important role in reliable event reporting.

1.5 Performance Metrics

In this thesis, the simulation results for performance analysis are obtained using simulation tools discussed in Appendix A. The followings are the performance metrics used for analysis.

- *Average energy consumption:* It is the average consumed energy per node per unit time. It is an important metric to measure the life of network since WSN has limited energy after deployment.

- *Success rate:* It is also important to evaluate the efficiency of a MAC protocol as higher success rate corresponds to less collisions. It is defined as the ratio of the number of received reports by the sink to the number of transmitted reports by the nodes in the event area.

- *Distortion:* It is the average distortion measured at the sink using the distortion function derived in the present work. The distortion is calculated over all the received reports at the sink.

- *Coverage ratio:* It is used to evaluate the performance of the node selection schemes. It is defined as the ratio of total area of field covered by the correlation regions of representative nodes to the area covered by all potential reporting nodes (i.e., event area). Higher coverage ratio denotes better event reporting as larger fraction of event area is covered by the selected reporting nodes.

- *Event latency:* It is the average transmission latency i.e. the time elapsed since the transmission of event reports by nodes to the instant of successful reception by the sink.

- *Detection latency:* It is defined as the time duration between the occurrence of an event and the time instant when the sink gets the first notification of it (as an alarm

message [24]). Unlike event latency, it includes the delay in primary decision and transmission delay in the routes of reports using routing procedure.

1.6 Thesis Organization

Chapter 2 presents the literature survey. In Chapter 3, a basic correlation model has been introduced to represent the correlation characteristics between sensors with omni-directional coverage capabilities. Inspired by the model framework, we extend the work of [14] and [15] using our sensing model and a new correlation function (i.e., distrotion fucntion to determine the relationships between reporting nodes and information reliability). Followings are the new contributions provited by this thesis at the both the MAC and the Network layers:

- At the network layer, the densely deployed WSN can be partitioned into non-overlapping correlated clusters of nodes by the use of proposed GCC algorithm.
- One can achieve required event reliability through proper tuning of sensing range (ϑ) and the correlation threshold (ξ) for reliable data delivery assuming a fixed node distribution for a given period of time.
- For the estimation of the event source, the minimum node density within the event area can be found to achieve the lower observed event distortion. In addition, a minimum reporting nodes can be determined through proper tuning of ϑ for estimated maximum event distortion limit (D_{max}).

The performance of the proposed model in terms of achieved distortion at the sink has been compared with the existing methods in [14] and [15]. The impact of various parameters like spatial node density, the number of selected nodes, and the nodes' sensing range has been investigated. The comparative study shows that the proposed correlation model outperforms the existing correlation models.

In Chapter 4, the idea of collision-free channel access in the influencing region only has been introduced. The proposed Event-MAC accomplishes the following:

- Event-MAC mitigates the effects of collisions and over-hearing by using local TDMA scheduling in the influencing region of an event. In addition, correlation based decision on node selection and prioritized scheduling provides QoS for the event reporting for Event-driven wireless sensor networks. Thus, based on correlation and priority, the minimum distortion using lesser number of reporting nodes is achieved by collection of reports only from nodes closer to the event.

- Upon occurrence of an event, a node will switch over to TDMA scheduling mode by running Event-MAC. A dynamic slot allocation technique is activated in the influencing region that assigns time-slots to the reporting nodes according to correlation and priorities. So all reporting nodes within event area operate scheduled channel access with priorities and the nodes outside the event area perform contention based channel access (CSMA/CA). Thus, Event-MAC ensures collision-free channel access and fast delivery of reports from the event area.

- Event-MAC considers priorities according to the events' observations (i.e., the received energy level from the event source) and other parameters such as packet type, queue length, packet delay and residual energy. Thus, Event-MAC removes redundancy in information and provides minimum energy consumption, high packet delivery ratio, and bounded delay for EWSNs.

Based on proposed model framework, Pre-Selected NLEACH algorithm has been also proposed in Chapter 5. This algorithm is modified version of popular routing protocol LEACH [25] in WSNs. PS-NLEACH divides the entire sensor field into several correlated sub-regions and selects a subset of nodes as the representatives of the regions. These representative nodes later execute a dynamic clustering N-LEACH algorithm to

collect data during each round. For correlation-based efficient data collection in sensor network applications, the Pre-Selected NLEACH algorithm achieves the following:

- The proposed routing protocol exploits the concept of spatial correlation during data collection by selecting the nodes using proposed . correlation model.
- In each round, the representative nodes do the data collection and data processing throughout the network while non-representative nodes switch over to the sleep mode. Further, every node gets equal chance of being a representative node in the network periodically.
- There is a trade-off between energy consumption and the information reliability, where required information reliability is the function of the number of nodes representing the network.

Simulation results show that PS-NLEACH reduces the energy consumption of nodes and improves the system's life-span under the given data resolution tuned by the correlation threshold parameter.

Finally, the major conclusions of the present work and the scope for future work have been given in Chapter 6.

Chapter 2

Literature Review

2.1 Correlation in Sensor Networks

Correlation in WSNs and its exploitation has been studied very widely [1], [11], [12], [13], [14], [15], [26], [27]. Also, investigations to have efficient communication protocols on the basis of correlation have been done [13, 15, 28]. Several methods have been proposed to allow a smaller number of sensing nodes to transmit their reports from the event area, so as to reduce the power expenditure and to avoid unnecessary contention among correlated nodes [14, 15, 16].

Correlated data collection in WSN has been considered as spatio-temporal correlated processes [29, 30, 31, 32, 33, 34, 35]. The expressions for the distortion function have been derived to capture the correlation in data for different network topologies. The distrotion analysis in one-dimensional and two-dimensional grid scenarios have been done in [29] while authors in [30] proposed for two-dimensional network topology as a wheel. The Pearson Correlation Coefficient was used as a kind of spatial correlation model to measure the correlation of sensor observations [31]. Only the strength and the direction of the linear relationship between two variables are measured. Thus, it only describes the linear correlation depedency in data and requires more data to be sent to

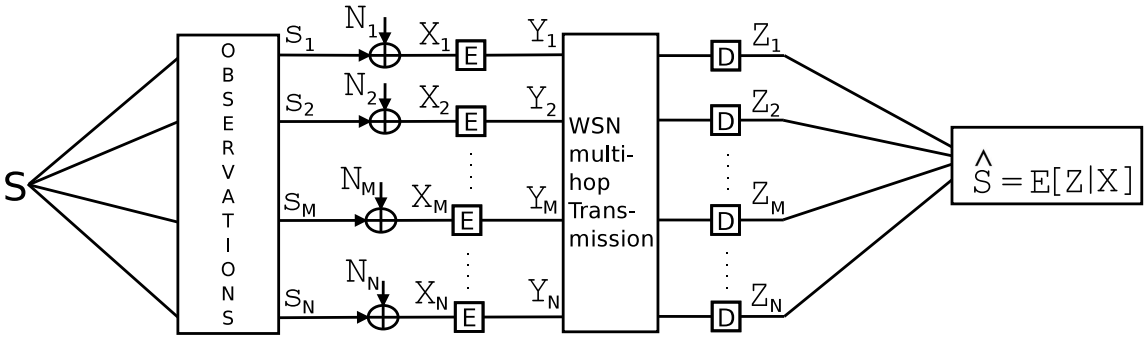


Figure 2.1: *WSN Model for Event Source S estimation [14, 15].*

the sink. However, it was not considered for the design of communication protocols and algorithms. In order to design an energy-efficient MAC protocols, researchers [14, 15] have used correlation models [36] to determine the relationship between the spread of sensor nodes and event estimation reliability. In some other correlation models [31, 11, 33, 35], statistical features were introduced based on sensors' observations. Authors in [11, 14] modelled the event source S as a random process $s(t, x, y)$ at a time t with spatial coordinate (x, y). It is assumed that a node n_i observes the event S as S_i and the sensed data as X_i. The node n_i encodes X_i and transmits it to the sink through multi-hop transmission. The sink node decodes all the received X_i's to estimate the S_i's. The encoder and decoder are denoted by E and D respectively as shown in Fig. 2.1. The variance of event information X_i is assumed to be same as variance of event source S. After collecting the event information X_i from node n_i, the sink estimates values of event information using the estimator. Using this estimation model, a distortion function was derived to investigate the impact of spatial correlation between nodes and event estimation reliability.

For event estimation, the conventional correlation models have been used to determine the correlation coefficients. Four types of spatial correlation models were given. These are *Spherical, Power exponential, Rational quadratic,* and *Matérn* [36]. However,

the sink node needs to collect much more data from the sensors for simulating the correlation in data using these models and thus the computation complexity is also higher. Furthermore, they do not consider the real network conditions such as sensing range, location of nodes, and distance between the nodes. Based on the correlation in data, the spatio-temporal characteristics are derived in terms of distortion functions using models of point and field sources [19]. Based on this analysis, efficient medium access [14] and reliable event transport [20] protocols have been proposed to exploit the spatial and temporal correlations respectively.

Vuran and Akyildiz [11] were the first to determine numerically that based on spatial correlation among nodes, only a the minimum number of representative nodes need to be activated to achieve the required information reliability at the sink. Accordingly, enhancements of the IEEE 802.11 DCF MAC protocol [14, 15] were proposed to activate only the representative nodes which need to transmit their reports. The non-representative nodes switch to sleep mode to save energy. Thus correlation in data is filtered out by allowing reports from representative nodes only. CC-MAC [14] supports the random representative node selection while the SC-MAC [15] proposed ordered node selection. However, Both the protocols had drawbacks. These protocols are contention-based and have over-hearing problems. Correlation in data may still exist due to representative node selection based on over-hearing. MAC protocols using correlation for estimation open up new directions to build access mechanisms. However, most of these protocols need the detailed event information for the estimation and exploitation of correlation in event reports. In event-driven application, more emphesize should be to identify whether the event happens or not and where it happened.

An event is detectable if at least one node lies within its sensing range. Any event outside of the sensing range of the nodes can not be detected. Sensor scheduling methods based on sensing coverage have also been proposed [3, 37] to remove redundancy

in sensor observations. The coverage and connectivity problems have been considered jointly to avoid the redundant nodes [5, 38, 39]. In [39], the redundancy based on sensor coverage is checked by the nodes to achieve differentiated coverage with respect to their neighbors. After this checking rule, a distributed and localised scheduling scheme is used to disseminate the scheduling information throughout the network. Thus, the nodes decide cooperatively to switch over the sleep mode for a certain period of time. Most of the work considers only the Boolean sensing model. YR Tsai [40] had studied sensing coverage in shadow-fading environment. Boolean sensing model and shadow-fading sensing model were considered in the analysis of network coverage for the randomly deployed nodes. Another sensing model given in the literature is known as Elfes sensing model [41]. In [42], Efles sensing model has been used to investigate the network coverage. Recently, in [43] a comperative study of three types of sensing models viz. Boolean sensing model, shadow-fading sensing and Elfes sensing model have been investigated for coverage analysis in WSN.

In this thesis, our concern is to find out whether the event happened or not and also where is it happened for the event-driven monitoring applications. To estimate the event source, the reports containing the event indication are of interest without any other details of the occurring event. The sensing range of the nodes has been used for accurate estimation and reconstruction of the event source. The location of nodes in event-driven sensor network is assumed to be known because accurate reporting of event within short period of time is a necessary requirement for most of the critical events (e.g., enemy detection/tracking in battlefield).

2.2 MAC Protocols in WSN

Medium Access Control (MAC) coordinates the channel access among nodes in fair and efficient manner. A MAC protocol plays important role in energy-efficiency, packet

delivery, delay and QoS (Quality of Service). For WSNs, the research on MAC protocol design has been focused mainly on energy-latency trade-offs [44]. S-MAC [45] is designed to save the energy by using listen and sleep periodically with collision avoidance facilities of IEEE 802.11 MAC standard [46]. S-MAC uses synchronization mechanism to form virtual clusters of sleep/wakeup schedule to avoid overhearing problem. Many extension protocols of S-MAC have been proposed to further decrease the energy consumption. These are D-MAC [47], T-MAC [48], DS-MAC [49], P-MAC [50], DW-MAC [51] etc. These all variants deal with major source of energy wastage such as idle-listening, overhearing and collisions problems. B-MAC [52] is low power listening mechanism , best suitable for low data rate wireless sensor networks. S-MAC and T-MAC were based on the concept of hybrids of CSMA and TDMA, but TDMA slots are much smaller than the normal TDMA slots. The synchronization failures in these protocols did not lead to communication failure because of RTS/CTS mechanism. Seamlessly adjusting the behavior of MAC between TDMA and CSMA according to the contention level was explored in [53, 54]. They used loose synchronization and topology information to work under high contention. Under low contention, these are not particularly well suited and reliable. Similar to S-MAC, and T-MAC, B-MAC also uses periodic ON and OFF of radio transceiver. However, it is un-synchronized duty cycle protocol. In order to transmit the packets, nodes sent long preamble before actual data transmission. These MAC protocols have been designed for general sensor network applications where latency is not considered as critical parameter. However, none of these have been designed to filtering out correlated information in observed data.

For Event-driven wireless sensor network (EWSNs), very little work has been proposed in the literature. In event-based reporting systems, sensor nodes transmit reports only when they detect the events. The events are unpredictable and appear spontaneous in sensor field. When an event of interest happens in particular zone, spatially correlated

information is observed and transmitted by surrounding nodes of the event. For Event based WSNs, EB-MAC [55] schedules the reports from the influencing region by initiating a round to transmit its own reports when an event is detected. Initially, all triggered nodes elect a leader node amongst themseleves, then the time slot is assigned to a leader node for data transmissions. EB-MAC exhibits a big overhead in maintaining the synchronization of round timers of different nodes with first leader node. Since the nodes belonging to different groups may have different schedules at the same time, EM-MAC may not work well in multihop scenario. To handle the dynamic forwarding in multihop scenario, ECR-MAC [56] works well for event-driven applications. It also handles spatially-correlated contention when an event is detected by multiple nodes. In ECR-MAC, multiple potential forwarders are maintained by each node to transmit the reports as soon as possible and the contention is minimized by dispersing the paths the sender will take.

SIFT MAC [57] were designed for event-driven WSN. The objective of SIFT MAC protocol is to minimize the latency when spatially-correlated contention occurs in a monitored area. Jamieson et. al [57] had argued that only R nodes out of N nodes that report to a common event are sufficient to be successful to transmit the event information to the sink node. SIFT MAC uses non-uniform geometric distribution to choose slot number for picking up a slot for transmission within fixed-size contention window (32 slots). QS-Sift [58] and HS-Sift [59] divided the event area into sub-regions based on the concept of different SNR observation records. They used a simple estimator to determine the SNR distribution with respect to measured distortion. Accordingly, different channel access schemes were used with different priorities, which were based on low and high SNR observations. QS-Sift protocol uses contention-based random channel access scheme while HS-sift were designed by combining TDMA and CSMA schemes.

The CC-MAC [14] was proposed by Vuran and Akyldiz to investigate more on event

distortion for EWSNs. In CC-MAC, Iterative Node Selection (INS) algorithm was proposed to calculate correlation radius (r_{corr}) based on correlation model. Only one node is allowed to transmit event information within a correlation radius. In this way, CC-MAC suppresses the transmissions by other nodes within same correlation radius. This single node is referred as representative node selected by CSMA mechanism during each transmission within r_{corr} radius. In first contention phase, all nodes within r_{corr} radius contend to access the channel like in any other contention based protocol. As a result, only node winning the contention is selected as representative node while other nodes turn to sleep. In CC-MAC, there is no control on selection of representative node to further saving the energy. It is unpredictable that which node will win the contention.

Following similar philosophy of CC-MAC protocol, two efficient node selection MAC protocols, [60] and [15] have been proposed. In these protocols, use of RTS/CTS mechanism increases extra overhead and is not suitable for small sized packet in WSNs. Overheads reduces the protocol efficiency. These are also not adaptive for traffic variations when no activity is observed by the nodes due to absence of the events. In such case, a mechanism of periodic sleep-wake cycle is required to save unnecessary energy consumption. One can assign the time-slots only to the nodes, which are within the influencing region of occurring the event. The remaining nodes in the network can have simple low power channel access scheme to route the event reports to the sink.

2.3 Routing Protocols in WSN

Several cluster-based routing protocols have already been proposed for cluster-based WSNs. In WSNs, routing protocols can be classified according to their mode of functioning and network structure. Hierarchical routing is one of the classifications based on network structure. For example, LEACH [25], TEEN [61], and APTEEN [62] are based on hierarchical network structures used for energy-efficient routing in WSNs. In cluster-

based routing protocols, the idea is to group the nodes into a set of disjoint clusters. In each cluster, a cluster head (CH) is selected randomly to receive data from other cluster members. The CH then aggregates the received data and transmits it to the sink directly. This mechanism is a widely accepted solution for organizing sensor nodes into clusters for energy conservation. CHs can suppress the collected observations from their members that are most likely to be highly correlated in a densely deployed network. In addition, this mechanism provides better control over the activities of cluster members, and over scalability issues (e.g. communication and routing costs). In these protocols, higher-energy nodes act as CHs to process the received data and send it to the sink after processing (i.e. aggregation). The cluster members are low-energy nodes that perform the sensing of physical phenomena. Other types of classifications include centralized clustering ([25, 63, 64, 65, 66, 67]), distributed clustering ([25, 68, 61, 62, 69, 70, 71]), probabilistic clustering ([72, 73, 74, 75, 76]), and non-probabilistic clustering ([77, 78, 79, 80, 81]).

To model spatial correlation and take advantage of it for energy-efficient routing, various methods have been proposed such as in [11, 12, 13, 26, 27]. These protocols consider the fact that spatial distance between nodes leads to spatial correlation in the data observed by them. Li and AlRegib [82] investigated the determination of optimal cluster size for energy-efficient cluster-based distributed estimation in WSNs. Using BLUE estimator, CHs make a local estimation of collected data from their respective member nodes. The member nodes send only quantized versions of measured data. At the end, the sink collects all estimated data from CHs directly and makes final estimation using another estimator called quasi-BLUE. The saving in energy has been claimed using the minimum number of clusters according to distributed estimation of cluster size. The joint effects of data correlation, distance, and network density have been studied in [83]. In this correlated data gathering protocol, the optimal cluster size has been derived for linear network by considering heterogeneous-sized clusters in the network. However,

protocol were not investigated for mutlihop scenario and event based systems.

Chapter 3

Spatial Correlation Modelling for Sensor Network Applications

In WSNs having a high density of sensor nodes, transmitted measurements are spatially correlated, often redundantly, whenever an event of interest is detected. In this chapter, a correlation model is proposed. It enables energy-efficient methodologies to exploit the spatial correlation at the Network and MAC layers. Two model frameworks are developed to investigate the impact of spatial correlation between nodes using various parameters such as correlation threshold, sensing range, selected number of nodes, spatial node density and distortion constraint. Section 3.1 gives the motivation about this work. A novel spatial correlation function is introduced and discussed in Section 3.2. Few of the applications which use the proposed correlation model, with simulation results, are discussed in Sections 3.3 and 3.4. In addition, Section 3.5 gives a comparative study with existing correlation models. Lastly, Section 3.6 presents the chapter summary.

3.1 Motivation

In this chapter, a correlation model using omni-directional sensor nodes have been presented. Its advantage for energy-efficient communication protocols have been dis-

cussed in details. It should be noted that correlation between sensory data of two nodes is related to spatial correlation between them [1], [11], [13], [14], [15], [34], [28]. The spatial correlation can be estimated based on sensory coverage of nodes [84, 85]. The base station (i.e. the sink) needs to know the locations of nodes and sensing range r, to estimate the correlation. Thus, base station can estimate a more accurate sensed parameter by considering the estimated spatial correlation as the correlation between the received sensory data while combining the received data.

We consider a dense WSN where a large number of sensor nodes are scattered in an area where events are to be observed. All nodes have been assumed to have equal initial energy and similar capabilities (communication, processing and sensing of events). All the sensor nodes are capable of omni-directional sensing, in which each node can sense the event at every point in a r-radius disk area with node at its center. A sensor node can detect all the events within such a disk, whereas no event outside the disk can be detected [2, 23]. Each node has variable sensing range for detecting the events and fixed transmission range for communication. The communication range is usually much larger than the sensing range. The sink node is assumed to be interested only in a collective report about the event from all the nodes and not in the individual reports from each node.

In general, a high density of sensor nodes is required to provide full coverage of the area under observation. When the number of nodes is more than the minimum required number of nodes in the field, multiple nodes detect the same event simultaneously, resulting in the generation of the redundant reports about the detected event. As discussed previously, some nodes can be redundant for reporting the event in the sensor field. The high density of redundant nodes in a WSN permits higher data resolution but at the expense of energy used in transmitting extra data to the sink. It is desirable to limit the number of reporting nodes to just enough to achieve the desired resolution. Motivated by

this, we formulate the problem statement as follows.

For a certain area of interest, suppose there is a set of N omni-directional sensor nodes that can detect an event within sensing range r. We denote them as $\mathcal{N} = \{n_1, n_2, n_3, ...\}$ with the spatial coordinates being denoted as $\{s_1, s_2, s_3, ...\}$. There exists correlation among these nodes which can be estimated based on their spatial coordinates. This spatial correlation is to be exploited for the design of more energy-efficient communication protocols in WSN by avoiding transmissions from redundant nodes.

Correlation in WSNs has been discussed in [11] and its exploitation has been studied in [1], [12], [13], [14], [15], [26], and [27]. Sensor scheduling methods based on sensing coverage have also been proposed in the literature. Lot of research efforts have been devoted to coverage-based scheduling methods to reduce redundancy, based on the sensing range of nodes [3, 37]. Recently, coverage and connectivity problems have been considered jointly in [5, 38]. Also, investigations to build efficient communication protocols on the basis of correlation have been done [13, 15, 28]. Several methods have been proposed to allow a smaller number of sensing nodes to transmit their reports from the event area, so as to reduce the power expenditure and to avoid unnecessary contention among correlated nodes [14, 15, 16].

One can also have a scenario when sensor nodes have variable sensing range. We show that energy can be saved in the data transmission process when nodes change their sensing range according to their positions and neighbourhood. When omni-directional sensors are used, sensing range of sensor nodes can be optimized based on the common overlapping coverage area with its neighbours. This avoids redundancy by reducing the correlation in reported data. The allowed maximum correlation in measured information observed by the nodes while detecting an event, is decided by the application in terms of reliability/fidelity. Thus, the required event reliability at the sink can be maintained by dynamic optimization of sensing range, i.e., sensing range of nodes can be decreased or

increased depending on the high or low density of nodes respectively, in its vicinity.

In this chapter, we investigate energy-efficient methods that exploit the spatial correlation at the Network and MAC layers. The interactions among various parameters such as distortion constraints, spatial node density, node selection, sensing range and their impacts on the reconstruction performance are quantitatively studied.

3.2 Correlation Model

Assuming that all the received sensory data from reporting nodes is jointly Gaussian, the covariance between the two measured values from nodes n_i and n_j at location s_i and s_j respectively can be expressed by

$$\mathbb{C}ov\{s_i, s_j\} = \sigma_S^2 K_\vartheta(\|s_i - s_j\|). \tag{3.1}$$

Here σ_S^2 is variance of sample observation from sensor nodes and $\mathbb{C}ov(.)$ represents mathematical covariance. We assume σ_S^2 to be same for all the reporting nodes. $\|.\|$ denotes the Euclidean distance between nodes n_i and n_j. $K_\vartheta(.)$ denotes the correlation function with $\vartheta = (\theta_1, \theta_2, ..., \theta_c)$ as the set of control parameters which will be discussed later in this section.

3.2.1 Mathematical Model

Symbols and notations used in Fig. 3.1 are given in Table 3.1. If $d_{(i,j)} < 2r$, then $\mathcal{S}_i$, $\mathcal{S}_j$ will have an overlap. We can define the correlation as

$$\rho_{(i,j)} = K_\vartheta(d) = \frac{A_i^j + A_j^i}{A}. \tag{3.2}$$

Table 3.1: *Notations used in Fig. 3.1*

Symbol	*Description*
$\mathcal{S}_i$	Sensing region of node n_i of disk with radius r centered at the node n_i
$\mathcal{S}_j$	Sensing region of node n_j of disk with radius r centered at the node n_j
A	Area of the sensing region of a node
$d_{(i,j)}$	distance between nodes n_i and n_j located at s_i and s_j
(P_{ij}^1, P_{ij}^2)	Intersection points of two disk of radius r centered at the nodes n_i and n_j
L_{ij}	Length of common chord length. It is equal to length of the line segment joining two intersection points P_{ij}^1 and P_{ij}^2
A_i^j	Area of region surrounded by arc denoted by $\overset{\frown}{P_{ij}^1 P_{ij}^2}$ for S_i and chord denoted by $\overline{P_{ij}^1 P_{ij}^2}$ as shown by shaded area
A_j^i	Area of region surrounded by arc denoted by $\overset{\frown}{P_{ij}^1 P_{ij}^2}$ for S_j and chord denoted by $\overline{P_{ij}^1 P_{ij}^2}$
$K_\vartheta(\|.\|)$	correlation function computed by a node with its neighbourseparated by Euclidean distance d

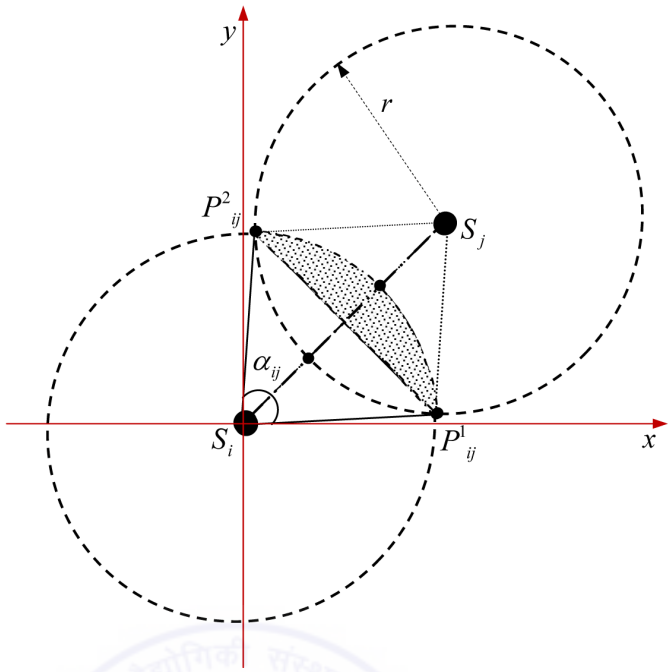


Figure 3.1: *The spatial correlation model*

Here $K_\vartheta(d)$ is the decreasing function with distance $d_{(i,j)}$, following the limiting value of 1 at $d = 0$ and of 0 for $d \geq 2r$. Areas A_i^j and A_j^i are same due to the symmetry as shown in Fig. 3.1.

$$A_i^j = A_j^i = r^2 \cos^{-1}\left(\frac{d_{(i,j)}}{2r}\right) - \frac{L_{ij} d_{(i,j)}}{4}. \tag{3.3}$$

Here, L_{ij} is length of chord formed by intersection of two circular disks and is given by $L_{ij} = 2.\sqrt{(r^2 - \frac{d^2(i,j)}{4})}$.

From (3.2), we get

$$\rho_{(i,j)} = \frac{2\cos^{-1}\left(\frac{d_{(i,j)}}{2r}\right)}{\pi} - \frac{d_{(i,j)}}{\pi r^2}.\sqrt{(r^2 - \frac{d^2_{(i,j)}}{4})}. \tag{3.4}$$

Let ϑ be a control parameter equal to $2r$. Then, (3.4) can be simplified as

$$\rho_{(i,j)} = \frac{2\cos^{-1}(\frac{d_{(i,j)}}{\vartheta})}{\pi} - \frac{2d_{(i,j)}}{\pi\vartheta^2}.\sqrt{(\vartheta^2 - d^2_{(i,j)})}. \tag{3.5}$$

We see that when $d_{(i,j)} = 2r$, the correlation model gives zero value. It means that there is no correlation between sensor nodes. For this reason, we introduce a control parameter ϑ equal to $2r$, as a variable to control the degree of correlation between nodes. The correlation model can be rewritten in a general form as

$$\rho_{(i,j)} = K_\vartheta(d) = \begin{cases} \frac{2\cos^{-1}(\frac{d_{(i,j)}}{\vartheta})}{\pi} - \frac{2d_{(i,j)}}{\pi\vartheta^2}.\sqrt{(\vartheta^2 - d^2_{(i,j)})}, & \text{if } 0 \leq d_{(i,j)} < \vartheta. \\ 0, & \text{if } d_{(i,j)} \geq \vartheta. \end{cases} \tag{3.6}$$

It can be seen from Equation (3.6) that when correlation function $K_\vartheta(d)$ is 0, it means that there is no correlation between the sensor nodes n_i and n_j, located at a distance $d_{(i,j)}$ from each other. If the correlation function $K_\vartheta(.)$ is equal to 1, the sensor nodes are perfectly correlated. The control parameter, ϑ is twice the sensing range of nodes. For the set of sensing ranges being $r = (2,\ 4.5,\ 6,\ 7.5,\ 9,\ 10)$, the set of control parameters ϑ, are $(\theta_1 = 4, \theta_2 = 9, \theta_3 = 12, \theta_4 = 15, \theta_5 = 18, \theta_6 = 20)$ respectively for analysis in next section.

3.2.2 Correlation Function, $K_\vartheta(.)$ - examples

In this subsection, in order to study connectivity with variation in sensing range, we simulate 200 randomly distributed nodes in a $150 \times 150\ m^2$ area as shown in Figs. 3.2(a) and 3.2(b), 150 nodes in a $150 \times 150\ m^2$ area as shown in Fig. 3.2(c), and 30 nodes in a $50 \times 50\ m^2$ area as shown in Fig. 3.2(d). The correlation between nodes is studied with variations in the control parameter ϑ. If the value of $\rho_{(i,j)}$ between two nodes is greater than zero, then they are shown using a connected solid line. In Figs. 3.2(a) and

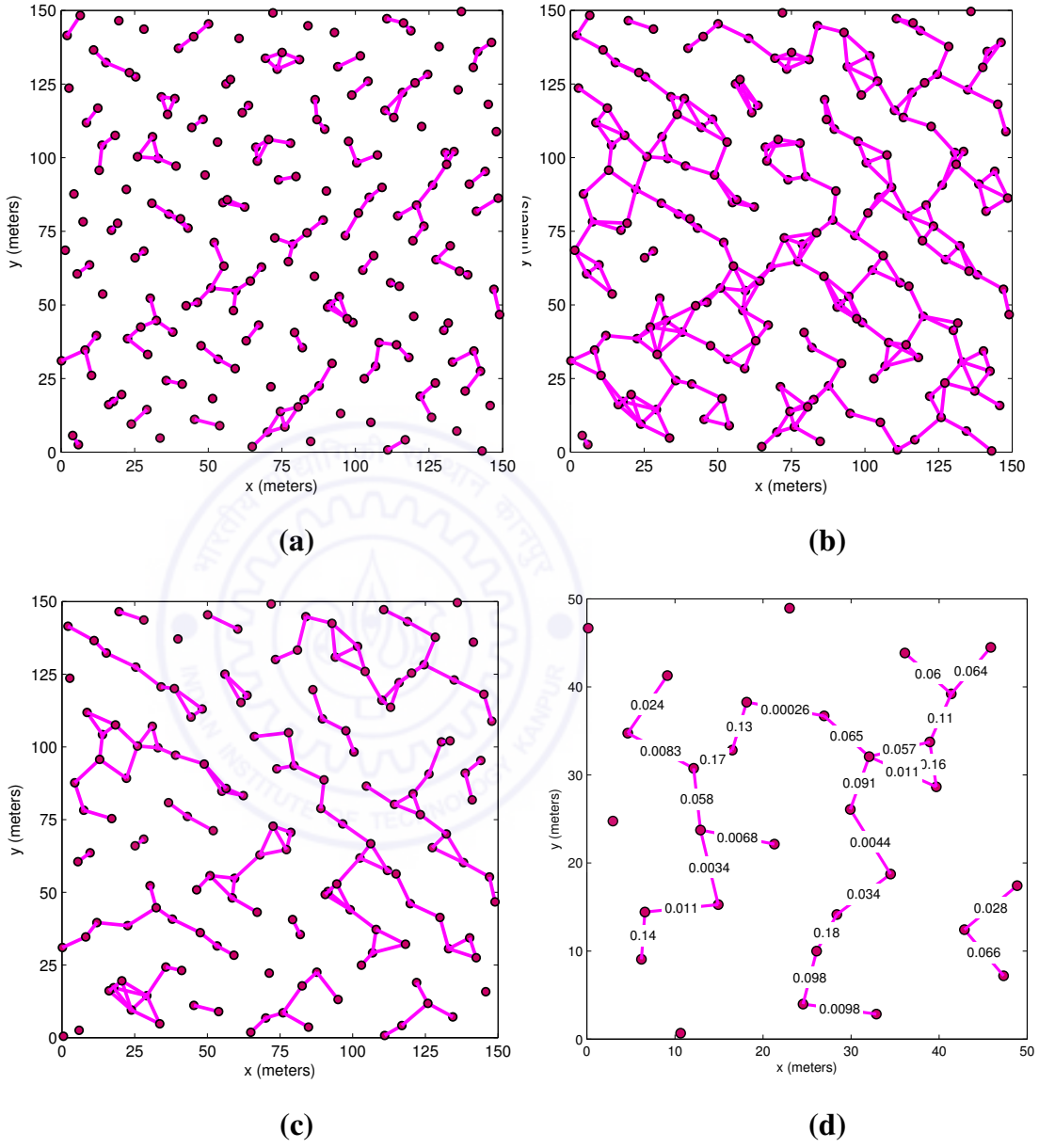


Figure 3.2: *Random distribution for 200 nodes with (a) $\theta_1 = 9\ m$, (b) $\theta_2 = 12\ m$, (c) 150 nodes with $\theta_2 = 12\ m$ and (d) 30 nodes with $\theta_1 = 9\ m$.*

3.2(b), when a node-pair does not show any correlation ($\rho_{(i,j)}$ is equal to zero) both the nodes are out of each other's sensing range and there is no connecting line between them. Fig 3.2(a) with a distribution of 200 nodes with $\theta_1 = 9\ m$ (or $r = 4.5$ m) shows only a few connected lines indicating that very few nodes are correlated with their neighbours. For the same location of the nodes, if we change the sensing range to $\theta_2 = 12\ m$ (Fig. 3.2(b)), more connected lines appear. This indicates that more nodes are now correlated with their neighbouring nodes. When node density changes for a given fixed sensing range (i.e. θ_1), more neighbouring nodes are correlated (Figs. 3.2(b) and 3.2(c)). Fig. 3.2(d) shows the correlation values on the lines connecting any two nodes. The results in tabular format for the same simulation scenario are given in Appendix B.1 using same simulation scenario. Furthermore, the detailed analysis has been provided using correlation function, $K_\vartheta(.)$ in Appendix B.2. Power consumption by the selection of appropriate nodes in the event area with different sizes for detection and reporting applications.

3.2.3 Discussion

In sensor network applications, as long as the area of interest is specified, the correlation characteristics of the sensory observations of nodes can be obtained as described in the previous section. Omni-directional sensing model is the traditionally-used sensing model for WSNs. Generally, sensor nodes can be equipped with temperature, humidity, magnetic field sensors etc. as required. These sensors can operate in a range of 360° around the node. The proposed model is a generic correlation model that can be applied to all sensor network applications [2]. The knowledge of distances between sensor nodes is needed for our model. This can be done by estimating the distances by the nodes based on received signal strength. Thus exact locations are not needed in our model which may require complicated sensor localization mechanisms. The proposed

correlation model can be used for designing energy-efficient communication protocols for most of the WSNs. Whole network can be partitioned into correlated clusters using the presented model. To gain more insight into energy-efficient operation at Network layer using proposed correlation model has been presented in section 3.3.

In order to design an energy-efficient MAC protocols, researchers [14, 15] have used correlation functions [36] to determine the relationship between the spread of sensor nodes and event estimation reliability. In [14, 15], it has been argued that achieved distortion strongly depends on the following two factors: (i) the positive effect of the *correlation coefficient*, $\rho_{(i,j)}$, between each pair of representative nodes n_i and n_j as the distortion decreases with increase in the distance between nodes due to mere independent observation and (ii) the negative effect of the *correlation coefficient*, $\rho_{(s,i)}$, between the event source S and the representative node n_i sending the event information as the distortion increases with increase in the distance between event source and the node. In this case, the location of the event source S is necessary requirement to estimate the $\rho_{(s,i)}$. Unlike existing models [14, 15] that consider the exact location of S, a framework has been developed in section 3.4 using the proposed model which does not require exact loation. Using this framework, the event information can be reconstructed at the sink without knowledge of location of the sensed event. Thus our design presented in section 3.4 will be more suitable for real network scenarios for the given sensing range, location of nodes and distances between them. A comparative study on event distortion has also been carried out using the presented correlation model and the earlier proposed correlation models [14, 15].

3.3 Exploiting Spatial Correlation at the Network Layer

Using the proposed model, we introduce the possible approaches that takes the advantage of spatial correlation for energy-efficient communication protocols. This model

has been used at the networking layer in this section, and the MAC layer in the next section.

3.3.1 Partitioning of WSN as Correlated Clusters

Larger the overlap area between nodes, stronger will be spatial correlation between them. We can define a correlation threshold ξ ($0 < \xi \leq 1$). If $K_\vartheta(d) \geq \xi$, then node n_i and node n_j are strongly correlated. If $K_\vartheta(d) < \xi$, then node n_i and node n_j are weakly correlated. The average number of nodes in a correlated cluster will depends upon sensing radius r (i.e., $\vartheta/2$) and node density in the event area. For highly correlated n_i and n_j,

$$K_\vartheta\{\|s_i - s_j\|\} \geq \xi. \tag{3.7}$$

Using (3.6), we get

$$\frac{2\cos^{-1}\left(\frac{d_{(i,j)}}{\vartheta}\right)}{\pi} - \frac{2d_{(i,j)}}{\pi\vartheta^2}\cdot\sqrt{(\vartheta^2 - d_{(i,j)}^2)} \geq \xi. \tag{3.8}$$

When correlation is ξ, $d_{(i,j)} = R_{corr}$, hence

$$\xi = \frac{2\cos^{-1}\left(\frac{R_{corr}}{\vartheta}\right)}{\pi} - \frac{2R_{corr}}{\pi\vartheta^2}\cdot\sqrt{(\vartheta^2 - R_{corr}^2)}. \tag{3.9}$$

For $d_{(i,j)} < R_{corr}$, $K_\vartheta\{\|s_i - s_j\|\} \geq \xi$.

If any node n_k is lying within R_{corr} from node n_i, it is within a strongly correlated distance from n_i. Otherwise, the correlation is weak. The value of ξ can be determined according to the requirements of the application i.e., data resolution with desired reliability at the sink. A larger value of ξ implies clusters with stronger correlation. For the given parameters ϑ and ξ, a densely deployed WSN can be partitioned into disjoint clusters of

size R_{corr}. The cluster identification is basically the clique-covering problem over correlation graph, which is NP-hard [86]. A correlation graph G can be created where each node is represented by vertex, and edge (u, v) is drawn if the value of $\rho_{(u,v)}$ between node u and v is greater than or equal to ξ. As long as the correlation threshold ξ and the correlation function $K_{\vartheta}(.)$ are given, the sets of correlated clusters are determined through a hierarchical clustering process. The algorithm known as Greedy Correlated Clustering (GCC) algorithm, is presented in the Fig. 3.3.

3.3.2 Correlation based Energy-efficient Data Collection

When a GCC algorithm is used, sets of correlated clusters are formed and all the nodes within a cluster are considered highly correlated. The information is observed by multiple sensor nodes in the event area creating redundant reports. This can be eliminated by exploiting spatial correlation. Only a few nodes need to report their sensory data, and the remaining nodes can remain in a silent state to save energy. The best way to make the proposed correlation model work is to partition the entire event area into correlated, disjoint and equal-sized hexagons with radius R_{corr}.

By exploiting redundancy in sensor observations within correlated cluster, saving of energy is possible at the network layer. Since all the nodes in a cluster can be treated equally and only a small fraction of nodes (or at least one node) are allowed to be active to serve as the representative(s) of the whole cluster. The remaining nodes in the cluster can be kept in the sleep state. Thus network lifetime can be increased when nodes within a cluster share the workload.

Since all the nodes in each correlated cluster most likely observe the same event information, it is desirable to schedule only one node at a time to pass on that information to the sink, in order to save the energy within the cluster. Within each cluster, TDMA scheme can be implemented to assign different time slots for sensing to each node that

Table 3.2: *Results of R_{corr} for different values of ϑ and ξ*

ϑ $(m.)$ \ ξ	0.2	0.4	0.5	0.6	0.8
9	5.78	4.11	3.51	2.3	1.25
12	8.03	5.5	4.7	3.05	1.7
15	10.04	6.23	5.87	3.86	2.16
18	12.07	8.23	7.05	4.67	2.59
21	14.0	9.6	8.23	5.44	3.02
24	16.13	11.07	9.39	6.22	3.46

will schedule the work sequentially in a cluster. A round-robin scheduling method can be adopted within the clusters [38]. For a cluster with k sensor nodes, the period T can be divided into k time slots. In every T time duration, every sensor node will work only for a $\frac{T}{k}$ time period. Initially, the sensor nodes should be time synchronized, and then the sink node can randomly assign a working schedule to the sensor nodes for each cluster.

We simulated the correlation-based clustering using the proposed Greedy Correlated Clustering (GCC) algorithm (see Fig. 3.3) by placing 60 nodes with a sensing range of 20 meters for different values of ξ. The results are shown in Fig. 3.4. It is clearly seen that the number of correlated clusters and the number of member nodes in a cluster depend on the value of the correlation threshold ξ. If ξ is smaller, a larger number of nodes falls in a correlated cluster on an average. On the other hand, a larger value of ξ, allows a lesser number of correlated nodes in a cluster. Similarly, for different values of ϑ with fixed node density, the sensor network is partitioned into different-sized correlated regions. Hence, the required information reliability/fidelity can be achieved through proper clustering by controlling parameters ϑ and ξ in data collection. Table 3.2 shows the results of R_{corr} for different values of ϑ and ξ.

The event area observed by the sensor nodes might change due to a change in environmental conditions. In this case, based on the reliability requirements, the sink node should decide whether the current clusters are valid or not. Then it should re-adjusts the

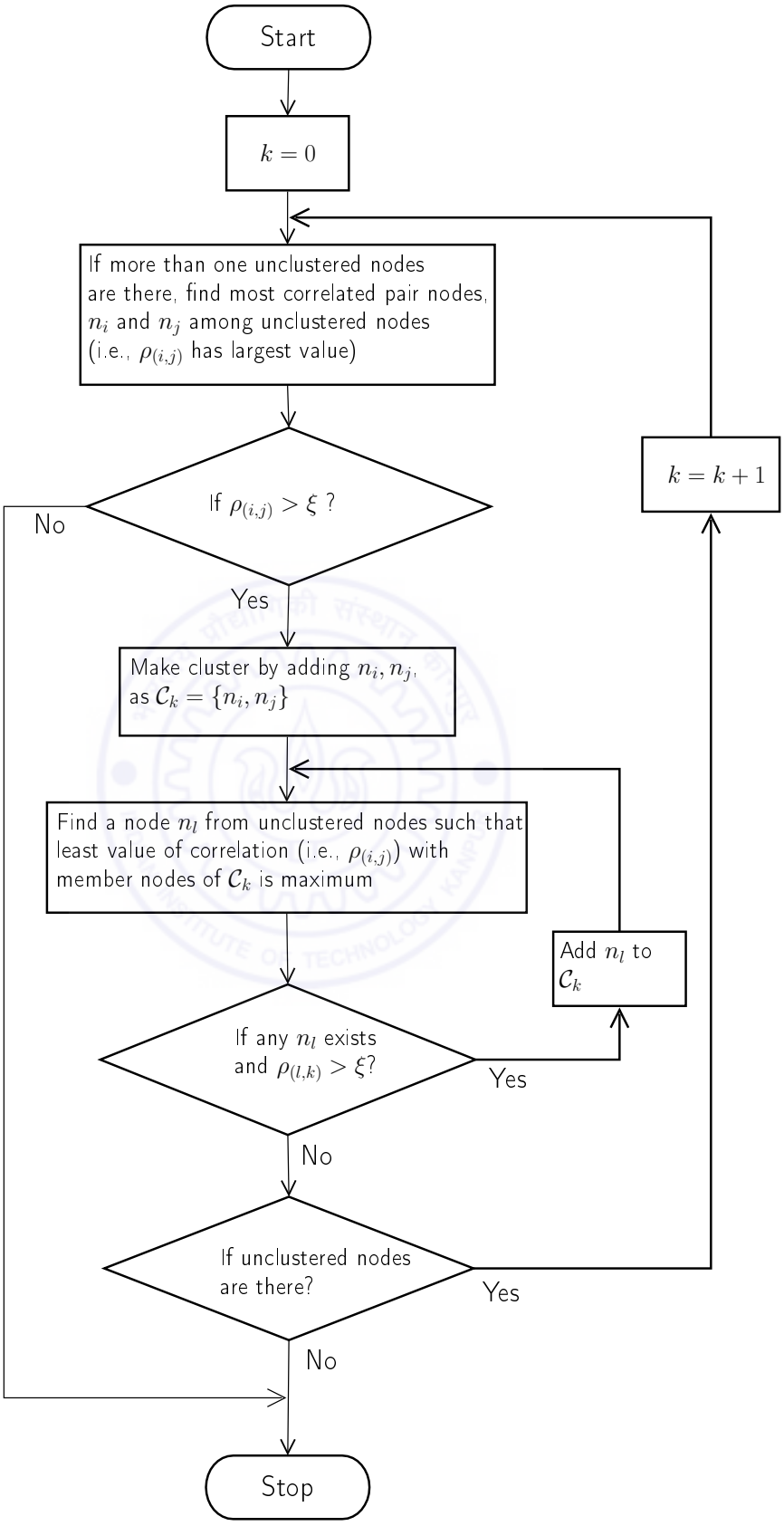
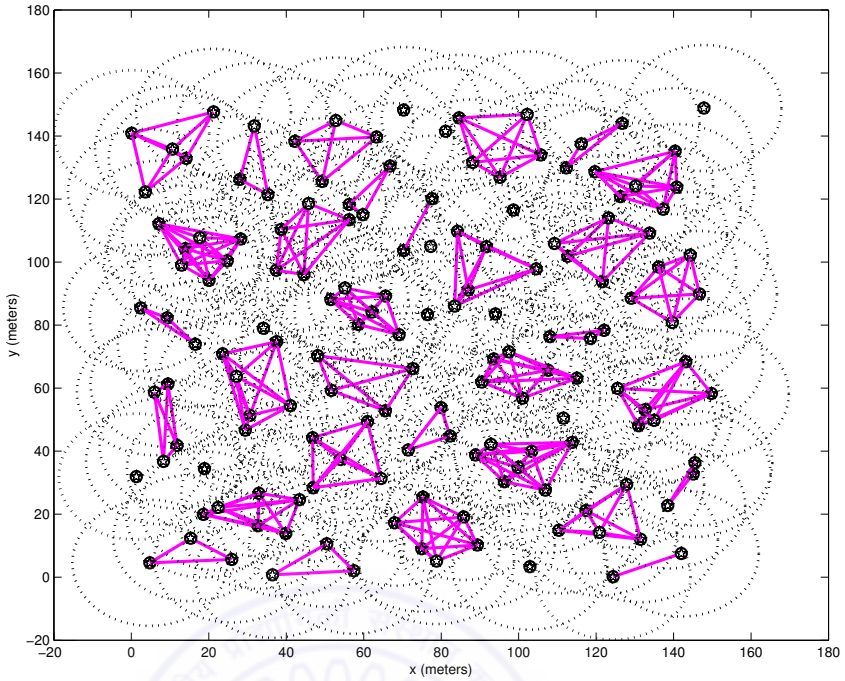
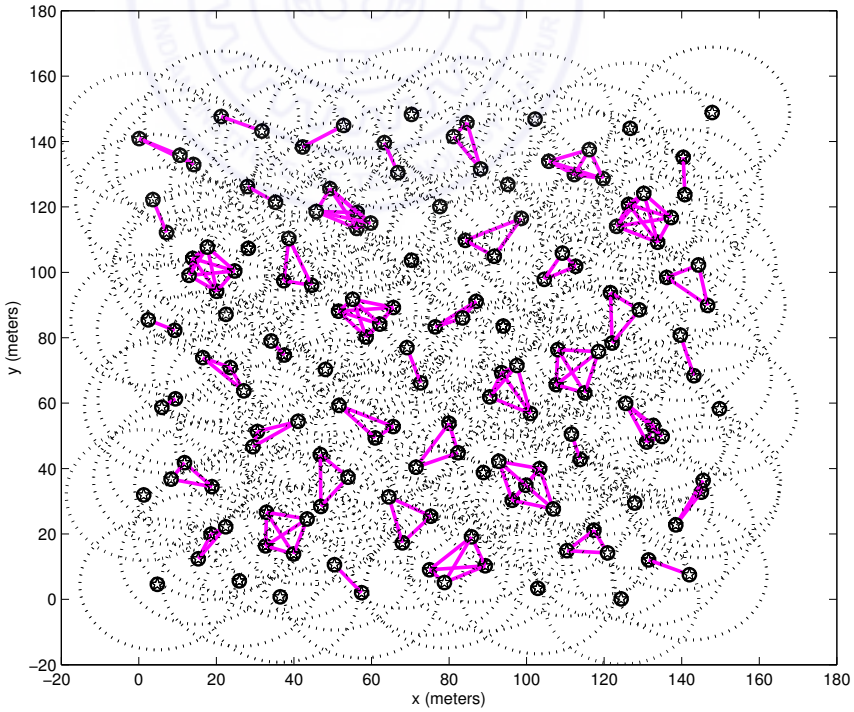


Figure 3.3: *Flow chart of Greedy Correlated Clustering (GCC) Algorithm*



(a)



(b)

Figure 3.4: *Correlation-based clustering results using GCC algorithm given in Fig. 3.3 for* 150 *sensor nodes with* $\theta_1 = 40\ m$ *for (a)* $\xi = 0.25$ *and (b)* $\xi = 0.5$.

clusters as quickly as possible by tuning the parameters ϑ and ξ and re-executing GCC algorithm. Thus, the data reliability can be improved by re-partitioning of the whole network.

3.4 Exploiting the spatial correlation at MAC layer

Several MAC protocols have been proposed to take advantage of spatial correlation. Some of them have been analysed in [11], [14], [15] and [16]. In this section, we first present a framework, using our correlation model to compute the estimation error in terms of event distortion. Here, our objective is to choose the minimum number of nodes inside an event area such that the final achieved event distortion is within the tolerable distortion constraints. A comparative case study has also been presented with the existing models.

3.4.1 WSN Model

Consider a WSN, where N sensor nodes are evenly distributed over a 2-dimensional measurement field $\mathcal{D}$ ($\mathcal{D} \subset \mathcal{R}^2$). The event source S occurs inside the field $\mathcal{D}$ and is located in the center of the event area. The nodes falling in the event area (called reporting nodes) will have threshold exceeded and thus send the sensory information to the sink. The WSN should be designed to minimize mean square error (MSE) distortion between event source S and its estimation $\hat{S}$ from the sensors' collective observations. The MSE distortion of the event is defined as a reliability measure required by the sensor applications. It is given as

$$D_E = \mathbb{E}[d_E(S, \hat{S})]. \tag{3.10}$$

Here $d_E(S, \hat{S})$ is the event MSE distortion measure and $\mathbb{E}(.)$ represents mathematical expectation. In this thesis, we refer D_{max} as the maximum tolerable information distortion

allowed in the sensor network as per the application's requirement. The sink reconstructs an estimate of the event source by exploiting the spatial data correlation while ensuring the given distortion constraint D_{max} (i.e., $D_E \leq D_{max}$). Based on similar assumptions as given in [14] for the estimation of the event source, we extend the framework using our proposed correlation function. The block diagram for Event Source S estimation is shown in Fig. 2.1 of Chapter 2. By considering this model, we can approximately estimate the (x, y) as centroid of all the reporting nodes. The weighted average of the information sensed by all the reporting nodes will give the estimate $\hat{S}$. We also assume that N nodes sense the event and send the information to the sink.

As the sensor nodes send the event information only when it is above a threshold, the mean value of data received from the reporting nodes is expected to be above threshold. As Fig. 2.1 shows, once the event of interest happens, node n_i observes the signal X_i with noise, at time t. Since the samples are temporally independent, we can neglect the time index t.

$$X_i = S_i + N_i; \quad i \in N. \tag{3.11}$$

Here, measured information X_i is assumed to be Jointly Gaussian Random Variable (JGRV) and N_i is zero mean Gaussian noise with variance σ_N^2. Hence,

$$\mathbb{E}[S_i] = m_i, \quad \mathbb{V}ar[S_i] = \sigma_S^2, \quad \mathbb{V}ar[N_i] = \sigma_N^2, \ i = 1, 2, ...N,$$

and

$$\rho_{(i,j)} = K_\vartheta(d_{(i,j)}) = \frac{\mathbb{E}[(S_i - m_i)(S_j - m_j)]}{\sigma_S^2}. \tag{3.12}$$

Here $\rho_{(i,j)}$ is the correlation coefficient between nodes n_i and n_j located at coordinates s_i and s_j. The $d_{(i,j)}$ is the distance between these nodes and $\mathbb{V}ar(.)$ represents mathematical variance. The function $K_\vartheta(.)$ can be determined using the correlation model

discussed in the section 3.3.1.

In the literature, *Spherical, Power exponential, Rational quadratic,* and *Matérn* [36] correlation models have also been proposed. For the case of sensor network with omni-directional sensor nodes, these conventional models do not consider the real network conditions such as sensing range, location of nodes, and distance between the nodes. Our proposed model first needs the placement coordinates of their node and their sensing range. While the earlier proposed ones do not use these parameters. It should be noted that when node placement and sensing range are unknown, our model cannot be used. Therefore, the correlation models must be chosen according to the properties of the physical phenomena and the sensed events' features. The events' features vary significantly for the energy-radiating physical phenomenon originating in the field. In this work, we have considered those events which trigger all the nodes within a radius of sensing range. They do not trigger the nodes which are out of range irrespective of value of S.

Assume, S represents energy-radiating signal that propagates. The signal strength of event source S decays with distance and it follows isotropic attenuation power model, given by

$$S_i = \frac{S_0}{1 + \beta d_i^{\alpha}}, \tag{3.13}$$

where S_0 is signal strength of event source, β is constant, and d_i is Euclidean distance between the node n_i and event source. The signal attenuation exponent is represented by α and it varies from 2 to 3. For simplicity, we use β = 1 and α = 2. Since we have used average of sensed information by all the reporting nodes, the estimated value $\hat{S}_{avg}$ will be smaller than actual S. Let us assume that an event S triggers N reporting nodes within event radius R_E and $\hat{S}_{avg}$ will be average of all the measured event information $\hat{S}_i$ from all these N nodes. Let this $\hat{S}_{avg}$ be equal to decayed version of S at a distance of R_c (where $R_c = R_E/\sqrt{2}$). Since we have used a simpler approach, $\hat{S}_i, \forall i$, is estimated using MMSE at the sink. Using $\hat{S}_i$, $\hat{S}_{avg}$ is estimated, which is further scaled up to get

$\hat{S}$ by a factor obtained from Equation (3.13) for $d_i = R_c$. It should be noted that the estimation $\hat{S}_i$, $\forall i$, is also a JGRV with same properties, as the actual event information S_i , $\forall i$, is also a JGRV. The sink will receive N measured values, as sample observations from N reporting nodes as given by

$$\mathbf{X} = \mathbf{AZ} + \mathbf{W}. \tag{3.14}$$

Here, the vector, $\mathbf{X} \in \mathcal{R}^{N\times 1}$ is sample observation, $\mathbf{Z} \in \mathcal{R}^{N\times 1}$ is random vector for physically sensed event information, $\mathbf{A} \in \mathcal{R}^{N\times N}$ is known transformation matrix and $\mathbf{W} \in \mathcal{R}^{N\times 1}$ represents noise vector. For simplicity, $\mathbf{A}$ is taken as identity matrix. Here $\mathcal{R}$ is set of real numbers. We also have

$$\mathbb{E}[\mathbf{Z}] = [m_1,\ m_2,\ \ldots,\ m_N]^T \in \mathcal{R}^{N\times 1}, \quad \mathbb{C}ov[\mathbf{Z}] = \mathbf{R_{zz}} = \sigma_s^2\mathbf{C_{ss}} \in \mathcal{R}^{N\times N}, \tag{3.15}$$

where

$$\mathbf{C_{ss}} = \begin{pmatrix} \rho_{(1,1)} & \rho_{(1,2)} & \cdots & \rho_{(1,N)} \\ \rho_{(2,1)} & \rho_{(2,2)} & \cdots & \rho_{(2,N)} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{(N,1)} & \rho_{(N,2)} & \cdots & \rho_{(N,N)} \end{pmatrix}. \tag{3.16}$$

The elements of correlation matrix $\mathbf{C_{ss}}$ is defined by our proposed correlation model Equation (3.6). The MMSE estimate, $\hat{\mathbf{Z}}$ of $\mathbf{Z}$ can be formulated using the affine function [87] of $\mathbf{X}$ as

$$\hat{\mathbf{Z}} = \boldsymbol{K}\mathbf{X} + \boldsymbol{b}, \tag{3.17}$$

where $\boldsymbol{K}$ is a matrix and $\boldsymbol{b}$ is a vector. It minimizes the scalar MSE criterion. According to orthogonality principle [87], $\hat{\mathbf{Z}}$ will be a MMSE estimate if $\mathbb{E}[(\mathbf{Z} - \hat{\mathbf{Z}})] = 0$ and

$\mathbb{E}[(\mathbf{Z} - \hat{\mathbf{Z}})\mathbf{X}^T] = 0$. Thus, we have

$$\boldsymbol{b} = \mathbb{E}[\mathbf{Z}] - \boldsymbol{K}\mathbb{E}[\mathbf{X}], \quad \boldsymbol{K} = \mathbf{R_{zz}}(\mathbf{R_{zz}} + \sigma_N^2\mathbf{I}_N)^{-1}. \tag{3.18}$$

Using Equations (3.14), (3.15) and (3.18), Equation (3.17) can be simplified as

$$\hat{\mathbf{Z}} = \mathbb{E}[\mathbf{Z}] + \mathbf{C_{ss}}(\mathbf{C_{ss}} + \frac{\sigma_N^2}{\sigma_S^2}\mathbf{I}_N)^{-1}\left(\ \mathbf{X} - \mathbb{E}[\mathbf{X}]\ \right). \tag{3.19}$$

The error covariance matrix, $\mathbf{R_{ee}}$ can be calculated as

$$\mathbf{R_{ee}} = \mathbb{E}[(\mathbf{Z} - \hat{\mathbf{Z}})(\mathbf{Z} - \hat{\mathbf{Z}})^T] = \mathbf{R_{zz}} - \mathbf{R_{zz}}(\mathbf{R_{zz}} + \sigma_N^2\mathbf{I}_N)^{-1}\mathbf{R_{zz}} = (\mathbf{R_{zz}}^{-1} + \sigma_N^2\mathbf{I}_N)^{-1}, \tag{3.20}$$

where, second quantity is obtained using an identity $\mathbf{A^{-1}}+\mathbf{A^{-1}C}(\mathbf{D} - \mathbf{BA^{-1}C})^{-1}\mathbf{BA^{-1}} = (\mathbf{A} - \mathbf{CD^{-1}B})^{-1}$ [88].

Here $\hat{\mathbf{Z}} = [\hat{S_1},\ \hat{S_2},\ \ldots,\ \hat{S_N}]^T \in \mathcal{R}^{N\times 1}$. The estimated values, $\hat{S_i}$'s are spatially correlated because the observed event information X_i's are spatially correlated. The too-much redundancy in data is not needed once the estimates is within distortion constraint (denoted by D_{max}) as decided by application requirement. If somehow, only a subset of nodes is allowed to transmit, it will be good enough to meet this constraint D_{max}. Therefore, we investigate the achieved distortion when only M out of N nodes are selected to send event information to the sink. Since the estimator uses the average of estimate $\hat{S_i}$ with scaled version obtained from Equation (3.13), $\hat{S}$, the estimate of S is given as

$$\hat{S}(M) = \frac{(1 + R_c^2)}{M}\sum_{i=1}^{M}\hat{S_i} = \frac{(1 + R_c^2)}{M}[1,\ 1,\ \ldots,\ 1][\hat{S_1},\ \hat{S_2},\ \ldots,\ \hat{S_M}]^T. \tag{3.21}$$

The event distortion achieved by selecting M packets, when each reporting node transmits only one report to the sink, can be estimated from average value of sensed infor-

mation. The achieved event distortion D_E is calculated as MSE [88, 87] that is equal to $\frac{(1+R_c^2)}{M^2}\ trace(\mathbf{R_{ee}})$ using Equations (3.20) and (3.21). Here $trace(\mathbf{A})$ is the trace of the matrix **A**. The calculation of distortion $D_E(N, M)$ can be simplified by selecting M nodes ($M < N$) through vector expansion in an alternative form as

$$D_E(N, M) = \mathbb{E}\left[\left(S(N) - \hat{S}(M)\right)^2\right]. \tag{3.22}$$

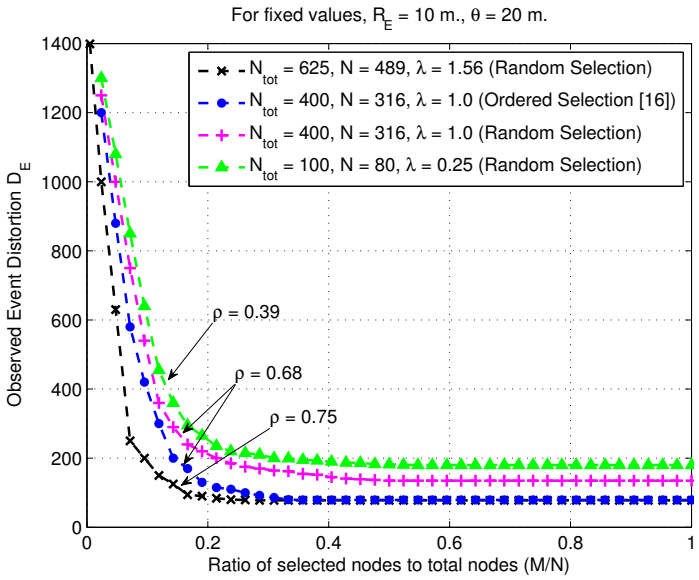
Expanding Equation (3.22) using Equations (3.11), (3.12) and (3.21),

$$\begin{aligned} D_E(N, M) = {} & \frac{(1+R_c^2)^2}{M^2}\left(\frac{M\sigma_S^4}{(\sigma_S^2+\sigma_N^2)} + \frac{\sigma_S^6}{(\sigma_S^2+\sigma_N^2)^2}\sum_{i=1}^{M}\sum_{\substack{j=1\\ j\neq i}}^{M}\rho_{(i,j)} + \sum_{i=1}^{M} m_i^2\right) \\ & -\frac{2(1+R_c^2)^2}{NM}\left(\frac{\sigma_S^4}{(\sigma_S^2+\sigma_N^2)}\sum_{i=1}^{N}\sum_{\substack{j=1\\ j\neq i}}^{M}\rho_{(i,j)} + \sum_{i=1}^{N}\sum_{\substack{j=1\\ j\neq i}}^{M} m_i m_j\right) \\ & +\frac{(1+R_c^2)^2}{N^2}\left(N\sigma_S^2 + \sigma_S^2\sum_{i=1}^{N}\sum_{\substack{j=1\\ j\neq i}}^{N}\rho_{(i,j)} + \sum_{i=1}^{N} m_i^2\right). \end{aligned} \tag{3.23}$$

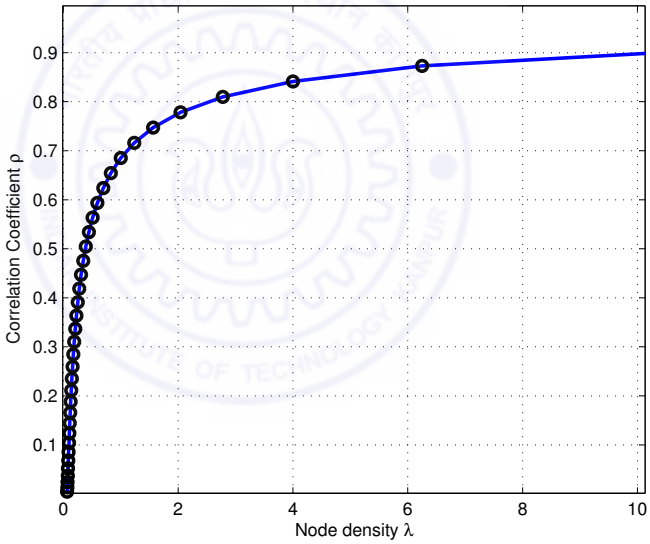
Here the correlation coefficients $\rho_{(i,j)}$ are the elements of correlation matrix $\mathbf{C_{ss}}$ which takes into account the sensing range and node separation. The detailed derivation of above Equation is given in Appendix C. Using the above distortion function, a case study has been presented in the next section.

3.4.2 Case study

The achieved event distortion at the sink is computed using Equation (3.23) which is the function of M reporting nodes, the correlation matrix $\mathbf{C_{ss}}$ and R_c, while other parameters are constants such as $N, m_i,\ \sigma_S,\ \sigma_N$. In this section, we study the impacts of node density, selected node number and sensing range (where $r = \vartheta/2$) on the distortion performance using Equation (3.23). As discussed in section 3.2.3, the achieved distortion strongly depends on the correlation coefficient, $\rho_{(i,j)}$. In this case, the knowledge of



(a)



(b)

Figure 3.5: *(a) The average distortion for different λ values according to changing the number of selected nodes in the event area. (b) The change of correlation coefficient with node density λ for fixed $\theta = 4\ m$ in grid topology.*

location of event source is not a necessary requirement. Unlike existing models [14, 15], our model does not need actual location of event source and thus will be more suitable for real network scenarios.

We conduct simulations over $\sqrt{N_{tot}}d \times \sqrt{N_{tot}}d\ m^2$ grid area where N_{tot} sensor nodes are located on the grid. The node density λ, as the number of nodes per unit area is $\frac{1}{d^2}$. We place an event source in the center of grid area where N nodes are activated within the event radius R_E. The simulations are performed with 1000 trails and the results are produced by averaging them as shown in Figs. 3.5 and 3.6. Fig. 3.5(a) shows the results of the event distortion for different values of node density λ as fraction of ratio of the number of selected nodes to the reporting nodes (i.e., M/N) for a given event area. It is observed that the distortion approaches to relatively constant when the reporting nodes (i.e., M) are more than a tpercentage of nodes sensing the event (i.e., N). We have chosen the reporting nodes such that they should be as closer as possible to center of area within which N nodes have sensed the information, while selecting the reporting nodes are as far as possible from each other [85]. When node density λ increases, the spatial correlation coefficient also increases for fixed value of ϑ as shown in Fig. 3.5(b). It means that a higher node density indicates a stronger correlation among data sample resulting in better estimation reliability. Hence, the event distortion is a monotonically decreasing function of λ as shown in Fig. 3.5(a). For node density 0.25 node density, 20% of total reporting nodes are sufficient enough to report the event reliability. This further implies that optimum performance can be achieved by allowing much lesser number of reporting nodes M to be activated in the event area ($M < N$). Thus, the energy can be saved by exploiting spatial data correlation to limit the reporting nodes while ensuring acceptable distortion constraint D_{max}. Fig. 3.6 shows the results of the M reporting nodes for different values of correlation coefficient ρ when the observed event distortion does not exceed D_{max} at the sink.

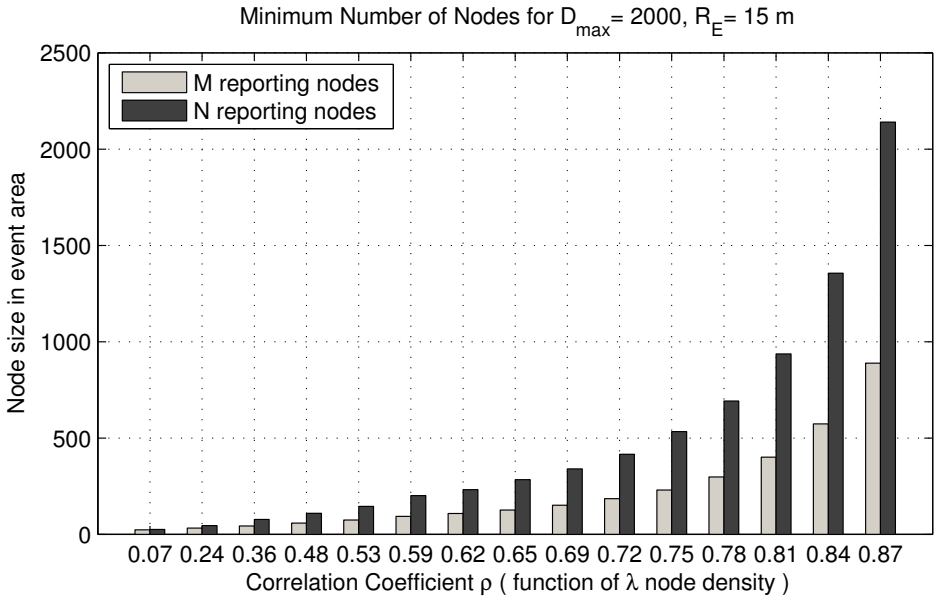


Figure 3.6: *Optimal value of N_{min} reporting nodes with different correlation coefficients ρ.*

3.5 Comparative study with existing correlation models

In this section, a comparative study has been provided to investigate the impacts of correlation coefficients on distortion performance. We extend the work of [14] and [15] using an alternative event driven sensing model and a new correlation function. In [14] and [15], the Power exponential model [36] is used as the correlation model to compute the event distortion function. However, it does not consider the real network conditions such as the position of nodes, sensing coverage and distance between nodes. Earlier work have argued that achieved distortion strongly depends on two factors; (i) the positive effect of the correlation coefficient, $\rho_{(i,j)}$ and (ii) the negative impact of the correlation coefficient, $\rho_{(s,i)}$. Both factors play a major role for making more reliable decision in estimating the event source at the sink.

A correlation model must be chosen according to the properties of the physical phenomena and the sensed events' features. The events' features vary significantly for the energy- radiating physical phenomenon originating in the field. We use the Power-

exponential model of the event source S, while our proposed correlation function is used to model the correlation characteristics between the nodes. The value of $\rho_{(i,j)}$ between each pair of representative nodes n_i and n_j, can be estimated using our model given for omni-directional sensor nodes, that is restated from (3.6) as

$$\rho_{(i,j)} = \frac{\cos^{-1}\left(\frac{d_v(i,j)}{\vartheta}\right)}{\pi} - \frac{d_v(i,j)}{\pi\vartheta^2}\cdot\sqrt{(\vartheta^2 - d_v^2(i,j))}, \tag{3.24}$$

where $\vartheta = (\theta_1,\ \theta_2,\ ...,\ \theta_c)$. The value of $\rho_{(s,i)}$, between the event source S and the representative node n_i sending the event information, can be estimated using Power exponential model [36] as

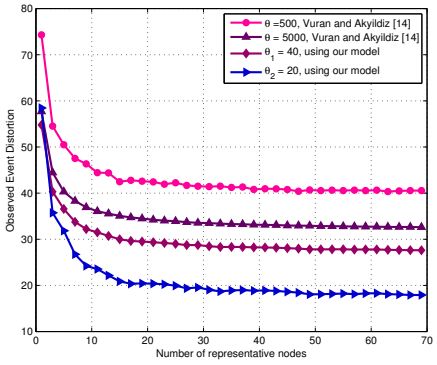
$$\rho_{(s,i)} = \mathrm{e}^{(-d/\theta_1^s)^{\theta_2^s}}, \quad \text{for } \theta_1^s > 0, \quad \theta_2^s \in (0, 2]. \tag{3.25}$$

Here, $\theta_2^s = 1$ for power exponential model. The Vuran and Akyildiz [14] and Zheng and Tang [15], both provided a different distortion function, $D(M)$ as the observed event information distortion at the sink when M representative nodes (out of total N nodes in the event area) transmit reports to the sink. Here, it is assumed that each node transmits only one packet as a report to the sink. Various distortion functions are given below.

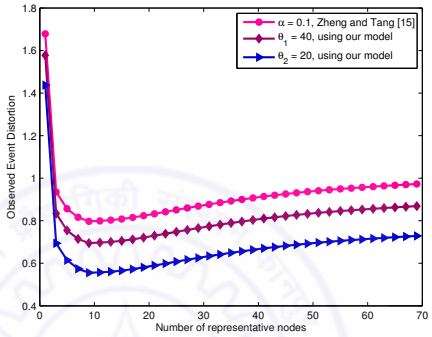
(i) Vuran and Akyildiz [14], worked it out as

$$\begin{aligned} D(M) = \sigma_s^2 - \frac{\sigma_s^4}{M(\sigma_s^2+\sigma_N^2)}\left(2\sum_{i=1}^{M}\rho_{(s,i)} - 1\right) \\ + \frac{\sigma_s^6}{M^2(\sigma_s^2+\sigma_N^2)^2}\sum_{i=1}^{M}\sum_{j\neq i}^{M}\rho_{(i,j)}. \end{aligned} \tag{3.32}$$

(ii) Zheng and Tang in [15] proposed it as



(a)



(b)

Figure 3.7: *Comparison of average event distortion performance using (a) proposed correlation model with random selection and the work of Vuran and Akyildiz [14], (b) proposed correlation model with ordered selection and the work of Zheng and Tang [15].*

$$D(M) = \sigma_s^2 + a^2 - \frac{2\sum_{i=1}^{M} E_i[\rho_{(s,i)}\sigma_s^4/(\sigma_s^2+\sigma_N^2) + aa_i]}{\sum_{i=1}^{M} E_i^2} + \left\{\sum_{i=1}^{M}\sum_{j\neq i}^{M} E_i E_j\left(\frac{\rho_{(s,i)}\sigma_s^6}{\sigma_s^2+\sigma_N^2} + a_j a_i\right)\right\}\left(\sum_{i=1}^{M} E_i^2\right)^2 + \frac{\sum_{i=1}^{M} E_i^2[\sigma_s^4/(\sigma_s^2+\sigma_N^2) + a_i^2]}{\left(\sum_{i=1}^{M} E_i^2\right)^2}. \quad (3.33)$$

Using the same assumptions and expressions given by (3.32) and (3.33), we extend the work of [14, 15] by applying the proposed correlation model (3.24) for the purpose of

Table 3.3: *Results of ϑ for different values of N_{min} and D_{max}*

N_{min} \ D_{max}	22	24	26	28	30	32
15	10	28	49	69	87	111
17	18	36	58	80	99	118
19	26	47	67	89	110	127
21	38	58	77	97	119	139
23	47	69	89	108	128	148
25	54	74	98	116	137	157

computing event distortion. To gain more insight into the event reconstruction distortion using our correlation model, we conducted a comparative analysis with the methods suggested in [14] and [15] respectively. In a 1000×1000 m^2 grid area, we have deployed 70 nodes randomly and a fixed event source in the center of the grid area. The event reconstruction distortion has been calculated over a range of $0 - 70$ sensor nodes that are sending event information to the sink. The simulation was performed using a fixed topology with 1000 trials. The event distortion was evaluated using the proposed as well as the earlier correlation models [14, 15]. These results are plotted in Fig. 3.7.

For the given total nodes N, the distortion decreases logically with the increase in representative nodes (M) that are activated to send the event information as shwon in Fig. 3.7(a). This happens as more event information is received at the sink with larger M. However, using our model, at the same number of M representative nodes, even lesser distortion is achieved (Table 3.3). It is shown that achieved distortion can be preserved at the sink using our model by allowing N_{min} nodes ($< M$) to be active. If somehow, only selected N_{min} nodes transmit event information, we can reduce the energy consumption. Fig. 3.7(b) shows a comparison using the method suggested in [15] and our model. It indicates the same behaviour, that reduced achieved distortion allows a lower value of N_{min} which is less than M. Hence, using our correlation model, energy consumption in data transmissions and collisions can be minimized at the MAC layer.

Furthermore, the observed event distortion decreases with decreasing sensing range as shown in Figs. 3.5 and 3.7. To illustrate this, we can have a predefined D_{max} value as allowed by the sensor network application. The results obtained, (Table 3.3) shows that for the lower D_{max}, θ_1 has to be 10 m while for higher D_{max}, θ_2 can be 111 m for same N_{min} (i.e., 15) representative nodes. Higher θ_2 implies that less number of nodes are used in same field. Hence, for same N_{min} nodes and lesser distortion D_{max}, the sensing range of the nodes needs to be kept small through the control parameter (i.e. ϑ). Consequently, we can state based on our correlation model that the power consumption can be minimized by somehow allowing smaller number of nodes to report the event while attaining the desired event reliability as specified by D_{max}. Both Fig. 3.7(a) and Fig. 3.7(b) indicate a lower bound on the minimum number of representative nodes that need to send the event information based on our correlation model.

3.6 Summary

In this chapter, a basic correlation model has been introduced to represent the correlation characteristics between sensor nodes. Using the proposed correlation model, various key elements have been discussed at both Network and MAC layer. It is found that the densely deployed WSN can be partitioned into non-overlapping correlated regions by using proposed GCC algorithm at network layer. Thus, only single node needs to be allowed to report data rather than every node from a given correlated region, resulting in significant energy savings during data collection. Furthermore, our results show that required event reliability can be achieved through proper tuning of ϑ and ξ, i.e., both the sensing range and the correlation threshold for reliable data delivery. In the context of MAC, a new model framework is developed to investigate the impacts of node density, number of selected nodes and nodes' sensing range on the achieved event distortion at the sink. For the estimation of event source, the optimum node density within an event

area can be found at minimum observed event distortion. In addition, a comparative study showed that the proposed correlation model outperforms existing correlation models. Using the proposed model, it is observed that optimal value of N_{min} nodes can be achieved through proper tuning of ϑ for given D_{max} as per the application requirement. In this context, it is also concluded that most of the redundant transmissions can be filtered out on the basis of presented model to make a more energy-efficient MAC design for sensor network applications.

Chapter 4

A localized, Spatial correlation based collision-free MAC for EWSNs

In Event-driven Wireless Sensor Networks (EWSNs) having a high density of sensor nodes whenever the phenomena of interests (PoIs) or events are detected and transmitted, the transmitted measurements are spatially correlated, often redundantly. Since each event triggers a large number of nodes for sensing and transmission, a large volume of data may be generated for transmission. In such a scenario, sensors encounter severe contention in the event area. By effectively exploiting spatial correlation and careful node scheduling in the event area, unnecessary energy costs for both correlated transmissions and collisions can be largely reduced. In this chapter, a novel hybrid TDMA/CSMA MAC, known as Event-MAC is proposed. It takes into account the spatial correlation and prioritized node scheduling for efficient event reporting. By activating the Event-MAC, only when an event is detected, its dynamic time slot allocation guarantees collision-free channel access with full channel utilization in the event area. Section 4.1 gives the motivation about this work. The related work for Medium Access Control (MAC)

protocols in WSNs is provided in Section 2.2. Based on the framework developed in the section 3.5 of Chapter 3, the node selection techniques are first introduced here in Section 4.2. A novel correlation based Hybrid MAC called as Event-MAC is then developed. It's detailed design and implementation is discussed in section 4.3. The operation of Event-MAC using an example is illustrated in Section 4.4. Section 4.5 discusses the Event-MAC performance analysis along with its comparative study. Finally, the chapter summary is provided in Section 4.6.

4.1 Motivation

Event-driven wireless sensor networks (EWSNs) is a spacial class of sensor networks that are composed of large number of sensor nodes distributed in a terrain to sense the PoIs. Such networks are considered to be event-driven where sensor nodes generate data to send only when an event occurs in the monitoring area. Users expect accurate event notification as fast as possible from these networks. Thus, EWSNs should provide efficient reporting of critical data such as location of the event. Since each event triggers a large number of nodes for sensing and transmission, a large volume of data may be generated in a short span and is available for transmissions. In such scenarios, sensors encounter severe contention in the event area.

Many existing MAC protocols (e.g. S-MAC [45], T-MAC [48], P-MAC [50] etc.) utilize the IEEE 802.11 DCF mechanism [46] in order to handle hidden terminal, exposed terminal and network congestion problems. For EWSNs, IEEE 802.11 DCF [46] is not suitable as the larger value of contention window size is not always necessary. Usually traffic is very low and becomes intense only when any activity is observed in a specific zone of sensor field. Therefore, the performance of IEEE 802.11 DCF is not as desired when network traffic is frequent or correlated depending on the nodes' interaction with physical environment. Tay, Jamieson, and Balakrishnan [89] have introduced

the characteristics for event-driven sensor network applications as follows.

- An event-driven sensor network encounters spatially-correlated contention on the occurrance of an event in a particular region/zone of monitored area. The multiple nodes of same neighborhood sense it and send data to report it to the base station. As a result, a synchronized burst of transmissions happens.

- In many applications, all the packets need not be treated as equally important. It is enough if only some of the nodes are successful in transmitting the data.

- The number of nodes getting activated by an event in a particular region changes with time. For example, when a target enters into a sensor field, the number of active sensing nodes could become large very quickly.

In these types of traffic patterns, channel access delay and system throughput are the performance limiting factor when number of reporting nodes are large in number for an event. These characteristics pose a challenge in design of MAC protocol for EWSNs.

In this chapter, a novel hybrid TDMA/CSMA MAC, known as Event-MAC is proposed that takes into account the spatial correlation and priorities on node scheduling for the efficient event reporting. By activating Event-MAC only when an event is detected, its dynamic time slot allocation guarantees collision-free channel access with full channel utilization in the event area. The time is divided into control slot period and data slot period. Data slots are assigned according to priority levels which are based on the received energy level of the event source and parameters such as packet type, queue length, packet delay and battery power level. Simulation results demonstrate that the proposed correlation based protocol outperforms the existing MAC protocols and has unique features such as efficiency, simplicity, energy-efficiency, fairness and QoS provisioning.

4.2 WSN Model used for MAC design

Consider a WSN, where N sensor nodes are evenly distributed over a measurement field $\mathcal{D}$ ($\mathcal{D} \subset \mathcal{R}^2$). The event source S is assumed to be located in the center of the event area inside the field $\mathcal{D}$. The nodes falling in the event area will have the threshold exceeded (also called reporting nodes) and thus send the sensory information to the sink. Each sensory node is assumed to be able to sense an event happening within the radius r around it. Thus all the nodes within a radius r around the event location will be able to sense and hence can act as reporting nodes. The event-based WSN scenario is shown in Fig. 4.1.

As per discussions in the section 3.5 of Chapter 3 [90], it was found that the optimum performance can be achieved by activating lesser number of reporting nodes M in the event area ($M < N$). Here, N is number of nodes which could sense the event. Thus, the energy can be saved by exploiting spatial data correlation to limit the reporting nodes while ensuring acceptable distortion constraint D_{max}. To apply these results, a node selection technique is required to select the optimal number of reporting nodes based on D_{max}, resulting in energy saving. We introduce a node selection technique in next paragraph, which is used by the sink in order to select the reporting nodes and their locations based on distortion constraint, D_{max}. Accordingly, the sink can determine average distance between selected reporting nodes and this computed average distance is broadcast into the network. The nodes in the event area perform Event-MAC using this average distance as explained in the section-4.3.

4.2.1 Node selection techniques

Based on results obtained on the distortion function given by Equation (3.23) in previous chapter, one can select reporting node ($< N$) while achieving acceptable distortion constraint, D_{max}. If we choose the reporting nodes such that (i) they are as closer as

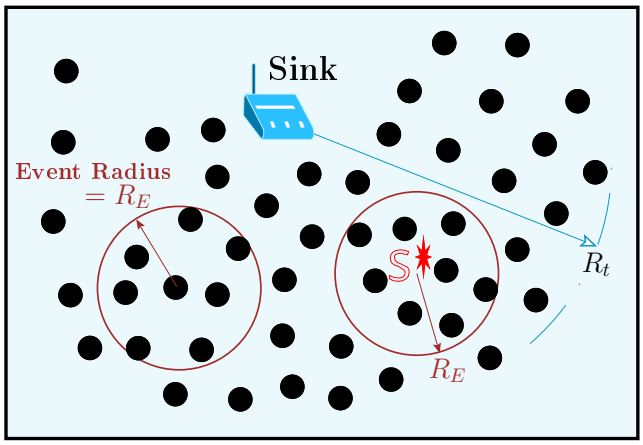


Figure 4.1: *Event-driven scenario with Source S.*

possible to center of area where N nodes sensed the event information and (ii) they are also as far as possible from each other, then even lower than M number of nodes provides the same D_{max}. These facts are taken into considerations in our proposed MAC design as discussed in section-4.3.

As stated earlier, in order to use the correlation model, the problem is the determination of the optimal number of reporting nodes and their locations for the given information reliability in term of D_{max}. In coding theory, analysis of correlated signal field and selection of correlated points in the signal field based on distortion constraint have been investigated with well-established Vector Quantization (VQ) methods [91]. Using VQ approach, Iterative Node Selection (INS) algorithm has been proposed for current problem [14]. The INS algorithm determines the minimum number of reporting nodes and their locations based on distortion constraint, D_{max}. In other words, the average minimum distance, denoted by R_{corr} between selected reporting nodes is derived using VQ approach. This R_{corr} value is later used by CC-MAC to filter out correlated event information by suitably selecting the nodes [14].

In our proposed MAC design, we use a different algorithm to derive the R_{corr} for a distortion constraint, D_{max}. The flow chart of node selection algorithm is shown in

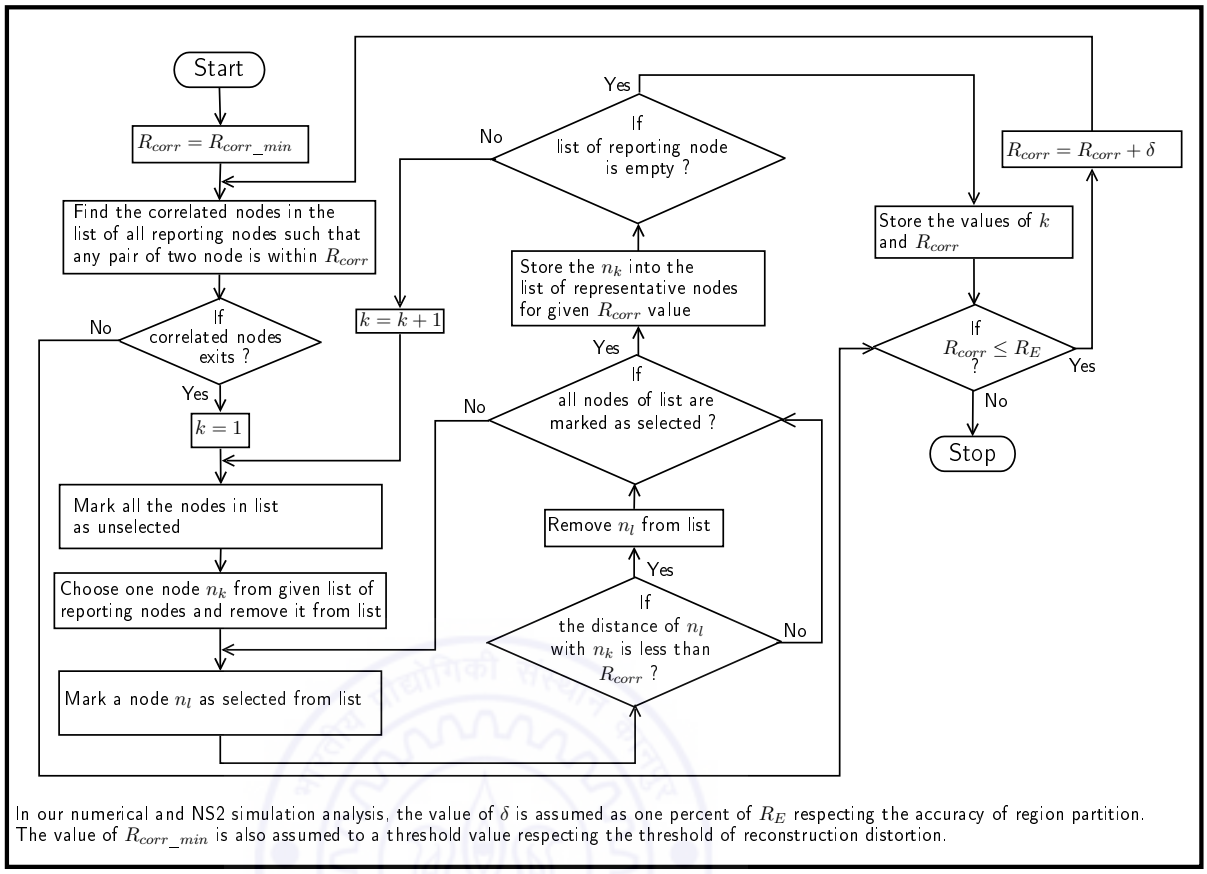


Figure 4.2: *Flow chart of representative node selection Algorithm*

Fig. 4.2. The algorithm finds the value of M for given R_E, R_{corr} and N. It should be noted that the predefined value of D_{max} provides the optimal number of reporting nodes denoted by M. For example, for given D_{max} to be equal to 100 (using Fig. 3.5), the minimum reporting nodes, M can be 24. For 24 reporting nodes, R_{corr} will be 13 using our algorithm. The event radius, R_E is chosen to be $50m$ as shown in Table 4.1 for these calculation.

Table 4.1: *Results of our node selection algorithm ($R_E = 50m$, $N = 49$)*

Minimum selected reporting nodes (M)	49	43	28	24	19	16	14	10
Minimum distance between nodes (R_{corr})	1	6	11	13	15	18	19	21

The node selection algorithm is executed at the sink node after the nodes' deployment. It requires location information and sensing range of omni-directional sensors. The knowledge of distances between sensor nodes is needed here. This can be done by distances estimated by the nodes based on received signal strength. The exact locations are not needed in our design since only relative distances between nodes are sufficient. Thus, complicated sensor localization mechanisms are not needed. Any event will trigger all the nodes within a radius of sensing range. Unlike CC-MAC [14], our design is more realistic as it takes the advantage of all type of events where sensing ranges play major role in the event detection process. In addition, proposed MAC protocol provides collision free channel access by scheduling the packets based on their priorities and correlations. In the next section, we present correlation based TDMA scheduling approach, also called localized, event-oriented MAC for event-driven sensor network applications.

4.3 Event-MAC: A localized, event-oriented MAC

The sensing process is usually omni-directional [2]. A node can detect an event within its limited sensing range. The event detection depends on the sensed strength of emitted event signal, properties of physical phenomena, sensed event's features and the hardware of the node. An event may trigger large number of nodes for sensing and transmission. For example, in surveillance applications, a target in sensor field will be detected by the several nodes in its proximity. These nodes become active and send reports about the target, to the sink. We call these nodes as source nodes or reporting nodes for the detected target or event. The area covered by these nodes is referred to as the influencing region or the event area. Two type of packets can exist in the event area: newly generated reporting packets by source nodes and forwarded reporting packets sent by intermediate nodes, also called the route-thru packets. When a node is successful in injecting its own newly generated packets into the network, other nodes will treat these

as route through packets.

When an event appears in a specific region, the Event-MAC localized TDMA is triggered by the reporting nodes, which are suddenly activated for reporting the event to the sink. The Event-MAC is 'event-oriented' in the sense that the TDMA frame only operates locally inside the event area. Sensor nodes outside of the event area perform pure CSMA based channel access as a base line scheme to route event packets to the sink.

The nodes that detect the event consider themselves to be in the event area. These nodes are then deemed as *E*-nodes. all *E*-nodes perform TDMA while the other nodes that do not sense the event (e.g., nodes outside the event region) perform CSMA. As soon as a node detects an event, it becomes an *E*-node. Since an event triggers multiple nodes in same neighbourhood at the same time, *E*-nodes appear to be synchronized to each other.

Every *E*-node keeps a timer that expires if an *E*-node does not have its own generated packets for transmissions. When timer expires, the *E*-nodes waits for route-thru packets for a wait time equal to the control slot period. It ensures that TDMA *superframe* operates inside the event area till the packets from all *E*-nodes are sent successfully out of event area. It means that the TDMA scheduling is responsible to provide collision-free transmission of both type of packets (i.e., newly generated packet and route-thru packet) inside the event area. TDMA *superframe* vanishes when all packets leave the event area and traverse the other nodes to reach the sink.

Our protocol minimizes the number of unused slots by dynamically assigning slots to *E*-nodes depending on the priority of packets. Furthermore, Event-MAC eliminates the redundant transmissions by stopping the correlated nodes as specified by R_{corr}.

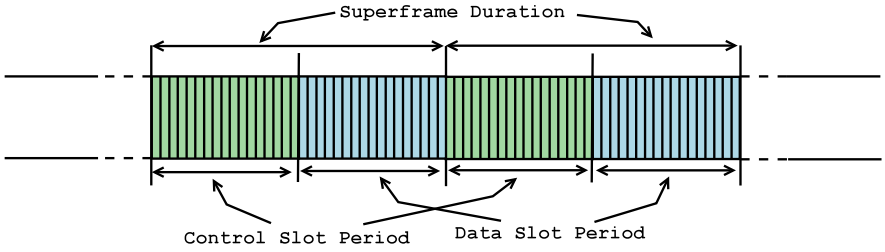


Figure 4.3: *Structure of* superframe *in event-oriented MAC*

4.3.1 Framing Structures in Event-MAC

The TDMA frame structure, also called as *superframe*, is shown in Fig. 4.3. Several *superframes* may be repeated to schedule the packets over the channel. The channel is divided into uniform time slots. A *superframe* is composed of a control slot period and data slot period. The slots in control slot period are allocated for control packet transmission, while data slots are used for data packet transmission. In each control packet, synchronization information is used to ensure the synchronization of *E*-nodes. Guard period is also used in each slot to reduce the error in reception.

Initially, sensor nodes follow CSMA-based channel access scheme with low power listening (LPL) mode (like B-MAC) when there is no activity. B-MAC [92] uses the interfaces of LPL and clear channel assessment (CCA) for higher energy-efficiency. Once an event appears in a specific region or zone, triggered sensor nodes start TDMA schedule on the fly in that zone. Each *E*-node transmits over the allocated time slots. Determining the number of data time slots is adaptive because this number is determined from information received using control packets and on the available number of active *E*-nodes in the event area.

Fig. 4.4 shows the structure of the control packet used in Event-MAC. A control packet consists of synchronization (syn) bits, source (Src) ID, destination (Dst) ID, priority, indicator bits (see Table 4.2), Source packet (SP) field, received signal strength of event source (i.e., energy level), and transmit power. The control slot period includes

Table 4.2: *Description of Indicator Bits*

Indicator Bits	Description
000	Control Packet
001	Data Packet
010	Acknowledgement to data Packet

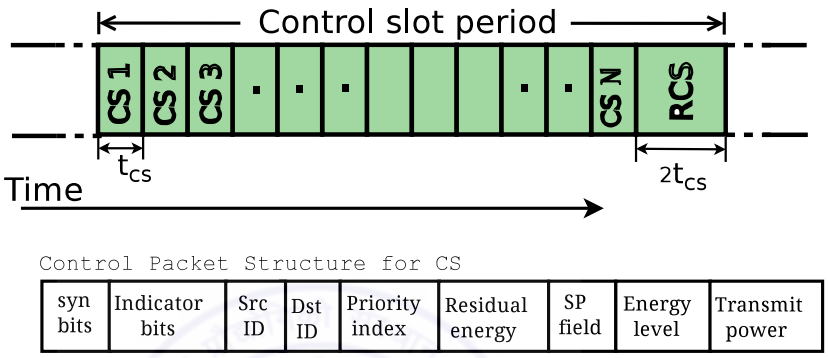


Figure 4.4: *Control Packet structure for control slot period*

the sequence of time slots for control packets $\{CS_1, CS_2, ..., CS_N\}$ and one reserved control slot (RCS) for a new *E*-node to join the *superframe*. RCS is reserved to initiate the communication if an *E*-node has a data packet for transmission. In order to take part in communication, a *E*-node joins the *superframe* by broadcasting a control packet in RCS as a request for data slot. To capture RCS, all *E*-nodes first perform CSMA based scheme to contend for the channel. Therefore, only a contention winner can broadcast its control packet into RCS. We define the slot length of RCS to be long enough to broadcast a control packet. The slot length of RCS is twice that of CS' s slot length. When an *E*-node broadcasts its control packet successfully on RCS, the RCS in next *superframe* is kept vacant for another new *E*-node by moving the successful *E*-node to subsequent control slot (i.e., CS_1 in Fig. 4.4). If an *E*-node wants to leave the channel, its vacant CS is healed by shifting the nodes in corresponding to successive control slots in sequence to fill the gap in the next *superframe*.

The number of control slots depends on the number of active *E*-nodes in the event

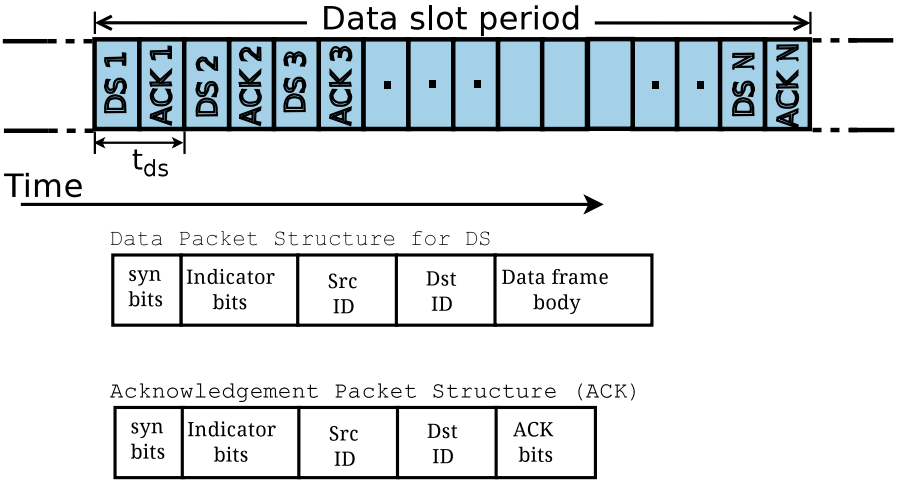


Figure 4.5: *Structures for Data and ACK packets.*

area. Since each *E*-node sends its request for data transmission by a control packet, all active *E*-nodes have the record of the number of *E*-nodes (i.e., reporting nodes) taking part in communication. It is assumed that transmission range is greater than the sensing range or event radius. So the *E*-nodes inside the event area form fully connected sub-network. When an *E*-node gets its control slot, data slot is assigned during data slot period based on the priority of packet. If a node captures the RCS to create the *superframe* first time, there is no need to calculate priority. After RCS, data slot DS_1 is scheduled for data packet transmission. The destination node receives this data packet because it knows when to wake up to receive after getting the control packets in control slot period. The data packet scheduling order is determined by each node receiving the *superframe*, after calculating the priority based on the information received in the control packets. The priority to the packets are given according to type of packet, queue length, packet delay, received energy level of the event source and residual energy of nodes as explained later in Section 4.3.3. The data packet scheduling in data slots would be $DS_1, DS_2, DS_3, ..., DS_N$, where N is total number of *E*-nodes requesting data slots through control packets during the control slot period. By calculating the priority, all the *E*-nodes which are listening to the control packets, are able to know the data schedule in

the data slot period. They can now sleep and wake up only to transmit the packet on the designated data slot only. Fig. 4.5 shows the structures for Data and ACK packets. In the section 4.3.3, we show how the priority levels of the packets are determined and the data slots are assigned accordingly. In addition, we also show how to schedule the nodes based on correlation as well as the received signal strength of the event source to achieve better information reliability.

4.3.2 Time synchronization

In TDMA scheme, two nodes can communicate to each other only if both the nodes agree on the transmission timing. This necessitates the time synchronization, but our protocol does not rely on any synchronization scheme. Once an event occurs, it activates many nodes to be activated for transmissions. Since these nodes wake up at the same time upon occurrence of an event, all activated nodes appear to be synchronized to each other. In addition, time synchronised nodes broadcast synchronization information in their control packets. If a node is trying to join the *superframe*, it has to receive this timing information for synchronization with the *superframe*.

4.3.3 Event-oriented, Correlation based scheduling

The CC-MAC [14] provides energy-efficient channel access by exploiting the effect of spatial correlation on achieved distortion at the sink. Based on the similar philosophy, our MAC design aims to provide the collision-free channel access and priority scheduling within the event area based on correlation and packet's sub-priorities. It guarantees reliable event reporting with minimum energy consumption and low end-to-end latency.

In this section, we first describe some observations about CC-MAC. Based on the observations, we discuss corresponding innovations in our MAC design. Next, we consider the prioritized event reporting in the event region. In the analysis presented in Section 2,

it was shown that the representative node selection based on its location from the event source rather than the random selection improves the achieved event distortion significantly. Furthermore, the priority levels should be assigned to packets according to the type of packet, queue length, packet delay, received energy level from a source and the residual energy of nodes.

• Exploiting the spatial correlation in data

Based on the discussions presented in Section 4.2, the number of representative E-nodes is less than total E-nodes in the event area for a given maximum distortion constraint, D_{max}. The selection of representative E-nodes can be determined using our proposed node selection algorithm shown in Fig. 4.2. The node selection algorithm, which is running at the sink, determines the correlation radius R_{corr} for given D_{max} and R_E. This R_{corr} value is then informed to each sensor nodes in the network during network setup.

If the *E*-nodes have the record of the number of source nodes for an event during control slot period, correlated data transmission can be eliminated by allowing transmissions from selective *E*-nodes based on the spatial correlation. According to available information from control packets, an *E*-nodes can easily determine whether a source node has the request for a data slot or not. The source nodes which are in correlation radius of R_{corr} are denoted as the correlation neighbours. To avoid the redundant transmission of data due to correlation, only one source node within R_{corr} radius is assigned the data slot for transmission while others can sleep. In our design, before transmitting event reports, every node requests for a data slot by sending a control packet in the control slot period. So potential source nodes can stop transmitting their requests if these nodes receive control packets belonging to correlation neighbours. After receiving the control packets, *E*-nodes extract the information about inter-node distance. If the distance is found to be

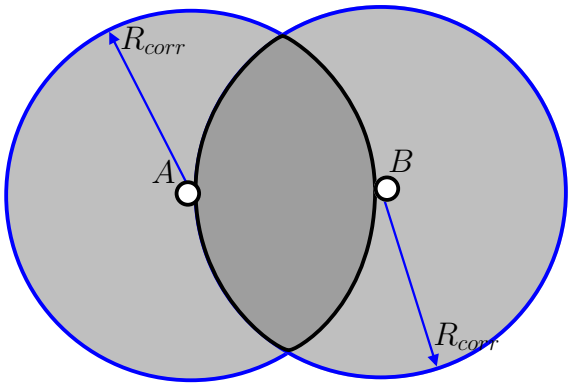


Figure 4.6: *Redundancy in data using inter-distance between nodes.*

less then R_{corr}, *E*-nodes discard their own reports from queue and may cancel out the requests. A similar scheme has been used in CC-MAC [14] for handling correlation-based event reports, but it filters out the correlation based on overhearing which causes additional power consumption.

Based on CC-MAC [14], we have a scenario to show that the correlation in data will still be present if nodes decide to send reports according to the distance between them [93]. We argue that CC-MAC fails to filter out correlation completely using overhearing. The scenario is shown in Fig. 4.6, where node B finds itself out of A's correlation region. So node B has to be a representative node. But, we can see that the correlation regions of both nodes overlaps significantly about 39%. The CC-MAC will allow B to send its own reports, because the distance between A and B is found to be greater than R_{corr}. To handle the problem, we have provision to allow transmission according to correlation more efficiently.

As shown above, the E-nodes near the boundary still send their reports. In our design, they also try to join the *superframe* using control packets. These *E*-nodes still have Redundancy in data based on distance with other representative *E*-nodes. To further remove the redundancy in data due to correlation, we schedule these partially overlapped correlated region with other correlated region based on the amount of correlation. For

simplicity, we assume that the control packets of source nodes in the buffer are denoted by $\{s_1,\ s_2,\ ...,\ s_k\}$, where k is number of packet from the source nodes. According to the distance between s_i and s_j, $\forall i, j,\ 1 \leq i, j \leq k$, an *E*-node can easily determine the control packets belonging to the *E*-nodes whose correlation regions are partially overlapping with it. Let us take the distance d_{ij}, between the node s_i and s_j, where $i = 1, j = 2, 3, 4, ..., k$, and $i \neq j$. Following two steps are executed by *E*-node s_i. If $d_{(i,j)} \geq R_{corr}$, but $< 2R_{corr}$, then s_j node has partially overlapped correlation area. These nodes are then scheduled according to increasing order of the amount by $INTC(d_{ij})/\pi(R_{corr})^2$, where $INTC(d_{ij})$ denotes the intersection area of two circles of the radius R_{corr} centered at two points located at s_i and s_j. $INTC(d_{ij})$ can be determined in advance using $4\int_{d_{ij}/2}^{R_{corr}} \sqrt{(R_{corr})^2 - x^2}dx$ during network setup.

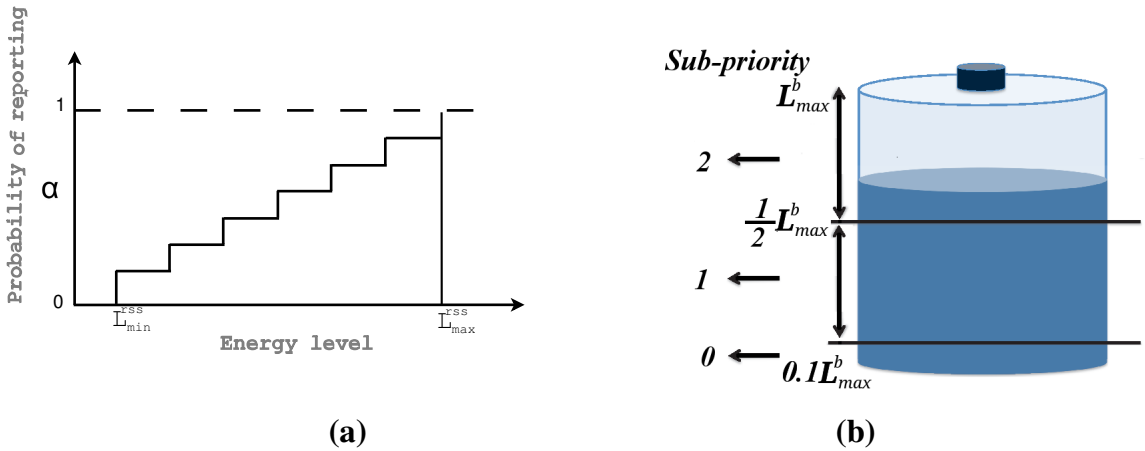
In the next section, we discuss the strategy to assign the data slots based on control packet requests. As per amount of correlation, two tables (say Table-A and Table-B) are maintained by each *E*-node by extracting information from received control packets. Table-A contains the records of those nodes which are representative nodes among its correlation neighbours without any overlap with the other ones. Table-B will have record of representative nodes whose correlation regions are partially overlapping with others. Using priority calculation, the data slot assignment is first performed on the representative nodes belonging to the records in Table-A. Remaining data slots are assigned to the representative nodes belonging to Table-B according to increasing order of the amount of correlation as discussed above. *E*-nodes calculate priority index (PI) depending on packet type, queue length, packet delay, received energy level of event source and residual energy of nodes as explained in next section. Priority in data slot assignment will be in decreasing order of Net Priority Index (NPI) which is weighted sum of current PI and last calculated PI (i.e., in last *superframe*) for *E*-nodes requesting data slots in current *superframe*. Since, each node will have same inputs for calculations, they all will compute

the same schedule and find their place in it.

• Priority calculator and Packet Schedular

A packet scheduler used in our design is function of MAC-layer queue only, and it can be implemented easily in reality. Based on the practical assumptions, wireless scheduling can be performed by MAC layer and the scheduler does not have to access transport layer information [94]. In this section, we first define mapping rules according to received energy level of the event source and residual energy of a node. We then introduce the rules of sub-priority indexing for the type of data packet, queue length and packet delay. These rules are used to calculate the Priority Index (PI) of a node requesting a data slot after getting the slot for control packet in the *superframe*.

We assume a simple detection model in which an event is detectable if received signal strength is greater than a threshold, known as the sensitivity. At the beginning of the event reporting, all *E*-nodes sample the energy level of detected event signal. Each *E*-node calculates a number α ranging from 0 to 1, depending on how strong is the received signal strength (RSS) compared to the threshold. Higher, the received energy level than the threshold, higher will be the number. We propose a stair case function $f(p)$ shown in Fig. 4.7(a) to map the received energy level of a node to α. This mapping gives the selection order of the nodes for event signal reconstruction. As seen from Fig. 4.7(a), when energy level is higher than L_{max}^{rss}, α is equal to 1 and for energy level lower than L_{min}^{rss}, α is zero. This criteria ensures that if *E*-nodes are close enough to the event source, they have more chance to send their reports according to received energy level of event source. Furthermore, it is mandatory to balance the energy consumption among *E*-nodes. To maintain the balanced energy consumption inside the event area, we also map the remaining battery power of an *E*-node to sub-priorities (denoted by (BP) Battery Power) as show n in Fig. 4.7(b) and Table 4.3. Here, L_{max}^{b} is maximum battery power



(a) **(b)**

Figure 4.7: *(a) Mapping of RSS of emitted event source with α, (b) Mapping of Battery power into sub-priorities for a node.*

level.

As discussed earlier, two type of packets are considered i.e., newly generated reporting packet and route-thru packets. The route-thru packets are more important than newly generated reporting packets. We use Source packet (SP) field to differentiate these type of packet. One bit field SP in the control packet determines whether the data packet to be transmitted is originating at this node or forwarded by it. When an *E*-node needs to transmit a packet generated by it, the SP field is set to 1 in control packet. Otherwise, SP field is set to 0 to treat it as route-thru packet.

The priority to packets can be given according to type of packet, queue length, packet delay, sensed energy level of the event and remaining battery power level at transmitting node. Table 4.3 shows the sub-priority index values. Priority index value (PI) for each data packet is computed by each node and data slot sequences are assigned accordingly. Based on above mentioned rules, PI is calculated as

$$PI = 2^{\alpha}(PT + QL + PD)BP. \tag{4.1}$$

As seen from Eq. (4.1), PI calculation takes care of node's remaining battery power

Table 4.3: *Sub-priority index values for BP, PT, QL and PD*

Packet Type	Sub-priority (PT)
Reporting Packet with SP =1	1
Reporting Packet with SP =0	2

(4a)

Battery Power Level	Sub-priority (BP)
$< 0.1L^b_{max}$	0
Range $[\ 0.1L^b_{max}, \frac{1}{2}L^b_{max}\ ]$	1
Range $[\ \frac{1}{2}L^b_{max}, L^b_{max}\ ]$	2

(4b)

Queue Length	Sub-priority (QL)
1 – 5 Packets	0
5 – 10 Packets	1
>10 Packets	2

(4c)

Packet Delay	Sub-priority (PD)
<5 *superframes*	0
5 – 10 *superframes*	1
>10 *superframes*	2

(4d)

(i.e. BP) and sensed energy level (i.e. α) of event source. So *E*-node with high received energy level and more residual energy have larger PI irrespective of PT (Packet Type), QL (Queue Length) and PD (Packet Delay). The net priority index (NPI) of each packet will be sum of current PI and calculated PI in last *superframe*. The last calculated PI value is sent to other nodes using control packet. If a *E*-node does not have last calculated PI, the default value of PI is assumed to be 3 in last *superframe*.

After getting PI values from all *E*-nodes in the control packet broadcast, an *E*-node can find current NPI for each *E*-node. It then arranges the NPI in decreasing order. The data slot sequence is allocated to data packets according to the decreasing order of NPI value. The *E*-node with highest NPI is assigned with first data slot DS_1 and subsequent data slots are assigned as per decreasing NPI value. The nodes with same NPI will have data slot order according to the order of received control packet in control slot period. Since each *E*-node that has data to send, joins the *superframe* by broadcasting a control packet with all the information, all other nodes have complete record for applying the above rules in NPI calculations. Thus, the calculated NPI by an *E*-node will be same for

all the others also.

4.4 Working of Event-MAC

In this section, the operation of Event-MAC is discussed with an illustrative example. Since multiple nodes detect an event at a time, these *E*-nodes appear to be synchronized to each other. Every *E*-nodes first contends to capture the RCS for joining the *superframe*. Initially, first contention winner of RCS creates the *superframe*. Other *E*-nodes then join the *superframe* one by one and the *superframe* will grow. The working of Event-MAC consists of the following four steps.

Step 1 – Entry of new reporting nodes: When a new node has data to send into the network where the communication is already in progress through the running *superframe*, it has to send its control packet over RCS. A node trying to join the *superframe* first examines the running *superframe* to know the time of RCS. A new node can synchronise with the *superframe* by listening the beacons sent from others in the control slots. Then it contend for the channel for transmitting control slot in RCS using CSMA/CA. Upon success of RCS, it occupies subsequent control slot, CS when the *superframe* repeats next time. Similarly, other potential sender nodes occupy the control slots (i.e., $CS_i, CS_{i+1}, ..$, where i is a slot number).

Step 2 – Leaving the superframe: When a node which has no packet to send and its buffer is empty, it can leave the running *superframe* by not sending any control packet in the assigned CS. All receiving nodes find corresponding CS of this node to be vacant (i.e., they do not receive the control packet). Hence they assume that the node has withdrawn. So they update their control slot positions by shifting places when next *superframe* starts.

Step 3 – Transmission over data slots: After all the nodes have sent their control packets and thus received the control packets from all the other nodes, they calculate NPI to determine the order in which these nodes would send the data packets during

data slot period. Each node in the *superframe* sends its data packet over the channel in its own data slot. After the successful reception of a data packet, the receiver sends an acknowledgement packet to complete the communication for a single data slot. After complection of transmissions of data packets during the data slot periods, the *superframe* again continues to repeat over the channel. The nodes follow the sequence which is the continuation of the previous control packet sequence. These control packets are then followed by the data packets as stated earlier.

Step 4 – Transmission Power control: The transmission power of the control packet is sent along with the control packets. The nodes can sense the power of the received control packet. With the information about transmitted power and sensed received power, the nodes can compute the distance of the sender. Thus the nodes can minimise their transmission power while ensuring that the packet should reach the destination with required SNR. The advantage is the less power consumption by the nodes. If there are large number of reporting nodes present in the event area and all of them are not taking active part in communication, the nodes can go to sleep mode. The reporting nodes which have no packets to transmit can temporarily leave or pretend to leave the *superframe*. This is done by stopping the transmission of control packets. In the sleep mode, a reporting node does not transmit any packet but it receives all the control packets from other active reporting nodes in the event area. As discussed earlier, the nodes waits for a time which is equal to duration of control slot period, So they may know about the packets destined to it.

4.4.1 Illustrative Example

Consider a simple WSN network as shown in Fig. 4.8. Nine sensor nodes detect the occurrence of the event. They are denoted as $n_1,\ n_2,\ n_3,\ $..., and n_9. Consider that only 3 nodes are sufficient to report the event information for the given distortion constraint.

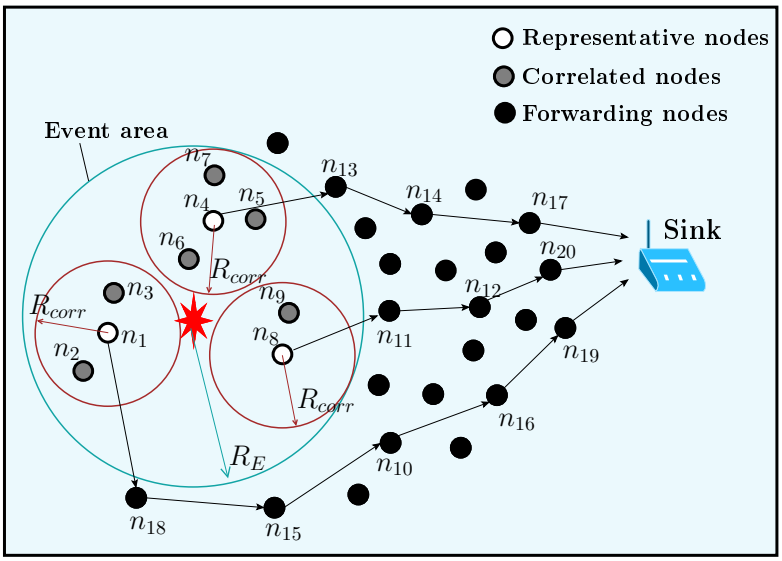


Figure 4.8: *Simple WSN example.*

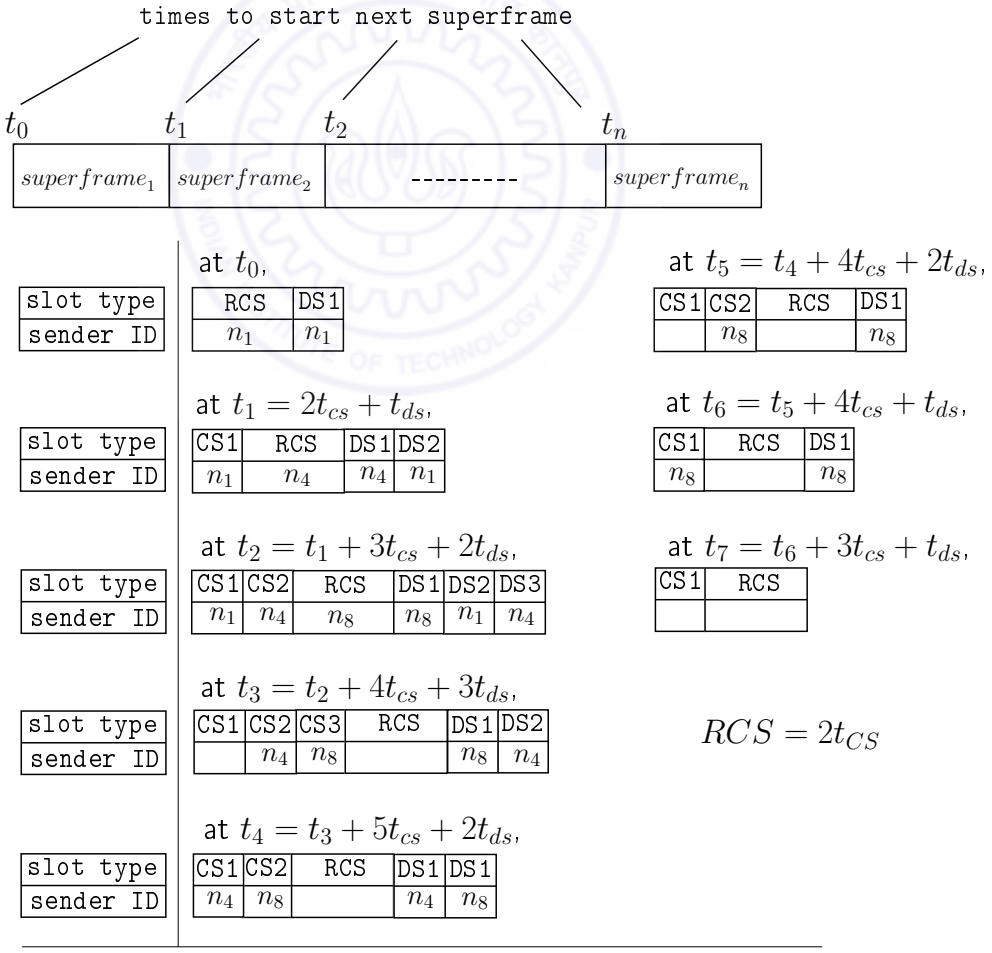


Figure 4.9: *Illustration of the running superframes by nodes using example in Fig. 4.8*

It means these correlation regions of R_{corr} radius are to be formed on the given event radius, R_E. The stepwise operation is as given below.

- Event-MAC operation begins by starting the *superframe* at time t_0. At t_0, all 9 nodes initiate by contending the channel access for their control packets using CSMA. Suppose n_1 is the contention winner to broadcast a control packet into RCS. This packet is received by the following nodes, $n_2,\ n_3,\ ...,\ n_9$, and n_{18}.

- Since n_1 creates the *superframe*, it occupies DS_1 for its data packet. Now next *superframe* will repeat at time, t_1 which would be $2t_{cs} + t_{ds}$.

- After examining the control packet from n_1 node, the nodes n_2 and n_3 discard their own generated packets and go into sleep mode. If these nodes have route-thru packet to send, then they would try to get RCS during next *superframe*.

- n_1 and remaining nodes outside of R_{corr} region of n_1 (i.e., $n_4,\ n_5,\ ...,\ n_9,\ n_{18}$ store the identity of n_1 in their own generated reporting lists.

- n_1 sends its data packet to the sink through following route, $[n_1,\ n_{18},\ n_{15},\ n_{10},\ n_{16},\ n_{19},$ Outside the event area, nodes perform LPL-based CSMA channel access. So after receiving the control packet of n_1, n_{18} as a destination node wakes up to receive data packet for passing it to other node using normal CSMA channel access.

- At t_1, next *superframe* starts. the control packet of n_1 moves to CS_1 and RCS is kept vacant for new potential sender in this *superframe*.

- Suppose that n_4 captures RCS by successfully sending a control packet. Now, nodes n_1 and n_4 get the slots in control slot period, the data slots are assigned according to priority which is determined by NPI calculation for each nodes. Suppose node n_4 has first priority then n_1. So DS_1 and DS_2 will be given to n_4 and n_1 respectively.

- The control packet of n_4 is received by all other nodes in the event area. By examining the packet, if nodes find themselves within R_{corr} region, they go to sleep (i.e.,$n_5,\ n_6,\ n_7$). The n_4 sends its data packet to the sink through following route, $[n_4,\ n_{13},\ n_{14},\ n_{17},\ sinknode]$.

- Next *superframe* continues to repeats on time t_2 which would be $t_1 + t_{cs} + 2t_{cs} + 2t_{ds}$. Similarly, all active reporting nodes join the *superframe* as illustrated in Fig. 4.9.

4.5 Performance Evaluation

This section first analyzes the performance of the proposed Event-MAC protocol. Thereafter the performance comparison with the correlation-based channel access schemes, CC-MAC [14], SC-MAC [15] and simple wireless MAC scheme i.e. IEEE 802.11 [46] has been done.

Our simulation results used the models and metrics in NS–2 network simulator [95]. The default network parameters and Event-MAC design in NS–2 are discussed in Appendix A. The performance metrics used in our simulations are given in Chapter 1.

Our analysis considers the Random Correlated Event Traffic (RCET) model [85] to simulate bursty traffic triggered by spatially correlated nodes within an event area. This is commonly used model for event-driven applications such as detection, tracking etc. RCET picks a random point at (x, y) location in 2-dimensional plane. If every node has r-radius disk sensor (e.g., omni-directional sensors), the nodes within circular area of sensing range r centered at (x, y) are allowed to generate packets as reports for this event. These reports traverse towards the sink through multi-hop communication. We simulated 50 random networks by placing 100 nodes randomly over the $1000 \times 1000\ m^2$ area. For each network, the sink node is placed at one of the corner of the area and

RCET model is used to generate multiple events, each within 50 seconds interval. The simulations are conducted using these 50 random networks and the results are produced by averaging them to minimize the statistical noise.

Since EWSNs should provide fast and accurate event notification of the PoI, we need a fast routing protocol. The routing protocols for sensor networks are mainly classified into three types: Pro-active routing, Reactive routing and Geographical routing. We adopted geographical routing [96] in our simulations. Geographical routing does not require prior route construction and maintenance. Thus, no additional overhead is needed for dissemination of the topology information. Hence, it offers low-overhead and low-latency which is best suited for event-driven applications. Actual geographical positions of nodes are used to forward the packets from one location to another. In other words, packets are transmitted/forwarded to the location of destination node without use of network address or routing table. The location information can be obtained by GPS (Global Positioning System) or localization algorithms [97, 98]. Thus, the routing protocol traffic is not included in our simulations.

4.5.1 Event-MAC performance analysis

Property 1. *For $\forall i, j$, if priority l_i is larger than l_j, then there exists $t_i < t_j$.*

Proof. The theorem shows that the nodes with higher NPI have first chance to schedule their data. In Event-MAC, a node with higher NPI gets first slot if its request arrives during control slot period. According to definition-1, if higher NPI corresponds to higher priority l_i, $t_i < t_j$. It means *E*-node with high received energy level and more residual energy will get first chance to transmit data than the other nodes. □

Property 2. *Event-MAC is a soft real-time MAC protocol.*

Proof. Event-MAC provides collision-free channel access and priority based scheduling

for event reporting. Upon occurrence of an event, activated nodes perform Event-MAC and start joining the *superframe*. Initially, first contention winner node creates the *superframe*. Each reporting node joins the *superframe* by sending a control packet on *superframe* as a request for data slot. The reports are not transmitted immediately after sending the request during control slot period. There are minimum and maximum transmission delay for m reports.

Once the contention is resolved, the minimum transmission delay, D_{min} for a report will be when first data slot is assigned in next *superframe* i.e.,

$$D_{min} = 2t_{cs} + mt_{cs} + t_{ds}. \tag{4.2}$$

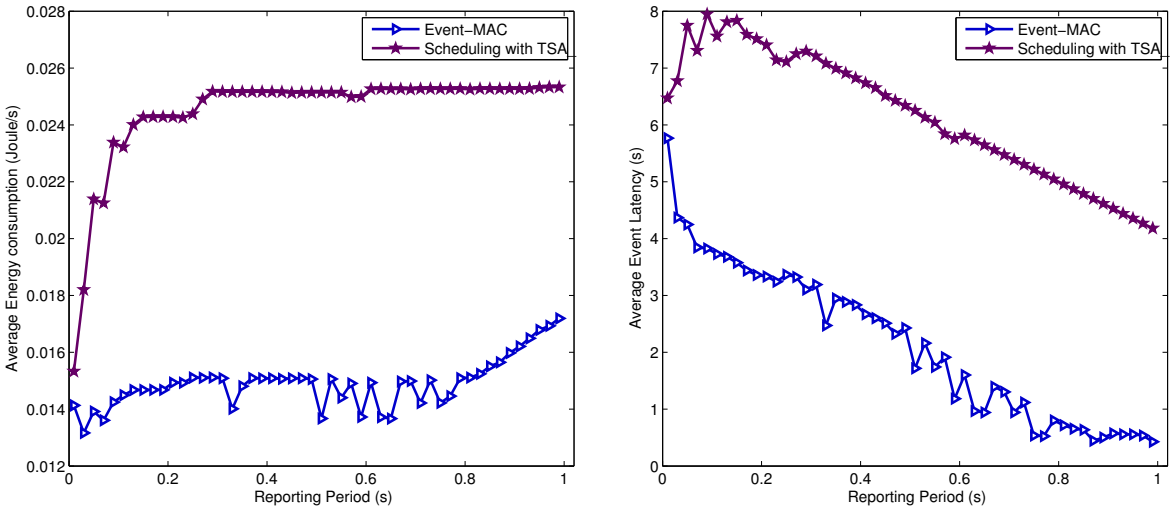
Similarly, the maximum transmission delay D_{max} after the contention win for a report will be when last data slot is assigned i.e.,

$$D_{max} = 2t_{cs} + mt_{cs} + mt_{ds}. \tag{4.3}$$

Based on above analysis, Event-MAC offers bounded time delay and hence supports real-time event reporting. Therefore, it is a soft real-time MAC for EWSNs. □

To get more insight into the protocol operation, We consider the two cases in our simulations for Event-MAC performance. In case-1, the Event-MAC is compared with traditional slot assignment scheme (TSA) as used in conventional TDMA schemes based on graph coloring techniques. In case-2, we evaluate the performance of the protocol by varying load of traffic by changing the reporting interval.

Fig. 4.10, shows the results of case-1 where we measure average energy consumption and event latency in the network to report an event using different reporting intervals for the node deployment scenario given in Section 4.5 (default parameters given in Table A.1). In this case, it can be seen from Fig. 4.10 that different slot assignment strategies



(a) **(b)**

Figure 4.10: *Comparison of Event-MAC with traditional slot assignment scheme (TSA). (a) Average energy consumed by network to report an event, (b) average event latency.*

will have different performance. Event-MAC performance is better in terms of energy consumption rate and event latency when traffic is variable in the event area. Since no slot is left vacant in Event-MAC, significant amount of energy saving is possible compared to TSA scheme. On an average, 70 % energy is conserved for the given traffic range as shown in Fig. 4.10(a). Further Event-MAC allocates time slots dynamically and hence has lesser delay (Fig. 4.10(b)). One can observe 45% reduction in the delay for reporting interval of $0.05s$ and 89.7% for reporting interval of $1s$. Hence by dynamically assigning data slots, the event-MAC ensures full channel utilization and hence supports real-time reporting with much less event latency.

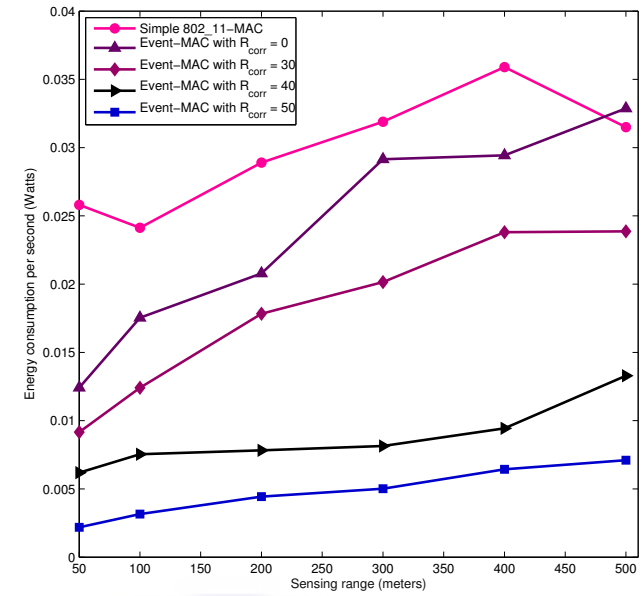
Table 4.4: *Average number of potential senders for each event under different sensing ranges*

Event sensing range (m)	50	100	200	300	400	500
Reporting nodes (N)	6	18	24	38	42	56

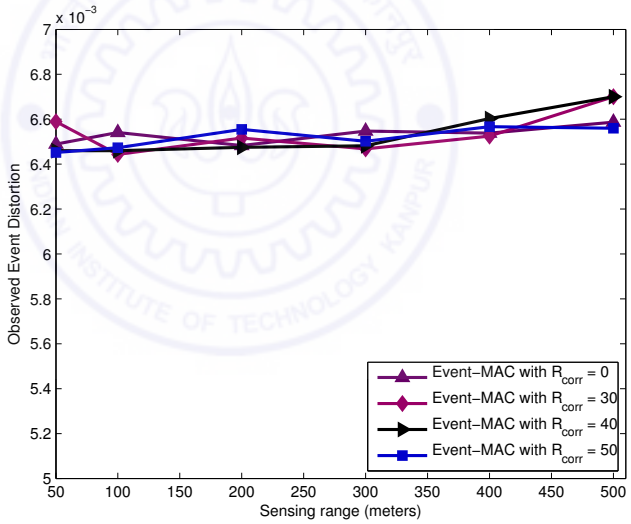
For the case-2, we adjust the event sensing range r from 50 to 250 meters using

same network setup. So the number of nodes sending the event reports will change more creating bursty traffic. Table 4.4 shows the average value of potential reporting nodes for different values of sensing range r. Multiple events are chosen randomly with maximum interval of 50 seconds during the simulation time. We first measure the total energy consumed per unit time by the network to transmit fixed number of reports over a given simulation time i.e. packet transmission rate. This gives energy consumption rate as jouls per second or Watts. Using Event-MAC, slower traffic will lead to lower rates of energy consumption. Further, the average distortion achieved at the sink is measured from the received reports of representative nodes.

Fig 4.11 shows the results for case-2 for four different correlation radius values i.e., $R_{corr} = 0,\ 30m,\ 40m$, and $50m$. It can be seen from Fig 4.11(a) that the rate of energy consumption increases as the reporting nodes are increased. This is expected since more traffic is generated when event area spreads due to increased sensing range. For increasing values of R_{corr}, energy consumption is less as fewer representative nodes are chosen to send the event information to the sink. As an example, energy saving of 78.9% is possible by increasing the value of R_{corr} from $30m$ to $50m$ for $500m$ sensing range. It can also be clearly seen that Event-MAC (schedule based) is more energy-efficient than contention based MAC. In Fig 4.11(b), effect of correlation radius, R_{corr} on measured distortion is shown with different sensing ranges. As seen from Fig 4.11(b), average distortion is relatively constant as the sensing range increases which causes participation of more reporting nodes in the event reporting. The measured distortion does not show the impact of correlation radius. This demonstrates that results using Event-MAC are consistent with theoretical results presented in the section 3.4 of Chapter 3. Because Event-MAC provides priority based node selection similar to ordered selection case in Fig. 3.5. So the nodes closer to the event source would have higher priority to schedule data than the other nodes. Thus, the reports from closest nodes around the event source



(a)

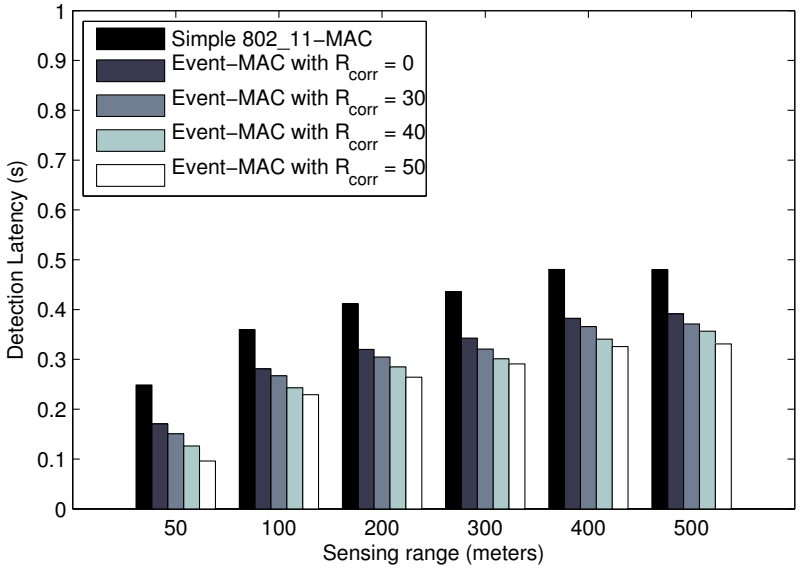


(b)

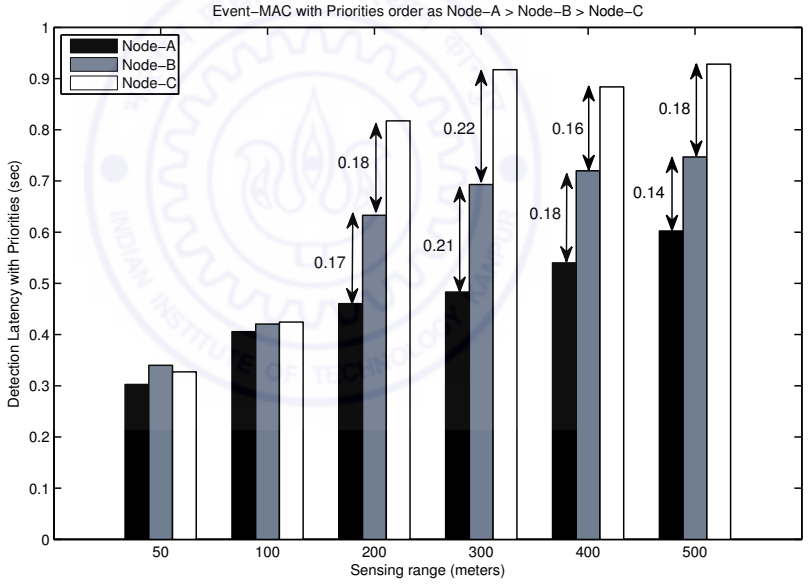
Figure 4.11: *Performance of Event-MAC on energy consumption and achieved distortion. (a) Average energy consumption per unit of time, (b) Achieved event distortion at the sink.*

would be scheduled first to reach the sink. This is the reason that the measured distortion stays constant to a minimum value.

In event-driven applications, network traffic load may be low due to infrequent oc-



(a)

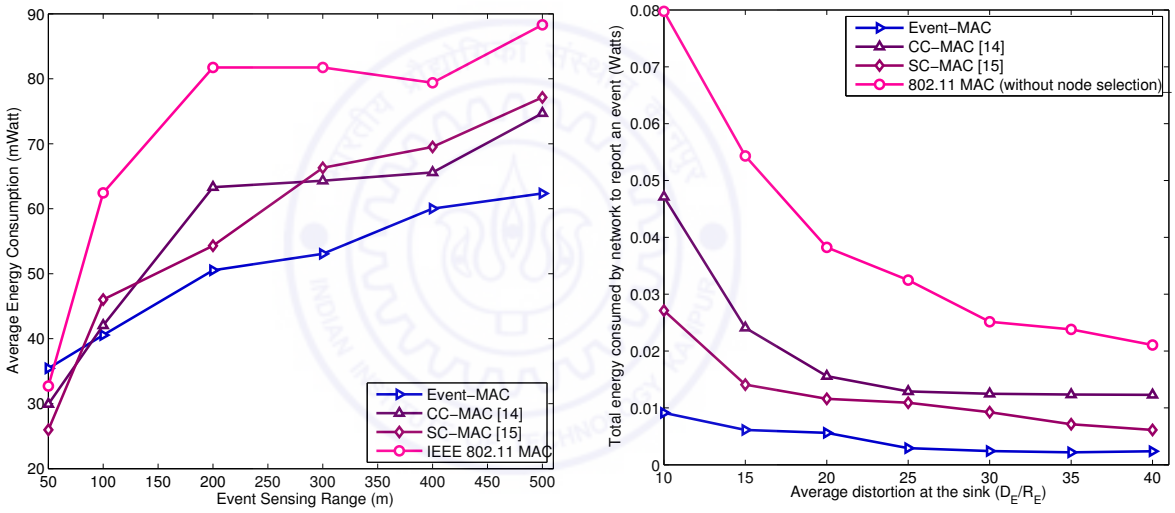


(b)

Figure 4.12: *Event-MAC performance on detection latency. (a) with different correlation radius R_{corr}, (b) with different priority.*

currence of events, but critical data such as event notification packets are of interest. So we measure the delay taken by first packet to reach the sink (i.e., detection latency). As seen from Fig. 4.12(a), this metric shows that Event-MAC has smaller detection latency

which is desirable. To get more insight into the protocol operation, Fig. 4.12(b) shows the performance of Event-MAC with priority. Three nodes (*Node-A*, *Node-B* and *Node-C*) are arranged in such a way that *Node-A* would have higher priority than other two nodes as it is much closer to event source. Similarly, *Node-B* has higher priority then *Node-C*. Thus, the priority would be in the order: *Node-A*>*Node-B*>*Node-C*. In Fig. 4.12(b), the node with higher priority shows much shorter delays than the nodes with lower priority. The difference in delays is shown for different sensing range. This demonstrates that the Event-MAC provides tolerable detection latency. As per theorem-2 also, we can conclude that Event-MAC supports real-time event reporting with higher success rate.



(a) **(b)**

Figure 4.13: *Comparison of various MAC protocols for energy-efficiency. Average energy consumed by network (a) with spread of event area, (b) with achieved distortion at the sink.*

4.5.2 Comparative analysis

In this subsection, we compare the simulation results for Event-MAC with CC-MAC [14], SC-MAC [15] and IEEE 802.11 [46] using same node deployment scenario as given in the previous subsection. First a comparative study of energy-efficiency for each

of the five protocols has been presented. Then, the performance using the two metrics success rate and coverage ratio has been given.

As the main advantage of proposed Event-MAC, we show the gain in energy conservation and the measurement distortion at the sink with the number of reporting nodes (i.e., sensing range) in Fig. 4.13. It can be seen that significant energy is conserved by applying Event-MAC compared to CC-MAC and SC-MAC schemes. This is because the Event-MAC relieves collisions and over-hearing which is also evident in Fig. 4.11(a) even with different correlation radii. The CC-MAC and SC-MAC have poor erformance due to higher collisions and over-hearing. As an example, 16.52% energy can be conserved by applying Event-MAC as compared with CC-MAC for larger number of reporting nodes in occurrence (i.e., for $500m$ sensing range). Similarly, as compared with SC-MAC, 19.16% energy is conserved by Event-MAC for $500m$ sensing range. Fig. 4.13(b) shows the performance with measured distortion at the sink. We can see that energy conservation using Event-MAC is even higher for achieving a larger range of distortion at the sink. The reason is efficient decision in node selection and node scheduling in the event area. Thus, minimum number of reporting nodes which are sufficient to achieve required distortion at the sink has been activated by Event-MAC.

Fig. 4.14 shows the performance in terms of success rate and coverage. The contention-based protocols CC-MAC and SC-MAC show poor lower performance, especially when event area becomes larger for both the metrics i.e., the success rate as well as the coverage as seen in Fig. 4.14(a) and Fig. 4.14(b). This is because the probability of packet drop is higher due to collisions and contention. Thus, fewer reports reach the sink node as expected. In Fig. 4.14(b), coverage performance is poor for CC-MAC and SC-MAC. One can observe that the Event-MAC has better performance in terms of both success rate and coverage. Event-MAC provides high coverage (i.e., low redundancy in data) and high success rate. Thus, our node selection scheme performs much better to further

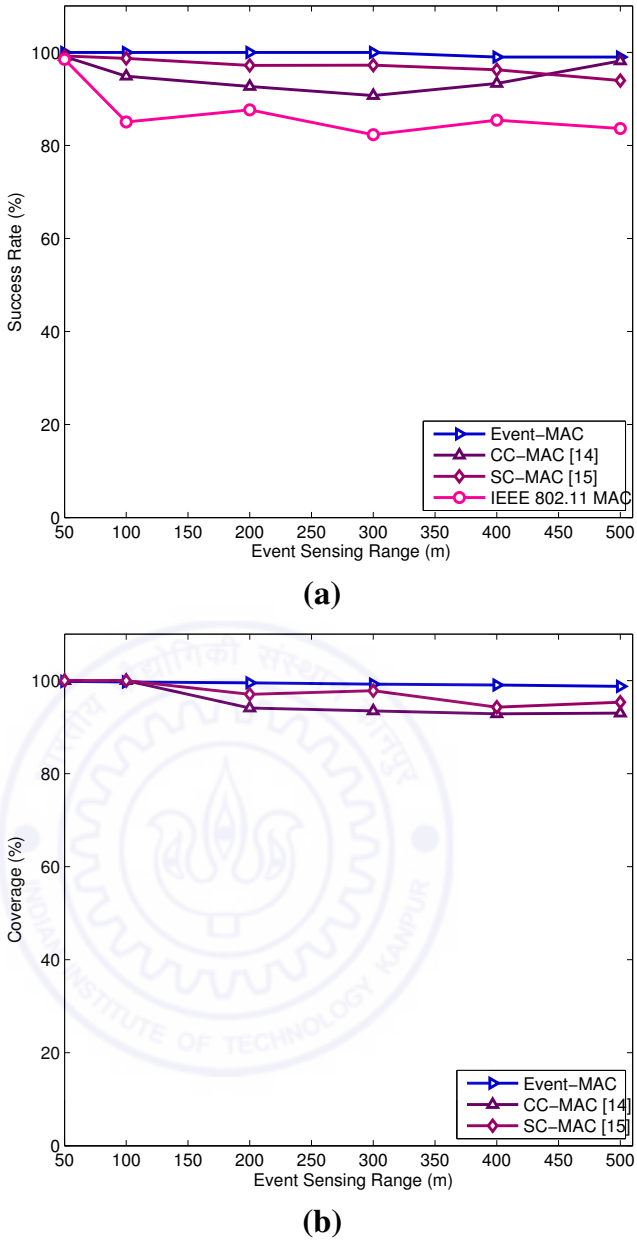


Figure 4.14: *Comparison of various MAC protocols. (a) The success rate, (b) Coverage of the event area by representative nodes.*

minimize the correlated information present in WSNs with high coverage. Furthermore, efficient dynamic data scheduling used in Event-MAC achieves high success rate.

4.6 Summary

Based on the new sensing framework, correlation based hybrid TDMA/CSMA protocol called Event-MAC is proposed in which the nodes switch over to TDMA scheduling mode only when an event is detected. Otherwise, simple low power CSMA/CA channel access operates network-wide. Furthermore, correlation based efficient decision on node selection and prioritized node scheduling make the Event-MAC a novel correlation-based MAC for EWSNs. Proposed protocol is compared with existing correlation-based MAC protocols to examine its performance. The performance improvement in term of energy consumption, distortion, success rate, detection latency, and coverage were investigated. It has been demonstrated that our proposed Event-MAC offers minimum distortion, high energy-saving, acceptable latency, low correlation (or higher coverage) and high success rate. In addition, proposed Event-MAC can adapt to different event sensing ranges i.e., large and small event areas.

Chapter 5

Correlation-based Energy-efficient data collection Scheme for WSN

In this chapter, spatial correlation-based clustering technique is proposed. Based on the framework developed in Chapter 3, some observations and innovations are discussed to develop a correlation-based energy-efficient data collection method, termed PS-NLEACH for WSNs. PS-NLEACH is an extension of the cluster-based routing method called N-LEACH which provides dynamic clustering and data collection in rounds for WSNs. PS-NLEACH exploits spatial correlation to minimize energy consumption and prevent superfluous correlated data collection, without degrading the information reliability. Section 5.1 gives the motivation for this work. In Section 2.3, related existing work on routing protocols for WSNs had already been reviewed. In Sections 5.2, a novel correlation-based PS-NLEACH is proposed. The simulation results are discussed in section 5.3. Finally, Section 5.4 summarizes the chapter.

5.1 Motivation

Since WSNs are usually operated unattended in harsh environments, monitoring activities cannot be managed efficiently without human intervention. In such environments, sensor nodes are randomly deployed in large numbers (that is, hundreds or thousands of nodes) to ensure fine-grained monitoring. The routing and data-gathering protocols should be energy-aware in order to maximize the life-span of the network. In WSNs, the grouping of nodes into clusters has been widely adopted along with routing and data-gathering protocols to achieve high energy efficiency and to maximum the life-span of the network [1, 72, 25]. These hierarchical data-gathering protocols provide data fusion and aggregation through cluster-based organization of the nodes. Thus they minimize the energy consumption of nodes to maintain long-term operation of WSN.

Current literature on cluster-based routing protocols [17, 72, 25, 68, 99, 100, 101] for WSNs focuses mainly on energy efficiency, system scalability and life-span of the network. There is very little research effort on the design of networking protocols that are capable of exploiting the spatial correlation directly. The correlation in data can be considered in the design process of clustering protocols. The cluster formation procedure should generate correlation-based clusters according to the given constraint of information reliability, so that the required data resolution can be preserved by the cluster head (CH) nodes while using minimum number of member nodes. Thus, energy optimization in the process of clustering is a fundamental design challenge for many sensor network applications.

Motivated by the results and discussions in the section 3.4, a correlation-based clustering technique is proposed to reduce redundant information during data collection for hierarchical routing protocols. As discussed in Chapter 3, exploiting the spatial correlation at the network layer can lead to significant performance improvement of routing protocols. Intuitively, data collection from spatially uncorrelated nodes provides more

useful information than the collection of data from close-by correlated nodes. Hence, a smaller number of nodes should be sufficient enough to gather data rather then every node doing so in the network, and this would not degrade information reliability.

In this chapter, the alternative sensing model discussed in earlier chapters has been used to capture details of spatial correlation to enable the design of an energy-efficient routing protocol. Based on the results obtained using the framework given in Chapter 3, a correlation-based Pre-selected N-LEACH algorithm (PS-NLEACH) follows.

5.2 Correlation-based Pre-selected N-LEACH algorithm

LEACH [25] is one of the best cluster-based routing protocols for WSNs. In LEACH, cluster heads (CHs) are selected with probability p among all the sensor nodes of the network in each round. After the CHs are selected according to p, each sensor node which is not a CH in the current round, finds an closest CH to join as a member. The CHs collect data from their members and aggregate it. Finally, the CHs transmit the aggregated data to the sink node directly. Due to the random selection of CHs in LEACH, the number of CHs is different in each round. The varying numbers of CHs in each round will also have dissimilar numbers of nodes in every cluster. As a result, there will be uneven energy consumption in each round. N-LEACH [102] is one of the solutions to this problem. In N-LEACH, balanced energy consumption is provided by selecting a cluster head based on the number of supported nodes in the previous round. N-LEACH achieves a consumption of approximately 1 epoch of energy on an average in each round. However, N-LEACH does not provide any scheme which exploits the spatial correlation during data collection.

Based on the framework developed in section 3.3 of Chapter 3, when a greedy correlated clustering (GCC) algorithm is used, sets of correlated clusters are formed and all nodes in a cluster are considered highly correlated. It is shown that correlation regions

(also called correlated clusters) can be formed by fine-tuning the correlation parameters ϑ and ξ. Table 3.2 in Chapter 3 indicates the results of R_{corr} for different values of ϑ and ξ. A CH can re-adjust the correlated cluster members to an optimum level very quickly by adjusting ϑ and ξ and re-executing the GCC algorithm developed in Chapter 3.

The Pre-selected N-LEACH combines N-LEACH with the Greedy Correlated Clustering algorithm [90]. In [90], the spatial correlation function is derived and correlated regions are formed using that function. A GCC algorithm is proposed to divide the sensor nodes into correlated regions of radius R_{corr}. Following [90], PS-NLEACH protocol divides sensor nodes into several correlated regions of Rcorr, where all the nodes in a region are considered as redundant nodes. Only one sensor node located at the center of the R_{corr} circular region can be selected as a representative of the region. Since these representative nodes are selected before dynamic routing in each round, the proposed protocol is named the Pre-selected N-LEACH. The dynamic routing N-LEACH algorithm is then executed by these representative nodes only to perform the actual operation of routing, while other non-representative nodes go into sleep-mode during the current round.

Based on the framework developed in the section 3.3, the relationship between correlation and the sensor nodes is derived in terms of the parameters ϑ, ξ and R_{corr}. It is restated as

$$\xi = \frac{\cos^{-1}\left(\frac{R_{corr}}{\vartheta}\right)}{\pi} - \frac{R_{corr}}{\pi\vartheta^2}.\sqrt{(\vartheta^2 - R_{corr}^2)}. \tag{5.1}$$

The following are the concluding remarks using equation (5.1) based on the parameters ϑ, ξ and R_{corr}.

- If any node n_k lies within R_{corr} from node n_i, both n_k and n_i are found to be strongly correlated to each other. Otherwise, the correlation between them is considered weak.

- A densely deployed WSN can be partitioned into disjoint circular correlated clusters of size R_{corr} for the given values of ϑ and ξ. Only a single node from this correlated cluster can be selected as a representative of the region covered by these nodes.

- A circular region of radius R_{corr} depends on the control parameter ϑ and ξ. The value of ξ controls the number of nodes which forms a R_{corr} - sized correlation region. The larger value of ξ implies stronger correlation. The value of ϑ controls the extent of the area from which event information is to be gathered.

Based on the above observations, the parameter ξ can be a function of data resolution in sensor measurements or in the achieved information reliability at the sink. To get more insight into using the above parameters, simulations were conducted over a $100 \times 100\ m^2$ area where 200 nodes are distributed randomly by varying the values of ϑ and ξ. Fig 5.1 and Fig 5.2 show the simulation results where red and green color nodes indicate representative and non-representative nodes respectively. The Voronoi partitioning of the network using different R_{corr} (i.e., $9\ m, 14\ m, 20\ m$) is shown in Fig 5.1. It can be seen that the correlated sub-regions spread out when the Rcorr value increases from $R_{corr} = 9\ m$ to $R_{corr} = 20\ m$. So the number of representative nodes decrease with an increase of R_{corr}. Hence a reduced number of nodes can represent the whole network leading to significant energy saving for a given large value of R_{corr}.

For different values of correlation threshold ξ, Fig 5.2 shows Voronoi partitioning into correlated sub-regions. It can be observed from the figure that as the correlation threshold ξ increases, the number of representative nodes increase due to the shrinking sub-regions. Inversely, with a decreased value of ξ, the correlated sub-regions spread out, thus leading to a lesser amount of energy consumption due to the lesser number of representative nodes. A larger value of ξ means shrinking of the size of correlated regions and, in turn, a larger number of representative nodes, since there is one for each region.

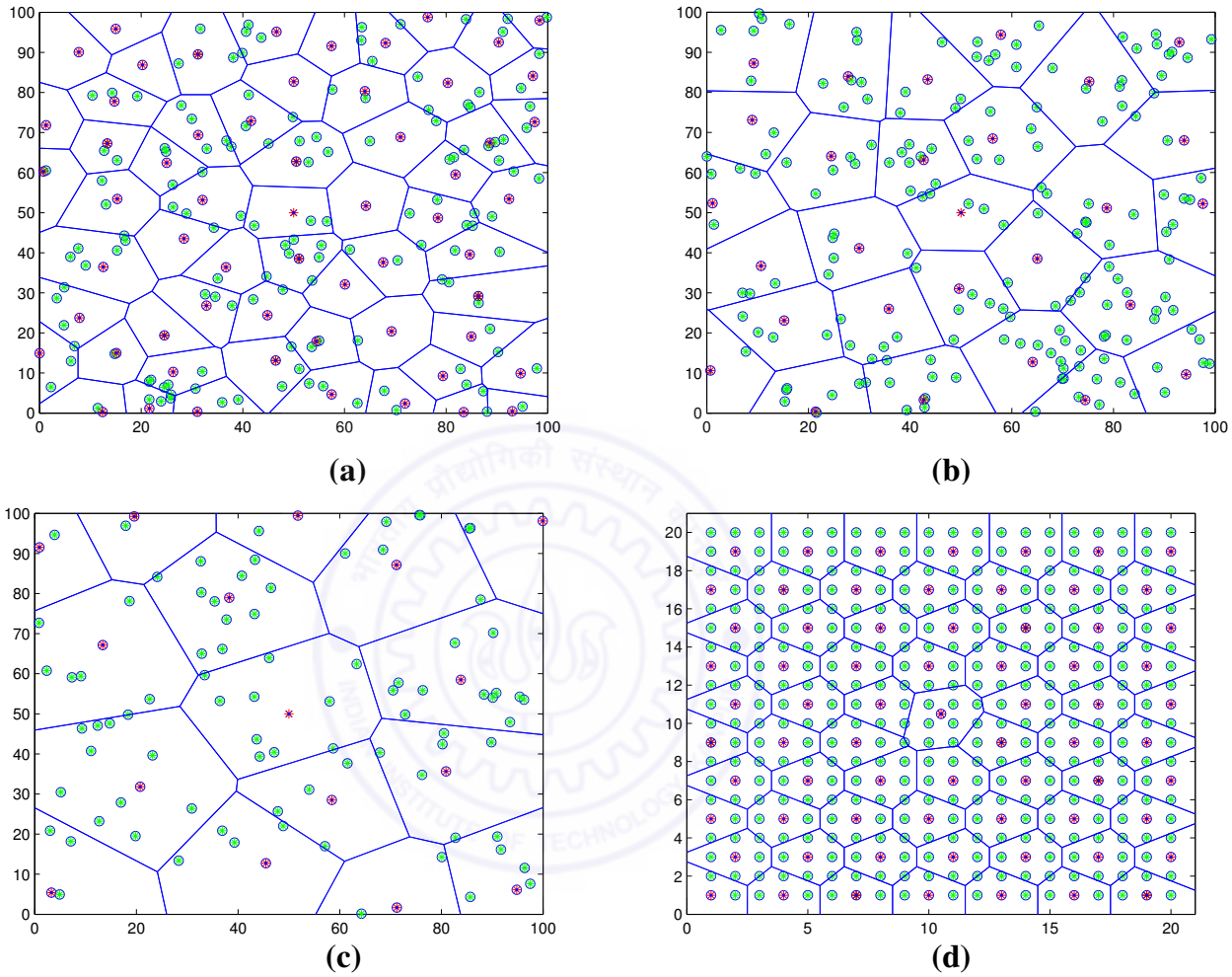


Figure 5.1: *Region divided into correlated sub-regions by Voronoi partitioning for different R_{corr}, given $\xi = 0.21$, where green sensors and red sensors show correlated nodes and representative nodes in its sub-region respectively.*
(a) 200 sensors are randomly distributed with $\theta = 7\ m$ which is partitioned into correlated regions of size, $R_{corr} = 9\ m$
(b) 200 sensors are randomly distributed with $\theta = 22\ m$ (i.e. $R_{corr} = 14\ m$)
(c) 200 sensors are randomly distributed with $\theta = 32\ m$ (i.e. $R_{corr} = 20\ m$)
(d) 200 sensors are uniformly distributed with $\theta = 1.5\ m$ (i.e., $R_{corr} = 2.16\ m$).

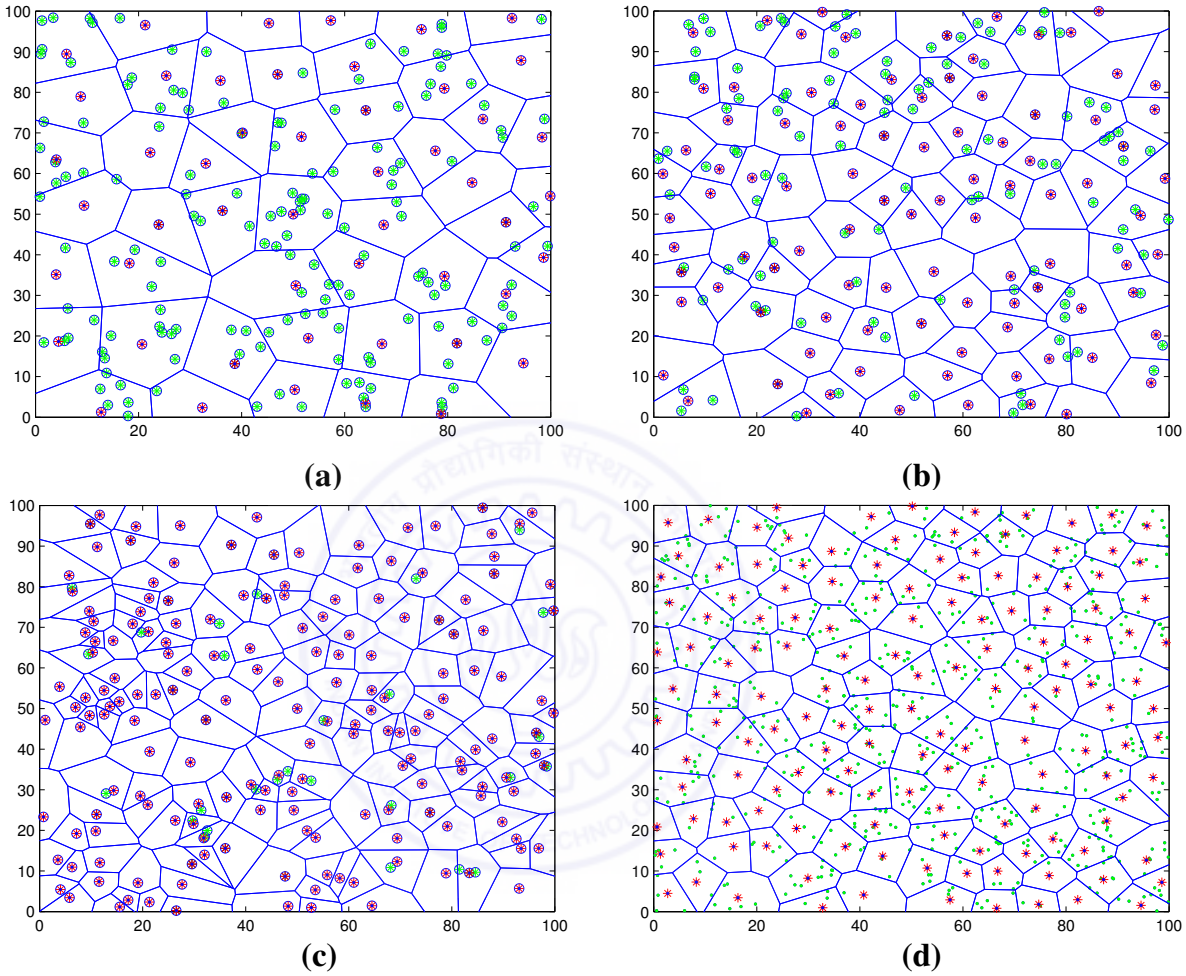


Figure 5.2: *Region divided into correlated sub-regions by the voronoi partitioning for different correlation threshold ξ, given total* 200 *sensors with* $\theta = 15$*, where green sensors and red sensors show correlated nodes and representative nodes respectively.*
(a) The correlation threshold, $\xi = 0.21$ *(i.e.,* $R_{corr} = 10\ m$*)*
(b) The correlation threshold, $\xi = 0.51$ *(i.e.,* $R_{corr} = 5.81\ m$*)*
(c) The correlation threshold, $\xi = 0.82$ *(i.e.,* $R_{corr} = 2.16$ *)*
(d) 800 sensor nodes distribution with correlation threshold, $\xi = 0.82$ *(i.e.,* $R_{corr} = 2.16\ m$*).*

So almost every node becomes active to represent the region when ξ goes down. This leads to higher energy consumption in the network.

Based on these observations, energy consumption can be improved to an optimum level by fine-tuning the parameter ξ, so that just the required number of representative nodes can provide reliable data collection with minimum redundancy. It is also observed from Fig 5.1 that different sensing ranges (i.e., ϑ) corresponding to different sized event areas will have different numbers of representative nodes which need to be activated in the network. So, further energy optimization can be achieved according to different sized event areas which depend on the features of the events and the sensor hardware.

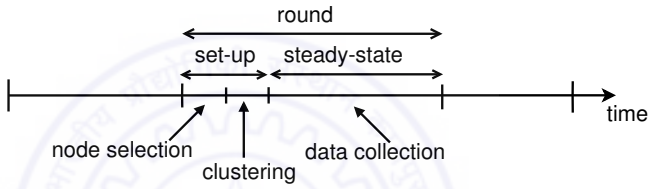


Figure 5.3: *Timeline of Pre-selected N-LEACH*

Based on observations from the above results, a new protocol, PS-NLEACH is proposed which is an extension of N-LEACH [102]. The idea is similar to the protocol given in [17], but we used different sensing framework here. Using proposed sensing framework, the selection of representative nodes can be decided according to the sensing range and the desired data resolution as per application requirements. In our PS-NLEACH, the representative node selection phase is inserted into the set-up phase just before the cluster formation phase as shown in Fig. 5.3. Initially, a base station broadcasts R_{corr} values to all the nodes according to a pre-defined required reliability. At the beginning of all rounds, the selection phase of representative nodes starts before the actual procedure of LEACH or N-LEACH. The steps are the following:

Step-1: For the given requirements of the application, the R_{corr} is pre-determined

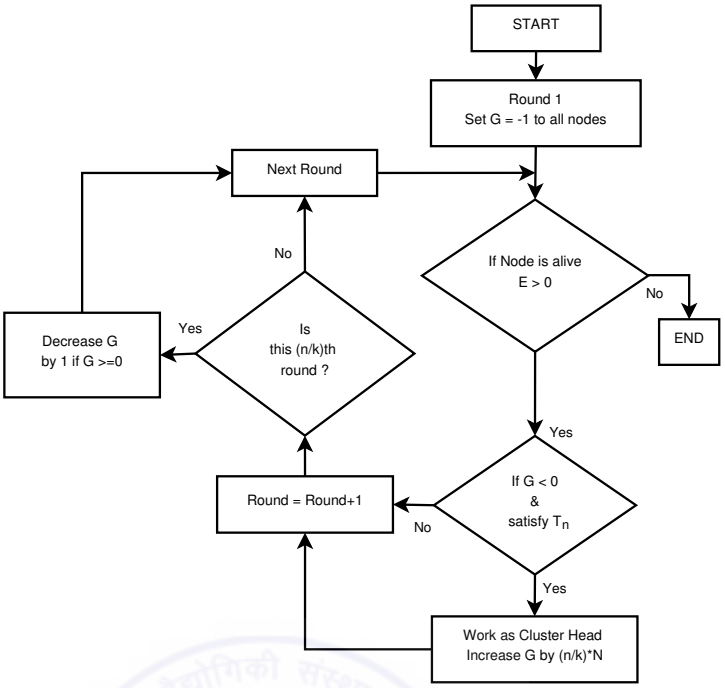


Figure 5.4: *N-LEACH Algorithm Flowchart.*

according to the sensing range and the desired data resolution. The sink broadcasts R_{corr} to all nodes in the sensor field.

Step-2: Each node adjusts its transmission power according to R_{corr}, so that only the correlation region is covered.

Step-3: To select a representative node in the correlation region, each node performs a carrier-sense multiple access (CSMA MAC) protocol by broadcasting a short packet header. If a node is successful in transmitting this packet, it would be the winner and hence will be the representative node in the current round. If other nodes overhear this packet, it means they are within correlation region of the selected representative node which has won the channel first. So they quit from data transmission in the current round and go to sleep.

Step-4: After the representative node selection phase, the N-LEACH procedure is executed to select the cluster heads among all the awake nodes (i.e. pre-selected nodes). Then, the data-collection procedure starts according to the actual routing operation of

LEACH. A complete procedure of N-LEACH is described in [102] and the flowchart is shown in Fig. 5.4.

In this way, the representative nodes, named as the pre-selected nodes based on spatial correlation, avoid the transfer of redundant data to the sink. A complete flowchart of the proposed Pre-selected N-LEACH is shown in Fig. 5.5. Fig. 5.6 shows how different representative nodes are selected in each round using PS-NLEACH that are R_{corr} distance away from other representative nodes.

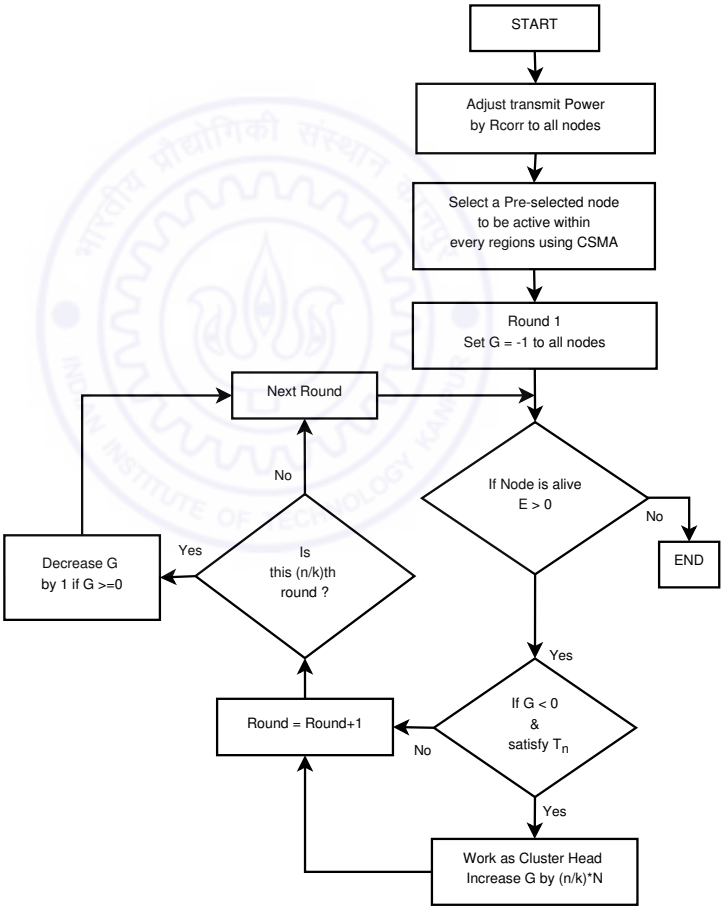


Figure 5.5: *Proposed Pre-selected N-LEACH algorithm Flowchart.*

5.3 Simulation Results and Analysis

In this section, the simulation results using the proposed Pre-selected N-LEACH protocol are discussed to evaluate the performance. In the MATLAB simulation tool written by us, a total of 500 sensor nodes are deployed randomly in 100×100 square meters area. The sink node is located at the centroid of the given area. Results are produced by comparing the proposed protocol with the existing routing protocols.

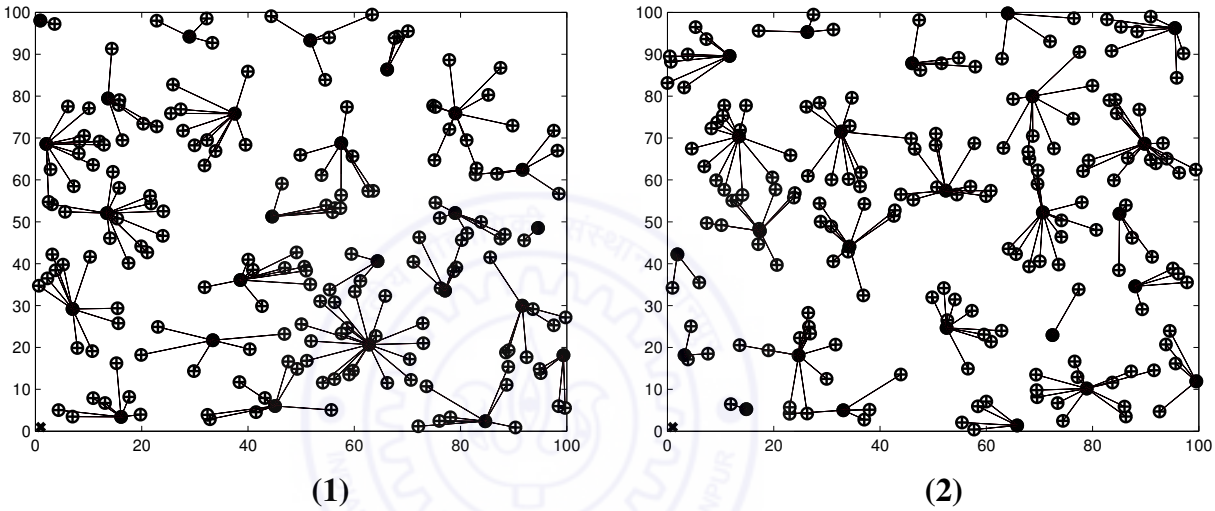


Figure 5.6: *Correlated member nodes selected by their representatives during i^{th} and $(i+1)^{th}$ rounds using Pre-Selected N-LEACH. A connected line shows the members connected with the representative node in the plots.*

Table 5.1: *Total operational rounds for different protocols until alive node left are given as in top row.*

Protocols \ total alive nodes	499	450	250	0
N-LEACH [88]	60	100	200	258
PS-NLEACH with $R_{corr} = 14$ m	65	200	700	1800
PS-NLEACH with $R_{corr} = 18$ m	77	400	1560	3691
PS-NLEACH with $R_{corr} = 24$ m	128	1000	3500	17893

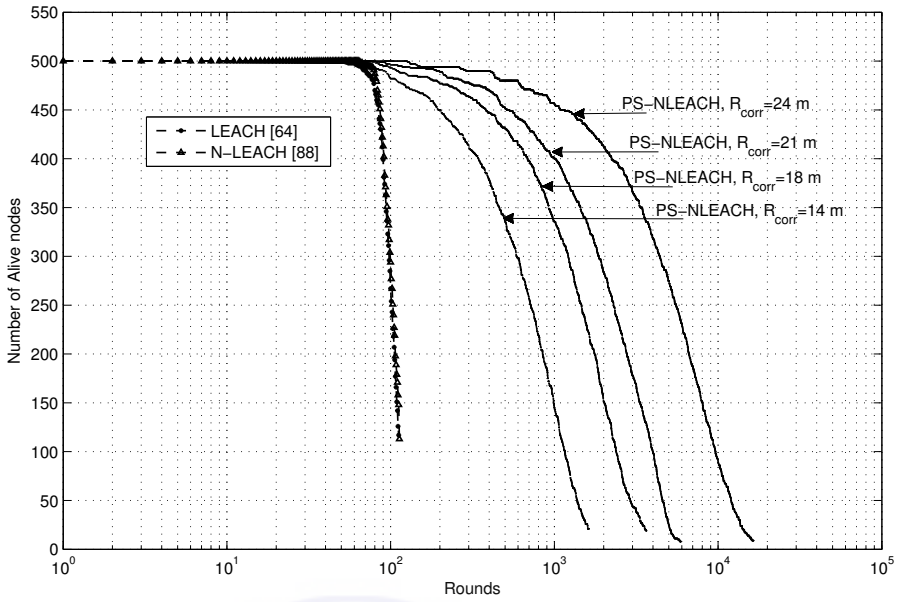


Figure 5.7: *The number of nodes remaining alive over the rounds in a sensor network for different protocols.*

5.3.1 Energy Model

In the simulations, the same radio model and energy parameters are used as discussed in [25], which indicates that if a node transmits k number of bits, the energy consumption in transmission will be

$$\begin{aligned} E_{TX}(k,d) &= E_{TX-elec}(k) + E_{amp}(k,d), \\ &= \begin{cases} k.E_{elec} + k.E_{fs}d^2 & \text{if } d < d_0; \\ k.E_{elec} + k.E_{mp}d^4 & \text{if } d > d_0, \end{cases} \end{aligned} \tag{5.2}$$

and the energy used in the reception will be

$$E_{RX}(k) = E_{RX-elec}(k) = k.E_{elec}. \tag{5.3}$$

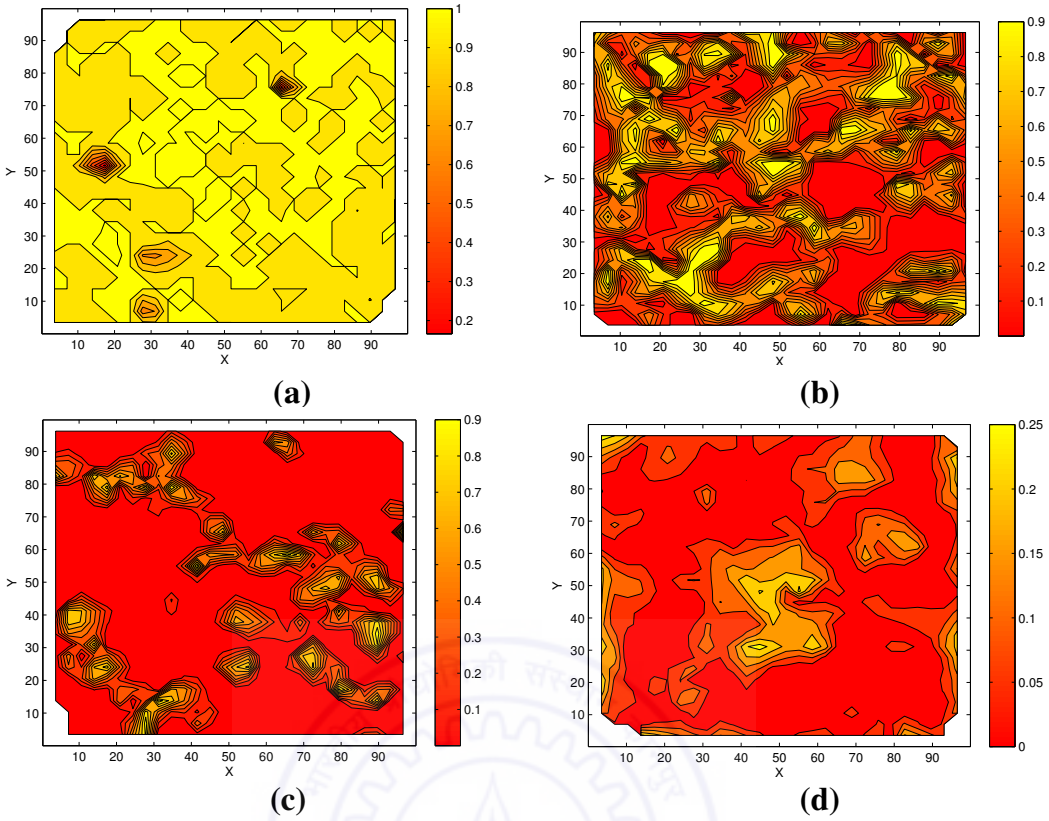


Figure 5.8: *Contour plots showing residual energy of nodes distributed over the region, for* 400 *sensors with* $\theta = 15$, $R_{corr} = 10\ m$. *The color bars show residual energy of nodes for given initial energy* $= 1$ *joule.*
(a) After 40 *rounds first node has died*
(b) After 650 *rounds,* 50 *% nodes have died*
(c) After 1158 *rounds,* 80 *% nodes have died*
(d) Using N-LEACH, after 105 *rounds,* 50 *% nodes are died.*

Here, E_{elec} is the energy consumed in radio electronics for transmission and reception. The consumed amplifier energy, E_{amp} for 'free space (fs) loss' and 'multi-path (mp) loss' models is $E_{fs}d^2$ and $E_{mp}d^4$ respectively. The values of radio parameters are set as $E_{elec} = 50\ nJ/bit$, $E_{fs} = 10\ pJ/bit/m^2$, $E_{mp} = 0.0013\ pJ/bit/m^4$ and $d_0 = 87\ m$.

5.3.2 Simulation Results

Fig 5.7 shows the number of live nodes (still retaining some battery power) over the simulation time. It can be seen from the plot that the life-span of a WSN using PS-

NLEACH is nearly 4 times longer than that of N-LEACH for $R_{corr} = 14\ m$ and 6 times longer for $R_{corr} = 24\ m$. For example, when a total of 450 nodes remain viable in the network, N-LEACH completes a total of 100 rounds while for the same number of viable nodes, PS-NLEACH achieves a total of 200 rounds with $R_{corr} = 14\ m$ and 1000 rounds with $R_{corr} = 24\ m$. Similarly, by the time a network is completely dead, the N-LEACH completes a total of 225 rounds while the PS-NLEACH achieves a total of 1800 rounds with $R_{corr} = 14\ m$ and 17893 rounds with $R_{corr} = 24\ m$. The results of the plot in Fig 5.7 are summarized in Table 5.1.

In PS-NLEACH, after the node selection phase during each round, only the representative nodes remain active while non-representative nodes go to sleep. So the number of active nodes is much lesser than that of LEACH. Therefore, the amount of energy consumption will reduce, and in turn the network's life-span will increase. It is also seen in Fig 5.7 that as the value of R_{corr} increases, the number of representative nodes decreases, which will lead to a further increased life-span of the network. Since power of the nodes is adapted according to R_{corr}, the PS-NLEACH further reduces the power consumption of the nodes. Consequently, it is shown from Figs. 5.1, 5.2 and 5.7 that the life-span of the network gets longer with an increase of R_{corr}, because the number of representative nodes decreases; thus, fewer collisions and lesser redundancy in data would lead to more energy savings.

It is shown by the contour plots given in Fig 5.8 that the energy consumption is quite balanced throughout the network. The plots indicate balanced energy consumption over the simulation time as shown by the distribution of the residual energy of nodes. It is also seen from Fig 5.6 that the representative nodes and the non-representative members in their sub-regions are chosen differently in every round using the proposed PS-NLEACH algorithm. Hence, the balanced energy consumption can be achieved.

5.4 Summary

In this chapter, a correlation-based LEACH i.e. PS-NLEACH is proposed that exploits the concept of spatial correlation during data collection. Motivated by the framework and node selection algorithm developed in Chapter 3, the proposed protocol divides the entire sensor field into several correlated sub-regions and selects a subset of nodes as the representatives of the regions. In each round, only the representative nodes collect data using the dynamic clustering protocol. Simulation results show that PS-NLEACH reduces the energy consumption of the nodes and increases the life-span of the system with the desired resolution in data (i.e., using the correlation threshold). Result analysis using two new parameters R_{corr} and ξ has been discussed in this chapter to achieve the required data resolution which is the function of the number of nodes representing the network.

Chapter 6

Conclusions and Future work

In this Chapter, the main contributions and limitations of the research work presented in this thesis are highlighted. Finally, the scope for the future work has also been presented.

6.1 Summary of Contributions

To summarize, the main contributions of the thesis are as follows.

- Development of model for spatial correlation of reported data collected by the sink from reporting nodes detecting the same event. The correlation model has been analysed, evaluated and varified using simulations. Further, MAC and routing protocols has also been proposed based on this framework.
- Investigation of node selection techniques which exploit the spatial correlation without compromising the quality of received information and its use for efficient design of communication protocols for EWSNs.
- Derivation of a new distortion function which gives the relationship between the

quality of information (i.e., information reliability) and the number of reporting nodes.

- Development of Event-MAC protocol to address the problems of correlated event reporting and generated collision-prone traffic when the correlated nodes detect the same event. Event-MAC protocol uses the outcome of the proposed model framework for reducing collisions.

- Development of Pre-Selected NLEACH algorithm based on model framework for correlation-based efficient data collection in WSNs.

6.2 Important Findings

In Chapter 3, a basic correlation model has been introduced to represent the correlation characteristics between Omni-directional sensor nodes in WSNs. Inspired by the model framework via numerical analysis, simulation results and discussions, following conclusions have been drawn at the both the MAC and Network layers:

- At the network layers, the densely deployed WSN can be partitioned into non-overlapping correlated clusters of nodes by the use of proposed GCC algorithm.

- One can achieve required event reliability through proper tuning of ϑ and ξ, i.e., both the sensing range and the correlation threshold for reliable data delivery if the distribution of node is considered to be fixed for a given period of time.

- For the estimation of the event source, the optimum node density within the event area can be found for the maximum observed event distortion. An optimal value of minimum reporting nodes can be determined through proper tuning of ϑ for given D_{max} as per the application requirement.

In Chapter 4, introducing the idea of collision-free channel access in the influencing region only, Event-MAC accomplishes the following:

- Event-MAC mitigates the effects of collisions and over-hearing by using local TDMA scheduling in the influencing region of the event. In addition, correlation based decision on node selection and prioritized scheduling provides QoS in the event reporting for EWSNs. Thus, based on correlation and priorities, the minimum distortion using lesser number of reporting nodes is achieved by Event-MAC where the sink collects reports only from nodes closer to the event.

- Upon occurrence of an event, a node will switch over to TDMA scheduling mode by running Event-MAC. So all reporting nodes within event area operate scheduled channel access with priorities and the nodes outside the event area perform contention based channel access (CSMA/CA). Thus, Event-MAC ensures collision-free channel access and fast delivery of reports from the event area.

- Event-MAC considers priorities according to the events' observations (i.e., the received energy level from the event source) and network components such as packet type, queue length, packet delay and residual energy. Thus, Event-MAC provides low redundancy, minimum energy consumption, high packet delivery ratio, and bounded delay for EWSNs.

Simulation experiments were conducted using NS-2, Network Simulator [95]. The advantage of spatial correlation and prioritized scheduling in event reporting are compared with the existing protocols. Comparison with other correlation-based MAC protocols shows that results using Event-MAC are consistent with theoretical results on minimum distortion. The reason is efficient decision in node selection and node scheduling in the event area.

Similarly, based on proposed model framework, Pre-Selected NLEACH algorithm has been proposed in Chapter 5. This algorithm is the modified version of popular routing protocol LEACH [25] in WSNs. For correlation-based efficient data collection in sensor network applications, the Pre-Selected NLEACH algorithm achieves the following.

- The proposed routing protocol exploits the concept of spatial correlation during data collection according to node selection based on proposed model framework.
- In each round, the representative nodes do the data collection and data processing throughout the network, while non-representative nodes switch over to the sleep mode. Further, every node gets equal chance of being a representative node in the network.
- There is a trade-off between the energy consumption and the information reliability, where required information reliability is the function of the number of nodes representing in the network.

For PS-NLEACH, the simulation experiments have been conducted using software written in MATLAB [103]. Simulation results show that PS-NLEACH reduces the energy consumption of the nodes and increases the life-span of the system with the desired resolution in data achieved using the appropriate correlation threshold. Almost half of energy is conserved by using PS-NLEACH instead of simple LEACH algorithm.

6.3 Limitations

There are many limitations of protocols proposed in this thesis. The assumptions and models used throughput the work done in the thesis are highlighted in Chapter 1. We revisit some of the limitations here. First the proposed protocols are designed for specific type of events only. The correlation model used with protocols needs placement

coordinates of node and their sensing range. We have considered those events which trigger all the nodes within a radius of sensing range. They do not trigger nodes which are out of range irrespective of signal strength of the event. Therefore, the proposed protocols are not suitable when node placement and sensing range are unknown.

Second, since the proposed protocols require the location information of nodes, the network should be static. The nodes can join or leave the network and can be dead at any point of time but nodes otherwise remain at the same location. Thus, the proposed protocols may not work for sensor networks with mobility of nodes.

Our proposed protocols utilise the header section of the packet for storing important information such as location, sensing range, correlation threshold value etc. When the required reliability/fidelity is not achieved because of a change in environmental conditions, users can disseminate the tuning information to the nodes in the network. However, they can do this at the expense of additional power consumption and protocol overhead.

Finally, the proposed protocols do not support the multiple simultaneous events in WSNs. There should be sufficient interval between any two consecutive events for working of the protocols. However, the work presented in this thesis can be extended for sensor network with multiple events.

6.4 Future Work

Many research problems still need to be investigated based on the correlation in wireless sensor network. The following research directions are identified as the scope for future work.

- For the estimation of the event source, we consider two-dimensional spatial signal field. One can also view this problem in the context of random/irregular sampling and thus, appropriate sampling strategies can be further investigated.

- In event-driven sensor network applications, convergecast type communication patterns are usually found in the network. One can assign predetermined potential forwarders to route the reports as fast as possible from the event area. Candidate forwarding schemes can be used with Event-MAC protocol in order to make it suitable for time-critical events. However, it needs a cross-layer architecture to utilize the MAC layer information in the forwarding decisions at Network layer.

- Event-MAC protocol can be extended for sensor network with multiple events. One can identify a particular event using associated time-stamps or location information. The multi-channel extension of Event-MAC can also be a possibility to handle the multiple events.

- The interpolation/extrapolation techniques can also be used in the context of spatial correlation in data for WSNs.

Appendix A

Simulation Tools

For simulation and analytical results in this thesis, two simulation tools have been used, the ns-2 network simulator [95] and our own simulator code written in MATLAB [103]. Event-MAC protocol has been implemented using the ns-2 network simulator. A short overview of ns-2 is given in next section. MATLAB is an effective tool to provide a technical computing environment that enables numerical computation and data visualization. The routing protocol implementation and analytical development of correlation framework have been done using MATLAB.

A.1 ns-2 Network Simulator – an Overview

Using the network simulator (ns-2) version 2.30 [95], the design of Event-MAC protocol has been implemented to evaluate its performance. The ns-2 is widely used standard network simulator to simulate wired and wireless network. Its object-oriented design allows easy implementation of a new protocols. Because of the popularity and availability

as open source, larger number of different protocols have been built and contributed by the research community as the extensions of ns-2. It is written in two languages – C++ and OTcl (Object-oriented Tool Command Language). Both the languages are linked together to provide interaction between various parameters. C++ is used to define internal mechanism and OTcl provides user interface. In other words, C++ and OTcl handle the data path and control path implementation respectively [104].

Wireless multi-hop networking support in ns-2 is developed by CMU Monarch group [105]. Each wireless node module of ns-2 comprises of different components based on the ISO network stack as shown in Fig. A.1. These components are described in next sections. Event-MAC protocol is implemented as an extension of MACtdma module of ns-2.

A.1.1 Wireless Node Design in ns-2

The schematic representation of a wireless node module in ns-2 [95] is shown in Fig. A.1. The physical layer modeling is defined by PHY module which represents the properties of radio and physical layer protocol. It includes radio interfaces, models of receiver sensitivity and antennas. The link layer model consists of medium access control (MAC) protocol with address resolution protocol (ARP) and interface queue (IFq). The function of ARP is to maintain mapping between the MAC address and the IP address. Once the MAC address of next hop is determined by ARP, the packet is queued in the IFq. The MAC module gets packets from queue (i.e., IFq) and sends them according to protocol operation. The network layer modeling is represented by a routing agent. At the upper level, the classifiers are responsible to pass the packet to its final destination sink agent or the routing agent. If the incoming packet has reached its final destination, then it will pass from address classifier to port classifier for delivering it to the sink agent. Otherwise, the routing agent will handle this packet to forward it to next hop.

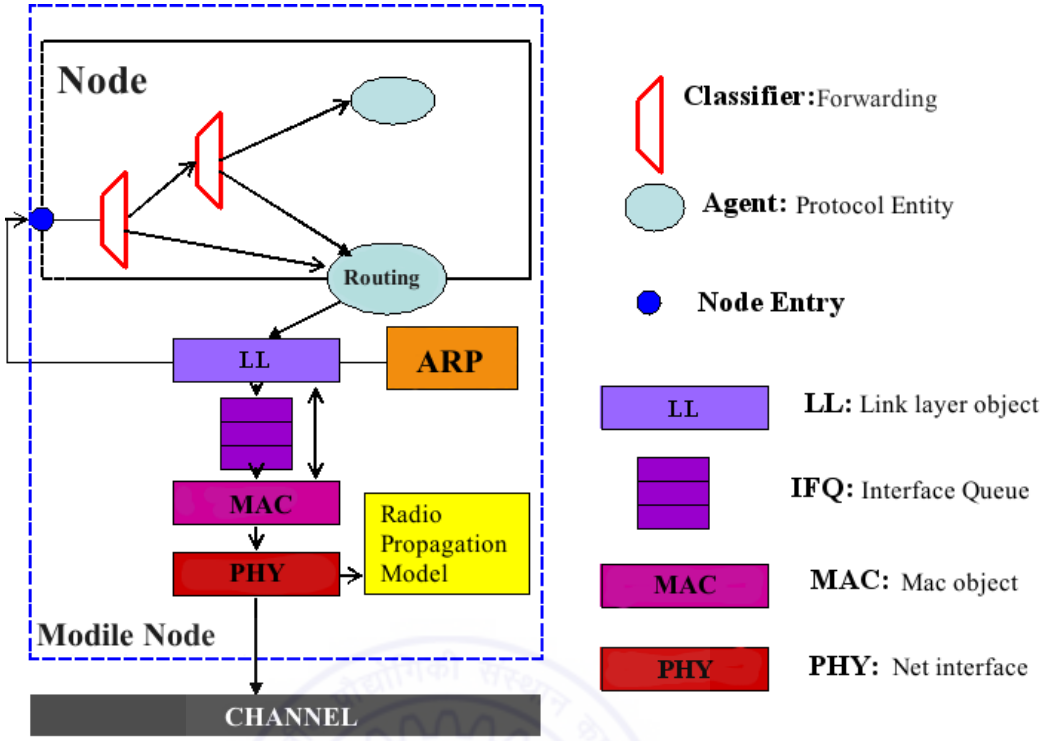


Figure A.1: *Schematic representation of a wireless node module in ns-2 (CMU Monarch implementation [31].*

A.1.2 Event-MAC Simulation Design in ns-2

The schematic representation of Event-MAC simulation design is shown in Fig. A.2. Event-MAC is implemented as an extension of MACtdma module of ns-2. The components used in Event-MAC simulation are described as follows.

PHY module: The PHY modules of all sensor nodes are connected to a common wireless channel. A common propagation model is also linked to PHY module to calculate the received power of a incoming packet. The *downtarget_* and *uptarget_* objects provide the interface between all ISO components of a sensor node. Two functions *send* and *recv* are implemented using *downtarget_* and *uptarget_* objects respectively. These functions are used to pass the packet between PHY and Channel as well as between PHY and MAC. When a sender node transmits a packet on the channel, it stores packet information such as transmitted power, location of the node in the header section of the

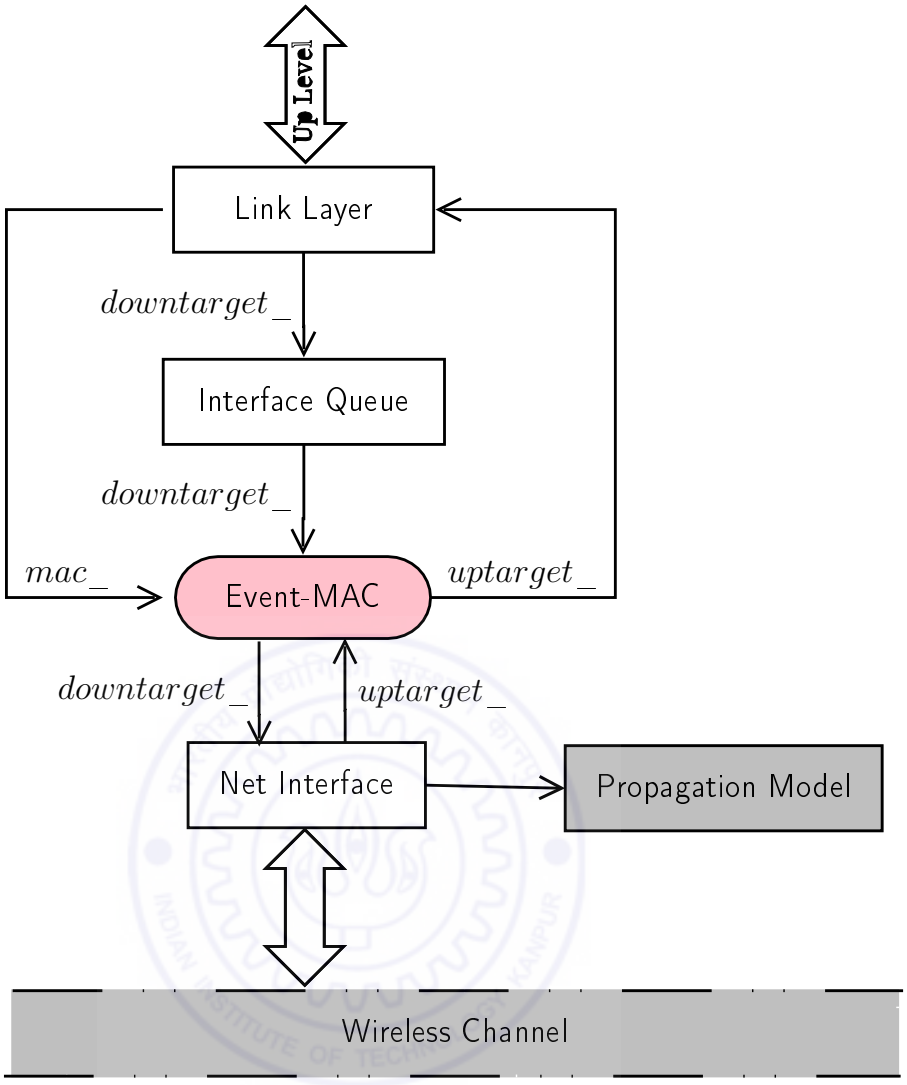


Figure A.2: *Event-MAC Simulation design in ns-2.*

packet. PHY module then triggers an event for transmission. Wireless channel module receives the packet and passes it to all other PHY modules. Upon notification of this event, PHY module passes necessary packet information such as transmitted power to the Propagation model. The Propagation model then calculates the received power of the packet and passes this value to all other PHY modules. PHY modules of other sensor nodes discard this packet if the received power is lower than the predefined carrier sense threshold power. Otherwise, the packet is delivered to MACtdma module and then Event-MAC module for further processing.

MACtdma module: MACtdma module is implemented in ns-2 as a simple TDMA protocol. It is preamble-based centralized TDMA protocol (see ns-2 user manual [106]). For each node in the network, different time slots are allocated to send and receive the packets. A TDMA superframe is composed of many time slots that depends on the number of nodes in the network. A TDMA superframe is divided into preamble and data transmission slots. Within preamble, each node has its own slot to broadcast the destination ID of outgoing packet. Other nodes listen this information in preamble and record the time for receiving the packets. Event-MAC utilizes the TDMA frame structure of MACtdma module. The Event-MAC protocol combines the TDMA with CSMA/CA scheme. So it uses RTS/CTS/DATA/ACK mechanism of MAC802.11 with MACtdma module in ns-2. Since TDMA operation is activated by an event using Event-MAC, the modified version of MACtdma will run inside the event area when an event occurs.

Propagation model: There are three types of Propagation models implemented in ns-2. These are two-ray, free space and shadowing model. It has been shown by Rappaport [107] that two-ray model gives more accurate prediction at a long distance than the free space model. So two-ray ground model is used in all our simulations. The received power at the distance d is given by.

$$P_r(d) = \frac{P_t G_t G_r h_t^2 h_r^2}{d^4}. \tag{A.1}$$

Where,

$P_r(d)$ is the received power at the given transmitter-receiver distance,

P_t is the transmission power,

G_t is the antenna gain of the transmitter,

G_r is the antenna gain of the receiver,

h_t is the height of transmitting antenna above ground,

h_r is the height of receiving antenna above ground,

Table A.1: *Simulation Parameters*

ns2 energy model	$EnergyModel$ used in NS2 [95]	Event sensing range	400 unit
Initial power	100 J	Transmission range	200 unit
Tx power	66 mW	Event generation interval	50 s
Rx power	39 mW	Carrier Sensing range	550 unit
Idle power	35 mW	Size of RTS/CTS/ACK	10 B
Bandwidth	20 Kbps	Correlation radius	50 m
DIFS	10ms	Size of Data	100 B
SIFS	5ms	Contention Window	64 ms
Retry Limit	5	Duty Cycle	5 %

d is the distance between the transmitter and the receiver.

Energy Model: ns-2 provides a module for the energy model to maintain total energy in the simulation. Using the energy model module, the energy consumption (in Watts) details at a given time are stored in ns-trace file. These are power consumption in SLEEP and IDLE state as well as consumption in transmitting and receiving the packets. Table A.1 shows the default parameters used in the simulations.

A.1.3 RCET model design in ns-2

In the ns-2 network simulator, the Random Correlated Event traffic (RCET) model [85] has been created to simulate bursty traffic triggered by spatially correlated nodes within an event area, commonly observed in event-driven applications such as detection, tracking etc. Let an event area be a round region of radius R_c with center at (x, y) location in two-dimensional plane. The nodes within the circular area of radius R_c are allowed to be activated for generating the traffic (i.e. UDP packets). All the UDP packets are routed to the sink via intermediate nodes. The sink node is placed at one of the corners of the network. For an occurring event, the RCET model picks a random point at (x, y) location within the given simulation area. Only the nodes inside the R_c-radius around (x, y) are

the potential UDP sources for event reporting. Since the simulations are performed for occurrance of many events, the minimum interval between two consecutive events is set to 50 seconds. The maximum duration between two consecutive events is 100 seconds.

Appendix B

Correlation function $K_\vartheta(.)$

B.1 Sample results

The results for correlation model for $\theta = 18$ (i.e., sensing range, $r = 9$) using same scenario as given in section 3.2.2, are shown in Table B.1.

B.2 C-INS algorithm using $K_\vartheta(.)$

The correlation function $K_\vartheta(.)$ can be used as a model of spatial correlation in existing MAC and routing protocols in event-driven WSNs. Based on correlation function $K_\vartheta(.)$, a Correlation-based Iterative Node Selection (C-INS) algorithm is used (**Algorithm 1B**) to determine the least correlated nodes as representative nodes within an event area that exploit the spatial correlation. In C-INS algorithm, $\mathcal{N}$ denotes a set of reporting nodes inside the event area and $K_\theta\{N_i, N_j\}$ is the correlation co-efficient between node

Table B.1: *Simulation Results using Correlation Model*

Nodes	Separation distance (m)	Correlation Coefficient	Nodes	Separation distance (m)	Correlation Coefficient
n_2, n_5	15.8441	0.04891	n_5, n_3	5.2761	0.6322
n_4, n_1	19.1921	0	n_9, n_{14}	17.3858	0.0075
n_7, n_3	0.7142	0.9495	n_{12}, n_{15}	11.5941	0.2408
n_5, n_3	16.9826	0.0160	n_{11}, n_{17}	10.9972	0.2737
n_9, n_{14}	15.844	0.0489	n_2, n_5	2.8991	0.7958
n_{12}, n_{15}	18.6793	0	n_4, n_1	12.4411	0.1962
n_{11}, n_{17}	6.0605	0.5137	n_7, n_3	10.2650	0.3154
n_2, n_5	6.5541	0.5468	n_4, n_1	9.8173	0.3417
n_4, n_1	0.7124	0.8791	n_7, n_3	9.7851	0.3436
n_7, n_3	3.9224	0.7251	n_5, n_3	6.7544	0.5337
n_5, n_3	17.8183	0.00121	n_9, n_{14}	15.6050	0.0571
n_9, n_{14}	10.944	0.2766	n_{12}, n_{15}	7.7948	0.4664
n_{12}, n_{15}	1.6887	0.8807	n_{11}, n_{17}	11.5042	0.2457

N_i and N_j given by our proposed correlation function.

B.3 Comparative analysis using C-INS algorithm[1]

The simulation scenario contains 49 nodes grid network distributed into $2000 \times 2000\ m^2$ area. A random event will be generated after a fixed time interval (i.e. every 100 sec.) using RCET model as introduced in Chapter 2. Total N nodes will be activated to transmit reports about an event to the base station (i.e., sink node). The least correlated reporting nodes (i.e. M) out of N nodes are selected inside the event area using the proposed correlation model. Except for the networking parameters discussed in Chapter 2, we have used default settings as used for evaluations of CC-MAC, RMAC, RI-MAC, and DW-MAC by the respective authors. In NS-2, we have simulated CC-MAC [14], RMAC [108], RI-MAC [109], and DW-MAC [51] to show the efficiency of

[1] R.K.Shakya, Y.N.Singh, and N.K.Verma, "A Correlation Model for MAC Protocols in Event-driven Wireless Sensor Networks", In Proc. of the 2012 IEEE Region 10 Conference, IEEE TENCON 2012, pp. 1–6, November 2012.

Algorithm 1B Correlation-based Iterative Node Selection (C-INS)

S = NodeSelectCorr($\{N_1, N_2, ...\}, \{K_\theta(.)\}_{N*N}, M$)
begin
S=$\{\varnothing\}, \mathcal{N} = \{N_1, N_2, ...\}, K_\theta\{N_i, N_j\} = K_\theta(i, j)$.
Find the (N_i, N_j) = $\arg\min_{N_i, N_j \in \mathcal{N}} \{K_\theta(N_i, N_j)\}$;
/*Find the least correlated pair of nodes.*/
Add the corresponding N_i and N_j into S.
if $M > 2$ **then**
for $k = 1$ to $M - 2$ **do**

$$N_m = \underset{N_m \in \mathcal{N}, N_m \notin S}{\arg\min} \{K_\theta(N_m, S)\};$$

Add N_m into S.
end for
end if
return $S = \{N_1, N_2, ..., N_M\}$
end

our correlation model in terms of energy and latency for event-driven workload. These protocols are either contention based or contention based combined with scheduling for WSNs. We first simulated RCET model to determine the total reporting nodes (i.e. N) inside an event area according to an event trigger. We then apply our correlation model to determine the least correlated reporting nodes (i.e. M) using Algorithm-I with $\theta = 140$ meters.

Figs. B.1(a) and B.1(b) show the end-to-end event latency corresponding to number of reporting nodes (achieved by varying the event radius). Basically it is the average time required to report an event. In other words, event latency is the time required to send all UDP packets generated by the nodes which are located inside an event area. On an average 637.45 packets are generated for an event. It is shown that CC-MAC achieves low latency compared to other protocols in both cases because it is minor modification of IEEE 802.11 without duty-cycle mode and forwarding of packets is done based on priority handled by E-MAC. In the case of least correlated reporting nodes (see Fig. B.1(b)), event latency has been reduced to 20 - 40 % for all the protocols because of

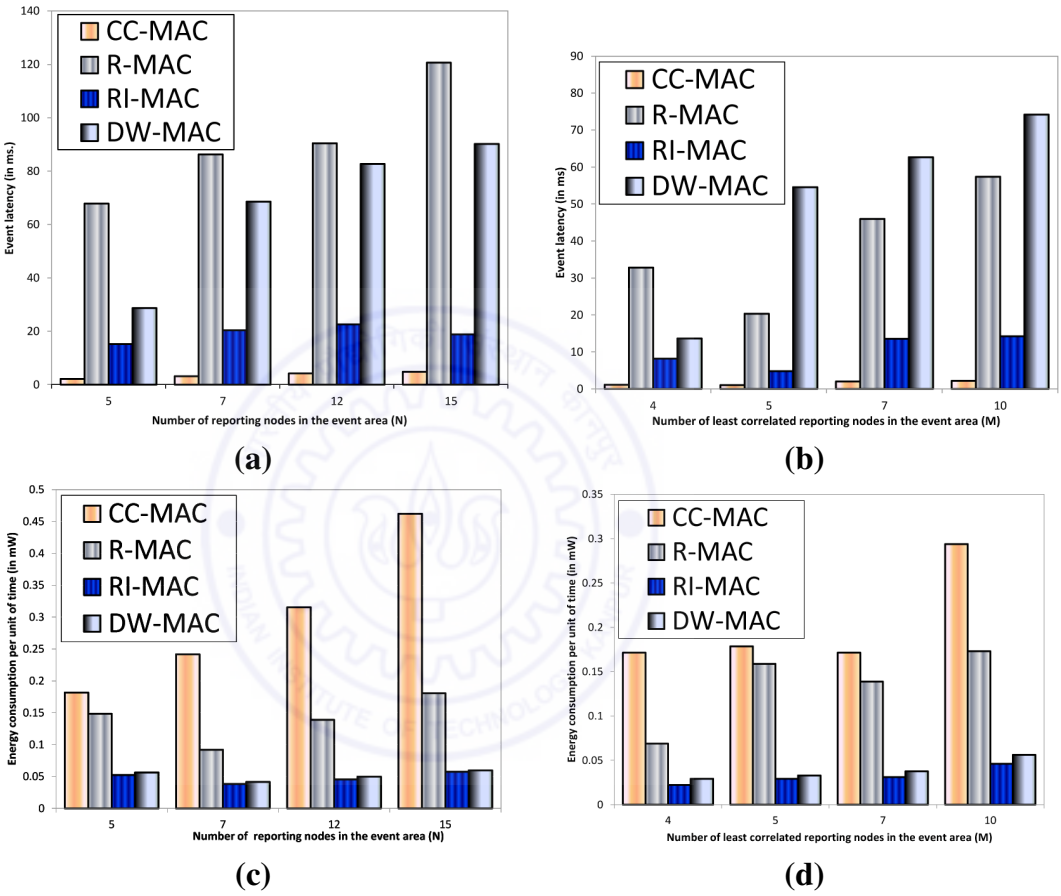


Figure B.1: *Average time required to report an event ((a) without using correlation model and (b) using correlation model), average amount of energy consumed by the network to transmit an event ((c) without using correlation model and (d) using correlation model).*

reduction of number of reporting nodes to M based on the location. It decreases the traffic and hence correlated contention is minimized.

Figs. B.1(c) and B.1(d) show the average amount of energy consumed per unit time for an event by varying number of reporting nodes (i.e., varying the event radius). We can see the impact of reporting nodes on energy consumption due to increased traffic load. With our correlation model, energy consumption is obviously reduced as lesser number of nodes (M) are selected for same scenario taking spatial correlation into consideration. Hence, we see that energy consumption and event latency both are reduced by exploiting the spatial correlation.

Appendix C

Derivation for distortion function, $D_E(N, M)$

The estimated value, $\hat{S}_i$ using Equation (3.19) is simplified as

$$\hat{S}_i = \mathbb{E}[S_i] + \frac{\mathbb{E}[S_i X_i] - \mathbb{E}[S_i]\mathbb{E}[X_i]}{\mathbb{E}[X_i^2] - \mathbb{E}[X_i]\mathbb{E}[X_i]} \left(X_i - \mathbb{E}[X_i] \right). \tag{C.1}$$

Here, the following properties are assumed to be true.

$$X_i = S_i + N_i,$$
$$\mathbb{E}[S_i] = m_i, \qquad \mathbb{V}ar[S_i] = \sigma_S^2, \quad \mathbb{V}ar[N_i] = \sigma_N^2, \ i = 1, 2, ...N,$$

and

$$\begin{aligned}
&\mathbb{E}(a^2) = [\mathbb{E}(a)]^2 + \mathbb{V}ar(a), \qquad \mathbb{E}(ab) = \mathbb{C}ov(a,b) + \mathbb{E}(a)\mathbb{E}(b),\\
&\mathbb{V}ar(kc) = k^2\mathbb{V}ar(c), \qquad \mathbb{E}(kc) = k\mathbb{E}(c), \text{ where } k \text{ is constant,}\\
&\mathbb{V}ar(\sum_{i=1}^{N} a_i) = \sum_{i=1}^{N} \mathbb{V}ar(a_i) + \sum_{i=1}^{N}\sum_{\substack{j=1\\ j\neq i}}^{N} \mathbb{C}ov\{a_i, a_j\},\\
&\mathbb{C}ov\left\{ \sum_{i=1}^{N} k_i a_i, \sum_{j=1}^{M} k_j a_j \right\} = \sum_{i=1}^{N}\sum_{\substack{j=1\\ j\neq i}}^{M} k_i k_j \mathbb{C}ov\{a_i, a_j\}.
\end{aligned} \tag{C.2}$$

Hence,

$$\begin{aligned}
\mathbb{E}[S_i X_i] &= \sigma_S^2 + m_i^2,\\
\mathbb{E}[X_i^2] &= m_i^2 + \sigma_S^2 + \sigma_N^2.
\end{aligned} \tag{C.3}$$

Using Equation (C.3), Equation (C.1) can be simplified as

$$\hat{S}_i = \frac{\sigma_N^2 m_i + \sigma_S^2 (S_i + N_i)}{(\sigma_S^2 + \sigma_N^2)}. \tag{C.4}$$

Now, the calculation of distortion $D_E(N, M)$ is given by Equation (3.22) as

$$D_E(N, M) = \mathbb{E}\left[\left(S(N) - \hat{S}(M) \right)^2 \right], \tag{C.5}$$

where,

$$S(N) = \frac{(1+R_c^2)}{N} \sum_{i=1}^{N} S_i, \quad \hat{S}(M) = \frac{(1+R_c^2)}{M} \sum_{i=1}^{M} \hat{S}_i. \tag{C.6}$$

Now, Equation (C.5), using Equation (C.6) can be given as

$$\begin{aligned}
D_E(N, M) &= \mathbb{E}\left[\left(Y_1 - Y_2 \right)^2 \right]\\
&= \frac{(1+R_c^2)^2}{N^2} \underbrace{\mathbb{E}[Y_1^2]}_{\text{(I)}} + \frac{(1+R_c^2)^2}{M^2} \underbrace{\mathbb{E}[Y_2^2]}_{\text{(II)}} - \frac{(1+R_c^2)^2}{NM} \underbrace{2\mathbb{E}[Y_1 Y_2]}_{\text{(III)}},
\end{aligned} \tag{C.7}$$

where,

$$Y_1 = \sum_{i=1}^{N} S_i, \quad Y_2 = \sum_{i=1}^{M} \hat{S}_i. \tag{C.8}$$

$$(I) = \mathbb{E}[Y_1^2] = \mathbb{V}ar(Y_1) + [\mathbb{E}(Y_1)]^2, \tag{C.9}$$

$$(II) = \mathbb{E}[Y_2^2] = \mathbb{V}ar(Y_2) + [\mathbb{E}(Y_2)]^2, \tag{C.10}$$

$$(III) = 2\mathbb{E}[Y_1 Y_2] = 2\mathbb{C}ov(Y_1, Y_2) + 2\mathbb{E}(Y_1)\mathbb{E}(Y_2). \tag{C.11}$$

Now, Equation (C.9) can be simplified using Equation (C.2) as

$$\begin{aligned}
(I) &= \mathbb{V}ar\{\sum_{i=1}^{N} S_i\} + \{\mathbb{E}[\sum_{i=1}^{N} S_i]\}^2 \\
&= \sum_{i=1}^{N} \mathbb{V}ar\{S_i\} + \sum_{i=1}^{N} \sum_{\substack{j=1 \\ j\neq i}}^{N} \mathbb{C}ov\{S_i, S_j\} + \sum_{i=1}^{N} m_i^2 \\
&= \sum_{i=1}^{N} \sigma_S^2 + \sum_{i=1}^{N} \sum_{\substack{j=1 \\ j\neq i}}^{N} \rho_{(i,j)} \sqrt{\mathbb{V}ar\{S_i\}\mathbb{V}ar\{S_j\}} + \sum_{i=1}^{N} m_i^2 \\
&= N\sigma_S^2 + \sigma_S^2 \sum_{i=1}^{N} \sum_{\substack{j=1 \\ j\neq i}}^{N} \rho_{(i,j)} + \sum_{i=1}^{N} m_i^2,
\end{aligned} \tag{C.12}$$

Equation (C.10) can be simplified using Equations (C.2) and (C.4) as

$$\begin{aligned}
(II) &= \mathbb{V}ar\{\sum_{i=1}^{M} \hat{S}_i\} + \{\mathbb{E}[\sum_{i=1}^{M} \hat{S}_i]\}^2 \\
&= \sum_{i=1}^{M} \mathbb{V}ar\{\hat{S}_i\} + \sum_{i=1}^{M} \sum_{\substack{j=1 \\ j\neq i}}^{M} \mathbb{C}ov\{\hat{S}_i, \hat{S}_j\} + \sum_{i=1}^{M} m_i^2 \\
&= \sum_{i=1}^{M} \frac{\sigma_S^4}{(\sigma_S^2+\sigma_N^2)} + \sum_{i=1}^{M} \sum_{\substack{j=1 \\ j\neq i}}^{M} \rho_{(i,j)} \sqrt{\mathbb{V}ar\{\hat{S}_i\}\mathbb{V}ar\{\hat{S}_j\}} + \sum_{i=1}^{M} m_i^2 \\
&= \frac{M\sigma_S^4}{(\sigma_S^2+\sigma_N^2)} + \frac{\sigma_S^6}{(\sigma_S^2+\sigma_N^2)^2} \sum_{i=1}^{M} \sum_{\substack{j=1 \\ j\neq i}}^{M} \rho_{(i,j)} + \sum_{i=1}^{M} m_i^2,
\end{aligned} \tag{C.13}$$

and Equation (C.11) can be simplified using Equations (C.2) and (C.4) as

$$\begin{aligned}(III) &= 2\mathbb{C}ov\left\{ \sum_{i=1}^{N} S_i, \sum_{j=1}^{M} \hat{S}_j \right\} + 2\mathbb{E}(\sum_{i=1}^{N} S_i)\mathbb{E}(\sum_{j=1}^{M} \hat{S}_j) \\ &= 2\sum_{i=1}^{N}\sum_{\substack{j=1\\ j\neq i}}^{M} \mathbb{C}ov\{S_i, \hat{S}_j\} + 2\sum_{i=1}^{N}\sum_{\substack{j=1\\ j\neq i}}^{M} m_i m_j \\ &= 2\sum_{i=1}^{N}\sum_{\substack{j=1\\ j\neq i}}^{M} \rho_{(i,j)}\sqrt{\mathbb{V}ar\{S_i\}\mathbb{V}ar\{\hat{S}_j\}} + 2\sum_{i=1}^{N}\sum_{\substack{j=1\\ j\neq i}}^{M} m_i m_j \\ &= \tfrac{2\sigma_S^4}{(\sigma_S^2+\sigma_N^2)}\sum_{i=1}^{N}\sum_{\substack{j=1\\ j\neq i}}^{M} \rho_{(i,j)} + 2\sum_{i=1}^{N}\sum_{\substack{j=1\\ j\neq i}}^{M} m_i m_j.\end{aligned} \tag{C.14}$$

Thus, expanding Equation (C.7), using Equations (C.7), (C.12), (C.13) and (C.14),

$$\begin{aligned}D_E(N,M) = \frac{(1+R_c^2)^2}{M^2}&\left(\tfrac{M\sigma_S^4}{(\sigma_S^2+\sigma_N^2)} + \tfrac{\sigma_S^6}{(\sigma_S^2+\sigma_N^2)^2}\sum_{i=1}^{M}\sum_{\substack{j=1\\ j\neq i}}^{M}\rho_{(i,j)} + \sum_{i=1}^{M} m_i^2\right) \\ -\frac{2(1+R_c^2)^2}{NM}&\left(\tfrac{\sigma_S^4}{(\sigma_S^2+\sigma_N^2)}\sum_{i=1}^{N}\sum_{\substack{j=1\\ j\neq i}}^{M}\rho_{(i,j)} + \sum_{i=1}^{N}\sum_{\substack{j=1\\ j\neq i}}^{M} m_i m_j\right) \\ +\frac{(1+R_c^2)^2}{N^2}&\left(N\sigma_S^2 + \sigma_S^2\sum_{i=1}^{N}\sum_{\substack{j=1\\ j\neq i}}^{N}\rho_{(i,j)} + \sum_{i=1}^{N} m_i^2\right).\end{aligned} \tag{C.15}$$

Bibliography

[1] I. F. Akyildiz and M. C. Vuran, *Wireless Sensor Networks*. John Wiley & Sons, 2010.

[2] A. Ghosh and S. K. Das, "Coverage and connectivity issues in wireless sensor networks: A survey," *Pervasive and Mobile Computing*, vol. 4, no. 3, pp. 303–334, 2008.

[3] Y. Xiao, H. Chen, K. Wu, B. Sun, Y. Zhang, X. Sun, and C. Liu, "Coverage and Detection of a Randomized Scheduling Algorithm in Wireless Sensor Networks," *IEEE Trans. Comput.*, vol. 59, no. 4, pp. 507–521, 2010.

[4] E. Shih, S.-H. Cho, N. Ickes, R. Min, A. Sinha, A. Wang, and A. Chandrakasan, "Physical layer driven protocol and algorithm design for energy-efficient wireless sensor networks," in *Proceedings of the 7th annual international conference on Mobile computing and networking*, ser. MobiCom '01. New York, NY, USA: ACM, 2001, pp. 272–287.

[5] N. Tezcan and W. Wang, "TTS: a two-tiered scheduling mechanism for energy conservation in wireless sensor networks," *International Journal of Sensor Networks*, vol. 1, no. 3, pp. 213–228, 2006.

[6] M. Miller and N. Vaidya, "A MAC protocol to reduce sensor network energy con-

sumption using a wakeup radio," *IEEE Transactions on Mobile Computing*, vol. 4, no. 3, pp. 228–242, 2005.

[7] W. Ye, J. Heidemann, and D. Estrin, "Medium access control with coordinated adaptive sleeping for wireless sensor networks," *IEEE/ACM Trans. Netw.*, vol. 12, no. 3, pp. 493–506, 2004.

[8] S. Bhattacharjee and S. Bandyopadhyay, "Lifetime maximizing dynamic energy efficient routing protocol for multi hop wireless networks," *Simulation Modelling Practice and Theory*, vol. 32, pp. 15–29, 2013.

[9] J. Park and S. Sahni, "An online heuristic for maximum lifetime routing in wireless sensor networks," *IEEE Transactions on Computers*, vol. 55, no. 8, pp. 1048–1056, 2006.

[10] K. Chakrabarty, S. Member, S. S. Iyengar, and E. Cho, "Grid coverage for surveillance and target location in distributed sensor networks," *IEEE Transactions on Computers*, vol. 51, pp. 1448–1453, 2002.

[11] M. C. Vuran, O. B. Akan, and I. F. Akyildiz, "Spatio-temporal correlation: theory and applications for wireless sensor networks," *Computer Networks Journal (Elsevier)*, vol. 45, pp. 245–259, 2004.

[12] S. Yoon and C. Shahabi, "Exploiting spatial correlation towards an energy efficient clustered aggregation technique (CAG)," in *Proceedings of International Conference on Communications, ICC '05*, 2005, pp. 3307–3313.

[13] S. P. B. Krishnmachari and R. Govindan, "The impact of spatial correlation on routing with compression in wireless sensor networks," in *Proceedings of IPSN,04*, 2004, pp. 28–35.

[14] M. C. Vuran and I. F. Akyildiz, "Spatial Correlation based Collaborative Medium Access Control in Wireless Sensor Networks," *IEEE/ACM Trans. On Networking*, vol. 14, no. 2, 2006.

[15] G. Zheng and S. Tang, "Spatial Correlation-Based MAC protocol for Event-Driven Wireless Sensor Networks," *Journal of Networks*, vol. 6, no. 1, 2011.

[16] M. Zhao, Z. Chen, L. Zhang, and Z. Ge, "HS-Sift: hybrid spatial correlation-based medium access control for event-driven sensor networks," *IET Communications*, vol. 1, no. 6, pp. 1126–1132, 2007.

[17] J. Qian, Z. Yu, S. Liu, and J.-W. Chong, "A Pre-Selected Clustering Protocol Based on Spatial Correlation," in *Proceedings of 4th Int. Conf. on Wireless Communications, Networking and Mobile Computing, WiCOM '08*, 2008, pp. 1–4.

[18] A. Dabirmoghaddam, M. Ghaderi, and C. Williamson, "Cluster-Based Correlated Data Gathering in Wireless Sensor Networks," in *Proceedings of IEEE Int. Symposium on Modeling, Analysis Simulation of Computer and Telecommunication Systems (MASCOTS'10)*, 2010, pp. 163–171.

[19] M. Vuran and O. Akan, "Spatio-temporal Characteristics of Point and Field Sources in Wireless Sensor Networks," in *Proceedings of the IEEE International Conference on Communications (ICC) 2006*, vol. 1, 2006, pp. 234–239.

[20] O. B. Akan and I. F. Akyildiz, "Event-to-sink reliable transport in wireless sensor networks," *IEEE/ACM Trans. Netw.*, vol. 13, no. 5, pp. 1003–1016, 2005.

[21] F. Oldewurtel, J. Ansari, and P. Mahonen, "Cross-Layer Design for Distributed Source Coding in Wireless Sensor Networks," in *Proceedings of Second Int. Conf. on Sensor Technologies and Applications, SENSORCOMM '08*, 2008, pp. 435–443.

[22] J. Barcelo-Llado, A. Perez, and G. Seco-Granados, "Enhanced Correlation Estimators for Distributed Source Coding in Large Wireless Sensor Networks," *IEEE Sensors Journal*, vol. 12, no. 9, pp. 2799–2806, 2012.

[23] M. Amac Guvensan and A. Gokhan Yavuz, "On coverage issues in directional sensor networks: A survey," *Ad Hoc Netw.*, vol. 9, no. 7, pp. 1238–1255, 2011.

[24] L. Liang, D. Gao, H. Zhang, and O. W. W. Yang, "Efficient Event Detecting Protocol in Event-Driven Wireless Sensor Networks," *IEEE Sensors Journal*, vol. 12, no. 6, pp. 2328–2337, 2012.

[25] W. Heinzelman, A. Chandrakasan, and H. Balakrishnan, "An application-specific protocol architecture for wireless microsensor networks," *IEEE Transactions on Wireless Communications*, vol. 1, no. 4, pp. 660–670, 2002.

[26] S. Anna and S. S. D., "On the interdependence of routing and data compression in multi-hop sensor networks," in *Proceedings of the 8th annual international conference on Mobile computing and networking*, ser. MobiCom '02, 2002, pp. 140–147.

[27] S. S. P. J. Kusuma and K. Ramchandran, "Distributed Compression in a Dense Microsensor Network," in *Proceedings of the IEEE Signal Processing Magazine*, vol. 19, no. 2, 2002, pp. 51–60.

[28] M. Dong, L. Tong, and B. Sadler, "Impact of Data Retrieval Pattern on Homogeneous Signal Field Reconstruction in Dense Sensor Networks," *IEEE Transactions on Signal Processing*, vol. 54, no. 11, pp. 4352–4364, 2006.

[29] X. Zhang, H. Wang, F. Nait-Abdesselam, and A. A. Khokhar, "Distrotion Analysis for Real-Time Data Collection of Spatially Temporally Correlated Data Fields in

Wireless Sensor Networks," *IEEE Transactions on Vehicular Technology*, vol. 58, no. 3, pp. 1583–1594, 2009.

[30] D. A. K. Qiu, "Decentralized Random-Field Estimation for Sensor Networks Using Quantized Spatially Correlated Data and Fusion-Center Feedback," *IEEE Transactions on Signal Processing*, vol. 56, no. 12, pp. 6069–6085, 2008.

[31] C. Carvalho, D. G. Gomes, N. Agoulmine, and J. N. de Souza, "Improving prediction accuracy for WSN data reduction by applying multivariate spatio-temporal correlation," *IEEE Sensors Journal*, vol. 11, no. 11, p. 1001010037, 2011.

[32] R. Cristescu and M. Vetterli, "On the optimal Density for Real-Time Data Gathering of Spatio-Temporal Processes in Sensor Networks," in *Proceedings of 4th International Conference, IPSN 2005*, 2005, pp. 159–164.

[33] Y. Ma, Y. Guo, X. Tian, and M. Ghanem, "Distributed clustering-based aggregation algorithm for spatial correlated sensor networks," *IEEE Sensors Journal*, vol. 11, no. 3, p. 641648, 2011.

[34] N. Li, Y. Liu, F. Wu, and B. Tang, "WSN Data Distortion Analysis and Correlation Model Based on Spatial Locations," *Journal of Networks*, vol. 5, no. 12, 2010.

[35] J. Yuan and H. Chen, "The optimized clustering technique based on spatial-correlation in wireless sensor networks," in *Proceedings of IEEE Youth Conf. Inf., Comput. Telecommun. YC-ICT*, 2009, p. 411414.

[36] J. O. Berger, V. D. Oliveira, and B. Sanso, "Objective Bayesian Analysis of Spatially Correlated Data," *JOURNAL OF THE AMERICAN STATISTICAL ASSOCIATION*, vol. 96, pp. 1361–1374, 2000.

[37] M. Zahraie, A. Farkhady, and A. Haghighat, "Increasing Network Lifetime by Optimum Placement of Sensors in Wireless Sensor Networks," in *Proceedings of*

11th International Conference on the Computer Modelling and Simulation, UKSIM '09, 2009, pp. 611–616.

[38] C. Liu, K. Wu, and J. Pei, "A Dynamic Clustering and Scheduling Approach to Energy Saving in Data Collection from Wireless Sensor Networks," in *Proceedings of Second Annual IEEE Communications Society Conf. on Sensor and Ad Hoc Communications and Networks, IEEE SECON 2005*, 2005, pp. 374–38.

[39] A. Gallais, J. Carle, D. Simplot-Ryl, and I. Stojmenovic, "Localized Sensor Area Coverage with Low Communication Overhead," *IEEE Transactions on Mobile Computing*, vol. 7, no. 5, pp. 661–672, 2008.

[40] Y.-R. Tsai, "Sensing coverage for randomly distributed wireless sensor networks in shadowed environments," *IEEE Transactions on Vehicular Technology*, vol. 57, no. 1, pp. 556–564, 2008.

[41] A. Elfes, "Occupancy grids: a stochastic spatial representation for active robot perception," in *Autonomous Mobile Robots: Perception, Mapping and Navigation, IEEE Computer Society Press*, 1991, pp. 60–70.

[42] A. Hossain, S. Chakrabarti, and P. K. Biswas, "Node Placement, Sensing Coverage and Information Generation in a Wireless Sensor Network," in *Proceedings of 3rd International Conference on Wireless Communication and Sensor Networks (WCSN-2007)*, 2007, pp. 13–15.

[43] A. Hossain, P. K. Biswas, and S. Chakrabarti, "Sensing Models and Its impact on Network Coverage in Wireless Sensor Network," in *Proceedings of 2008 IEEE Region 10 Colloquium Conference*, 2008, pp. 1–5.

[44] J. Q. G. Z. Z. P. G. CM, "Overview of MAC protocols in wireless sensor networks," *Journal of Software*, vol. 19, no. 2, pp. 389–403, 2008.

[45] W. Ye, J. Heidemann, and D. Estrin, "An energy-efficient MAC protocol for wireless sensor networks," in *Proceedings of IEEE Twenty-First Annual Joint Conference of Computer and Communications Societies, INFOCOM 2002*, vol. 3, 2002, pp. 1567–1576.

[46] "Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specification,," IEEE Std. 802.11, 2007.

[47] G. Lu, B. Krishnamachari, and C. Raghavendra, "An adaptive energy-efficient and low-latency MAC for data gathering in wireless sensor networks," in *Proceedings of 18th International Parallel and Distributed Processing Symposium*, 2004.

[48] T. van Dam and K. Langendoen, "An adaptive energy-efficient MAC protocol for wireless sensor networks," in *Proceedings of the 1st international conference on Embedded networked sensor systems, SenSys '03*, 2003, pp. 171–180.

[49] P. Lin, C. Qiao, and X. Wang, "Medium access control with a dynamic duty cycle for sensor networks," in *Proceedings of IEEE Wireless Communications and Networking Conference, WCNC 2004*, vol. 3, 2004, pp. 1534–1539.

[50] T. Zheng, S. Radhakrishnan, and V. Sarangan, "PMAC: an adaptive energy-efficient MAC protocol for wireless sensor networks," in *Proceedings of IEEE 19th Int. Parallel and Distributed Processing Symposium*, 2005.

[51] Y. Sun, S. Du, O. Gurewitz, and D. B. Johnson, "DW-MAC: a low latency, energy efficient demand-wakeup MAC protocol for wireless sensor networks," in *Proceedings of the 9th ACM international symposium on Mobile ad hoc networking and computing*, ser. MobiHoc '08, 2008, pp. 53–62.

[52] J. Polastre, J. Hill, and D. Culler, "Versatile low power media access for wireless

sensor networks," in *Proceedings of the 2nd international conference on Embedded networked sensor systems*, ser. SenSys '04, 2004, pp. 95–107.

[53] A. Ephremides and O. Mowafi, "Analysis of a Hybrid Access Scheme for Buffered Users-Probabilistic Time Division," *IEEE Transactions on Software Engineering*, vol. 8, no. 1, pp. 52–61, 1982.

[54] I. Rhee, A. Warrier, M. Aia, J. Min, and M. Sichitiu, "Z-MAC: A Hybrid MAC for Wireless Sensor Networks," *IEEE/ACM Transactions on Networking*, vol. 16, no. 3, pp. 511–524, 2008.

[55] Z. Merhi, M. Elgamel, and M. Bayoumi, "Eb-mac: An event based medium access control for wireless sensor networks," in *Proceedings of IEEE Int. Conference on Pervasive Computing and Communications, PerCom 2009*, 2009, pp. 1–6.

[56] Y. Zhou and M. Medidi, "Energy-efficient contention-resilient medium access for wireless sensor networks," in *Proceedings of IEEE Int. Conference on Communications, ICC '07.*, 2007, pp. 3178–3183.

[57] K. Jamieson, H. Balakrishnan, and Y. C. Tay, "Sift: A MAC Protocol for Event-Driven Wireless Sensor Networks," in *Proceedings of the 3rd European Workshop on Wireless Sensor Networks (EWSN06)*, 2006, pp. 260–275.

[58] M. Zhao, Z. Chen, and Z. Ge, "QS-Sift: QoS and spatial correlation-based medium access control in wireless sensor networks," *International Journal of Sensor Networks*, vol. 2, no. 3, pp. 228–234, 2007.

[59] M. Zhao, Z. Chen, Z. Ge, and L. Zhang, "HS-Sift: a Hybrid Spatial Correlation-based MAC for Event-driven Wireless Sensor Networks," in *Proceedings of Int. Conf. on Communications and Networking in China, ChinaCom '06*, 2006, pp. 1–5.

[60] F. Bouabdallah, N. Bouabdallah, and R. Boutaba, "Efficient reporting node selection-based MAC protocol for wireless sensor networks," *Wireless Networks*, vol. 19, no. 3, pp. 373–391, 2013.

[61] A. Manjeshwar and D. Agrawal, "TEEN: a routing protocol for enhanced efficiency in wireless sensor networks," in *Proceedings of 15th Int. Parallel and Distributed Processing Symposium*, 2001, pp. 2009–2015.

[62] ——, "APTEEN: a hybrid protocol for efficient routing and comprehensive information retrieval in wireless," in *Proceedings of Int. Parallel and Distributed Processing Symposium, IPDPS 2002*, 2002, pp. 14–17.

[63] F. Bajaber and I. Awan, "Dynamic/Static Clustering Protocol for Wireless Sensor Network," in *Proceedings of Second UKSIM European Symposium on Computer Modeling and Simulation, EMS '08*, 2008, pp. 524–529.

[64] S. D. Muruganathan, D. C. Ma, R. I. Bhasin, and A. O. Fapojuwo, "A centralized energy-efficient routing protocol for wireless sensor networks," *Comm. Mag.*, vol. 43, no. 3, pp. 8–13, 2005.

[65] S. Tarannum, S. Srividya, D. S. Asha, R. Padmini, L. Nalini, K. R. Venugopal, and L. Patnaik, "Dynamic hierarchical communication paradigm for Wireless Sensor Networks : A centralized, energy efficient approach," in *Proceedings of IEEE Int. Conf. on Communication Systems, ICCS 2008*, 2008, pp. 959–963.

[66] H. Soude and J. Mehat, "Energy Efficient Clustering Algorithm for Wireless Sensor Networks," in *Proceedings of Int. Conf. on Wireless and Mobile Communications, ICWMC '06*, 2006, pp. 7–8.

[67] M. Zhang, C. Gong, Y. Feng, C. Liu, and H. Feng, "A Novel Energy-Efficient Dynamic Voting Cluster Algorithm for Wireless Sensor Networks," in *Proceed-*

ings of Int. Conf. on Networks Security, Wireless Communications and Trusted Computing, NSWCTC '09, vol. 2, 2009, pp. 631–634.

[68] O. Younis and S. Fahmy, "HEED: a hybrid, energy-efficient, distributed clustering approach for ad hoc sensor networks," *IEEE Transactions on Mobile Computing*, vol. 3, no. 4, pp. 366–379, 2004.

[69] H. Chan and A. Perrig, "ACE: An Emergent Algorithm for Highly Uniform Cluster Formation," in *Proceedings of the First European Workshop on Sensor Networks (EWSN*, 2004, pp. 154–171.

[70] M. Qin and R. Zimmermann, "VCA: An Energy-Efficient Voting-Based Clustering," *Journal of Universal Computer Science*, vol. 13, 2007.

[71] D. Kumar, T. C. Aseri, and R. Patel, "EEHC: Energy efficient heterogeneous clustered scheme for wireless sensor networks," *Computer Communications Journal*, vol. 32, no. 4, pp. 662 – 667, 2009.

[72] W. R. Heinzelman, A. Chandrakasan, and H. Balakrishnan, "Energy-Efficient Communication Protocol for Wireless Microsensor Networks," in *Proceedings of the 33rd Hawaii International Conference on System Sciences - Volume 8*, 2000, pp. 1–10.

[73] S. Bandyopadhyay and E. Coyle, "An energy efficient hierarchical clustering algorithm for wireless sensor networks," in *Proceedings of Twenty-Second Annual Joint Conference of the IEEE Computer and Communications, INFOCOM 2003*, vol. 3, 2003, pp. 1713–1723.

[74] V. Loscri, G. Morabito, and S. Marano, "A two-levels hierarchy for low-energy adaptive clustering hierarchy (TL-LEACH)," in *Proceedings on IEEE Vehicular Technology Conference, VTC-2005*, vol. 3, 2005, pp. 1809–1813.

[75] A. Youssef, M. Younis, M. Youssef, and A. Agrawala, "Distributed Formation of Overlapping Multi-hop Clusters in Wireless Sensor Networks," in *Proceedings of IEEE Conf. on Global Telecommunications Conference, GLOBECOM '06*, 2006, pp. 1–6.

[76] Y. Jin, L. Wang, Y. Kim, and X. Yang, "EEMC: An energy-efficient multi-level clustering algorithm for large-scale wireless sensor networks," *Comput. Netw.*, vol. 52, no. 3, pp. 542–562, 2008.

[77] S. Selvakennedy, S. Sinnappan, and Y. Shang, "A biologically-inspired clustering protocol for wireless sensor networks," *Comput. Commun.*, vol. 30, no. 14-15, pp. 2786–2801, 2007.

[78] J. Kamimura, N. Wakamiya, and M. Murata, "A Distributed Clustering Method for Energy-Efficient Data Gathering in Sensor Networks," pp. 113–120, 2006.

[79] C.-Y. Wen and W. A. Sethares, "Automatic decentralized clustering for wireless sensor networks," *EURASIP J. Wirel. Commun. Netw.*, vol. 2005, no. 5, pp. 686–697, 2005.

[80] L. Yu, N. Wang, W. Zhang, and C. Zheng, "GROUP: A Grid-Clustering Routing Protocol for Wireless Sensor Networks." in *Proceedings of Int. Conf. on Wireless Communications, Networking and Mobile Computing, WiCOM '06*, 2006, pp. 1–5.

[81] H. Zhang and A. Arora, "GS3: scalable self-configuration and self-healing in wireless sensor networks," *Comput. Netw.*, vol. 43, no. 4, pp. 459–480, 2003.

[82] J. Li and G. AlRegib, "Energy-efficient Cluster-based Distributed Estimation in Wireless Sensor Networks," in *Proceedings of the 2006 IEEE Conference on Mil-*

itary Communications, ser. MILCOM'06. Piscataway, NJ, USA: IEEE Press, 2006, pp. 600–606.

[83] A. Dabirmoghaddam, M. Ghaderi, and C. Williamson, "Cluster-Based Correlated Data Gathering in Wireless Sensor Networks," in *Proceedings of 2010 IEEE Int. Symposium on Modeling, Analysis Simulation of Computer and Telecommunication Systems (MASCOTS),*, 2010, pp. 163–171.

[84] R. K. Shakya, Y. N. Singh, and N. K. Verma, "A novel spatial correlation model for wireless sensor network applications," in *Proceedings of the Ninth IEEE International Conference on Wireless and Optical Communications Networks (IEEE WOCN-2012)*, 2012, pp. 1–6.

[85] ——, "A correlation model for MAC protocols in event-driven wireless sensor networks," in *Proceedings of the 2012 IEEE Region 10 TENCON Conference*, 2012, pp. 1–6.

[86] M. R. Garey and D. S. Johnson, *Computers and Intractability; A Guide to the Theory of NP-Completeness*. New York, NY, USA: W. H. Freeman & Co., 1990.

[87] R. Bhatia, *Matrix Analysis*. New York:Springer-Verlag, Graduate Texts in Mathematics, 1991, vol. 169.

[88] S. M. Kay, *Fundamentals of Statistical Signal Processing, Vol. I: Estimation Theory*. Prentice-Hall, 1993.

[89] Y. C. Tay, K. Jamieson, and H. Balakrishnan, "Collision-minimizing CSMA and its applications to wireless sensor networks," *IEEE J. Select. Areas Commun.*, vol. 22, no. 6, pp. 1048–1057, 2004.

[90] R. K. Shakya, Y. N. Singh, and N. K. Verma, "Generic Correlation Model for

Wireless Sensor Network Applications," *IET Wireless Sensor Systems*, vol. 3, no. 4, pp. 266–276, 2013.

[91] Y. Linde, A. Buzo, and R. Gray, "An Algorithm for Vector Quantizer Design," *IEEE Transactions on Communications*, vol. 28, no. 1, pp. 84–95, 1980.

[92] J. Polastre, J. Hill, and D. Culler, "Versatile low power media access for wireless sensor networks," in *Proceedings of the 2th international conference on Embedded networked sensor systems*, 2004, pp. 95–107.

[93] C.-Y. Lin, Y.-C. Tseng, and T.-H. Lai, "Exploiting Spatial Correlation at the Link Layer for Event-driven Sensor Networks," *International Journal of Sensor Networks*, vol. 1, no. 4, pp. 197–212, 2012.

[94] J. Ghaderi, T. Ji, and R. Srikant, "Connection-level scheduling in wireless networks using only MAC-layer information," in *Proceedings of IEEE International Conference on Computer Communications, INFOCOM, 2012*, 2012, pp. 2696–2700.

[95] *The Network Simulator – ns-2*. [Online] Available: http://www.isi.edu/nsnam/ns/.

[96] L. Karim and N. Nasser, "Reliable location-aware routing protocol for mobile wireless sensor network," *IET Communications*, vol. 6, no. 14, pp. 2149–2158, 2012.

[97] W.-Y. Chiu, B.-S. Chen, and C.-Y. Yang, "Robust Relative Location Estimation in Wireless Sensor Networks with Inexact Position Problems," *IEEE Transactions on Mobile Computing*, vol. 11, no. 6, pp. 935–946, 2012.

[98] K. Ota, M. Dong, H. Zhu, S. Chang, and X. Shen, "Traffic information prediction in Urban Vehicular Networks: A correlation based approach," in *Proceedings of*

IEEE Wireless Communications and Networking Conference, WCNC-2011, 2011, pp. 1021–1025.

[99] G. Shah and M. Bozyigit, "Exploiting Energy-aware Spatial Correlation in Wireless Sensor Networks," in *Proceedings of 2nd Int. Conf. on Communication Systems Software and Middleware, COMSWARE 2007*, 2007, pp. 1–6.

[100] M. Xiang, Z. Luo, and P. Wang, "Energy-Efficient Intra-Cluster Data Gathering of Wireless Sensor Networks," *Journal of Networks*, vol. 5, no. 3, 2010.

[101] T. Gao, R. Jin, J. Song, T. Xu, and L. Wang, "Energy-Efficient Cluster Head Selection Scheme Based on Multiple Criteria Decision Making for Wireless Sensor Networks," *Wireless Personal Communications*, vol. 63, no. 4, pp. 871–894, 2012.

[102] R. K. Tripathi, Y. N. Singh, and N. K. Verma, "N-LEACH, a balanced cost clusterheads selection algorithm for Wireless Sensor Network," in *Proceedings of National Conference on Communications (NCC-2012),*, 2012, pp. 1–5.

[103] *MATLAB 7.13, (R2011b)*. The MathWorks Inc., Natick, MA 01760-2098, 2011.

[104] *Introduction to Network Simulator NS2*. Springer Science+Business Media, LLC, 233 Spring Street, NY 10013, USA, 2009.

[105] "The CMU Monarch projects wireless and mobility extensions to ns," (1999), CMU Monarch project, Computer Science Department, Canergie Mellon University, Pittsburgh.

[106] *The ns Manual*. [Online] Available: http://www.isi.edu/nsnam/ns/ns-documentation.html.

[107] T. S. Rappaport, *Wireless Communications: Principle and Practice*. IEEE press Piscataway, NJ, USA, 1996.

[108] S. Du, A. Saha, and D. Johnson, "RMAC: A Routing-Enhanced Duty-Cycle MAC protocol for Wireless Sensor Networks," in *Proceedings of 26th IEEE International Conference on Computer Communications (INFOCOM '07)*, 2007, pp. 1478–1486.

[109] Y. Sun, O. Gurewitz, and D. B. Johnson, "RI-MAC: a receiver-initiated asynchronous duty cycle MAC protocol for dynamic traffic loads in wireless sensor networks," in *Proceedings of the 6th ACM conference on Embedded network sensor systems*, 2008, pp. 1–14.

List of Publications

International Journal Publications (peer-reviewed)

[A1] Rajeev K. Shakya, Yatindra N. Singh, Nishchal Verma, "Generic Correlation Model for Wireless Sensor Network Applications," *IET Wireless Sensor Systems journal,* vol. 3, No. 4, pp. 266–276, Dec 2013.

[A2] Rajeev K. Shakya, Yatindra N. Singh, Nishchal Verma, "Event-MAC: A localized, Spatial correlation based collision-free MAC for Event-driven Wireless Sensor Networks," *revision submitted in IEEE/ACM transactions on Networking, 2014.*

International Conference Publications (peer-reviewed)

[B1] Rajeev K. Shakya, Yatindra N. Singh, Nishchal Verma, "A Novel Spatial Correlation Model for Wireless Sensor Network Applications," *in Proc. of the Ninth IEEE International Conference on Wireless and Optical Communications Networks (IEEE WOCN-2012), Indore, INDIA,* Sept., 20-22, pp. 1–6, 2012.

[B2] Rajeev K. Shakya, Yatindra N. Singh, Nishchal Verma, "A Correlation Model for MAC Protocols in Event-Driven Wireless Sensor Networkss," *in Proc. of the 2012 IEEE Region 10 Conference, IEEE TENCON-2012, Cebu, Philippines,* Nov 19-22, pp. 1–6, 2012.

भारतीय प्रौद्योगिकी संस्थान कानपुर
INDIAN INSTITUTE OF TECHNOLOGY KANPUR