SPATIO-TEMPORAL PROBABILISTIC QUERY GENERATION MODELS AND SINK-ATTRIBUTES ANALYSIS IN ENERGY EFFICIENT WIRELESS SENSOR NETWORKS

Thesis

Submitted in partial fulfillment of the requirement for the degree of

DOCTOR OF PHILOSOPHY

By

PRAMOD KUMAR

DEPARTMENT OF ELECTRICAL AND ELECTRONICS ENGINEERING

NATIONAL INSTITUTE OF TECHNOLOGY KARNATAKA

SURATHKAL, MANGALORE-575025

APRIL, 2017

DECLARATION

By the Ph.D. Research Scholar

I hereby declare that the Research Thesis entitled **“SPATIO-TEMPORAL PROBABILISTIC QUERY GENERATION MODELS AND SINK-ATTRIBUTES ANALYSIS IN ENERGY EFFICIENT WIRELESS SENSOR NETWORKS”** which is being submitted to the **National Institute of Technology Karnataka, Surathkal** in partial fulfillment of the requirements for the award of the Degree of **Doctor of Philosophy** in **Department of Electrical and Electronics Engineering** is a **bonafide report of the research work carried out by me**. The material contained in this Research Thesis has not been submitted to any University or Institution for the award of any degree.

PRAMOD KUMAR
(090698EE09P01)
Department of Electrical and Electronics Engineering

Place: NITK- Surathkal
Date:

CERTIFICATE

This is to certify that the Research Thesis entitled **"SPATIO-TEMPORAL PROBABILISTIC QUERY GENERATION MODELS AND SINK-ATTRIBUTES ANALYSIS IN ENERGY EFFICIENT WIRELESS SENSOR NETWORKS"** submitted by **PRAMOD KUMAR**, (Register Number: 090698EE09P01) as the record of the research work carried out by him, is accepted as the Research Thesis submission in partial fulfillment of the requirements for the award of degree of **Doctor of Philosophy.**

Dr. Ashvini Chaturvedi
(Research Guide)

Dr. Vinatha U.
(Chairperson – DRPC)

ACKNOWLEDGMENTS

I would like to express sincere gratitude to my guide Dr. Ashvini Chaturvedi, Associate Professor, Department of Electronics and Communication Engineering, for giving me an opportunity to work under his guidance which is invaluable. His unflinching support, suggestions, directions have helped in smooth progress of the Ph.D work. He has been a constant source of inspiration in all possible ways for successful completion of my research work.

I am extremely grateful to my beloved HOD Dr. Vinatha U, Department of Electrical and Electronics Engineering, National Institute of Technology Karnataka, Surathkal for his encouragement and providing me with sufficient computational facilities to successfully complete the research work. I would also like to express my deepest gratitude to research progress assessment committee members, Prof. G. Ram Mohana Reddy, Dept. of Information Technology and M.N. Satyanarayan, Associate Professor, Department of Physics for their valuable guidance, suggestions, and support throughout my research work.

I express my heartfelt thanks to all the teaching and non-teaching staff of the Department of Electrical and Electronics Engineering for full co-operation and assistance. I also extend my thanks to the library staff.

Its my pleasure to thank my Father Sri Bindhyachal Singh, Mother Ram Dulari Devi and my family members for their support, encouragement and love they gave me. They are the one who kept me on high spirits in hard times; especially I would like to thank my wife, Puja Kumari for supporting and inspiring me at all times. And to my children Nyasha, Preksha and Neelesh who inspires me a lot.

I thank to Dr. Ujjwal Verma, Asst Professor, MIT, Manipal who helped me a lot in Latex to compile the thesis and all my friends who helped me directly and indirectly with discussion and useful timely suggestions. I praise and glorify the name of God, the Creator who creates all these nice people and these pleasant opportunities.

PRAMOD KUMAR

ABSTRACT

Rapid advancement in Micro-Electronic-Mechanical-Systems (MEMS) and distributed computing infrastructure along with compactness and economic viability in IC technology has accelerated the versatile growth and deployment of wireless sensors networks (WSNs). Owing to its sensing and subsequent parameters estimation abilities while maintaining higher spatial resolution led to a prominent position for WSNs in the networking paradigm. The main contribution of this thesis is to provide a modeling and analysis based on probabilistic framework for varieties of query generation scenarios in WSNs. In broad regime of query based WSNs, the query generation dynamics owe significantly to the associated spatial and/or temporal parameters. To encompass varying degree of uncertainties associated with spatio-temporal parameters; these parameters are treated as Fuzzified intervals, thus address quantum of uncertainties with the finest approximation or accuracy. Further, to ensure reliable network operation having fairly uniform coverage over the stipulated lifetime; usage of varieties of clustering schemes, sink attributes and spatial-fusion concept are explored.

From operational aspects; the most important issues in wireless sensor networks (WSNs) are the coverage and the efficient usage of sensor nodes limited energy reserve. Irrespective of the applications served; owing to difficulty associated with battery replenishment, proper energy usage has been at centre-stage in WSNs operations, that ultimately influence the lifetime of WSNs. In this thesis; during various case-studies, square shape service-areas of varying area-dimension, different sensor node specifications and hierarchical network architecture are considered. These cases differ in terms of sink-attributes such as single/multiple, and stationary/portable. Sink is an important interface between the end user and the remote entities (sensor nodes), thus the proposed algorithms are formulated on considering sink at the center-stage. Further, the performance of WSNs also depends upon the topological structure of the network, usually it is hierarchical one. To realize hierarchical structure; the usage of clustering schemes namely static k-means (SKM), static fuzzy c-means (SFCM), dynamic k-means (DKM) and dynamic fuzzy c-means (DFCM) are explored for the clusters formation and the subsequent selection of cluster heads (CHs).

Selection of CHs is done using residual energy status (RES) of participating sensor nodes. Cartesian-coordinate of these sensor nodes appropriately weighted with RES estimate decides the energy-centroid (EC) location for each cluster. The Euclidean distance measure between a sink and the ECs is used to identify new appropriate location for sink while complying with principal motive of energy conservation. Initially, during few query generation scenarios; different quadrants of the service-area observe distinct pattern of query spatial distribution. These query dissemination patterns are modelled using amplitude and angle modulated vectors. Later, the probabilistic approach is used to model the query generation process. The parameters of probabilistic models are regulated using the associated spatio-temporal aspects.

In this thesis; usage of uniform probability mass function (PMF) and Poisson PMF models are investigated and analyzed to replicate query generation process. Further, the lifetime of WSNs depends upon sensor nodes energy dissipation pattern that is non-uniform in terms of spatial distribution over any short epochs. Thus, integrating the spatio-temporal aspects with the Poisson PMF model appears more reasonable.

In mainstream probabilistic models; the associated control parameters are treated as crisp numbers, which fail to encompass uncertainties associated with the modelled parameters. To incorporate these uncertainties, Fuzzified interval-bound values of spatio-temporal parameters are considered to model the control parameter in Poisson PMF expression. Further, exploring the solutions in higher-dimensional space always entrust its superiority over the solutions that are derived from lower-dimensional space. This rationale is exploited using the concept of spatial (quadrants)-fusion in anticipation of improved profile of network performance measures.

With these motivations, the thesis explores: (1) uses of energy efficient clustering schemes, (2) incorporation of spatio-temporal parameters uncertainties into probabilistic model of query generation using fuzzy-intervals bound, (3) importance of sink attributes, and (4) exploiting heuristic framework based on spatial-fusion concept to enhance network lifetime or to meet desired service norms over the stipulated network lifetime. For various network surveillance scenarios; the performance measures namely average residual energy status (ARES) of entire sensor network, critical residual energy status (CRES) of individual nodes as well as that of entire sensor network, fraction of sensor nodes attaining CRES mark and the network service-time-duration (STD) are estimated and analyzed.

Key words:-Average Residual Energy Status (ARES), K-means, Fuzzy c-means (FCM), Critical Residual Energy Status (CRES), Service Time Duration (STD), Poisson distribution, Gaussian Distribution, Spatio-temporal, Fusion, Fuzzy Intervals, Wireless Sensor Networks(WSNs).

DEDICATED TO

My

Parents

&

Family

Contents

List of Figures

List of Tables

Nomenclature

Symbol	Meaning
λ	Control Parameter of Poisson PMF
ω	Weight factor
d	Euclidean distance
J	Joule
$\boldsymbol{E}$	Energy vector
$\boldsymbol{e}$	Energy of each node
μ	Degree of belongingness
θ	Angle of a phasor
K	Iteration count
α	Spatial parameter
β	Temporal parameter

Abbreviations

MAC	Media Access Control
QoS	Quality of Service
FDMA	Frequency Division Multiple Access
TDMA	Time Division Multiple Access
RTS	Right to Send
CTS	Clear to Send
ACK	Acknowledgement
LEACH	Low Energy Adaptive Clustering Hierarchy
LMAC	Light Weight Medium Access
UDG	Unit Disk Graph
RES	Residual Energy Status
FCM	Fuzzy c-means
CH	Cluster Head
EC	Energy-Centroid
SN	Sensor Node
IER	Initial Energy Reserve
CRES	Critical Residual Energy Status
SSS	Single stationary sink

SPS	Single Portable Sink
FSS	Four Stationary Sinks
SKM	Static k-means
SFCM	Static Fuzzy c-means
DKM	Dynamic k-means
DFCM	Dynamic Fuzzy c-means
PMF	Probability Mass Function
FOI	Finite Observation Interva
AM	Arithmetic Mean
GM	Geometric Mean
AMI	Arithmetic Mean Index
GMI	Geometric Mean Index
SL	Lower Support Interval
SU	Upper Support Interval
AMDI	Arithmetic Mean Deviation Index
GMDI	Geometric Mean Deviation Index
FCI	Fusion Count Index
OMH	Operational Mechanism and Heuristic
STPI	Spatio-temporal-Plane-Interval
STFI	Spatio-temporal-Fuzzy-Interval

Chapter 1

OVERVIEW OF WIRELESS SENSOR NETWORKS

1.1 Introduction

Wireless Sensor Networks (WSNs) contain several tens or hundreds of tiny sensor nodes. In technical literature, these sensor nodes are also known as motes and are capable of communicating with each others or directly to an external base-station (BS)/sink node. Compared to other conventional networks, relatively higher densities of these sensors (motes) facilitate job of sensing over larger geographical regions with greater accuracy. Figure 1.1 shows the functional schematic diagram of sensor node architecture. Basically, each sensor node comprises sensing, processing, transmission, mobilizer, position finding system, and power units (some of these components are optional like the mobilizer and position finding system). Figure 1.1 also shows the flow of data communication in WSN using directional arrows. Usually, the sensor nodes are laid down in a scattered manner in the given service-area. Sensor nodes coordinate among themselves to produce high-quality spatio-temporal information about the physical environment. Each of these scattered sensor nodes has the capability to collect and route data either to other sensors or back to an external sink(s). Sink nodes may be stationary/portable/mobile and are capable of connecting the sensor network to the outside world using the internet gateways, thus felicitate a ubiquitous access to users about the measured data [(Karaki and Kamal, 2004)]. In terms of topological structure, the sensor network architecture can be classified into following two types:

1. Flat architecture

2. Hierarchical architecture

In a flat architecture, sensor network comprises of homogeneous sensor nodes and one or multiple sink node(s).

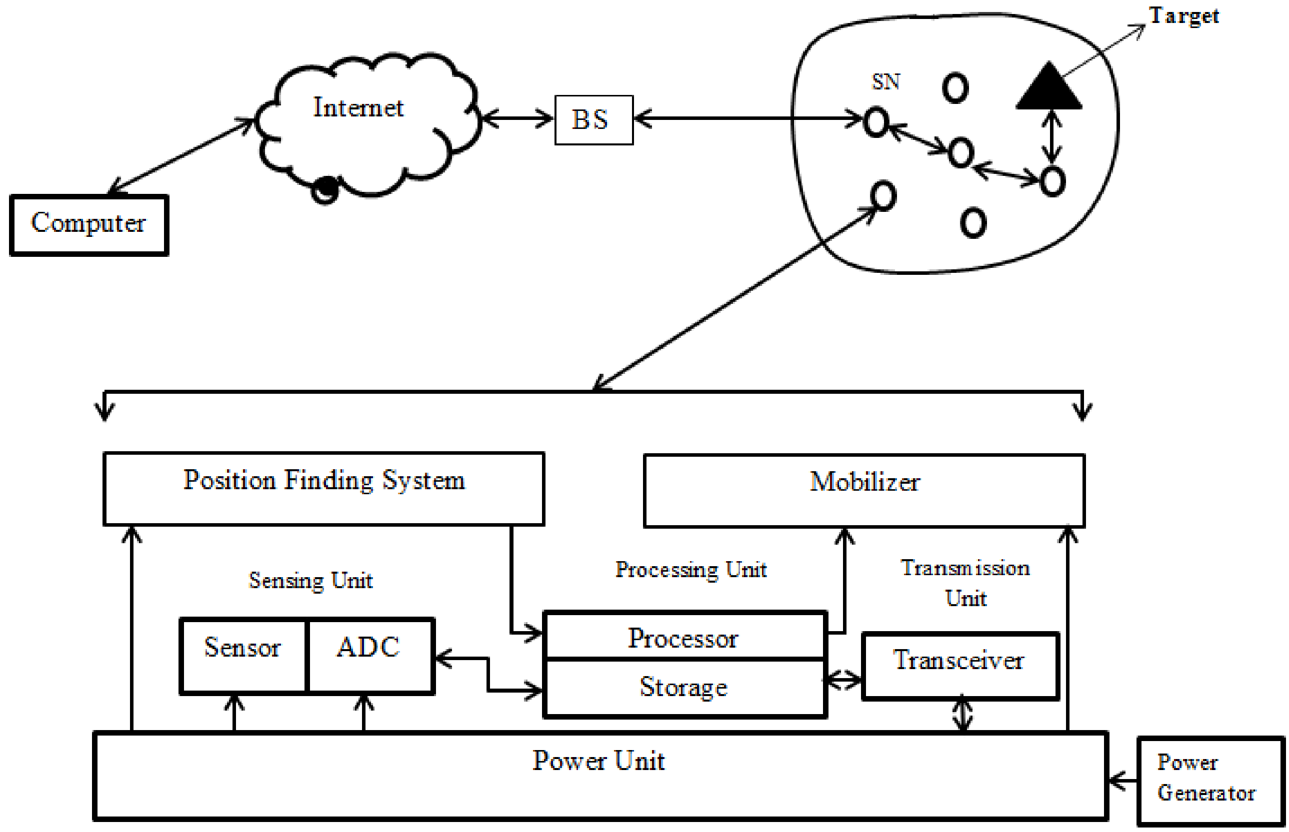


Figure 1.1: The System architecture of a Sensor Node.

In a hierarchical architecture, based on multimodal sensing attributes; sensor nodes can be homogeneous/heterogeneous/mixed in nature. Usually, hierarchical topology based sensor networks deploy clustering mechanism; and comprises of simple sensor nodes, cluster heads and sink nodes(s) as its basic building blocks. In both of these architectures the sink node(s) can be stationary or mobile. A schematic for hierarchical sensor network architecture is shown in Figure 1.2; it consists of a group of clusters (here cluster is a smaller sub-region within a service-area that possess few sensor nodes and for each cluster a particular sensor node is deputed as cluster head (CH). The inter cluster bidirectional communication is performed by cluster heads (CHs) Within a cluster, nodes communicate with each other and establish a connection with CH using single/multiple hop wireless links. The queries generated by the sink make its way through CHs to field deployed sensor nodes. In response to it sensor nodes convey the measured attributes of physical variables via CHs to the sink.

1.1.1 Sensor Network Constraints and Challenges

To sustain the desired service norms in WSNs regime, one has to deal with several challenges as the sensor network architecture has many inherent limitation or constraints. Invariably, most of the WSNs have the following features:

- Sensor nodes are densely deployed in a given service-area, it requires careful strategy to deal with issues such as packets collision and congestion control.
- For outdoor surveillance applications; owing to harsh environmental conditions, sensor nodes face stringent failure threats.
- Deployment of sensors in hostile and/or inaccessible service-area leads to situations wherein replenishing or replacing the battery happens to be a difficult task, thus energy conservation remains a primary parameter of interest.
- Owing to energy reserve constraint and relatively large overhead associated with addressing scheme, the global indexing of sensor nodes is a complex task.
- Network topology changes frequently, this in turn demands dynamic strategies for reliable service operations.

  To comply with all these features, any careful network design strategy must include the following considerations:

- Wide range of nodes density: Depending upon the nature of the service-area namely indoor or outdoor to be served; number of sensor nodes deployed varies from few tens to several hundreds. Thus, scalability is a major concern and must be given due consideration during algorithm design phase.
- Energy reserve constraint: Owing to limited battery reserve, the energy consumed during sensing, data communication and communication must be handled with utmost care so as to maintain requisite requirement of network coverage and longevity of sensor network.
- Data Aggregation: High density of sensor nodes could lead to congestion problem within the network. However, reasonably good spatio-temporal correlation among sensor nodes manifests the uses of data aggregation or fusion techniques.
- Network self-organization: Network nodes are always vulnerable to life threat owing to associated finite energy supply and hostile ambient conditions. Thus, few sensor nodes may encounter failure state while to supplement extended service-area surveillance, new nodes may join the network. To accommodate this kind of nodes

dynamics, the network must possess self organizing capability. Need based it could be exercised periodically or sporadically.

- Collaborative signal processing: Contrary to the mobile ad hoc networks (MANETs), the principal task in sensor networks is detection of events or measurements of physical parameters of interest, not merely exchange of information among network entities. To accomplish reasonably good detection performance, sensed data from multiple sensors are fused. Data fusion operation requires the transmission of data and control (overhead information in a form of time-stamped and location-indexed data) messages.

- Querying ability: There are two types of addressing modes in sensor network; data-centric, and address centric. In data-centric mode; a query is sent to all network nodes, in response to it only a few nodes that comply with query's attributes establish communication links with sink. In address-centric mode; queries are sent to a set of few sensor nodes that are positioned at particular locations within a service-area [(Intanagonwiwat et al., 2002)]. Many protocols and algorithms have been proposed for wireless ad hoc networks. However, these algorithms do not suit to the specific requirements of WSNs regime and thus needs to be altered.

Wireless sensor networks are designed and deployed to cater applications specific needs, in general WSNs architecture have some inherent advantages, which are listed as:

- WSNs impart an improved accuracy by distributed processing of measured sensing information that could be seismic data, acoustic data, high-resolution images, etc. On the basis of mutual cooperation and sharing, these sensor nodes provide enriched information about the sensed data.

- Redundancy in number of sensor nodes deployed bids a fault tolerance characteristic by virtue of which sensor network continues its surveillance operation in a reliable manner.

- Energy consumption pattern is unique as during the short epochs a fraction of network nodes is engaged in sensing and subsequent communication.

- On several occasions; sensor nodes are deployed in hostile and inaccessible terrains that restrict the human intervention. In such situations, the WSN architecture emerges as viable technology.

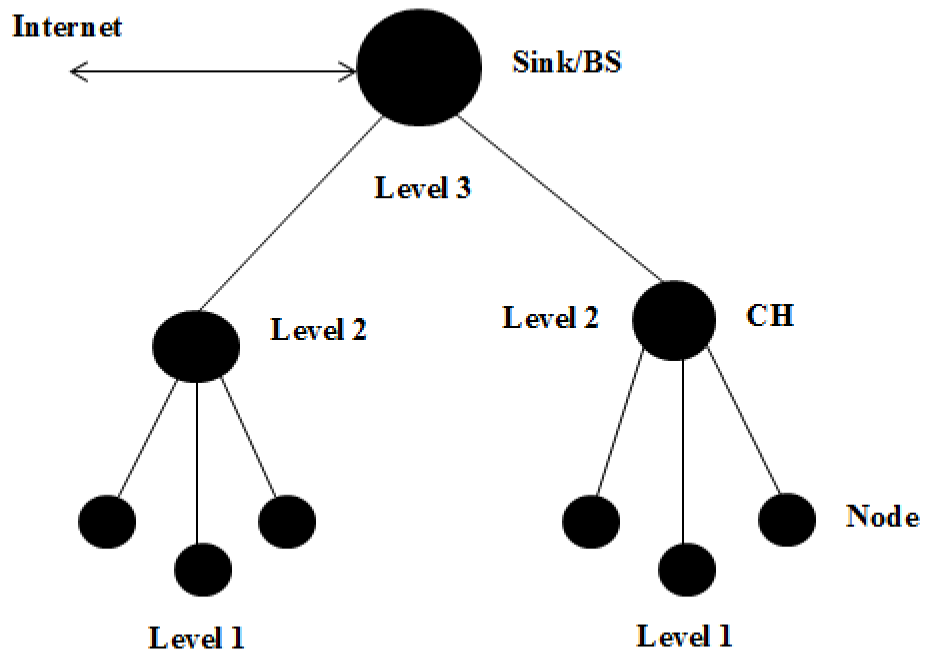


Figure 1.2: Hierarchical Sensor Network Architecture.

1.2 Protocol Stack Architecture for wireless sensor networks

The typical architecture of the protocol stack used in wireless sensor network regime is shown in Fig 1.3. It combines variety of issues such as cooperation among the nodes, inferences drawn on correlation of measured variables, power budget distribution for sensing ,computing and communication, congestion control, route discovery and route maintenance etc. The protocol stack comprises of physical layer, data link layer (media-access control (MAC)), network layer , transport layer, application layer and three management planes which are power management plane, mobility management plane, and task management plane. Varieties of tasks performed at these specific layers and mechanisms exploiting the cross-layer concepts are discussed briefly. The issues such as uses of appropriate modulation schemes, transceiver circuit design issue, and data aggregation are resolved at the physical layer. As the ambient environment is noisy and there could be simultaneous transmission of data packets by neighboring nodes. This conflict leads to collisions among network nodes broadcast. The MAC layer resolves it by minimizing the collisions. The roll of network layer is to find the appropriate routing path, maintaining & updating them. It operates on the data supplied by the transport layer. Depending upon applications specific requirements, the transport layer regulates and maintains the flow of the data packets. To perform different sensing tasks and to meet applications

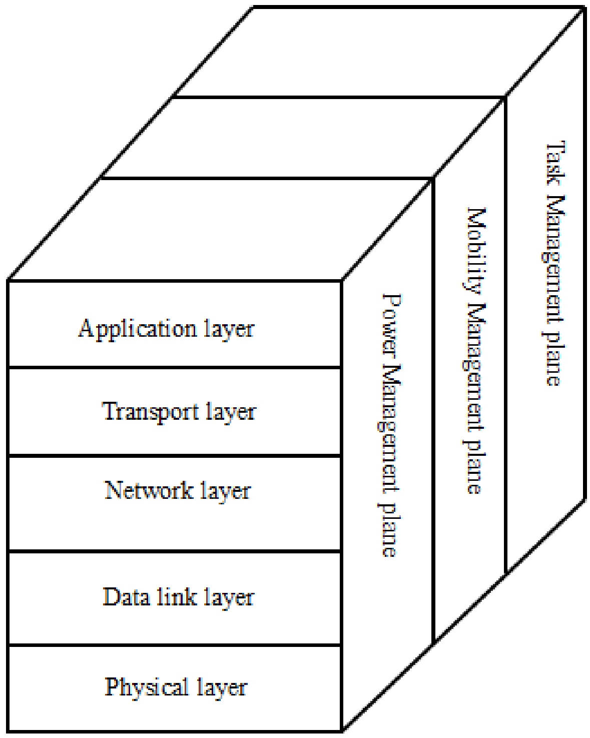


Figure 1.3: Protocol stack of wireless sensor networks

specific needs, different types of application software are used at applications layer. Further, the power, mobility, and task management planes monitor the power consumption, movement and task assignment of network elements respectively. The power management plane administers and observes the manner in which sensor nodes utilize its battery power. For instance periodically or sporadically, it turns off the transmitter or receiver for short epochs. In situations when the residual power of the sensor node is low, the sensor node broadcasts this update to all of its neighbors and thus relinquishes the job of relay node and keeps the remaining power reserve for sensing only. The mobility management plane detects the movements of sensor nodes and utilizes this information to update the routing path details. Mobility of network nodes always imparts frequent changes in network topological structure, thus mobility management plane also facilitates network nodes to keep track of their neighbors. The task management plane takes care for scheduling the tasks in terms of their geographical distribution and timings of its execution. Owing to the high degree of correlation, in a localized area not all the sensors are required to perform the sensing task simultaneously. All the three management planes work harmonically so as to ensure efficient utilization of scarce energy resource by means of dynamic route selection strategy. In aggregate terms good coordination among these management plane results in an energy efficient operation of sensor networks. Brief summary of layers specific protocols is presented next.

1.2.1 Physical-layer protocols

Principal tasks such as identifying appropriate modulation schemes belong to the physical layer. The physical also address the design strategies of simple transceiver circuits. Owing to large scale deployment, cost involved and tiny size of sensor nodes, sensor networks require uses of low power modulation schemes and compact hardware circuits that could be in-house in very small volume. In last two decades, several energy-efficient protocols have been proposed to prolong the network lifetime. In WSNs paradigm estimation of lifetime is a complicated issue since it depends on multifaceted factors such as network architecture, channel characteristics, energy consumption model, data collection initiation and types of protocols used at different layers.

[(Chen and Zhao, 2005)] proposed a general formula to estimate the network lifetime that holds irrespective of the underlying network model. This framework makes use of two physical layer parameters which are the channel state and the residual energy of sensors, both of these are crucial to network lifetime. Different aspects lead to different basis for the physical layer technologies in typical sensor networks paradigm. For instance, based on bandwidth considerations, three different technologies namely, Narrow-band, Spread spectrum and Ultra-Wideband (UWB) exist. In addition to transceiver circuit, sensing and processing unit too have stringent requirements of compact size, low-power and low-cost. A snap-shot of some of the methodologies that have dealt with physical layer issues are presented briefly.[(Hartwell et al., 2007)] presented a scheme that performs estimation of energy consumption associated with frame retransmission in physical layer. The proposed method utilizes estimate of an optimal target SER (symbol error rate) that minimizes overall energy consumption that takes place at physical layer. The optimization framework stimulate a balance between reducing transmit energy that leads to increased SER value and minimizing the energy consumed on retransmissions. In most of the sensor network applications; Quality of service (QoS) requirements presumes error free transmission, which indicates that any frames received erroneously are retransmitted. This appears in a form of a tradeoff while deciding the stipulated SER target. An increase in target SER lowers the signal to noise ratio (SNR) at the receiver and reduces the transmit energy budget. However, there are investigations [(Li et al., 2005),(Cui et al., 2004),(Cui et al., 2005)] which make use of a fixed bit error rate (BER) or SER target. [(Holland et al., 2009)] investigated the problem of energy-efficient data transmission over a noisy channel. In this process, principal emphasis is given to physical layer parameters setting. Further, the authors derived a metric "energy per successfully received bit" which offers an estimate of energy required to transmit a bit successfully over a certain distance under a noisy channel model. Corresponding to uses of different modulation schemes; minimizing this metric results in estimate of the energy-optimal relay distance and the optimal transmit energy having functional dependency on channel noise condition and path loss exponent.

[(Ward and Younis, 2012)] presented a metric to measure sink anonymity in WSNs. Post sink identification, authors considered the possible attacks that an adversary could initiate against a WSN. To counter such malicious attacks, a mechanism is required that ensures network security. One intuitive rational is to increase the anonymity of sink such that it becomes difficult to distinguish a sink from other sensor nodes. So far many techniques have been proposed to improve sink anonymity, however only a few of them addressed it as a physical layer issue. To conquer it authors have proposed an algorithm called "Intercept, correlate and follow" (ICF) to characterize physical layer anonymity in a WSN. In a slightly different perspective,[(Ozdemir et al., 2009)] presented a framework at the physical layer that deals with target tracking problem using particle filters in a WSN. The proposed framework makes use of network physical layer design parameters along with channel imprecision between sensors and the fusion center. The particle filter algorithm is validated for different receiver architectures and the wireless channel models. Finally, the tracking performance is evaluated against the posterior Cramer-Rao lower bounds.

1.2.2 Medium Access Control (MAC)-layer protocols

The MAC layer performs the typical tasks such as providing access control, channel assignment neighbor's update and power control.

In WSNs regime; at sensor node the energy consumption mainly attributes to sensing, data processing and communications. Out of overall energy expenditure, communication (transceiver module) consumes major chunk of it. Transmitted radio signal power in transmission process is controlled by MAC protocols, thus MAC protocols plays an important role in energy consumption of sensor nodes. More carefully spending energy at MAC layer subsequently influences the network lifetime.

Broadly MAC protocols can be classified as contention less and contention-based. Contention less MAC protocols are realized using FDMA or TDMA schemes, whereas contention-based MAC protocols are based on IEEE802.11standard [(IEEE, 2012)].

1.2.2.1 Contention-based MAC protocols

S-MAC: Sensor-MAC [(Ye et al., 2002b),(Ye et al., 2003)] protocols use IEEE 802.11 CSMA/CA mechanism to avoid collision. However, in addition to conventional IEEE

802.11 standard, S-MAC uses the following specific aspects.

1. It avoids idle listening by implementing low-duty-cycle scheme in multi hop networks that results in significant reduction in energy consumption.

2. Avoidance of overhearing by allowing interfering nodes to switch its mode to sleep mode based on RTS or CTS packet.

3. Invoking message passing to reduce amount of control overhead information and data latency. The principal drawback of S-MAC is high value of message delivery latency. It owes to inherent feature of S-MAC that gives priority to energy savings over the latency.

T-MAC: Timeout- MAC protocol [(Dam and Langendoen, 2003)] makes use of synchronization scheme that is used in S-MAC. Sensor nodes communicate with each other using right to send (RTS), clear to send (CTS), DATA and ACK frames that provide collision avoidance and thus ensure reliable data transmission. In comparative terms, energy-efficiency outcome of this scheme is better than that of the S-MAC. However, T-MAC suffers from early sleeping problem that limits its latency and throughput.

PCS-MAC: For WSNs, a power controlled sensor-MAC protocol [(Caylrcl, 2005)] is a distributed power controlled contention-based protocol. It is an advance version of S-MAC protocol having transmission power control capabilities. To conserve the energy, PCS-MAC protocol makes usage of varying power levels while transmitting RTS, CTS, DATA and ACK frames instead of transmitting these frames at rated maximum power. Further, in addition to offering reduced power consumption, PCS-MAC maintains the other silent features which are collision and overhearing avoidance properties.

1.2.2.2 TDMA-based MAC protocols

Time division multiple access (TDMA) scheme has an inherent advantage over the contention-based protocols. Owing to reduced duty-cycle of radio unit and absence of contention-oriented collisions and overhead, synchronization is achieved easily using reservation and scheduling. Though it suffers from a drawback under which it fails to meet scalability property, as it is not as good as that supported by contention-based protocols. [(et al, 2000)] proposed a self-organizing Medium Access Control for sensor networks (SMACS). It operates on a flat topology of sensor networks and is a distributed protocol that enables a set of nodes to determine their neighbors. It also establishes transmission/reception schedules without any intervention of any master node. SMACS too also suffers from some drawbacks which are: it is not a location aware so the estimated neighbors are not necessarily nearest and a node has to wait for its turn to transmit even if the channel is idle. [(Heinzelman et al., 2002)] proposed low-energy adaptive clustering hierarchy (LEACH) protocol, it combines TDMA/CDMA schemes. In a cluster, each node communicates to a dynamically elected cluster by a single-hop link while access is regulated by TDMA scheme. Cluster heads establish information sharing with sink using a CDMA based single-hop wireless link. To avoid energy-well (a localized area having sensor nodes that possess severe shortage of energy) occurrences, LEACH incorporates randomized rotation of high-energy cluster-head. [(van Hoessel and Havinga, 2004)] presented a Light Weight Medium Access (LMAC) protocol for WSNs. It mainly exploits the physical layer properties. The principal motive for designing LMAC is to minimize the usage of number of transceiver switches that leads to limiting the implementation complexity. [(Kim et al., 2015)] presented an overview that contains both MAC-layer and physical layer techniques for sensor network design. In general, physical layer comprises of two main tasks: (i) Energy efficient strategy for transceiver design and (ii) for applications such as event detection and target tracking, design of simple data fusion framework. To achieve stringent requirement for energy-efficient operation and reliable transceiver design, network configuration must comprises of a low cost and low-powered sensor nodes that are deployed in hostile environment thus suffer from high failure rate. Towards this design aspect following techniques have been proposed.

1.2.2.3 Spatial Diversity Technique

Reliable communications at the physical layer have been attained by using varieties of diversity gains. The space-time block code (STBC) and beam forming are the two versatile techniques that have been in use to achieve the optimal spatial diversity gain. Both the technique, i.e., STBC and beamforming techniques have common feature in that to achieve the diversity benefit and/or signal to noise ratio (SNR) gain, these techniques make use of multiple spatially separated antennas and special forms of signals. However, these two technique exhibit differences as well. The prominent difference between these techniques is that implementation of beamforming requires channel state estimation (CSI) at the transmitter, whereas, the STBC does not necessarily depend upon the known CSI at the transmitter.

A1:[(Alamouti, 1998)] proposed a simple transmit diversity technique for wireless communication and is the most widely used STBC scheme. In its basic architecture, the transmitter unit consists of two transmit antennas and the receiver has one receiver antenna. The scheme makes use of specific channel coefficients from transmit antennas to the receiver antenna. The received signal is expressed as linear combination of two transmits signals and the complex Gaussian noise.

A2:[(Larsson and Stoica, 2008)] presented a scheme that offers equal gain transmission and is the most commonly used beamforming technique. In a multiple input single output (MISO) system environment, it is a technique to send multiple coherent replicas of a symbol using multiple antennas. During equal gain transmission approach, each beamforming weight is constrained and offers diversity gain as well as the SNR or beamforming gain.

In contrast to it in a distributed sensor network environment conventional STBC or beamforming fails to deliver and distributed STBC and/or distributed beamforming technique are devised. In the distributed STBC or beamforming, the principal challenge is attaining the global synchronization, which involves carrier synchronization and time synchronization.[(Viswanathan and Varshney, 1997)] proposed time synchronization algorithms that offer fine precision. In literature several practical algorithms for carrier frequency synchronization have been proposed. In this context, [(Mudumbai et al., 2007)] presented master-slave (MS) architecture. Though, in practice the typical tasks such as time and/or phase synchronization consumes a considerable amount of energy, thus these tasks weaken their utility in energy constrained scenarios especially in sensor networks.

1.2.3 Network-layer Protocols

In WSNs, routes determination and subsequently maintaining it is a nontrivial task as the energy constraint and abrupt changes in sensor node state leads to frequent and unpredictable topological restructuring. Depending upon the network structure namely flat, hierarchal or location-based, the routing protocols are classified. The performance of a routing protocol depends heavily on the network architecture, which involves following design issues

1. Dynamics of networks elements (sensor nodes and sink) and the dynamics associated with monitored events or phenomenon.

2. Nature of nodes deployment that is either deterministic or self-organizing.

3. Hope-count dependent energy consumptions.

4. Pattern of data delivery to the sink, it could be continuous, event-driven, query-driven and hybrid.

5. Nodes functional capabilities; homogeneous and heterogeneous to support multi-modality features.

6. Incorporation of data aggregation/fusion at network nodes to realize processing within the sensor networks.

Depending upon the manner in which a source node determines a route to the intended destination, routing protocols can be classified into three categories namely reactive [(Lindsey and Raghavendra, 2001)], proactive [(Heinzelman et al., 1999b)] and hybrid protocols [(Braginsky and Estrin, 2002),(Intanagonwiwat et al., 2003a)]. In proactive protocols, all the probable routes are computed a priori, while in reactive protocols, based on demand of service, the routes are determined. Hybrid protocols combine these two schemes. Owing to limited storage space associated with sensor node and exceptionally higher number of routes requirement in proactive protocols, reactive and hybrid routing protocols gained attraction of sensor network research community. Further, depending upon whether one makes uses of location details or not, these protocols can be classified into location aware and location-less routing protocols. As per aforementioned classification basis, a brief summary of some routing protocols is presented here. [(Stojmenovic et al., 2005)] presented strategy to design network layer protocols in that in place of unit disk graph (UDG) model; realistic physical layer model is used. In this framework, physical, MAC, and network layers exchange information containing probability of a bit or packet reception as a function of distance between nodes instead of using the transmission radius in the UDG model. The sensor nodes used were homogeneous in a sense that they spent same amount of transmission power to send a message. A packet is considered as successfully received when all its bits are received correctly. Probability of correct reception is utilized by the MAC layer on regulating the number of acknowledgements and/or retransmission. [(Heinzelman et al., 1999a)] proposed sensor protocols for information via negotiation (SPIN) for flat network architecture. The framework makes use of high-level descriptors or meta-data and meta-data are exchanged among sensor nodes via a data advertisement mechanism prior to actual transmission of information bearing data. Incurred energy efficiency is reasonably good as the SPINs meta-data negotiation overcomes the typical problems such as overlap and resource blindness. Though SPIN suffers some demerits which are (i) lack of scalability (ii) nodes having proximity to sink are vulnerable in terms of their energy state. Thus, in applications such as intrusion detection that demands reliable delivery of data packets over regular intervals, SPIN is not a right choice.

Low Energy Adaptive Clustering Hierarchy (LEACH) [(Heinzelman et al., 2000)] is a routing protocols specifically designed for clusters based network configuration. To maintain the even distribution of energy dissipation among the senor nodes, LEACH utilizes randomized rotation of the cluster-heads. It presumes that all the network nodes are homogenous and sink is stationary and positioned outside the sensing fields. The principal components that lead to energy saving in LEACH are combination of data compression and uses of specific routing strategy. However, the LEACH also has some drawbacks as well , which are (i) Unable to meet stringent latency requirement in time critical applications, (iii) more likelihood of hot spot creation and (iii) difficulty associated with realization of specific sink attributes.

[(Lindsey and Raghavendra, 2001)] proposed a power-efficient Gathering in Sensor Information Systems (PEGASIS) routing protocol, it is a chain-based protocol that operates on LEACH structure. It outperforms LEACH by eliminating the overhead information exchange associated with dynamic cluster formation. Though it also suffers from some infeasible assumptions such as presuming all nodes have location information about all other nodes.

[(Manjeshwar and Agarwal, 2001)] presented Threshold Sensitive Energy Efficient Sensor Network protocol (TEEN). It is also a cluster-based routing protocol and makes use of LEACH framework, thus it employs LEACHs strategy to form clusters. During its implementation phase, TEEN presumes that the sensor network comprises of sensor nodes having the same initial energy and the sink has ample power, thus it can communicate to all the network nodes using single-hop wireless links. TEEN also suffers from drawbacks of LEACH as well as some more demerits which stem from its inherent features.

[(Intanagonwiwat et al., 2003a)] proposed a data-centric routing protocol "Directed Diffusion" is one in which communication takes place for named data. Its entire framework is based on four elements; interests, data messages, gradients and reinforcements. Compared to the previously mentioned protocol, this protocol achieves energy efficiency somewhat better, though it also has some associated limitations as well. For example, to perform data aggregation it makes use of time synchronization technique that is not easy to implement in sensor networks.

[(Braginsky and Estrin, 2002)] presented Rumer Routing that combines event flooding protocols with query flooding randomly. The rational for this protocol is preference to arbitrary paths over the shortest paths. In algorithm design certain assumptions are made about network dynamics that include uses of densely distributed stationary sensor nodes, and short range transmission links.

[(Wang et al., 2003)] proposed ad hoc back up node setup routing protocol (ABRP) that primarily makes use of ad hoc networks intrinsic properties and treats quality of routing as important metric. The main feature of the proposed algorithm is provision of a backup node mechanism to reconnect quickly.

[(Kannan and Iyengar, 2004)] presented Game-theoretic models to determine routing paths in wireless sensor networks. During its framework authors considered path length, path reliability and energy constraint of sensor nodes as three major parameters. The routing schemes make use of two payoff functions and justify that computing optimally reliable energy-constrained paths is NP-hard problem. Further, the path weakness measure is used to determine the qualitative performance of different routing schemes.

[(Alwan and Agarwal, 2009)] presented a survey on different fault tolerant routing techniques in WSNs. In that the fault tolerance mechanisms are classified into two major schemes: retransmission and replication. Retransmission schemes utilizes framework of directed diffusion, which typically includes following phases: Interest transmission, Gradients setup, Path reinforcement and Reinitiate reinforcement. The principal theme in implementing replication schemes is reliable information forwarding using multiple paths.

1.3 Other Miscellaneous Issues in Wireless Sensor Networks

Owing to severe energy constraint, sensor data fusion techniques poses them as a legitimate strategy. Motive for applying data fusion is to reduce the energy consumption associated with network overhead while complying with required estimation performance. In a last decade, uses of low complexity data fusion techniques have been investigated. Broadly these techniques are classified as distributed quantized estimation (DQE) and distributed compressed estimation (DCE). DQE can be realized by mean-square-error (MSE) minimization [Xiao and Luo (2005)] or by maximizing likelihood estimate between the true and estimated value of measured variable [(Dogandzic and Qiu, 2008),(S.Kar et al., 2012)]. DCE mechanism underlies determination of linear compression functions that minimize the estimation error at the data fusion center.

[(Gastpar et al., 2006)] proposed a distributed Karhunen-Loeve (KL) transform to perform the optimal dimension reduction. In a slightly distinguish manner,[(Fang and Li, 2010)] presented a study that operates based on joint dimension allocation and suppress matrix optimization.

[(Mendes and Rodrigues, 2011)] presented a survey on cross-layer solutions for wireless sensor networks. Cross-layer design approach operates on philosophy that parameters of two or more layers can be regulated based on inferences drawn from participating layers to achieve an optimization objective. The concept of cross-layering has been first exploited for TCP/IP networks [(V and M, 2005)], later it has attracted attention of researchers from WSNs regime. Verities of solutions have been proposed in literature that comprises of test-bed setups, analytical framework or simulation studies. In WSNs, some of the prominent objectives practiced based on cross-layer optimizations are efficient routing [(Y. et al., 2005)], QoS provisioning [(Yuan et al., 2006)], energy consumption [(S et al., 2006)] and optimal scheduling [(T and M, 2009)].

With a slightly different perspective,[(Vijay et al., 2010)] presented a survey on making usage of cognitive approaches in WSNs. The cognitive task is implemented by introducing the concept of knowledge plane in the conventional management planes based protocol stack of sensor network architecture. The survey in principal discusses neural networks based models, game theoretic approaches, adaptive modulation and sleep-scheduling to incorporate cognitive framework.

[(Fei et al., 2016)] presented a comprehensive survey about usage of multi-objective optimization in WSNs. In that the authors discussed varieties of optimization algorithms to cater the specific needs of sensor network environments, usage of appropriate metrics and open problems to explore.

1.4 Motivation

Over the last two decades many diverse algorithms have been proposed and investigated to address important operational and performance issues of wireless sensor networks. The reported methods encompass solution strategies to layer specific issues, performance measures optimization based on cross-layers concepts, uses of evolutionary computing techniques to enhance application specific performance measures, developing the fault-tolerant routing algorithms, energy-conservation schemes based on data aggregation and distributed fusion in sensor networks, and graph theoretic approach to devise routing path etc. Further, the efficacies of some of these approaches have been tested for query-driven as well as for the event driven sensor networks. However, the importance of uncertainties associated with network parameters have been acknowledged but still this particular issue hasnt properly analyzed and remains unexplored in some sense. The motivation for work carried out in this thesis mainly consider uncertainty issue at centre-stage and explore inferences of parametric uncertainties to analytical framework and its impact on key network performance measures. In Wireless Sensor Networks regime, invariably in all applications lifetime of the network have been a one of the important performance measures. Throughout the thesis during different case studies, the network lifetime is considered as one of the prime performance measures.

1.5 Objective of Research Work

The enhancement of wireless sensor networks lifetime is achieved by exploring the following multifaceted strategies, which focus on:

- To investigate the more appropriate clusters formation schemes for hierarchical network architecture, and subsequently devising the cluster-heads (CHs) selection strategy that maintains a reasonable balance among the residual energy status (RES) of participating sensor nodes.

- To explore sink-attributes; mainly its multiplicity and location aspects, so as to keep uniform RES gradient across the service-area. This result in avoidance of islands/hot-spots formation over the stipulated lifetime duration and at the same time ensuring fairly uniform coverage across the service-area.

- To formulate and analyze the mathematical model of query generation that owes significantly to inherent uncertainty associated with the query-generation process and its dependency upon the corresponding spatio-temporal parameters.

- To address and deal with the spatio-temporal uncertainties in a highly precise manner so as to encompass wide range of parameters values. Subsequently, it helps the network designer to decide the types and specifications of requisite network elements in more appropriate way. This also yields precise estimates of network's energy performance measures.

- To investigate and determine the timings of sink(s) relocation aspect and exploring the energy-threshold driven strategy that allows fusion of eligible quadrants, thus exploits spatial fusion aspect. The fusion exercise results in inclusion of topological hierarchy that offers more balanced energy gradient across the sensor network, which in turn leads to the network lifetime enhancement.

1.6 Thesis Outline and Contributions

In this section, the outline and contribution of the thesis is given. The thesis is organized as follows. The main contribution of the thesis is presented in five chapters. Chapter 2 is on exploring the use of k-means and fuzzy c-means algorithms for hierarchical sensor network architecture and a geometry based framework to determine the optimal location of a sink. Chapter 3 is on modeling the query generation process using appropriate probabilistic approaches that owe considerably to the associated spatio-temporal parameters. Chapter 4 is on modeling the query generation mechanism using Poisson probability mass function that undergoes parametric variations as per the interval form of the spatio-temporal parameters. Chapter 5 is an extension of chapter-4 and to treat uncertainties inherent with chosen spatio-temporal parameters intervals, these intervals are transformed using fuzzification process. Chapter 6 is on investigating the hybrid framework that includes dependence of the query generation process on Fuzzified spatio-temporal parameters intervals and exploring the spatial-fusion heuristic in anticipation of enhancement in energy performance metrics. Chapter 7 concludes the thesis. The outline of the thesis is as follows.

1.6.1 Chapter 2: FUZZY ALGORITHMS BASED CLUSTERING AND SINK POSITIONING

This chapter presents uses of k-means and fuzzy c-means (FCM) clustering algorithms to facilitate energy efficient clustering for a typical WSN scenario. The chapter also includes a brief summary of reported work carried out by other researchers to investigate impact of clustering mechanism in WSNs. Mathematical framework to implement FCM clustering algorithm is presented. Clusters obtained on using k-means and FCM schemes are further classified based on Euclidean distance measure with respect to a stationary sink position. On using these clustering schemes; the operational performance of the WSN is validated in terms of energy estimate driven performance measures. The performance measures include average residual energy status (ARES) estimation for two different clustering schemes resultant network topologies and average energy consumption per query on using these two schemes over the stipulated lifetime duration.

1.6.2 Chapter 3: PROBABILISTIC MODELS OF QUERY GENERATION

This chapter presents usage of probabilistic methods to model the query generation dynamics since deterministic models fail to capture the genesis of the query generation mechanism. Query generation process is modelled using Uniform probability distribution function and the Poisson probability mass function (PMF). The chapter also presents a survey of other reported methodologies that are based on probabilistic framework. To realize the hierarchical network topology uses of four different cluster formation schemes are investigated. The typical energy performance measures include (i) Residual energy status (RES) estimation for all the network nodes, (ii) Average residual energy status (ARES) estimation of entire network nodes. Further, these measures are estimated based on mathematical framework which is primarily governed by energy-centroid (EC) estimation and the optimum location of a sink. The network performance are estimated and compared for three different sink driven network scenarios. On considering Poisson PMF model; its single parameter (λ) is regulated as per the spatial and temporal aspects associated with query generation process.

1.6.3 Chapter 4: SPATIO-TEMPORAL POISSON DISTRIBUTION MODELS OF QUERY GENERATION

In this chapter, to incorporate uncertainties associated with the spatial and temporal parameters of query generation process in a manner that broaden its scope to cover wide uncertainties range, parametric model of Poisson PMF makes usage of interval-bound values of spatial and temporal parameters instead of considering crisp value. In the proposed methodology, value of the control parameter (λ) in Poisson PMF model is regulated by interval-form values of spatial and temporal parameters that strongly govern query dissemination and its inter-arrival-time-rate respectively. Network operational efficacy is verified in terms of performance measures such as ARES of entire network over the stipulated lifetime and the service-time-duration (STD) estimate that result on combining variety of sink driven attributes with the usage of four different clustering schemes. Chapter also reports a brief summary of varieties of methodologies investigated in last one and half decade.

1.6.4 Chapter 5: UNIFIED FUZZY INTERVALS POISSON DISTRIBUTION MODELS OF QUERY GENERATION

In this chapter, to deal with parametric uncertainties that recline within spans of chosen intervals, three-stage mathematical framework is presented. Stage-1 incorporates the modeling of query associated spatio-temporal parameters singleton form using appropriately chosen interval-spans. Stage-2 assimilates the fuzzy-triangular characteristics into chosen interval-spans by means of scaling the elements of spatio-temporal parameters interval-spans with corresponding degree of belongingness. Finally, in stage-3; using arithmetic mean estimation, arithmetic mean index (AMI) is proposed that aggregates the impact of interval-spans and degree of belongingness. This exercise subsequently yields spatio-temporal parameters in scalar form that offers the ease of mathematical computation against the intervals-bound intricate computation. The effectiveness of the proposed scheme is validated through simulation. The desired network service norms are estimated in terms of performance measures such as ARES of network over the stipulated lifetime, and the service-time-duration (STD) by which network ARES attains predetermined critical RES (CRES) mark on combining four different clustering schemes with the chosen sinks attributes. The chapter also presents brief discussions about varieties of modalities driven by spatio-temporal aspects and their aftereffects in WSNs paradigm in literature survey section.

1.6.5 Chapter 6: FUZZY INTERVALS MODELS AND SPATIAL FUSION FOR WIRELESS SENSOR NETWORKS

In this chapter, analysis of network performance measures obtained in previous chapters indicates more likelihood of further enhancement in energy metrics on exploring network surveillance by deploying multiple portable sinks. To validate it, an additional case study is investigated which makes use of four portable sinks for network surveillance. Further, as increase in dimension (here it signifies inclusion of more hierarchy in network topology in a restricted manner) always results into more likelihood of getting the optimal solution. To exploit this rationale, the concept of spatial (sub-areas/quadrants)-fusion is investigated. Implementation of spatial-fusion relies upon compliance of energy index based inequality. A heuristic comprises of mathematical framework is proposed to execute the task of spatial-fusion. The key performance measures of the sensor network namely, ARES and STD are estimated and analyzed for the varieties of sink attributes driven network scenarios with and without enabling the spatial-fusion concept. The literatures survey section of the chapter presents some of the recent attempts that deal with sink mobility issue in WSNs regime.

1.6.6 Chapter 7: CONCLUSIONS AND FUTURE WORK

This chapter summarizes the contribution of the thesis and discusses the possible future extensions.

Chapter 2

FUZZY ALGORITHMS BASED CLUSTERING AND SINK POSITIONING

2.1 Introduction

The major issues in wireless Sensor Networks (WSNs) are efficient uses of network's limited computing and communication capabilities and devising an appropriate routing strategy under severely constrained energy scenarios. To address these issues; it is necessary to work out algorithms that optimizes the consumption of scarce energy resources and select appropriate routing strategy that enhances the network lifetime under the given energy constraint. In this work, the use of k-means and fuzzy c-means (FCM) algorithms are investigated to form clusters and for the subsequent selection of cluster heads (CHs). For all the clusters; selection of CHs is done using corresponding sensor nodes residual energy status (RES) measure and Euclidean distance estimate. Depending upon the Euclidean distance measure between the sink node and center of gravity of clusters; these clusters are classified into five types. Further, the RES is measured for cluster heads as well for the sensor nodes in iterative manner, where iteration count depends upon number of queries generated within the network. For the periodic occurrences of queries/events; the fixed amount of generated queries could be transformed into time-interval elapsed as number of hours/months etc.

2.2 Literature Survey

Many protocols have been developed and reported to form the set of clusters for WSNs environment; some of the prominent protocols are Direct Communication (DC), Minimal Transmission Energy (MTE), Low Energy Adaptive Clustering Hierarchy (LEACH), K-Means, and Fuzzy C-Means [(Tan et al., 2008)]. In K-Means clustering; the nodes near the edges of the clusters are affected largely because the possible association of sensor nodes to a particular cluster is based on the degree of belongingness.

The degree of belongingness is either zero or one, although, the edge nodes may have equal degree of belongingness to more than one cluster. Thus, it creates misperception to sensor nodes about deciding their association with neighboring clusters [(Tilak et al., 2002)]. To overcome this problem, Fuzzy C-Means (FCM) clustering algorithm is used. It is an unsupervised, nondeterministic iterative method for an optimal cluster formation. During FCM implementation, the degree of belongingness could be any real number spanning between zero and one. The sensor nodes are assigned to a cluster based on their degree of belongingness and it is practiced repeatedly for all the sensor nodes over moderate epochs [(A. Paulraj and D.Gore, 2003)].

In this chapter, for the query based network, the uses of two different schemes namely K-means and Fuzzy C-means (FCM) algorithms are investigated. It is observed that the FCM scheme handles more number of queries than the K-means till the network attains stipulated lifetime. Thus, the FCM scheme is more suitable for applications at which primary concern is maximizing the network lifetime. Enhancement in network lifetime or network longevity is achieved by prolonging RES of individual sensor node which in turn mitigate "hot spot" phenomenon effectively [(K.Kalpakis et al., 2002)]. In [(P and Z, 2009)], authors formed the clusters and selected respective cluster heads by using k-means algorithm and energy status of each sensor nodes is investigated. In [(B.N et al., 2010)], authors described a dynamic approach to reduce broadcast radius that in turn results in reduced collision. In [(Alla et al., 2011)], authors proposed Hierarchical Adaptive Balanced energy efficient Routing Protocol (HABRP) to decrease probability of failure of sensor nodes to prolong the lifetime of sensors till these nodes drain out completely. In [(M.S. et al., 2011),(Li et al., 2011),(T. and A.H., 2011)], the authors proposed geographical and power based clustering algorithm (GPCA); a heterogeneous-aware clustering protocol in that the sensor nodes are identified by their global positioning system (GPS). In [(Wang et al., 2009),(Novak et al., 1999)], the authors presented an analytical approach to determine the optimal number of clusters in dense wireless sensor network using cross layer optimization approach. In [(Raghuvanshi et al., 2010)], authors developed routing algorithm based on FCM clustering and subsequently compared optimal and random cluster in anticipation of energy efficient routing that maximizes the network lifetime.

In [(H et al., 2009)], the authors proposed a distributed WSN data stream clustering algorithm to minimize sensor nodes energy consumption to improve the network lifetime. It is based on subtractive fuzzy cluster means "(SUBFCM)", and analyzes its energy efficiency as well as clustering performance in comparison with the standard data clustering algorithms.

In [(Chen, 2012)], author presented cluster-based hierarchical routing protocol based on LEACH (Low-Energy Adaptive Clustering Hierarchy) algorithm and a FCM. In that all the network nodes are divided into number of clusters and the associated cluster heads (CHs) are selected using RES measure of participating sensor nodes. In [(S and S.S, 2013)], authors proposed data mining process that reduces the energy consumption of sensor nodes in sensor network. In that a clustering algorithm is composed of three stages: first stage forms clusters and selects cluster heads (CHs), second stage comprises of transferring data from sensor nodes of a cluster to its CH. Finally, CH relays only one of the aggregated or compressed data packets to sink.

Several protocols have been proposed by researchers for optimum routing path and optimum location of sensor nodes to deliver the qualitative data in WSNs. The main aim of these protocols is to minimize the nodes energy consumption in order to increase the lifetime of the network. The sensor nodes which are in close proximity to the sink have to share relatively higher volume of queries. While doing so, the energy of these nodes depletes much faster compared to other nodes as most of the time these nodes act as relay entity. This situation is called "sink neighborhood problem". To mitigate this problem; possible solution comprises of mobility of the sink in a predefined path or finding an optimal path so as to facilitate the sink with dynamic set of neighboring nodes. This exercise results in a balanced energy gradient across the network till the network attains its stipulated lifetime.

Exploiting sink relocation aspect, several protocols have been proposed in the literature to show remarkable improvements in network lifetime by movement of the sink through optimum route locus, and by devising optimum positioning of sensor nodes in a geographical area [(Gandham et al., 2003),(Luo and Hubaux, 2003),(Papadimitriou and Georgiadis, 2005),(Wang et al., 2005)]. Movement of the sink can follow different patterns in a given service-area such as random, predetermined/fixed path or controlled movement, in order to improve the lifetime of WSNs. In [(Kinalis and Nikoletseas, 2006)], authors described a single sink driven sensor network that uses a pull strategy to collect data from sensor nodes. The principal demerit of single sink is its poor coverage. To alleviate coverage problem, the uses of multiple sinks have been investigated. In [(Ye et al., 2002a), (Kinalis and Nikoletseas, 2007), (Marta and Cardei, 2009)], authors described the path coordination strategies using the moveable sink in scenarios under which other sinks also participate in service-area surveillance. In [(Yu et al., 2010)], authors investigated a scheme in which to track the location of random movement of sink; the sink periodically transmits a beacon message which contains location information. In [(Giannakos et al., 2009)], it is reported that the reactive data forwarding using pull strategy is based on request message transmitted by the sink. In [(Kumar et al., 2012)], authors observed that the minimum RES can only be achieved, if the mobility trajectory of the sink is set close to the periphery of the service-area.

This chapter presents following strategies (i) The clusters formation and selection of CHs is done based on k-means & fuzzy c-means algorithms (ii) Euclidian distance estimate to determine optimal location of the sink. Both of these approaches are investigated for a particular sensor network scenario.

2.3 K-means and Fuzzy c-means clustering

In query based sensor networks, usually a prior knowledge about number of queries a sensor node can handle before it enters into hibernation or completely drains out its energy reserve is partially available. In this chapter, to get an approximate estimate about query handling capability, a cluster head node that is closest to the sink node is arbitrary chosen; and it is presumed that this particular node remains CH over the entire service-period. With this strategy; this particular CH is used exhaustively; in tern it drains out in a speedy manner and thus the query handling capacity of a node is inferred. This strategy helps out in getting approximate estimate about the minimum number of queries supported by network nodes. In this work, it is assumed that the sink is stationary and positioned at the center of the service-area periphery. For two principal tasks, i.e., for the formation of clusters and subsequently to select CHs; we adopted two different schemes namely, K-means and Fuzzy C-means (FCM) algorithms. It is observed that the network topology resulted from FCM algorithm handles more number of queries than the one obtained on implementing k-means algorithm. Thus, FCM approach is more suitable for applications in which primary concern is longevity of network lifetime.

The x and y coordinates of energy-centroid (EC) of each cluster is estimated using equation (2.1). The Euclidean distance between i^{th} sensor node of a cluster and the EC of that cluster is estimated using equation (2.2). In a particular cluster; sensor node maintaining close proximity to the estimated EC location performs the duties of the cluster head (CH).

$$EC_{jk}(x, y) = \frac{1}{n}\sum_{i=1}^{n} \omega_i(x_i, y_i) \tag{2.1}$$

$$D \equiv d_{EC_i,k} = \sqrt{(x_{EC} - x_i)^2 + (y_{EC} - y_i)^2} \tag{2.2}$$

Here, n is number of participating sensor nodes in a cluster, i is sensor node index (i = 1, 2... n), j is the cluster index (j = 1, 2...p), k is iteration count and ω_i is the weight factor of i^{th} sensor node. The weight factors are chosen as real numbers from the range [0, 1], the typical value of ω_i is inferred from the residual energy status (RES) estimate.

2.4 Euclidean distance based cluster classification

Under this scheme, the clusters are formed using two different schemes namely the k-means and Fuzzy c-means algorithms. Implementation of these algorithms results in classified cluster. In this chapter; the clusters are classified into five types based on Euclidean distance (d) measure from the sink to the centroid of clusters. The distance estimation results into cluster classification and the outcomes of k-means and fuzzy c-means algorithms implementations are listed in Table 2.1 & Table 2.2 respectively.

Table 2.1: Clusters classification based on Euclidean distance using k-means algorithm

Cluster Type	**Distance measure**	**Number of clusters**
I	Nearest ($D \leq 50$)	11
II	Nearer ($50 < D \leq 100$)	42
III	Near ($100 < D \leq 150$)	57
IV	Moderate ($150 < D \leq 200$)	66
V	Far ($200 < D \leq 250$)	74
	Total number of clusters=250	

Table 2.2: Clusters classification based on Euclidean distance using fuzzy c-means algorithm

Cluster Type	**Distance measure**	**Number of clusters**
I	Nearest ($D \leq 50$)	14
II	Nearer ($50 < D \leq 100$)	48
III	Near ($100 < D \leq 150$)	65
IV	Moderate ($150 < D \leq 200$)	75
V	Far ($200 < D \leq 250$)	84
	Total number of clusters=286	

2.5 Network simulation parameters

In this chapter during simulation run, the initial energy of each sensor node and their current consumption rate is chosen as 1.725 Joules and 575 mAh (milli-Ampere-Hour) respectively. To comply with these specifications, each sensor node is equipped with two 1.5 Volt alkaline batteries. The energy required to transmit and receive a single binary digit (bit) is taken as 1μJ and 0.5μJ respectively. In context to application served, the data frame size (as packetization of bits) of query messages and response (reply) messages vary arbitrarily. In present work, the size of these two data frames is considered as 240 bits and 1200 bits for query and response respectively [(Sha and Shi, 2005)]. To test and validate the usage of cluster formation schemes; the chosen network parameters are listed in Table 2.3.

Table 2.3: Network simulation parameters

S. No.	**Network Parameters**	**Value**
1	Service area; square in shape having its area	250X250 m^2
2	Maximum No. of query a node can handle before it drains out completely	2160
3	Number of Sensor Nodes deployed	1250
4	Number of clusters & respective heads (varies with time & chosen scheme)	250
5	Average number of sensor nodes in each cluster (varies with time & chosen scheme)	5
6	Initial energy of each sensor node	1.7250 J

2.6 Test Results and Analysis

In this section; simulation results obtained on using k-means and Fuzzy c-means algorithms are presented and discussed in two separate sections. During entire simulation exercise it is presumed that the service-area surveillance is done using a single stationary sink. In both the sections; the sink is arbitrarily stationed at the center of the square shape service-area periphery.

2.6.1 Results and Analysis on using k-means Clustering Algorithm

During the simulation exercise, sensor nodes are deployed uniformly across the service-area. Initially, cluster formation and the CH selection tasks are done on random basis. In subsequent epochs; the clusters are formed randomly, however, the CH selection for newly formed clusters is realized based on the RES estimate for participating sensor nodes and the node having highest RES is elected as CH. For the chosen network parameters as listed in Table 2.3; clusters are classifieds based on Euclidean distance measure from the arbitrary located stationary sink to the ECs of clusters. Out of five classified cluster types; for a set of each cluster type (class), RES measure of CH and other sensor nodes is estimated. With respect to increasing queries; RES measure for five arbitrary nodes that belong to different cluster types are shown in Fig. 2.1 and Fig. 2.2 for nodes acting as simple sensing entity and CH respectively. In both of these figures; captions used for a CH and a simple sensor node belonging to a particular cluster type is given as C_{p-i-CH} and C_{p-i-SN} respectively. Here, p denotes cluster index and in this chapter based on Euclidean distance measures five different classes of clusters are considered.

The RES measure of randomly selected cluster head (CH) namely 5^{th} node in cluster type-I, 25^{th} in cluster type-II, 37^{th} in cluster type-III,48^{th} in cluster type-IV, 68^{th} in cluster type-V are shown in Fig. 2.1. As anticipated the observations infer that the energy depletes at much faster rate for nodes (whether it is a sensor nodes or cluster head) belong to cluster type-I compared to types (II-V). It is also observed that with increasing Euclidean distance measures between a stationary sink and network nodes; the RES measure of network nodes remains at much better energy reserve status, while nodes communicate to the sink through multi-hop wireless links.

Further, if the set of same index nodes (here arbitrary chosen as number 5, 25, 37, 48, and 68) act as simple sensor nodes, their RES measure is sketched in Figure 2.2. Most critically for CHs that belong to type-I cluster and thus happen to be in proximity with the sink, RES measure is estimated. Here, it is observed that on average on handling 2160 queries; these CHs attain 10 percentage of initial energy reserve (IER) or deplete 90 percentages of IER. Hence, in this characteristic, the energy statuses of these arbitrary nodes are plotted on completion of 2160 processed queries. In the rest of this thesis, 10 percentages of IER is treated as critical residual energy status (CRES). Thus the nodes which have attained CRES mark are no longer continuing operation as CH and on the other hand, nodes that maintain better RES perform the CH task. However, the nodes that belong to cluster type-II to V are relatively far from sink obviously maintain much better RES measure and thus still get ample opportunity to operate as CHs.

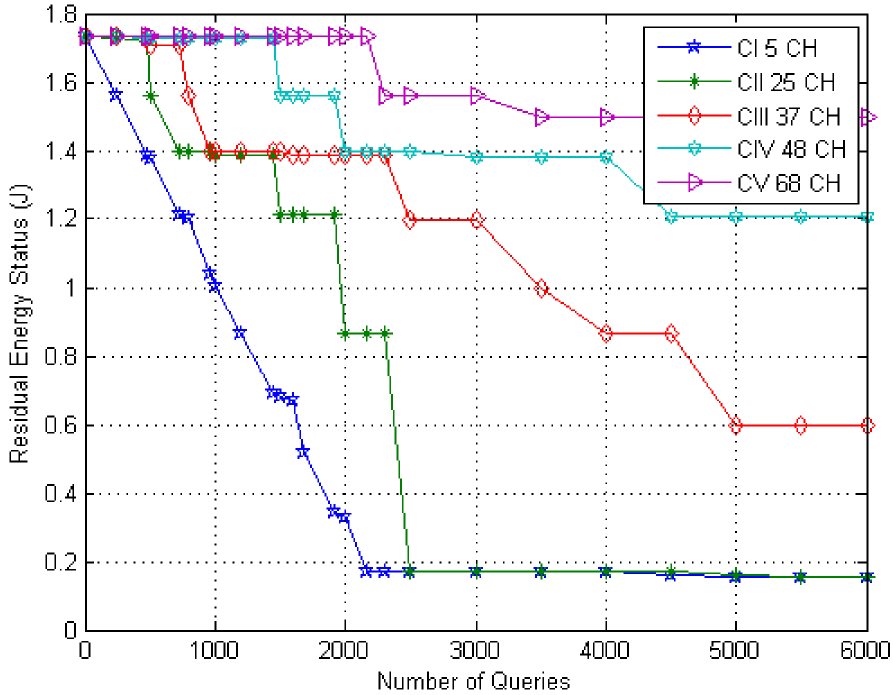


Figure 2.1: RES variations of CHs belonging to different clusters on using k-means algorithm

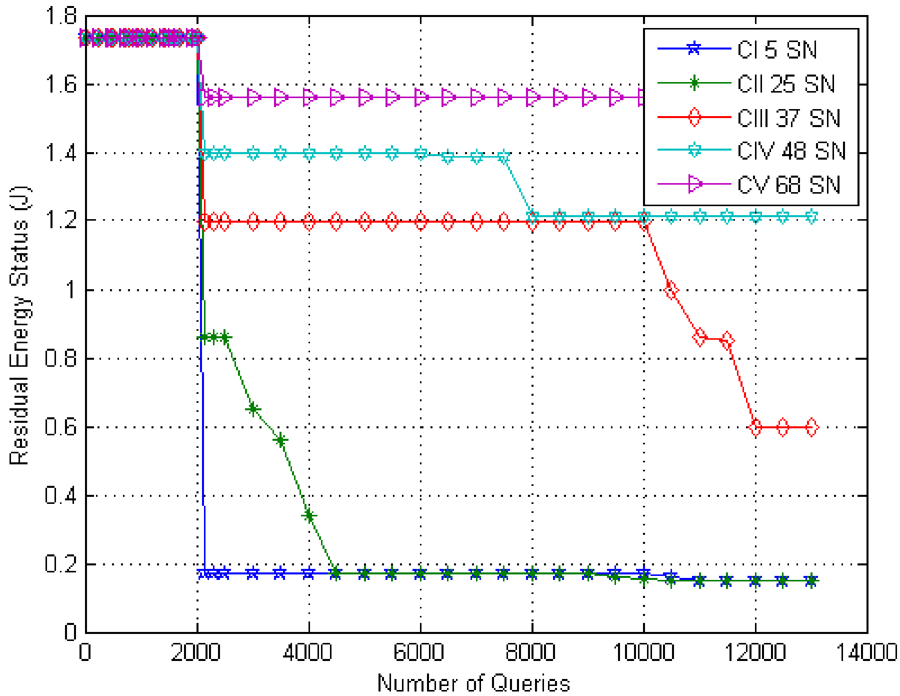


Figure 2.2: RES variations of arbitrary sensor nodes belonging to different clusters on using k-means algorithm

2.7 Fuzzy c-means (FCM) Algorithm

The principal limitation of k-means algorithm is the difficulty of assigning sensor nodes association to particular cluster for nodes that lie at the edges of clusters. To overcome this problem, uses of fuzzy c-means (FCM) algorithm is investigated and reported. The typical feature of fuzzy c-means is innate flexibility under which a node may belongs to more than one cluster with associated belongingness (membership-value) that varies from zero to one. Membership value near zero shows relatively poor association, whereas, its likelihood near one resembles strong association for a particular cluster. Thereby, exploiting this specific feature of FCM, clusters are formed and dynamically updated with time elapsed.

2.7.1 Formulation of FCM Algorithm

To get the optimal number of clusters and strategy about the cluster heads (CHs) selection; proposed algorithm makes use of the following variables and parameters: N=Total number of sensor node in an M X M m^2 service-area, n = number of sensor nodes in a cluster, C = cluster, i = node index (1, 2..N), j = cluster index (1, 2...C_{max}), **E** = energy vector, e = energy of each sensor node, J = fuzzy c-means objective function, EC = energy centroid of any arbitrary cluster, m = weight factor of membership function, β = set threshold, l = iteration count. μ_{ij} is a degree of belongingness with which node i associates with cluster j. In this work, it is presumed that the total number of N sensor nodes is uniformly deployed in M X M m^2 service-area. Sensor nodes mapped to a cluster of cluster index C_j where j=1, 2, 3...C_{max}. The cluster index and thus maximum number of simultaneous clusters is governed by the equality $2 \leq C_{max} \leq$ N. To determine the optimum value of C_{max} using fuzzy clustering approach; we have considered the energy vector **E** and its element as e_i (RES of node i), where i = 1, 2, 3.... N. The algorithm execution is based on the estimation of fuzzy energy matrix while complying with the following constraints:

$$\mu_{ij} \in [0...1], 1 \leq j \leq C_{max}, 1 \leq i \leq n \tag{2.3}$$

for any network node

$$\sum_{j=1}^{C_{max}} \mu_{ij} = 1 \tag{2.4}$$

The optimal Fuzzy c-means objective function J is defined as a function of energy vector **E** and C_{max} as:

$$J = f(\mathbf{E}, C_{max}) = \sum_{i=1}^{N} \sum_{j=1}^{C_{max}} (\mu_{ij})^m \parallel n_i - C_j \parallel^2 \tag{2.5}$$

Where,$\|\|$ is the Euclidean distance between i^{th} sensor node and centroid of j^{th} cluster and $'m'$ is the weight factor. In fuzzy c-means scheme, based on RES membership value associated with an arbitrary node the energy centroid estimation for a cluster is exercised using following pair of expressions

$$EC \triangleq (EC_x, EC_y) \tag{2.6}$$

$$EC_x = \frac{1}{n} \sum_{i=1}^{n} \mu_i.x_i \tag{2.7}$$

$$EC_y = \frac{1}{n} \sum_{i=1}^{n} \mu_i.y_i \tag{2.8}$$

The step by step procedure to implement FCM is summarized as below:

1. Select a value of N, C_{max}, m and, β a small positive constant. Initialize randomly a Fuzzy c-means objective function which is satisfying above conditions.

2. Set iteration,$l = 1, 2, 3..$ and compute the centroid of each cluster iteratively.

3. Compare J^l and J^{l-1}. If $\mid j^l - j^{l-1} \mid < \beta$(threshold).

4. Stop the process.

2.7.2 Results and Analysis on using FCM clustering Algorithm

The RES measure for a set of randomly selected nodes 5, 25, 37, 48 and 68 paired with cluster types I to V respectively is shown in Fig. 2.3. These RES characteristics distinguish that the energy consumption rate or RES decay pattern for CHs belonging to cluster type-I is significantly higher compared to CHs associated with other cluster types. A common observation is that as Euclidean distance measure increases; the RES measure of node maintains much better energy reserve compared to the nodes that proximate to the sink. Instances in which the set of same nodes (indexed with number 5, 25, 37, 48, 68) act as simple sensor node, the corresponding RES measure is shown in Fig. 2.4. Using fuzzy c-means algorithm, in average terms the query handling capacity of arbitrary sensor node associated with cluster type-I approximately reach 4300 with respect to chosen CRES mark. Here too, the general observation is that the nodes associated with higher index cluster types sustain relatively better energy reserve as they maintain progressively increasing Euclidean distances. Thus, on handling approximately 4300 queries most of the sensor nodes associated with cluster type-I are at the verge of CRES thereby inhibiting the CH tasks, whereas, the sensor nodes belonging to other types of cluster still perform typical tasks of CH. The performance outcomes of k-means and fuzzy c-means algorithms are compared in terms of energy consumption per query and are shown in Figure 2.5. The bar graph in this figure shows that the energy required to process a single query on using k-means algorithm is 0.7229 mJ, whereas, it amounts to 0.3614 mJ on using FCM algorithm. Thus, on average basis to handle a single query the FCM algorithm requires approximately 50% of the energy than that of the k-means algorithm. Therefore, for a given limited energy reserve, FCM algorithm is more efficient as it is capable of processing approximately two fold amount of queries. In aggregate terms this scaled query handling capacity can be calibrated in terms of duration served and conclusively we can say that the use of fuzzy c-means algorithm extends the network lifetime approximately double than that obtained on using k-means algorithm.

2.7.3 Heuristic for Queries dissemination

In this section, for service-area surveillance a single stationary sink is used and is located at the center of square shape service-area. To address spatial aspect at much better resolution; the service-area is divided into four uniform quadrants and is shown in Fig 2.6 To maintain the diversity in query spatial distribution, it is presumed that in four different quadrants query propagates in distinguish manner and thus generates different spatial patterns, that evolve with time on recursive basis. As an example, in quadrant-1; query initiate close to sink location and progress outward along the off-principal diagonal, OC. In quadrant-II; query instigate from the corner A of the service-area and subsequently it progresses radially inwards along the principal diagonal AO. Queries originate in a close by location of sink and evolve with time along the diagonal OG in quadrant-III. Whereas, in quadrant-IV; query initiate from two separate locations, which are in vicinity with edges OD and OF and advance perpendicular to these edges towards the diagonal OE. These query propagation dynamics are arbitrary chosen and its specific pattern applied to particular scenarios must be dealt on case to case basis. Strategy to devise

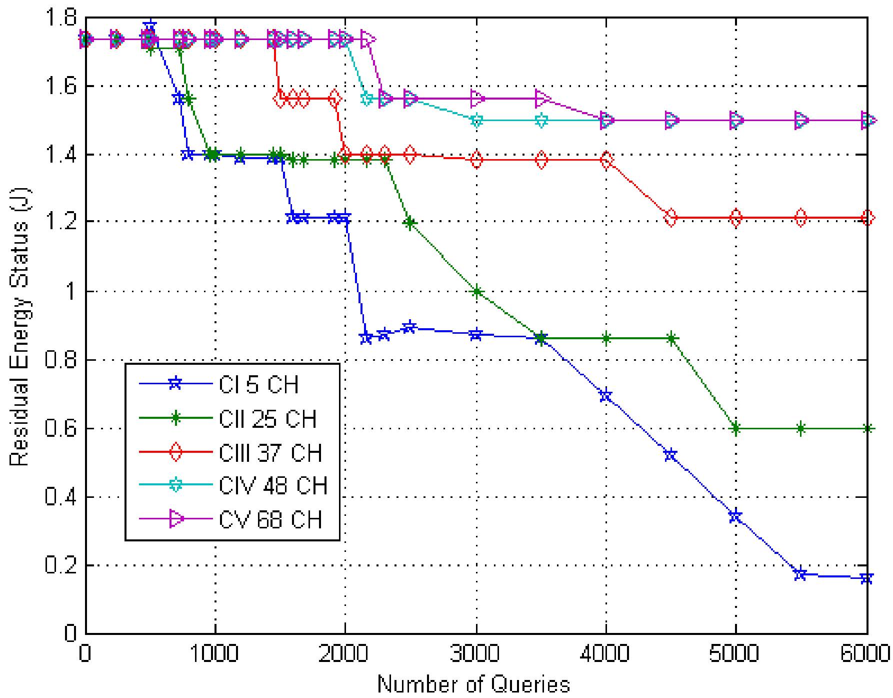


Figure 2.3: RES variations of CHs belonging to different clusters on using fuzzy c-means algorithm

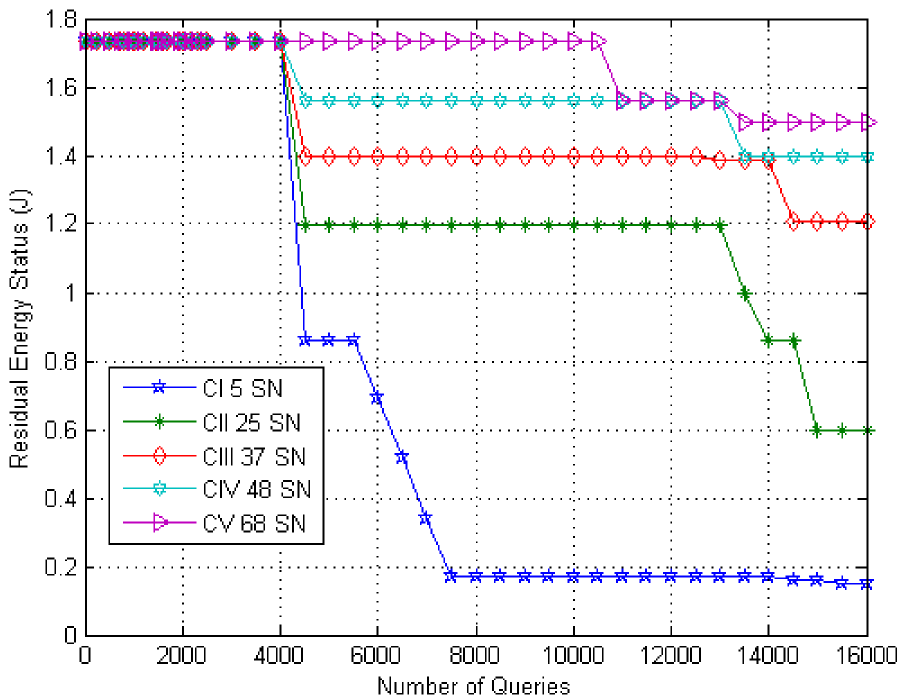


Figure 2.4: RES variations of arbitrary sensor nodes belonging to different clusters on using fuzzy c-means algorithm

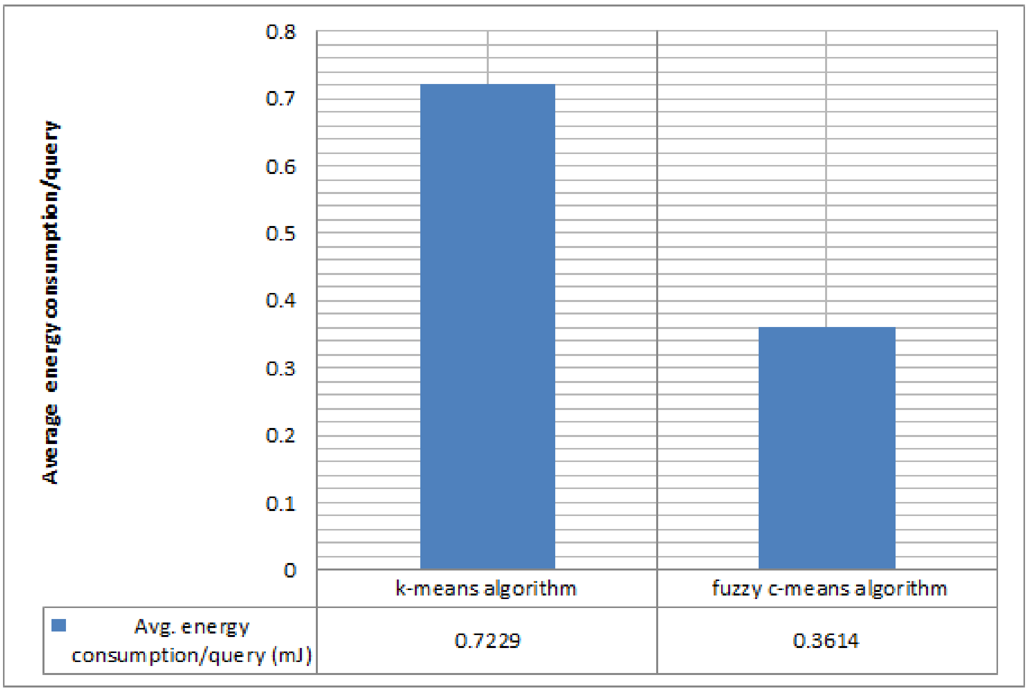


Figure 2.5: Comparison of energy consumption per query using k-means and fuzzy c-means algorithms

optimal location of sink that yields enhanced energy metrics is discussed in the following section.

2.7.4 Estimating Optimum Locations of Sink

Initially, it is presumed that the sink is positioned at the center of the service-area, designated with coordinate (x_c, y_c). The Euclidean distance between sink and an arbitrary node having its coordinates as (x_i, y_i) is given by

$$r_i = \sqrt{(x_c - x_i)^2 + (y_c - y_i)^2} \tag{2.9}$$

Where, i =1, 2.. N are number of sensor nodes in the entire network.

As time evolve, the sink must be relocated to new locations that surround sensor nodes maintaining relatively better RES measure compared to other nodes. In complementary manner, position aspect using the rectangular Cartesian coordinate could also be expressed in polar coordinate, (r, θ) form. Imparting motion to a sink, the optimum locations of the sink primarily rely upon the measures namely, angular direction (θ_i) and the radial distance (r_i) which are estimated for all the network nodes. These two measures θ_i and r_i are transformed into the Cartesian coordinates to estimate the nodes

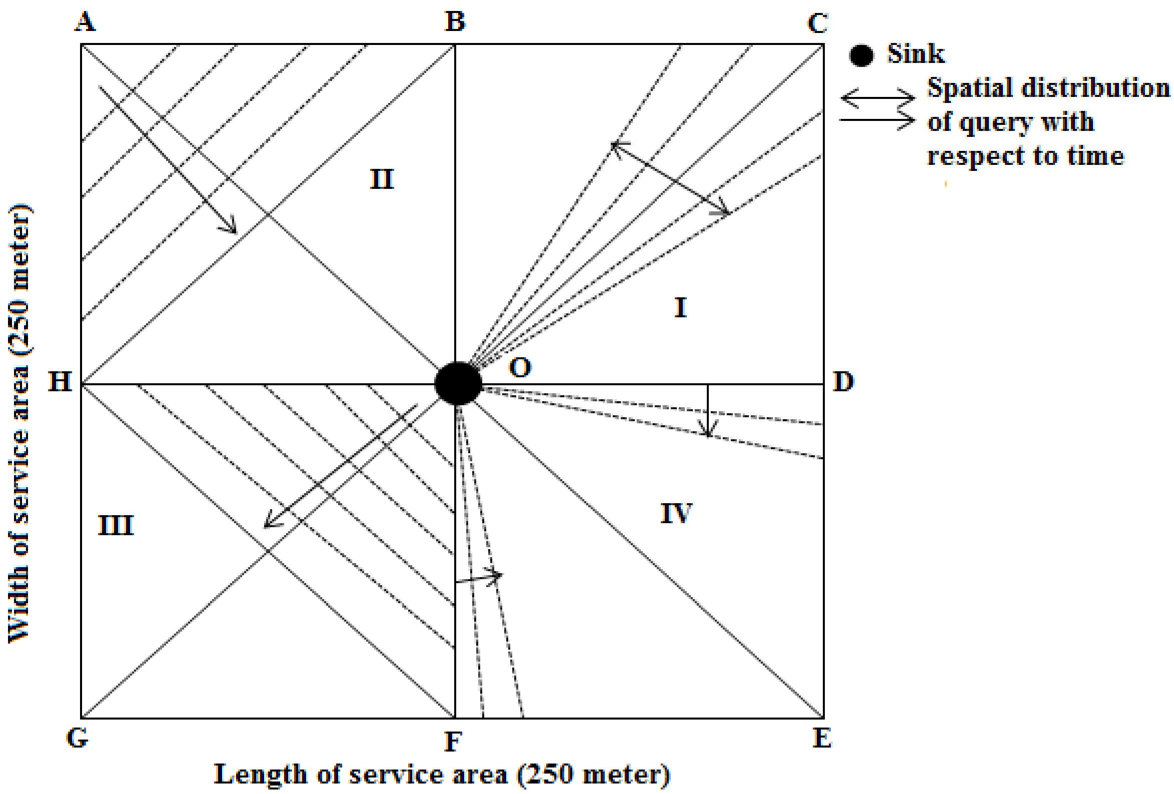


Figure 2.6: Service-area exhibiting different query generation patterns in different quadrants

positions using the following relations:

$$x_i = r_i \cos\theta_i, 0 \le \theta_i \le 2\pi \tag{2.10}$$

$$y_i = r_i \sin\theta_i, i = 1, 2, 3.....N \tag{2.11}$$

Where, r_i is the length of phasor (joining the sink node to the i^{th} sensor node) and θ_i is the angle made by phasor in counter clockwise direction with respect to the positive x-axis. Using 2-tuple update, i.e., (r_i, θ_i) for all the network nodes; the optimal location of a sink is designated as (r_{avg},θ_{avg}) where expressions for r_{avg} , θ_{avg} are given as:

$$r_{avg} = \frac{1}{N}\sum_{i=1}^{N} r_i \tag{2.12}$$

$$\theta_{avg} = \frac{1}{N}\sum_{i=1}^{N} \theta_i \tag{2.13}$$

2.8 Conclusions

The sensor network lifetime principally rely upon two components. First one is the energy resource available at individual nodes and the associated energy usage pattern. Whereas, another equally important component that also govern overall performance of the network is the proper and dynamic strategy of routing path selection. The simulation results indicate that in the case of a single stationary sink based surveillance scenario; the fuzzy c-means algorithm outperforms the k-means algorithm. It can be seen by observing the RES pattern of simple sensor nodes as well as the CHs. In addition to that, the estimation of average energy consumption on per query basis also put forth advantage in favour of fuzzy c-means algorithm. This indirectly affects the network lifetime as the reduced energy consumption results into enhancing the amount of the query handling capacity and thereby prolongs the network lifetime. The chapter concludes with a geometry based mathematical framework that estimates the optimal location of a sink under imbalanced energy gradients across the sensor network.

Chapter 3

PROBABILISTIC MODELS OF QUERY GENERATION

3.1 Introduction

Wireless Sensor Networks (WSNs) architecture consists of large number of tiny sensor nodes typically in the order of several hundred or few thousands. These sensor nodes are randomly disbursed or form a systematic topology such as array or grid/mesh within a service-area. These sensor nodes share the observed signal parameters attributes with each other or to an external base-station (BS)/sink node using multi-hop wireless communication links. Depending upon the applications, the sensor nodes are homogeneous or heterogeneous and support multi-modality sensing attributes. Compare to other conventional networks, relatively higher densities of sensor nodes facilitate job of sensing with a reasonable accuracy over large geographical areas. Typically, the sensor node architecture comprises of sensing, processing, transmission, mobilizer, position finding system, and power units (some of these components are optional like the mobilizer and position finding system). Sensor nodes coordinate among themselves to produce high-quality information about the physical environment under surveillance. Each of these scattered sensor nodes generate/relay the acquired data to other sensor nodes or to an external sink node(s). Sink nodes can be stationary or mobile (relocation aspects) and connect the WSN to the outside world ubiquitously using internet gateway, thus the end users access the reported data [(Karaki and Kamal, 2004)].

In a majority of WSNs architectures, sensor nodes possess limited resource of energy (battery), and thus the energy consumption must be dealt with proper care. In turn the lifetime of the WSN depends on the optimum utilization of scarce energy reserve and the adopted routing strategy for packets transmissions. For a query based network, the network lifetime depends on the amount of query generated, inter-arrival-time-rate of query generation (frequency of query generation) and spatial distribution of query. In most of the WSN applications, the query generation/arrival pattern is discrete in nature; owing to it most of the time sensor nodes remain in sleep mode. Hence, the energy dissipation pattern is inherently discrete and if it is systematically scheduled among the neighboring nodes it would definitely leads to enhancement of the network lifetime. Several researchers have investigated the importance of energy aspects in WSN environment, some of these are discussed here.

3.2 Literature Survey

[(Intanagonwiwat et al., 2003b)] proposed a data-centric energy efficient routing protocol using existing wireless local area network (WLAN) technologies. [(Gharavi and Ban, 2003)] presented a cluster-based ad hoc routing scheme for multi-hop wireless sensor network.[(Kwon et al., 2003)] reported an on-demand clustering mechanism, passive clustering to overcome limitations of limited scalability and inability to adapt high-density sensor distributions. In another investigation by [(Kumar et al., 2002),(Gharavi and Kumar, 2003b)]; the important aspects of WSNs have been investigated. These include distributed data compression and transmission, and collaborative signal processing. In a WSN; detection, classification, and tracking of target require collaboration among sensor nodes. Distributed signal processing in a sensor network reduces the amount of communication required in the network, lowers the risk of network node failures, and prevents the fusion center from being overwhelmed by huge amount of raw data from sensors. [(Chair and Varshney, 1986),(Varshney, 1997)] obtained the optimum fusion rules under the conditional independence assumption. Decision fusion with correlated observations has been investigated in [(K.Willett et al., 2000),(Drakopoulos and C.Lee, 1991),(Kam et al., 1991),(Kam et al., 1992)]. Many findings on the problem of distributed detection with constrained system resources are reported in the literature [(Rago et al., 1996),(Gini et al., 1998),(Yu and Varshney, 1998b),(Yu and Varshney, 1998a),(Kasetkasem and Varshney, 2001),(Hu and Blum, 2001)].

[(Chamberland and Veeravalli, 2003),(Niu et al., 2004),(Niu and Varshney, 2005)] proposed framework with a random number of sensors, a decision fusion rule uses the total number of detections reported by local sensors as a statistic for hypothesis testing. Herein, the other reported observations indicate that the signal power attenuates as a function of the distance from the target, the number of sensors follows a Poisson distribution, and the locations of sensors follow a Uniform distribution within the region of interest (ROI).[(Wang and Fang, 2010)] compared Poisson and Gaussian distribution of sensor nodes for object tracking in wireless sensor networks for randomly deployed sensor nodes.[(Niu and Varshney, 2005)] presented a concept of region of interest ROI, in which the network architecture was typically an ad hoc network.

[(Oka and Lampe, 2008)] in reported work improved the reliability of the detection/ estimation of hidden data of WSN. The localized hidden data are defined by a hidden Markov model (HMM). The propagated probability mass function is also observed by using Gibbs sampler (GS).[(Liu et al., 2011a)] have addressed energy-efficient data gathering issues in WSNs; an energy aware probability-based clustering algorithm (EPC). It has high scalability and flexibility for large scale WSNs. [(Liu et al., 2011b)] presented an event detection application of WSN and work mainly comprises of a multiple event detection scheme using compressed sensing (CS). Three algorithms of CS are used in the proposed scheme to manifest the associated merits of detection probability over the traditional decentralized detection methods using Bayesian approach.

[(Xu et al., 2011)] proposed Bayesian Compressive Sensing (BCS) theory with hierarchical Bayesian analysis model to investigate the process of wideband spectrum detection and data fusion for Cognitive Wireless Sensor Network (C-WSN). [(Mousavi et al., 2013)] proposed a Spatio-temporal event detection algorithm. The algorithm provides a probabilistic graphical model (PGMs) of WSNs. The algorithm incorporates the Markov chains in temporal dependency and Markov random field's theory in the spatial dependency of sensors in a distributed fashion.[(Jensen et al., 2006)] presented the Spatio-temporal indexing problem that includes a benchmark for the performance evaluation and comparison of Spatio-temporal indexes. [(Naymat et al., 2007)] presented the performance of spatial indexing structures which drastically deteriorates in a high dimensional space. On dimensionality reduction, another reported methodology comprises of a preprocessing strategy which involves a random projection to reduce the transformed space.

[(Yang et al., 2015)] presented uses of microbial fuel cells (MFCs) as renewable energy source to support environmental monitoring. The paper reports mathematical models for optimal duty-cycling that minimizes the probe packet reception time.[(Sun et al., 2014)] presented a primate-inspired mobility model for intermittently connected mobile networks. To overcome some of the inherent drawbacks of the proposed model a scheme called primate-inspired adaptive routing protocol (PARP) is presented. In this chapter, uses of four different cluster formation schemes are investigated. In a service-area; query generation and its spatial distribution pattern are approximated by two types of probability mass function (PMF) models which are Uniform and Poisson. To improve the network lifetime, the limited energy reserve to be used precisely and the routing protocol must be devised dynamically in tune with RES of network nodes. With this motivation, we present four different schemes for clusters formation while adopting different query generation patterns. On deploying all the four clustering schemes namely, SKM, SFCM, DKM, and DFCM in a given service-area; WSN performance is evaluated on the basis of following measures (i) Residual energy status (RES) estimation for all the network nodes, (ii) Average residual energy status (ARES) estimation of entire network nodes. Need based these measures are estimated periodically/sporadically and on attaining stipulated network lifetime of WSN and (iii) Critical residual energy status (CRES) estimation with progressive time; in terms of percentage of network nodes that attain predetermined threshold energy level, it is a fraction of sensor node initial energy reserve (IER).

3.3 Estimation of Energy centroid and Sink Optimum Locations

In this section procedure for estimating two vital components that decide the energy efficiency aspects of a WSN is presented. These include (i) estimation of energy-centroid (EC) of clusters and (ii) a necessary mathematical framework to estimate the optimal locations of sink(s). Let N_1 = number of sensor nodes per cluster, (x, y) = coordinates of EC of clusters. N = total number of sensor node in a service-area, j = cluster index (j = 1, 2, 3..M), i = sensor node index ($i = 1, 2...N_1$), k = iteration count (here it represents uniform time epochs), ω_i = normalized weight factor associated with i^{th} sensor node and its value depends upon RES of that node.

The rectangular coordinates (x, y) of EC of each cluster and the distance between member sensor nodes of a cluster and the EC of that cluster are estimated using equations (3.1) and (3.2) respectively as:[(Kumar et al., 2012)].

$$EC_{jk}(x, y) = \frac{1}{N_1} \sum_{i=1}^{N_1} \omega_i(x_i, y_i) \tag{3.1}$$

$$d_{EC_k} = \sqrt{(x_{EC} - x_i)^2 + (y_{EC} - y_i)^2} \tag{3.2}$$

In complementary terms; the pseudo-code for the cluster EC estimation is summarized as:

1. Initially clusters are arbitrary formed so as to exhaustively cover all the network nodes.

2. For every cluster the location of EC is estimated using equation (3.1).

3. Euclidean distance is measured from EC of a cluster to its associated nodes and it is repeated for every cluster.

4. To ensure acceptable network services for the stipulated lifetime; maintaining a balanced energy gradient within the clusters and network is of primary concern. For that matter, step (iii) outcomes are arranged in ascending order and the node which is closet to EC location owes the responsibility of cluster head (CH).

3.3.1 Mathematical framework for the optimal location of sink(s)

A step by step procedure to identify the optimal locations of sink(s) is presented as:

1. ARES for each of the dynamically formed cluster is estimated.

2. EC for every cluster is estimated using equation (3.1).

3. Select CH for each cluster based on the RES estimation.

4. For all the clusters (j=1,2..M)
   Estimate the Euclidean distance (r) of EC of a cluster to associated sensor nodes in that cluster using

$$r = \sqrt{(x_{EC} - x_i)^2 + (y_{EC} - y_i)^2} \tag{3.3}$$

   Where,
   $i = 1, 2, 3, 4....N_1$

5. For all the clusters (j = 1,2,..M)
   Estimate the Euclidean distance between EC of every cluster to the initial location of sink having its coordinate (x_s, y_s) as

$$r_{js} = \sqrt{(x_j - x_s)^2 + (y_j - y_s)^2} \tag{3.4}$$

6. For all the clusters (j = 1,2,3...M)
   Estimate angle θ (angle made by a vector joining (x_j, y_j) to (x_s, y_s) with respect to positive x-axis) using

$$\theta_{js} = \arctan(\frac{y_j - y_s}{x_j - x_s}) \tag{3.5}$$

Steps (5) and (6) gives a result in a vector space representation, where number of vectors in a modeled space are equal to number of clusters formed, vector refers a straight line locus joining sink(s) with EC of a cluster. Each of these vectors is expressed in polar form as:

$$v_j = r_{js} e^{i\theta_{js}} \tag{3.6}$$

Here, i is the imaginary operator

7. For all the clusters (j = 1,2,..M); using polar form representation, we estimate r_{avg} and θ_{avg} as per the following pair of expressions

$$r_{avg} = \frac{1}{M} \sum_{i=1}^{M} r_{is} \tag{3.7}$$

$$\theta_{avg} = \frac{1}{M} \sum_{i=1}^{N} \theta_{is} \tag{3.8}$$

8. Equations (3.7) and (3.8) outcomes are used to estimate the rectangular coordinates for the optimal location of sink (x_{opt} , y_{opt}) as:

$$x_{opt} + jy_{opt} = r_{avg} * e^{j\theta_{avg}} \tag{3.9}$$

$$x_{opt} + jy_{opt} = r_{avg} \cos\theta_{avg} + jr_{avg} \sin\theta_{avg} \tag{3.10}$$

For the possible movement of a sink, the optimum location of a sink depends primarily on r_{avg} and θ_{avg} estimates as per equation (3.10). For the illustration purpose few plausible locations of the sink in the first quadrant of a service-area are shown in Fig.3.1.For an arbitrary location denoted by point **P**,its polar form representation is given as $r_i.e^{j\theta_i}$. All these locations are represented in vector form with respect to center (0, 0) of the service-area.

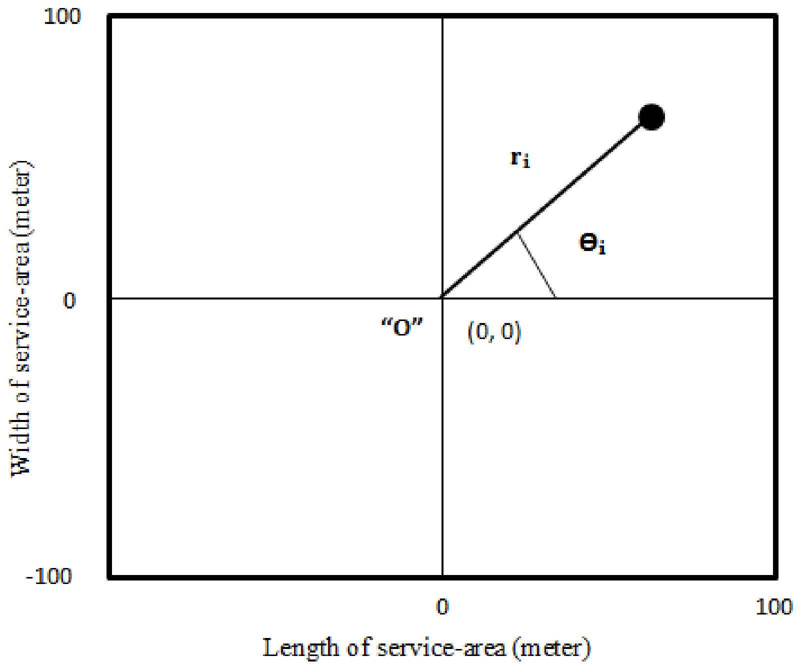


Figure 3.1: Polar form representation of vectors in a service-area

3.4 Mathematical model for query generation

In a broad perspective for the query based network paradigms; the query generation mechanism can be classified into two types and these are (a) Event based querying, and (b) model based querying. In both of these classes the generated queries follow the distributed quad tree (DQT) that is inherently a discrete pattern [(Demirbas et al., 2002)]. The challenging task in a sensor network is to implement querying in a distributed, lightweight, resilient and energy efficient manner so as to ensure reliable service norms for the stipulated network lifetime. In this chapter, we propose two different probabilistic distribution functions to model the query distribution pattern and these are Uniform and Poisson distribution functions. The analytical frameworks to make use of these distribution functions are presented next.

3.4.1 Uniform probability distribution model of query generation

Under this framework; we presumed that the queries are generated uniformly in a given service-area. Further, the uniform distribution implications can be seen in terms of spatial distribution of queries and their consistent occurrences in short epochs. The typical network's performance measures such as ARES and the percentage of network nodes that attain predetermined CRES are estimated for three different network scenarios. These scenarios are homogeneous in a sense that the service-area size, numbers of sensor nodes deployed in it and computing, transceiver and battery specifications of these nodes are same. However, the scenarios considered differ in terms of sink attributes. Here, we present three scenarios that involve a single stationary sink, a single portable sink and multiple (four) stationary sinks to perform network surveillance.

In all these three case studies; observations are drawn at two different levels namely, micro-level (cluster-level) and macro-level (network-level). The micro-level computation involves tasks such as RES estimation for all sensor nodes and cluster formation (Initially on random basis and subsequently based on RES of individual sensor nodes). Following it the cluster head (CH) selection is exercised for every cluster and the process is repeated till the network attains its stipulated lifetime. Further, for each cluster on using the Cartesian coordinates of associated sensor nodes and RES of these participating nodes, the location of energy-centroid (EC) is estimated. At network-level computations; outcomes of micro-level exercise are used to estimate the optimal locations of sink(s).

During the clusters formation phase, the term “static” refers to a situation in which clusters composition is time-invariant, since clusters once formed their topological structure do not change or update with progressive time. More precisely, the clusters composition in terms of participating sensor nodes (SNs) remains time-invariant. Whereas, the term “dynamic” refers to changing scenarios in which cluster formation takes place time to time (not necessarily periodic) and it is primarily based on the RES of sensor nodes.

Further, in k-means algorithm association of arbitrary sensor nodes with a specific cluster is formed based on a hard decision. In contrast, in fuzzy c-means algorithm decision about associating any arbitrary sensor node with a specific cluster is exercised based on the degree of belongingness by which a sensor node may join a specific cluster. It is usually based on inferences drawn from an approximately chosen fuzzy distribution characteristics. In comparison to the hard decision based k-means algorithm, the fuzzy c-means algorithm involves gradually varying multiple threshold levels in a continuum range from zero to one.

3.4.2 Poisson Probability Mass Function model of query generation

In a discrete sample space for a majority of diversified practical scenarios, the arrival pattern of random events is mathematically modeled using Poisson probability mass function (PMF).In WSNs environment, the query arrival and spatial distribution pattern are random, and owing to discrete nature of query generation, the Poisson PMF appears as a suitable candidate to model the query generation dynamics. Thus, during this case study, we presumed that the total number of queries (Q) generated in a finite observation interval (FOI) is modeled using Poisson PMF. The single control parameters (λ) driven Poisson PMF to model the random generation of queries (Q) is given as:

$$P(Q = k) = \frac{\lambda^k}{k!} e^{-\lambda} \tag{3.11}$$

Where, k= 0,1,2,3....
To incorporate the dependency of control parameter (λ) upon spatial and/or temporal parameters, three different strategies to regulate λ are presented in the following sections.

3.5 Adopted Strategies to regulate control parameter (λ)

In few specific WSN scenarios; $4 \sim 8$ queries are generated per minutes [(Karakaya, 2013)]. In this chapter, it is presumed that in a given FOI that typically encompasses expected lifetime span; on an average 400 queries are generated in the first quadrant of a service-area. The amount of query is arbitrarily chosen and depending upon applications specific need it can be altered. In the subsequent subsections; three different strategies are presented to regulate the λ in equation (3.11) and thereby the probabilities of queries distribution are estimated. Regulation criteria for λ includes: (a) arbitrary selection of λ, (b) to inculcate λ dependency upon spatial parameters; λ varies in proportion to fraction of network radial distance that signifies effective area of queries occupancy. In a way, this exercise implicitly shows λ sensitivity towards spatial-variance, and (c) for temporal parameter driven dependency; λ is modeled as a function of length of the observation time-interval/epochs that in turn governed by query inter-arrival-time-rate. Thus exhibits sensitivity of λ on temporal-variance. These regulation strategies are discussed in the following sub-sections.

3.5.1 Arbitrary selection of control parameter (λ)

In this case; probability values are estimated using equation (3.11) for two arbitrary values of λ; λ=5 and λ=150 [(singh and sapre, 2009)]. Corresponding to these λ values; the PMF characteristics are shown in Fig.3.2. These characteristics reveal that for the lower value of λ the shape of characteristics is Poisson, while for higher value of λ ; the shape resembles the Gaussian characteristics having mean value λ (here $\lambda = 150$).

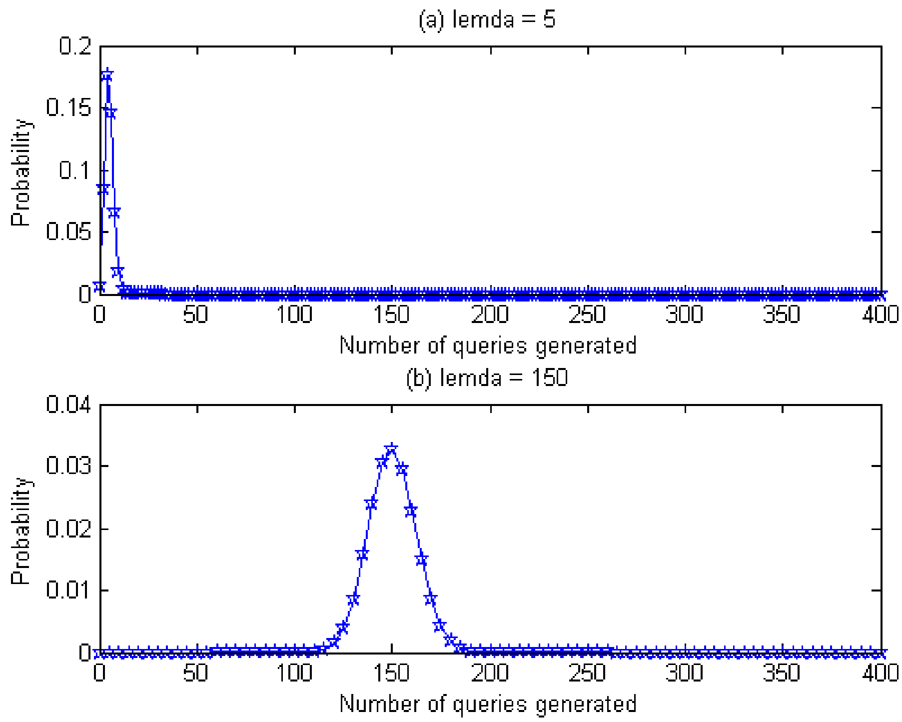


Figure 3.2: Poisson's PMF characteristics for two arbitrary values of λ

3.5.2 Spatial Parameter driven λ Regulation

During this case study; it is presumed that in a given FOI, the Euclidean distance (range) measures of queries disembark points with respect to sink(s) location regulates the typical values of λ used in the Poissons PMF expression to estimate the amount of query generation. Unlike the previous case the λ does not vary arbitrary rather it depends upon range (r_i) measures. To implement the distance (range) measurement based strategy, the entire service-area is divided into four quadrants and the radial distance (principal diagonal length) measure of the first quadrant is 141.2 m and is shown in Fig.3.3.

In simulation exercise; five different ranges are considered and details about these ranges estimation are given in Table 3.1. Corresponding to these five ranges, we have five different areas (sections) into which quadrant-I is decomposed, and these areas are marked as a_1 to a_5 in Fig.3.3. In this figure, the typical value of r_{max} is $100\sqrt{2}$ and is used to estimate the ranges as given in Table 3.1. The number of queries generated in any section (a_1 to a_5) of quadrant-I is a fraction (α) of queries generated in the whole quadrant-I. The α is estimated as:

$$\alpha = \frac{\text{area estimation of } a_i}{\text{total area of a quadrant - I}} \tag{3.12}$$

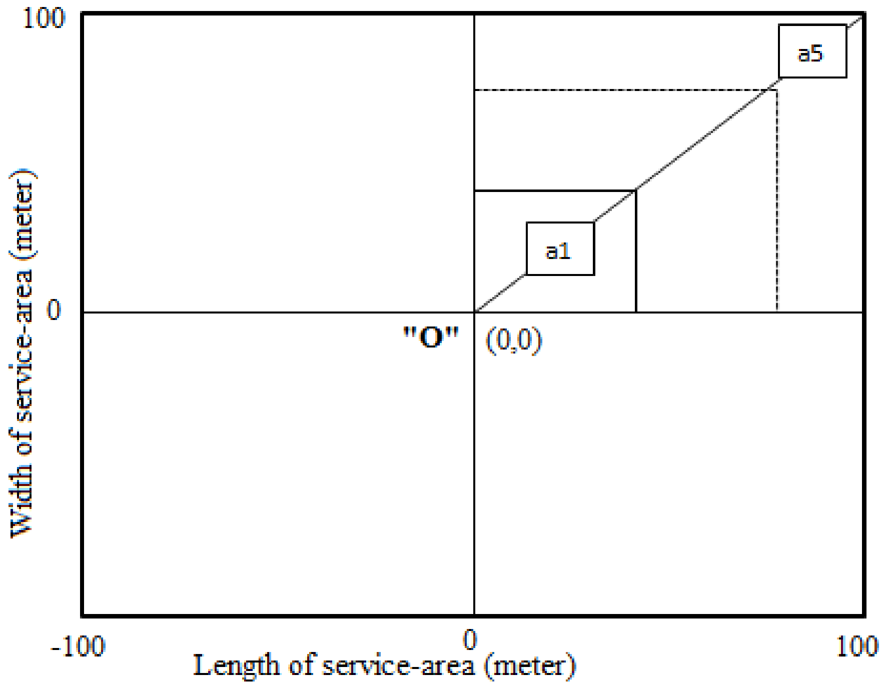


Figure 3.3: Pattern of queries density in the five sectors of 1^{st} quadrant (a_1 to a_5)

Where, $i = 1, 2, 3...5$

Table 3.1: Classification of radial distance measures

Radial distance measures classification	Ranges (m)	Average value (m)
Lowest	$0.2 * r_{max}$	5
Low	$0.4 * r_{max}$	30
Medium	$0.6 * r_{max}$	70
Moderate	$0.8 * r_{max}$	110
High	$0.9 * r_{max}$	136

These different ranges measures (r_i) are substituted in place of λ in equation (3.11) to obtain the probability values and these PMF characteristics are shown in Fig.3.4.

3.5.3 Temporal parameter driven λ Regulation

In this scheme, the temporal parameter associated with query generation process governs the values of λ that in turn estimate probability of query generation using equation (3.11). Its implementation is illustrated using an exemplary situation and is shown in Fig.3.5. Here all the four quadrants of the service-area consist of distinct query propagation locus. These loci propagate with progressive time in all the four quadrants in a manner which ensures that queries are not ubiquitous during any short time intervals (epochs).

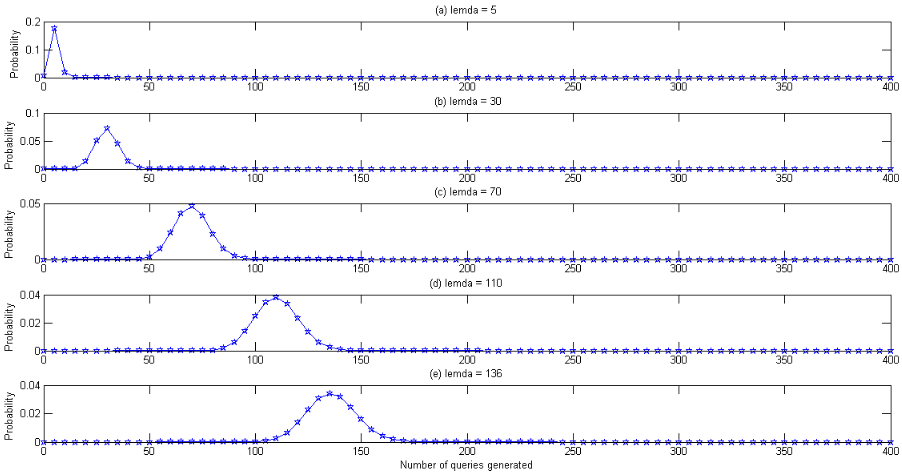


Figure 3.4: Poisson's PMF characteristics for five different value of λ having dependency on range (r_i) measure

For example, in quadrant-I, the query propagation pattern repeats many cycles with progressive time. Here each cycle exhibits propagation of query locus away from the origin towards north-east corner of quadrant-I. It is also shown in the query propagation locus that with progressive time as locus moves away from the origin it spans wider cone or covers larger angular displacement and thus effectively queries occupy large area.

Entirely different pattern is presumed for quadrant-IV, in that with progressive time query locus moves towards origin from the south-east corner of the quadrant-IV and gradually spans much wider cone in the proximity of the origin. Two other distinct query propagation patterns are shown in Fig.3.5 for quadrant-II and quadrant-III. Further, it is highly unrealistic to assume that on instantaneous basis, the query propagation pattern (more precisely its spatial distribution pattern) maintains uniformity in different parts of service-area. In aggregate terms over the network stipulated lifetime; the nature of query spatial distribution may be approximated as uniform probability distribution. However, as time progresses the dynamic behavior of query spatial distribution in shorter FOIs (epochs) resembles more closely a Poisson PMF as compared to the uniform PMF. Aforementioned aspects (spatial and temporal both) include dynamics associated with sensing field and in this chapter we propose an approach to incorporate these issues of practical concern be an integral part of query generation model.

On implementing the dynamics involved with query arrival pattern, the resultant spatial distribution of queries for quadrant-I is discussed first. In first quadrant, the locus of query moves away from the origin and the associated angular span of query dissipation become larger with progressive time. The process could be periodic or sporadic and leads to generation of query at different locations along the off-principal diagonal in Fig.3.5.

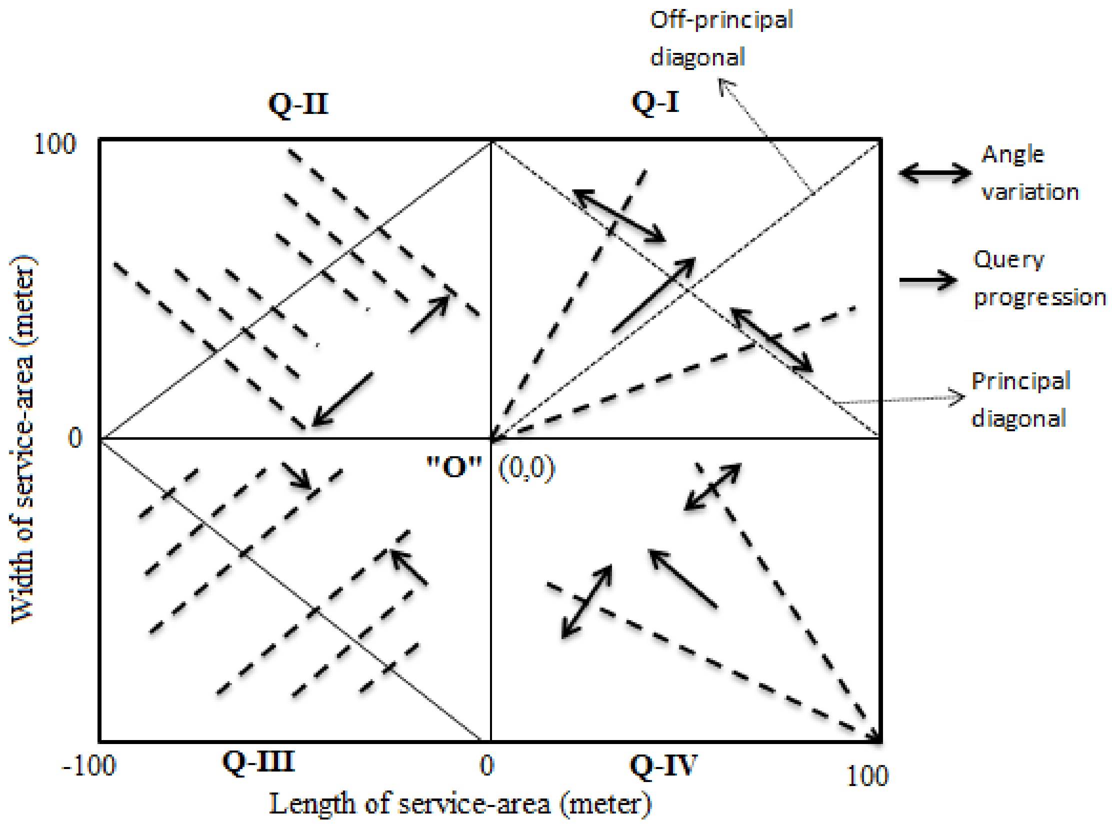


Figure 3.5: The distinct query generation loci in all four quadrants

On partitioning the spatial span along the off-principal diagonal, the total radial distance (range) is divided into ten parts, i.e., (r_1, r_2, ...,r_9, r_{10}), where, $r_{10} > r_9 >....r_1$. The range segments estimation used during simulation exercise is presented in Table 3.2.

Table 3.2: Radial distance (range) estimation along off-principal diagonal

Radial distance classification	Ranges (m)	Average value (m)
Lowest	(r_1 to r_2)	5
Low	(r_3 to r_4)	30
Medium	(r_5 to r_6)	70
Moderate	(r_7 to r_8)	110
High	(r_9 to r_{10})	136

Contrary to the quadrant-I, the query's arrival pattern and the corresponding spatial distribution in quadrant-IV is entirely different. In the quadrant-IV, queries are generated during different time intervals along the principal diagonal as shown in Fig.3.5 and locus moves towards the origin. With progressive time; the spatial locations of queries disembark spots are decomposed into ten different ranges, these are r_1, r_2,.. r_9, r_{10}; where, $r_1 > r_2 .. > r_{10}$. Using average computation; five different classified range intervals are estimated and reported in Table 3.3.

The time dependent/driven range intervals as listed in Table 3.2 and 3.3 regulates the value of control parameter λ in the Poisson PMF expression and thereby in turn estimate the probability of query spatial distribution.

Table 3.3: Radial distance (range) estimation along principal diagonal

Radial distance classification	Ranges (m)	Average value (m)
High	(r_{10} to r_9)	136
Moderate	(r_8 to r_7)	110
Medium	(r_6 to r_5)	70
low	(r_4 to r_3)	30
Lowest	(r_2 to r_1)	5

For the remaining two quadrants, i.e., quadrant-II and quadrant-III of the service-area, the query propagation pattern differs from earlier discussed pattern of quadrant-I and quadrant-IV. As shown in Fig.3.5; in quadrant II, the query locus commences from the centre of quadrant-II and progress away from the centre along the off-principal diagonal. During this traversing phase query locus encompasses wider cone (angular span) and exhibits periodic or sporadic repetitions. The typical angular displacement values used during simulation exercise are presented in Table 3.4.

In quadrant-III; queries instigate from two scattered locations along the principal diagonal and as time progress, the query locus traverses towards the center of quadrantIII. In contrast to query distribution pattern of quadrant-II, the query locus covers wider angular spans in proximity to center of the quadrant. Query locus driven angular span intervals used during simulation test are reported in Table 3.5. For both of these quadrants, i.e., quadrant-II and quadrant-III, the values of angular span interval listed in Table 3.4 and Table 3.5 respectively are used to regulate the value of control parameter λ .Subsequently these values of λ are used in Poisson PMF to estimate the probability of query spatial distribution.

For all the four quadrants, the PMF characteristics have the trends similar to one given in Fig.3.4, i.e., for narrower angular span the characteristics are Poisson, while for wider angular span it resemble Gaussian characteristics. The different patterns of query spatial distribution in four quadrants of a service-area are chosen on arbitrary basis, and it could be altered depending upon application scenarios. However, the inherent distinguishable variations in these patterns match the typical diversity of query distribution pattern in WSN regime. Having generated these varieties of query distribution patterns, the subsequent task is sensing and communicating the typical attributes associated with observed parameters of interest.

The amount of query generated and their spatial distributions derived from the uniform and Poisson distribution function models just discussed above are used during simulation exercise. Simulation is done using MATLAB programming language. The simulation results based on uniform PMF and Poisson PMF models of query spatial distribution are presented in next two sections separately. Both of these query probabilistic models are tested for WSN paradigm in which service-area surveillance is performed using a single stationary sink, a portable sink and four stationary sinks. For all these case studies; the network parameters used during simulation exercise are listed in Table 3.6 on considering network stipulated lifetime as eight years.

Table 3.4: Angle variation in second quadrant

Angle classification	Ranges (degree)	Average angle (degree)
Lowest	5-10	7
Low	11-25	18
Medium	26-46	36
Moderate	47-79	63
High	80-90	85

Table 3.5: Angle variation in third quadrant

Angle classification	Ranges (degree)	Average angle (degree)
High	80-90	85
Moderate	47-79	63
Medium	26-46	36
Low	11-25	18
Lowest	5-11	7

3.6 Simulation results for Uniform PMF model of query generation

For three different sink attributes driven case studies, we presumed that in a given service-area the spatio-temporal parameters dependent query generation pattern closely follows the uniform PMF model. The typical measures namely RES, CRES of sensor nodes, ECs location within clusters and ARES estimations for the entire network are estimated in three different subsections. These case studies corresponding to scenarios under which the network surveillance is coordinated by single stationary sink, single portable sink and four stationary sinks. Each of these network scenarios comprises uses of four different clustering schemes, and the impact of clustering schemes implementation are analyzed and compared in terms of performance measures such as ARES and CRES.

3.6.1 Network surveillance by a single stationary sink (SSS)

In this scenario; we presumed that the location of a single sink is arbitrary fixed at the center (considered as origin) of the square shape service-area. On deploying the four variants of cluster formation schemes namely static k-means (SKM), static fuzzy c-means (SFCM), dynamic k-means (DKM) and dynamic fuzzy c-means (DFCM) algorithms, the energy dissipation of individual sensor nodes as well as that of the entire network is observed as the network service time progresses towards the stipulated lifetime. The ARES characteristics (in Joules) of entire network and the number of network nodes (in percentage) that attain predetermined CRES mark are shown in Fig. 3.6 (a) and 3.6 (b) respectively. On analyzing the ARES measure, Fig. 3.6 (a) clearly indicates that on deploying DFCM the network nodes remain in good health (in terms of relatively better energy-reserve profile of network nodes). The network as a whole attains CRES at 6th year in progress compared to 4^{th} year in progress when SKM is used for cluster formation. Further during initial 75% duration of network's lifetime, the network enjoys relatively good energy profile on deploying DFCM compared to other three clustering schemes. Another noteworthy feature of DFCM is that the entire set of network nodes attains CRES mark approximately a year later compared to the network topology that stems from SKM as shown in Fig.3.6 (b).

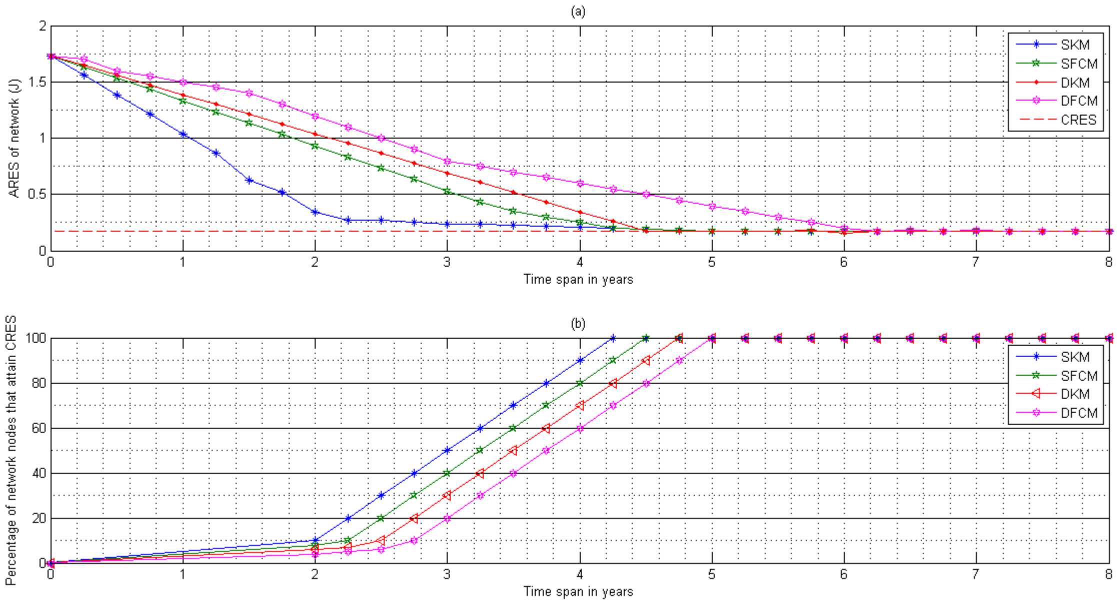


Figure 3.6: (a) ARES measure and (b) Percentage of network nodes attaining CRES in SSS scenario

3.6.2 Network surveillance by a single portable sink (SPS)

Herein, we presumed that initially the location of a single sink is arbitrary fixed at the center (origin) of a given service-area. As time progresses, the sink is allowed to move at optimal locations within the service-area so as to keep energy efficiency at its premium. Alike single stationary sink scenario, here too we deploy four different clustering schemes namely SKM, SFCM, DKM and DFCM. The overall network energy reserve status is evaluated in terms of two measures (i) ARES of entire network in Joules, and (ii) the fraction of network nodes in percentage attaining CRES mark. The DFCM superior performance can be validated as the ARES approaches predetermined CRES mark at 6^{th} year in service as against the SKM scheme resultant network topology as shown in Fig. 3.7 (a). The trend of percentage of network nodes attaining CRES is also in close proximity with that of single stationary sink scenario. These observations were recorded periodically and are shown in Fig. 3.7 (b). This also clearly reflects the superior performance of DFCM compared to other three clustering schemes. The vary characteristics also demonstrated that for any fraction/percentage of network nodes, the

DFCM scheme resultant network topology attains CRES mark approximately a year later compared to the network topology obtained on using SKM scheme.

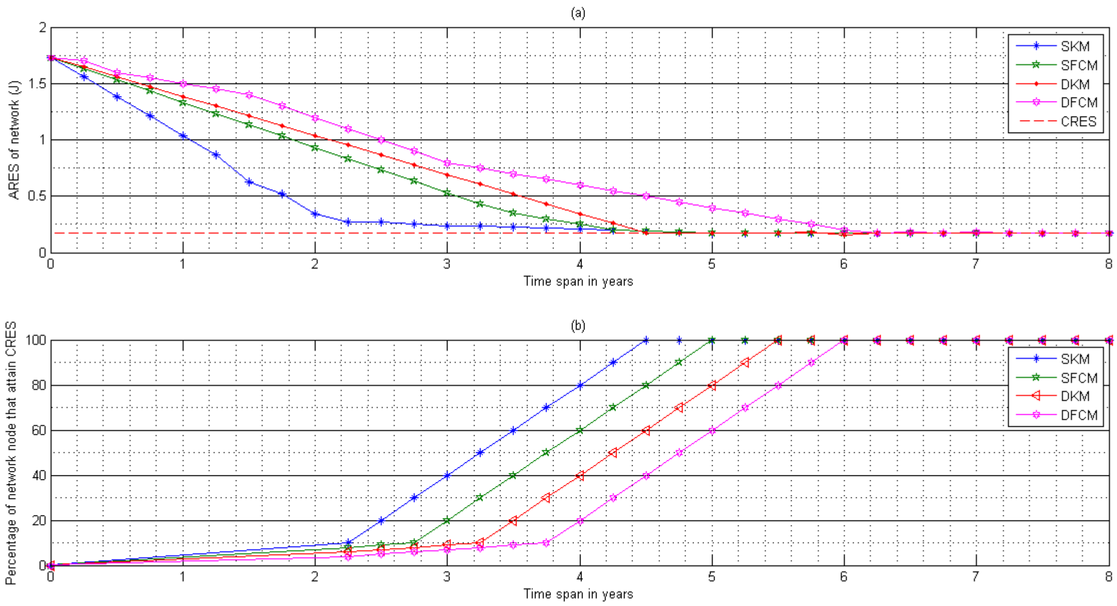


Figure 3.7: (a) ARES measure and (b) Percentage of network nodes attaining CRES in SPS scenario

3.6.3 Network surveillance by four stationary sinks (FSS)

In this scenario; locations of the four sinks are fixed at the center of periphery of a given service-area. Analogous to previous two scenarios, four different clustering schemes are deployed, and network performance measures namely, the ARES of entire network and fraction of network nodes attaining CRES are estimated periodically over the stipulated lifetime span. The ARES characteristics are shown in Fig. 3.8 (a). Among the four tested clustering schemes; the DFCM outperforms other three schemes. On comparing this ARES pattern with ARES characteristics as shown in Fig.3.6 (a) and Fig. 3.7 (a), the outcomes of the DFCM scheme deployment seems to be more promising as here network ARES attains predetermined CRES mark almost at the verge of network lifetime. The observation about fraction of network nodes attaining CRES mark is shown in Fig. 3.8 (b). The characteristics clearly manifest merits of DFCM as on using this scheme, all network nodes attain CRES mark while 7^{th} year is in progress compared to the outcome for a network structure that generated on using SKM scheme. As uses of SKM scheme leads to a situation where all network nodes attain predetermined CRES mark even earlier than 5^{th} year of service in progress.

In comparative terms, on combining four clustering schemes with three different sink attributes driven scenarios, the performance of ARES measure is summarized in Fig.3.9. These histograms indicate that a particular combination of DFCM clustering scheme with four stationary sinks based surveillance scenario results into the most optimal value of ARES measure at the stipulated network lifetime.

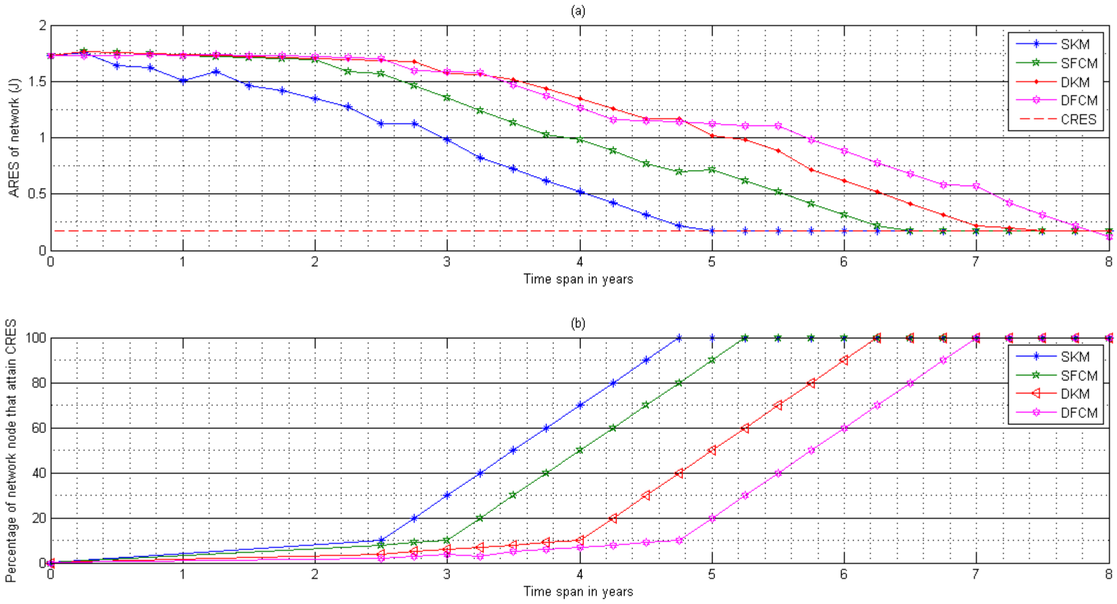


Figure 3.8: (a) ARES measure and (b) Percentage of network nodes attaining CRES in FSS scenario

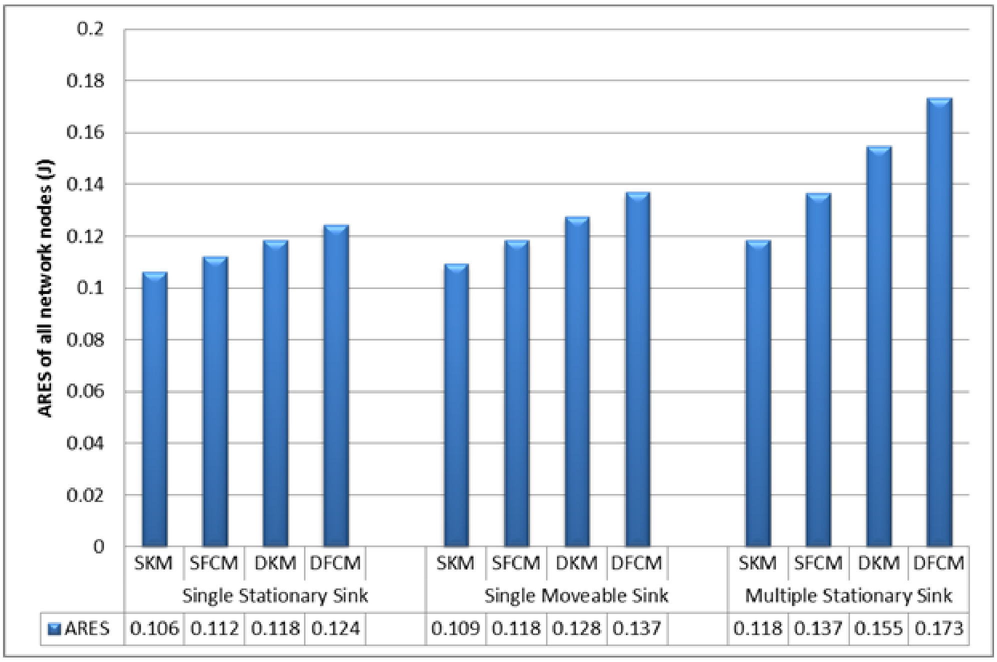


Figure 3.9: Summary of ARES measure at stipulated lifetime for uniform PMF model

3.7 Simulation results for Poisson PMF model of query generation

In this network paradigm, we presumed that the Spatio-temporal distribution of query closely follows the Poisson PMF characteristics. In other words, the Poisson PMF estimates the likelihood of query loci spatial and temporal aspects. The only controllable parameter (λ) of Poisson PMF expression is regulated using three different control strategies as discussed in sections 3.5.1, 3.5.2 and 3.5.3. In a manner similar to the uniform PMF model of query generation, here also we considered the uses of four cluster formation schemes namely SKM, SFCM, DKM and DFCM and three different scenarios, wherein the network surveillance is coordinated by a single stationary sink (SSS), a single portable sink (SPS) and four stationary sinks (FSS). For these three network scenarios, simulation results are presented and discussed in next three subsections.

3.7.1 Network surveillance by a single stationary sink (SSS)

Analogous to section 3.6.1, we presumed that arbitrarily the location of a single sink is fixed at the center (origin) of a square shape service-area. On using arbitrary two extreme values of parameters λ; lower value (λ=5) and higher value (λ=150), the network ARES estimation over the stipulated lifetime span is shown in Fig. 3.10 (a) and Fig. 3.10 (c) respectively. For lower λ (λ = 5), the resultant PMF shape is Poisson and corresponding to uses of DFCM as clustering scheme, the network ARES attains predetermined CRES mark at an instant that arrive when network enters into 5^{th} year of service. Corresponding to the higher value (λ=150); the Poisson PMF characteristic shape more closely approximate a Gaussian characteristics. In this case also the DFCM clustering scheme emerges as the most efficient one and uses of DFCM yields network ARES measure attaining a CRES mark two years later than that resulted from uses of SKM scheme.

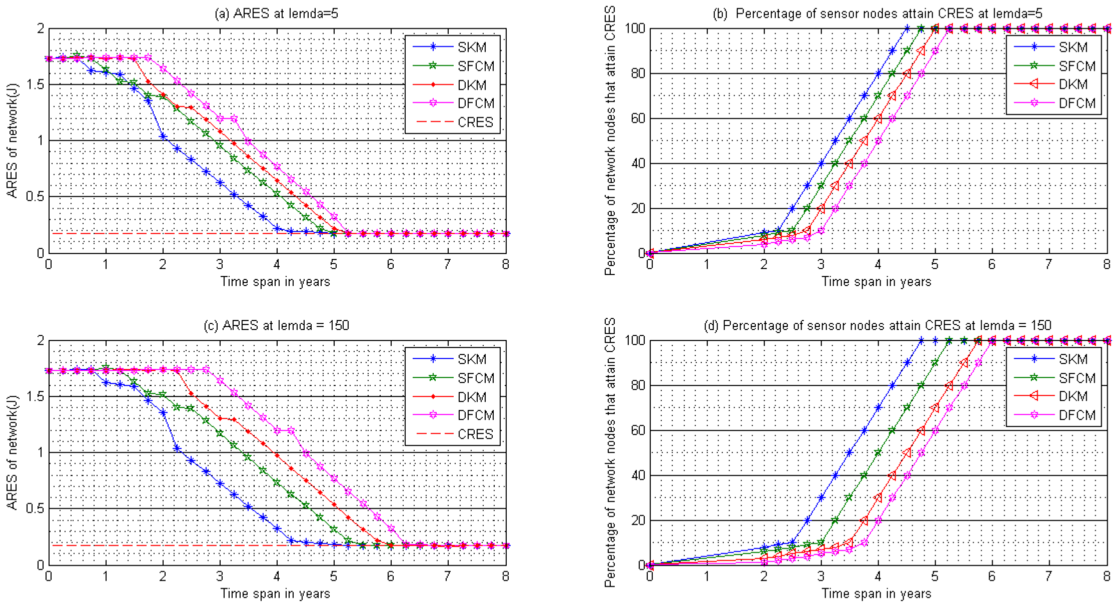


Figure 3.10: (a & c) ARES measure and percentage of network nodes attaining CRES (b & d) for SSS at $\lambda = 5$ & 150 respectively

3.7.2 Network surveillance by a single portable sink (SPS)

In this scenario; we presumed that at the onset of network operation, arbitrarily the sink is positioned at the center (origin) of the square shape service-area and subsequently depending upon the nodes energy consumption pattern or RES, it is relocated to suitable locations so as to maintain the network sustainability in terms of its energy reserve. Here, the λ varies in direct proportion to the spatial aspects of query distribution and thus the network radial distance range intervals regulate λ. However, for consistency and sake of uniformity to compare the network's performance measures such as ARES and the fraction of network nodes attaining predetermined CRES mark, in this section as well as in section 3.6.2; simulation results are presented and analyzed for query generation probabilities that are derived using $\lambda = 5$ and 150 . Substituting these values in Poisson PMF expression, we do estimate the probability of query distribution and thereby amount of query with associated spatial aspects. On using $\lambda = 5$ and $\lambda = 150$, the ARES characteristics and the characteristics highlighting the details about the fraction of network nodes attaining predetermined CRES mark are shown in Fig. 3.11 (a) to Fig. 3.11 (d). All these characteristics unanimously support the DFCM as the most efficient technique for cluster formation.

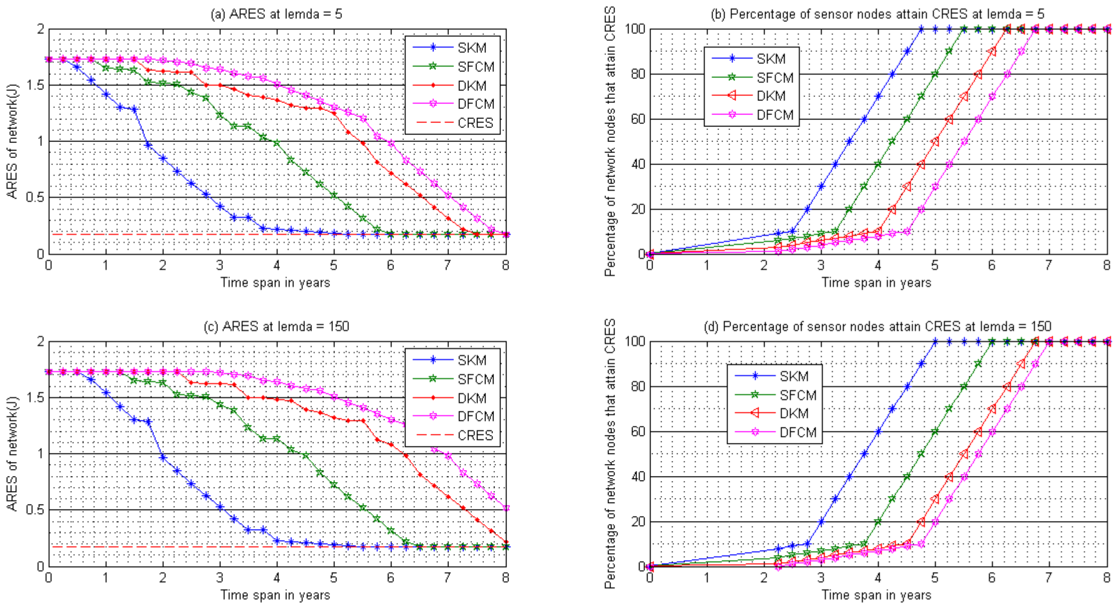


Figure 3.11: (a & c) ARES measure and percentage of network nodes attaining CRES (b & d) for SPS at $\lambda = 5$ & 150 respectively

3.7.3 Network surveillance by four stationary sinks (FSS)

In this case the network architecture in terms of sensor node specifications and sink specifications is same as the one presented in section 3.6.3. It is assumed that the query spatial distribution follows a Poisson's PMF characteristic, wherein the control parameter λ depends upon the temporal aspects of query generation which subsequently regulates the query spatial distribution pattern. However, as mentioned in previous section, during this case study also for consistency we use the similar lower and upper bound of λ values, i.e., 5 and 150 respectively. Uses of λ lower and upper bound values in Poisson PMF expression results into set of probability values having not only explicit temporal dependence but also the spatial dependence. On using these lower and upper bound values of λ, the network performance measures; the ARES estimation and the percentage of network nodes attaining the set CRES mark are shown in Fig. 3.12 (a & c) and

Fig. 3.12 (b & d) respectively. Here also DFCM inherits the status of the most energy efficient clustering scheme. In summary, on combining the multiple aspects such as sink attributes, and the uses of four clustering schemes; the network ARES states are shown in Fig.3.13 and Fig. 3.14 for $\lambda = 5$ and $\lambda = 150$ respectively. These results once again ensure the superior performance of the DFCM as promising technique for cluster formation.

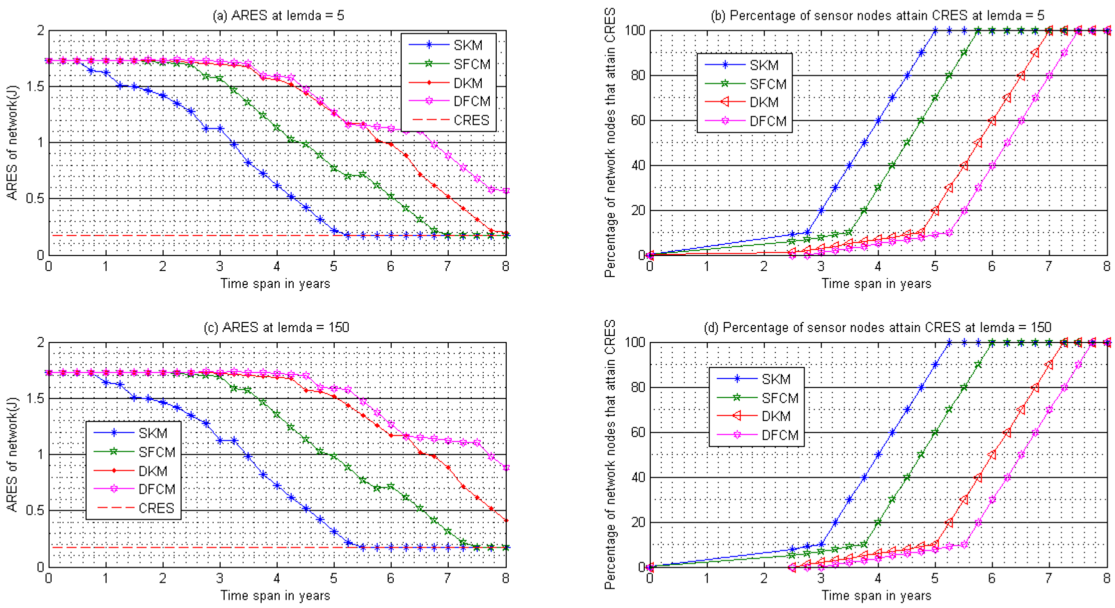


Figure 3.12: (a & c) ARES measure and percentage of network nodes attaining CRES (b & d) for FSS at $\lambda = 5$ & 150 respectively

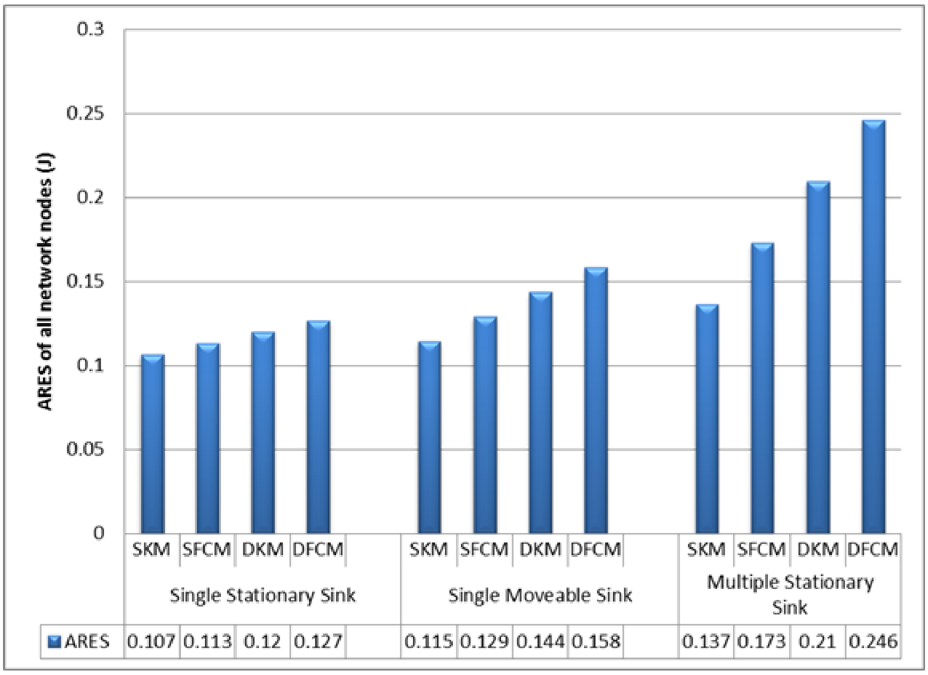


Figure 3.13: ARES measure comparison on using four clustering schemes at stipulated lifetime for $\lambda = 5$

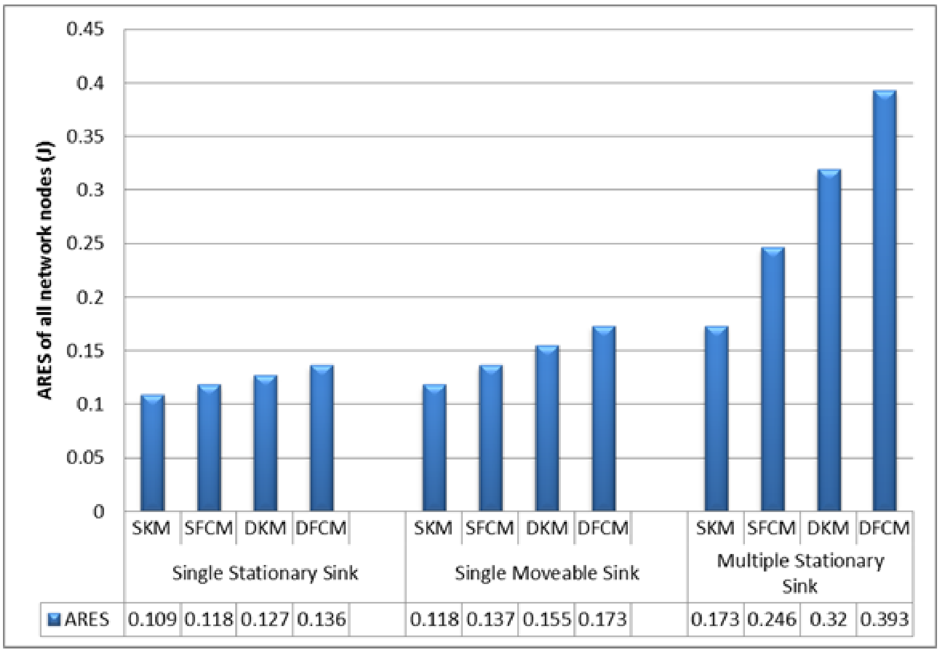


Figure 3.14: ARES measure comparison on using four clustering schemes at stipulated lifetime for $\lambda = 150$

3.8 Conclusions

In principles, the three major issues that affect the operational performance of WSN are addressed and discussed in this chapter. It includes (i) exploring the uses of four different clustering techniques, (ii) approximating the spatio-temporal dependency of query distribution pattern using appropriate probability PMF models, and (iii) Enhancing the sensor network overall energy reserve status that in turn improve the longevity of network on exploiting the sink node(s) attributes. Sink node(s) specifications are treated as basis for three types of case studies. The residual energy of sensor network status is evaluated using two different performance measures namely, the network ARES, and the percentage of network node attaining predetermined CRES mark as time progresses towards the stipulated network lifetime. Invariably, the DFCM emerges as the most energy efficient clustering scheme. Further, on comparing the outcomes of various case studies, the WSN architecture having four stationary sinks with DFCM exhibits the superior performance in a form of much improved overall network ARES.

It is assumed that N sensor nodes are randomly deployed in a given service area, and these sensors form the clusters as per the k-means algorithms and fuzzy c-means. The details about important network parameters such as number of nodes, number of clusters, nodes density in an arbitrary cluster, initial energy of sensor nodes and threshold set for CRES are given in Table 3.6.

Table 3.6: Network Simulation Parameters

Sr. No.	Network Parameters	Value
1	Service area; square in shape having its area	$200X200m^2$
2	Number of sensor nodes deployed	2308
3	Number of clusters & respective heads at arbitrary instant (t_0)	278
4	Average number of sensor nodes in each cluster at t_0	8
5	Initial energy of each sensor nodes	1.732 J
6	Critical residual energy status(CRES)	10% of total energy Thus it is 0.1732J

Chapter 4

SPATIO-TEMPORAL POISSON DISTRIBUTION MODELS OF QUERY GENERATION

4.1 Introduction

In a majority of wireless sensor networks (WSNs) architecture to support verities of sensing/measurements activities, sensor nodes possess limited energy resource (battery power), and thus it must be utilized with utmost care. Subsequently, the stipulated lifetime of the WSN depends upon the optimum utilization of network resources, and the adopted routing strategy for packet transmissions. Owing to widespread penetration of sensors based networks into plethora of surveillance applications; optimal utilization of scarce energy resource has attracted an unprecedented interest in WSN researchers community. For query based WSNs, the network lifetime depends on the amount of query generated, inter-arrival-time-rate of query occurrences (frequency of query generation) and the spatial distribution of queries. Usually, in large number of WSNs paradigms; query generation pattern is inherently discrete. Owing to discrete pattern, most of the time sensor nodes remain in sleep-mode. Therefore, appropriately designed mode-control strategies could lead to the enhancement of the network lifetime.

Usually, the sensor nodes are deployed randomly in a given service-area. Sensor nodes coordinate with each other to produce high-quality spatio-temporal information about the attributes of physical variables under observations. Each of these sensor nodes collect and route data either to other sensors or back to an external sink node(s). These sink node(s) may be stationary/portable/ mobile and connect the sensor network to the outside world using internet gateway, and thus the users have access to the desired data set [(Karaki and Kamal, 2004)].

Alike in other domains, where investigations are based on multidimensional data set WSNs also involve sensing task which predominantly depends upon spatial and temporal aspects associated with the physical variables to be sensed or measured. Many researchers have addressed the issues that impact the energy efficiency aspect of WSN and importance of spatial-temporal aspects. Some of these methodologies are briefly discussed here.

4.2 Literature Survey

[(Intanagonwiwat et al., 2003b)] proposed a data-centric energy efficient routing protocol using existing wireless local area network (WLAN) technologies. [(Gharavi and Ban, 2003)] presented a cluster-based ad hoc routing scheme for a multi-hop wireless sensor network. [(Kwon et al., 2003)] reported an on-demand clustering mechanism, passive clustering to overcome limitations of limited scalability and inability to adapt high-density sensor distributions . In a study proposed by [(Kumar et al., 2002)]; in context to WSN, the important operational aspects such as distributed data compression and transmission, and collaborative signal processing were investigated. In a WSN; detection, classification, and tracking of targets require collaboration among sensor nodes. Distributed signal processing in a sensor network reduces the amount of communication required in the network, lowers the risk of network node failures, and prevents the fusion center from being overwhelmed by huge amount of raw data from sensors. [(Gharavi and Kumar, 2003a)] have summarized the importance of collaborative signal processing, distributed networking, mobility, and ad hoc routing aspects of the sensor network.

[(Niu et al., 2004)] proposed a scheme in that the number of sensors follows a Poisson distribution, and the locations of sensors follow a Uniform distribution within the region of interest (ROI). [(Wang and Fang, 2010)] compared Poisson and Gaussian distribution of sensor nodes for object tracking in wireless sensor. [(Karakaya, 2013)] presented a novel technique for mobile sink(s) for a near real time application, namely, query-based data gathering with deadlines. The algorithm balances the system throughput and energy consumption by optimizing number of hops and duration of response time. [(Yu, 2010)] reported a scheme abbreviated as HIBOR (HIstogram with Bit vectOR); in this scheme, the sequences of data collected by sensors in a period are modeled by histograms and bit vectors. [(Liu et al., 2011a)] have addressed energy-efficient data gathering issues in WSNs; an energy aware probability-based clustering algorithm (EPC). It has high scalability and flexibility for large scale WSNs.

[(M.M.A. and A.A., 2011)] have proposed a dynamic indexing system for spatio-temporal queries in WSNs. The thesis addresses the issue of power consumption and maintains it at lowest possible value on indexing the sensed data. The indexing of data sequence helps in answering the queries.[(Guvenc Degirmenci and Prokopyev, 2014)] proposed an algorithm to maximize the lifetime of query based WSN. In that to maximize the overall network lifetime, a mathematical programming model was proposed that operates on selection of optimal transmission range, and controls the active/sleep mode schedules of sensor nodes. Precisely the principal focus was about the scheduling modes control while not addressing issues such as: dependency of query generation process on spatio-temporal parameters and impact of sensor nodes proximity to sink node on non-uniform energy consumption.

[(Munir et al., 2015)] presented a mathematical framework based on Markov process; in that reliability aspects were investigated at node, cluster and network level. Later these outcomes were utilized to analyze the fault detection and fault tolerance in WSNs. Here too, the associated spatial and temporal parameters were kept aside. [(Yoon and Shahabi, 2007)] presented clustered Aggregation (CAG) algorithm that forms clusters of sets of nodes sensing similar values with chosen thresholds over spatial aspects and time. The algorithm was formulated based on assumption that the sensed data in WSNs exhibit the spatial and temporal correlation of physical attributes existing in the sensing field. Authors proposed CAG for two modes of operation so as to cover its uses for different environments.

Owing to the universal importance of energy efficient operations in WSN; many researchers have proposed schemes that ensure optimal energy usage. It has been predominantly achieved by scheduling operation modes of sensor nodes and by carefully controlling the energy spend on computational and communication aspects. Further, clustering based unsupervised techniques were used to partition the whole sensor network into finite number of manageable smaller networks. Invariably, in all of these clustering schemes residual energy status (RES) of sensor nodes has been the key measure to derive estimate about other performance measures.

In this thesis, we make an attempt to address the dependency of generated query on spatial and temporal parameters in query based sensor networks. Owing to uncertainties associated with spatial and temporal aspects of query generation process and due to operational and/or physical constraints of network elements, the holistic approach is proposed that includes multivariate aspects with reasonable approximation. To keep the spatio-temporal aspects of query generation at centre stage; we propose probabilistic models for query generation and its spatial distribution. Towards it; the spatial and temporal parameters driven dependency of query generation process is modeled by Poisson probability mass function (PMF). In that control parameter (λ) of the Poisson PMF is made sensitive to associated spatial parameter (Δa) and temporal parameter (Δt) of query generation. Further, to devise a generic model that can accommodate broad spectrum of sensing based attributes and diversified applications, the revised control parameters (λ, Δa and Δt) in PMF expression are expressed as interval bounds instead of having a crisp value.

To resolve a given service-area, four different clustering schemes are deployed, and for the resultant clusters specific sensor network topologies; the performance measures of sensor network are estimated. These performance measures include averages residual energy status (ARES) and the critical residual energy status (CRES). CRES is a predetermined energy threshold usually kept constant but may be treated as variable need based, here for simple algorithmic framework; we treated CRES as fixed threshold level. These measures are estimated periodically and at stipulated network lifetime, while considering dynamics associated with spatio-temporal parameters of query generation process.

4.3 Query generation perspective and modeling

In general, for a majority of query driven sensor networks; the query generation mechanism is approximated using (a) event based querying and (b) model based querying. Both types of query follow the distributed quad tree (DQT) and the DQT leads to discrete pattern [(Demirbas et al., 2002)]. Owing to, inherent discrete random pattern of query arrivals, and constraint of energy reserve in WSN paradigm, the querying modeling aspect in a distributed, lightweight, resilient and energy efficient manner becomes a challenging task.

4.3.1 Associated spatio-temporal aspects of query generation process

For a given service-area during event monitoring/surveillance by a WSN, the two extremities are largely true. At one hand the area or spatial locations wherein events (physical phenomenon) may happen is usually large. Whereas, on the other side, the spatial locations or sub-areas where events take place (actually happen) are relatively sparse. Owing to these facts; the sensors are densely distributed in a given service-area, so as to support reasonably good coverage of sensing field and thereby avoid the islanding of any location within the service-area (misdetection of events). However, due to the cost involved and some other practical constraints, sensing a field with relatively higher density of sensor nodes becomes an infeasible task. Further, looking into the severity of environmental conditions; some of these sensor nodes breakdowns. Thus, the accurate detection of events with limited number of sensors remains a stimulating task in practical WSN. Spatio-temporal parameters integrated probabilistic models of query generation process are discussed in next two subsections.

4.3.2 Poisson distribution model of query generation

In any arbitrary paradigm; to model the random arrival of events with fair degree of approximation, the Poisson's probability mass function (PMF) manifest itself as one of the best contender. In WSN scenarios; two important aspects (spatial and/or temporal) associated with query generation process are random and thereby modeled by appropriately chosen PMFs. So far, in the reported studies; the query generation patterns in WSNs have been modeled as uniform PMF and Poisson's PMF with associated parameter λ having a crisp-value. Owing to random dynamics involved with query inter-arrival-time-rate (temporal aspects) and query density (spatial aspects); in aggregate terms, modeling of query generation process as the uniform PMF may be a good approximation towards the end of stipulated network lifetime but it lacks in following the spatial and temporal dynamics involved with query generation pattern on instantaneous basis. Whereas, these spatial and/or temporal dynamics can be addressed in a much better by modeling query generation process as Poisson's PMF. Thus, in this work, we presumed that over finite observation interval (FOI)the query generation process is modeled using Poisson's PMF. The Poisson's PMF expression to estimate the probability of query (Q) generation is given below for arbitrary chosen control parameter (λ) as:

$$P(Q = k) = \frac{\lambda^k}{k!} e^{-\lambda} \tag{4.1}$$

Where, k = 0,1,2,3....
The next two sub-sections include the details about the dependency of parameter λ on spatial and temporal parameters associated with query generation process.

4.3.2.1 Control strategy for λ regulation

In several instances during uses of Poisson PMF expression, its sole control parameter (λ) is treated as an arbitrary crisp number. Corresponding to two different values of λ resultant PMFs are of two types, namely (i) Poisson PMF for smaller λ and (ii) Gaussian PMF for larger λ.

In this chapter, the control parameter (λ) is inferred from a knowledge domain, which principally comprised of two important aspects of query generation, i.e., the spatial and temporal resolution aspects. Further, owing to inherent uncertainty; considering the spatial & temporal resolution value as crisp one (single scalar value) does not resemble associated spatio-temporal dynamics of query generation scenarios with a reasonable degree of accuracy. In this context; instead of working with crisp parameters, appropriately chosen intervals for considered parameters could lead to a better approximation towards the modeling of query generation having spatio-temporal dependence. To incorporate these spatio-temporal dynamics dependence equation (4.1) is suitably amended and so obtained models having explicit dependence on spatio-temporal parameters are discussed next.

4.3.2.1.a Spatial parameter driven λ regulation

In WSN paradigm; we presumed that the sensor nodes are uniformly distributed in a given service-area. To estimate the amount of queries that has dependency on spatial aspect, the entire service- area is resolved into finite number of uniform size sub-areas and these sub-areas are referred as clusters. Owing to feasible sensing range of ordinary sensor nodes (approximately 15-35 m), we considered the cluster coverage area as approximately 32 X 32 square meter, i.e., approximately 1000 m^2. Depending upon requisite spatial resolution this coverage area could be altered. Due to inherent randomness associated with query generation process; it is quite unusual that all of these clusters entertain same amount of queries during an arbitrary FOI. Thus, concurrently queries arrive to few clusters only.

In equation (4.1), the spatial aspect is included by modifying it to estimate the probability of k queries in an area-span of Δa as

$$P(Q = k) = \frac{(\lambda \Delta a)^k}{k!} e^{-\lambda \Delta a} \tag{4.2}$$

Where, k= 0,1,2,3....

To encompass a wider set of surveillance applications; value of λ is treated as interval bound instead of a crisp value. The λ-interval denoted as, λ_I, comprises of lower bound and upper bound values that we denote as λ_{Il} and λ_{Iu} respectively. In this chapter, we consider two λ-intervals, i.e., one smaller values interval and other one larger values interval so as to extend the scope of covering query generation mechanism for broad-ranging applications. Here, for simulation purpose; arbitrary chosen λ-intervals are [2, 10] and [100, 200]. Having considered intervals bounds for parameter λ in equation (4.2), we express spatial parameter Δa also in interval form. Corresponding to two different amount of query density that stem from two extreme scenarios, we consider lower and upper density of queries as 80 queries/Km^2 and 1250 queries/Km^2 respectively [(Huang et al., 2007)]. During simulation exercise; for the chosen service-area of size 250*250 m^2; on scaling proportionally, the lower and upper crisp values of queries density are 5 and 78. Owing to uncertainties associated with query density, it is always desirable to work with interval bounds. Thus, instead of considering crisp value for lower and higher density of query as 5 and 78; we span these values using lower density intervals, Δa_l = [2, 10] and higher density interval, Δa_u = [60, 90] respectively.

On approximating the parameters λ and Δa in equation (4.2) using appropriately chosen intervals, we estimate the probability. The product computation of the $(\lambda_I \Delta a_I)$term in equation (4.2) is done as per the interval arithmetic rules [(Alefeld and Herzberger, 1983)] as:

$$(a, b) * (c, d) = min \ \& \ max \ of \ \ (ac, ad, bc, bd) \tag{4.3}$$

On using the notation α for the product term $(\lambda_I \Delta a_I)$, i.e., $\alpha = \lambda_I \Delta a_I$ in equation (4.2); we obtain the α_l (on combining λ_{Il} and Δa_l) and α_u (on combining λ_{Iu} and Δa_u) as:

α_l = min. & max. of (2*2, 2*10, 10*2, 10*10)
= (4, 100) for entire service area
= (0.06, 1.63) per cluster

In a similar way, α_u is (6000, 18000) for the entire service-area and it is (98.36, 295) per cluster. Similarly on combining the query inter-arrival-time- rate with λ-intervals; the product terms are denoted as β, this process leads to temporal parameter driven bounded intervals and these intervals are listed in Table 4.1.

Table 4.1: Bounded values of α and β inferred from λ_I, Δa_I and Δt_I

Spatial parameters driven query amount (α) per cluster				**Temporal parameters driven query amount (β) per hour**			
λ_{Il} combined with Δa_{Il}		λ_{Iu} combined with Δa_{Iu}		λ_{Il} combined with Δt_s		λ_{Iu} combined with Δt_f	
Lower value	Upper value	Lower value	Upper value	Lower value	Upper value	Lower value	Upper value
0.06	1.63	98.36	295	0.22	2.22	166.66	472.22

4.3.2.1.b Temporal parameter driven λ regulation

Analogous to previous case study; query inter-arrival-time-rate (Δt) is modelled using interval bounds instead of having crisp value. Further, in a broad regime of civilian and military domains WSN infrastructures; surveillance applications seek different requirements about the query inter-arrival-time-rate and is highly application specific. For two distinct WSN scenarios; the typical lower and higher values of query inter-arrival-time-rate are considered as 120 queries and 1440 queries per month respectively [(Schreiber, 2007)]. To support wide-range of applications, we encompass these two crisp values of query arrival rate with appropriate intervals. In this work; the symmetrical lower interval [80, 160] and upper interval [1200, 1700] are considered for the lower and upper bounds of slower (Δt_s) and faster (Δt_f) query generation rate. To incorporate temporal aspect into query generation process, the equation (4.1) is amended; wherein, the probability of k queries arriving within an epoch of Δt is expressed as:

$$P(Q = k) = \frac{(\lambda \Delta t)^k}{k!} e^{-\lambda \Delta t} \tag{4.4}$$

Similar to spatial parameter driven λ-regulation case, here too we consider the smaller and larger intervals of λ value as [2, 10] and [100, 200] respectively. On combining interval bound of λ_{Il} with Δt_s and λ_{Iu} with Δt_f; we obtain β_l and β_u respectively which are listed in Table 4.1. Subsequently, these β values are used in equation (4.4) to estimate the probability of temporal parameter driven query amount.

The probability estimations based on equation (4.2) and (4.4) are sketched as PMF characteristics (PMFs) and are shown in Fig.4.1 and Fig.4.2 respectively. Both of these characteristics possess the following unique features; for the lower bound α and β values, the characteristics-shape resembles Poisson distribution, whereas, for higher bound α and β values these Poisson distribution transform into Gaussian distribution.

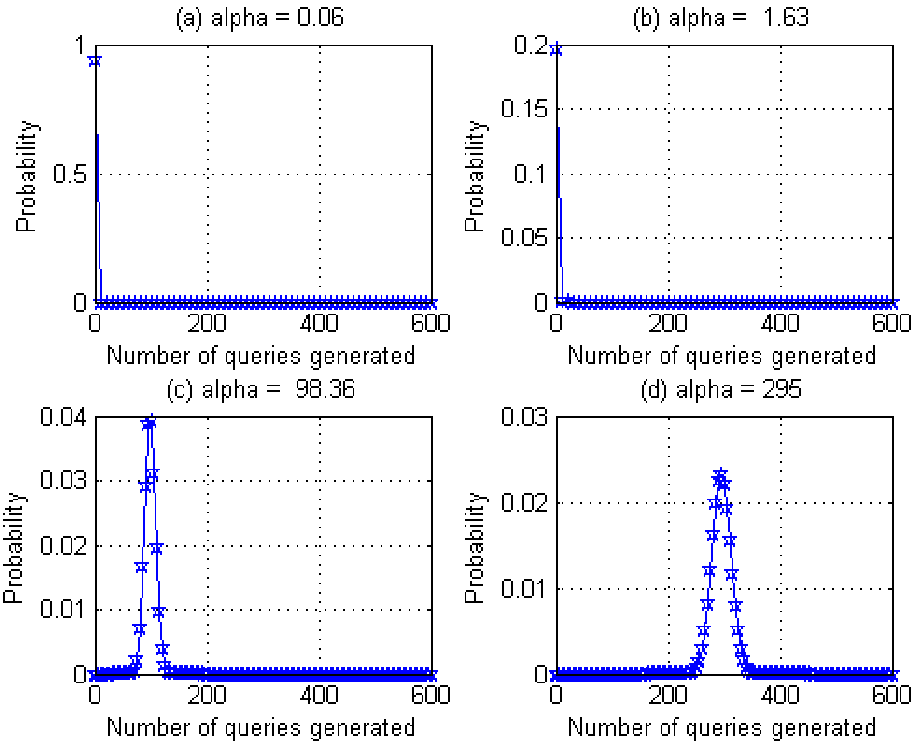


Figure 4.1: PMFs inferred from α on aggregating λ_I with lightly and highly dense queries

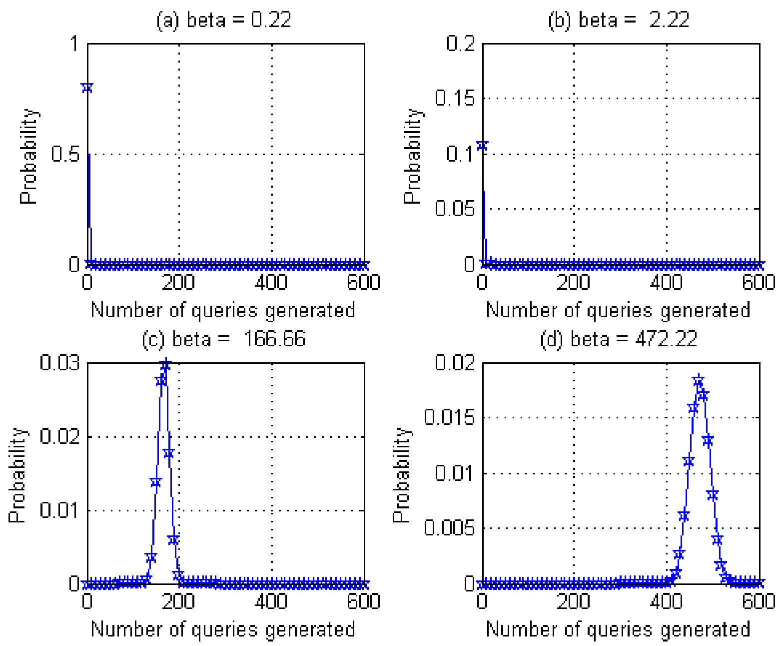


Figure 4.2: PMFs inferred from β on aggregating λ_I with slower and faster queries inter-arrival-time-rate

4.4 Clustering schemes and WSN network parameters

The probabilities estimation using equations (4.2) and (4.4) put forward inference about spatial and temporal parameters dependent amount of query generated respectively. It is presumed that N sensor nodes are randomly deployed in a given service-area and on using the four clustering schemes namely, static k-means (SKM), static fuzzy c-means (SFCM), dynamic k-means (DKM) and dynamic fuzzy c-means (DFCM) these sensor nodes form the clusters. Subsequently by means of multi-hop wireless communication through the cluster heads; sensor nodes and sink node exchange query and response in WSNs. In this thesis, the network performance is measured in terms of optimal energy utilization, which is primarily governed by strategies such as cluster-head selection and relocations of sink node. The network nodes energy dissipation pattern is measured using energy reserve indices and is presented in next two sections for spatial and temporal parameters driven query generation process. During simulation exercise for dichotomy of query generation mechanism; we presumed that the whole service-area is monitored by a single stationary sink, a single portable sink and four stationary sinks. Network surveillance by these sink models lead to three distinct case studies. The other sensor network specifications such as size of service-area, number of sensor nodes deployed in a service area (N) , the number of cluster formed at an arbitrary instant, the initial residual energy of each sensor node are listed in Table 4.2. In all these case studies; the network stipulated lifetime is considered as 8 years.

4.5 Hypothesis and Simulation Results (HSR) of Spatio-temporal parameters dependent query generation process

This section comprises of two different sub-sections which include hypothesis and simulation test results for three different sink attributes driven scenario under the impetus of spatial and temporal parameters associated with query generation process.

4.5.1 HSR for spatial parameter driven query generation process

The energy performance measures are analyzed at two levels namely, micro-level (cluster-level) and macro-level (network-level). At cluster-level, the computation exercise includes tasks such as clusters formation (initially randomly) and later based on residual energy status (RES) of individual sensor nodes, and residual energy estimation of each sensor node at periodic/sporadic time instants. Subsequently, the cluster heads (CHs) are deputed for each cluster dynamically till the network attains its stipulated lifetime. Further, the energy-centroid (EC) of each cluster is estimated. The EC estimation depends on the spatial aspects (rectangular coordinates) of associated sensor nodes and their RES status [(Kumar and Chaturvedi, 2015)]. To determine energy performance measures at the network-level; based on dynamicity of EC, the optimum locations of sink(s) are estimated. Positioning of sink(s) at optimum locations avoids occurrence of "hot spot phenomenon" while ensuring all requisite service norms of sensor network till it attains stipulated lifetime. In all the subsequent case studies; during the clusters formation phase, term "static" refers to a situation in which clusters once formed they remain structurally unchanged. More precisely, the clusters composition in terms of participating sensor nodes (SNs) remains time invariant. Thereby, a specific set of sensor nodes remains integral part of a particular cluster during entire operation period. In contrast, the term "dynamic" refers to changing scenarios in which cluster formation takes place from time to time (not necessarily periodic) and it is primarily based on the RES of sensor nodes. On using four different clustering schemes for three different sink paradigms; the energy performance measures are analyzed and compared in the following three subsections.

4.5.1.1 Sensor network surveillance by a single stationary sink (SSS)

In broad perspective; the spatial parameter driven query density can be classified into two types, i. e., sensor networks having low density and high density of queries. In these case studies; we presumed that the location of a single stationary sink is fixed at the center (origin) of the square-shape service-area. The network energy profile is evaluated in terms of performance measures considering four different clustering schemes namely, SKM, SFCM, DKM and DFCM.

The network's energy performance measures include ARES estimation of all network nodes and CRES estimation of individual sensor nodes. Corresponding to lower and upper values of α-bounds, i. e., α_l and α_u as listed in Table 4.1; the ARES characteristics (measured in joules (J)) are shown in Fig. 4.3(a & b) and Fig. 4.3(c & d) respectively. These observations were recorded periodically till the network attains stipulated lifetime. On analyzing these characteristics we observed that in terms of energy efficiency aspect, the DFCM performs better than other three clustering schemes. Uses of DFCM ensures that the network ARES status remains in relatively healthy state with respect to predetermined CRES mark as the network service time approach stipulated lifetime. As an exception, for lower α-bounds; the network ARES attains CRES mark prematurely (with reference to stipulated lifetime), otherwise the uses of DFCM ensures that the network is fully operational till it attains the expected lifetime duration.

For diversified WSN scenarios that comprise of uses of four clustering schemes and lower and upper values of α-bounds, service-time-duration (STD) by which network ARES attains predetermined CRES mark is estimated and summarized using histograms in Fig. 4.4. These histograms give network STD estimate in years for specific combinations of clustering schemes and α-bound intervals. These histograms also endorse DFCM as the most promising scheme for clusters formation. Further, these histograms establish that corresponding to relatively higher values of query density that stem from higher value of α ; it is the only DFCM scheme which results into the STD attaining the stipulated network lifetime.

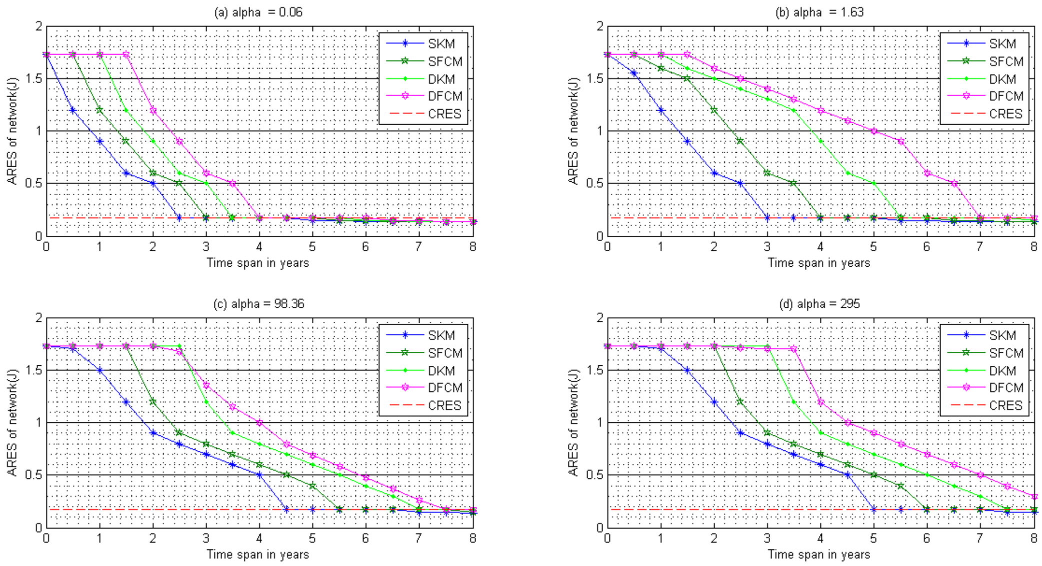


Figure 4.3: Network ARES for lightly and highly dense queries under SSS scenario

4.5.1.2 Sensor network surveillance by a single portable sink (SPS)

In this surveillance scenario; we presume that initially the sink is stationed at the center (considered as origin) of the service-area and as time progresses it can be relocated to arbitrary optimum locations. The process repeats itself and the likely new location of the sink relies upon the past energy consumption pattern of the network nodes. In principles, the ARES of network nodes and thereby the energy centroid (EC) of clusters primarily decides relocation aspect of the sink. Analogous to the single stationary case; here too, we deploy four clustering schemes namely SKM, SFCM, DKM and DFCM.

Similar to SSS based network scenario; all these four clustering schemes are tested for

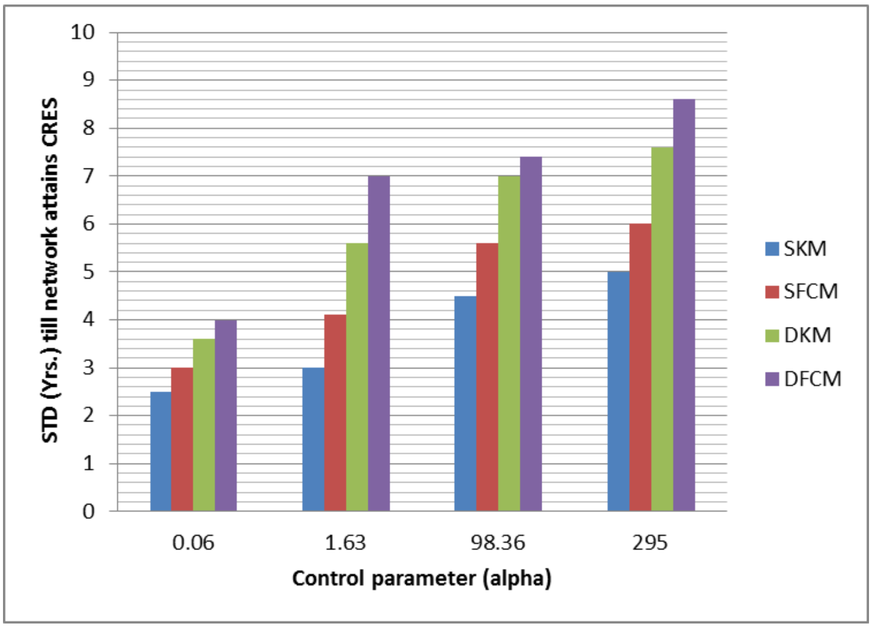


Figure 4.4: Network STD corresponding to four clustering schemes and various α-bounds under SSS scenario

the lower and upper α-bound intervals. For these lower and upper values of α-bounds; the ARES estimations are shown in Fig. 4.5 (a & b) and (c & d) respectively. In this scenario also, the ARES characteristics clearly indicate that DFCM performs better compared to the rest of the three clustering schemes. Further, in particular Fig. 4.5 (c & d) shows that for the upper α-interval, i.e., $\alpha_u = \{98.36, 295\}$; uses of DFCM results into a network state, wherein the ARES still maintains sufficient margin with respect to predetermined CRES mark even at the lifetime instant. Further, on combining the four clustering schemes with α-bound intervals the network STD histograms are shown in Fig. 4.6. It also validates that on using the specific combination of DFCM and the upper α-intervals i.e., $\alpha_u = \{98.36, 295\}$ the network STD exceeds the stipulated lifetime span of 8 years while still maintaining ARES well above the predetermined CRES mark.

4.5.1.3 Sensor network surveillance by four stationary sinks (FSS)

As compared to the previous two scenarios, in this case; instead of a single sink, the service-area is monitored by four stationary sinks. The locations of four sink nodes are fixed at the center of periphery of the service-area at the onset of surveillance operation. Corresponding to the lower and upper bound α-intervals; on deploying four clustering schemes, the ARES characteristics are shown in Fig. 4.7 (a & b) and (c & d) respectively. All these ARES characteristics clearly indicate that among the four clustering schemes, DFCM emerges as the most efficient clustering scheme. Further, for the upper α-bound interval; on comparing ARES characteristics with that of Fig. 4.3 (c & d) and Fig. 4.5(c

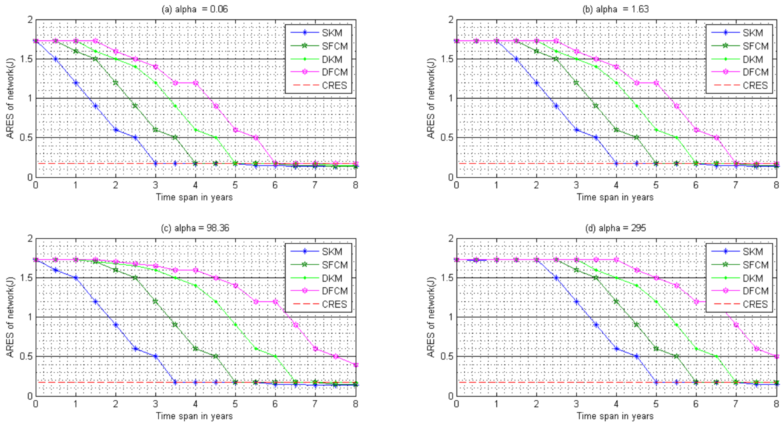


Figure 4.5: Network ARES for lightly and highly dense queries under SPS scenario.

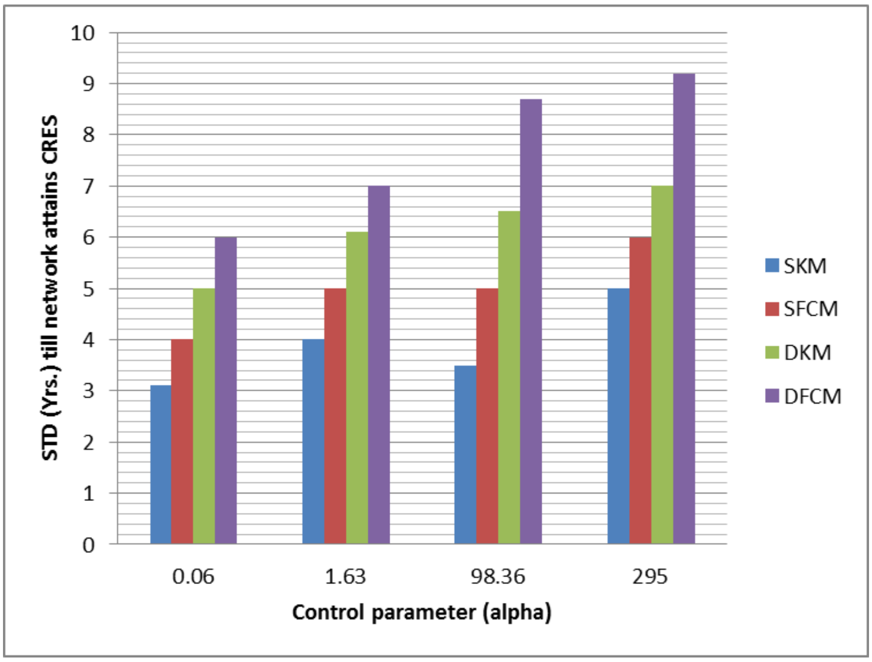


Figure 4.6: Network STD corresponding to four clustering schemes and various α-bounds under SPS scenario

& d) establishes two perspectives: (i) among the four clustering schemes; DFCM delivers the superior energy performance and (ii) ARES or energy-reserve status of entire network maintains relatively better profile under the four stationary sinks based surveillance compared to single stationary/portable sink based surveillance scenarios.

Observation drawn at lifetime mark of 8^{th} year for upper α-bound interval also reveals that on deploying DFCM scheme, the energy difference margin between the ARES and the CRES mark monotonically increases as we move from single stationary sink based surveillance to single portable sink based surveillance and finally to four stationary sinks based surveillance. This margin not only ensures maintaining the operational aspects at the stipulated lifetime mark but also indicates a sort of surveillance capability beyond the expected lifetime.

In summary, for the lower and upper α-bound intervals; the approximate STD in years are shown as histograms in Fig. 4.8 on deploying four clustering schemes under four stationary sinks based surveillance. Here, few noteworthy observations are: (i) DFCM always outperforms over other three clustering schemes irrespective of α-bound values, hence no dependency on query density amount, (ii) STD increases linearly with α irrespective of clustering schemes. Although, exact value of STD has dependency on both, i. e., chosen α as well as the type of clustering schemes used.

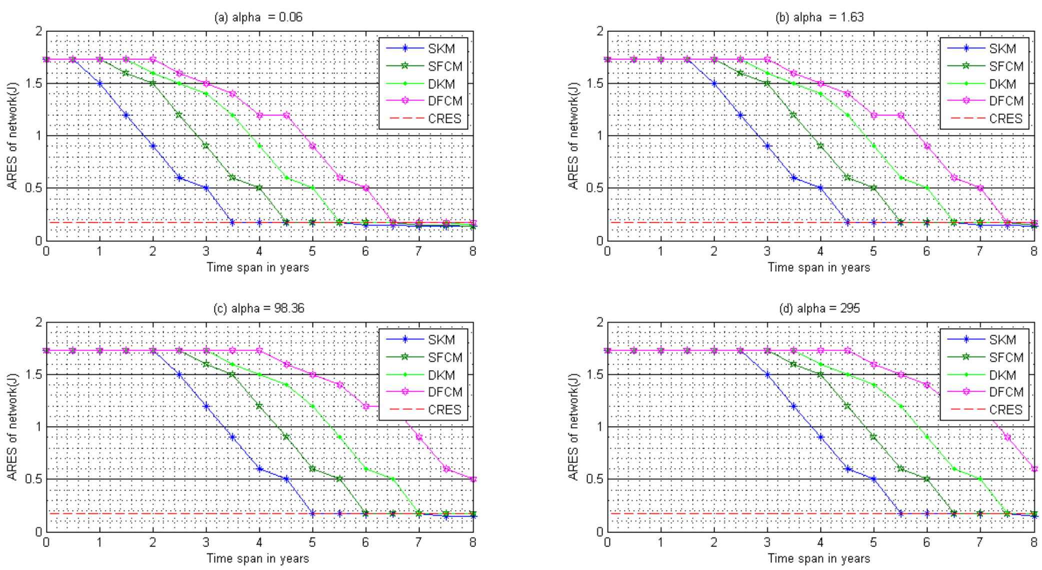


Figure 4.7: Network ARES for lightly and highly dense queries under FSS scenario

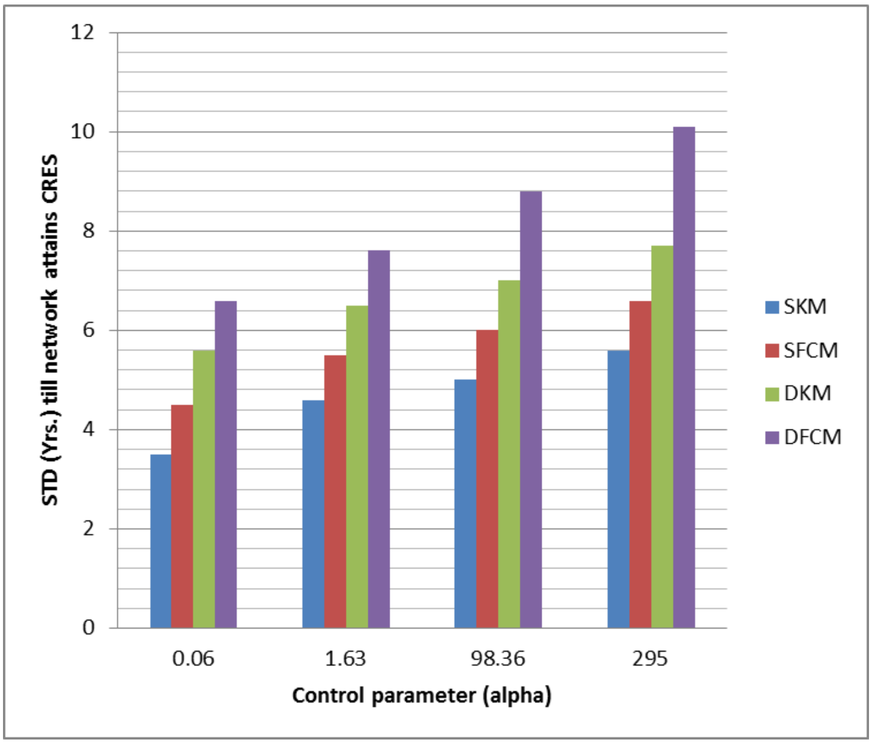


Figure 4.8: Network STD corresponding to four clustering schemes and various α-bounds under FSS scenario

4.5.2 HSR for temporal parameter driven query generation Process

In this section, under three different sink attributes driven scenarios; the query arrival rate or query inter-arrival-time-rate (Δt) is classified into two types (i) query generated at a slow rate (Δt_s) and (ii) query generated at a faster rate (Δt_f). Similar to previous case studies, here also we consider uses of four different clustering schemes. To analyze the network energy performance, we estimate the measures such as RES of individual nodes, ARES of network, relative value of these measures with CRES mark, and service-time-duration (STD). All these performance measures were obtained in a backdrop of scenarios wherein the probability of query generation using equation (4.4) relies upon the associated temporal parameter. The estimate of all these performance measures corresponding to three different sink driven scenarios is presented in the following sub-sections.

4.5.2.1 Sensor network surveillance by a single stationary sink (SSS)

Analogous to section 4.5.1.1, it is presumed that a single sink is positioned at the centre (origin) of the service-area. On using four different clustering schemes; for the lower and upper β-bound interval values, the ARES characteristics are shown in Fig. 4.9 (a

& b) and Fig. 4.9 (c & d) respectively. Predetermined CRES mark is also indicated in all these ARES temporal characteristics as dashed line and the energy margin difference between instantaneous ARES and CRES mark is treated as network's energy health indicator. More the energy margin better is the network energy reserve status. These ARES characteristics as shown in Figures (4.9a - 4.9d) indicate that undoubtedly among the four clustering schemes, the DFCM performs better than its contemporaries. During this single-stationary sink based surveillance scenario; ARES maintains healthy status for relatively longer duration of time as the network experiences higher amount of query density. However, from reliable service perspective with respect to stipulated lifetime, uses of DFCM scheme along with relatively faster generation of queries (here in Fig. 4.9 d, $\beta = 472.22$) is the only peculiar combination that maintains network's ARES closer to the predetermined CRES mark.

On combining the uses of four different clustering schemes with the lower and upper bound β- intervals; the network STD in a form of histograms are shown in Fig. 4.10. These histograms reveal some important aspects about the network's sustainable operation over the stipulated lifetime. Irrespective of query inter-arrival-time-rate (Δt_s or Δt_f) considered in this section with extreme β values such as $\beta_{min.} = 0.22$, $\beta_{max.} = 472.22$; none of these clustering schemes are able to attain an STD that approximates stipulated lifetime. Even the uses of DFCM scheme with $\beta_{max.} = 472.22$; also fails short by substantiate margin with respect to expected lifetime mark of 8^{th} year. As a consequence of it only remedy is either to forcibly cut short the lifetime span of the network for the reliable operation or to compromise with the operational reliability feature of the network as the hot-spots creations are inevitable during the later stage of stipulated lifetime. Subsequently, these aftereffects could lead to premature attainment of CRES mark. Only thing that favors DFCM is that on using it as clustering scheme, we can defer the hot-spots phenomenon by the maximum possible extent but still it could not be avoided over the stipulated lifetime span especially when it matures. To overcome some of these difficulties, we explore the two other types of sinks attributes driven surveillance scenarios in subsequent sections.

4.5.2.2 Sensor network surveillance by a single portable sink (SPS)

To investigate the impact of sink mobility on network's performance measures, it is presumed that at the onset of network operations the sink is stationed at the centre of the service-area and subsequently depending upon dynamicity of energy gradient among the network nodes, it can be relocated to arbitrary optimum locations. On exploiting the aforementioned four clustering schemes with the lower and upper bound β- intervals; the network ARES characteristics are shown in Fig. 4.11(a & b) and Fig.4.11 (c & d)

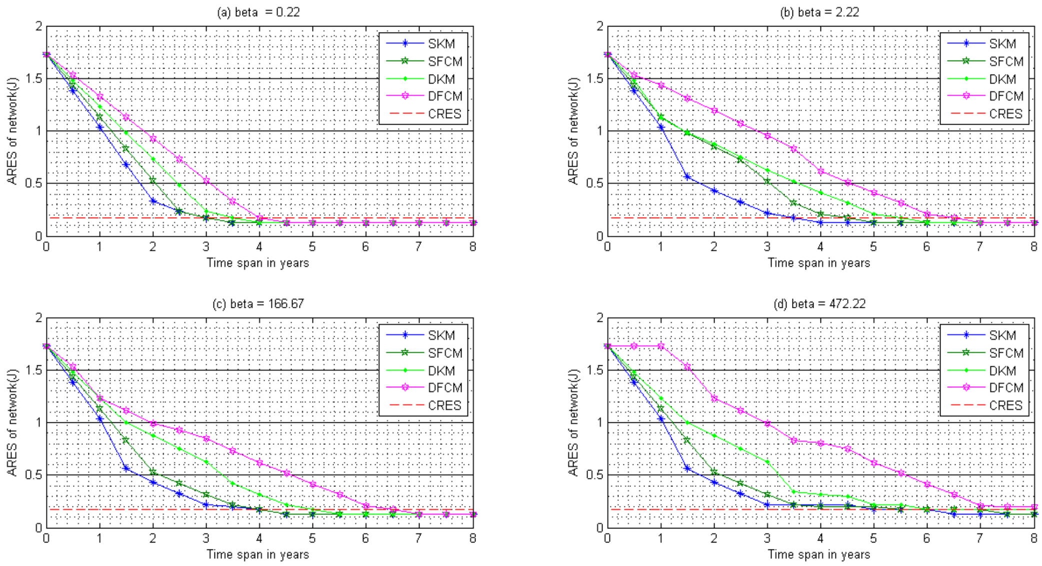


Figure 4.9: Network ARES for slower and faster query generation rate under SSS scenario

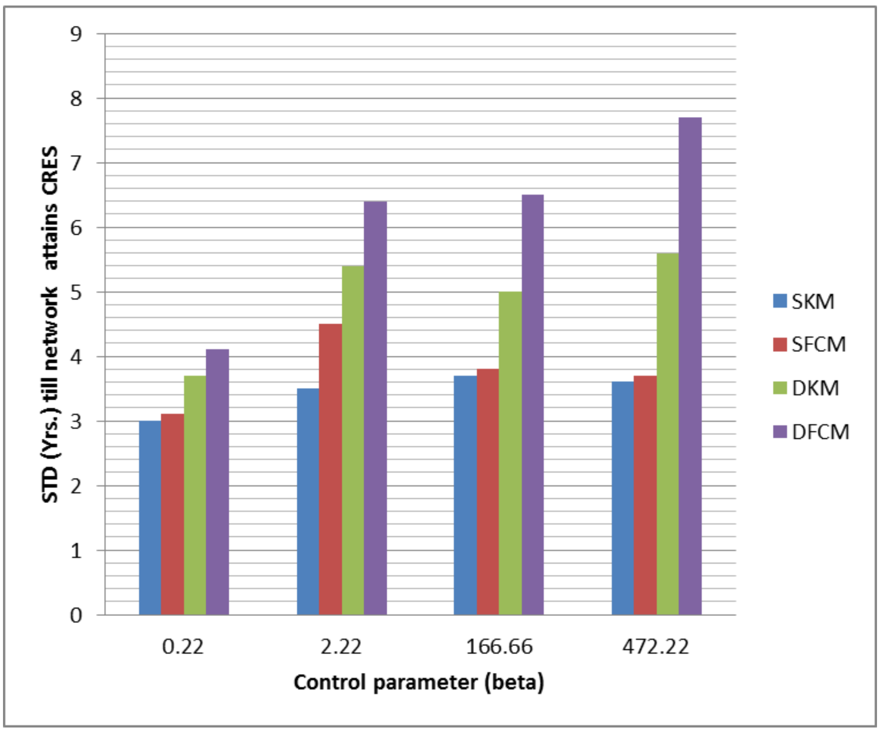


Figure 4.10: Network STD corresponding to four clustering schemes and various β-bounds under SSS scenario

respectively. Unlike the ARES measure obtained from a single stationary sink driven network scenario as shown in Fig. 4.9, in this case; except for extremely low value of β, i.e., $\beta = 0.22$, for other values of chosen β, the uses of DFCM ensures that the energy margin difference between the predetermined CRES estimate and the instantaneous ARES measure at stipulated lifetime mark remains a positive real quantity. These ARES characteristics also reveal that irrespective of β values, DFCM always ensures maximum service span as compared to other three clustering schemes. In a complementary manner, network ARES based STD measure is shown as histograms in Fig. 4.12 on combining four clustering schemes with the lower and upper bound β- intervals. These histograms clearly indicate that except for $\beta = 0.22$, deploying DFCM ensures the service operative norms up to and beyond the stipulated lifetime mark of 8th year. This specific observation yields inference about the network longevity on using DFCM as clustering scheme and thereby substantiates a sort of assurance that avoids plausible islands creation within the service-area over the stipulated network lifetime. In the next case-study; we explore implementation of four-stationary sinks for service-area surveillance and its impacts on network performance measures.

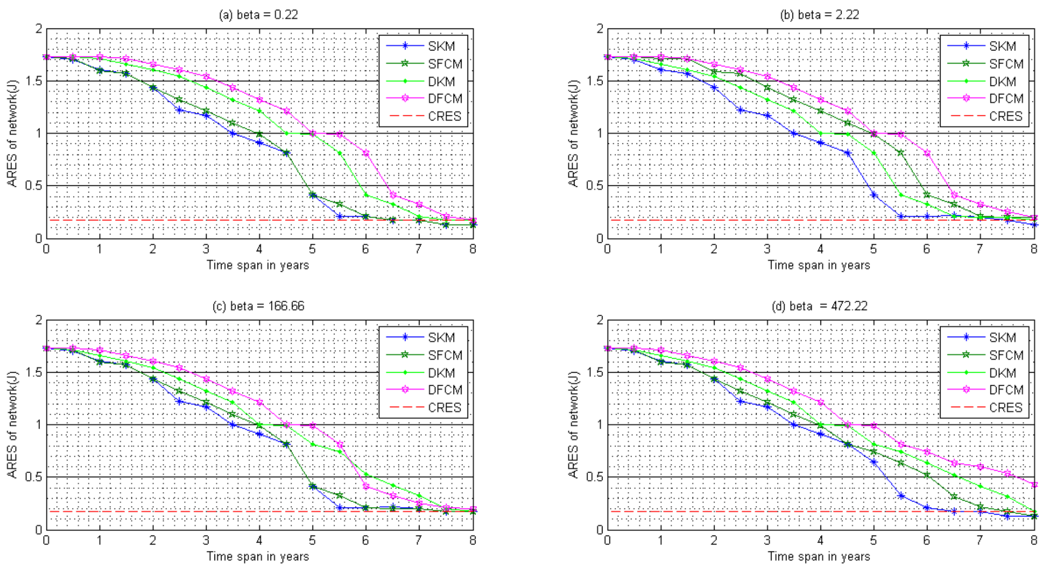


Figure 4.11: Network ARES for slower and faster query generation rate under SPS scenario

4.5.2.3 Sensor network surveillance by four-stationary sinks (FSS)

In this surveillance scenario, the sink nodes physical specifications as well as other operating conditions are identical to that of narrated in section 4.5.1.3, except the difference that during this case study the amount of query generated depend upon the temporal

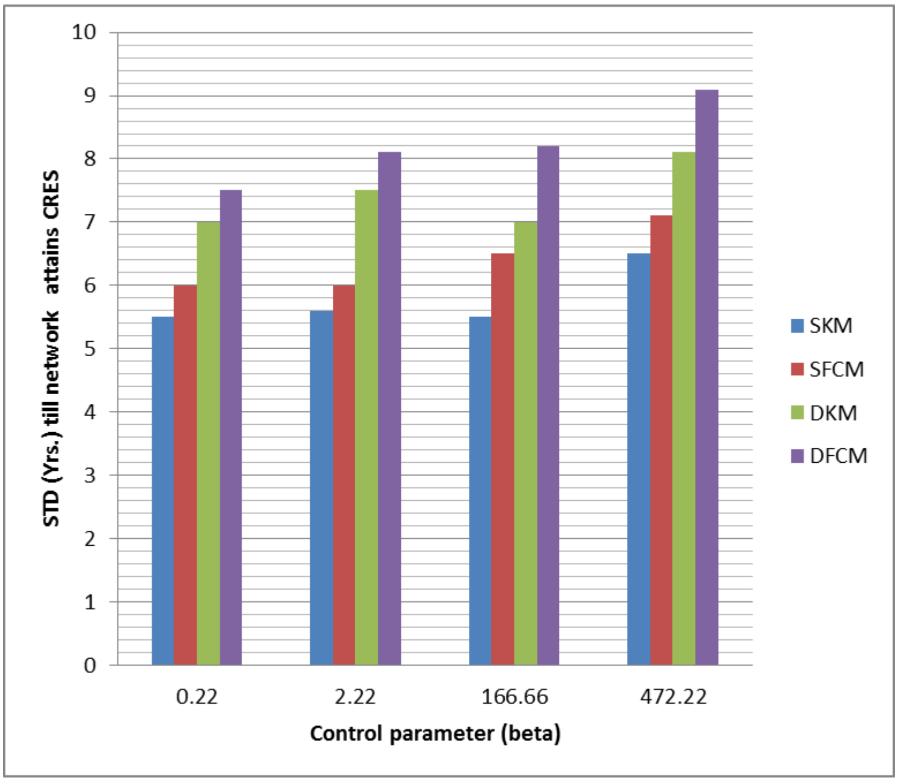


Figure 4.12: Network STD corresponding to four clustering schemes and various β-bounds under SPS scenario

parameter (inter-arrival-time-rate) associated with query generation process. Under the four stationary sinks based surveillance scenario; the typical combination of four clustering schemes with the lower and upper bound β- intervals result into ARES measure estimation and are shown in Fig. 4.13(a & b) and Fig. 4.13(c & d) respectively. Contrary to all previously reported ARES characteristics, interesting ARES characteristics surface in Fig.4.13. In all of these ARES characteristics, irrespective of β values, the implementation of DFCM scheme assures reliable network operation till the network attains stipulated lifetime.

A close analysis of all these ARES characteristics indicate that at the time of stipulated lifetime maturity, in addition to ensuring service operative norms, uses of DFCM also assures that the energy margin difference between the predetermined CRES mark and ARES measure increases with incremental β values. The improved energy margin and thus the ARES profile indicate sustainability of network's operational aspects for a period beyond the expected lifetime. As a complementary measure for the network energy health indicator, STD measures in a form of histograms are shown in Fig. 4.14. Relatively these high rising histograms establish merits of four stationary sinks based network surveillance over the single stationary/portable sink based network surveillance.

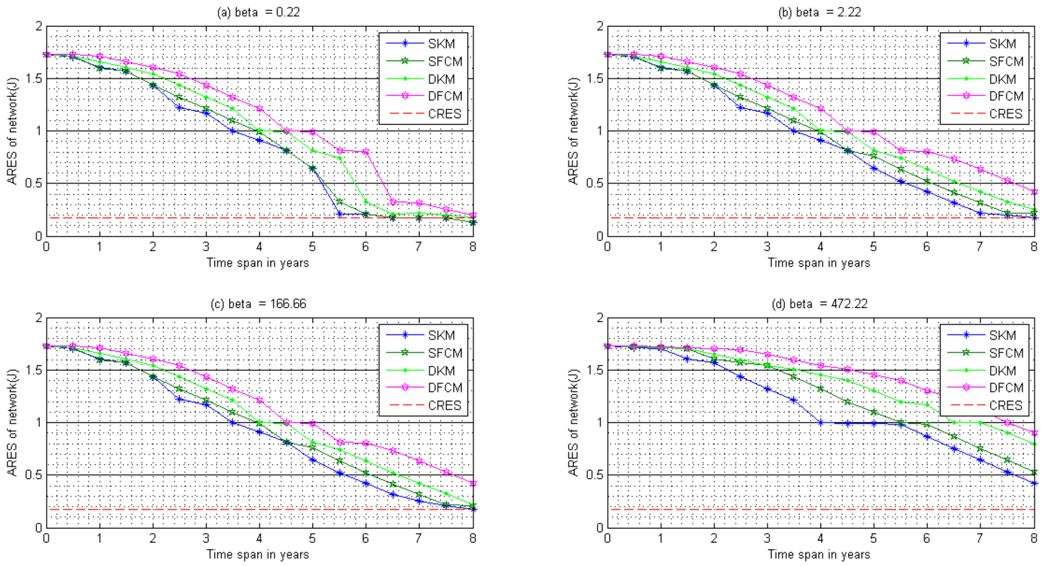


Figure 4.13: Network ARES for slower and faster query generation rate under FSS scenario

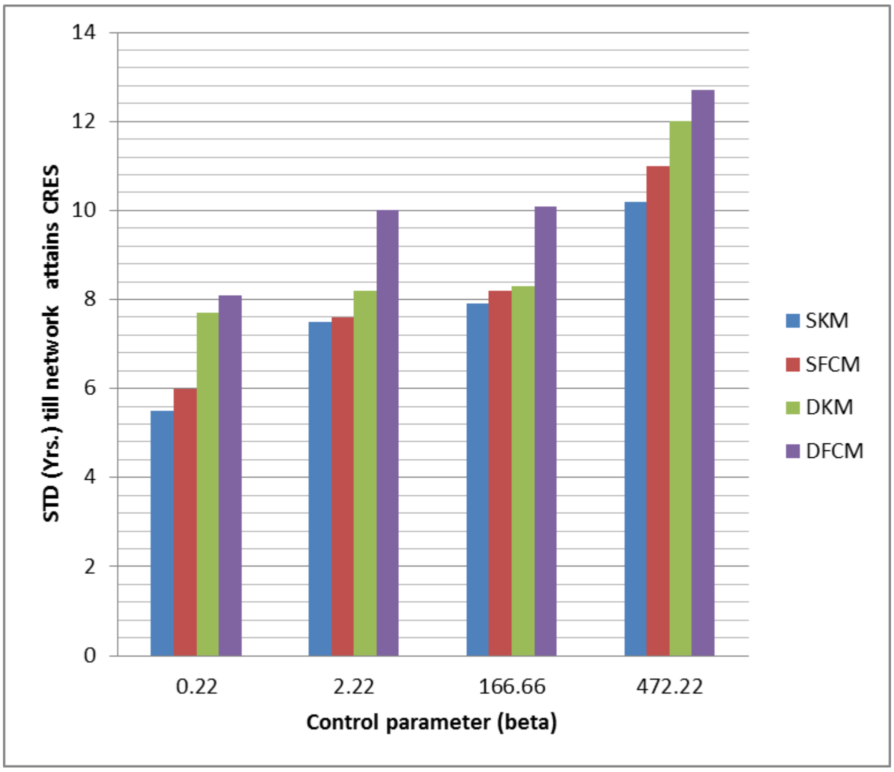


Figure 4.14: Network STD corresponding to four clustering schemes and various β-bounds under FSS scenario

4.6 Conclusions

A large number of WSN architecture based sensing/measurement applications cover wide range of query density as well as the queries inter-arrival-time-rate. Both of these spatial and temporal aspects that govern the amount of query generated possess a reasonable degree of uncertainty. Owing to this uncertainty and non-deterministic nature of query arrival process; we presented few amendments in the universally adopted Poisson PMF model of query generation process. In that, the control parameter (λ) is modified so as to incorporate its dependency on spatial and temporal parameters. Further, to encompass the uncertainty associated with numbers/scalars, rather than using crisp value of these parameters, parameters are modeled using appropriately chosen intervals. Other two significant contributions of this chapter are: (i) validation of DFCM as the most prominent clustering scheme and (ii) exploring the sink attributes for optimal uses of energy; in that four-stationary sinks based network surveillance outperforms over the single-stationary/portable sink based network surveillance. Further, the two principal measures, the ARES estimation with respect to predetermined CRES mark and the network STD with respect to stipulated lifetime establish the fact that for a given service-area; combination of DFCM as the cluster formation scheme and four-stationary sinks based service-area surveillance appears an optimal choice ensuring that the sensor network attains the stipulated lifetime.

During the entire simulation work, it is presumed that the sensor nodes are randomly deployed in a given service-area, and these sensors form the clusters as per the k-means and fuzzy c-means algorithms. The details about important network parameters such as number of nodes, number of clusters, nodes density in an arbitrary cluster, initial energy of sensor nodes and the predetermined energy level for CRES are given in Table 4.2.

Table 4.2: Network Simulation Parameters

Sr. No.	Network Parameters	Value
1	Service area; square in shape having its area	250X250 m^2
2	Number of sensor nodes deployed	2308
3	Number of clusters & cluster heads at an arbitrary instant (t)	278
4	Average number of sensor nodes in each cluster at t	8
5	Initial energy of each sensor nodes	1.732 J
6	Critical residual energy status(CRES)	10% of initial energy; 0.1732J

Chapter 5

UNIFIED FUZZY INTERVALS POISSON DISTRIBUTION MODELS OF QUERY GENERATION

5.1 Introduction

Proliferation in Micro-Electro-Mechanical-Systems (MEMS) technology along with advancement in distributed computing infrastructure has facilitated the versatile usage and deployment of wireless sensors networks (WSNs) in last one and half decades. WSNs support large number of applications from the civilian and military regimes. Irrespective of these regimes; owing to difficulty associated with battery replenishment, proper energy usage has been at centre stage in WSNs operations. In a majority of wireless sensors networks (WSNs) architectures; sensor nodes possess limited resource of energy (battery), thus it must be utilized with proper care. The lifetime of WSNs depends upon the optimum utilization of battery, which in turn governed by the routing strategy adopted for packet transmissions. For query based WSNs, network lifetime depends upon: the amount of generated query, inter-arrival-time-rate of query occurrences (frequency of query generation), and the spatial distribution of query.

In query based WSNs; sensors establish bidirectional communication with sink(s) either using single/multiple-hop links. Although, during downlink (from outside world to specifically located sensor via internet gateway and sink(s)), and uplink (from a sensor to the outside world via sink(s) and internet gateway) commissioning; the participating sensor nodes may form disjoint sets. Typically in sensor networks; sensors energy dissipation pattern is non-homogeneous with respect to spatial distribution over any short epochs. The genesis behind this non-homogeneity is random generation of queries which owes to application specific spatio-temporal parameters. Importance of spatio-temporal parameters is ubiquitous in WSNs paradigm and uncertainties are inevitable with these parameters, although the degree of uncertainties varies in accordance to applications served. Thus, from network design perspectives, precision involved with spatio-temporal aspects must be given due priority to obtain a mathematical model that maintains a good rapport with realistic query generation process.

5.2 Literature Survey

[(Akylidiz et al., 2002)] presented a survey that contains important operational dynamics of sensor network. Usually, sensors are deployed in a scattered manner in a given service-area. Mutual cooperation among these sensors results in high-quality spatio-temporal information about the local physical entities under observation/surveillance. Each of these scattered sensors collects/aggregates the received information (data/packets) and en-route it either to other sensors or back to an external sink(s). These sink(s) may be stationary/portable/mobile and connect the WSNs to the outside world using Internet gateway, which enable users an easy access to the reported data [(Karaki and Kamal, 2004)]. A Comprehensive snap-shot about some of the research methodologies is presented next.

[(Yoon and Shahabi, 2007)] presented clustered Aggregation (CAG) algorithm; it was formulated based on assumption that the sensed data in WSNs exhibit the spatial and temporal correlation of physical attributes existing in the sensing field. [(Takruri et al., 2011)] proposed a spatio-temporal model that takes into consideration drift associated with measurement process. To predict the future measurements, support vector regression algorithm was used, whereas, to detect and correct associated drift and random error Kalman filter is used to prolong the network lifetime.[(Jain et al., 2005)] observed that the WSNs can significantly improve the quality of spatio-temporal data monitoring. The authors describe a communication architecture that supports distributed query processing to evaluate spatio-temporal queries within the network.

[(Mousavi et al., 2013)] proposed a Spatio-temporal event detection algorithm. The algorithm deliberates probabilistic graphical models (PGMs) of WSNs; it incorporates the Markov chains in temporal dependency and Markov random field's theory in the spatial dependency of sensors in a distributed fashion.[(Cheng et al., 2003a),(Cheng et al., 2003b)] reported probabilistic query evaluation based upon uncertain data. In that classification of probabilistic queries is done using attributes namely, value-based non-aggregate class, entity-based non-aggregate class, and entity based aggregate class etc. and obtained measurements were presented in interval from. [(Dutta et al., 2005)] proposed an experimental work to detect rare, random, and ephemeral events based on observations drawn from WSN.

[(Deshpande et al., 2005)] presented a model based approximate querying in WSNs; it reports a selection strategy for best sensor reading to acquire live data with commendable degree of confidence. However, in work reported in this thesis; the inferences drawn from probabilistic model have primary focus on uncertainty associated with spatio-temporal parameters and thus maintain substantiate differences. [(Wang et al., 2004)] presented a mathematical model, wherein the localization uncertainty is estimated and compared using Bayesian bound (BB) with Cramer-Rao bound (CRB). Here, mainly the localization uncertainty is attributed to the network topology.

[(Wang et al., 2014)] proposed a lifetime enhancement method for WSNs; the method operates on sink relocation aspect to prolong the occurrence of *hot spot*. The reported work emphasises on time complexity analysis and numerical analysis for sink relocation mechanism. [(Nachabe et al., 2015)] reported a unified data model for WSN in that the proposed solution comprises of a semantic open data model. The whole emphasis was on processing, storage and transmission of data, while inherent uncertainties of data were ignored, which in a way affect these capabilities. In summary, the above mentioned algorithms utilize diversified approaches such as uses of specific routing strategies and exploiting the spatio-temporal information with emphasis on energy-efficient operation of WSNs.

In principal, this chapter investigates energy efficiency aspect of network on incorporating three diversified issues, which are (i) exploring energy efficiency associated with different clustering schemes, (ii) incorporating query generation dependence on spatio-temporal aspects to predict network performance more precisely, and subsequently devising a strategy to utilize network resources more appropriately, and (iii) exploiting sink multiplicity and location issues so as to maintain balanced energy gradient across the network. The first issue is addressed by deploying four clustering schemes namely static k-means (SKM), static fuzzy c-means (SFCM), dynamic k-means (DKM) and dynamic fuzzy c-means (DFCM) algorithms. The second issue needs a mechanism that makes query generation process sensitive to spatio-temporal parameters; it is achieved by incorporating the spatio-temporal parameters in Poisson PMF expression. Towards the end to address third issue, we need balancing mechanisms that ensure avoidance of sensor nodes *"islanding"* thereby ensuring uniform coverage. To attain it, the network is monitored by single-stationary/portable sink and four-stationary sinks. In portable-sink driven network surveillance; sink relocation aspect is governed by residual energy status

(RES) estimates of sensor nodes. In variety of network surveillance scenarios; network's energy performance is estimated in terms of the following measures: average residual energy status (ARES) estimation of all participating sensor nodes at cluster as well as at network level, and the network service-time-duration (STD) by which the network attains the predetermined critical residual energy status (CRES). These energy measures are estimated periodically and at the stipulated lifetime of the network.

5.3 Overview and General Framework for Query Generation Process

Two inevitable aspects associated with query generation models are its spatial distribution and temporal occurrences. During any short epochs, queries are not ubiquitous or in alternative term, within entire network, spatial distribution of queries is not concurrent. Further, accuracy of query generation model depends upon inclusiveness of the precision associated with spatio-temporal parameters. Thereby, incorporation of spatio-temporal parameters with finest possible granularity would definitely result in a well approximate model of query generation process. Primarily, the inference drawn from spatial-temporal parameters and uses of probabilistic approach integrated with fuzzy-intervals to obtain more approximate model of query generation mechanism are presented in this chapter.

5.3.1 Universal aspect and Poisson distribution model of query generation process

In a plenty of service-area surveillance; the two persistent facts are: at one hand the spatial locations where events (physical phenomenon) may happen is usually large, whereas, on the other hand, the spatial locations where events take place (actually happen) are relatively sparse. Owing to these facts, the sensors are usually densely distributed in the service-area to support reasonably good coverage, and thereby avoid possibilities of *"islands"*(*a small geographical location deprived of connectivity*) occurrences at any arbitrary locations within the service-area.

In variety of network paradigms; to model the random arrival of events, the Poisson probability mass function (PMF) manifests itself as one of the best contender. In wide ranging WSNs scenarios; spatio-temporal parameters associated with query generation process are random, thus it can be approximated using an appropriate PMF. So far, for typical WSNs scenarios; the query generation patterns have been modeled as uniform PMF and Poisson PMF with associated parameter λ having a crisp value. Owing to the inherent randomness with query inter-arrival time-rate (temporal aspects) and query density (spatial aspects); the query dynamic distribution as the uniform PMF may be a good approximation on aggregate basis over the stipulated network lifetime. However, it lacks in following the spatio-temporal dynamics involved with query generation process. Thus, it is presumed that the query generation in any finite observation interval (FOI)) is modeled using Poisson PMF, on treating Q as a random variable for generated queries the Poisson PMF having dependence on control parameter λ is given by

$$P(Q = k) = \frac{\lambda^k}{k!} e^{-\lambda} \tag{5.1}$$

Where, k = 0,1,2,3....

5.3.2 Spatio-temporal parameters dependent Poisson PMF model

In Poisson PMF expression given in equation (5.1); contrary to conventional way of choosing λ as an arbitrary scalar, a mathematical framework is proposed that establishes functional dependency of control parameter λ on the spatio-temporal aspects associated with a query generation process. For it, equation (5.1) is amended as:

$$P(Q = k) = \frac{\beta^k}{k!} e^{-\beta} \tag{5.2}$$

Where, k=0,1,2,3...
Here, β is a parameter inferred from spatio-temporal parameters. Depending upon crisp or bound interval-values of spatial and temporal parameters; the parameter β in equation (5.2) is modeled as:

$$\beta = \delta_s \lambda a + \delta_t \lambda \tau \tag{5.3}$$

$$\beta = \delta_s \Delta\lambda \Delta a + \delta_t \Delta\lambda \Delta\tau \tag{5.4}$$

Variables used in equation (5.3) and (5.4) are defined as: δ_s and δ_t are weight/scaling factor associated with spatial and temporal parameters respectively. $\Delta\lambda$ is the chosen interval for λ parameter, i.e., a closed bound set of λ, a and τ are the chosen crisp value associated with query density and query inter-arrival-time-rate [(Cheng et al., 2003a),(Huang et al., 2007) and (Cheng et al., 2003b)] respectively. Intervals Δa and $\Delta\tau$ comprise the lower and upper bound interval-values of query density, and query inter-arrival-time-rate respectively. In equation (5.4) on performing the product; the final outcomes are the intervals having lower and upper bound values of β (as per the product rule of interval arithmetic [(Rokne, 2001)]) as:

$$[a_1\ a_2] * [b_1\ b_2] = minimum\ \&\ maximum\ of\ [a_1b_1, a_1b_2, a_2b_1, a_2b_2] \tag{5.5}$$

Relative significance of spatial and/or temporal parameters governs the weight factors δ_s and δ_t. Usage of bound interval-values around crisp/scalar variables though addresses the vagueness/uncertainty associated with these scalar variables. However, it has the associated drawbacks as well, as the apprehension for intervals formation is based on hypothesis that during whole interval-span the degree of uncertainty presumes uniform distribution characteristic. Thus, to cope with imprecision involved with these ambiguous variables; uses of fuzzy-triangular distribution characteristic is presented next.

5.4 General Fuzzy Triangular Characteristics

The typical fuzzy-triangular distribution characteristic is shown in Fig. 5.1 and is mathematically expressed as:

$$A(x : a, m, b) = max \left\{ min \left[\frac{(x-a)}{(m-a)}, \frac{(b-x)}{(b-m)} \right], 0 \right\} \tag{5.6}$$

In this expression, the parameters a, b locate the *"feet"* and the parameter *"m"* locates the peak of the triangular characteristics as shown in Fig.5.1. Alternatively, scalar *"m"* can be interpreted as singleton/crisp-number having no uncertainty and interval [a, b] treated as lower support and upper support value of an interval corresponds to zero $\alpha - cut$. For better spatio-temporal resolution and thus to address uncertainty with fine granulation, the triangular characteristic is segmented into finite number of base spans having specific $\alpha - cuts$ (in this chapter, five base spans and correspondingly five $\alpha - cuts$ factors are considered) [(Zimmermann, 2001)].

5.4.1 Fuzzified intervals based spatio-temporal integrated Poisson PMF

Contrary to presuming that all the scalar values within a given interval bound follow a uniform distribution probability model, to address the associated uncertainties of chosen intervals at microscopic level; it is considered that interval bounded parametric values of $\Delta\lambda$, Δa and $\Delta\tau$ are appropriately scaled using $\alpha - cuts$ inferred from independent fuzzy-triangular distribution characteristics. On incorporating it, the PMF expression of

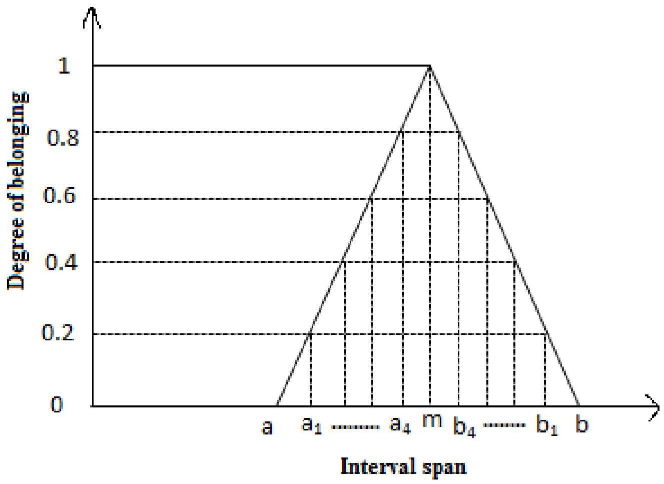


Figure 5.1: A typical fuzzy-triangular distribution characteristic

equation (5.2) is amended as:

$$P(Q = k) = \frac{e^{-\beta_{F-I}}}{k!}(\beta_{F-I})^k \tag{5.7}$$

With β_{F-I} is modeled as:

$$\beta_{F-I} = \delta_s \Delta\lambda_{F-I} \Delta a_{F-I} + \delta_t \Delta\lambda_{F-I} \Delta\tau_{F-I} \tag{5.8}$$

Here, β_{F-I} represents spatio-temporal parameters driven fuzzy-interval bounds and F-I as subscripts with $\Delta\lambda$, Δa and $\Delta\tau$ denote fuzzification of $\Delta\lambda$, Δa and $\Delta\tau$ respectively. To accomplish this task, a mechanism is devised in that the extreme lower and upper values of these intervals are scaled by appropriately chosen $\alpha - cuts$ factors. Let the associated $\alpha - cuts$ factors for intervals $\Delta\lambda$, Δa and $\Delta\tau$ are represented by notations α_i, α_j and α_k respectively. Incorporation of inferences drawn from fuzzy-triangular characteristics about associated $\alpha - cuts$ result in

$$\beta_{F-I} = \delta_s(\alpha_i \lambda_i)(\alpha_j a_j) + \delta_t(\alpha_i \lambda_i)(\alpha_k \tau_k) \tag{5.9}$$

Here, each of these $\alpha - cuts$ factors belongs to independent and uncorrelated closed intervals of zero and unity ($0 \leq \{\alpha_i, \alpha_j \; and \; \alpha_k\} \leq 1$). Whereas, λ_i, a_j and τ_k are arbitrary element values spanned by the chosen lower and upper bound intervals that represent the base-span of triangular distribution characteristics. The base-span of triangular distribution for α_i, α_j and α_k encompasses the ranges $(\lambda_u - \lambda_l)$, $(a_u - a_l)$ and $(\tau_u - \tau_l)$ respectively. The estimated fuzzy-intervals along with simple chosen intervals are given in Table-5.1.

The outcomes inferred from pair of equations (5.9 & 5.7) yield probability estimation. Depending upon chosen spans of $\Delta\lambda$, Δa, $\Delta\tau$ and number of $\alpha - cuts$ (α-levels set); the probability estimation task becomes cumbersome. To overcome it; two different indices, which in aggregate terms includes the basic features of considered interval-bounds and subsequent imposition of fuzzification procedure on these intervals are proposed and discussed next.

5.4.2 Estimation of Scalar inferences from fuzzy-interval bounds

The computational tasks for probability estimation based on equations (5.2 and 5.7) are certainly in favor of making uses of scalar variables. For interval-bound variables, the probability estimation task becomes highly intensive owing to voluminous interval-bounds data sets. Hence, uses of variables in interval-bounds form (simple or fuzzy) for probability estimation definitely lead to computationally poor algorithm. To overcome the rigorous computational exercise; two indices namely arithmetic mean index (AMI) and geometric mean index (GMI) are proposed, these indices yield scalar outcomes. The indices estimation is based on arithmetic mean (AM) and geometric mean (GM) operations performed on lower support fuzzy-intervals (SL_i) and upper support fuzzy-intervals (SU_i). In this work, five different support intervals corresponds to five distinct α-cuts value, i. e., $\alpha_i = \{0.2, 0.4, 0.6, 0.8, 1\}$ are considered. Variables used during the AMI and GMI estimations are defined as:
Where,
i = class index represents finite (say M) distinct classes associated with support intervals, i. e., i = 1...M.
M = total number of sub-intervals-spans that constitute a given interval span.
$\alpha_i = i^{th}$ α-cut factor inferred from fuzzy triangular distribution characteristics.
SL_i = lower value of support interval-span maintaining at least α_i factor.
SU_i = upper value of support interval-span maintaining at least α_i factor.

Using these variables, the intermediate parameters namely, lower-support and upper-support aggregate, of AMI (LSA-AMI & USA-AMI) and GMI (LSA-GMI & USA-GMI), are estimated as:

$$LSA - AMI = \frac{1}{M}\sum_{i=1}^{M}\alpha_i SL_i \tag{5.10}$$

$$USA - AMI = \frac{1}{M}\sum_{i=1}^{M}\alpha_i SU_i \tag{5.11}$$

$$LSA - GMI = (\prod_{i=1}^{M}\alpha_i SL_i)^{\frac{1}{M}} \tag{5.12}$$

$$USA - GMI = (\prod_{i=1}^{M}\alpha_i SU_i)^{\frac{1}{M}} \tag{5.13}$$

Subsequently, equations (5.10 & 5.11) and equations (5.12 & 5.13) are used to estimate aggregate AMI (AMIagg.) and aggregate GMI (GMIagg.) respectively and are given as:

$$AMI_{agg.} = \frac{1}{2}[LSA - AMI + USA - AMI] \tag{5.14}$$

$$GMI_{agg.} = \sqrt{(LSA - GMI) * (USA - GMI)} \tag{5.15}$$

The chosen lower and upper bound values of simple intervals of control parameters $\Delta\lambda$, $\Delta\tau$, and Δa are given in Table-5.1. Considering five sub-intervals-spans, i. e., M = 5 with associated α-cuts factors in equations (5.10-5.13), intermediate parameters $AMI_{agg.}$ and $GMI_{agg.}$ are estimated and listed in Table-5.1.

Arithmetic mean (AM) and geometric mean (GM) operations are performed on the chosen simple intervals of $\Delta\lambda$, $\Delta\tau$, and Δa as well. These statistical measures are used to estimate two different deviation indices namely arithmetic mean deviation Index (AMDI) and geometric mean deviation index (GMDI), which are given as:

$$AMDI = AM - AMI_{agg.} \tag{5.16}$$

$$GMDI = GM - GMI_{agg.} \tag{5.17}$$

On estimating the AMDI and GMDI, a heuristic is proposed as; higher the value of a deviation index; it facilitates relatively improved potential of uncertainty variations comprehensiveness. Rationality of it is reflected in Table-5.1; it is observed that AMDI is greater than GMDI. Therefore, in subsequent analysis and during network simulation exercise, only $AMI_{agg.}$ based inferences are considered.

Invariably, spatial and/or temporal parameters play vital role in query generation process. On relative basis depending upon parameters mutual significance; three different cases could be stimulated, where query generation mechanism trends as: (i) govern by spatial parameter largely (ii) having relatively high dependency on temporal parameter and (iii) having equal emphasis on spatial and temporal parameters. However, for illustration purposes, third case-study is examined, wherein spatial and temporal parameters maintain equal priority.

Using parametric values of LSA-AMI and USA-AMI corresponding to the smaller and the larger intervals of $\Delta\lambda$, $\Delta\tau$, and Δa as listed in Table-5.1; a set of expressions is formulated that bid lower and upper bound scalar value of β as:

$$\beta_l = \delta_s(\Delta\lambda_l)(\Delta a_l) + \delta_t(\Delta\lambda_l)(\Delta\tau_l) \tag{5.18}$$

$$\beta_u = \delta_s(\Delta\lambda_u)(\Delta a_u) + \delta_t(\Delta\lambda_u)(\Delta\tau_u) \tag{5.19}$$

In the above expressions; $\Delta\lambda_l$ & $\Delta\lambda_u$, $\Delta\tau_l$ & $\Delta\tau_u$, Δa_l & Δ_u are LSA-AMI and USA-AMI of $\Delta\lambda$, $\Delta\tau$, and Δa respectively. Depending upon the relative importance between the spatial and temporal aspects and the situation in context, weight factors δ_s and δ_t are appropriately chosen. Using relevant parameters as listed in Table-5.1 in equations (5.18 and 5.19); resultant lower and upper bound values of β i.e., β_l and β_u are given in Table-5.2. These values are estimated corresponding to three different scenarios as: (i) Only spatial parameter taken into considerations ($\delta_s = 1$ and $\delta_t = 0$) (ii) Temporal parameter dominates over the spatial parameter ($\delta_s = 0$ and $\delta_t = 1$) and (iii) Spatial and Temporal parameters encompass equal emphasis (with $\delta_s = \delta_t = 0.5$). These scalar values of β are subsequently used in equation (5.2) to estimate the probability of query generation, which inherent spatio-temporal dynamics. Probabilities estimation based on equation (5.2) results in PMF characteristics which are shown in Fig. (5.2). These PMFs are corresponding to spatial aspect only Fig. 5.2(a & b), exclusively temporal aspect Fig. 5.2(c & d), and spatio-temporal aspects Fig. 5.2(e & f).

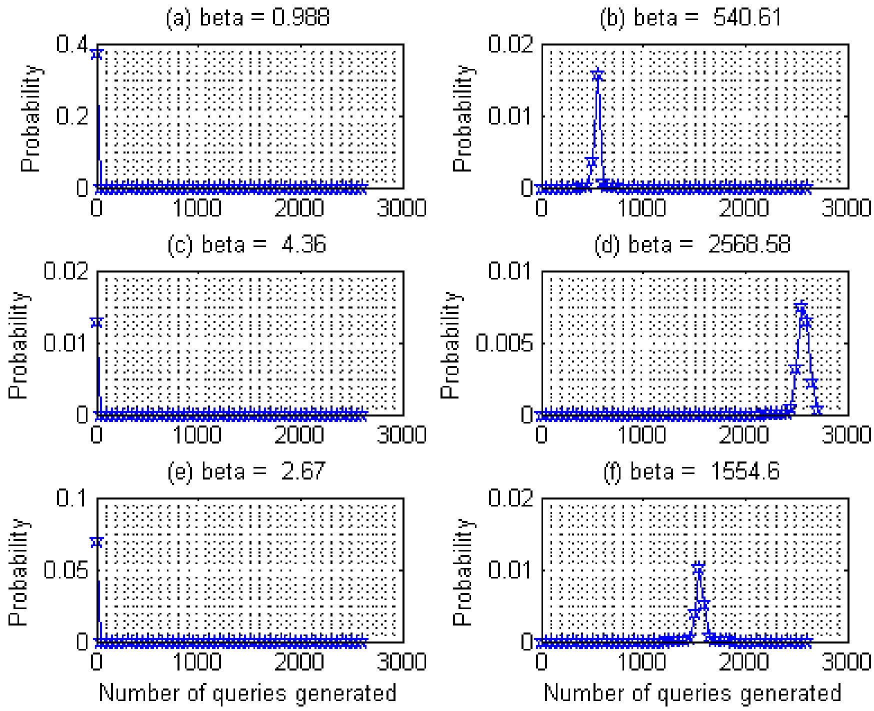


Figure 5.2: PMFs for various β inferred from spatial aspect: (a & b), temporal aspect: (c & d), and spatio-temporal aspects: (e & f)

5.5 Service-area and Sensor Network Specifications

To validate the mathematical framework given in sections 5.3.2 and 5.4 during simulation; an arbitrary service-area of $250 * 250$ square meter comprises of 100 homogeneous sensor nodes is considered. Presuming $6mW$ average power consumption per query; each sensor node is equipped with two AA alkaline batteries that support 1000 hours of continuous operation [(Cheng et al., 2003b)]. Thus, each sensor node has an initial energy budget of 6 Watt hour (Whr). During the assessment of energy performance measures for three different sink attributes driven case studies; CRES threshold mark is arbitrary set as 10% of initial energy. In all the case studies; the network stipulated lifetime is considered as 45 days. Values of other parameters used during simulation exercise are given in Table-5.1 [(Cheng et al., 2003a), Prabaldutta2005 and (Huang et al., 2007)].

*Crisp values corresponding to slow arrival and fast arrival rate taken from [(Dutta et al., 2005)] are 120 and 1440 queries per month. Around these values; symmetrical intervals are [80, 160] and [1200, 1700] for slow and fast rate respectively, later these intervals are transformed for hourly scale.

@Crisp values corresponding to low density and high density of queries considered are 5 and 78 per square km. as given in [(Huang et al., 2007)]. Typical intervals considered are

Table 5.1: Simple and fuzzy-intervals of control parameters $\Delta\lambda$ [(Cheng et al., 2003a)], $\Delta\tau$[(Dutta et al., 2005)]* and Δa[(Huang et al., 2007)]@

Control parameters	Scalar ($\Delta\lambda$)		Rate of generation of query($\Delta\tau$)		Spatial Parameters(Δa	
Intervals/Intermediate	Smaller Interval	Larger Interval	Smaller Interval/hour	Larger Interval/hour	Smaller Interval/cluster	Larger Interval/cluster
	[2,8]	[10,40]				
Estimated intervals as per rule of interval arithmetic (using equation 5.5)	..	..	[0.64, 6]	[216.4, 263.2]	[0.26, 1.31]	[39.34, 59]
LSA-AMI	2.6	12.6	1.56	140.1	0.38	27.92
USA-AMI	3.48	17.4	2.42	147.62	0.56	31.07
Aggregate AMI	**3.04**	**15**	**1.99**	**143.86**	**0.47**	**29.49**
LSA-GMI	1.92	9.64	1.09	120.01	0.282	23.52
USA-GMI	3.19	15.99	2.25	129.77	0.52	27.63
Aggregate GMI	**2.55**	**12.81**	**1.67**	**124.89**	**0.4**	**25.57**

[2, 10] and [60, 90] respectively, which are subsequently scaled down for a cluster area of 32*32 m^2.

Table 5.2: Scalar β inferred from LSA-AMI and USA-AMI for the Smaller and Larger intervals

	Smaller Intervals		Larger Intervals	
Intervals	Lower bound	Upper bound	Lower bound	Upper bound
Spatial	0.988	1.94	351.79	540.61
Temporal	4.36	8.42	1765.38	2568.58
Spatio-temporal	**2.67**	5.18	1058.58	**1554.6**

5.6 Clustering mechanism

The whole service-area is resolved into four uniform quadrants, which are further decomposed into several non-uniform clusters. Each quadrant is monitored using 25 sensor nodes. At the time of random deployment, minimum distance between any two nodes is constrained as 10 meters, whereas, the coverage radius of each sensor node is considered as 30 meters. Each of these clusters possesses few sensor nodes that sense/measure physical parameters in their vicinity and communicate the measured attributes to sink node through multi-hop wireless links. For cluster formation; four clustering schemes namely SKM, SFCM, DKM and DFCM are implemented. Post clustering; key energy performance measures are estimated at cluster level as well as network level. Initially, the clusters are formed randomly and subsequently formation is based on RES of associated sensor nodes. At sporadic sampling instants for each cluster; node having maximum RES is elected as cluster head (CH). Further, the energy-centroid (EC) positioning for each cluster is estimated, which is primarily governed by Cartesian coordinates of sensor nodes and their associated RES [(Kumar and Chaturvedi, 2015)]. Having known ECs locations at the network level, the optimum locations of sink(s) are estimated that avoids occurrence of *hot-spot/islanding* till the network attains the stipulated lifetime.

5.7 Query Characterization

Queries nature considered in this chapter is of pull and unstructured WSNs (prior to sensing) [(Rachuri et al., 2008)]. In that sink nodes send simple and one shot queries to detect presence and identification of target type. Queries correspond to inquiry about the presence of civilians, soldiers, and vehicles. For that matter, multi-modal sensors (in this work three different type of sensors [(Dutta et al., 2005)] are used. In a prototype program, three different fields are created in query packets, which are features extracted from infrared, magnetic and acoustic sensors [(Dutta et al., 2005)]. The movement of these targets (civilians/soldiers/vehicles) is considered to maintain an average speed of 1.5 m/sec., $3m/sec.$ and 15 $m/sec.$ Owing to different speed these targets remain in coverage radius of a sensor node for different short epochs. During 20 different simulation runs; on average basis, each of these three different types of targets remain in coverage field of one particular sensor node for 2.05 sec. short duration epoch.

Each sensor consumes on average 0.4 mJ for transmission and reception [(Crossbow Technology, 2007)]. Thus transceiver consumes 0.8 mJ. For three different types of onboard sensor using average contact duration of 2.05 seconds, approximately 4.8 mW power is consumed by each mote (which supports transceiver module of three different types of sensors). Usually, the transceiver module consumes significant proportion of power compared to the other subsidiary subsystems. For querying and sensing the presence of these targets an aggregate power budget of 1.2 mW is considered [(Rachuri et al., 2008)]. For simplicity costs resulting from the other power consuming components are ignored.

The network's energy metrics in terms of performance measures are estimated and illustrated for a network scenario having equal emphasis on the spatio-temporal aspects. In this scenario, the inference derived from the spatio-temporal parameters leads to the lower and upper bound β values of [2.67, 5.18] and [1058.58, 1554.6] for the smaller and larger intervals respectively and are listed in Table-5.2. For all the three sink driven network scenarios; owing to highly correlated trends for performance measures; these measures are estimated and analyzed for two extreme values of β corresponding to extracted lower and upper values from the smaller and larger intervals respectively and is referred as β-set in the rest of the chapter, i.e., β-set = [2.67, 1554.6].

5.8 Simulation Setup and Test Results

In order to investigate the performance of the proposed approach that relies upon inclusion of spatio-temporal uncertainties appropriateness, clustering schemes deployed and sink(s) driven attributes simulations are performed using MATLAB programming language. These simulation instances are performed on MATLAB version 7.9.0.529 (with clock speed 2.27 GHz and 32 bit operating system). The simulator is a discrete event driven simulator that comprises surveillance of three different targets viz. civilians, soldiers, and vehicles. In each simulation instance, 20 different random runs were considered following which average value of the time epochs over which these targets remain in coverage field of a sensor node is estimated. For three different sink attributes driven network scenarios, estimate and analysis of performance measures are presented next.

5.8.1 Network Surveillance by a Single-Stationary sink

In this scenario; a sink is located at the center of the service-area. End users exchange queries and responses (replies) with field deployed sensor nodes using multiple-hop wireless communication links. In a process of establishing these links; sensor nodes deplete their energy reserve that in turn reduces the network ARES.

For the spatio-temporal aspects driven β-set; the network energy reserve status is determined in terms of ARES and is shown in Fig. 5.3 (a & b). Each of these figures comprises of four ARES characteristics corresponding to four different clustering schemes. It indicates that the ARES maintains relatively much better energy status in Fig. 5.3(b) (for β = 1554.6) in comparison with Fig. 5.3(a) (for β = 2.67). For β = 1554.6, the ample margin between CRES threshold mark and the resultant ARES value on using DFCM scheme at stipulated lifetime instant i.e., at 45^{th} day of surveillance operation further confirms the surveillance task beyond the lifetime mark. In complementary terms, the STD attained on using four clustering schemes are shown as the leftmost group of histograms in Fig. 5.6 and Fig. 5.7 for β = 2.67 and 1554.6 respectively. It exhibits dependency of STD on chosen clustering scheme, and values of β considered which in turn governed by spatio-temporal parameters. Fig. 5.6 depicts that none of the clustering schemes are able to meet operational service requirement for a stipulated lifetime span of 45 days (1080 hours). Superiority of DKM and DFCM over SKM and SFCM is shown in Fig. 5.7; as the resultant network configurations attain the stipulated lifetime.

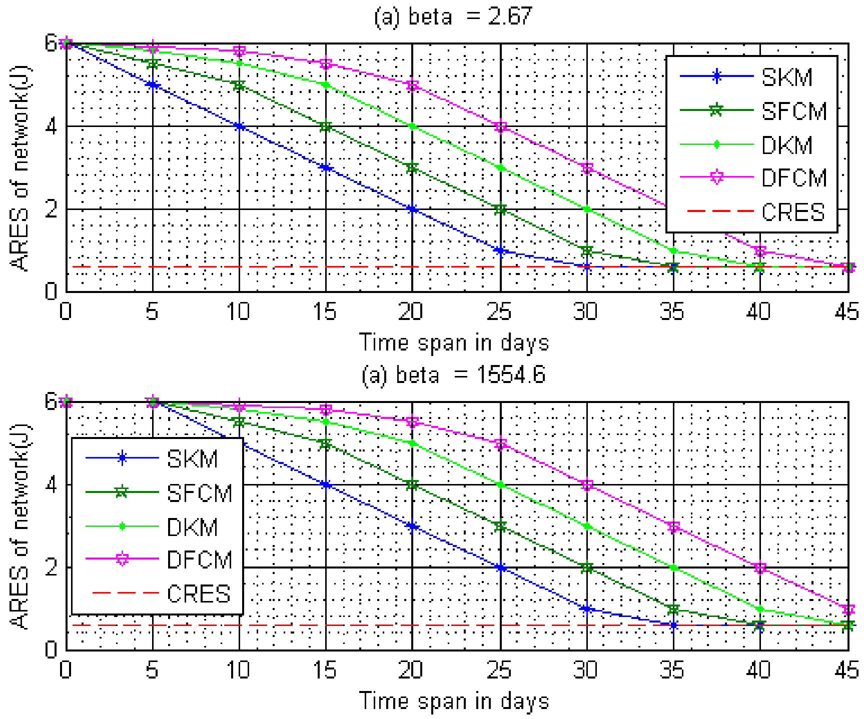


Figure 5.3: ARES of network with single-stationary sink (a) for low β (b) for high β.

5.8.2 Network Surveillance by a single-portable sink

In this case; initially the sink is located at the center of the service-area. Subsequently, depending upon query spatial distribution and its inter-arrival-time-rate; the RES of sensor nodes depletes unevenly. Sink relocation mechanism is solely governed by EC location so as to minimize the energy consumption associated with bidirectional single/multiple-hop communication links. Corresponding to the uses of four clustering schemes; the network ARES characteristics are shown in Fig. 5.4(a & b) for $\beta = 2.67$ and 1554.6, respectively. Alike previous case study, the network ARES have dependency upon types of clustering schemes deployed and β values. The difference margin between the stipulated CRES mark and the instantaneous network ARES yields network overall energy profile. More the margin better is the network energy reserve. In complementary terms, the period of reliable and sustainable service is estimated as STD measure. It is shown as histograms grouped at the middle of Fig. 5.6 and Fig. 5.7 for $\beta = 2.67$ and 1554.6 respectively. Subset of histograms in Fig. 5.6 (for $\beta = 2.67$) indicates that only DKM and DFCM are able to comply with service norms on attaining CRES mark at or beyond the lifetime mark of 1080 hours. Whereas, for $\beta = 1554.6$ (in Fig. 5.7); except for SKM, uses of other three clustering schemes lead to network topologies that ensure sustainable service norms well beyond the stipulated lifetime. Among the chosen clustering schemes, DFCM outperforms the others and assures the surveillance task over an additional serving time of 11%.

5.8.3 Network Surveillance by four-stationary sinks

To enable energy centric network surveillance and anticipating lifetime longevity, each quadrant of the service-area is monitored by a single-stationary sink. In each quadrant; the sensor nodes participate to form clusters based on SKM, SFCM, DKM, and DFCM. Implementing these clustering schemes result into the network topologies for which ARES characteristics are shown in Fig. 5.5 (a & b) for $\beta = 2.67$ and 1554.6 respectively. Fig. 5.5(a) shows that except for SKM, all other clustering schemes result into network configurations that assure reliable surveillance task over a stipulated lifetime span of 45 days. Further, the uses of DKM and DFCM lead to outcomes maintaining ample energy buffer with respect to chosen CRES mark at lifetime instant. Interestingly, for $\beta = 1554.6$; irrespective of clustering schemes deployed, the resultant ARES characteristics maintain a healthy energy profile at lifetime instant and is shown in Fig. 5.5(b).

As a complementary measure, the network energy profile is estimated in terms of STD. Corresponding to the uses of four clustering schemes; it is shown in the rightmost group of histograms in Fig. 5.6 and Fig. 5.7 for $\beta = 2.67$ and 1554.6 respectively. In Fig. 5.6, for four-sinks based surveillance, the STD (in hours) is shown in an array appended at the bottom. The STD measure (in hours) in an array as $\{965, 1085, 1104, 1179\}$ shows that DFCM emerges as the most energy efficient clustering scheme as its implementation offers 100 hours extendable operation beyond the stipulated lifetime mark of 1080 hours. The STD measure in another array $\{1080, 1180, 1150, 1250\}$ shown in Fig. 5.7 validates that irrespective of the clustering schemes used, the network is able to deliver surveillance task for the lifetime span. Here too, the DFCM outperforms the other clustering schemes, as comparison with the stipulated lifetime span of 1080 hours with STD obtained from DFCM uses infer a much better network energy reserve status. Here, the uses of DFCM ensure 15% more service span over and above the stipulated lifetime span of 1080 hours.

5.8.4 Performance analysis of spatio-temporal Parameters Regulated β-set

A comprehensive analysis of three different sink-attributes driven surveillance scenarios substantiates importance of β-set, as it leads to spatial-temporal parameters inclusive probability estimation of query generation. The performance measures ARES and STD depend on estimated probabilities; which in turn draw inference from spatio-temporal

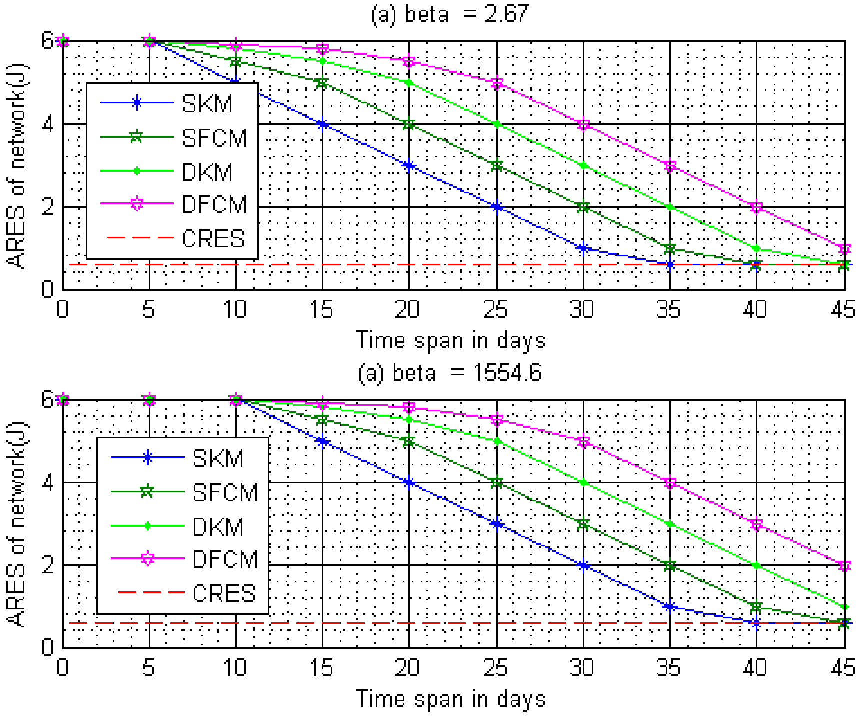


Figure 5.4: ARES of network with single-portable sink (a) for low β (b) for high β.

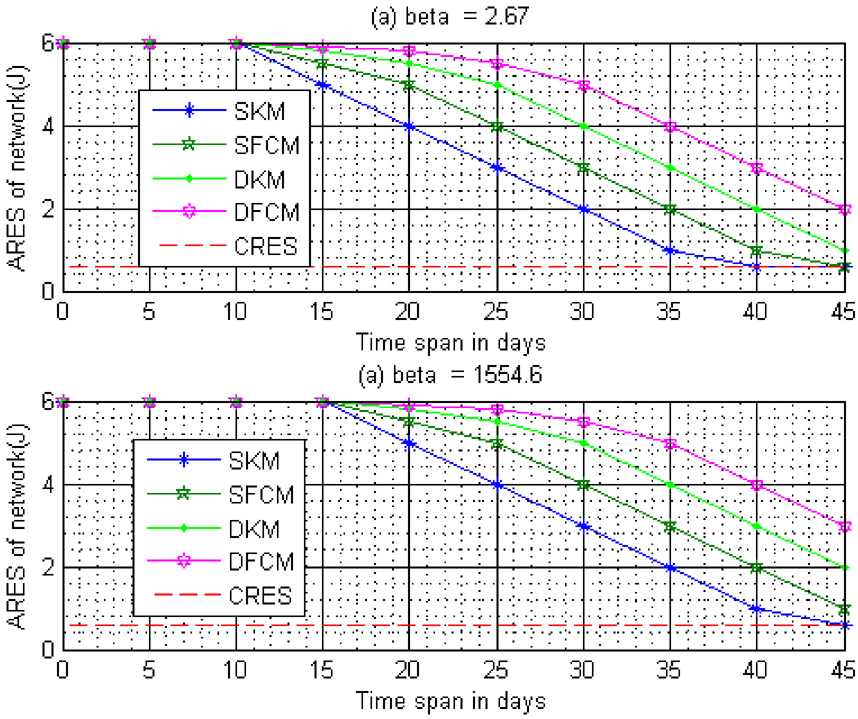


Figure 5.5: ARES of network with four-stationary sinks (a) for low β (b) for high β

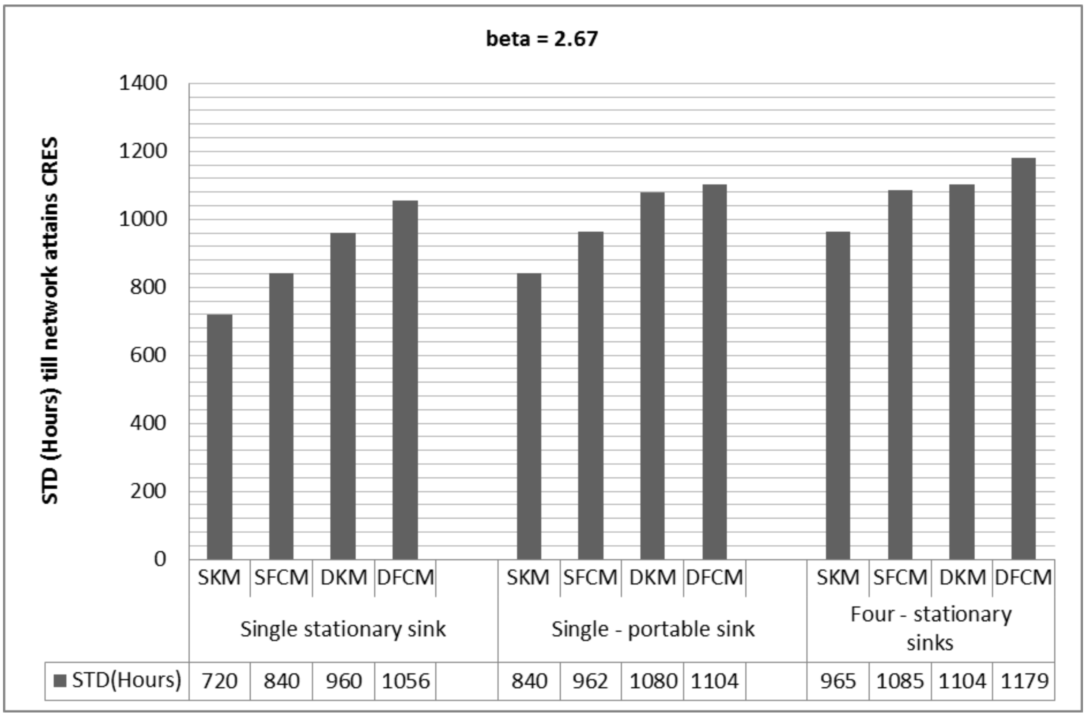


Figure 5.6: STD of network for different clustering schemes and sink attributes for low β

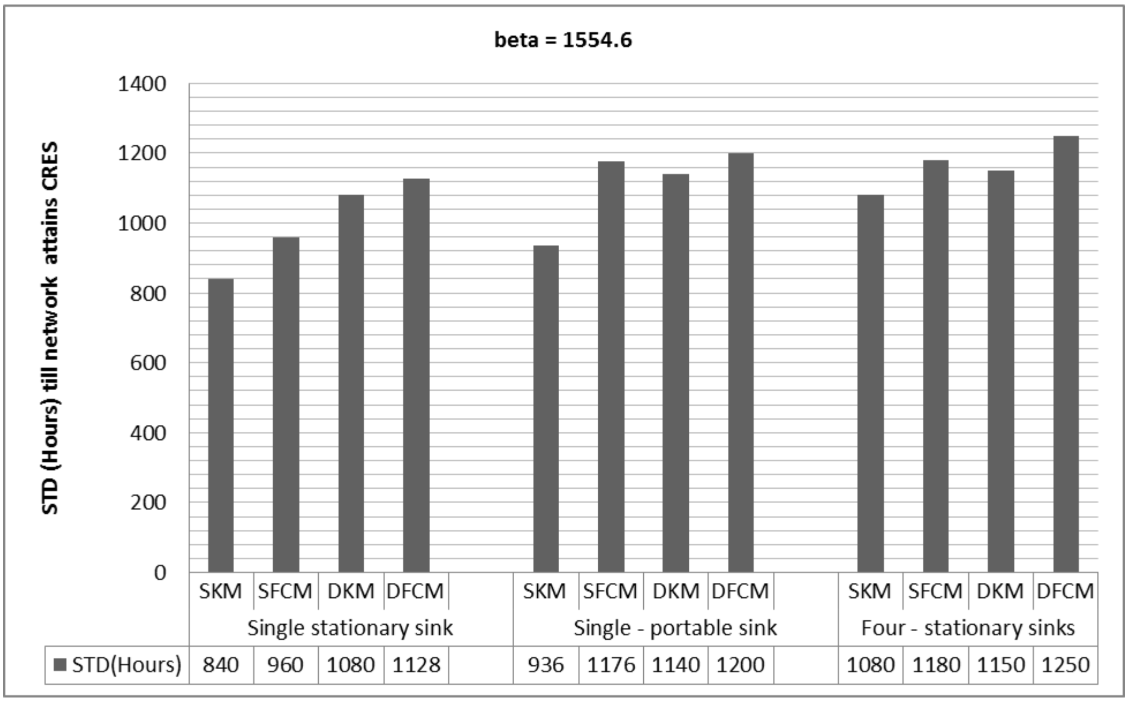


Figure 5.7: STD of network for different clustering schemes and sink attributes for high β

parameters driven β-set. To appreciate importance of β-set elements value; pairing of the poorest and the best contenders from the sink-attributes driven model-space and a set of clustering schemes is considered. It leads to combining SKM with single-stationary sink and DFCM with four-stationary sinks.

On using β-set; the absolute difference between the STD is $\{1128 - 840 = 288\ hours\}$, $\{1056 - 720 = 336\ hours\}$, $\{1179 - 965 = 214\ hours\}$ and $\{1250 - 1080 = 170\ hours\}$. These difference estimates relative to stipulated lifetime span of 1080 hours maintain an extensively varying ratio of 15.74% to 31.12%. The STD outcome of these extreme pairs is sometime unable to meet the nominal service span of 1080 hours, whereas, at other times it supplements the predetermined span of 1080 hours. Thus, the importance of β-values could be comprehended, as correctly chosen spatio-temporal parameters ultimately yield an estimate about stipulated lifetime of the WSNs. Subsequently, the approximate value of stipulated lifetime becomes one of the important inputs during network design phase; which strategically guide network designer about type and specifications of network components on which network infrastructure rests. The analysis of performance measures ARES and STD reported in this chapter are based on query-density and query inter-arrival-time-rate, which are predominantly application specific.

In this chapter, network lifetime estimate owes significantly to issues such as quantum of spatio-temporal uncertainties, uses of clustering schemes and sink attributes. In contrast to it [(Wang et al., 2014)] have reported network lifetime estimate in that variation in lifetime is observed with respect to parameters namely, number of sensor nodes, initial battery reserve, service-area size, and transmission range. Further,[(Perillo et al., 2005)] have also studies lifetime variations while treating uses of different transmission power, network radius and number of cluster head deployed for varying network radius as a set of regulated parameters. In both of these work; the parameters regulation demonstrates a sort of control over physical/hardware specifications. Whereas, the work reported in this chapter has purely algorithmic emphasis and highlights the importance of appropriate means to address the inevitable uncertainties and its impetus on network lifetime.

5.9 Conclusions

In this chapter, three different issues in context to energy efficient operation of WSN are investigated. These three issues include: (i) determining most energy savvy clustering technique, (ii) empowering the query generation model with spatio-temporal parameters, while giving due considerations to associated uncertainties so as to predict query density and its inter-arrival-time-rate more sensibly that could equip the network planning agencies to frame decision policies which ensure proper utilization of network resources, and

(iii) analysis about sink node multiplicity and impact of its location aspect being in stationary or portable state. Apparently these issues seem to be quite indifferent; however, they operate in synergistic manner towards achieving the precious objective of minimizing the energy consumption along with specific details about associated spatio-temporal resolution. Another noteworthy feature of the work is two-tier of spatio-temporal uncertainties incorporation, initially using interval bound and later invoking *fuzzification* of these interval bounds. The proposed algorithm can be modified on tailor-made basis to include application specific requirements. Further, uses of four different clustering schemes in three distinct sink attributes driven scenarios establish following key observations; (a) DFCM emerges as the most promising clustering scheme, and (b) network surveillance using four-stationary sinks renders the healthy ARES and prolongs STD measures.

Chapter 6

FUZZY INTERVALS MODELS AND SPATIAL FUSION FOR WIRELESS SENSOR NETWORKS

6.1 Introduction

Rapid advancement and growth in micro-electro-mechanical-systems (MEMS) technology, multimodality and low cost development in sensors combined with feasibility of distributed computing has changed the impetus and diversified applications of wireless sensor networks (WSNs) in a variety of networking and sensing paradigms. WSNs infrastructure comprises of tiny sensor nodes which are most often stationary and sinks that could be stationary or mobile. These sinks are interface between field deployed sensor nodes and the external world. Depending upon the nature of application and spatial resolution; the typical number of sensor nodes deployed in a service-area varies in an order of from few tens to few thousands. These sensor nodes are deployed either inside the physical environment or in close proximity to it. In entire network service operation varieties of tasks such as sensing, preprocessing, computation and communication are performed. All these functionalities require certain minimum energy to carry out intended task. Owing to the difficulty associated with battery charging or replenishing it; battery has been a vital resource that must be utilized with the utmost care. Thus, invariably in all WSNs based network infrastructure, limited battery supply influences the finite lifetime over which reliable service norms are expected. Maintaining the service norms in wireless sensor networks (WSNs) till it attain a stipulated lifetime is a major issue that impacts the application or utility of such networks.

For query based WSNs, network lifetime depends upon: the amount/quantum of generated query, inter-arrival-time-rate of query occurrences (frequency of query generation), and the geographical distribution of query. [(Akylidiz et al., 2002)] presented a survey that focuses on variety of important operational dynamics of sensor network. The sink(s) may be stationary/portable/mobile and connect the WSNs to the outside world using Internet gateway to enable users an easy access to the reported data [(Karaki and Kamal, 2004)]. A brief summary about the reported research methodologies addressing issues related with spatio-temporal aspects, mobility of sink nodes and their locations aspects are presented next.

6.2 Literature Survey

[(Yoon and Shahabi, 2007)] presented clustered Aggregation (CAG) algorithm; genesis for its formulation is based on assumption that the sensed data in WSNs exhibit the spatial and temporal correlation of physical attributes existing in the sensing field. [(Takruri et al., 2011)] proposed a spatio-temporal model that considers drift associated with measurement process. To predict the future measurements, support vector regression algorithm is used, whereas, to detect and correct associated drift and random error Kalman filter is used to prolong the network lifetime. [(Jain et al., 2005)] observed that the WSNs can significantly improve the quality of spatio-temporal data monitoring. The authors described a communication architecture that supports distributed query processing to evaluate spatio-temporal queries for the network considered. [(Mousavi et al., 2013)] proposed a spatio-temporal event detection algorithm. The algorithm provides probabilistic graphical models (PGMs) of WSNs; it incorporates the Markov chains in temporal dependency and Markov random fields theory in the spatial dependency of sensors in a distributed fashion. [(Cheng et al., 2003a),(Cheng et al., 2003b)] reported probabilistic query evaluation based upon uncertain data; classification of queries is done using attributes namely, value-based non-aggregate class, entity-based non-aggregate class, entity based aggregate class and value-based aggregate class. Measurements obtained from sensors were presented in interval from. [(Dutta et al., 2005)] proposed an experimental work to detect rare, random, and ephemeral events based on observations drawn from WSN.

[(Deshpande et al., 2005)] presented a model based approximate querying in WSNs. The authors reported a strategy about selection of best sensor reading to acquire live data with commendable degree of confidence. However, in this thesis; the inferences are drawn from probabilistic model that have primary focus on uncertainty associated with spatio-temporal parameters, thus it maintains substantiate differences. [(Wang et al., 2004)] presented a mathematical model, wherein the localization uncertainty is estimated and compared using Bayesian bound (BB) with Cramer-Rao bound (CRB). Here, mainly the localization uncertainty is attributed to the network topology.

[(Wang et al., 2014)] proposed a lifetime enhancement method for WSNs; the method operates on sink relocation aspect as it prolongs the occurrence of hot spot. The work address time complexity analysis and numerical analysis for sink relocation mechanism.

[(Gandham et al., 2003)] reported a study that makes use of mobile base station to prolong the network lifetime. The authors decomposed the lifetime into uniform short-duration epochs. To determine new locations for the base-stations and for energy efficient routing an integer programming approach and flow-based routing protocols are used respectively.

[(Shah et al., 2003)] proposed three-tier architecture for sparse sensor networks. The proposed scheme exploits the presence of mobile entities deployed in the environment. These mobile entities on coming in close range with sensors, gathers data, buffer it and finally convey it to wired access points. Short range communication lead to substantial amount of power saving. The entire framework is based on two-dimensional random walk for mobility and involves key variables namely number of entities, sensors and access points.

[(Younis et al., 2003)] presented a base-station repositioning scheme to enhance network performance. Repositioning is exercised based on a heuristic search that makes use of the current network topology and traffic pattern. The reported method operates under operational constraint that during the relocation of base-station/gateway node no data should be lost. Further, the gateway performs a tradeoff analysis between the reduced energy consumption that results from its positioning at a new location and the overhead burden that the relocation imposes on sensors. It is presumed that the gateway undergoes a limited mobility.

[(Gatzianas and Georgiadis, 2008)] presented a distributed approach to maximize the network lifetime of WSNs having a mobile sink. The problem of lifetime maximization is posed as a linear program and subsequently on reducing it to a simpler form solved using dual decomposition. In problem formulation; the sink sojourn times and the routing flow vector for each sink location are treated as unknowns.

[(Luo and Hubaux, 2010)] reported a hybrid approach that principally exploits routing strategy and sink mobility to enhance the lifetime of the WSNs. The proposed approach formally establishes NP-hardness of the problem and investigates induced subproblems. Initially a framework is developed to solve subproblem involving a single sink using an efficient primal-dual algorithm and later it was generalized to deal scenarios involving multiple sinks. Obtained results put light on benefit of involving sink mobility and also imply the desirable moving traces of a sink.

Strictly considering the optimal locations of sinks is an NP-hard problem, thus to overcome complexity associated with this enormous complex task a heuristic scheme is proposed which offers quasi-optimal solutions that in turn facilitate estimation of possible locations to which sinks could be relocated for the sake of optimal consumption of energy.

In principal, this chapter investigates four diversified issues in an unified manner and these are (i) varying energy efficiency incurred with the uses of four different clustering schemes, (ii) incorporating query generation dependence on spatio-temporal aspects to envisage network energy performance measures more precisely, and subsequently using these outcomes to devise a strategy so as to utilize network resources more appropriately, (iii) modeling the problem in high-dimension space by notion of spatial (quadrants) fusion concept, which limits the extent of clusters formation across the quadrants, and (iv) exploiting sink multiplicity and location issues so as to maintain energy balance across the network, thereby avoiding possibility of encountering hot spots or islands prematurely. The first issue is addressed by deploying four clustering schemes namely static k-means (SKM), static fuzzy c-means (SFCM), dynamic k-means (DKM) and dynamic fuzzy c-means (DFCM). The second issue needs a mechanism that owes query generation process dependence upon spatio-temporal parameters; it is achieved by incorporating the spatio-temporal aspects in the Poisson PMF expression of probability estimation. For third issue; a common observation from the optimization realm generalizes that the solutions inferred from high-dimension space in any circumstances are no worse than that derived from low-dimension subspace. This rational is implemented using the notion of quadrants fusion that effectively gets activated subjected to satisfying the energy metrics driven inequality norm. Towards the end to address the forth issue, we need balancing mechanisms that ensure avoidance of sensor nodes *islanding* which stems from non-uniform energy gradient thereby ensures reliable coverage. To achieve remedial solution that counter acts the non-uniform energy gradient, the network is monitored by single stationary sink (SSS), single portable sink (SPS), four stationary sinks (FSS) and four portable sinks (FPS). In portable-sink(s) based network scenarios; sink(s) relocation is done based on residual energy status (RES) estimate of sensor nodes.

In all the network scenarios presented in this chapter; network overall energy performance is evaluated using variety of energy metrics. These energy measures include; time dependent residual energy status (RES) of individual sensor nodes, average residual energy status (ARES) estimate of all participating sensor nodes at cluster/quadrant/network level, and the network service-time-duration (STD) estimate by which the network attains the predetermined critical residual energy status (CRES) threshold mark. These energy measures are estimated periodically and at stipulated lifetime instant of the network. Other auxiliary results include ARES variations with fusion count index (FCI), FCI variation and locus of sinks observed over the network lifetime on pairing sink attributes driven network scenarios with and without quadrants fusion mechanism. Further, details of the participating quadrants in fusion process for few arbitrary chosen FCI under the SPS and FPS cases are also presented.

The effectiveness of proposed scheme is validated through simulation. The proposed algorithm is validated using MATLAB programming. The random runs of algorithm are performed using MATLAB version 7.9.0.529 on a system having clock speed 2.27 GHz and 32 bit operating system.

6.3 Attributes and Probabilistic Framework for Query Generation Process

Uncertainties with query generation process are inevitable and it dependence on spatio-temporal parameters that influence the query generation considerably. Observing network operations during any short epochs, queries are not ubiquitous or in other words within entire network, spatial distribution of queries is not concurrent. Accuracy of query generation model hinges on inclusiveness of the precision associated with spatio-temporal aspects. Thereby, incorporation of spatio-temporal aspects with finest possible granularity could result in a well approximate model of query generation process.

In plethora of network paradigms; to envisage the random arrival of events, the Poisson's probability mass function (PMF) manifests itself as one of the appropriate model. In wide ranging WSNs scenarios; two important aspects (spatial and/or temporal) associated with query generation process are random and thereby approximated using an appropriate PMF. So far, in WSNs regime; the query generation patterns have been modeled as uniform PMF and Poisson's PMF with associated parameter having a scalar crisp(singleton) value. Owing to the inherent randomness with query inter-arrival time-rate (temporal aspects) and query density (spatial aspects); the query dynamic distribution as the uniform PMF may be a good approximation on aggregate basis towards the end of predetermined network lifetime. However, it lacks in following the spatial and temporal dynamics involved with query generation process. Thus, we presume that the query generation mechanism in any finite observation interval (FOI)) is modeled as per Poissons PMF. Thereby on treating the number of queries generated as a random variable, Q, the expression for PMF is given by

$$P(Q = k) = \frac{\lambda^k}{k!} e^{-\lambda} \tag{6.1}$$

Where, k = 0,1,2,3....

Where k is an independent variable indicates number of queries and λ is a control parameter. Dependency of on spatio-temporal parameters is treated next.

6.3.1 Spatio-temporal parameters Integrated Poisson PMF model

In equation (6.1); contrary to usual way of choosing control parameter (λ) as an arbitrary scalar entity, a mathematical framework is proposed that establishes functional dependency of λ on the spatio-temporal parameters associated with the query generation process as:

$$P(Q=k)=\frac{\beta^k}{k!}e^{-\beta} \tag{6.2}$$

Where, k = 0,1,2,3....

Here, β is a parameter inferred from spatio-temporal parameters. Depending upon crisp or intervals- bound values of spatio-temporal parameters; the parameter β in equation (6.2) is defined as:

$$\beta=\delta_s\lambda a+\delta_t\lambda\tau \tag{6.3}$$

$$\beta=\delta_s\Delta\lambda\Delta a+\delta_t\Delta\lambda\Delta\tau \tag{6.4}$$

Variables used in equation (6.3) and (6.4) are defined as: δ_s and δ_t are weight/scaling factor associated with spatial and temporal parameters respectively, $\Delta\lambda$ is the chosen interval for λ parameter, a and τ is the chosen crisp value associated with query density and query inter-arrival-time-rate respectively. Intervals Δa and $\Delta\tau$ comprises the lower and upper bound interval-values of query density, and query inter-arrival-time-rate respectively. In equation (6.4) on performing the product; the final outcomes are the intervals having lower and upper bound values of β (as per the product rule of interval arithmetic [(Rokne, 2001)]) as:

$$[a_1\ a_2]*[b_1\ b_2]=minimum\ \&\ maximum\ of\ [a_1b_1,a_1b_2,a_2b_1,a_2b_2] \tag{6.5}$$

Weight factors δ_s and δ_t are governed by relative significance of spatial and/or temporal parameters values. Usage of bound interval-values around crisp/singleton values though addresses the vagueness/uncertainty associated with these scalar parameters.

However, it has the associated drawbacks as well, as the apprehension for intervals formation is based on hypothesis that during whole interval-span the degree of uncertainty presumes uniform distribution characteristic. Thus, to cope with imprecision involved with these ambiguous parametric values; uses of fuzzy-triangular distribution characteristic is presented next.

6.4 General Fuzzy Triangular Characteristics

The typical fuzzy-triangular distribution characteristic is mathematically expressed in equation (6.6) and is shown in Fig. 6.1.

$$A(x : a, m, b) = max\left\{min\left[\frac{(x-a)}{(m-a)}, \frac{(b-x)}{(b-m)}\right], 0\right\} \tag{6.6}$$

In this expression, the parameters a, b locate the *feet* and the parameter m locates the peak of the triangular characteristics as shown in Fig.6.1. To incorporate better spatio-temporal resolution so as to address uncertainty with fine granulation, the triangular characteristic is segmented into finite number of base spans having specific α-cuts (in this chapter; five base spans and correspondingly five α-cuts factors are considered) [(Zimmermann, 2001)].

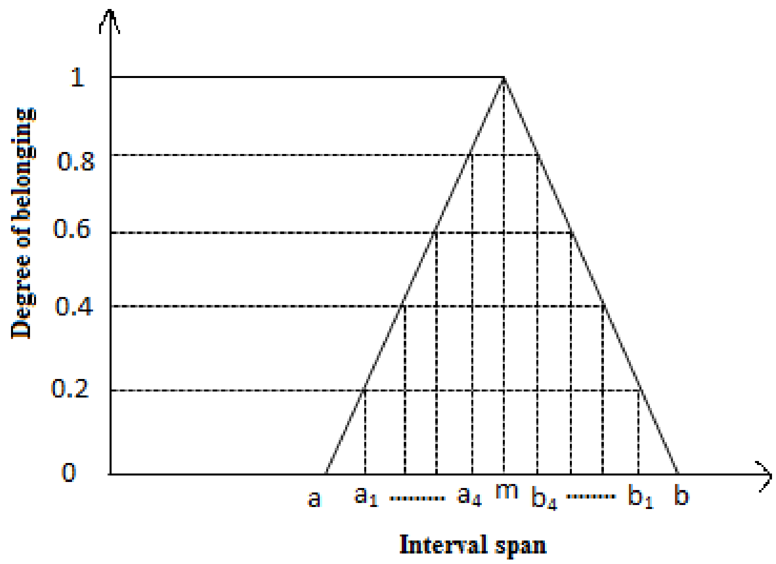


Figure 6.1: A typical fuzzy-triangular distribution characteristic

6.4.1 Fuzzified intervals based spatio-temporal integrated Poisson's PMF

Contrary to presuming that all the scalar values within a given interval bound follow a uniform distribution probability model; it is presumed that intervals bound parametric values of $\Delta\lambda$, Δa and $\Delta\tau$ are appropriately scaled using α-cuts inferred from independent fuzzy-triangular distribution characteristics. Incorporating this rational amends equation (6.2) as:

$$P(Q = k) = \frac{e^{-\beta_{F-I}}}{k!}\beta_{F-I}^{k} \tag{6.7}$$

With β_{F-I} is modeled as:

$$\beta_{F-I} = \delta_s \Delta\lambda_{F-I} \Delta a_{F-I} + \delta_t \Delta\lambda_{F-I} \Delta\tau_{F-I} \tag{6.8}$$

Here, β_{F-I} represents spatio-temporal parameters driven fuzzy-interval bounds and F-I as subscripts with $\Delta\lambda$, Δa and $\Delta\tau$ denote fuzzification of intervals $\Delta\lambda$, Δa and $\Delta\tau$ respectively. To accomplish it, these parameters values spanned by the lower and upper bound intervals are scaled by appropriately chosen α-cuts factors. Let the associated α-cuts factors for intervals $\Delta\lambda$, Δa and $\Delta\tau$ are represented by notations α_i, α_j and α_k respectively. Incorporation of inferences drawn from fuzzy-triangular characteristics about associated α-cuts result in

$$\beta_{F-I} = \delta_s(\alpha_i\lambda_i)(\alpha_j a_j) + \delta_t(\alpha_i\lambda_i)(\alpha_k\tau_k) \tag{6.9}$$

Here, each of these α-cuts factors belongs to independent and uncorrelated closed intervals of zero and unity ($0 \leq \{\alpha_i, \alpha_j \; and \; \alpha_k\} \leq 1$). Whereas, λ_i, a_j and τ_k are arbitrary element values spanned by the chosen lower and upper bound intervals that represent the base-span of triangular distribution characteristics. The base-span of triangular distribution for α_i, α_j and α_k encompasses the ranges $(\lambda_u - \lambda_l)$, $(a_u - a_l)$ and $(\tau_u - \tau_l)$ respectively. The estimated fuzzy-intervals along with plane intervals are given in Table-5.1. The obtained outcomes on using equations (6.7 and 6.9) lead to probability estimation. Owing to intervals form of $\Delta\lambda$, Δa, $\Delta\tau$ and number of α-cuts (α-levels set); the probability estimation task becomes cumbersome. To overcome the painstaking computational task, an arithmetic mean index (AMI) is proposed, which in aggregate terms includes the basic features of considered interval-bounds and subsequent imposition of fuzzification procedure on these intervals and is discussed next.

6.4.2 AMI for Scalar inferences from fuzzy-interval bounds

The computational task for probability estimation using equations (6.2 and 6.7) is simpler on making uses of scalar values. For intervals-bound parameters, the probability estimation task transforms into highly intensive one. Thus, uses of parameters in intervals-bound form (simple or fuzzy) for probability estimation definitely lead to computationally poor algorithm. To overcome the rigorous computational burden; an arithmetic mean index (AMI) is presented that yields scalar outcomes. The index estimation is performed on lower support fuzzy-intervals (SL_i) and upper support fuzzy-intervals (SU_i). In this chapter, five different support intervals corresponding to five distinct α-cuts value, i. e., $\alpha_i = \{0.2, 0.4, 0.6, 0.8, 1\}$ are considered. Variables used in AMI formulation are defined as:

i = class index represents finite distinct classes associated with finite base span support Intervals, i. e., i = 1.....M
M = total number of sub-intervals-spans that constitute a given interval span. $\alpha_i = i^{th}$ α-cut factor inferred from fuzzy triangular distribution characteristics. SL_i = lower value of support interval-span maintaining at least α_i factor. SU_i = upper value of support interval-span maintaining at least α_i factor.

Using these variables; the lower-support and upper-support aggregate, of AMI (LSA-AMI & USA-AMI) are estimated as:

$$LSA - AMI = \frac{1}{M}\sum_{i=1}^{M} \alpha_i SL_i \tag{6.10}$$

$$USA - AMI = \frac{1}{M}\sum_{i=1}^{M} \alpha_i SU_i \tag{6.11}$$

Subsequently, equations (6.10 & 6.11) are used to estimate aggregate AMI ($AMI_{agg.}$) as:

$$AMI_{agg.} = \frac{1}{2}[LSA - AMI + USA - AMI] \tag{6.12}$$

The used lower and upper bound values of chosen simple intervals of control parameters $\Delta\lambda$, $\Delta\tau$, and Δa are given in appendix in Table-5.1. Considering five sub-intervals-spans, i. e., M =5 with associated α-cuts factors in equations (6.10 & 6.11), AMIagg is obtained and listed in Table-5.1.

Using parametric values of LSA-AMI and USA-AMI corresponding to the smaller and the larger intervals of $\Delta\lambda$, $\Delta\tau$, and Δa as listed in Table-5.1; a set of expressions that offer lower and upper bound scalar value of β is formulated as:

$$\beta_l = \delta_s(\Delta\lambda_l)(\Delta a_l) + \delta_t(\Delta\lambda_l)(\Delta\tau_l) \tag{6.13}$$

$$\beta_u = \delta_s(\Delta\lambda_u)(\Delta a_u) + \delta_t(\Delta\lambda_u)(\Delta\tau_u) \tag{6.14}$$

In the above expressions; $\Delta\lambda_l$ & $\Delta\lambda_u$, $\Delta\tau_l$ & $\Delta\tau_u$, Δa_l & Δa_u represent LSA-AMI and USA-AMI of intervals $\Delta\lambda$, $\Delta\tau$, and Δa respectively. Depending upon situations, weight factors δ_s and δ_t are appropriately chosen. In variety of applications; spatio-temporal parameters influence the query generation process by different extent. On importing relevant estimated parameters as listed in Table-5.1 in equations (6.13 and 6.14); resultant lower and upper bound values of β i.e., β_l and β_u are given in Table-6.4. These scalar values of β are subsequently used in equation (6.2) to estimate the probability of query generation, which obviously inherent spatio-temporal dynamics associated with it. These values are estimated corresponding to three different scenarios as: (i) Only spatial aspect taken into considerations ($\delta_s = 1$ and $\delta_t = 0$) (ii) Temporal aspect dominates over the spatial aspect ($\delta_s = 0$ and $\delta_t = 1$) and (iii) Spatial and Temporal aspects encompass equal emphasis (with $\delta_s = \delta_t = 0.5$). For illustration purpose, during simulation exercise to estimate network performance measures a case is considered wherein spatial and temporal parameters assert equal priority.Further,the energy consumption of individual nodes as well as of the entire network configuration hinges upon operational mechanism and heuristics adopted to form clusters, to select cluster heads and sink attributes etc. and are discussed in the next section.

6.5 Operational Mechanism and Heuristic for Multiple Entities

Operational mechanism and heuristic (OMH) for clustering, cluster head (CH) selection, energy-centroid (EC) estimation, sink attributes, time scheduling for sinks relocation and a strategy to enable quadrants fusion are discussed next in sub-sections.

6.5.1 OMH for Clustering, CH and EC estimates

For better spatial resolution; the whole service-area is divided into four uniform quadrants, which are further decomposed into several non-uniform clusters. Each quadrant possesses total number of 25 sensor nodes. Minimum distance between any two nodes is constrained as 10 meters, and the coverage radius of each sensor node is considered as 30 meters. Sensor nodes that belong to different clusters sense/measure attributes of physical parameters in their vicinity and communicate these measured attributes to sink(s) through multi-hop wireless links. For cluster formation; four clustering schemes namely SKM, SFCM, DKM and DFCM are implemented. Initially, the clusters are formed randomly and subsequently it is executed based on RES of associated sensor nodes. At sporadic sampling instants for each cluster; node having maximum RES is elected as cluster head (CH). Further, the energy-centroid (EC) positioning for each cluster is estimated, which is primarily governed by Cartesian coordinate of sensor nodes and their associated RES [(Kumar and Chaturvedi, 2015)]. Using these known ECs locations at the network level, subsequently the optimum locations of sink(s) are estimated in anticipation that it might avoid occurrence of *hot-spot/islanding* till the network attains the stipulated lifetime.

6.5.2 OMH for Sinks Attributes and Scheduling Timings for Sinks Relocation

In terms of sinks attributes implications, four different case studies are investigated that include monitoring of a given service-area using (i) a single-stationary sink (SSS), (ii) a single-portable sink (SPS), (iii) four-stationary sinks (FSS) and (iv) four-portable sinks (FPS). In SSS and SPS, initially a sink is arbitrary located within the service-area and is permissible to relocate in SPS scenario. Inference about new plausible locations is governed by EC estimate. Whereas, in FSS and FPS; at the onset of network operation, each quadrant of the service-area is monitored by exclusively one sink and its initial location is arbitrary fixed within the quadrant or at the periphery of the quadrant. In FPS scenario; sinks relocation aspects are regulated by RES measure driven EC estimate.

For much better spatial resolution and to avoid the energy-hole problem during FSS and FPS; each quadrant of the service-area is monitored by a sink. In FPS based surveillance operation; the sinks remain positioned at their initial arbitrary locations till the ARES of all the quadrants is more than 95% (arbitrary chosen) of the initial energy reserve (IER) estimate that prevails at the onset of network services. With progressive time (t)/iteration count (n); as ARES for a quadrant attains 95% value of IER, the specific quadrant's sink is relocated to a new plausible location, which maintains close proximity to EC while satisfying relevant physical and operational constraints. Subsequently, this exercise is repeated whenever ARES attains (100 5n)% of the IER. The vary exercise is practiced for every quadrant of the service-area.

6.5.3 OMH Strategy for Quadrants Fusion

Step by step procedure to enable quadrants fusion module is as follows:

1. In concatenation ARES for, the clusters of each quadrant, whole quadrant and the entire network are estimated. The ARES estimate for the four quadrants's clusters and entire network is denoted as $ARES_{cls-q-1}$, $ARES_{cls-q-2}$, $ARES_{cls-q-3}$, $ARES_{cls-q-4}$, and $ARES_{nw}$ respectively.

2. The normalize ARES for all the four quadrants is estimated as

$$ARES_{norm-i} = \frac{ARES_{cls-q-i}}{ARES_{nw}} \tag{6.15}$$

Where i is quadrants index and i = 1, 2, 3 and 4.
Subsequently the aggregate normalize ARES is estimated as

$$ARES_{norm-agg.} = \frac{1}{4}\sum_{i=1}^{4} ARES_{norm-i} \tag{6.16}$$

Using the above mentioned ARES based normalized energy measures; for every quadrant pair (i, j), the following inequality norm is evaluated:

$$\mid ARES_{norm-i} - ARES_{norm-j} \mid \langle ARES_{norm-agg.} \tag{6.17}$$

Here i $\neq$ j, and i, j = 1, 2, 3 and 4.

Subjected to satisfactorily compliance of equation (6.17); no change takes place in the clustering dynamics of that particular quadrants pair. Further, the extension of this criterion to all the combinations of quadrants pairs leads to a situation wherein, clusters formation takes place in usual sense in all the quadrants. Depending upon degree of belongingness, sensor nodes can be member of multiple clusters and such clusters could span either any one quadrant or may spread across two or more than two quadrants. The other routine mechanisms such as dynamic cluster formation within a quadrant or across quadrants and the subsequent selection of cluster-heads (CHs) for all the clusters work in a usual manner.

However, at other observation instants for a particular quadrants pair (i, j), if the inequality norm of equation (6.18) satisfies as:

$$| ARES_{norm-i} - ARES_{norm-j} | \geq ARES_{norm-agg.} \tag{6.18}$$

Then that particular quadrants pair (i, j) is fused into a single entity (sub-area). Whereas, the other two quadrants that flout the equation (6.18) norm operate independently. Implications of the quadrants fusion process could be interpreted as: the two quadrants that participate in the fusion process now represent a single continuum that signifies half of the service-area. It instigates restrictions on the cluster formation mechanism as: the clusters that belong to a sub-area or fused quadrants could no longer belong to or spread in other two quadrants. Thus, the quadrants pair (i, j) that satisfies equation (6.18) create dichotomy with two other quadrants pair (k, m) which fails to comply with equation (6.18). Further, the clusters that are spread across the quadrants pair (k, m) bear no such restriction and could belong (continue their associations) to either k or m or both the quadrants based on participating sensor nodes degree of belongingness. Collectively, the typical modes of query generation, specification of sensor nodes and the service-area under surveillance ascertain the desired network service norm over the stipulated lifetime and these issues are presented in the next section.

6.6 Query Characteristics and Network parameters specification

The characteristics of typical queries considered in this chapter along with sensor nodes and service-area specifications are discussed in this section.

6.6.1 Query characterization and sensor nodes specifications

Queries nature considered in this chapter is of pull and unstructured WSNs (prior to sensing) [(Rachuri et al., 2008)]. In that sink nodes send simple and one shot queries to detect presence and identification of target type. Queries correspond to inquiry about the presence of civilians, soldiers, and vehicles. For that matter, multi-modal sensors (in this chapter, models of three different type of sensors [(Dutta et al., 2005)] are considered. In the prototype program three different fields are created in the query packets, which are features extracted from infrared, magnetic and acoustic sensors as used in [(Dutta et al., 2005)]. The movement of these targets (civilians/soldiers/vehicles) is considered to maintain an average speed of 1.5 m/sec., 3m/sec. and 15 m/sec. Owing to different speed these targets remain in coverage radius of a sensor node for different short epochs. During simulation execution over 20 different random runs; on average basis, each of these three different types of targets remain in coverage field of one particular sensor node for 2.05 sec. short duration epoch.

Each of these three different types of sensors consumes on average 0.4 mJ for transmission and reception [(Crossbow Technology, 2007)]. Thus transceiver consumes 0.8 mJ. For three different types of onboard sensor and based on average contact duration of 2.05 seconds, it amounts to approximately 4.8 mW power consumed by each mote (which supports transceiver module of three different types of sensors). Usually, the transceiver module consumes significant proportion of power compared to the other subsidiary subsystems. For querying and sensing the presence of these targets an aggregate power budget of 1.2 mW is considered [(Rachuri et al., 2008)]. For the sake of simplicity costs resulting from the other power consuming components are ignored.

6.6.2 Service area and β Specifications

To validate the proposed mathematical framework; an arbitrary service-area of 250*250 square meter comprises of 100 homogeneous sensor nodes is considered. Presuming 6mW average power consumption per query; each sensor node is equipped with two AA alkaline batteries that support 1000 hours of continuous operation [(Dutta et al., 2005)]. Thus, each sensor node has an initial energy budget of 6 Watt hour (Whr). In the subsequent simulation analysis; during the assessment of energy performance measures, CRES threshold mark is arbitrary set as 10% of initial energy while, the network stipulated lifetime is considered as 45 days. Values of other parameters used during simulation exercise are given in Table-5.1. [(Cheng et al., 2003a), (Dutta et al., 2005) and (Huang et al., 2007)].

For illustration to show efficacy, the energy performance measures and other auxiliary measures are estimated for five different scenarios; each of which considers equal emphasis to spatio-temporal parameters to approximate probability of query arrival and its spatial distribution. As listed in Table-6.4; the spatio-temporal aspects driven process leads to the lower and upper bound β values of [2.67, 5.18] and [1058.58, 1554.6] for the smaller and larger intervals respectively. For all the network scenarios; owing to highly correlated trends obtained about performance measures; these measures are estimated and analyzed for two extreme values of β corresponding to extracted lower and upper values from the smaller and larger intervals respectively and is referred as β-set in the rest of the chapter, i.e., β-set = [2.67, 1554.6]. β-set values are used to estimate the probability of query which is inferred from associated spatio-temporal parameters.

6.7 Simulation Results

To validate the proposed scheme; simulation is carried out for the network specifications as given in sections 6.6.1 and 6.6.2. The proposed analytical framework comprises multi-facet aspects such as uses of different clustering schemes, sinks driven surveillance while exploiting associated multiplicity and motion related attributes, and the types of models (plane-interval/fuzzy-interval) deployed to encompass parametric uncertainties which are subsequently used to draw inference about spatio-temporal parameters governed query generation process. Further, exploring the solutions in higher-dimensional space always entrust its superiority over the solutions that are derived from lower-dimensional space [(Wolsey and Nemhauser, 1988)]. To validate this well-established corollary from the optimization realm, a concept of quadrants fusion (as discussed in section 6.5.3) is implemented. It operates on the principles of satisfying the energy gradient based inequality norms.

On incorporating these above mentioned hybrid approaches; various case-studies/network surveillance scenarios are envisioned. In a variety of network surveillance scenarios, the simulation results of interest are: (i) ARES estimation at node/cluster/quadrant /network-level, (ii) Time since onset of network operation by which network ARES attains predetermined CRES mark, (iii) ARES temporal characteristics over a lifetime period corresponding to uses of four clustering schemes and determining the most energy efficient clustering scheme, (iv) ARES estimate at sporadic time instants corresponding to instants at which quadrants fusion takes place, (v) Information about quadrants that participate in fusion process (thus create sub-areas) at some arbitrary time instants, (vi) Locus of single-portable-sink (SPS) and four-portable-sink (FPS) in entire service-area or in specific quadrants with/without exercising quadrants fusion mechanism, and (vii) Service-time-duration (STD) in units of time (in this chapter it is treated as number of days) till network ARES attains predetermined CRES mark.

On the basis of sink attributes; aforementioned performance measures that directly or indirectly affect the network overall energy status and thus the longevity of sensor network are estimated for the following cases:

(a) Network surveillance by a single-stationary-sink with quadrants-fusion (SSSWQF)
(b) Network surveillance by a single-portable-sink with quadrants-fusion (SPSWQF)
(c) Network surveillance by four-stationary-sinks with quadrants-fusion (FSSWQF)
(d) Network surveillance by four-portable-sinks without quadrants-fusion (FPSWOQF)
(e) Network surveillance by four-portable-sinks with quadrants-fusion (FPSWQF)

Further, to utilize space meticulously; few of the network performance measures are summarized briefly in a tabular form. Two extreme profiles of the performance measures corresponding to network surveillance by SSSWQF, FPSWOQF, and FPSWQF are presented and analyzed next.

6.7.1 Network surveillance by a SSSWQF

In all the case studies; a square-shape service area of 250*250 square meters is considered. The sink is arbitrarily located at the periphery of the service-area, which is resolved into four uniform size quadrants. The sink is located at the boundary interface between the outside world and any of these quadrants boundaries. For a stipulated lifetime of 45 days, on using four different clustering schemes the ARES characteristics are shown in Fig. 6.2. These ARES characteristics corresponding to a situation in which inference about the associated spatio-temporal parameters are drawn from plane-interval based Poisson PMF. During this scenario depending upon compliance with energy metric based inequality norm, the eligible quadrants participate in fusion process. On using four clustering schemes; time duration in number of serving days by which network ARES attains predetermined CRES mark is listed in Table 6.1 in the leftmost column.

The uses of none of these clustering schemes meet the demand for a stipulated lifetime period of 45 days. Among the four clustering schemes; DFCM emerges as the most energy savior, though its uses also fail to deliver the service longevity norm as ARES attains CRES mark by the 20th day of operation. Since intervals of parameters also exhibit reasonably good degree of uncertainty, so the fuzzification of intervals-span in a way could render more approximate values. With this rational; the uncertainties associated with spatio-temporal parameters are incorporated in the fuzzy-interval form and its subsequent uses offer statistics about the number of serving days in 2nd column of Table 6.1. On comparing first two columns of Table 6.1; it shows marginal rise in number of serving days for spatio-temporal-fuzzy-interval (STFI) model driven inferences over the results obtained from spatio-temporal-plane-interval (STPI). Another observation about network ARES by keeping fusion count index (FCI) as one of the variables is shown in Fig. 6.3(a). Information about time instants by which fusion process gets activated is shown in Fig. 6.4(a). In spite of using the optimal clustering scheme (here, it is DFCM), fusion strategy and inference from STFI, the network is unable to reach close to the stipulated lifetime mark of 45^{th} day. This speculates to other situations wherein SPS, FSS, and FPS are deployed for network surveillance and are presented next.

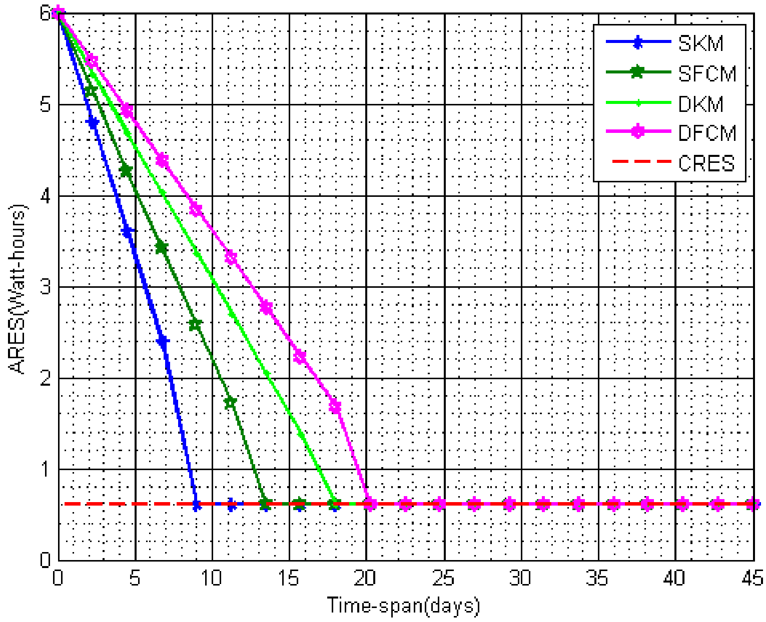


Figure 6.2: ARES using SSSWQF with inference from STPI

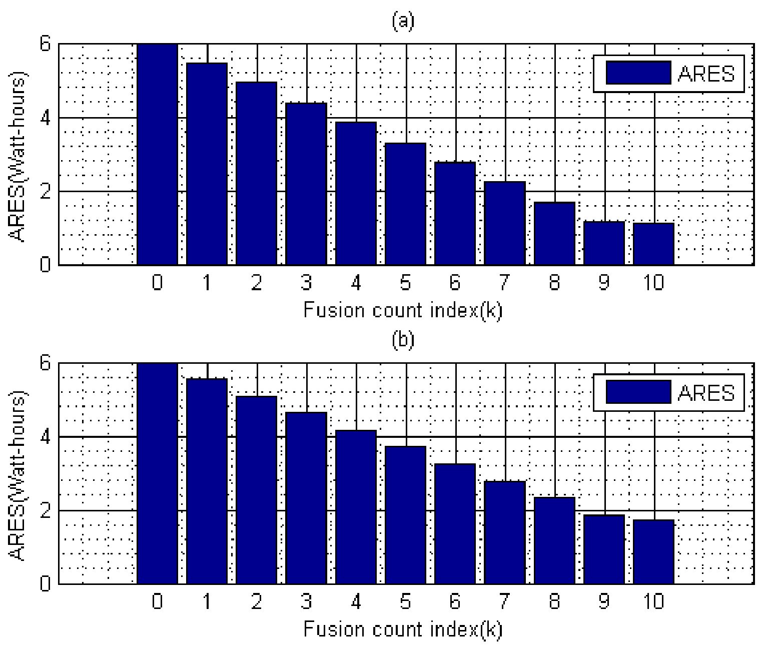


Figure 6.3: ARES variations with FCI (k) using STFI for (a) SSS and (b) SPS

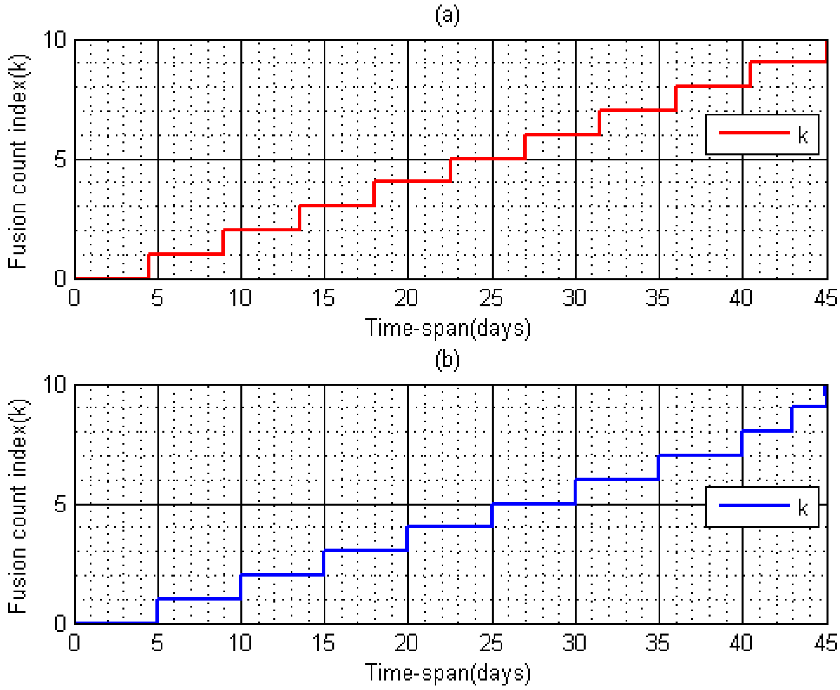


Figure 6.4: Time dependent incremental FCI (k) using STFI for (a) SSS (b) SPS

6.7.2 Network surveillance by a SPSWQF

In this scenario; initially the sink is located at the periphery or inside the service-area. Subsequently based on energy gradient spatial distribution, it is relocated to appropriate locations to maintain the balanced profile of energy consumption across the network nodes [(Kumar and Chaturvedi, 2015)]. Following it, the criteria for quadrants fusion is implemented that owes to satisfaction of energy inequality norm and is based on relative normalize ARES estimate of the quadrants. Implication of uses of four different clustering schemes, drawing inference about parametric uncertainties using STPI and STFI and quadrants fusion module results into desired service norms for a finite period. The rightmost column of Table 6.1 contains details about number of serving days by which network ARES attains predetermined CRES mark. Here too, the network topologies resulting from uses of four clustering schemes fail to attain stipulated lifetime of 45 days.

Treating FCI (k) as heuristically governed variable; the typical variations in network ARES is shown in Fig. 6.3(b). It may be of interest to keep track of time instants at which quadrants get fused. For the SPSWQF scenario; variation in FCI (k) over the lifetime is shown in Fig. 6.4(b). In deciding the overall performance of WSN, a sink is an important entity and its significance is hard to overlook. Apart from many sink related attributes, one of the important features of interest is its locations and associated constraints that limit relocation. The locus of a single sink in SPS scenario corresponding to without quadrants fusion (SPSWOQF) and with quadrants fusion (SPSWQF) is shown in Fig. 6.5 and 6.6 respectively.

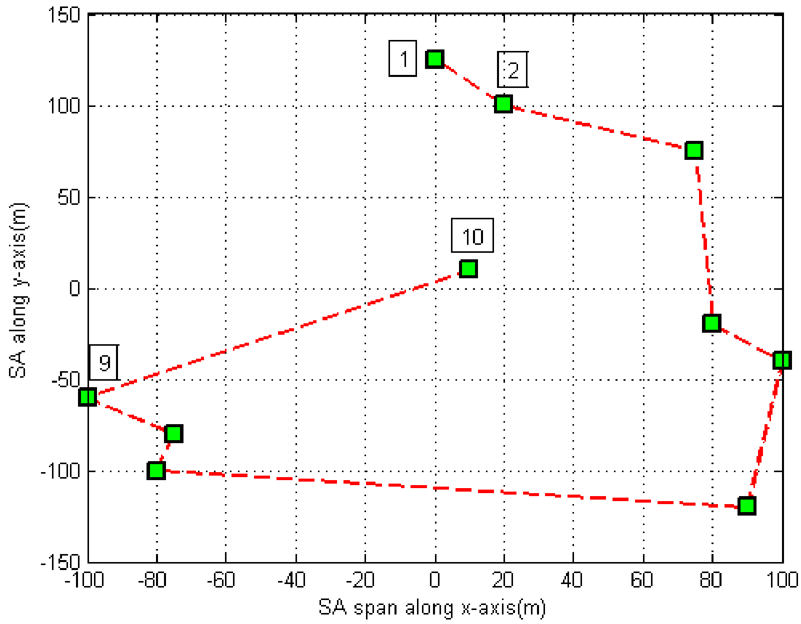


Figure 6.5: Locus of a single sink without quadrants fusion (SPSWOQF)

6.7.3 Network surveillance by FSSWQF

In this case study; one sink is arbitrary fixed at the periphery of each quadrant, thus to monitor whole service-area four stationary sinks are deployed in anticipation of alleviating some of the shortcomings experienced in previous two cases. In an analogous manner to previous case studies; a combination of four clustering schemes, inference based on STPI and STFI models and quadrants fusion module results into a specific scenario for which performance measures are estimated. Corresponding to the uses of different clustering schemes; the number of serving days by which the network ARES attains CRES threshold is listed in the leftmost column of Table 6.2. The typical variation in network ARES with incremental change in FCI (k) is shown in Fig. 6.7(a). The information about different time instant (days) at which service-area undergoes the quadrants fusion process is depicted in Fig. 6.8(a). The obtained results here too are disappointing one as the uses of all the four clustering schemes combined with STPI or STFI fail to offer reliable service over a stipulated lifetime. However, in comparative terms outcomes are better than that of obtained from SSSWQF and SPSWQF.

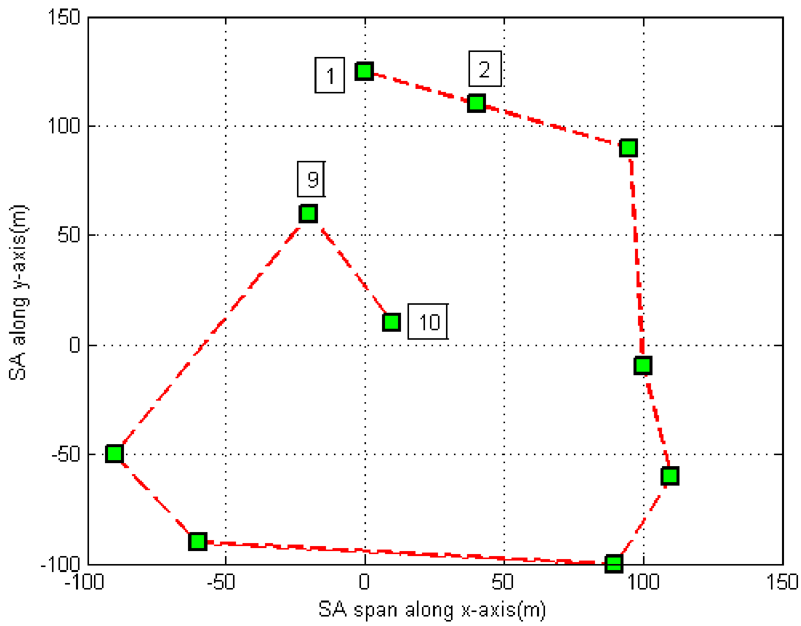


Figure 6.6: Locus of a single sink with quadrants fusion (SPSWQF)

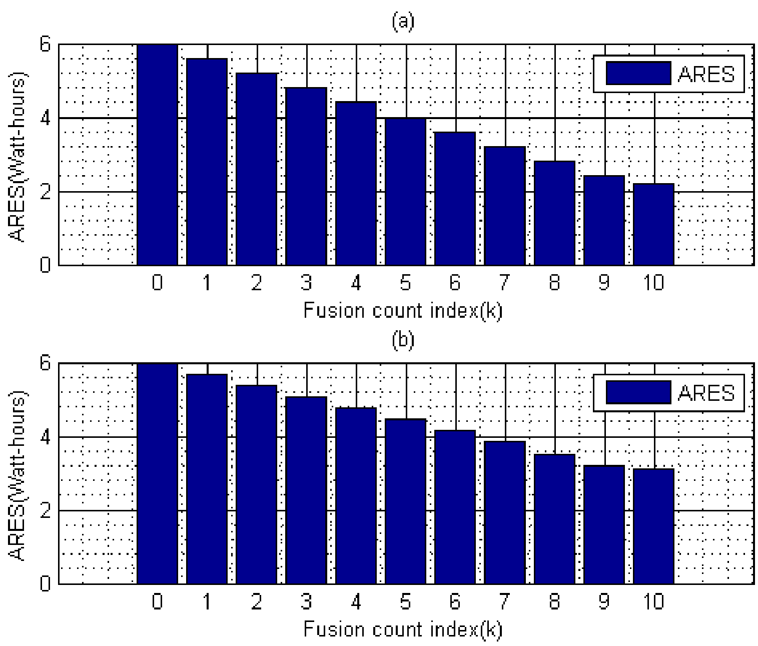


Figure 6.7: ARES variations with FCI (k) using STFI for (a) FSS and (b) FPS

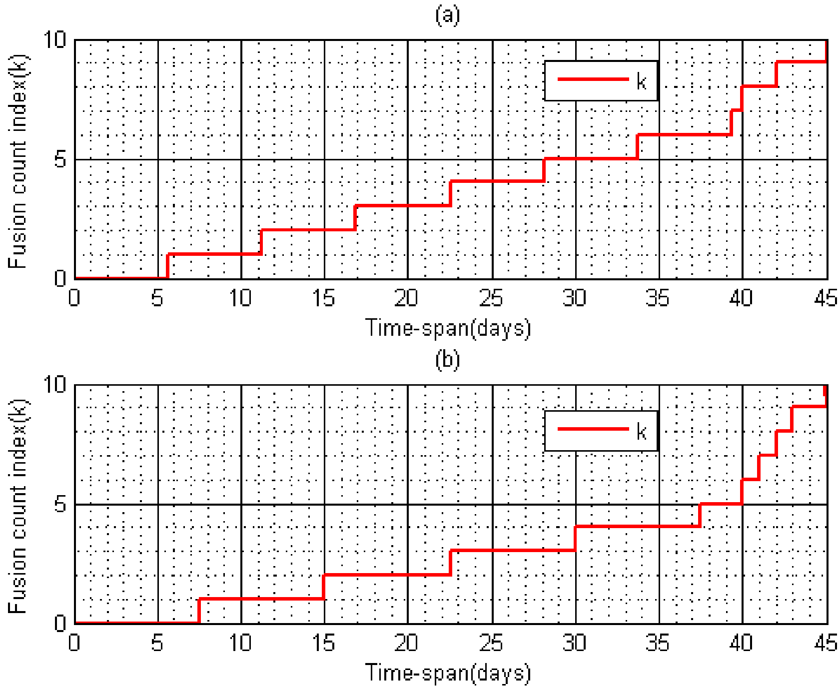


Figure 6.8: Time dependent incremental FCI (k) using STFI for (a) FSS (b) FPS

6.7.4 Network surveillance by FPSWOQF

To perform surveillance task in this scenario, each quadrant of a given service-area commences its operation with a dedicated sink that is arbitrary stationed at periphery or inside the quadrant. Thus, initially each quadrant comprises of one randomly deployed sink and later based on the energy metric driven heuristic these sinks could be relocated. Implementing four clustering schemes in conjunction with inferences from STPI and STFI models result into number of serving days statistics and is reported in the rightmost column of Table 6.2.

Interestingly, the last two rows of the Table 6.2 give useful insight about the tradeoff behavior between the sink's stationary/mobility aspects and compliance/noncompliance status of quadrants fusion strategy. The uses of two most energy efficient clustering schemes, i.e., DKM and DFCM result in nearly same statistics about number of serving days for the following attributes pairs: (STPI, FSSWQF) with (STPI, FPSWOQF) and (STFI, FSSWQF) with (STFI, FPSWOQF). This indicates overall energy profile of the network maintains identical status. The tradeoff can be envisaged in terms of saving incurred with sinks stationary state against implementing quadrants-fusion module and spending energy to impart motions to sinks against a network state wherein the quadrants-fusion module is implemented. The network ARES characteristics corresponding to uses of four clustering schemes with disabled quadrants fusion mechanism are shown in Fig. 6.9 and Fig. 6.10 for STPI and STFI models respectively. In both of these figures, the most promising clustering scheme (DFCM) results in ARES profile that attains CRES mark quite prematurely. In current (FPSWOQF) scenario; over a lifetime span corresponding to ten heuristically driven time instants, the locus of four independent sinks in four different quadrants of a service-area are shown in Fig. 6.11 (a-d). In all the quadrants; the sinks are randomly deployed and their initial locations are marked as 1, while final locations of sinks are denoted as 10. The journey from initial location-1 successively to location-10 depends upon ARES estimate regulated EC locations [(Kumar and Chaturvedi, 2015)]. Even if the initial positions of the sinks in respective quadrants are mandatorily fixed at some specific predetermine location, owing to distinct spatio-temporal parameters driven query generation process, the locus of sink movements within a quadrant exhibits distinct contour patterns.

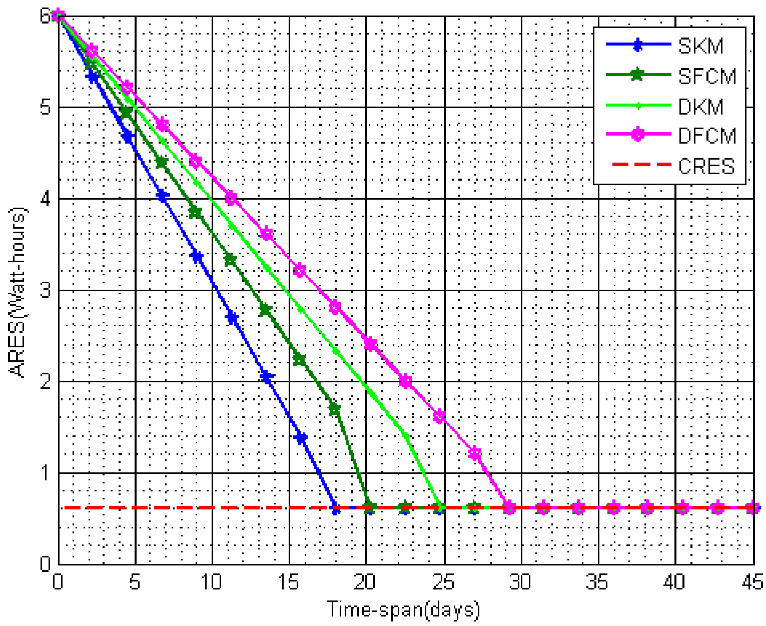


Figure 6.9: ARES using FPSWOQF with inference from STPI

6.7.5 Network surveillance by FPSWQF

Similar to previous case study, initially four sinks are deployed randomly, one each in every quadrant. Subsequently as time progress, depending upon ARES based measures as discussed in section 6.5; variety of task such as detection of appropriate timing for sinks relocations exercise, dynamic cluster formation, CHs selection and the energy inequality driven decision policy for quadrants fusion are performed.

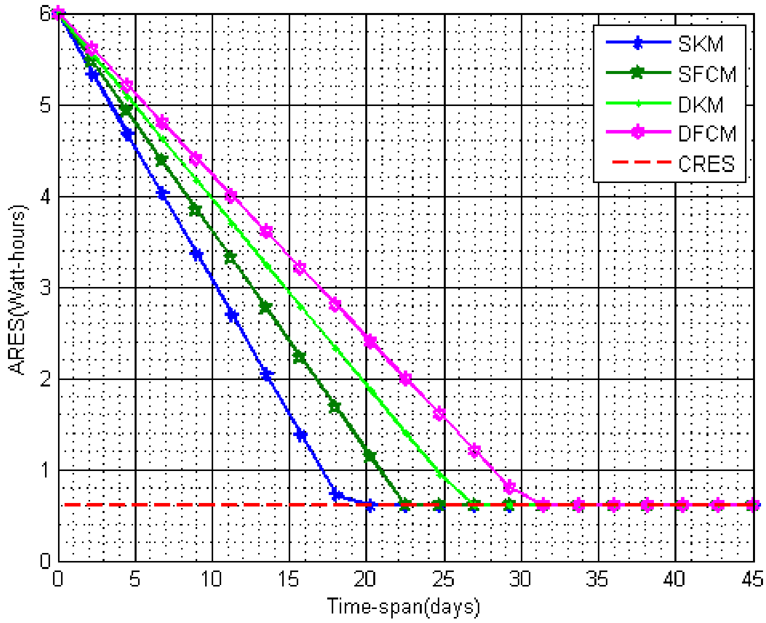


Figure 6.10: ARES using FPSWOQF with inference from STFI

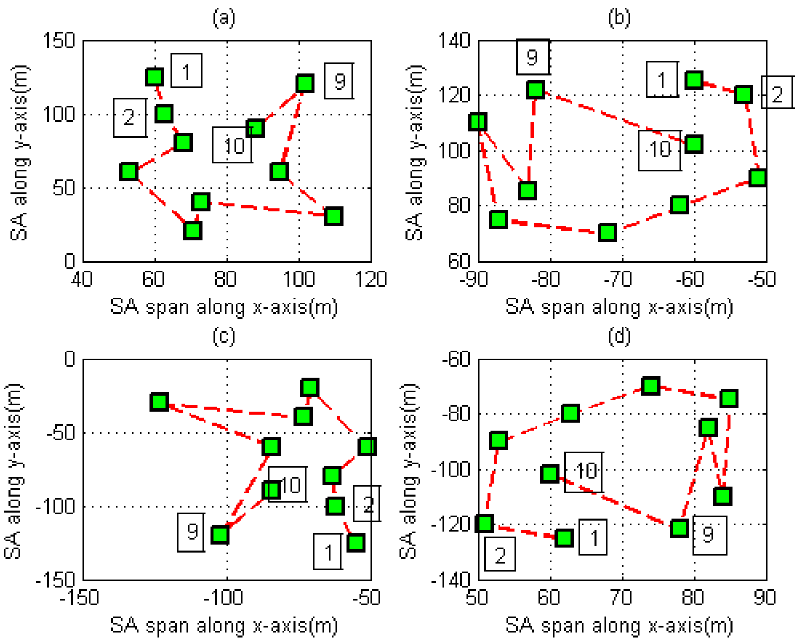


Figure 6.11: Locus of Four sinks without quadrants fusion (FPSWOQF) for (a) Sink-1 in quadrant-I, (b) Sink-2 in quadrant-II, (c) Sink-3 in quadrant-III, and (d) Sink-4 in quadrant-IV.

Corresponding to the uses of four clustering schemes; network ARES characteristics over the lifetime span associated with STPI and STFI driven inferences are shown in Fig. 6.12 and Fig. 6.13 respectively. In both of these figures, DFCM invariably outperforms over the other three clustering schemes. Uses of DFCM combined with inferences from STFI model facilitates reliable surveillance operation for a span of forty plus days, thus misses the stipulated lifetime by a narrow margin. This particular outcome is obtained with all possible care about modeling the uncertainty using appropriately chosen interval bound parameters value and their associated scaling on incorporating fuzzy-triangular distribution characteristics. Thus, in spite of taking into considerations the best possible care for uncertainty modeling aspect, the network fails to cater surveillance task for the entire lifetime span. It suggests insufficient IER (here it is taken as 6 Whr for each node) for the stipulated lifetime of 45 days. Hence, IER of all the sensor nodes must be carefully chosen while giving due considerations to the analytical model that addresses uncertainties associated with spatio-temporal parameters of query generation process.

In entire lifetime span; the network ARES variation with successive fusion mechanism is shown in Fig. 6.7(b). A comparison of Fig. 6.3 (a & b) and Fig. 6.7 (a & b) histograms especially with higher value of FCI, i.e., for $k \geq 8$ also validates merits of FPSWQF. A successive time instant at which quadrants fusion takes place is shown in Fig. 6.8(b). For the FPSWQF scenario; corresponding to ten different time instants the locus of four sinks in four different quadrants (one sink in each quadrant) are shown in Fig. 6.14 (a-d). Initially position of sinks in four quadrants is denoted as 1 and is randomly chosen. Later EC plays a significant role to regulate the movement of sinks and thereby establishes the unique locus.

In complementary terms, service-time-duration (STD) by which ARES of network attains CRES threshold mark is shown in Fig. 6.15 for five different combinations of sink attributes, status of quadrants fusion strategy, two best clustering schemes namely DKM, and DFCM and inference from STFI model. This also validates a superior performance on using DFCM along with inference from STFI model for FPSWQF scenario.

Further, based on heuristic about implementing the fusion mechanism in a lifetime span, fusion takes place at ten different time instants and is referred as fusion count index (FCI). As an illustration, information about which particular quadrants participate in fusion process is given in Table 6.3 for SPSWQF and FPSWQF scenarios corresponding to three arbitrary FCI (k = 1, 5 and 10). The graphical representation of quadrants pairs that participate in fusion process for SPSWQF (as listed in Table 6.3) are shown in Fig. 6.16. Initially at early stage of network operation, fusion doesn't take place so the origin of the entire service-area is indicated as "O". The notions O^1, O^5 and O^{10} are used to indicate the origin of sub-area that instigate from fusion of quadrants pair

$(Q-1, Q-4), (Q-3, Q-4)$ *and* $(Q-2, Q-3)$ corresponding to the fusion count k = 1, 5 and 10 respectively.

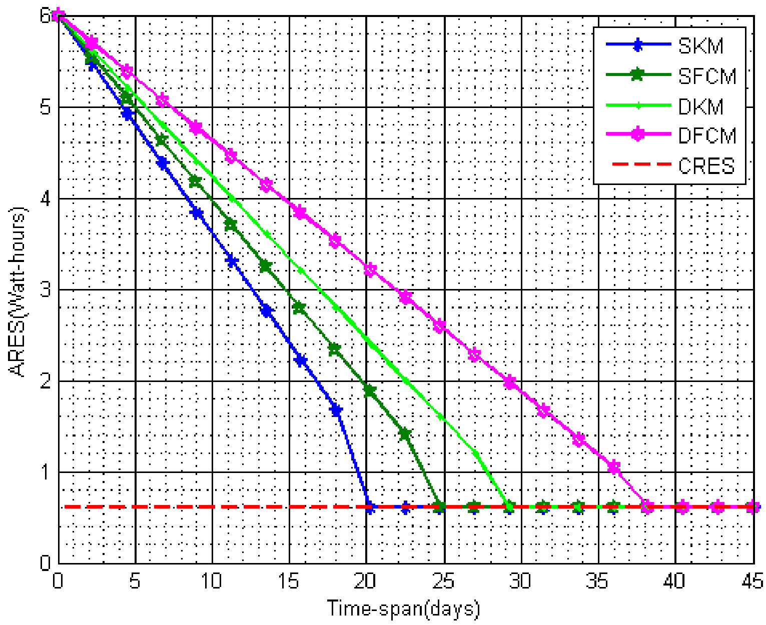


Figure 6.12: ARES using FPSWQF with inference from STPI

Table 6.1: Number of serving days by which ARES attains CRES threshold using a single sink

Sink attributes mixed with fusion status	**SSSWQF**		**SPSWQF**	
Spatio-temporal Inference model / Clustering schemes deployed	**STPI**	**STFI**	**STPI**	**STFI**
SKM	9	11	13.5	16
SFCM	13.5	16	18	20
DKM	17	21	20	22.5
DFCM	20	22.5	24.5	27

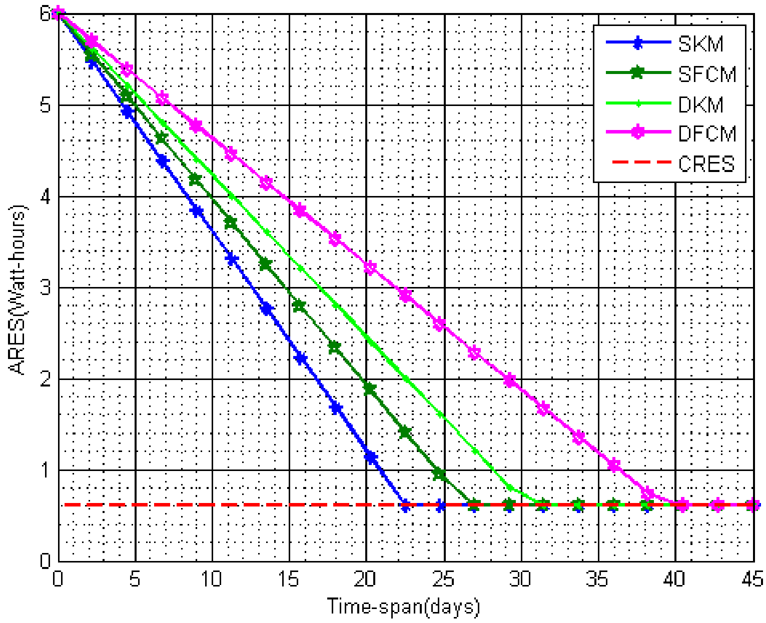


Figure 6.13: ARES using FPSWQF with inference from STFI

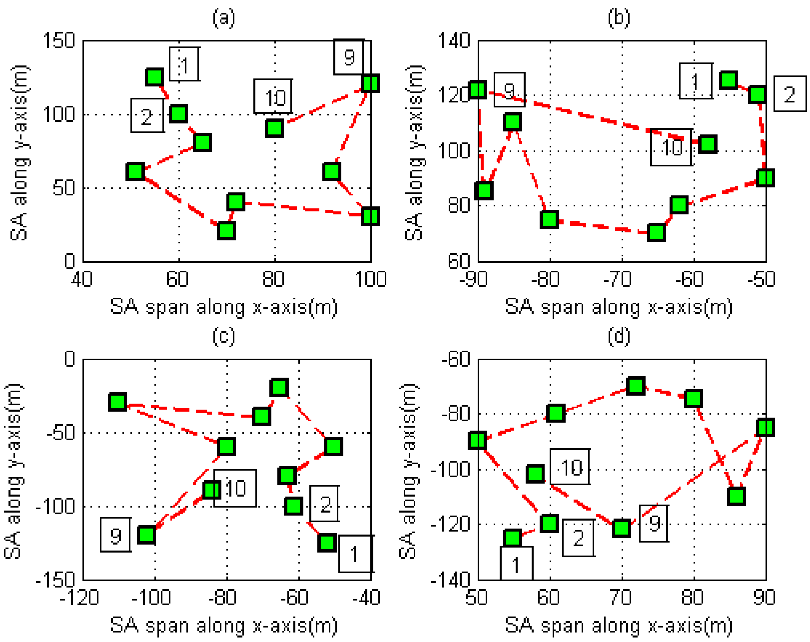
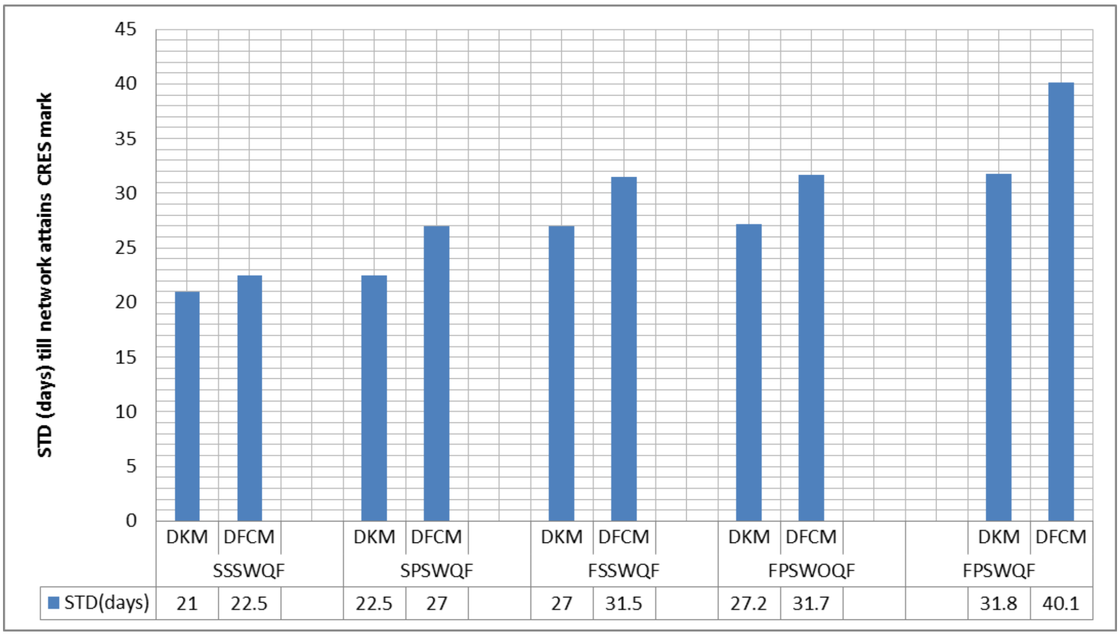


Figure 6.14: Locus of Four sinks with quadrants fusion (FPSWQF) for (a) Sink-1 in quadrant-I, (b) Sink-2 in quadrant-II, (c) Sink-3 in quadrant-III, and (d) Sink-4 in quadrant-IV.



	SSSWQF DKM	SSSWQF DFCM	SPSWQF DKM	SPSWQF DFCM	FSSWQF DKM	FSSWQF DFCM	FPSWOQF DKM	FPSWOQF DFCM	FPSWQF DKM	FPSWQF DFCM
STD(days)	21	22.5	22.5	27	27	31.5	27.2	31.7	31.8	40.1

Figure 6.15: A comparison of reliable STD for five different cases corresponding to the uses of DKM and DFCM, sink driven attributes and inference from STFI

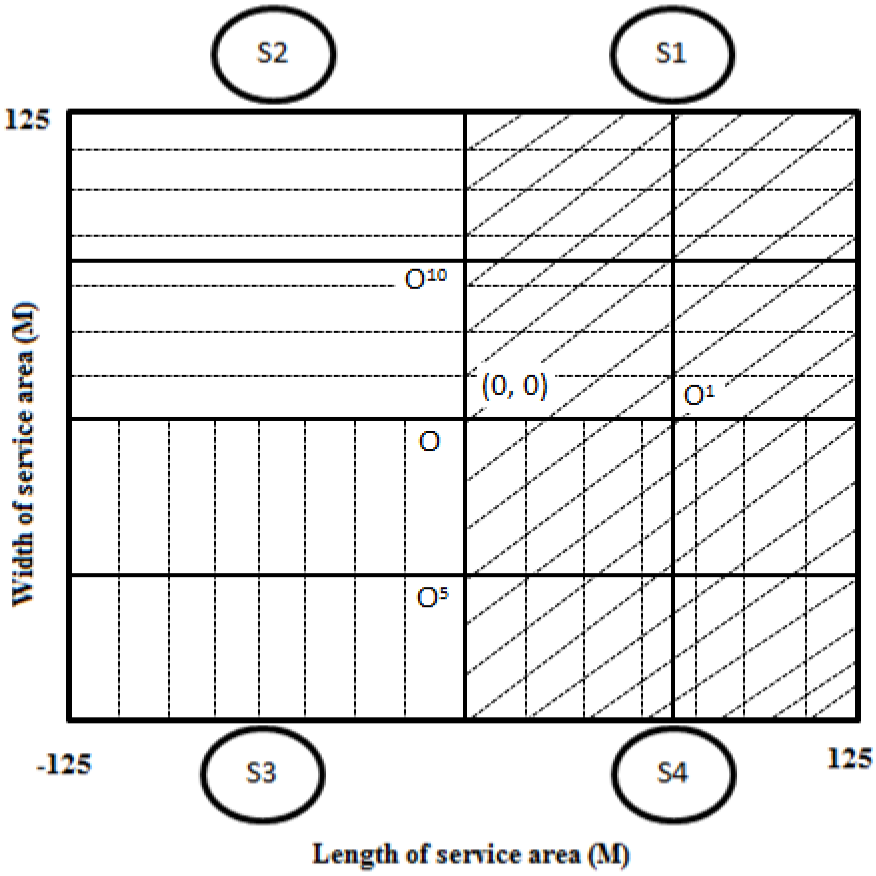


Figure 6.16: Geometrical representation of quadrants fusion process for k = 1, 5 and 10 in SPS scenario

Table 6.2: Number of serving days by which ARES attains CRES threshold using four sinks

Sink attributes mixed with fusion status	**FSSWQF**		**FPSWOQF**	
Spatio-temporal Inference model / Clustering schemes deployed	**STPI**	**STFI**	**STPI**	**STFI**
SKM	18	20	18	20.5
SFCM	22.5	24.5	20	22.5
DKM	24.5	27	24.5	27.2
DFCM	29	31.5	29.5	31.7

Table 6.3: Details of participating quadrants corresponding to three arbitrary values of k

Fusion count index (k)/ Sink attributes	**K = 1**	**K = 5**	**K = 10**
Single-portable sink	(Q1,Q4)	(Q3,Q4)	(Q1,Q2)
Four-portable sinks	(Q1,Q4)	(Q2,Q3)	(Q2,Q4)

Table 6.4: Scalar β inferred from LSA-AMI and USA-AMI for the Smaller and Larger intervals

	Smaller Intervals		Larger Intervals	
Intervals	Lower bound	Upper bound	Lower bound	Upper bound
Spatio-temporal	**2.67**	5.18	1058.58	**1554.6**

6.8 Conclusions

This thesis addresses spatio-temporal uncertainty issue acknowledged but largely overlooked by research community. To address it; query generation process is modelled using Poisson PMF with its single control parameter (λ) having dependency on plane/fuzzy-interval form of spatio-temporal parameters. Intervals forms are computationally intensive, to alleviate the computational burden AMI index is proposed that transforms intervals form to scalar values. Sink is an important entity between field deployed sensor nodes and outside world, to enhance the network lifetime sink multiplicity and motion aspects are investigated. From optimization realm; modeling a problem into higher dimension space always leads to a result that is definitely better than one that could be achieved in a reduced dimension subspace. Further, another degree of dimensionality enhancement is incorporated by a notion of quadrants fusion concept that works on energy metric driven inequality norms and create dichotomy of quadrants in which clusters formation mechanism follows different strategies. Again a significant improvement in all energy metrics is observed in network scenarios with an active quadrants fusion module. Invariably, irrespective of sinks attributes and the status of quadrants fusion strategy, the DFCM outperforms over other three clustering schemes. Finally, the chapter concludes with a note that appropriate considerations to spatio-temporal parameters driven uncertainties could ensure surveillance operation till the network attains stipulated lifetime. It also leads to auxiliary information about the sufficiency of the chosen IRE budget that could equip network designers with an advance estimate for battery selection thus voiding possibilities of network failures or poor network coverage issue owing to insufficient energy reserve.

Chapter 7

CONCLUSIONS AND FUTURE WORK

The Principal motive to undertake the research work is devising a framework that addresses the uncertainties which are inevitable with most of the physical phenomenon and its aftereffects on key performance measures. The main contribution of this thesis is to endow with heuristic and modeling framework, which rely upon hybrid mechanisms that comprise of the following strategies: usage of appropriate clustering schemes, energy-centroid (EC) location estimate to select cluster head (CH) for each cluster, queries spatial-distribution or dissemination using vector modulation approach, investigating sink(s) attributes in particular multiplicity and motion aspects, usage of prominent probabilistic models to exhibit dynamics associated with random arrivals of events/queries, during in situ usage of parametric Poisson probability mass function (PMF); regulating parameter of Poisson PMF in tune with associated spatio-temporal aspects that govern criticality of query generation dynamics, to incorporate varying degree of uncertainties treating spatio-temporal parameters as crisp entities, bound-values intervals and Fuzzified bound-values intervals, and exploiting the concept of spatial-fusion to transform the problem from lower-dimensional space to higher-dimensional space. Perhaps, these diversified approaches at first glance surface to meet different objective, however, all these mechanisms work harmonically to meet two interdependent objectives, which are (i) over the stipulated lifetime duration; maintaining the balanced energy gradient across the network that ensure uniform coverage and (ii) assuring that under the randomly varying scenarios, network surveillance task remains consequential at least for the stipulated network lifetime. During varieties of case studies reported in the thesis; to validate the efficacy of the above mentioned techniques, following performance measures are estimated and analyzed: residual energy status (RES) of individual sensor nodes, average residual energy status (ARES) of entire network, critical residual energy status (CRES) of individual sensor nodes, fraction of sensor nodes attaining CRES mark and the network service-time-duration (STD).

Summary

In chapter 2, the typical findings include: (i) usage of k-means and FCM clustering algorithms to realize hierarchical network topologies and subsequently electing cluster heads (CHs) and (ii) Euclidian distance measures based geometrical framework to determine optimal location of the sink and the classification of clusters with respect to arbitrary location of the sink using Euclidean distance measures. The typical performance measure RES is evaluated for dichotomy sets of sensor nodes that include simple sensor nodes and the cluster heads under two different network scenarios in which topological structure of sensor network is derived using k-means and FCM algorithms. In aggregate terms, simulation results yields that to handle a single query the FCM algorithm requires approximately 50

Chapter 3 presented the three major issues that affect the operational performance of WSN, which include (i) uses of four different clustering techniques namely SKM, SFCM, DKM and DFCM and their impact on energy estimate driven performance measures , (ii) approximating the spatio-temporal dependency of query distribution pattern using parametric uniform and Poisson PMF models, in that inference about parameter values is drawn from associated spatial and temporal aspects and (iii) enhancing the sensor network overall energy reserve status that in turn improve the longevity of network on exploiting the sink(s) attributes, particularly sink multiplicity and motion aspects that are treated as basis for three types of case studies. The efficacy of the proposed algorithm is evaluated using two different performance measures namely, the network ARES and the percentage of network nodes attaining predetermined CRES mark as time progresses towards the stipulated network lifetime. Invariably, the DFCM emerges as the most energy efficient clustering scheme. Further, on comparing the outcomes of considered case studies, the sensor network configuration having four stationary sinks with DFCM exhibits the superior performance as during this scenario network attains improved ARES.

Chapter 4 presented a mathematical framework that governs the dependency of generated query on spatial and temporal parameters in event/query based sensor networks. Owing to uncertainties associated with spatial and temporal aspects of query generation process, the holistic approach is proposed that keeps the spatio-temporal aspects of query generation at centre stage. The Poisson probability mass function (PMF) model is used that incorporates dependency of query generation process on the associated spatial and temporal parameters. Thus, the control parameter () of the Poisson PMF is regulated by spatial parameter (a) and temporal parameter (t) of query generation process. Further, to encompass widely varying degree of uncertainties; the bounded-values intervals of control parameters (, a and t) are considered instead of treating them as a crisp entity. The effectiveness of the proposed algorithm is validated on estimating two principal measures, which are: the ARES estimation with respect to predetermined CRES mark and the network STD with respect to stipulated network lifetime. Simulation outcomes established the inference that the combination of DFCM clustering scheme and four-stationary sinks based surveillance task appears an optimal choice ensuring that the sensor network attains the stipulated lifetime.

In Chapter 5, heuristics based mathematical framework is presented that considers approach proposed in chapter 4 as base framework. Thus, analogous to chapter 4, four clustering schemes namely SKM, SFCM, DKM and DFCM are deployed and sink attributes are also exploited in the same manner. However, in this chapter, uncertainties associated with spatio-temporal parameters are treated with an added feature in that instead of using simple bound-values intervals for spatio-temporal parameters, Fuzzified intervals are considered. Arithmetic operations are quite intricate especially when the bound-values intervals are involved. To mitigate the rigorous computation concerned with intervals-form, two indices AMI and GMI are proposed that transformed back intervals-form to crisp number. In portable-sink driven network surveillance; sink relocation aspect is governed by RES estimates of sensor nodes. In varieties of sink attributes driven scenarios; sensor network energy performance is estimated in terms of the following metrics: ARES estimate for all participating sensor nodes at cluster as well as at network level, and the network STD by which the network attains the predetermined CRES mark. The rationale for the multifaceted approach is to devise the query generation model that draws inference from uncertainties associated spatio-temporal parameters. This result in predicting the query density and its inter-arrival-time-rate more sensibly that could equip the network planning agencies to make more approximate estimate about network resources so as to ensure desired service over the stipulated lifetime.

Chapter 6 presented an extension to work carried out in chapter 5, in that the significant accompaniments are: inclusion of multiple portable sinks for improved monitoring performance and modeling the problem in high-dimension space by notion of spatial (quadrants)-fusion concept in anticipation of attaining much better values of network performance measures. The later task is implemented based on compliance of energy metrics driven inequality norm. The entire exercise leads to an inference that appropriate considerations to spatio-temporal parameters driven uncertainties could ensure desired network operation till the network attains stipulated lifetime. It also leads to auxiliary information about the sufficiency of the chosen IRE budget that could equip network designers with an advance estimate for battery selection thus voiding possibilities of network failures or deprived of minimal coverage requirement owing to insufficient energy reserve.

Future Work

In the sense of protocol design; the state of the art existing research in the WSNs regime is highly applications and underlying platform specific. However, to make Internet of things (IoT) a living phenomenon, many WSNs are supposed to operate in an interactive manner that requires a common communication infrastructure. Therefore, the WSNs protocol needs to compatible to various heterogeneous platforms as well as to plethora of applications. However, the task of developing the individual protocols becomes highly intricate and expensive as the degree of heterogeneity increases. Even with the highly complex mesh structure that comprises of many constituents WSNs, the issue like inclusion of uncertainties becomes quite challenging. Thus, to overcome these challenges, the following issues could be explored:

a. Cross-layer optimization concept may overcome some of the sufferings experienced at individual layers.

b. In highly constrained energy scenarios; game-theoretic models may contribute significantly in terms of devising interactive strategies.

References

A. Paulraj, R. N. and D.Gore (2003). "Introduction to space-time wireless communications". *Cambridge Univ. PressCambridge, U.K.*

Akylidiz, I. F., Su, W., Sankarsubramaniam, Y., and Cayrici, E. (2002). "A survey on sensor network". *IEEE Communication Magazine*, 40(8):102 – 114.

Alamouti, S. (1998). "Alamouti1998a simple transmit diversity technique for wireless communications". *IEEE J. Sel. Areas Commun.*, 16(8):1451–1458.

Alefeld, G. and Herzberger, J. (1983). "Introduction to interval computations". *Academic Press.*

Alla, B., Tiwari, S., A, E., A, B., and M.L, H. (2011). "Hierarchical adaptive balanced energy efficient routing protocol (habrp) for heterogeneous wireless sensor networks". In *International Conference on Multimedia Computing and Systems (ICMCS)*, pages 1–6.

Alwan, H. and Agarwal, A. (2009). "A survey on fault tolerant routing techniques in wireless sensor networks". In *Third International Conference on Sensor Technologies and Applications*, pages 366–371.

B.N, M., M, T. V., and R.V, D. (2010). "Dynamic route and threshold level selection in energy efficient wsn". In *3rd International Conference on Emerging Trends in Engineering and Technology (ICETET)*, pages 797–802.

Braginsky, D. and Estrin, D. (2002). "Rumor routing algorithm for sensor networks". In *in Proc. of the 1st ACM International Workshop on Wireless Sensor Networks and Applications, Atlanta, Georgia, USA.*

Caylrcl, P. C. N. E. (2005). "PCSMAC: A power controlled sensor-mac protocol for wireless sensor networks". *0-7803-8801-1/05/2005/IEEE.*

Chair, Z. and Varshney, P. K. (1986). "Optimal data fusion in multiple sensor detection systems". *IEEE Transactions on Aerospace and Electronic Systems*, 22(1):98–101.

Chamberland, J. and Veeravalli, V. V. (2003). "Decentralized detection in sensor networks". *IEEE Trans. Signal Processing*, 51(2):407–416.

Chen, J. (2012). "Improving life time of wireless sensor networks by using fuzzy c-means induced clustering". In *World Automation Congress (WAC)*, pages 1–4.

Chen, Y. and Zhao, Q. (2005). "On the lifetime of wireless sensor networks". *IEEE Communications letters*, 9(11):976–978.

Cheng, R., Kalashnikov, D. V., and Prabhakar, S. (2003a). "Evaluating probabilistic queries over imprecise data". In *Proceedings of the 2003 ACM SIGMOD International conference on Management of data*, pages 551–562.

Cheng, R., Prabhakar, S., and Kalashnikov, D. V. (2003b). "Querying imprecise data in moving object environments". In *IEEE Conerence Proceding (ICDE)*, pages 1–3.

Crossbow Technology, I. (2007). *Wireless Sensor Networks:Product Reference Guide*.

Cui, S., Goldsmith, A., and Bahai, A. (2004). "Energy-efficiency of mimo and cooperative mimo techniques in sensor networks". *IEEE Journal on Selected Areas in communication*, 22(6):1089–1098.

Cui, S., Goldsmith, A., and Bahai, A. (2005). "Energy-constrained modulation optimization". *IEEE Transactions on Wireless Communications*, 4(5):2349–2360.

Dam, T. and Langendoen, K. (2003). "An adaptive energy-efficient mac protocol for wireless sensor networks". In *In Proceedings of the First ACM Conference on Embedded Networked Sensor Systems, Los Angeles*.

Demirbas, M., Lu, X., and Singla, P. (2002). "An in network querying framework for wireless sensor networks". *IEEE TRANSACTIONS ON PARALLEL AND DISTRIBUTED SYSTEMS*, 20(8):1202–1215.

Deshpande, A., Guestrin, C., Madden, S. R., Hellerstein, J. M., and Hong, W. (2005). "Model-based approximate querying in sensor networks". *The VLDB Journal*, 14(4):417–443.

Dogandzic, A. and Qiu, K. (2008). "Decentralized random-field estimation for sensor networks using quantized spatially correlated data and fusion-center feedback". *IEEE Trans. on Signal Processing*, 56(12):6069–6085.

Drakopoulos, E. and C.Lee, C. (1991). "Optimum multi sensor fusion of correlated local decisions". *IEEE Transactions on Aerospace and Electronic Systems*, 27(4):593–606.

Dutta, P., Grimmer, M., arora, A., Bibyk, S., and culler, D. (2005). "Design of a wireless sensor network platform for detecting rare, random, and ephemeral events". In *Fourth International Symposium on Information Processing in Sensor Networks(IPSN2005)*, pages 497–502.

et al, K. S. (October 2000). " Protocols for self-organization of a wireless sensor network". In *IEEE Pers. Commun*, pages 17–27.

Fang, J. and Li, H. (2010). "Optimal/near-optimal dimensionality reduction for distributed estimation in homogeneous and certain inhomogeneous scenarios". *IEEE Trans. On Signal Processing*, 58(8):4339–4353.

Fei, Z., Li, B., Yang, S., Xing, C., Chen, H., and Hanzo, L. (2016). " A survey of multi-objective optimization in wireless sensor networks: Metrics, algorithms and open problems". In *IEEE Communication Surveys and Tutorials.*

Gandham, S., Dawande, M., Prakash, R., and Venkatesan, S. (2003). "Energy efficient schemes for wireless sensor networks with multiple mobile base stations". In *Proceedings of IEEE Globecom,San Francisco,CA*, volume 1, pages 377–381.

Gastpar, M., Dragotti, P., and Vetterli, M. (2006). "The distributed karhunen-loeve transform". *IEEE Trans. on Signal Processing*, 52(12):5177–5196.

Gatzianas, M. and Georgiadis, L. (2008). "A distributed algorithm for maximum lifetime routing in sensor networks with mobile sink". *IEEE Transactions on Wireless Communications*, 7(3):984–994.

Gharavi, H. and Ban, K. (2003). "Multihop sensor network design for wide-band communications". *Proceedings of the IEEE*, 91(8):1221–1234.

Gharavi, H. and Kumar, S. P. (2003a). "Scanning the issue-special issue on sensor networks and applications". *Proceedings of the IEEE*, 91:1151–1153.

Gharavi, H. and Kumar, S. P. (2003b). "Special issue on sensor networks and applications". *Proceding of IEEE*, 91(8):1151–1153.

Giannakos, A., Karagiorgos, G., and Stavrakakis, I. (2009). "A message optimal sink mobility model for wireless sensor networks". In *Proceeding of 8th International Conference on Networks*, pages 287–291.

Gini, F., Lombardini, F., , and Verrazzani, L. (1998). "Decentralised detection strategies under communication constraints". *IEEE Proceedings - Radar, Sonar and Navigation*, 145(4):199–208.

Guvenc Degirmenci, J. P. K. and Prokopyev, O. A. (2014). " Maximizing the lifetime of query-based wireless sensor networks". *Transactions on Sensor Networks (TOSN)*, 10(4)(56):1–24.

H, S., A, A. A., and Hosseini, G. (2009). "Distributed wsn data stream mining based on fuzzy clustering". *Symposia and Workshops on Ubiquitous, Autonomic and Trusted Computing ,UIC-ATC '09*, pages 395–400.

Hartwell, J. A., Messier, G. G., and Davies, R. J. (2007). "Optimizing physical layer energy consumption for wireless sensor networks". In of Electrical, D., Engineering, C., and TRLabs, U. o. C., editors, *IEEE Proc.*, pages 1550–2252.

Heinzelman, W., Chandrakasan, A., and Balakrishnan, H. (2000). "Energy efficient communication protocol for wireless micro sensor networks". In *Proc. of the 33rd Annual Hawaii International Conf. on System Sciences*, pages 3005–3014.

Heinzelman, W., Kulik, J., and Balakrishnan, H. (1999b). "Negotiation- based protocols for disseminating information in wireless sensor networks". In *Proc. of the 5th Annual ACM/IEEE International Conf. on Mobile Computing and Networking.*

Heinzelman, W., Kulik, J., and Balakrishnan, H. (seattle,WA,August, 1999a). " Adaptive protocols for information dissemination in wireless sensor networks". In *Proceedings of the 5th Annual ACM/IEEE International Conference on Mobile Computing and Networking (MobiCom99).*

Heinzelman, W. B., Chandrakasan, A. P., Senior, and Balakrishnan, H. (2002). "An application-specific protocol architecture for wireless micro sensor networks". *IEEE Transactions On Wireless Communications*, 1(4).

Holland, TianqiWang, Tavli, B., Seyedi, A., and Hemzelman, W. (2009). "Optimizing physical layer parameters for wireless sensor networks". *ACM Journal.*

Hu, J. and Blum, R. S. (2001). "On the optimality of finite-level quantizations for distributed signal detection". *IEEE Transactions Information Theory*, 47(4):1665–1671.

Huang, H., Hartman, J. H., and Hurst, T. N. (2007). "Efficient and robust query processing for mobile wireless sensor networks". *International Journal of Sensor Networks*, 2(1/2):99–107.

IEEE (2012). "IEEE standard for information technology–telecommunications and information exchange between systems local and metropolitan area networks–specific requirements part 11: Wireless lan medium access control (mac) and physical layer (phy) specifications". *IEEE Std 802.11-2012 (Revision of IEEE Std 802.11-2007)*, pages 1–2793.

Intanagonwiwat, C., Govindan, R., and Estrin, D. (August 2002). " Directed diffusion: A scalable and robust communication paradigm for sensor networks". In *Proceedings of the Sixth Annual International Conference on Mobile Computing and Networks (MobiCOM 2000)*.

Intanagonwiwat, C., Govindan, R., Estrin, D., and Heidemann, J. (2003a). "Directed diffusion for wireless sensor networking". *IEEE ACM Transactions on Networking*, 11(1):2–16.

Intanagonwiwat, C., Govindan, R., Estrin, D., Heidemann, J., and Silva, F. (2003b). "Directed diffusion for wireless sensor networking". *IEEE/ACM Trans. Networking*, 11(1):2–16.

Jain, N., Biswas, R., Nandiraju, N., and Agrawal, D. P. (2005). "Energy aware routing for spatio-temporal queries in sensor networks". In *IEEE Wireless Communications and Networking Conference*, volume 3, pages 1860–1866.

Jensen, C. S., syte, D. T., , and sauskas, N. T. (2006). "The cost benchmarkcomparison and evaluation of spatio-temporal indexes". *Lecture Notes in Computer Science*, 3882:125–140.

Kam, M., Chang, W., and Zhu, Q. (1991). "Hardware complexity of binary distributed detection systems with isolated local bayesian detectors". *IEEE Transactions on Systems, Man, and Cybernetics*, 21(3):565–571.

Kam, M., Zhu, Q., and Gray, W. S. (1992). "Optimal data fusion of correlated local decisions in multiple sensor detection systems". *IEEE Transactions on Aerospace and Electronic Systems*, 28(3):916–920.

Kannan, R. and Iyengar, S. S. (2004). "Game-theoretic models for reliable path-length and energy-constrained routing with data aggregation in wireless sensor networks". *IEEE Journal on Selected areas in Communications*, 22(6):1141–1150.

Karakaya, M. (2013). " Deadline aware energy efficient query scheduling in wireless sensor networks with mobile sink". *The Scientific World Journal*, 2013(Article ID 834653).

Karaki, J. N. A. and Kamal, A. (2004). "Routing techniques in wireless sensor networks: A survey". *IEEE Wireless Communications*, 11(6):6–28.

Kasetkasem, T. and Varshney, P. K. (2001). “Communication structure planning for multi sensor detection systems”. *IEEE ProceedingsRadar, Sonar and Navigation*, 148(1):2–8.

Kim, T., Kim, H., Sun, Y., and Jin, Z. Y. (2015). “ Physical layer and medium access control design in energy efficient sensor networks: An overview”. *IEEE Transactions on Industrial Informatics*, 11(1):2–15.

Kinalis, A. and Nikoletseas, S. (2007). “Scalable data collection protocols for wireless sensor networks with multiple mobile sinks”. In *Proceedings of the 40th Annual Simulation Symposium Annual Simulation Symposium. IEEE Computer Society, Washington, DC*, pages 60–72.

Kinalis, I. C. A. . and Nikoletseas, S. (2006). “Sink mobility protocols for data collection in wireless sensor networks”. In *Proceedings of the international Workshop on Mobility Management and Wireless Access, MobiWac06, Terromolinos, Spain*, pages 52–59.

K.Kalpakis, K.Dasgupta, and P.Namjoshi (2002). “Maximum lifetime data gathering and aggregation in wireless sensor networks”. In *Proc. of IEEE Networks Conference*, pages 1–13.

Kumar, P. and Chaturvedi, A. (2015). “ Sink attributes analysis for energy efficient operations of wireless sensor networks under randomly varying temporal and spatial aspects of query generation”. *AEU International Journal of Electronics and Communications*, 69(7):1058–1069.

Kumar, P., Chaturvedi, A., and M.Kulkarni (2012). ”Geographical location based hierarchical routing strategy for wireless sensor networks”. In *IEEE International Conference on Devices Circuits and Systems (ICDCS-12),India*, pages 9–14.

Kumar, S., Zhao, F., and Shepherd, D. (2002). “Collaborative signal and information processing in microsensor networks”. *IEEE Signal Processing Magazine*, 19(2):13–14.

K.Willett, P., Swaszek, P. F., and S.Blum, R. (2000). “The good, bad and ugly : distributed detection of a known signal in dependent gaussian noise”. *IEEE Trans. Signal Processing*, 48(12):3266–3279.

Kwon, T. J., Gerla, M., Varma, V. K., Barton, M., and Hsing, T. R. (2003). “Efficient flooding with passive clustering an overhead free selective forward mechanism for ad hoc/sensor networks”. In *Proceding of IEEE*, volume 91, pages 1210–1220.

Larsson, F. G. and Stoica, P. (2008). *Space -Time Block Coding for Wireless Communications*. Cambridge Univ. Press.

Li, C., Pu, F., AipingGu, Liu, L., and Cao, Q. (2011). "A sink-controlled energy-efficient and geographical routing protocol in wireless sensor networks". In *International Conference on Computer and Management (CAMAN)*, pages 1–4.

Li, X., Chen, M., and Liu, W. (2005). "Application of stbc- encoded cooperative transmissions in wireless sensor networks". *IEEE Signal Processing Letters*, 12(2):134–137.

Lindsey, S. and Raghavendra, C. (2001). "PEGASIS: Power-efficient gathering in sensor information systems". In *International Conf. on Communications*.

Liu, X., Zhao, H., and Li, X. (2011a). "EPC: Energy-aware probability-based clustering algorithm for correlated data gathering in wireless sensor networks". In *IEEE International Conference on Advanced Information Networking and Applications*, pages 419–426.

Liu, Y., Zhu, X., Ma, C., and Zhang, L. (2011b). "Multiple event detection in wireless sensor networks using compressed sensing". *IEEE 18th international conference on Telecommunications*, pages 27–32.

Luo, J. and Hubaux, J. (2003). "Joint mobility and routing for lifetime elongation in wireless sensor networks". In *Proceedings of IEEE Infocom , Miami, FL*, volume 3, pages 1735–1746.

Luo, J. and Hubaux, J.-P. (2010). "Joint sink mobility and routing to maximize the lifetime of wireless sensor networks:the case of constrained mobility". *IEEE/ACM Transactions on Networking*, 18(3):871–884.

Manjeshwar, A. and Agarwal, D. (2001). "TEEN: a routing protocol for enhanced efficiency in wireless sensor networks". In *International Proc. of the 15th Parallel and Distributed Processing Symposium*, pages 2009–2015.

Marta, M. and Cardei, M. (2009). "Improved sensor network lifetime with multiple mobile sinks". In *Proc. of Pervasive and Mobile Computing*, pages 542–555.

Mendes, L. D. and Rodrigues, J. J. (2011). "A survey on cross-layer solutions for wireless sensor networks". *Journal of Network and Computer Applications*, 34:523–534.

M.M.A., M. and A.A., K. (2011). "Dynamic indexing system for spatio-temporal queries in wireless sensor networks". In *12th IEEE International Conference on Mobile Data Management (MDM)*, volume 2, pages 35–37.

Mousavi, A., Duckham, M., Kotagiri, R., and Rajabifard, A. (2013). “Spatio-temporal event detection using probabilistic graphical models (pgms)”. In *IEEE Symposium on Computational Intelligence and Data Mining (CIDM)*, pages 81–88.

M.S., J., A., S., B., A., W., U., and M.F.A., R. (2011). “Energy efficient geographical and power based clustering algorithm for heterogeneous wireless sensor networks”. In *17th Asia-Pacific Conference on Communications (APCC)*, pages 826–831.

Mudumbai, R., Barriac, G., and Madhow, U. (2007). “On the feasibility of distributed beamforming in wireless networks”. *IEEE Trans. Wireless Communications*, 6(5):1754–1763.

Munir, A., Antoon, J., and Gordon-Ross, A. (2015). “ Analysis of fault detection and fault tolerance in wireless sensor networks”. *Transactions on Embedded Computing Systems (TECS)*, 3(3).

Nachabe, L., Genet, M. G., and Hassan, B. E. (2015). “ Unified data model for wireless sensor network”. *IEEE sensors Journal*, 15(7):3657–3667.

Naymat, G. A., Chawla, S., and Gudmundsson, J. (2007). “Dimensionality reduction for long duration and complex spatio temporal queries”. In *Proceedings of the 2007 ACM Symposium on Applied Computing (SAC), Seol, Korea, March 11-15*, pages 393–397.

Niu, R. and Varshney, P. K. (2005). “Distributed detection and fusion in a large wireless sensor network of random size”. *EURASIP Journal on Wireless Communications and Networking,*, 4:462–472.

Niu, R., Varshney, P. K., Moore, M. H., and Klamer, D. (June-July 2004). “ Decision fusion in a wireless sensor network with a large number of sensors”. In *Proc. 7th IEEE International Conference on InformationFusion (ICIF 04), Stockholm, Sweden.*

Novak, V., Perfilieva, I., and Mockor, J. (1999). “Mathematical principles of fuzzy logic”. *The International Series in Engineering and Computer Science, Kluwer Academic Publishers, Norwell, MA.*

Oka, A. and Lampe, L. (2008). “Energy efficient distributed filtering with wireless sensor networks”. *IEEE TRANSACTIONS ON SIGNAL PROCESSING*, 5(5):2062–2075.

Ozdemir, O., Niu, R., and Varshney, P. K. (2009). “Tracking in wireless sensor networks using particle filtering: Physical layer considerations”. *IEEE Transactions on Signal Processing*, 57(5):1987–1999.

P, S. and Z, V. (2009). “Energy-efficient circular sector sensing coverage model for wireless sensor networks”. In *Third International Conference on Sensor Technologies and Applications*, pages 229–233.

Papadimitriou, I. and Georgiadis, L. (2005). “Maximum lifetime routing to mobile sink in wireless sensor networks”. In *Proceedings of the International Conference on Software, Telecommunications and Computer Networks, SoftCOM 2005, Split, Croatia.*

Perillo, M., Cheng, Z., and Heinzelman, W. (2005). “An analysis of strategies for mitigating the sensor network hot spot problem”. *The Second Annual International Conference on Mobile and Ubiquitous Systems: Networking and Services*, pages 474–478.

Rachuri, K., Franklin, A. A., and Murthy, C. S. R. (2008). “Energy efficient searching in delay-tolerant wireless sensor networks”. *ACM, Heter SANET 08,China*, pages 37–44.

Raghuvanshi, A. S., Tiwari, S., Tripathi, R., and Kishor, N. (2010). “Optimal number of cluster in wireless sensor networks: An fcm approach”. In *International Conference on Computer & Communication Technology (ICCCT10)*, pages 817–823.

Rago, C., K.Willett, P., and Shalom, Y. B. (1996). “Censoring sensors: a low-communication-rate scheme for distributed detection”. *IEEE Transactions on Aerospace and Electronic Systems*, 32(2):554–568.

Rokne, J. G. (2001). *Interval Arithmetic and Interval Analysis: An Introduction of Studies in Fuzziness and Soft Computing*, volume 70. Springer.

S, K., A, I., and C, R. (2006). “An address-light, integrated mac and routing protocol for wireless sensor networks”. *IEEE/ACM Transactions on Networking*, 14(4):793–806.

S, N. and S.S, K. (2013). “Data aggregation in wireless sensor network using node clustering algorithms a comparative study”. In *IEEE Conference on Information & Communication Technologies (ICT)*, pages 508–513.

Schreiber, F. A. (2007). “Automatic generation of sensor queries in a wsn for environmental monitoring”. *Proceedings of ISCRAM’ 2007*, pages 245–254.

Sha, K. and Shi, W. (2005). “Modeling the lifetime of wireless sensor networks”. *Sensor Letters*, 3(1):1–10.

Shah, R. C., Roy, S., Jain, S., and Brunette, W. (2003). “Data MULEs: Modeling a three-tier architecture for sparse sensor networks”. In *Proceedings of the First IEEE International Workshop on Sensor Network Protocols and Applications*, pages 30–41.

singh, R. P. and sapre, S. D. (2009). Analog and digital communication systems. *The Mc Graw-Hill companies*, 2:156–160.

S.Kar, Hao, C., and Varshney, P. (2012). “Optimal identical binary quantizer design for distributed estimation”. *IEEE Trans. on Signal Processing*, 60(7):3896–3901.

Stojmenovic, I., Nayak, A., and Kuruvila, J. (March,2005). “ Design guidelines for routing protocols in ad hoc and sensor networks with a realistic physical layer”. *IEEE Communications Magazine*, pages 101–106.

Sun, Q., Hu, F., Wu, Y., and Huang, X. (2014). “ Primate-inspired adaptive routing in intermittently connected mobile communication systems”. *Wireless Networks*, 20(7):1939–1954.

T., C. and A.H., Z. (2011). “Localized energy efficient routing for wireless sensor networks”. In *International Conference on Distributed Computing in Sensor Systems and Workshops (DCOSS)*, pages 1–7.

T, S. and M, K. (2009). “Energy-efficient power/rate control and scheduling in hybrid tdma/cdma wireless sensor networks”. *Computer Networks*, 53(9):1395–1408.

Takruri, M., Rajasegarar, S., Challa, S., Leckie, C., and Palaniswami, M. (2011). “Spatio-temporal modelling-based drift-aware wireless sensor networks”. *IET Wireless Sensor Systems*, 1(2):110–122.

Tan, L., Gong, Y., and Chen, G. (2008). “A balanced parallel clustering protocol for wireless sensor networks using k-means techniques”. In *IEEE International conference on sensor technologies and applications (SENSORCOMM 08)*, pages 300–305.

Tilak, S., Abu-Ghazaleh, N., and Heinzelman, W. (2002). “A taxonomy of wireless microsensor network model”. *ACM Mobile Computing and Communications Review*, 6(2):28–36.

V, S. and M, M. (2005). “Cross-layer design: a survey and the road ahead”. *IEEE Communication Magazine*, 43(12):112–119.

van Hoessel, L. and Havinga, P. (2004). “A lightweight medium access protocol (lmac) for wireless sensor networks: Reducing preamble transmissions and transceiver state switches”. In *Postbus 217, NL-7500 AE Enschede. THE NETHERLANDS.*

Varshney, P. K. (1997). "Distributed detection and data fusion". *Communications Engineering, Networks,Springer, New York, NY, USA,.*

Vijay, G., Bdira, E., and Ibnkhala, M. (2010). "Cognitive approaches in wireless sensor networks: A survey". In *25th Biennial Symposium on Communications*, pages 177–180.

Viswanathan, R. and Varshney, P. (1997). "Distributed detection with multiple sensors: Part i-fundamentals". In *Proc. IEEE*, volume 85, pages 54–63.

Wang, C. F., Shih, J. D., Pan, B. H., and Wu, T. Y. (2014). " A network lifetime enhancement method for sink relocation and its analysis in wireless sensor networks". *IEEE Sensor Journal*, 14(6):1932–1943.

Wang, H., Yip, L., Yao, K., and Estrin, D. (2004). "Lower bounds of localization uncertainty in sensor networks". In *Proceedings of IEEE International Conference on Acoustics,Speech, and Signal Processing (ICASSP'04)*, volume 3, pages 917–920.

Wang, L. C., Wang, C. W., and Liu, C. M. (2009). "Optimal number of clusters in dense wireless sensor networks: A cross-layer approach". *IEEE Transactions on Vehicular Technology*, 58(2):966 –976.

Wang, Y. and Fang, F. L. F. (2010). "Poisson versus gaussian distribution for object tracking in wireless sensor networks". In *IEEE Proceeding*.

Wang, Y.-H., Chuang, C. C., Hsu, C.-P., and Chung, C. M. (2003). "Ad hoc on-demand routing protocol setup with backup routes". In *IEEE Proc.*, pages 137–141.

Wang, Z., Basagni, S., Melachrinoudis, E., and Petrioli, C. (2005). "Exploiting sink mobility for maximizing sensor networks lifetime". In *Proceedings of the 38th Hawaii International Conference on System Sciences, Big Island Hawaii.*

Ward, J. R. and Younis, M. (2012). "A physical layer metric for measuring base station anonymity in wireless sensor networks". In *First IEEE International Workshop on security and Forensics in Communication Systems*, pages 6689–6693.

Wolsey, L. A. and Nemhauser, G. L. (1988). *Integer and Combinatorial Optimization.* Wiley.

Xiao, J. and Luo, Z. Q. (2005). ''Decentralized estimation in an inhomogeneous sensing environment". *IEEE Trans. on Information Theory*, 51(10):3564–3575.

Xu, X., Zheng, B., and Yan, J. (2011). “A sparse reconstruction algorithm with hierarchical bayesian analysis for wideband spectrum detection”. In *IEEE international conference on Wireless Communications and Signal Processing (WCSP)*, pages 1–5.

Y., C. J., S., K. H., I, B., and H, K. W. (2005). “Celll based energy density aware routing: a new protocol for improving the lifetime of wireless sensor networks”. *Computer Communications*, 28(11):1293–1302.

Yang, F., Wang, K., , and Huang, Y. (2015). “ Energy-neutral communication protocol for very low power microbial fuel cell based wireless sensor network”. *IEEE Transactions on Sensors*, 5(4):2306–2315.

Ye, F., Luo, H., Cheng, J., Lu, S., and Zhang, L. (2002a). “A two-tier data dissemination model for large-scale wireless sensor networks”. In *Proceedings of 8th Annual international Conference on Mobile Computing and Networking, MobiCom02, Atlanta, Georgia, USA*, pages 148–159.

Ye, W., Heidemann, J., and Estrin, D. (2002b). “An energy-efficient mac protocol for wireless sensor networks”. In *In Proceedings of the IEEE INFOCOM.*

Ye, W., Heidemann, J., and Estrin, D. (2003). “Medium access control with coordinated adaptive sleeping for wireless sensor networks”. *IEEE Transactions on Networking.*

Yoon, S. and Shahabi, C. (2007). “The clustered aggregation (CAG) technique leveraging spatial and temporal correlations in wireless sensor networks”. *Transactions on Sensor Networks (TOSN)*, 3(1)(3):1–39.

Younis, M., Bangad, M., and Akkaya, K. (2003). “Base-station repositioning for optimized performance of sensor networks”. In *IEEE 58thVehicular Technology Conference(VTC 2003-Fall)*, volume 5, pages 2956–2960.

Yu, C. and Varshney, P. K. (1998a). “Bit allocation for discrete signal detection”. *IEEE Transactions on Communications*, 46(2):173–175.

Yu, C. T. and Varshney, P. K. (1998b). “Paradigm for distributed detection under communication constraints”. *Optical Engineering*, 37(2):417–426.

Yu, F., Lee, E., Park, S., and Kim, S. (2010). “A simple location propagation scheme for mobile sink in wireless sensor networks”. In *IEEE Communications Letters*, volume 14, pages 321–323.

Yu, Y. (2010). “HIBOR: An efficient approach to sequence pattern query processing in wireless sensor networks”. *IET International Conference on Wireless Sensor Network*, pages 161–166.

Yuan, H., Ma, H., and Liao, H. (2006). “Coordination mechanism in wireless sensor and actor networks”. In *IEEE Proceedings of the First International Multi-Symposiums on Computer and Computational Sciences (IMSCCS'06).*

Zimmermann, H. J. (2001). *FUZZY SET THEORY AND ITS APPLICATION*. Springer science and Business Media, 4 edition.

Publications based on the thesis

International Journals

1. **Pramod Kumar**, Ashvini Chaturvedi, "Spatial-Temporal Aspects Integrated Probabilistic Intervals Models of Query Generation and Sink Attributes for Energy Efficient WSN", Wireless Personal Communications [Under Review, Paper ID: WIRE-D-15-02240], **Springer Publisher**.

2. **Pramod Kumar**, Ashvini Chaturvedi, "Fuzzy-Interval based Probabilistic Query Generation Models and Fusion Strategy for Energy Efficient Wireless Sensor Networks",Computer Communications[Under Review, Paper ID: COMCOM-D-16-00570], **Elsevier Publisher**.

3. **Pramod Kumar**, Ashvini Chaturvedi, "Spatio-temporal Probabilistic Query Generation Model and Sink Attributes for Energy Efficient Wireless Sensor Networks", **IET Networks**, [DOI: 10.1049/iet-net.2016.0014], Online ISSN 2047-4962,Vol. 5, Issue 6,pp. 170-177,September 2016.

4. **Pramod Kumar**, Ashvini Chaturvedi, "Probabilistic Query Generation and Fuzzy C-means Clustering for Energy Efficient operation in a Wireless Sensor Network", International Journal of Communication Systems. [DOI: 10.1002/dac.3112], Vol. 29, Issue - 8, pp. 1439-1450, 2016, **Wiley Publisher**.

5. **Pramod Kumar**, Ashvini Chaturvedi, "Sink Attributes Analysis for Energy Efficient Operations of Wireless Sensor Networks under Randomly Varying Temporal and Spatial aspects of Query Generation", International Journal of Electronics and Communications,Vol 69, issue 7,pp.1058-1069 2015, **Elsevier Publisher**.

6. **Pramod Kumar**, Ashvini Chaturvedi, "Performance Measures of Fuzzy C-means Algorithm in Wireless Sensor Networks", International Journal of Computer Aided Engineering and Technology (IJCAET),Vol.9,No. 1,pp. 84-101,2017, **Inder Science Publisher**, UK.

7. **Pramod Kumar**, Ashvini Chaturvedi, "Location Movement and Multiplicity Attributes of Sink in Query based Wireless Sensor Networks", International Journal of Computer Aided Engineering and Technology (IJCAET), (DOI: 10.1504/IJCAET.2015.072600), Vol. 7, No. 4, pp. 436-455, 2015, **Inder Science Publisher**, UK.

International Conferences

1. **Pramod Kumar**, Ashvini Chaturvedi "Performance Analysis of Energy Aware Routing Protocol for Wireless Sensor Networks", IEEE International Conference on Devices Circuits and Systems (ICDCS-14), March 6-8, 2014, India.

2. **Pramod Kumar**, Ashvini Chaturvedi "Life Time Enhancement of Wireless Sensor Network Using Fuzzy C-Means Clustering Algorithm" , IEEE International Conference on Electronics and communication System (ICECS-14), February.13-14,2014 Coimbatore, India.

3. **Pramod Kumar**, Ashvini Chaturvedi "An Energy Efficient Algorithm to avoid Hot Spot effects in Wireless Sensor Networks", IEEE International Conference on Signal Processing ,Image Processing and Pattern Recognition (ICSIPR13), pp 117-121, February 7-8, 2013,India.

4. **Pramod Kumar**, Ashivini Chaturvedi & M.Kulkarni "Geographical Location Based Hierarchical Routing Strategy for Wireless Sensor Networks", IEEE International Conference on Devices Circuits and Systems (ICDCS-12),pp 9-14,March 15-16,2012,India.

Brief Bio-data

PRAMOD KUMAR
Research Scholar
Department of Electrical & Electronics Engineering
National Institute of Technology Karnataka, Surathkal
P.O. - Srinivasanagar
Mangalore-575025
E-mail:- 79.pramod@gmail.com
Phone: (+91)9901449757, (+91)7259003124

Addresses

Communication	Permanent
PRAMOD KUMAR FLAT NO.:006, PRETTY ARCHANA 1st Main, 5th Cross, V P Nagar, Manipal-576104, Karnataka	PRAMOD KUMAR Village-Purushottampur, PO- Bagadhi Via-Noan, District- Kaimur (Bhabua) Bihar-802132

Qualification

B.E. (Electronics & Communication Engineering) at Golden Valley Institute of Technology (VTU), K.G.F, Karnataka.
M.Tech. (Microwave Engineering) at Madhav Institute of Technology and Science (RGPV), Gwalior, M.P.