COGNITION IN LARGE SCALE INFORMATION-CENTRIC SENSOR NETWORKS: NOVEL DEPLOYMENT AND DATA DELIVERY SOLUTIONS

by

Gayathri Tilak Singh

A thesis submitted to the Department of Electrical and Computer Engineering

In conformity with the requirements for

the degree of Doctor of Philosophy

Queen's University

Kingston, Ontario, Canada

(April, 2015)

Copyright © Gayathri Tilak Singh, 2015

Abstract

Smart Cities, enabled by Wireless Sensor Networks (WSNs), have emerged as one of the most promising applications of the Internet of Things (IoT). These Smart City environments require that the underlying sensor network infrastructure be enriched with smart devices, so that the network can understand and respond to requests from multiple users with diverse information requirements. Now, the use of artificial intelligence has enabled some amount of user-requirement awareness in sensor networks. However, there is no architectural framework around how cognition is incorporated in the network, or where the smart decision making is implemented. In addition, WSN implementations are mostly address-centric, where users must specify the location from where data must be gathered. But this is counter-intuitive to how users would like to access information in a smart city environment, especially in applications such as smart parking, where the network needs to provide the location of free parking spots to the user. Moreover, managing the large IP address space becomes problematic as the network size expands to vast counts of sensor nodes.

In this thesis, we propose an architectural framework called Cognitive Information Centric Sensor Network (CICSN), to introduce cognition in WSNs. Cognitive nodes, capable of *knowledge representation*, *learning*, and *reasoning*, along with an information-centric approach to data delivery, are central to the idea of the CICSN. We propose a deployment strategy for cognitive nodes in the network such that connectivity of sensor nodes with the sink is maintained, and the number of cognitive nodes is minimized as well. Knowledge representation is done using attribute-value pairs. In addition, a Quality of Information (QoI) aware data delivery strategy, with Analytic Hierarchy Process (AHP) as the reasoning technique, is used to identify data delivery paths that dynamically adapt to changing network conditions and user requirements. Latency, reliability, and throughput are the attributes used to identify the QoI along the delivery path. Further, heuristic learning techniques are explored to improve the success rate of data delivery to the sink. Simulation results show that the proposed architecture significantly improves the QoI as well as the success rate of data delivered by the network.

Co-Authorship

Chapter 2 has been co-authored by Dr. Elyes Bdira and Dr. Mohamed Ibnkahla. It has been published in IEEE Sensors Journal in its Special Issue on Cognitive Sensor Networks.

Chapters 3, 4, and 5 have been co-authored by Dr. Fadi Al-Turjman. Chapter 3 has been submitted to Wiley's Wireless Communication and Mobile Computing Journal. Chapter 4 has been accepted for publication in Elsevier's Computer Communication Journal, in its Special Issue on Current and Future Architectures, Protocols, and Services for the Internet of Things. Chapter 5 has been submitted to IEEE Internet of Things Journal's Special Issue on Large-scale Internet of Things: Theory and Practice.

Four conference articles have also been written with Dr. Fadi Al-Turjman, Dr. Mervat Abu Elkheir, Dr. Abd ElHamid Taha, Dr. Elyes Bdira, and Dr. Mohamed Ibnkahla as co-authors. These conference articles are listed as references at the end of some of the thesis chapters. The first and subsequent drafts for all of these coauthored manuscripts have been prepared by Gayathri T Singh (formerly Gayathri Vijay). Details of the publications associated with each chapter are listed below.

Chapter 2

- **G. Vijay,** E. Ben Ali Bdira, M. Ibnkahla, "Cognition in Wireless Sensor Networks: A Perspective," Sensors Journal, IEEE, SI on Cognitive Sensor Networks, vol.11, no.3, pp.582-592, March 2011.
- **G. Vijay,** E. Bdira, and M. Ibnkahla, "Cognitive approaches in wireless sensor networks: a survey," In *Proc. of the IEEE Queen's Biennial Symposium on Communications (QBSC)*, Kingston, ON., Canada, 2010, pp. 177-180.

Chapter 3

- **G. Singh,** and F. M. Al-Turjman, “Cognitive Node Architecture and Deployment for Future Sensor Networks”, Wireless Communication and Mobile Computing, Feb. 6, 2015. [Submitted]
- **G. Singh,** Mervat Abu-Elkheir, F. M. Al-Turjman, and Abd-Elhamid Taha, “Towards prolonged lifetime for large-scale Information-Centric Sensor Networks,” *In Proc. of the IEEE Queen's Biennial Symposium on Communications (QBSC),* Kingston, ON., Canada, 2014, pp. 87-91.
- **G. Vijay,** and M. Ibnkahla, "CCAWSN: A cognitive communication architecture for wireless sensor networks," In *Proc. of the IEEE Queen’s Biennial Symposium on Communications (QBSC),* Kingston, ON., Canada, 2012, pp. 132-137.

Chapter 4

- **G. Singh,** and F. M. Al-Turjman, “A Data Delivery Framework for Cognitive Information-Centric Sensor Networks in Smart Outdoor Monitoring,” Computer Communications, Available online 20 January 2015, ISSN 0140-3664, http://dx.doi.org/10.1016/j.comcom.2015.01.002.
- **G. Singh,** and F. M. Al-Turjman, “Cognitive Routing for Information-Centric Sensor Networks in Smart Cities,” *In Proc. of the International Wireless Communications and Mobile Computing Conference (IWCMC),* Nicosia, Cyprus, 2014, pp. 1124 - 1129.

Chapter 5

G. Singh, and F. M. Al-Turjman, “Learning Data Delivery Paths in QoI-aware Information Centric Sensor Networks”, IEEE Internet of Things Journal, Special Issue on Large-scale Internet of Things: Theory and Practice, April 2015. [Submitted]

To Shri. Bhagawan Singh, Smt. Shivarani Bai, and Param.

“You cannot connect the dots looking forward; you can only connect them looking backwards. So you have to trust that the dots will somehow connect in your future.”

— *Steve Jobs*

“Nothing ever goes away until it teaches us what we need to know.”

— *Pema Chodron*

Acknowledgements

First of all, I am grateful to God Almighty for his divine guidance, help, and countless blessings in my life. My deepest gratitude to my parents, who have been my first teachers, for believing in me, and for their constant encouragement and support. They have taught me that knowledge is power, and patience is a virtue that can help one face any challenge in life. To my husband Vijay, who has been with me through all the highs and lows of graduate life, for encouraging me throughout my PhD program; thank you for always being there for me. To my brother Darshan, for all our pep talks that have helped me keep my sense of humor, and reminded me of my abilities; you have been a constant source of motivation and unconditional support. To my little prince Param, whose unconditional love and funny ways have always brought joy to my heart and cheered up the gloomiest days; you have changed my life for the better! To my in-laws for supporting my Doctoral education; to Shri Harihar Iyer and Smt. Radha Iyer, for their blessings and teachings that have made a positive impact on my life. To my beloved Aunt Padmakshi, Aunt Krishnakumari, and Uncle Raghuveer, whom I lost during this doctoral journey, I know your love and blessings are always with me. To my grandparents, relatives, and friends, thank you for wishing me well and always believing in me.

My sincere gratitude to Dr. Kim McAuley for her kindness, understanding, encouragement, guidance, and support, without which this thesis would not have been completed. I would like to thank Dr. Hossam Hassanein, whose foresight and guidance laid the foundation to the ideas developed in this thesis. His encouragement and support have been invaluable to me. I would also like to thank Dr. Mohamed Ibnkahla for giving me the opportunity to pursue research on Wireless Sensor Networks at his WiSiP lab, and for all the opportunities for collaboration and publication. My heartfelt thanks to Dr. Woodhouse, Dr. Brian Surgenor, and Dr. Michael Greenspan for their support.

To Debie Fraser, for her warm, understanding and helpful ways that have been very valuable during my graduate life at ECE. Thank you Dr. Elyes Bdira and Dr. Abd-Elhamid Taha for the interesting discussions and valuable feedback. To all the members of WiSiP lab, thank you for all your feedback on my work and presentations.

Dr. Fadi Al-Turjman, your willingness to help and make a positive difference is inspiring. In your various roles as my colleague, mentor and technical advisor, I have always found you supportive and encouraging. I cannot thank you enough for the valuable discussions, reviews and guidance that have led to the completion of this thesis. Most of all, I would like to thank you for stepping in to help me when I needed it the most.

My heartfelt gratitude to Victoria Millious, for being my pillar of strength during difficult times, for her trust in me, and for all the encouragement she has given me.

To my friends in Kingston, Bushra, Vidisha, Nikhil, Shilpa, Smrithi, the Khan, Lokhande, and Kamat families, you have made Kingston feel like home away from home.

Table of Contents

List of Figures

List of Tables

List of Acronyms

AF	Application Framework
AHP	Analytic Hierarchy Process
AHPDD	Analytic Hierarchy Process based Data Delivery
AI	Artificial Intelligence
AM	Adaptive Modulation
AMS	Adaptive Modulation combined with Sleep scheduling
ANN	Artificial Neural Network
AODV	Ad hoc On-demand Distance Vector
APL	Application Layer
API	Application Program Interface
APS	Application Support Sublayer
AT	Average Throughput
BER	Bit Error Rate
BS	Base Station
CHAL	Cumulative Heuristic Accelerated Learning
CICSN	Cognitive Information Centric Sensor Network
CN	Cognitive Node (Except in Ch. 2 where CN refers to Cognitive Network)
CPN	Cognitive Packet Network
CR	Cognitive Radio
CSL	Cognitive Specification Language
CSMA	Carrie Sense Multiple Access
CSN	Cognitive Sensor Network
DB	Development Board

DCSN	Data Centric Sensor Network
DD	Directed Diffusion
DE	Data Entity
DRIP	Directional Reception Incremental Protocol
DSN	Destination Sequence Number
DVR	Distance Vector Routing
EB	Evaluation Board
EM	Evaluation Module
FIFO	First In First Out
GCN	Global Cognitive Node
GRISP	Gateway Relocation algorithm for Improved Safety and Performance
HRBDD	Higher Remaining Battery based Data Delivery
IA	Information Attribute
ICN	Information Centric Network
ICSN	Information Centric Sensor Network
IEEE	Institute of Electrical and Electronics Engineers
IoT	Internet of Things
IP	Internet protocol
ISM	Industrial, Scientific & Medical
IT	Instantaneous Throughput
ITU	International Telecommunication Union
KB	Knowledge Base
KP	Knowledge Plane
LCN	Local Cognitive Node

*LDDA**	Learning Data Delivery A*
*LRTA**	Learning Real-Time A*
LRTS	Learning in Real-Time Search
LRU	Least Recently Used
LVF	Least Valuable First
MAC	Medium Access Control layer
MDD	Multipath Data delivery
ME	Management Entity
MHR	MAC Header
MLDE	MAC Layer Data Entity
MLME	MAC Layer Management Entity
M-QAM	M is the parameter representing the modulation level in Quadrature Amplitude Modulation
NGN	Next Generation Network
NLDE	Network Layer Data Entity
NLME	Network Layer Management Entity
NP	Non-deterministic Polynomial-time
NR	Node Reliability
OADA	Observe Analyze Decide Act
OL	Observed Latency
OODA	Observe Orient Decide Act
OSI	Open Systems Interconnection
PHY	Physical layer
PLME	Physical Layer Management Entity
QoE	Quality of Experience

QoI	Quality of Information
QoS	Quality of Service
RF	Radio Frequency
RFID	Radio Frequency Identification
RL	Reinforcement Learning
RN	Relay Node
RQ	Request Classifier
RSSI	Received Signal Strength Indicator
SAN	Software Adaptable Network Layer
SAP	Service Access Point
SI	Sensed Information
SN	Sensor Node
SNR	Signal to Noise Ratio
SOM	Smart Outdoor Monitoring
SP	Smart Packets
SSP	Security Service Provider
SSR	Self-Selective Routing
TCP-IP	Transmission Control Protocol-Internet Protocol
TI	Texas Instruments
TTL	Time To Live
VoI	Value of Information
VoSI	Value of Sensed Information
WRAN	Wireless Regional Area Network
WSN	Wireless Sensor Network
ZDO	ZigBee Device Object

Chapter 1

General Introduction

Wireless Sensor Networks (WSN) have evolved from supporting application-specific deployments such as habitat monitoring, health care, and retail supply chains, to enabling multi-user platforms that simultaneously support multiple applications operating in the large-scale Internet of Things (IoT) environment. In the IoT paradigm, every object in the physical world is uniquely identifiable and traceable, and can communicate with other machines and objects over the internet [1]. The use of Radio Frequency Identification (RFID) devices makes the objects uniquely identifiable; and sensors and actuators embedded into these objects helps to capture information from their physical environment and bring it into the digital realm over the internet infrastructure. This interconnection between the physical and virtual realms paves the way for many new applications and services that benefit from such information sharing. Smart Cities are an example of such an IoT-enabled application platform, in which the end-user benefit from being able to access information from the physical world via the internet. These Smart City applications are expected to support multiple-users accessing information from a city-wide deployment of sensors that capture data from multiple applications. The users may include: (i) individual users trying to access information for their personal use, (ii) private data centers or public enterprises accessing information periodically to build an information base, and (iii) government agencies monitoring information to issue public alerts in case of emergencies. These users may want to access data from any of the following applications supported on the smart city application platform: (i) on-demand information about availability of free parking spots from the smart parking management application, (ii) periodic information from the air pollution and environment monitoring application, concerning air pollution levels, temperature, humidity, UV index, etc., and (iii) the city management may want to receive alerts in case of emergency situations such as major road accidents, which can in turn be used to manage traffic. Such a smart city application environment with multiple users and multiple applications on the same platform is illustrated in Figure 1-1. The central cloud in this figure represents the underlying sensor network that contains smart nodes, and supports the various applications such as air pollution monitoring, traffic monitoring, and emergency response on the same platform. The smart nodes communicate with the applications, the sensor network's nodes, and with each other, to gather requested information and transmit it to the central data sink. The sink in turn communicates with a gateway, which acts as the interface to the interconnecting infrastructure through which all the users interact with the network.

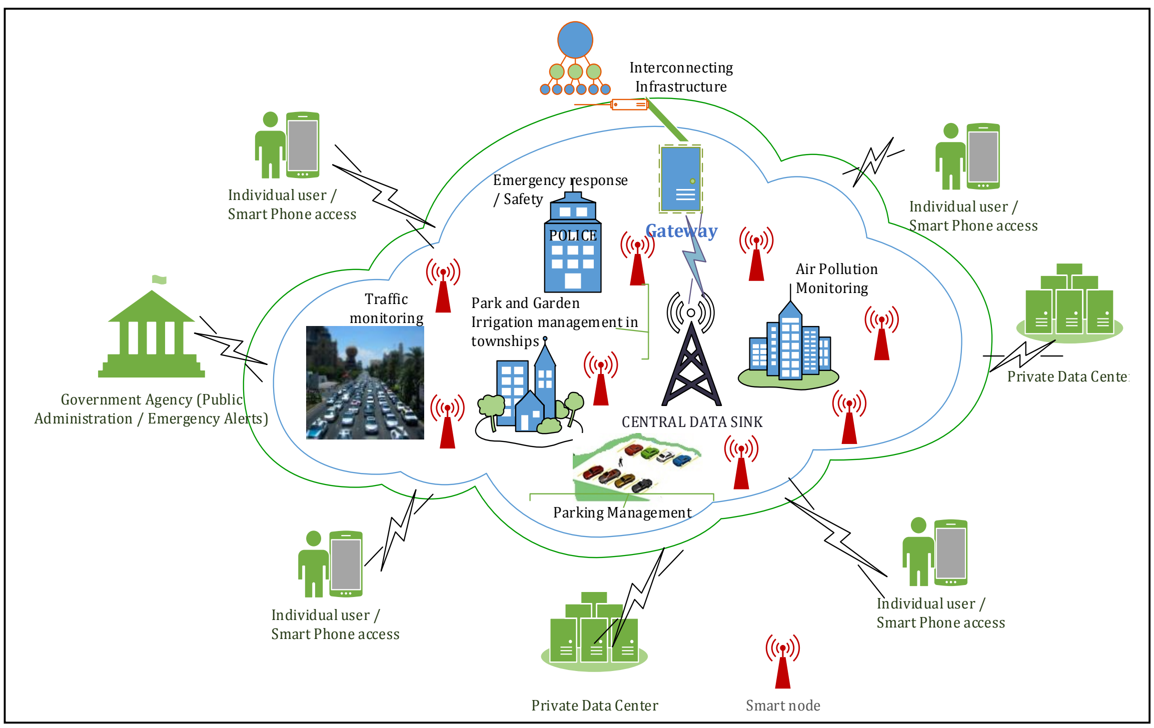


Figure 1-1: A Smart City application platform showing smart nodes, multiple-users, and multiple-application support.

Although deployments such as the SmartSantander facility [2], and the European smart cities [3] have a similar vision as depicted in Figure 1-1, they lack in their ability to support these multiple applications running simultaneously, while allowing multi-user access. WSNs are a suitable starting point for implementing the underlying network in these smart environments because of their inherent ability to observe information from the environment and communicate it wirelessly to end-users. However, their ability to deliver data according to the end-users requirements, in terms of attributes such as latency, reliability, relevance and timeliness, has not received sufficient research attention. This is primarily because the research on enabling WSNs to support smart environment applications of the IoT paradigm is still in the experimental stage. The complexity involved in handling heterogeneous traffic flows in the underlying sensor network results from multiple requests with diverse user requirements. To address this complexity, smart environments consider the deployment of smart nodes/devices that constantly work towards dynamically adapting and responding to requests coming in from multiple users [5, 6]. However, what is not well defined is, how these smart devices are implemented, and what specific features make these devices smart. Cook and Das formally defined a smart environment as "an environment that is able to acquire and apply knowledge about the environment and its inhabitants in order to improve their experience in that environment" [7]. From this definition, we identify two important features for smart devices. First, is the ability to observe the environment being sensed, and analyze the changes taking place in the network in terms of resources consumed, while responding to user requests. Second, is the ability to apply the knowledge acquired to decide on a course of action that will help the network adapt to end-user requirements when there are multiple users trying to access the network at the same time. This ability to observe, analyze, decide and act is best described by Boyd's loop for decision making [8]. So, we call these smart nodes *cognitive nodes* in the WSN. They must have elements of cognition that will enable them to gather data from the network, combine it with the knowledge gained from the observations made to convert it to useful information, and make informed decisions that will benefit the network in the short term, as well as the long term. Now, there has been work in the direction of building a cognitive sensor network, where researchers have made use of artificial neural networks, genetic algorithms, game theory and even software agents to implement distributed and intelligent decision making in sensor networks [9-12]. However, there does not exist a single framework that can be used to implement cognition in sensor networks in a way that is domain and application independent.

Another aspect of smart city environments is that the user is more interested in the information observed from the network, rather than the identity of the node that published the information. WSNs were originally designed to function as information gathering networks, instead of serving as point-to-point communication networks. However, the energy limited nature of the sensor network's nodes, the need to be compatible with the IP based internet, and ZigBee compliance, led to WSNs becoming address-centric. But a shift back to the information-centric approach is required for smart city applications, especially in large scale deployments, because of the large number of nodes used.

Although IPV6 supports a vast address space, there is a need for a shift from the address-centric to information-centric approach. This is because the IoT application environment needs to support billions of sensing devices, and serve multiple users who are more interested in the information from these sensors rather than their individual node IP address. For example, in applications such as street parking management in a busy city, users must be able to access information about free parking spots in a region without having to check each spot with its address, to know whether it is free or not. This requires decoupling the information from its publisher's identity, so that routing can be executed directly on a semantic name instead of maintaining an end-to-end connection between the client and the publisher [4].

Thus, cognition and an information centric approach to data gathering and delivery form the building blocks of the architecture we propose for future sensor networks in this thesis. In the following sections, we delve into the challenges involved in implementing such a Cognitive Information Centric Sensor Network (CICSN) for large scale deployment of sensor networks in the IoT application paradigm.

1.1 Motivation and Challenges

Smart City environments have attracted a lot of research attention from the industry and academia alike [5, 6, 13, 14, 15-18]. This is because of the practicality of these applications, and their ability to make a huge impact on how humans interact with their environment to get the information they require. Examples in the smart city application platform include smart parking for vehicles, irrigation and garden management to maintain the parks and gardens in the city, pollution monitoring to issue alerts to the public when required, and environmental monitoring for weather updates [19]. What makes these applications both interesting and challenging to design, is the constant evolution of the technologies that form the building blocks of a smart environment, which includes WSNs, smart devices, and the Internet. WSNs are being integrated with RFID devices to form an integrated RFID-Sensor Network architecture that can supersede the capabilities of both the individual technologies [20-23]. Internet is moving towards an IP-free, Information Centric Networking paradigm that is based on information access through named data objects, and caching to improve information availability and accessibility [24]. Furthermore, the concept of cognition in wireless networks has moved from cognitive radio at the physical layer to cognition at the network and application layers. This enables the network to dynamically adapt to varying network conditions and optimize the end-to-end performance objectives through learning and reasoning. Cognition has been applied to WSNs too, to achieve application specific goals [11, 12]. Although each of these advances is towards improving the performance of the individual technologies, putting them together in a way that improves the experience of the end-user in smart environments is a challenge.

WSNs have traditionally been designed for application-specific purposes, which usually include monitoring, tracking and emergency alert systems. The user requirements are assumed to be fixed and integrated into these application specific designs. However, for supporting IoT applications in smart environments, WSNs should be able to support multiple user access, and deliver data based on the possibly diverse user requirements. Although artificial intelligent techniques have been used in WSNs for distributed decision making at the network level, the design goals have mostly been application-specific [9, 10, 12]. The ability to support multiple users with varied requirements is a feature that is new to WSNs used in IoT applications. The challenge lies in interpreting the user requirements and mapping them to features that can be understood and supported by the network. For example, if a user wanted to request for information about the availability of parking space in a specific region, it would be expected that this information is delivered within a specified duration (minimum latency) and with a certain amount of reliability. Although Quality of Service (QoS) metrics have been used to objectively evaluate the performance of the operational aspects of a sensor network [26], and Quality of Experience (QoE) metrics are used to subjectively evaluate the ability of the network to match the user expectations, and how satisfied end-users are with the information they receive from a network

[27], there are no metrics that are associated with the characteristics of the sensory information made available to the end-user.

In addition to being aware of changing user requirements and supporting multiple users, the smart sensor network should also be able to function as a common platform that supports more than one application. For example, a smart sensor network deployed in the city should be able to provide periodic weather updates, let users access data on-demand regarding parking space availability and also issue alerts when there is vehicular traffic congestion in an area, all using the same network resources. This results in heterogeneous traffic flows in the network, which makes it challenging to identify data delivery paths that dynamically adapt to changing network conditions and user requirements.

In order to be responsive to multiple traffic types, the network nodes must constantly observe the changing network conditions (topology changes due to changes in channel conditions and node deaths, and resource availability), and keep track of user requirements. In addition, they must also be able to make intuitive decisions based on learning from past behavior and responses, and predict a course of action that best serves the user. These functions map closely to Boyd's observe, Orient, decide and act feedback loop [8]. We have mapped the "Orient" function to "Analyze" function, to make it easily understandable in the context of environment/situation analysis for sensor networks, thus calling it the Observe-Analyze-Decide-Act (OADA) loop as depicted in Fig. 1-2. However, cognition has not been implemented in WSNs in a way that can implement every aspect of this feedback loop. Although learning and reasoning have been identified as elements of cognition in cognitive wireless networks [25], the challenge lies in identifying whether these techniques are feasible for incorporation in sensor networks in terms of

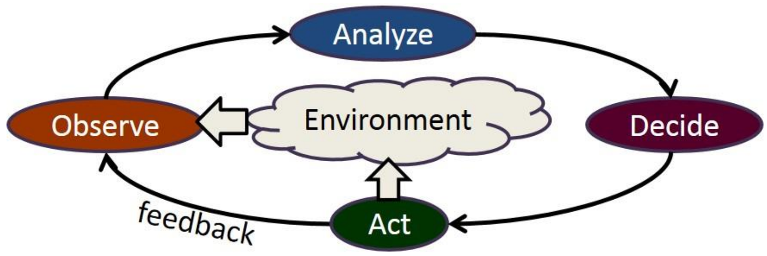


Figure 1-2: The OADA feedback loop.

their computation and time complexity, communication overhead, and resource overhead (in terms of memory and processing power). The decisions made to adapt to the dynamic network behavior must come quickly enough.

In other words, if either of the learning or decision making mechanism is slow, then the network state may change even before the system responds to a previous state, thus rendering the current response superfluous, or even incongruous. Thus a timely response is crucial to making the cognitive techniques work in favor of improving network performance.

Another aspect to consider is, where, or in which node cognition should be implemented: sensor, router or coordinator node. Sensor nodes can easily be ruled out, as they are already resource constrained. They must be least loaded with functions other than sensing and transmitting their observations. That leaves the router and/or co-ordinator nodes as choices that can be considered to implement cognitive functionality. In case both the nodes implement cognitive functionality, the router and co-ordinator will differ in the goals that cognition achieves at these nodes.

The challenge is to develop a generic framework that can be used to introduce cognition in WSNs, such that it can provide the intelligent infrastructure support required to implement Smart Cities. In developing such a framework, it would be useful to consider the user's perspective about the most convenient way to request and gather information from the environment. Users would most often not know where the requested information might be located in a smart environment. Since WSNs are primarily information gathering networks, a data centric approach would be more useful than an IP based address centric approach for request dissemination and data gathering. In addition, it would be more convenient for users to issues semantic requests that convey information about their requirements rather than pointing to an address where the information is located. Information Centric Networks (ICNs) provide a good framework for information centric data delivery [24]. However, the concept of ICN has been developed for the wired internet, which is a static and resource rich environment. Moreover, the use of cognition in ICNs is still in its nascence. Hence the ICN techniques also need to be enhanced to accommodate cognition, and extended to be applied to WSNs in smart environments.

In summary, the objective of this thesis is to develop an architectural framework to introduce cognition in WSNs that can be applied to Smart City environments. It must be able to support multiple end-users, multiple applications, and be responsive to diverse traffic types. The user must be able to request for information based on information attributes, rather than IP addresses. The network must be able to dynamically identify data delivery paths as the network conditions change, and make intelligent decisions that increase the probability of successful data delivery to the sink. It must also be able to satisfying the end-user's requirements in terms of the value of attributes such as latency, reliability and throughput, associated with the delivered data. Table 1-I provides a summary of comparison of WSNs, ICNs and the cognitive information centric sensor networks (CICSNs) proposed in this work. This table is used to show how the proposed approach improves the capabilities of a WSN by making use of Cognition and an Information centric approach to make it useful in an IoT environment.

Table 1-I: Comparative Summary of WSN, ICN, and CICSN systems.

	WSNs	**ICNs**	**CICSNs**
Components	Sink or Base Station, Relay or Router nodes, and Sensor nodes	Publishers, Subscribers, and Intermediate nodes	Sink or Global Cognitive Node, Local Cognitive nodes, Relay Nodes, and Sensor Nodes.
Network Topology	Dynamic due to node deaths, node mobility or variations in the wireless channel	Static, wired network for the Internet	Dynamic network, same as WSN.
Node deployment	Sensor nodes are typically randomly distributed. Relay nodes may be deterministically or randomly deployed based on targeted application and area to be monitored.	ICN use cases impact their deployment plan. Deployment in vehicular networks could be very different from deployment in home networks.	Sensor nodes deployed uniformly randomly; relay and cognitive nodes deployed deterministically based on probability of successful data reception.
Addressing	Data or address centric	Named data objects; Information centric	Information centric
Network Scalability	Scalable. Can support large-scale deployments	Limited by scalability of name resolution.	Designed to support large-scale IoT deployments.
Multiple user access support	Typically designed for specific user or set of users	Expected to support multi-user access	Will support multi-user access with diverse requirements on information and its quality.
Multiple application support on single platform	No. Typically designed for single application scenario.	Application is always to provide some content created or published to end-user.	Yes. Single deployment will be able to support multiple applications.
Use of Cognition for data delivery	Artificial Intelligent techniques or software agents used to improve data delivery decisions, but very application specific.	Still in its nascence	Elements of cognition identified and implemented in Cognitive nodes (CNs). These CNs cognitively adapt data delivery decisions according to network conditions and user requirements.

Content caching	Caching typically used to store data before delivery to sink. Not exploited well in data centric WSNs to make data available closer to sink.	Caching is extensively used to make requested information available closer to the Subscriber. Useful as content once published doesn't change. Only its popularity may vary.	Will use content caching to make data available closer to sink. However cached data may become obsolete as newly sensed data becomes available. New cache replacement strategy proposed.
Metric to measure end-user satisfaction	Quality of service (QoS)	Quality of Experience (QoE)	Quality of Information (QoI)

1.2 Thesis contributions

In this research, we propose and implement the cognitive information centric sensor network architecture for data delivery in large scale monitoring applications in smart city environments, within the IoT framework. We assume a static WSN, within which cognitive nodes are introduced. These cognitive nodes combine the use of cognition and an information-centric approach, to make data delivery decisions that dynamically adapt to changing network conditions and user requirements. The information centric approach enables the functioning of the sensor network as an information gathering network, rather than working as a point-to-point communication network. Accordingly, following are the main contributions of this thesis:

1. We introduce intelligent behavior in the router nodes of a ZigBee based experimental sensor network setup in hardware. By observing the network's ability to identify new data delivery paths before existing links fail, we establish the ability of the network to support intelligent behavior, and identify routers and the sink as the ideal candidates where cognitive behavior can be introduced in the sensor network.

2. We identify knowledge representation, reasoning and learning/planning as the elements of cognition, and associate it with the Observe-Analyze-Decide-Act (OADA) feedback loop to define cognition. Specifically, we identify techniques to implement each of these elements of cognition into specialized nodes called cognitive nodes. We use two types of cognitive nodes in the sensor network - global cognitive nodes (GCNs) and local cognitive nodes (LCNs). GCNs are used at the user interface of the network to interpret the user request for the network, and LCNs are used within the network to help with cognitive data delivery to the GCN. Together, these cognitive nodes profile the network traffic based on application and end-user requirements to identify the most suitable data delivery paths.

3. We combine the use of cognition with an information-centric approach to propose an architectural framework called COGNICENSE which can potentially support multiple applications simultaneously. Named identification of sensed data using attribute-value pairs, and caching of the gathered data at the LCNs makes this framework similar to information-centric networks (ICNs), which have emerged as one of the promising technologies for the future internet. This framework offers the following diversity: (a) Support for multiple clients/end-users (b) Support for multiple applications, and (c) Responsive to multiple traffic types. Although we do not delve into the details of node mobility, this framework can also support sensor node and LCN mobility because of the data caching capabilities of the LCNs, and due to the information-centric nature of data gathering.

4. We propose a 2-dimensional grid deployment plan for static relay and cognitive nodes in the CICSN framework in which each sensor node is guaranteed to be connected with at least one other relay node or LCN at the time of deployment. Sensor nodes are assumed to be uniformly, randomly deployed, and sensor and relay nodes have fixed communication range

and transmit power. However, the LCNs are assumed to have variable transmit power and communication range. The deployment plan is based on probability of successful reception of data at a given distance for a given transmit power.

5. We make use of the Quality of Information (QoI) metric to measure the level of satisfaction experienced by the end-user for the information received from the network [28, 29]. Attributes of latency, reliability and throughput are used to evaluate the QoI delivered by the network to the user. Network traffic is profiled into one of three types (periodic, on-demand or intermittent, and emergency or low latency traffic) based on the relative priorities of the QoI attributes associated with each of these flows. Data delivery paths are cognitively identified using an Analytic Hierarchy Process (AHP), which we classify as the reasoning technique, for each traffic type by analyzing the values of the QoI attributes observed on each link during previous transmission rounds. The details of the QoI attributes, the model for identifying the values of each of the attributes, and how these values are analyzed by the LCNs for cognitive data delivery using AHP analysis are explained. In addition, we propose a new definition for the Network lifetime of the CICSN based on the ability of the network to deliver QoI-aware data to the end-user.
6. We propose two heuristic based learning techniques that can be used with the reasoning technique, to improve the cognitive capabilities at the local cognitive nodes of the ICSN. The heuristic functions either help to choose paths that deliver data with good QoI to the sink, or identify data delivery paths that are more resource aware and considerate towards the energy consumption of the network. The impact of using these learning techniques on the success rate of data delivered to the sink, and the network lifetime are studied.

1.3 Thesis outline

The rest of this thesis is organized as follows. In Chapter 2 we delve into the related work in the field of cognition in wireless sensor networks. In Chapter 3 we provide the details of a 2-dimensional grid-based deployment plan for relay and cognitive nodes in the CICSN architecture. Chapter 4 provides the details of the elements of cognition in the COGNICENSE framework, and elaborates on implementation details of the knowledge representation and reasoning techniques in the cognitive nodes. We also compare the performance of the cognitive data delivery strategy of the CICSN network with two other commonly used techniques in data centric sensor networks. In chapter 5, we provide the details of the learning strategies proposed in this work, which work with the reasoning technique to improve the average rate of successful data delivery to the network's sink. QoI of the data delivered to the sink, and energy consumed in the network during the data delivery process are the metrics that are considered in the design of the learning strategies. We conclude the thesis in Chapter 6 with our perspective on the directions for future work.

1.4 Definitions

We define some frequently used terms in this thesis in the following section to provide the context of use in this work, and to serve as a quick reference resource for the reader.

1.4.1 Cognition

Cognition refers to the act or process of acquiring knowledge through perception, reasoning and intuition. In wireless networks, the cognitive process involves the following steps: (a) analyzing information observed from the environment, and from feedback of past actions, and (b) dynamically adapting decisions to enable the network to respond to changing user requirements. This feedback loop is implemented using the Observe, Analyze, Decide and Act (OADA) feedback loop to make cognitive decisions. This is based on Boyd's observe, orient, decide and act loop for cognition.

1.4.2 Cognitive Sensor Network

It is a sensor network comprised of intelligent agents that implement elements of cognition to achieve distributed information processing and dynamic decision making. These agents provide the infrastructure that enable applications to evolve around the user, and the network to evolve around the applications. The intelligence of this network lies in its ability to cope with large-scale deployments and adaptability to changing user requirements.

1.4.3 Cognitive Nodes

Nodes that implement the elements of cognition, such as knowledge representation, learning and reasoning, are called cognitive nodes. These nodes enable distributed intelligence and dynamic decision making in the network. They observe the changes in network connectivity, and learn from the feedback of past actions, to dynamically identify network paths that ensure that data is successfully delivered to the sink. Cognitive nodes are the smart nodes in smart environments.

1.4.4 Data versus Information

Data is transformed to useful information when it is combined with knowledge. For example, a string of 13 digits may not convey much information about the object associated with it. However, if the 13 digits are associated with an attribute, the ISBN, then it conveys the information that the object being referred to, is a book. The attribute-value pair together constitute the knowledge that transforms the 13 digit data into useful information.

1.4.5 Information Centric Sensor Network (ICSN)

An ICSN incorporates the features of an Information Centric Network [19] into the WSN by making use of the ideas of named data identification and information caching. Data objects are identified by their names and attributes rather than IP addresses, and the gathered information is stored in caches to make them readily available at multiple locations in the network, other than their source of publication. Data is processed to convert it to useful information by combining it with knowledge acquired from observations made by the network. This information is brought into the network layer, and routing is executed directly on a semantic name to find the nearest data source. This saves energy and reduces delay at the resource-constrained sensor nodes, which would otherwise require the mediation of a domain name server to discover the IP address of the information producer, i.e. the sensor node [2].

1.4.6 Data Centric Sensor Network (DCSN)

Data centric sensor networks are one of the earliest forms of content based data gathering in sensor networks. Here, users are interested in the collective information gathered from multiple sensors about a physical phenomenon, and routing algorithms focus on node-to-node data propagation rather than node-to-node packet switching [35, 36]. The ideas of data aggregation and caching at intermediate nodes existed even for DCSNs, similar to ICSNs. However, it was sensor nodes that performed all these tasks, despite their severely resource constrained nature in terms of energy availability, communication bandwidth, and computation capabilities. Their scalability in large-scale deployments was restricted because there was no global knowledge of the network, although decision making was fully distributed. Moreover, the energy consumption bias problem, in which nodes closer to the sink were eventually drained out of energy faster than nodes lying far away from the sink, also limited the network's scalability. Eventually, the DCSNs got replaced by address-centric sensor networks in order to conserve energy at sensor nodes and off-load the data communication task from sensors to relay nodes. The ZigBee protocol enabled hop-over-hop communication of data through relay nodes, thus addressing the scalability and energy consumption problems by segregating the sensing and data communication tasks.

1.4.7 Quality of Information (QoI)

QoI is defined as the quality experienced or perceived by the user with respect to the usefulness of information received from the network [28, 29]. Raw sensed data is considered to be useful information, and fit for use for a particular application, if it satisfies the user's requirements. A collection of attributes such as timeliness, accuracy, relevance, and reliability are used to capture information about the QoI delivered by a network [32]. In this work, we use latency, reliability and throughput as the attributes that determine the QoI. We also describe the terms Quality of Service (QoS), and Quality of Experience (QoE) associated with a communication network to provide the reader with a better understanding of how these terms are related, but different from each other.

1.4.8 Quality of Service (QoS)

QoS is a measure of the service quality that a network offers its applications/users [31]. The International Telecommunication Union (ITU) defines QoS in ITU Rec. E.800 as follows [30]: *"The QoS is the collective effect of service performances which determine the degree of satisfaction of a user of the service."* Network QoS support in Wireless Sensor Networks can be measured using parameters such as latency, packet loss, bandwidth and throughput. In addition, parameters such as aggregation delay, fault tolerance, and optimum number of active sensors are also used to measure application-specific QoS support in WSNs [26].

1.4.9 Quality of Experience (QoE)

QoE is a subjective measure of the overall value of service provided, from the user's perspective. It is dependent on the user's expectations, nature of content, and choice of terminal device used

[33]. It ties together user perception, experience, and expectations, to application and network performance, typically expressed by the QoS parameters [34]. QoE is related to QoS, but differs from it, as QoS more objectively measures the service delivered by the vendor.

1.4.10 Smart Environment

In this work, we define a smart environment as a network environment that uses intelligent infrastructure to provide data as required by the user. The environment we consider in this work is a large-scale sensor network, whose services are enhanced by cognitive nodes. These cognitive nodes form the network's intelligent infrastructure, which adapt the network's decisions during information acquisition and data delivery in a way that improves the quality of user's interaction with the network.

1.4.11 Smart City

In this work, we define Smart City as a connected self-aware city environment that includes several disparate applications such as, traffic monitoring, smart parking, city environment monitoring, and emergency alerts during extreme weather or accidents, on a single platform. The focus is on how to make the WSN infrastructure smart, so that it can integrate the physical city infrastructure to provide better services to the users.

1.4.12 Network Lifetime for the CICSN

Network lifetime of the Cognitive ICSN (or CICSN) is defined as the time or number of transmission rounds beyond which the network can no longer deliver useful information to the end-user. This is reflected by the network's inability to find a data delivery path with satisfactory values for QoI attributes (latency, reliability and throughput), as determined by the end-user, or when there is insufficient energy in the network nodes to deliver such data to the sink for any of the application generated requests.

1.5 References

[1] D. Miorandi, S. Sicari, F. De Pellegrini, and I. Chlamtac, "Internet of things: Vision, applications and research challenges," Ad Hoc Networks, vol. 10, no. 7, pp. 1497-1516, September 2012.

[2] Anonymous, SmartSantander Project, [Online]. Available: http://www.smartsantander.eu/index.php/testbeds/item/132-santander-summary.

[3] Anonymous, European Smart Cities 3.0 (2014). [Online]. Available: http://www.smart-cities.eu/?cid=01&ver=3.

[4] N. T. Dinh, and Y. Kim. "Potential of information-centric wireless sensor and actor networking." In *Computing, Management and Telecommunications (ComManTel), 2013 International Conference on*, pp. 163-168. IEEE, 2013.

[5] E. Avilés-López, and J. A. García-Macías, "Mashing up the Internet of Things: A Framework for Smart environments", EURASIP Journal on Wireless Communications and Networking *2012,* 2012:79.

[6] N. L. Fantana, T. Riedel, J. Schlick, S Ferber, J. Hupp, S. Miles, F. Michahelles, and S. Svensson, "IoT Applications — Value Creation for Industry", In *Internet of Things-Converging Technologies for Smart Environments and Integrated Ecosystems,* Edited by O. Vermesan, and P. Friess, pp-153-206, River Publishers, 2013.

[7] D.J. Cook, and S.K. Das, "How smart are our environments? An updated look at the state of the art", Journal of Pervasive and Mobile Computing, vol. 3, no. 2, pp-53–73, 2007.

[8] J. Boyd, "A discourse on winning and losing: Patterns of conflict", 1986.

[9] L. Reznik, and G. Von Pless, "Neural networks for cognitive sensor networks", *IEEE Int. Joint Conf. on Neural Network.,* IJCNN 2008, pp. 1235 - 1241, June 2008.

[10] P. Boonma, and J. Suzuki, "Exploring self-star properties in cognitive sensor networking**",** *Proc. of IEEE/SCS Int. Symp. on Performance Evaluation of Comput. and Telecommun. Syst.* (SPECTS), Edinburgh, UK, June 2008, pp.36-43.

[11] W. Youssef, and M. Younis, "A cognitive scheme for gateway protection in wireless sensor network", *Appl. Intell.J.*, vol. 29, no. 3, pp 216-227, 2008.

[12] K. Shenai and S. Mukhopadhyay , "Cognitive sensor networks", *IEEE Twenty Sixth Int. Conf. on Microelectronics* (MIEL), pp.315-320, May 2008.

[13] IBM, "A Smarter Planet", [Online]. Available: http://www.ibm.com/smarterplanet/ca/en/overview/ideas/

[14] Libelium, "50 Sensor Applications for a Smarter World", [Online]. Available: http://www.libelium.com/top_50_iot_sensor_applications_ranking/

[15] Escher Group, "Five ICT essentials for Smart Cities", Whitepaper, [Online]. Available: www.eschergroup.com/publications/five-ict-essentials-for-smart-cities.cfm

[16] G. Falconer, S. Mitchell, "Smart City Framework A systematic process for enabling Smart+Connected Communities", CISCO, Sept. 2012, [Online]. Available: http://www.cisco.com/web/about/ac79/docs/ps/motm/Smart-City-Framework.pdf

[17] T. Bakıcı, E. Almirall, J. Wareham, "A Smart City Initiative: The Case of Barcelona," Journal of Knowledge Economy, vol. 4, no. 2, pp. 135-148, 2013.

[18] A. Caragliu, C. Del Bo, P. Nijkamp, "Smart cities in Europe," Journal of Urban Technology, vol. 18, no. 2, pp. 65–82, 2011.

[19] Anonymous, ITU-T Y.2221 (01.2010), "Requirements for support of ubiquitous sensor network (USN) applications and services in the NGN environment", [Online]. Available: http://www.itu.int/ITU-T/recommendations/rec.aspx?id=10235.

[20] L. Zhang, and Z. Wang, "Integration of RFID into Wireless Sensor Networks: Architectures, Opportunities," In Proc. of the 5th Int. Conf. on Grid and Cooperative Computing Workshops (GCCW '06), Changsha, China, Oct. 2006, pp. 463–469.

[21] H. Liu, M. Bolic, A. Nayak, and I. Stojmenovic, "Integration of RFID and Wireless Sensor Networks," in Proc. of Sense ID Workshop at ACN SenSys, Sydney, Australia, 2007, pp. 6-9.

[22] C. Alcaraz, P. Najera, J. Lopez and R. Roman. "Wireless Sensor Networks and the Internet of Things: Do We Need a Complete Integration?" in Int. Workshop on the Security of the Internet of Things (SecIoT), Tokyo, Japan, Nov. 2010.

[23] F. Al-Turjman, A. Al-Fagih and H. Hassanein, "A Novel Cost-Effective Architecture and Deployment Strategy for Integrated RFID and WSN Systems," in the Proc. of the 1st IEEE Int. Conf. on Computing, Networking and Commun. (ICNC'12), Maui, Hawaii, 30 Jan.- 2 Feb. 2012, pp. 835-839.

[24] B .Ahlgren, C. Dannewitz, C. Imbrenda, D. Kutscher, B. Ohlman, "A survey of information-centric networking", Communications Magazine, IEEE , vol.50, no.7, pp.26,36, July 2012.

[25] D. H. Friend, R. W. Thomas, A. B. MacKenzie, and L. A. DaSilva, "Distributed learning and reasoning in cognitive networks: Methods and design decisions," in Cognitive Networks - Towards Self-Aware Networks (Q. H. Mahmoud, ed.), pp. 223-246, John Wiley & Sons, 2007.

[26] C. Dazhi, and P. K. Varshney, "QoS Support in Wireless Sensor Networks: A Survey." In *International Conference on Wireless Networks*, vol. 233, pp. 1-7. 2004.

[27] J. Shaikh, M. Fiedler, D. Collange, "Quality of Experience from user and network perspectives," J. annals of telecommunication, vol. 65, no. 1-2, pp 47-57, Feb. 2010.

[28] C. Bisdikian, J. Branch, K. K. Leung, and R. I. Young, "A letter soup for the quality of information in sensor networks." In *Pervasive Computing and Communications, 2009. PerCom 2009. IEEE International Conference on*, pp. 1-6. IEEE, 2009.

[29] V. Sachidananda, A. Khelil, N. Suri, "Quality of Information in Wireless Sensor Networks: A Survey", ICIQ (to appear), 2010.

[30] H.-L. Lu, "Quality of service standardization for next generation networks," in *ITU-T Workshop on NGN in collaboration with IETF*, ITU Headquarters, Geneva, Switzerland, May 1-2 May, 2005.

[31] B. Bhuyan, H. Sarma, N. Sarma, A. Kar, R. Mall, "Quality of Service (QoS) Provisions in Wireless Sensor Networks and Related Challenges," *Wireless Sensor Network*, vol. 2, pp. 861–868, 2010.

[32] C. Bisdikian, "On sensor sampling and quality of information: A starting point," 3rd IEEE Int'l Workshop on Sensor Networks and Systems for Pervasive Computing (PerSeNS 2007), 2007, pp. 19-23.

[33] Anonymous, "Quality of experience", [Online]. Available: http://en.wikipedia.org/wiki/Quality_of_experience, Wikipedia, accessed 26 April 2015.

[34] M. Fiedler, T. Hossfeld, P. Tran-Gia, "A generic quantitative relationship between quality of experience and quality of service," *Network, IEEE* , vol.24, no.2, pp.36-41, March-April 2010.

[35] Krishnamachari, B., Estrin, D., and Wicker, S.," Modelling data-centric routing in wireless sensor networks", IEEE Infocom, Vol. 2, pp. 39-44, Jun. 2002.

[36] Intanagonwiwat, C., Govindan, R., Estrin, D., Heidemann, J., and Silva, F., "Directed diffusion for wireless sensor networking", Networking, IEEE/ACM Transactions on, vol. 11, no. 1, pp. 2-16, 2003.

Chapter 2

Cognition in Wireless Sensor Networks: A Perspective

Preface

This chapter is from the publication in IEEE Sensors Journal paper titled: "Cognition in Wireless Sensor Networks: A Perspective". This was an invited paper and presents a survey on cognitive techniques applied to sensor networks and the goals they aim to achieve. The idea of knowledge plane in WSN is introduced, and an existing framework for cognitive networks that sensor networks can benefit from, is discussed. An experiment on ZigBee-based hardware platform was used to show that WSNs can accommodate cognitive behavior, and may benefit from cognitive techniques applied to various sensor network applications.

2.1 Abstract

Wireless Sensor Networks are believed to be the enabling technology for Ambient Intelligence. They hold the promise of delivering to a smart communication paradigm which enables setting up an intelligent network capable of handling applications that evolve from user requirements. Cognitive agents capable of making proactive decisions based on learning, reasoning and information sharing when interspersed in sensor networks, may help achieve end-to-end goals of the network even in the presence of multiple constraints and optimization objectives. Cognitive radio at the physical layer of such agents may enable the opportunistic use of the heterogeneous wireless environment. However, research efforts have been discrete and cognitive techniques have focused on improving specific aspects of the network or benefiting specific applications. The main contribution of this section is providing the vision and advantage of a holistic approach to cognition in sensor networks, which can be achieved by incorporating learning and reasoning in the upper layers, and opportunistic spectrum access at the physical layer. Rather than providing an ostensive survey of cognitive architectures applicable to sensor networks, this chapter provides the reader with a framework based on knowledge and cognition that can help achieve end-to-end goals of application-specific sensor networks.

2.2 Introduction

Technological advancements are often dictated by human needs. Providing enhanced health care services for the aging baby-boomer population, remote monitoring of forest environments to preserve and protect them from natural calamity such as forest fires, and protecting endangered species are examples of needs that have led to an increased focus on research in Wireless Sensor Networks (WSNs). Due to the energy-constrained nature of the sensor nodes, researchers have been working on devising techniques to improve network performance, its lifetime, data-rate, throughput and other such goals, depending on the application. Cross-layer design, energy scavenging, in-network data processing, and energy-aware routing are some of the many techniques employed to achieve these goals. However, these techniques are limited in their scope and do not work well when there is a need to simultaneously optimize multiple goals – typically, lifetime and rate. In recent times, cognitive techniques and tools of artificial intelligence are being increasingly used in WSNs to circumvent the limitation imposed by existing techniques. Knowledge based learning and reasoning is being introduced into the communication architecture in the form of a construct called Knowledge Plane (KP) [1]. "Cognition" refers to the process of knowing through perception, reasoning, knowledge and intuition with a focus on information available from the environment. When nodes with cognitive capabilities are introduced into an entire network of communicating sensor nodes, it gives rise to exciting new opportunities in sensor network research that could overcome the limitations imposed by current design techniques. Cognitive communication among sensor nodes could not only help meet end-to-end goals of the entire network, but also increase reliability of the network, reduce maintenance costs and increase the network lifetime. Hence, research efforts in recent times have been directed towards incorporating cognition, or some level of intelligence into sensor networks [2 - 5]. Currently the air interface of WSNs has been standardized by IEEE 802.15.4-2006 [6] (with two additional addendums, 802.15.4c/d in 2009) and the upper layers standardized by the ZigBee standard [7]. However, there is no definite framework or standard for cognitive communication within a cognitive sensor network. IEEE is working on 802.22 standard for Wireless regional Area Network (WRAN) [8] that provides for enabling communication among cognitive radio devices, but is restricted to the air interface and is not intended to cater to the low data rate, low energy consumption requirements of sensor networks. The contribution of this section lies in providing a vision and the advantages of a holistic approach to cognition in sensor networks, which can be achieved by incorporating learning and reasoning in the upper layers, and opportunistic spectrum access at the physical layer (PHY). It also suggests that a cognitive framework based on knowledge and cognition can help achieve the end-to-end goals of application-specific sensor networks.

This section has been organized into the following sections: Section 2.3 provides an overview of the communication architecture, typical research problems in WSNs and limitations of cross-layer design. In Section 2.4 the concept of Knowledge Plane and a cognitive architectural framework that could be applied to sensor networks to make it truly cognitive in all its actions is introduced. Cognitive techniques applied to sensor networks are summarized and analyzed in section 2.5, followed by section 2.6 which contains an illustration of a simple cognitive routing scheme based on Texas Instruments' ZigBee based CC2430 hardware platform. The advantages and challenges of the cognitive approach are presented in section 2.7 and section 2.8 concludes this chapter.

2.3 Wireless Sensor Networks

2.3.1 Communication architecture

Fig.2-1 illustrates an example of multi-hop communication in WSNs. Sensor nodes communicate data to relay nodes that act as sink for the sensed information. These relay nodes may be assumed to have more sustainable power sources than the sensor nodes that are often powered by non-replenishable power sources (primary batteries). The data is then forwarded to a base-station which may be located in the region of deployment of the sensor network. From here the data is wirelessly transmitted and made available to an end user who remotely monitors the observed data.

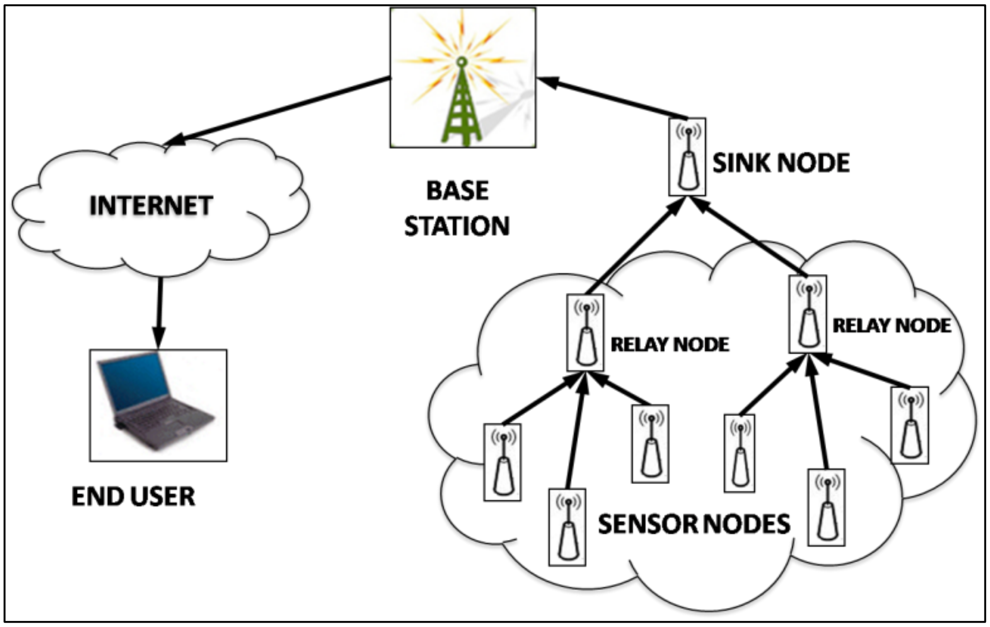


Figure 2-1: Example of Multi-hop Communication in Wireless Sensor Network.

These features of WSNs made them immensely popular in a wide range of applications such as health, military, environment monitoring, security and they are even believed to be the enabling technology for ambient intelligence [9]. Each of these applications requires sensing different events and optimizing performance based on a unique set of parameters and constraints to achieve specific goals for the network.

2.3.2 The WSN Protocol Stack

IEEE 802.15.4-2006 [6] is the standard that defines the physical and medium access layers for radio frequency communication among WSN devices operating in the 2.4GHz and 868/915MHz license free Industrial, Scientific & Medical (ISM) bands. ZigBee builds on top of the 802.15.4 standard, and defines the network and application layers and a security service provider (SSP) [7]. The ZigBee specifications were developed by a consortium of multiple vendors - the ZigBee Alliance in an effort to provide a common platform for the development of low cost, low data rate, short range devices with low power consumption for WSN applications. Thus the ZigBee stack, comprised of IEEE 802.15.4 based PHY & MAC & ZigBee based network & application layers forms the WSN protocol stack as shown in Fig 2-2. The different layers of the stack communicate with each other using Service Access Points (SAP) which are interfaced to the data entity (DE) or management entity (ME) services provided by a specific layer to the upper layers [7]. Three types of devices are defined within a ZigBee network: A Coordinator that starts and configures the network, a router that supports associations and forwards messages to other devices in the network and an end device that communicates the sensed data to other devices in the ZigBee network. The ZigBee Device Object (ZDO) within the application layer (APL) defines these roles and is also responsible for initiating or responding to binding requests and securing relationships between network devices.

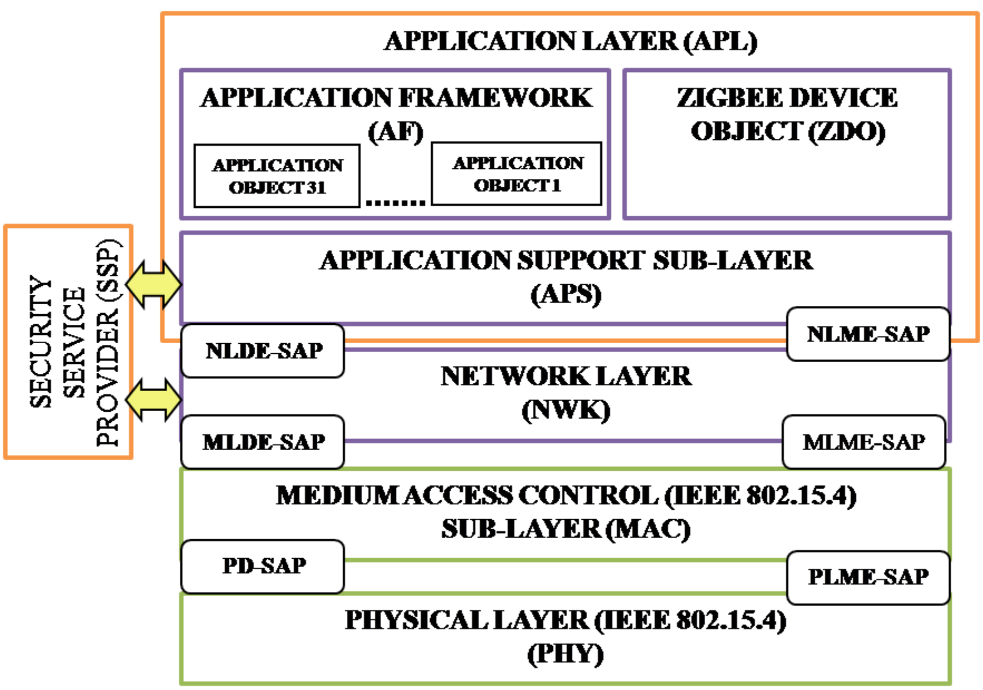


Figure 2-2: IEEE 802.15.4/ZigBee protocol stack.

Star, cluster tree and mesh are the supported topologies and the network layer supports all the functions related to starting a network, addressing, routing, and synchronization within the network. Manufacturer defined application objects within the Application Framework (AF) implement the actual applications in accordance with the ZigBee-defined application descriptions. The IEEE 802.15.4 PHY and MAC along with ZigBee's network and application layers provide reliable data transfers over short ranges at very low power consumption, thus making it convenient to deploy sensor networks in a variety of applications.

2.3.3 Research problems in sensor networks

The following problems receive considerable research attention in sensor networks:

i. *Network lifetime maximization*

This deals with the problem of maximizing the lifetime of energy-constrained nodes in sensor networks by computing optimal values of parameters such as transmission powers, rates, and link schedule [10].

ii. *Energy efficient routing*

Communication of information over a wireless channel consumes significant power. Efforts are made to find the best routing technique suitable to the specific sensor network applications to minimize energy consumption during routing, while guaranteeing data delivery and network connectivity [11].

iii. *Reliable event detection and transfer*

In event-based sensor networks, reliable detection of events and their transfer to a sink are important goals. While some researchers focus on end-to-end reliable event transfer schemes [12], others focus on event-to-sink reliable transport [13].

iv. Optimization among multiple, conflicting objectives

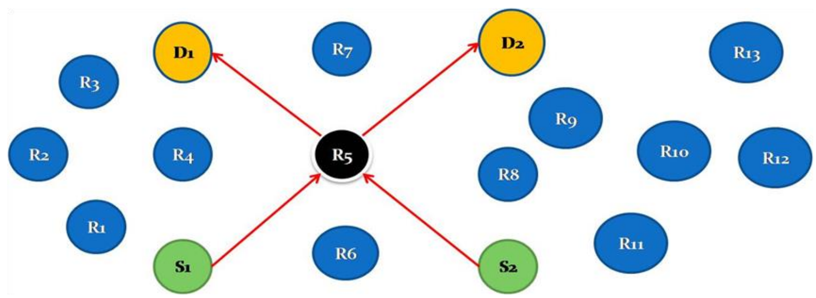
Sensor networks might have more than one goal to achieve such as optimal rate and network lifetime. New approaches are often researched to find means of maximizing network performance while trying to achieve conflicting goals [14].

v. Bringing flexibility into the application-specific design of WSNs

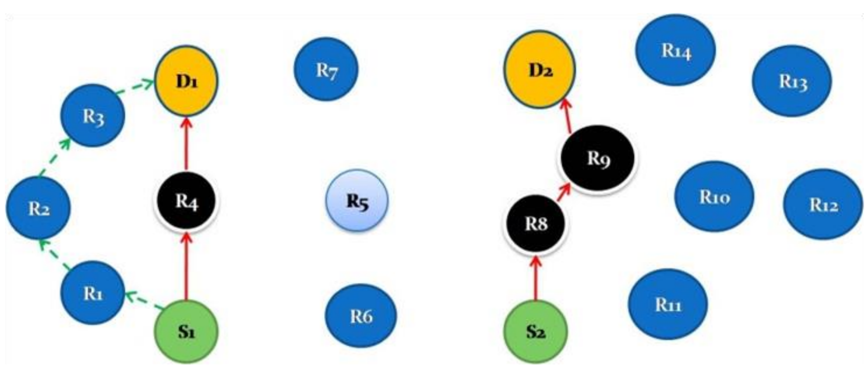
Software [15] or middleware [16] based Mobile Agents are used to dynamically deploy new applications in otherwise application specific sensor networks.

Most of the above design objectives require a cross-layer approach where there is direct communication / sharing of information among non-adjacent layers. Observations (about network performance, channel conditions etc.) are made available for adaptation at layers other than the ones providing the information. Although several authors have applied cross-layer techniques in solving design problems in the sensor network protocol stack, this technique has its own limitations. Cross-Layer design is focused on interactions among a subset of layers of the sensor network protocol stack where independent optimizations are performed without accounting for the network-wide performance goals. They are memory-less adaptations that react in the same way to a set of inputs (irrespective of how poorly the adaptation had performed in the past). The design does not evolve around the application requirements and is more data centric rather than being knowledge centric. Increased complexity and reduced modularity lead to adaptation loops [17] which in turn increase the maintenance cost of such networks. Here is where cognitive approaches come in with an upper hand. They provide the network with the ability to observe and learn from past behavior to improve network performance. The illustration in Fig. 3 shows the advantage of the cognitive approach.

In Fig.2- 3a, S1 and S2 are the source nodes that are trying to route data to destination nodes D1 and D2. Rn are all the available relay nodes that can help in transmitting the data from the source to the destination. Out of these relay nodes, it is determined that R5 has the lowest link outage probability to D1and D2. Hence S1 starts routing data through R5. In the meantime, S2 also starts routing a high traffic of data through R5 (Indicated by the solid paths). When multiple source nodes start routing their data through this node, the route through R5 may get congested. A cognitive network with learning capabilities may be able to identify the congestion at R5 (by observing the decrease in throughput at the source nodes).



a: Classical routing in a sensor network.



b: Cognitive routing in response to congestion.

Figure 2-3: (a) Classical routing and (b) Cognitive routing in WSNs.

Sharing this observation with all the nodes in the network, the cognitive network would be able to respond to the congestion proactively, by routing the data through a different path involving nodes R4, R8 and R9 as shown in Fig. 2-3b. This helps to preserve nodes like R5, thus maintaining network connectivity and providing reliable data transfer. Based on the application, the network may be able to choose between minimizing the number of hops by choosing route S1->R4->D1 or minimizing power, irrespective of the number of hops by choosing route: S1->R1->R2->R3->D1 for example. A similar example can be found in [17].

Knowledge based cognitive networks offer a new research paradigm to avoid the pitfalls of cross layer design and optimize system performance in the presence of multiple conflicting goals. The ability to tune node behavior to user requirements gives cognitive networks an end-to-end scope. The first mention of such a network was made by Clark et al. in [1] where they proposed the concept of a "knowledge plane" to accomplish cognitive tasks. Thomas et al. [18] built on this concept and proposed a framework for cognitive networks.

2.4 Cognitive Architectures

2.4.1 Knowledge plane

The concept of Knowledge Plane (KP) was first proposed by Clark et al. [1] for the internet in the wired communication domain. The objective was to break the barriers of the layered structure of the protocol stack and enable seamless communication across the layers, up the protocol stack as illustrated in Fig. 2-4. KP was proposed to be a pervasive system based on knowledge rather than tasks, so that observations from different parts of the network could be correlated to make judgments in the presence of incomplete, inconsistent or conflicting information in dynamic environments. According to the authors in [1], the KP was expected to make decisions in the presence of partial or conflicting information, automate decisions, respond to emergency situations and even foresee problems and proactively take corrective actions. The idea was to build a network that could assemble itself given high level instructions, adapt itself to changes, reassemble if required, discover problems and fix them or explain why it cannot be fixed. Basically, the KP was proposed as a solution to the limitations imposed by the cross-layered approach. Reasoning was expected to support the network's high-level goals and constraints and mediate between users or operators with conflicting goals and design constraints. The end-to-end goals of the entire network were kept in mind during optimizations, rather than those of a few interacting layers as in the cross-layer design approach. This required the system to have knowledge of the network and its actions and also have a learning mechanism to make decisions based on its past experiences. Hence, the tools of Artificial Intelligence (AI) & cognitive techniques of representation, learning and reasoning, were believed to be best suited to achieve the complex objectives of the KP as opposed to traditional algorithmic approaches.

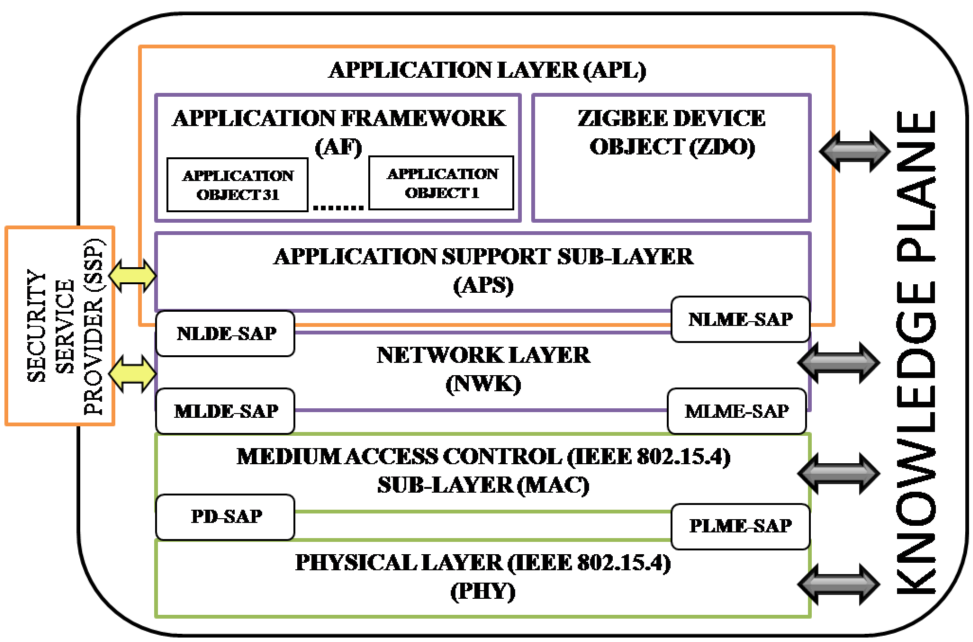


Figure 2-4: Knowledge Plane in WSN Protocol Stack.

Thus, the first step towards introducing intelligence and some kind of cognition into communication systems was made by Clark et al. Thomas et al. [17] adopted this idea of the knowledge plane from the wired communication world, to the wireless communication domain and proposed the concept of a Cognitive Network (CN).

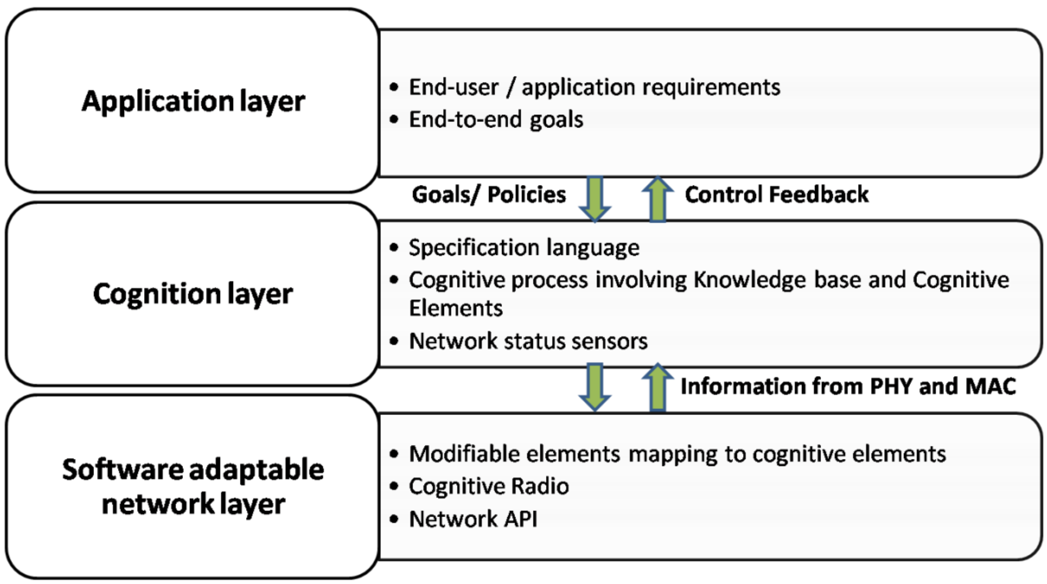
In the Cognitive Network (CN) paradigm, the authors spoke of the end-to-end scope of the network's goals, involving all elements within a data flow. It aimed at achieving these goals by breaking the layering up of the network stack & communicating with nodes across the whole network. The CN was defined to be a self-aware, self-organizing and adaptive network capable of making intelligent adaptations based on (a) observations of the network state made by individual elements, (b) information sharing among nodes beyond the limitations of the layered protocol architecture, (c) learning and reasoning before acting on its decisions to optimize network performance [17]. These CNs derive knowledge about the network performance from end users and applications and have an end-to-end scope.

2.4.2 The Cognitive framework

A three-layer framework was described by Thomas [19] in order to implement the goals of the CN. These three layers were constituted of the *Requirements layer, Cognition layer and Software Adaptable Network (SAN)* in order to implement the goals of the CN as described in Table 2-I.

The SAN is the architecture's interface with the Physical world. Configurable network elements such as directional antennas, cognitive radios (in which transmit power can be modified) etc. form the action elements of the cognitive process. These are known as the *modifiable elements*. For each modifiable element, there is a one-on-one mapping with *cognitive elements* in the Cognitive process in layer 2. These elements in the cognition layer help to distribute the operation functionally and spatially. Network status sensors provide partial knowledge of the network to the cognitive elements. At the highest level of abstraction is the Requirements layer that transforms end-to-end objectives to goals for each cognitive element of the Cognition Layer through a Cognitive Specification Language (CSL). The Cognition layer forms the central mechanism of this architecture. It learns about the system state, has knowledge about the current network goals and decides on an appropriate response to observed network behavior. It makes use of a feedback loop in which past interactions with the environment guide current and future interactions. An OODA loop [18] - Observe, Orient, Decide and Act is used as the feedback loop.

Table 2-I: Cognitive network architecture.



It is very important that the process is able to learn and converge to a solution faster than the network status changes and keep itself updated with the changes so that it can re-converge to a new solution quickly. To establish the effectiveness of the three-layered cognitive architecture of Table 2-I on a wireless network, the authors [18] experimented on the problem of providing a distributed cognitive network solution to maximizing the connection time for unicast and multicast communication between a source node and one or more destination nodes. The network consisted of multiple radios, both battery powered and wire-line supplied. The only modifiable network element in each radio was at the physical layer - the adjustable transmission power. Table 2-II provides a mapping of each of the network functions onto the three-layer cognitive framework of Table 2-I.

Table 2-II: Mapping of network goals/functions of a Multicast Flow to the three-layer cognitive framework.

Layer	Sub-layer	Goals/Functions
Application Layer	Requirements layer	Maximize the connection time of a unicast/multicast flow in a wireless environment
Cognitive Layer	Cognitive specification language	Map the goal of connection lifetime to a fitness function that describes the total power utilization of the chosen route.
	Cognitive process	Control output power from transmitting nodes
	Network status sensors	Measure the following parameters: power output at each node capacity at each node network connectivity
Software adaptable network layer	Network API	Provide the following to cognition layer: cognition limits on possible power settings a hook for setting requested power output Provide a bi-directional communication mechanism to share utilization ratios and calculate fitness function among nodes in the network.

The cognition process interacts with the environment and “learns” to optimize the network by adapting the power output of all interior tree nodes to a lower power utilization ratio while maintaining a fully connected multicast tree. The performance of this cognitive method was compared with the performance of a non-cognitive, heuristic method called DRIP (Directional Reception Incremental Protocol).The authors [18] found that the cognitive method was able to out-perform the heuristic method in terms of average power utilization. Moreover, across a given set of receivers, the heuristic method could find a workable solution only 86% of the time, whereas the cognitive method was able to find a solution 97% of the time. With these results, the architecture and potential of the cognitive network was illustrated and their capacity to handle data networks operating in heterogeneous, dynamic and complex environments using learning was established. This architecture can be extended to optimize multiple goals by enhancing the capabilities of the cognition layer. Hence, the KP approach combined with the cognitive framework suggested by Thomas may provide valuable performance improvements for WSNs. Several other cognitive architectures have been proposed in the past, with an attempt to achieve a high-level of cognition. SOAR [20] developed to implement general intelligent behavior, DUAL [21] a context-sensitive cognitive architecture, and ICARUS [22] a cognitive architecture for physical agents are examples of such cognitive architectures. Duch et al. present a critical survey of the state of the art in cognitive architectures in their work in [23] and also provide a simplified taxonomy of these architectures based on memory and learning. Symbolic, Emergent and Hybrid Paradigm architectures were the three main groups according to this classification.

SOAR [20] and ICARUS [22] were some of the architectures classified under the Symbolic type. They used a classical artificial intelligence based, top-down analytic approach for information processing. High-level symbols or declarative knowledge were used and centralized control was exercised over information flow. NOMAD [24] was an example of *Emergent Architecture.* These architectures typically used a bottom-up approach and consisted of a large number of processing units in the network through which low-level activation signals would flow. NOMAD was capable of real-time pattern recognition and ran on a set of powerful computers that simulated the nervous system. However, this architecture was not envisioned to be capable of achieving a higher level of cognition. ACT-R [25], DUAL [21], and LIDA [26] are examples of the *Hybrid architecture* that combine various aspects of both symbolic and emergent architectures. ACT-R used a top-down learning approach and was another example of a system that aimed at understanding and emulating human cognition. Although this architecture had been successfully applied in psychological studies and intelligent tutoring systems, applications of this system in problem solving and reasoning were still missing. LIDA made use of software agents and codelets to perform distributed tasks, and implemented several forms of learning from events to implement the cognitive system. While this architecture was capable of explaining many features of the mind, its competence in understanding language and reasoning were still to be established. Thus, it can be seen that many of the cognitive architectures proposed in the past have aimed at achieving a high-level of intelligence, but have had very few real world applications. Since the primary motivation behind introducing cognition into sensor networks is to improve network performance, rather than trying to imitate the human brain closely, the cognitive framework proposed by Thomas was found most suitable in the networking context.

In the next section, we take a look at the various research directions in sensor networks in recent literature that use techniques including artificial intelligence and cognitive radio.

2.5 Cognitive Techniques in Sensor Networks

This section presents the latest trends in WSN research, which largely includes cognitive approaches being adopted by researchers to improve the performance of WSNs.

2.5.1 Cognitive Radio in WSN:

Cavalcanti et al. [27] present a conceptual design for Cognitive Radio (CR) [28] based WSN and compare its performance with a standard ZigBee/802.15.4 WSN, both built on the standard model available in OPNET (ZigBee/802.15.4), operating in the 2.4 GHz band. In this experiment, the authors [27] assumed the 802.15.4 based CSMA access method in non-beacon mode at the MAC layer and ZigBee based protocols at the network layer (table-based mesh routing) and application layer for both the CR and 2.4GHz modes. Transmit and receive antennas were both assumed to have the same unit gain. The receiver sensitivities were set at -85dBm for the CR channel centered at 680MHz, as well as for the first channel in the 2.4GHz band. From these simulation results, the authors found that for the same transmit power, the maximum communication range in the CR channel is almost twice of what is obtained in the 2.4GHz channel. This increased range reduces the number of hops travelled per packet and hidden node problems, thus enhancing the efficiency of the multi-hop routing and the MAC. The overall application throughput was also found to be better in the CR mode.

Akan et al. [29] also talk about the main design principles, potential advantages, application areas, and network architectures of cognitive radio sensor networks. They explore the possibility of applying existing techniques for cognitive radio and WSN to such networks and identify the challenges in doing so.

2.5.2 Cognitive schemes using Neural Networks models

Reznik and Pless [4] establish the feasibility of using distributed intelligence to embed cognition into sensor networks by studying the problem of signal change detection[1]. They map Artificial Neural Network **(**ANN**)** architecture to sensor networks and experimentally prove the advantage of this approach in terms of reduction in resource consumption - network bandwidth, processor power and memory usage due to reduced connectivity and communication costs.

Youssef and Younis [30] propose GRISP (Gateway Relocation algorithm for Improved Safety and Performance), a neural network model to assess safety of gateway/sink node at various locations in a WSN environment trained using genetic algorithms. A “Threat index” for each location visited, along with the snapshot captured trains the neural network. This helps the neural network generate a “Risk Assessment factor” for making future safe relocation decisions.

2.5.3 Cognitive Sensor Networks (CSN)

Shenai et al. [31] have presented a distributed wireless sensor network-based control system for intelligent and reliable operation of large power grids. Here sensor data (voltage and power factor) is reported to motes that can be mapped to “modifiable elements” of Thomas’s Cognitive framework [21]. These motes are intelligent and communicate with nearby motes as well as a “control station” that can be mapped to the “requirements layer”. AUTOMAN, the software cognitive agents that motes and the control stations run, are said to have knowledge of local policies as well as awareness of the end-to-end operational requirements of the end application. Dynamic decisions are delivered and adequate information management is achieved by combining techniques of sensor coordination and intelligent data fusion. Thus, even under

[1] Definition of Change detection in [4] : “An identification of unforeseen change in general characteristics and parameters of the measured signals against the predicted ones”

changing environments, dynamic reconfiguration is possible without grid downtime and the system also ensures that the QoS requirements of the customer are always respected.

Boonma and Suzuki propose *MONSOON* [5], a biologically inspired framework to build cognitive WSN applications. This framework introspectively understands conflicting design objectives (data yield, data fidelity, power consumption), finds optimal tradeoffs with given constraints and autonomously adapts to the dynamics of the network. It models an application as a decentralized group of software agents that collect sensor data from individual nodes and carry them to the base stations. From simulations, the authors' show that agents adapt to network dynamics by satisfying conflicting objectives under given set of constraints and exhibit self-configuration, self-optimization and self-healing properties.

2.5.4 Game theoretic formulation of energy efficiency and security in WSNs

Machado and Tekinay present a survey of game-theoretic formulation to the problems of energy efficiency and security in WSNs [2]. They found that distributed decision-making capabilities of WSNs and the selfish behavior of the individual nodes is exploited by the game theoretic approach to optimize performance at the node-level (conserving battery power) as well as network level (maximizing the network utility which is directly proportional to the number of sensors involved). They discuss the use of game-theoretic approaches in performing distributed cross-layer optimization by making use of power control game at the physical layer and rate allocation game at the application layer. In dealing with security issues in WSNs, the authors' present work on game-theoretic models used to analyze situations where there are attacks by malicious nodes and outside intruders on WSNs. They also talk about the use of pursuit-evasion games in WSNs for model detection, tracking and surveillance applications.

2.5.5 Cognitive and Self-Selective routing

Gelenbe *et al*. [32] present Cognitive Packet Network (CPN) and Self-Selective Routing (SSR) algorithms that use different forms of learning as new approaches to achieving Quality of Service (QoS) routing in WSNs. CPN routing uses Smart Packets (SPs) for path discovery, along with random neural networks with reinforcement learning. It has the ability to adapt to varying traffic load and is scalable for networks with flows to many destinations. SSR technique on the other hand, makes use of pheromone based communication inspired by ants in a colony that communicate information about traversed paths to members of their kind. There is self-selection of routes at each node, which leads to additional overhead, but provide the network with the ability to adapt to conditions where there are unexpected link failures or the connections are unreliable. Both the algorithms are capable of supporting fault-tolerance to different degrees and their protocol structure allows for different levels of efficiency in diverse application contexts.

2.5.6 Cognitive schemes using adaptive modulation and sleep-scheduling

In [33], a framework for performance modeling and design of cognitive WSNs is presented. The approach makes use of adaptive modulation, sleep scheduling and energy-aware higher layer processing based on information sharing across layers. The cognition algorithm makes use of feedback about channel condition and modulation rate to determine the sleep time of each node. This concept is applied in the following simulation study assuming realistic parameters for Energy consumption of each node in four modes: sleep mode, receive mode, active mode (ready to transmit but not transmitting), and transmission mode. A simultaneous cognitive management of sleep cycles and M-QAM modulation schemes saves energy and extends network lifetime significantly.

Fig. 2-5a displays the mean node lifetime in a WSN using Adaptive Modulation (AM) only. In this case the modulation level (parameter M in M-QAM modulation) is chosen for each packet according to the channel condition (i.e. instantaneous SNR value). This case assumes no cognition between PHY and MAC layers and the sleep time is predetermined by the MAC layer according to the multiple access rules. Fig. 2-5b shows the mean node life time using an adaptive modulation scheme combined with scheduled sleep (AMS). Here cognition is employed at the MAC and PHY layers and sleep times are scheduled according to channel conditions and bit rates. In both cases two traffic patterns are studied using Poisson distributions: High traffic (mean packet arrival rate 90%) and low traffic (mean packet arrival rate 10%) and log-normal shadowing where shadowing variance takes values from 0 (no shadowing) to 6dB. The figures show that the cognitive approach significantly outperforms the non-cognitive one in terms of node lifetime.

Table 2-III: Node parameters used in simulation.

Parameter	Value
Current consumption in Sleep mode: I_{sleep}	1 µA
Current consumption in Receive mode: I_{rx}	20 mA
Current consumption in active mode (ready to transmit but not transmitting): I_{ac}	100 mA
Current consumption while transmitting	120mA
Traffic intensity	10% or 90%
Log-Normal Shadowing variance (σ)	0, 2dB, 4 dB or 6dB
BER required (QoS)	10^{-4}
RF Bandwidth used	200kHz
Adaptive Modulation Stages : 0 (no transmission), BPSK, 4QAM, 16QAM, 64QAM and 256QAM	6 stages

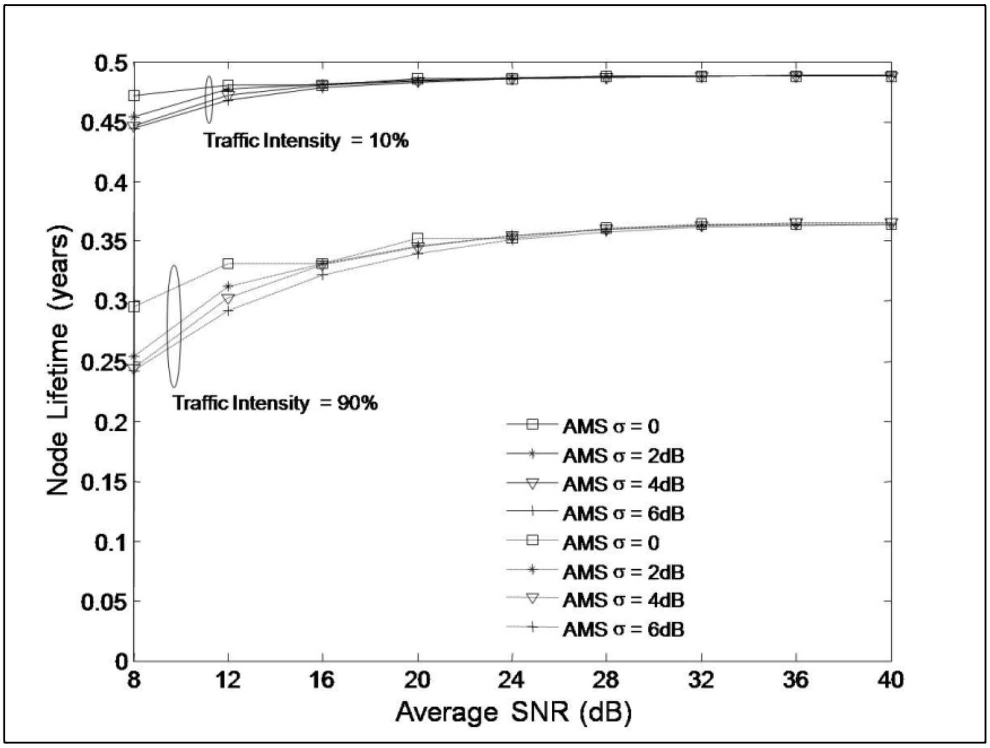


Figure 2-5a: Node lifetime using Adaptive Modulation.

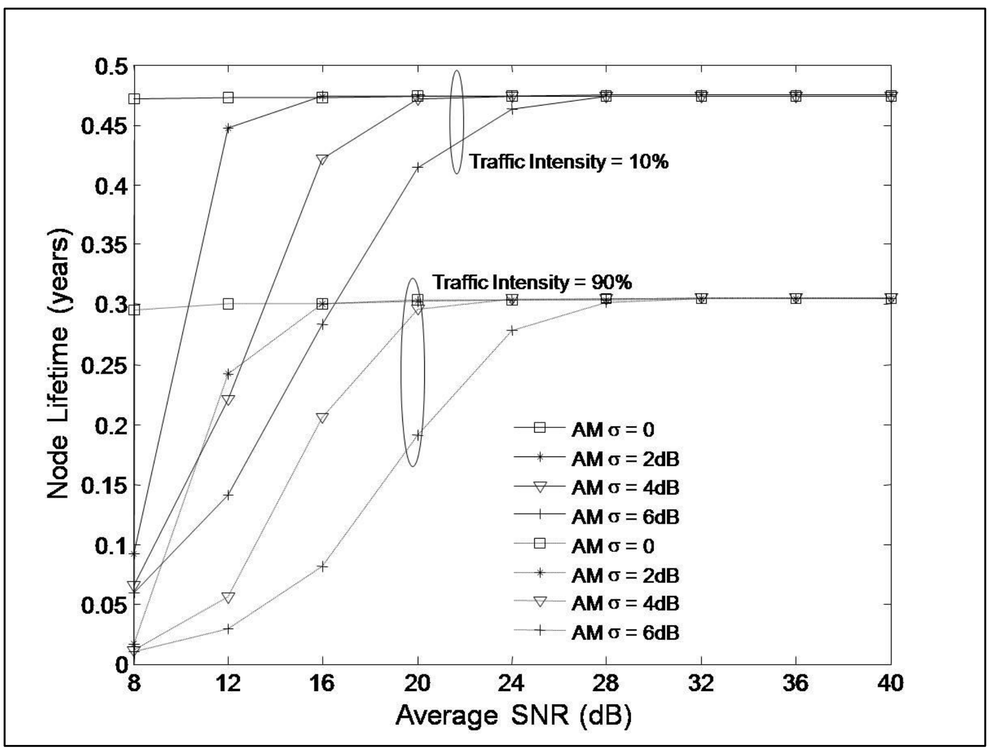


Figure 2-5b: Node lifetime using Adaptive Modulation and Adaptive Sleep.

This is because M-QAM modulation, when carefully chosen (with respect to channel conditions) will require less transmission time, the cognitive system exploits this information to modify sleep time accordingly. The improvement made by the cognitive approach is higher for both high traffic intensity and severe channel conditions (i.e. low SNR and high shadowing variance). The reason is that high traffic and bad channel conditions require frequent transmissions and re-transmissions, respectively. Here the AM scheme is penalized because it does not go to sleep immediately after each (re-) transmission. Instead, it stays active and goes to sleep only when the pre-scheduled MAC sleep time starts.

2.6 Summary and analysis from Table 2-IV

From the analysis in Table 2-IV, the approaches proposed by Shenai et al. [31] and Boonma and Suzuki [5] are the ones closest to the author's vision of applying cognition to sensor networks in order to achieve end-to-end goals of the network. The results presented in [33] suggest the advantages of sharing information seamlessly across the layers of the network. These cognitive techniques applied in sensor networks, definitely promise improvements over the cross-layer approach, especially because they are based on knowledge and learning. However, these techniques do not explore cognitive radio at the lower layers to opportunistically access the licensed spectrum, which could be an added benefit. This strengthens the author's perspective on the benefits that can be derived out of a holistic approach to cognition in sensor networks based on knowledge, learning and information sharing.

Table 2-IV: Comparison of cognitive techniques applied to sensor networks.

Technique	Goal achieved by cognition	Means of achieving the goals	Cognition based on knowledge/learning / reasoning/context awareness	Influence of cognition on network's end-to-end goals
CR in WSN [27]	Increased communication range and application throughput	Implementing CR at PHY	Context awareness	No
ANN [4]	Reduced resource consumption by reducing connectivity and communication costs	Implementing distributed intelligence by mapping ANN to WSN architecture	Reasoning	Yes to a limited extent
GRISP [30]	Make safe relocation decisions for Gateway node of WSN	Neural network model trained using genetic algorithms to assess safety of sink nodes	Learning	No
CSN [31]	Intelligent and reliable management and operation of large power grids, ensuring QoS requirements of end user are always respected.	Distributed sensors communicating with motes that have intelligent software agents called AUTOMAN, make system aware of end-user requirements and enable dynamic reconfiguration	Knowledge, context awareness	Yes
MONSOON [5]	Network exhibits self-configuration, self-optimization and self-healing properties by means of software agents	Decentralized group of software agents inspired by a biological framework that adapt to network dynamics by satisfying conflicting objectives under given set of constraints	Knowledge, context awareness,	Yes to a good extent
Game theoretic formulation [2]	Network and node level performance improvement by game theoretic	Power control games at PHY, and rate control games at Application layer for distributed cross-	May only be considered as an analysis tool; not cognition	No

	formulation of problems of energy efficiency and security in WSNs.	layer optimization; other game-theoretic models to analyze security-issues in WSNs		
Cognitive and self-selective routing [32]	QoS routing in WSNs under diverse application contexts	Implementing different forms of learning such as: RNNRL and pheromone based techniques for cognitive routing	Reinforcement learning	No
Adaptive modulation and sleep scheduling [33]	Better node lifetime in environment monitoring applications.	A cognition algorithm operating primarily on PHY and MAC layers, making use of feedback about channel conditions and modulation rate for adaptive sleep.	Learning from feedback	No

2.6.1 Illustration of a simple cognitive routing scheme on ZigBee based hardware platform

In order to illustrate the concept discussed in Fig. 2-3, we setup a simple experiment making use of Texas Instruments' (TI's) CC2430 ZigBee Development kit. The CC2430 is a widely used platform for WSN deployment and experimentation. However, there are no existing standards or platforms available for experimenting with the cognitive concepts presented in this section. As discussed in the section, most of the current work has been software or simulation based. Hence, we decided to make use of TI's ZigBee based platform and introduce cognition in the network by means of cognitive decision making during routing. This provides a means of remaining rooted to the 802.15.4/ZigBee standard while experimenting with the concept of cognition in the sensor network and see how it is able to benefit the same.

2.6.2 Experiment setup

The experimental setup is as shown in Fig. 2-6 and consists of the following: (i) A set of sensor nodes that sense some environmental parameters (such as temperature, humidity, etc.) and battery levels on end devices constituted by CC2430Development Board (DB) or CC2430Evaluation

Module (EM) on battery boards (ii) A pair of routers R1 and R2, constituted by two CC2430EMs on SmartRF04Evaluation Boards (EBs) that collect information from these sensor nodes & forward them to the final destination - the coordinator node (iii) The coordinator node sends the received packets over a serial port, which can be viewed on a PC through HyperTerminal by connecting a serial cable. IAR embedded workbench was used to program the nodes in this experiment.

A ZigBee network establishes new paths for data communication when existing links fail or break when nodes die (run out of battery). However, in this experiment, we propose to add intelligence to the router nodes to detect increased packet loss in successive transmissions and initiate rerouting of data before significant amount of data is lost. In this experiment, a packet loss threshold is taken as an indication of increasing congestion at a given node.

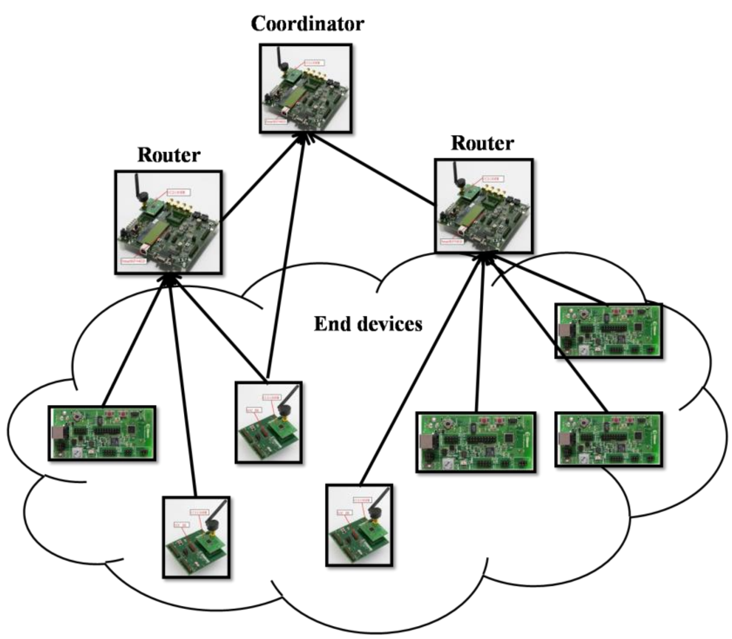


Figure 2-6: Experiment setup.

2.6.3 Detecting packet loss:

A one byte destination sequence number (DSN) in the MAC Header (MHR) is used to track the number of packets lost during transmission. A jump in the sequence number at the receiving end indicates that packets were lost during the communication, and the difference between the received packet's sequence number and the expected sequence number gives the number of packets lost [34]. When the packet loss exceeds a user-programmed threshold at any router in successive transmissions, it will immediately request to disassociate itself from the network.

2.6.4 Results and Inference:

The following observations were recorded when one of the intermediate routers turned its allow-bind mode off as part of the intelligent behavior:

i. Router redundancy is advantageous when introducing cognition into the network as it helps maintain connectivity when intermediate routers are congested.

ii. In case there is no router redundancy, the end devices were able to locate an alternate router between 6s and 13s or directly bind to the coordinator if it was in its communication range.

From the experimented approach of introducing intelligent decision making in the network, it was seen that the cognitive behavior was well supported by the ZigBee stack. New routes were decided based on the Received Signal Strength Indicator (RSSI) value, AODV protocol and the routing table (nearest neighbor).

Implementing the latest and more advanced sensor network protocols for routing will further enhance the cognitive behavior. Instead of a router disassociating itself from the network, it may be able to identify the nodes causing increased traffic/congestion and have such nodes find alternate or less loaded routers. Since cognition in sensor networks is an emerging field, hardware development lags behind software and simulation platforms. Once the platform development catches up, cognitive sensor networks will be capable of offering themselves to a whole new gamut of advanced sensor network applications.

2.7 Advantages and challenges

Research in [2-5, 27, 29-33] suggests the growing interest in applying cognitive techniques to WSNs. The optimization in a cognitive network is done in the light of intelligent adaptations based on learning, reasoning and information sharing among multiple nodes in the network to achieve end-to-end goals. This leads to improved performance of the network over an extended period of time even in the presence of conflicting goals. In a heterogeneous wireless environment comprised of nodes of a sensor network (battery operated sensor nodes, and rechargeable battery powered relay nodes), cellular phone and Bluetooth radio, data from the sensor network's source node may be opportunistically routed through the cellular and Bluetooth radio nodes before it reaches the sink. This can save the energy expended by intermediate sensor nodes in data communication. Such a cognitive sensor network setup would help extend the life of the sensor network. To benefit further from this approach, a cognitive framework could be setup for the sensor network similar to the one proposed in [19], where the software adaptable network layer implements CR functionality to detect spectrum opportunities for communication in a heterogeneous wireless environment. Learning and reasoning processes could aid in intelligent and adaptive decision making, thus enhancing the benefits of adding cognition to sensor networks. Thus the proposed cognitive framework provides an opportunity to develop an architectural foundation on which end-to-end goals of the network can be optimized. Such a knowledge based approach would bring sensor network applications closer to a ubiquitous computing environment, making this an interesting area of research.

However, applying cognitive techniques involving artificial intelligence and game theoretic approach to increase knowledge in the system has several challenges. As stated by the author in [35], nodes/agents may be cheating about their valuations or might have communication bounds, i.e. the number of possible subsets of resources might be very large, and therefore communicating an agent's valuation for each possible subset of the resources might become infeasible. It has been established through several research efforts [2-5, 27, 29-33] that cognitive techniques provide performance improvements in WSNs. However, establishing the feasibility of integrating Cognitive Radio into the "Dynamic Spectrum Access" [36] scheme at the physical layer, along with cognition in upper layers to achieve end-to-end performance goals is an open research problem. This approach means adding intelligence about spectrum handoff decisions to the knowledge plane of the sensor network protocol stack in order to save the energy overhead, as more spectrum handoffs mean more energy consumption. For such networks to be justifiable, the performance improvement must outweigh the cost in terms of overhead, architecture and operation.

For successful implementation of the cognitive process, the Cognitive Network would need information / decision on the implementation of the network i.e. distributed or centralized and the amount of network state information known to the process. An analysis on the amount of energy expended in information sensing (spectrum-hole [36] detection and network status detection) and communicating the same to neighboring nodes is essential in establishing the suitability of this approach to sensor networks. Since the information available to the network may be partial or incorrect, it may lead to security issues and hence, techniques to deal with such issues must also be identified. The proposed cognitive nodes could be static agents distributed in the network or could be mobile agents gathering information from remote locations of the sensor nodes, as depicted in Fig. 2-7.

Hence deciding on the optimal deployment architecture of the cognitive capability enhanced nodes is also a challenging problem.

Finally, though the experiment in the case study suggests that ZigBee stack supports intelligent behavior, in order to achieve a completely cognitive sensor network, standardization of such networks is essential. Protocols that define how the knowledge plane can be implemented to seamlessly access information from the end user and use it to make decisions at the physical layer, the cognitive specification language and the tools used in cognitive decision making must all be standardized to make such networks interoperate.

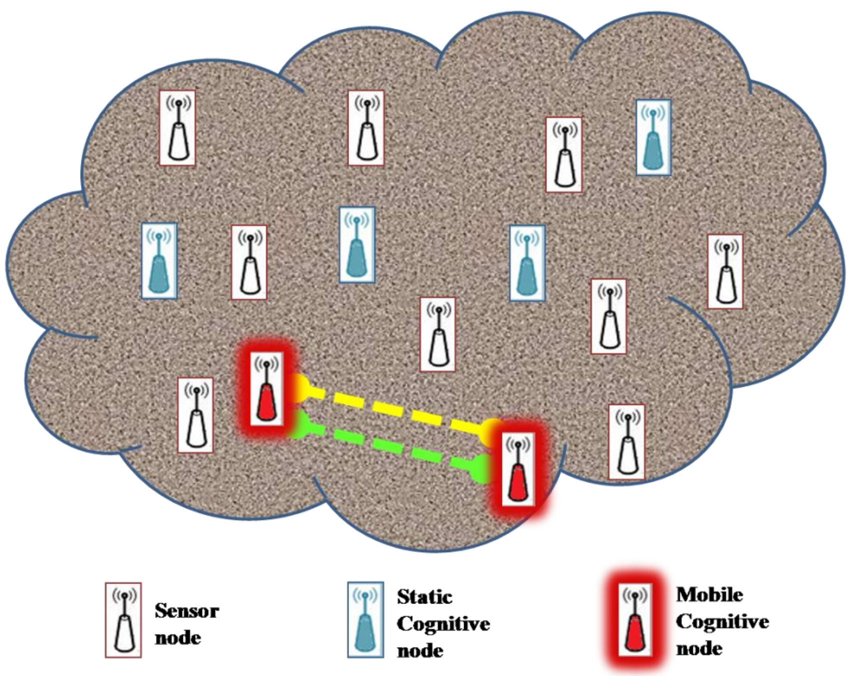


Figure 2-7: Placement of Static and Mobile Cognitive Agents in a WSN.

2.8 Conclusion

The goals of specific implementations of sensor networks are unique. Constraints in one application may be network goals to be optimized in another. Thus, the details of implementing a cognitive sensor network based on knowledge that achieves end-to-end goals of the network, offers itself to involved research.Cognitive communication in a sensor network could not only help meet end-to-end goals of the entire network, but also increase reliability of the network, reduce maintenance costs and increase the network lifetime.The cognitive network provides a knowledge based framework where network decisions are based on learning and reasoning and on information shared among the network nodes about the observations made. Such a framework applied to sensor networks could go a long way in achieving application specific objectives in the presence of multiple, sometimes conflicting optimization goals. The idea of a holistic approach to introducing cognition in heterogeneous sensor networks that combines the advantage of opportunistic spectrum access at the physical layer, with cognitive communication among sensor nodes seamlessly across the network promises to be advantageous over existing design techniques.

2.9 Acknowledgment

The authors thank Shandy Zhao for providing simulation results for section 2.5.6. They would also like to thank Ontario Centers of Excellence (OCE) and Ontario Ministry of Natural Resources for their support.

2.10 References

[1] D. D. Clark, C. Partrige, J.C. Ramming, J.T. Wroclawski, “A knowledge plane for the internet”, *Proc. of the SIGCOMM* 2003, Karlsruhe, Germany, Aug. 2003, pp. 3-10.

[2] R. Machado, S. Tekinay, “A Survey of game-theoretic approaches in wireless sensor networks”, *Comput. Netw. J.*, Elsevier, vol.52, no. 16, pp. 3047-3061, Nov. 2008.

[3] Y. B. Reddy, C. Bullmaster, “Application of game theory for cross-layer design in cognitive wireless networks”, *6th Int. Conf. on Inform. Technology: New Generations*, itng, pp.510-515, 2009.

[4] L. Reznik, G. Von Pless, “Neural networks for cognitive sensor networks”, *IEEE Int. Joint Conf. on Neural Network.*, IJCNN 2008, pp. 1235 - 1241, June 2008.

[5] P. Boonma, J. Suzuki, “Exploring self-star properties in cognitive sensor networking**”,** *Proc. of IEEE/SCS Int. Symp. on Performance Evaluation of Comput. and Telecommun. Syst.* (SPECTS), Edinburgh, UK, June 2008, pp.36-43.

[6] Anonymous "IEEE Standard for Information Technology- Telecommunications and Information Exchange Between Systems- Local and Metropolitan Area Networks- Specific Requirements Part 15.4: Wireless Medium Access Control (MAC) and Physical Layer (PHY) Specifications for Low-Rate Wireless Personal Area Networks (WPANs)," *IEEE Std 802. 15. 4-2006 (Revision of IEEE Std 802. 15. 4-2003),* pp. 0_1-305, 2006.

[7] P. Kinney, “ZigBee Technology: Wireless Control that Simply Works”, ZigBee Alliance, Oct. 2003. [Online]. Available: http://www.zigbee.org/en/resources/

[8] Anonymous, "IEEE Draft Standard for Information technology--Telecommunications and information exchange between systems--Local and metropolitan area networks--Specific requirements--Part 22.1: Standard to Enhance Harmful Interference Protection for Low Power

Licensed Devices Operating in the TV Broadcast Bands," *IEEE Unapproved Draft Std P802. 22. 1/D6, Feb 2009,* 2009.

[9] L. Benini, E. Farella, C. Guiducci, "Wireless sensor networks: Enabling technology for ambient intelligence", *Microelectronics J.*, Elsevier, vol. 37, no. 12, pp.1639-1649, Dec. 2006.

[10] R. Madan, S. Cui, S. Lall, A. Goldsmith, "Cross-layer design for lifetime maximization in Interference-limited wireless sensor networks", *IEEE Trans. Wireless Commun*., vol. 5, no.11, pp. 3142 – 3152, Nov. 2006.

[11] H. Hassanein, J. Luo, "Reliable energy aware routing in wireless sensor networks", *Second IEEE Workshop on Dependability and Security in Sensor Networks and Syst.*, DSSNS, pp. 54-64, Apr. 2006.

[12] N. Tezcan, E. Cayirci, M. Caglayan, "End-to-end reliable event transfer in wireless sensor networks", *Fifteenth IEEE Int. Symp. on Personal, Indoor and Mobile Radio Commun*., PIMRC 2004, vol. 2, pp 989-994.

[13] Y. Sankarasubramaniam, O. B. Akan, I. F. Akyildiz, "ESRT: End-to-end reliable transport in wireless sensor networks", *IEEE/ACM Trans. on Networking*, vol. 13, no.5, pp. 1003- 1016, Oct. 2005.

[14] J. Zhu, K-L. Hung, B. Bensaou, F. Nait-Abdesselam, "Rate-lifetime tradeoff for reliable communication in wireless sensor networks", *Comput. Netw. J*., Elsevier, vol. 52, no.1, pp. 25 - 43, Jan. 2008.

[15] M. Chen, T. Kwon, Y. Yuan, Y. Choi, V. C. C. Leung, "Mobile agent-based directed diffusion in wireless sensor networks", *EURASIP J. on Advances in Signal Processing*, Hindawi Publishing Corporation, vol. 2007, DOI:10.1155/2007/36871.

[16] D. Georgoulas, K. Blow, “Intelligent mobile agent middleware for wireless sensor networks: A real time application case study”, *Proc. of the 2008 Fourth Advanced Int. Conf. on Telecommun.*, AICT ‘08, pp. 95-100, 2008.

[17] R. W. Thomas, D. H. Friend, L. A. DaSilva, A. B. MacKenzie, “Cognitive Networks: Adaptation and learning to achieve end-to-end performance objectives”, *IEEE Commun. Mag.*, vol. 44, no. 12, pp. 51-57, Dec. 2006.

[18] R. W. Thomas, L. A. DaSilva, A. B. MacKenzie, “Cognitive networks”, *2005 First IEEE Int. Symp. on New Frontiers in Dynamic Spectrum Access Networks*, DySPAN 2005, pp. 352 - 360.

[19] R. W. Thomas, “Cognitive Networks”, Ph.D. Dissertation, Comput. Eng., Virginia Polytechnic and State Univ., Blacksburg, VA, June 15, 2007.

[20] J. F. Lehman, J. Laird, P. Rosenbloom. (2006). “A gentle introduction to SOAR, an architecture for human cognition: 2006 update”, University of Michigan. [Online]. Available: http://ai.eecs.umich.edu/soar/sitemaker/docs/misc/GentleIntroduction-2006.pdf

[21] B. N. Kokinov, “The context-sensitive cognitive architecture DUAL”, *Proc. of the Sixteenth Annual Conf. of the Cognitive Science Society*, Erlbaum, pp. 502–507, 1994.

[22] P. Langley, D. Choi, “A unified cognitive architecture for physical agents”, *Proc. of the Twenty-first AAAI Conf. on Artificial Intell.*, vol. 21. Menlo Park, CA, Boston: AAAI Press, pp. 1469-1474, 2006.

[23] W. Duch, R. J. Oentaryo, M. Pasquier, “Cognitive Architectures: Where do we go from here?”, *Proc. First Conf. on Artificial General Intell.*, Univ. of Memphis, IOS Press, pp. 122-136, Mar. 1-3, 2008.

[24] G.M. Edelman, “Neural Darwinism: Selection and reentrant signaling in higher brain function”, *Neuron, vol.* 10, no. 2, pp 115-125, Feb. 1993.

[25] J.R. Anderson, C. Lebiere, “The Newell test for a theory of cognition”, *Behavioral and Brain Science, vol.* 26, no. 5, pp 587-601, 2003.

[26] S. Franklin, “The LIDA architecture: Adding new modes of learning to an intelligent, autonomous, software agent”, *Proc. of the Int. Conf. on Integrated Design and Process Technology (IDPT),* San Diego, CA: Society for Design and Process Science, Jun. 2006.

[27] D. Cavalcanti, S. Das, J. Wang, K. Challapali, “Cognitive Radio based wireless sensor networks”, *Proc. Seventeenth Int. Conf. Comput. Commun..and Networks*, 2008. ICCCN '08, pp. 1 - 6.

[28] J. Mitola III and G. Q. Maguire, “Cognitive Radio: Making software radios more personal,” *IEEE Personal Commun.*, vol. 6, no. 4, pp. 13-18, Aug. 1999.

[29] O.B.Akan, O. Karli, O.Ergul, “Cognitive Radio sensor networks”, *IEEE Network*, vol. 23, no. 4, pp. 34 - 40, July-Aug. 2009.

[30] W. Youssef, M. Younis, “A cognitive scheme for gateway protection in wireless sensor network”, *Appl. Intell.J.*, vol. 29, no. 3, pp 216-227, 2008.

[31] K. Shenai and S. Mukhopadhyay , “Cognitive sensor networks”, *IEEE Twenty Sixth Int. Conf. on Microelectronics* (MIEL), pp.315-320, May 2008.

[32] E. Gelenbe, P. Liu, B. K. Szymanski, M. Lisee, K. Wasilewski. (2009, Aug. 14) “Cognitive and self-selective routing for sensor networks”, *J. of Computational Manage. Sci*. [Online], Berlin-Heidelberg: Springer, vol. 6, pp. 1-22. Available: www.springerlink.com/index/ j351352425w41v68.pdf

[33] E. Bdira, M. Ibnkahla, “Performance modeling of cognitive wireless sensor networks applied to environmental Protection”, presented at IEEE Globecom, Honolulu, Hawaii, 2009.

[34] Anonymous. (2009, Aug.), “CC2430 Software Examples User Guide”, SWRU178B, Texas Instruments. [Online]. Available: http://focus.ti.com.cn/cn/lit/ug/swru178b/swru178b.pdf

[35] M. Tennenholtz. (2002, Jan. 01), “Game theory and artificial intelligence”, in *Foundations and Applications of Multi-Agent Systems.* [Online], Berlin-Heidelberg:Springer, vol. 2403/2002, pp. 51. Available: www.springerlink.com/content/wj7eqnqgf66y33gu/fulltext.pdf

[36] I. F. Akyldiz, W. Y. Lee, M. C. Vuran, and S. Mohanty, “NeXt generation/dynamic spectrum access/cognitive radio wireless networks: A survey,” *Comput. Network J.,* Elsevier, vol. 50, no.3, pp. 2127–2159, Sept. 2006.

Chapter 3

Cognitive-Node Architecture and Deployment for Future Sensor Networks

Preface

This chapter has been submitted to Wiley's Wireless Communication & Mobile Computing Journal. Details about the conceptual architecture of cognitive nodes for use in future sensor networks, and their implementation details, are presented in this chapter. A grid-based strategy for the deployment of cognitive nodes in a Cognitive Information Centric Sensor Network (CICSN) is proposed. This strategy considers the energy, hardware, and communication costs involved in large-scale deployments of the CICSN for smart city environments.

3.1 Abstract

Cognitive Information-Centric Sensor Network (CICSN) represents the application platform that we propose for future sensor networks. It is comprised of intelligently networked sensors deployed over a large area, supporting multiple application scenarios, while catering to user-desired information quality. In this paper, we propose the use of cognitive nodes (CNs) in the underlying sensor network to provide intelligent information processing and knowledge-based services to the end-users using an information-centric approach. The CNs act on queries from the end-user, containing information about the nature of the request, without needing to know the address of the end node from where data has to be gathered. The main contributions of this paper are firstly, identification of the tools and techniques to implement the cognitive functionality, and secondly, formulation of a strategy for the deployment of CNs in the underlying sensor network in a way that ensures a high probability of successful data reception among communicating nodes. From Matlab simulations, we were able to verify that in a network with randomly deployed sensor nodes, CNs can be strategically deployed at pre-determined positions, to ensure that end-user's quality of information (QoI) requirements are met, even under heavy traffic conditions and high application payloads.

3.2 Introduction

Information Centric Sensor Networks (ICSNs) are a class of context-aware communication networks that provide an infrastructure for knowledge-based intelligent information service to anyone, anywhere and at any time [1]. IEEE 802.15.4/ZigBee based Wireless sensor networks (WSNs) provide the basic infrastructure to deliver sensed information to end-users in diverse application environments such as, agricultural monitoring in rural areas, structural health monitoring of buildings and bridges in urban areas, tracking items in industrial supply chain management applications, detection of forest fires, and even landmines detection in former war zones [2]. These ICSN applications that require a large-scale deployment of the sensor network in order to cover the large target areas and provide more sensing points in the region being monitored, the network topology changes dynamically due node deaths, changing node associations and varying environment conditions, thus affecting the network connectivity and information gathering and delivery capabilities. In addition, the network may have to deal with service requests coming from a variety of end-users including individual consumers, public enterprises, government organizations, and even machines that are information monitoring devices. It is very challenging for sensor networks in their current form, to provide a common platform that can support such diverse ICSN applications, while providing context-aware information to end-users that differ in their requirements on the attributes associated with the service-data such as reliability, latency and throughput. As in the case with Internet of Things (IoT) applications, WSNs are not equipped to handle the heterogeneous traffic, nor do they have adequate capacity to store the large volume of data generated as a result of the multiple requests being serviced by the network and need modifications in the infrastructure to support the functionality [3].

To improve the capabilities of the network that delivers data to IoT environment, we propose the use of Cognitive Nodes (CNs) based on the elements of learning, reasoning and knowledge representation in the underlying network. CNs will provide enhanced capabilities to the WSN to deal with the network connectivity and node dynamics in large-scale deployments. They will also provide space for local storage of data, before data gets delivered to the end user. This will help to maintain their availability at intermediate locations, other than their points of publication (i.e. the sensor nodes) for ease of access. To this end, the main contributions of this paper are as listed below:

i. We provide a description of the conceptual architecture of the cognitive node and the components that constitute its cognitive elements i.e. knowledge representation, learning and reasoning, and describe their functions.

ii. We identify a grid-based deployment strategy for relay and cognitive nodes in the large-scale WSN such that the probability of successful data reception between the communicating nodes is greater than 0.8.

iii. We also calculate the number of relay and cognitive nodes required to cover the target area while ensuring a high probability of successful data reception.

The remaining sections have been organized as follows: In section 3.3 we present the related work. The conceptual architecture of the cognitive node and description of the cognitive elements are provided in Section 3.4. Section 3.5 details the system models, and the deployment strategy for cognitive nodes in the network is presented in section 3.6. Simulation results are presented in section 3.7 before concluding the paper in section 3.8.

3.3 Related work

Sensor node deployment problem has been extensively studied in literature over the past decade. Researchers have considered various factors such as coverage, connectivity, energy-efficiency and fault-tolerance while proposing deployment strategies for sensor nodes (SNs) [6]. With the introduction of the ZigBee standard [11], the focus shifted from sensor node to relay node (RN) placement problem, as the RNs could serve to maintain connectivity of sensor nodes with their base station even when the network size scaled-up [13]. The RNs increased the communication range of SNs and also took over the energy demanding task of data communication within the network from the SNs. This in turn increased the lifetime of the SNs, thus improving the longevity of the network. However, as WSN applications evolved from simple event monitoring or tracking applications to complex applications such as monitoring a coal mine [18], network deployment and its operational complexity increased. The WSN had to not only provide periodically monitored data, but had to even respond to on-demand queries and emergency situations. The changing application requirements made the network traffic very heterogeneous, leading to load balancing issues among the nodes and traffic bottlenecks in the network. Recent research has even considered the use of mobile data collectors, traffic-aware relay node deployment and artificial intelligent (AI) techniques to manage the dynamic network [19]. But data latency and reliability become an issue when mobile data collectors are used, and AI techniques have targeted very specific applications [20]. They have not been architecturally developed and implemented in a way that can be extended to different WSN application platforms. Thus we say that in their current state, WSNs with SNs, RNs and Data Collector nodes will not be able to understand and respond to changing application requirements. The network will not be able to cater to performance attributes of latency, reliability, energy consumption and fault-tolerance while delivering data to the sink. We collectively call these attributes as Quality of Information (QoI) attributes [22], as they represent the attributes that the application layer would associate with the data delivered to the sink, to measure the application-awareness of the response generated by the network to the end user’s request. In order to make the network aware of the changing application requirements, and enable it to provide QoI aware data, we propose the use of special nodes called Cognitive Nodes (CNs) in the underlying WSN. These CNs when strategically deployed in the network will ensure data-delivery with user-desired QoI to the sink, in each round of data transmission throughout the lifetime of the network. We will refer to this network as an Information Centric Sensor network (ICSN) from this point forward, as it will draw on the features of Information Centric Networks (ICNs) in terms of named-data association, in-network caching and the use of CNs as intermediate nodes that will process and store the information within the network [23]. However, we must mention that the idea of named-data association in WSNs is not new. The idea has existed in data-centric sensor networks (DCSNs), which are a special class of WSNs that function as information-retrieval networks rather than serving as point-to-point communication networks [2], [25]. Sensor attributes are used for data gathering and delivery, which makes the use of node addresses inessential. This can lead to huge energy savings for the sensor network, as a single query can be broadcast throughout the network to gather all relevant data from different sources, as against multiple queries addressed to specific locations to gather the same data. This translates to energy savings for all the network nodes, leading to prolonged network lifetime.

Shifting our focus back to the cognitive nodes, the information centric approach to query dissemination used by these nodes helps in finding only relevant data and changes the way the network handles user requests. There is awareness in the network, about the specific information requested by the user; i.e. temperature data or humidity information from a specific geographic area, at a specific time in the present, or from sometime in the past. In addition, the CNs enable the network to understand the QoI with which it is expected to return the requested data to the sink and the network is able to adapt the use of its resources to find paths that are either reliable, have low latency or offer a high throughput. This way, the network is not always exerting itself to find the best path that satisfies all the attributes, but prioritizes the QoI attributes for each transmission round based on the end-user requirements, and finds a suitable path accordingly, thus prolonging the network lifetime. Now the challenge is in finding the best place to deploy these nodes in the ICSN. Optimal node placement is a very challenging problem and has been proven to be NP-hard [6]. With CNs, there are constraints on how many such nodes can be used in the network, and whether one can use only CNs or combine it with the use of RNs in the underlying network.

In this paper, we identify the cognitive functions of the CN, address its deployment problem and through simulations, identify the best combination of RNs and CNs that the network can benefit from to minimize energy consumption and prolong network lifetime, while catering to the QoI attributes of reliability and instantaneous throughput. Thus contributing to not only good quality of user experience, but also improving the lifetime of the network during which data is delivered to the end-user based on user-desired QoI attributes.

3.4 Cognitive Elements and Conceptual Architecture of the Cognitive Node

Cognition is the process by which knowledge is acquired through intuition, perception, planning, and reasoning. If a communication network is capable of observing the impact of its own actions on the environment, and learns to use the knowledge acquired from accumulating these observations to adapt the data delivery paths according to user requirements and changing network conditions, then it is said to be exhibiting cognitive behavior. From this perspective, we identify knowledge representation, learning, and the ability to infer from the knowledge acquired, as elements of cognition sufficient to achieve our proposed goals. Figure 3-1 represents the three major components that we define for our cognitive nodes: learning, reasoning and knowledge representation; and associate them with their respective functions. These elements of cognition, when incorporated in the network nodes of a WSN, help it in better understanding and catering to the requirements of the end-user.

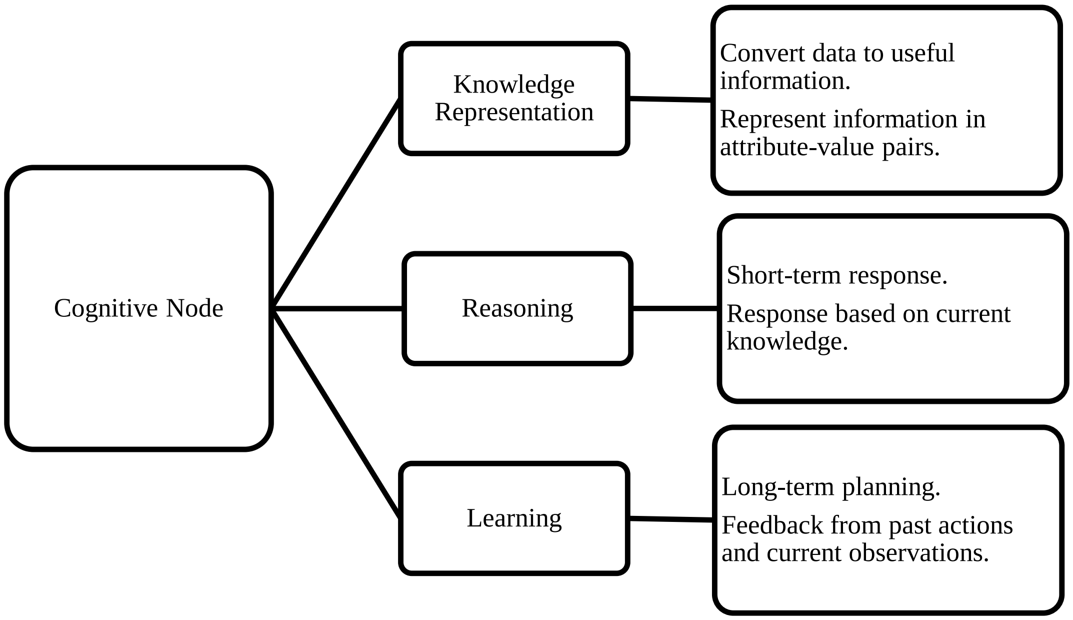


Figure 3-1: Cognitive Node and its elements.

We have briefly introduced cognitive nodes and proposed their node architecture in our work [30]. On similar lines, our expectation from the cognitive elements would be to cater to the following objectives:

i. For the short term: observe current network behavior and respond adaptively to changing network dynamics

ii. For the long term: learn from past behavior and plan for the future so as to make predictions and decisions that positively impact the network survivability and application QoI during its lifetime.

Using these elements, a conceptual architecture of the cognitive node is illustrated in Figure 3-2, which is on the lines of the discussion in our paper [32].

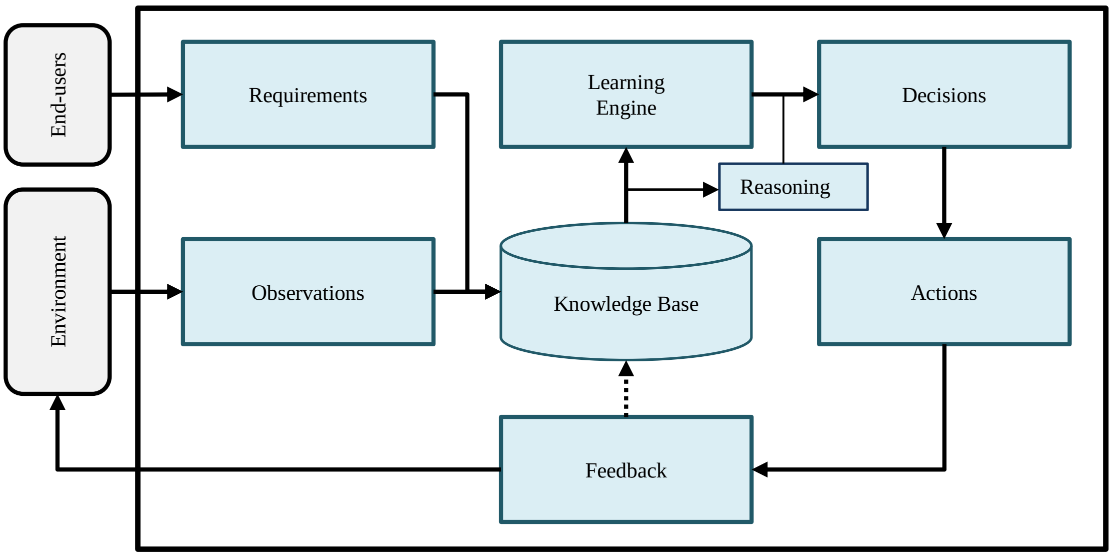


Figure 3-2*:* Cognitive node conceptual architecture.

The ultimate goal to be achieved by CNs is to reduce the periodic data transmission in favor of a planned data collection mechanism that is optimized to match the end-user's requirements while meeting the coverage, connectivity, and longevity requirements of the sensor network. In order to achieve that, the CNs will need to store information about the network status, as well as have a learning mechanism that enables decision making based on past experiences. End-user requirements will be received via a CN's transceiver, while environment observations are to be collected via the sensor units. The two sources of information are consolidated in a knowledge base, which will provide this information to a learning engine. The learning engine analyzes observations and requirements in order to form decisions as to how the requirements will be satisfied given the network's current status and its past response to similar situations from the knowledge base. The outcome is a set of actions that comprise the final plan that will serve a specific request.

The learning engine in a CN will perform various tasks such as data discovery, nodes coordination and task allocation, and prioritized routing as necessary [40]. The actual effects of carrying out the data requests on the underlying network at a given time is fed back to the learning engine in order to fine-tune its performance. Therefore, the CNs' knowledge base will need to keep track of the following information:

- Periodic network status information, such as energy levels, request rates, alive nodes etc.
- Node-specific information, such as named data generated by nodes, locations, and energy profiles.
- Request-specific information, such as energy consumed, delay, routes used, requested data, and QoI metrics.

The CNs will receive data requests from the Base Station (BS) as well as periodic network and data status information from sensor nodes. After this information is analyzed, the CNs will send execution and routing plans to relay nodes, and consolidate results into data reports that are sent back to BSs. CNs therefore act as brokers between the SNs, where data is generated and named, and the BSs, where requests originate and results are consumed. The large-scale deployment of the network will necessitate some form of distributed learning and decision making by the CNs.

3.5 System Models

In this section we describe the network, communication, cost, and energy consumption models of the network used in this work.

3.5.1 Network Model

Nodes in WSNs can be deployed in a flat, hierarchical or geographic location based strategy. In terms of energy conservation, hierarchical deployment strategies provide better performance for

WSNs [2]. We had proposed a hierarchical strategy for cognitive communication in WSNs in our earlier work [30, 31]. We make use of a similar approach here for a large-scale WSN in IoT applications. Cognitive Nodes (CNs), Relay Nodes (RNs), Sensor Nodes (SNs) and a sink node are the node–level entities of the network. CNs act as cluster heads for RNs and are the decision makers for the network. They process the requests received from the end-user by making use of the cognitive elements of knowledge representation, learning and reasoning to identify a data delivery path to the sink that meets the user's Quality of Information (QoI) requirements. RNs act as cluster heads for SNs and also participate in relaying information from CNs to the sink. Sensor nodes gather sensed data and forward it to both RNs and CNs lying within their communication range. The hierarchical deployment strategy helps to distribute the tasks between the relay and cognitive nodes and better manage the network connectivity. We assume that RNs and CNs have the same communication range for a given transmit power. However, the transmit power at RNs is fixed (0dB) at the time of deployment and CNs are allowed to adapt their transmission power from a pre-determined set of values (-3dB to +10dB) to achieve the desired transmission range and QoI.

3.5.2 Energy Consumption model

The voltage discharge characteristics of most Lithium AA batteries (irrespective of their chemistry) suggests that once the terminal voltage drops to about 30% of its original value, almost all of the battery's usable energy is depleted. Lithium batteries typically last 500-1000 cycles before the terminal voltage drops to this value, depending on the application and environment in which it is operated [33]. In our system, we assume that the batteries at RNs and CNs are capable of delivering consistent performance for about 500 cycles, after which they are assumed to be drained out of energy. Rechargeable batteries, or batteries of higher energy rating may be considered for use at CNs when compared to RNs, as their processors may consume additional energy during information processing, adaptive transmission and information feedback for the cognitive decision process. So, the actual cost of RNs and CNs may not be very different, but the energy consumption at CN's could be higher due to the following reasons: (i) CN's ability to adapt transmit power to alter their communication range during the data transmission phase, and (ii) CN's processors operating for longer (when the nodes are active, but not transmitting data), to perform additional computing and information processing required during the cognitive decision making process. If energy consumption at a RN is denoted by E_{RN}, and that at CN is denoted by E_{CN}, then following the energy consumption discussion above, we can say that energy consumption at cognitive node is greater than that at relay nodes, as depicted by Eq. 1.

$$E_{CN} > \quad E_{RN} \tag{1}$$

We set the initial battery cycle-life to 500 units and every time a node is involved in a data or control message communication, we reduce the node's battery cycle-life as shown in Table 3-I, based on the transmit power used for communication.

Table 3-I: Reduction in cycle life based on Transmit power.

P_t	Cycle life reduction (units)
<3dBm	1
3dBm – 5dBm	2
5dBm – 7 dBm	3

3.5.3 Communication Model

In this work, we use the ZunPhy transition region based communication model modified for outdoor environments [34]. Overstepping the binary disc shaped model, the transitional region model identifies a location between the connected and disconnected regions within which the probability of having the received signal strength above a threshold value is above 80%. We use the log normal shadowing path loss communication model, with values for path loss exponent n=4 and standard deviation of the zero-mean Gaussian random variable $X\sigma$, σ=4. The radios communicate in the ISM band at a data rate of 250kbps and the reference distance d0 is 100m. If *d* represents the distance between the transmitter and receiver, then the path loss (PL) at distance *d* is given by:

$$PL(d) = PL(d0) + 10nlog\left(\frac{d}{d0}\right) + X_\sigma \tag{2}$$

$$PL(d0) = 20\log(4\pi d0/\lambda) \tag{3}$$

The received signal strength P_{recv} at distance *d*, for transmit power P_t is given by:

$$P_{recv}(d) = Pt - PL(d) \tag{4}$$

$$\overline{PL(d)} = PL(d0) + 10nlog\left(\frac{d}{d0}\right) \tag{5}$$

$$\overline{P_{recv}(d)} = Pt - \overline{PL(d)} \tag{6}$$

3.5.4 Operational Cost Model

In a WSN deployment consisting of sensor, relay, sink, and cognitive nodes, the sensors used for application-relevant data gathering are typically the most expensive hardware component. Sensors are deployed on sensor nodes which are powered by single-use batteries which are typically not replaceable due to accessibility issues. They operate at low transmit powers (typically less than 3dB) and have a relatively small communication range (few hundred meters), when compared with RNs and CNs. To have the sensors remain operational for the longest duration possible, it is best to let them operate in sleep mode more often as discussed in chapter 2, section 2.5.6, and turn them on only when data needs to be gathered from their surroundings. Comparing relay and cognitive nodes, their hardware costs are very close in terms of the batteries and processors used. However, they differ in their energy consumption model, as discussed in section 3.5.2.

In addition, CNs may incur a slightly higher hardware cost in terms of having an additional flash memory storage, where it can store the data gathered from nearby sensors and relay nodes, and also the information observed from the network interactions. This way, CNs can better perform information caching functions in the information-centric sensor network environment. If we represent the hardware cost of RNs as C_{RN-HW}, and that of CNs as C_{CN-HW}, then following the above discussion, we can arrive at Eq. 7, which suggests that the cost of CN's hardware is equal to or greater than the cost of the relay node hardware, based on the use of the additional flash storage.

$$C_{CN-HW} \geq \; C_{RN-HW} \tag{7}$$

Thus, comparing the energy consumption model and hardware costs of the RNs and CNs from Eq. 1 and Eq. 7, we can say that the cost of operating CNs , or its operational cost (OC_{CN}) is more than the operational cost of RNs (OC_{RN}), as shown in Eq. 8.

$$OC_{CN} > \; OC_{RN} \tag{8}$$

Hence, it is important to consider reducing the number of cognitive nodes used in the deployment strategy, in order to reduce the operational cost of network and its maintenance thereafter.

3.5.5 Problem Definition

For large-scale WSNs, we define the node deployment problem as follows: Determine the number and location for the placement of relay and cognitive nodes in a given target area such that a) The probability that the received signal strength is above a threshold value (-101dBm*) is 0.8 or more b) The network is connected in such a way that there is a path from each SN to the sink through the RNs or CNs at the time of deployment and c) Provide high throughput and reliable data transmission over each hop till data is delivered to the sink.

The probability of the received signal strength being above a threshold value is defined as the probability of successful data reception (Pr). It is a function of the separation distance d between two communicating nodes. When the separation distance is correctly estimated, the probability that the signal strength is above a specified threshold γ_{th} (the receiver's sensitivity for instance) can be estimated using a Q-function (Eq. (9)) based on the work in [35].

$$Q(z) = \frac{1}{\sqrt{2\pi}} \int_{z}^{\infty} \exp(\frac{-x^2}{2})\, dx \tag{9}$$

Using the Q-function, the value of Pr can be estimated using a cumulative density function as follows:

$$\Pr[P_{recv}(d) > \gamma_{th}] = Q[\frac{\gamma_{th} - \overline{P_{recv}(d)}}{\sigma}] \tag{10}$$

As identified in the cost model, we want to minimize the number of CNs and keep their number lesser than the number of RNs to minimize the total cost to the network. We also want to ensure there is at least one RN/CN for each sensor node to deliver its information so that SNs are only involved in short-range, local communications that incur minimum cost. This way the network lifetime depends only on the RNs and CNs. In the following section we identify a strategy for the deployment of CNs and RNs for large-scale ICSN applications.

3.6 Deployment strategy for cognitive nodes

In order to determine the deployment strategy for the router and cognitive nodes in the network, we make the following assumptions:

i. Sensors nodes are deployed randomly, but uniformly throughout the target area. They have a fixed transmit power with a communication range of 175m.
ii. SNs can communicate and bind with RNs and CNs in a single-hop, but do not communicate with each other.
iii. RNs have a fixed transmit power and communication range, but the values are higher when compared with those of SNs.
iv. CNs can vary their transmit power to increase their transmission range to values higher than those of RNs. However, when their transmit power is the same value as that of RNs, they offer the same communication range.
v. We approximate the target area to be a square region, and divide the entire area into smaller squares of side L.
vi. The Sink is deployed at approximately the center of the target area.

Given these assumptions, the goal of the deployment strategy is to identify the length L of the side of each square grid cell and the position of the RNs and CNs on the grid such that the RNs can communicate with at least one CN, and $Pr > 0.8$ along each hop of the data delivery path from a source node to the sink. We device such a node deployment strategy by making use of the following basic properties of a square.

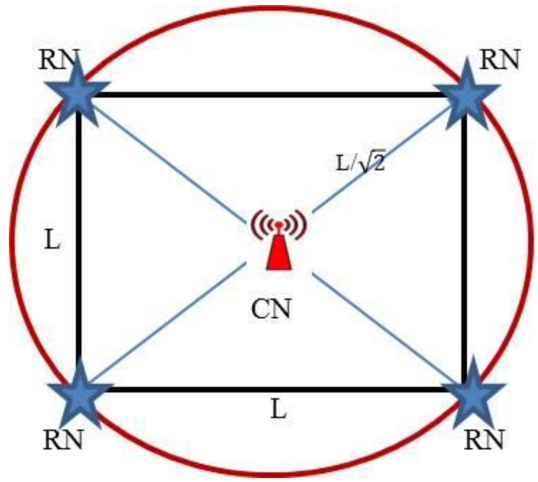


Figure 3-3: Square grid cell of side L having relay nodes (RN) at the corners and cognitive node (CN) at the center.

Property 1. *Center of a square of side 'L' is equidistant from each of its vertices, and has a length* $L/\sqrt{2}$.

For a square of side L, the diagonals intersect at the center of the square, and are perpendicular bisectors of each other. Using Pythagoras's theorem [41], we know that the length of the diagonal which forms the hypotenuse of the isosceles triangle formed by the two sides of the square, can be found as: $\sqrt{(L^2 + L^2)}$= $\sqrt{2}$*L. Thus the center of a square is equidistant from each of its vertices, and has a length $\mathrm{L}/\sqrt{2}$, as shown in Figure 3-3. □

Property 2. *The distance from the center of the square to any of its vertices is the maximum separation distance that can be achieved between the center and any other point on the square*.

If we draw a circle with radius $L/\sqrt{2}$, whose center lies at the center of a square of side L, we would be circumscribing a circle that passes through each of the vertices of the square. This circle does not touch any other point of the square other than the vertices. Thus the distance between the center of the square and its vertices is the maximum separation distance that can be achieved between the vertex and any other point within the square. □

Now let us consider the path loss (*PL)* and received signal strength P_{recv}. From the log-normal shadowing path loss communication model and Eq. 5, we know that the path loss increases as the separation distance between a transmitter and receiver node increases. Conversely, if the distance between two communicating nodes is reduced from $\sqrt{2}$*L across the diagonals of a square of side *L*, to $\mathrm{L}/\sqrt{2}$ at half the distance, the increase in received signal strength P_{recv} can be calculated from Eq. 4 as follows:

$$P_{recv}(\sqrt{2} * \mathrm{L}) - P_{recv}\left(\frac{\mathrm{L}}{\sqrt{2}}\right) = PL\left(\frac{\mathrm{L}}{\sqrt{2}}\right) - PL(\sqrt{2} * \mathrm{L}) \tag{11}$$

Expressing $P_{recv}\left(\frac{\mathrm{L}}{\sqrt{2}}\right)$ in terms of $P_{recv}(\sqrt{2} * \mathrm{L})$,

$$P_{recv}\left(\frac{\mathrm{L}}{\sqrt{2}}\right) = P_{recv}(\sqrt{2} * \mathrm{L}) + 1.5 * n \tag{12}$$

Thus the received signal strength P_{recv} at distance $\frac{\mathrm{L}}{\sqrt{2}}$ is increased by an amount of *1.5*n*, when compared with its value at double the distance $\sqrt{2} * \mathrm{L}$. Using the result from Eq. 12, we can say that for a 2-dimensional square grid cell with transmitter nodes placed at the vertices, an intermediate node placed at the center of the square can improve the probability of reliable reception by 1.5**n*, using a two-hop communication.

Next, we describe the algorithm for deployment of RNs and CNs in the target area.

In Algorithm 1, lines 1 to 5 describe the inputs required to come up with the deployment plan. The size of the target area, number of SNs available and their communication range and the location of the sink are essential to decide on the outputs described in Steps 7 and 8. They are the number and position of RNs and CNs required for maintaining connectivity of the SNs with the sink. In Step9, the receiver sensitivity of the RNs and CNs is set to -101dBm, which is typically the value in commercially available SNs and RNs. In step 10, a threshold value of signal strength γ_{th} is set such that it is 3 dBm above R_{sense}. This is to guarantee reception of signals that are stronger than the receiver's sensitivity, which is the least value of the signal that it can detect. Once these values are set, we use Eq. 10 to plot a graph of the variation of *Pr* as a function of *d* at transmit powers in the range (-5dBm to 10dBm). The transmitter and receiver represent the RNs and CNs. In line 13, the values obtained from the plot are tabulated to strategically identify a value of *d* in Step 14 to ensure that there is at least one RN or CN lying between any two SNs to guarantee connectivity of SNs with the sink across the entire network. In line 15, we ensure that the transmit power for the chosen *d* is able to support $Pr > 0.8$ for every link, at least under near-ideal conditions. Once the side of each square grid is identified in Step 16, Steps 17 – 21 describe the steps to identify the number of rows and columns in the square grid covering the target area and number and position of RNs and CNs in each grid cell. Thus, for SNs placed uniformly, randomly in a target region, Algorithm1 gives the deployment plan for placing RNs and CNs in the area. From this plan, we arrive at L = 350m, Number of CNs = 9, Number of RNs = 16. Figure 3-5 illustrates the deployment of the relay and cognitive nodes in a square grid using this strategy. Summarizing this deployment strategy, we state the following:

For a square target area of side A, and square grid of side L, we have $\sqrt{G}$ number of rows of square grids, G number of cognitive nodes and $(\sqrt{G}+1)^2$ number of relay nodes in the network, where G is A^2/L^2 rounded off to the nearest higher perfect square number.

Table 3-II shows the values of *Pt* and *d* for *Pr* > 0.9 and *Pr* > 0.8 are tabulated using values from the simulation results in Figure 3-4.

Table 3-II: Values of '*d*' for different Transmit powers for *Pr* > 0.8 and *Pr* > 0.9.

Pr > 0.9		*Pr* > 0.8
P_t (dBm)	d (m)	d
-3	150	190
0	200	225
3	250	280
5	275	300
7	300	350
10	360	400

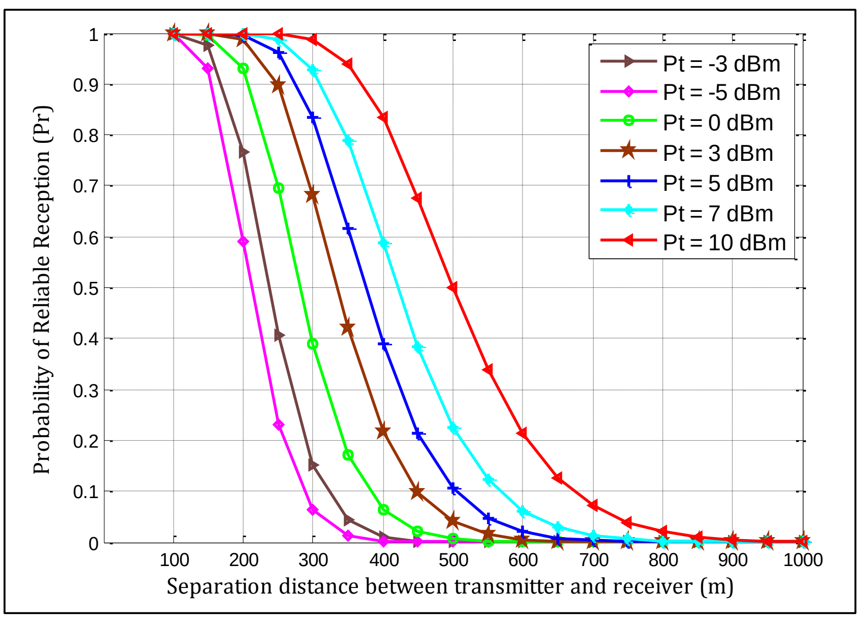


Figure 3-4: Plot of Probability of received signal strength versus separation distance'd'.

Algorithm1: Deployment plan for RNs and CNs in target area

1. **Inputs**:
2. Target area A^2: 1050m x 1050m
3. Number of sensor nodes: 1500
4. Sensor node communication range (r_{SN}): 175m
5. Sink position: Center of target area
6. **Outputs**:
7. Number of RNs and CNs for the target area
8. Position of RNs and CNs in the deployment region
9. **Begin:**
10. **Initialize**: Receiver sensitivity R_{sense}= -101dBm
11. Threshold signal strength γ_{th} = - 98dBm
12. Plot a graph of *Pr* against *d*, for different transmit powers
13. Tabulate values from the plot in Step 12
14. Identify a value of *d* such that $d \leq 2*(r_{SN})$ and $Pr > 0.8$
15. Choose *Pt* such that $0 \leq Pt \leq 10$, for $Pr > 0.8$ at *d*
16. Set *d* as the side *L* of each square grid in the target area A^2
17. Approximate number of square grids required to cover the target area x is A^2/L^2
18. Round off x to the nearest higher whole number G, such that G is the perfect square of a whole number
19. $\sqrt{G}$ is the number of rows and columns of square grids in the target area
20. G will be the total number of CNs in the network, each deployed at the center of a square grid at a distance $L/\sqrt{2}$ from the corners of the grid
21. $(\sqrt{G}+1)^2$ will be the total number of RNs in the network, placed at the corners of each of the square grid

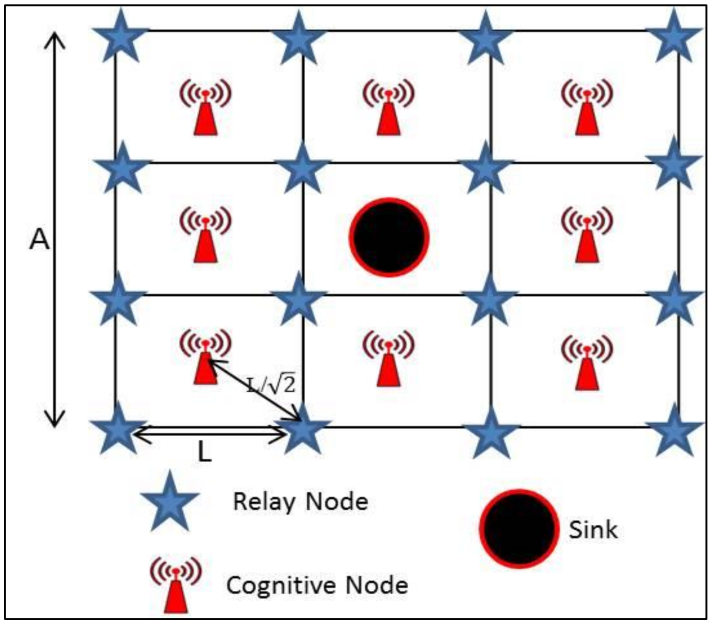


Figure 3-5: Relay and Cognitive Nodes in the grid.

Thus, to cover a square target area of side A, unit square grid cells of side L can be replicated over the entire area to ensure connectivity among all the nodes. Following the steps in Algorithm 1, we know that G is not the least number of such square grid cells required to cover the target area, but a value chosen to retain the approximation of the target area as a square one, while maintaining the size of unit grid cells constant. We know that the separation between nodes at the corners and central node is maximized and the central nodes will be equidistant when juxtaposed with similar cells, due to the properties of the square. Thus, we can say that the number of RNs and CNs identified in Algorithm1 provide a workable deployment plan.

3.7 Simulation results and discussion

Simulation results in this section are from Omnet++ and Matlab simulations. Our initial setup in Omnet++ considered the use of a planned 2D-grid deployment of CNs in a randomly deployed sensor network. No relay nodes were considered in this deployment. The grids were squares of side 200m each. The deployment plan is represented by Figure 3-6. We considered 2 techniques for data routing. The first technique (technique A) is based on the directed diffusion (DD) technique [39], where the data was returned along the same path along which it arrived. This technique was originally proposed for a network containing only sensor nodes. We apply the same strategy to the CNs in our network, with a restriction that the packets can travel only along the sides of the square grid, to maintain the same transmit power at all CNs. The second technique (technique B) allowed the CNs to choose a different data delivery path, compared to the request arrival path. We also let the CNs communicate across the diagonals of the square grid. To increase the probability of reliable data reception at nodes located at different separation distances from the source node, the CNs were allowed to vary their transmit power between 0dBm to 10dBm, with specific increment values of {0, 3, 5, 7, 10} dBm as described in Table 3-II. We let the simulations run till the first CN death for both techniques and found that Technique B is better than Technique A in terms of network lifetime (time to first CN death) and the remaining energy at LCNs after 1st node death. This indicated that technique B would support a more graceful degradation of the network, as the remaining energy at the LCNs was more uniformly distributed after 1st node death, compared to technique A. In addition, we found that for technique B, the time to first CN death depends on the frequency with which requests arrive in the network. Since ICSNs are expected to support multiple applications with varying rates of request arrivals from end-users, the lifetime of the network is completely dependent on the type and frequency of requests being served. Another observation that was made from the simulations, was that allowing communication across the diagonals of the square grid in technique B, helped to reduce the number of hops required to reach the sink from the edge of the network, when compared with technique A where data was routed only along the sides of the square. However, the transmit power had to be increased at the transmitting node to ensure reliability of data arriving at the receiving node. This led to an increase in the hop-over-hop energy consumption across diagonally connected CNs, increasing the overall energy consumption of the network.

To mitigate the need for increasing the transmit power to maintain reliable communication over longer distances, we considered the use of relay nodes (RNs) between diagonally communicating CNs, to serve as a multi-hop communication paths. But adding RNs in the current setup means increasing the total number of nodes deployed in the network, thus increasing the overall hardware cost of the network. So, neither increasing the transmit power at CNs, nor simply adding extra RNs in the deployment plan of Figure 3-6, provide elegant solutions to improving the connectivity, or reducing the energy consumption of the network. From Eq. (8), we know that the operational cost of CN is higher than that of RNs. This indicates that it is feasible to have more number of RNs than CNs in the network. Hence, we adopt the deployment strategy proposed in Algorithm 1 to maintain a lower ratio of CNs to RNs used in the network.

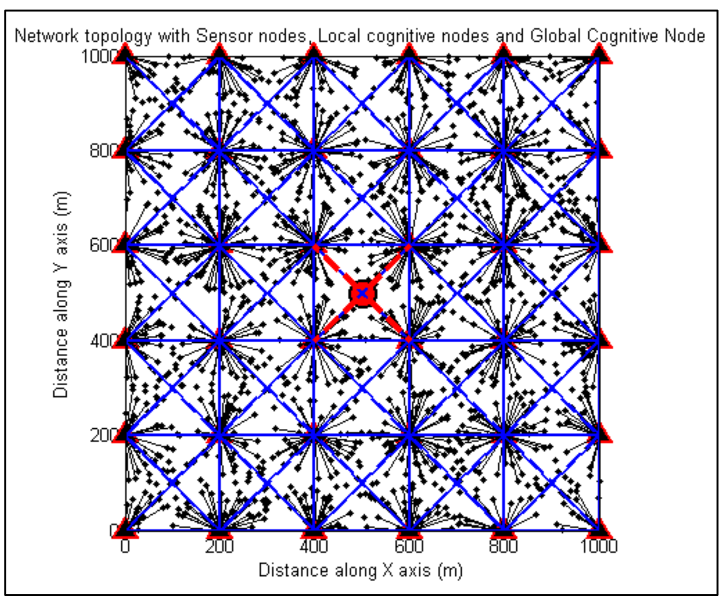


Figure 3-6: ICSN with only CNs and SNs in the network.

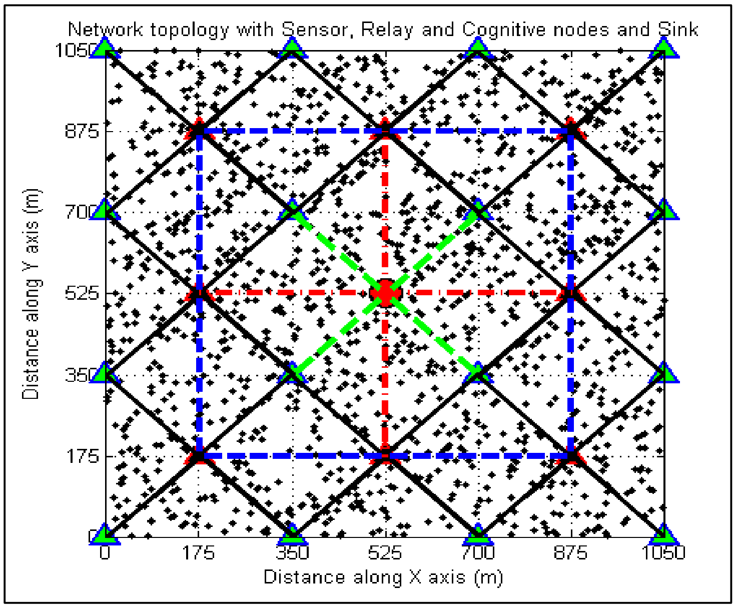


Figure 3-7: ICSN with CNs, RNs and SN in the network.

The transmit power of RNs and CNs, and their respective communication ranges are planned based on Table 3-II and Figure 3-4, and the resulting node deployment plan is as shown in Figure 3-7. This deployment plan was implemented in Matlab as described in the following subsections to allow more flexibility with the parameter setting and control over the CN's behavior during network operation. The simulations were used to study the impact of the node deployment and inter-node interactions on the Quality of Information (QoI) attributes of latency, reliability and throughput. Details of the definitions of the QoI attributes and the simulation setup are provided in the following sections.

3.7.1 The Quality of Information (QoI) attributes

We use Latency, Reliability and Instantaneous Throughput as QoI evaluation metrics for IEEE 802.15.4 PHY – MAC.

3.7.1.1 Node Reliability (NR) at the transmitting node

Node reliability is defined as the probability that a transmitting node is able to successfully deliver a data packet to its next hop neighbor. It is a function of the node's buffer capacity (blocking probability), and the channel conditions (based on the number of nodes trying to simultaneously transmit data) at the time of channel access/data transmission. Thus it inherently reflects upon the link reliability as well. This definition of reliability is based on the work in [36] for low-power nodes in the 802.15.4 PHY-MAC model. We apply the same definition to the cognitive nodes as well, as they will be interacting with sensor and relay nodes in the same setting.

$$NR = \left((1 - P_{blocking}) * (1 - P_{c-fail}) * (1 - P_{p-discard})\right) \quad (13)$$

Where $P_{blocking}$represents the blocking probability due to a buffer-full condition; P_{c-fail} is the channel access failure probability and $P_{p-discard}$ is the probability that a packet is discarded on reaching the maximum number of retries limit.

3.7.1.2 Instantaneous Throughput (IT) at the receiving node

The definition for Instantaneous throughput *(IT)* at a receiving node is based on the work in [37], [38]and is applied to both relay and cognitive nodes. It is defined as a ratio of the size of the frame payload at the physical layer (Overhead + application payload) *L* in bits, over the mean service time *M* in seconds.

$$IT = L/M \text{ bits/s} \quad (14)$$

3.7.1.3 Observed Latency (OL) at the receiving node

The observed latency at a receiving node accounts for delays due to the mean service time at the transmitting node, which is a function of the frame arrival rate.

In the following section, we evaluate the Node Reliability (NR), Observed Latency (OL), and Instantaneous Throughput (IT) for the proposed deployment strategy and communication range after initiating our simulation setups in Matlab.

3.7.2 Simulation setup

Using Matlab (R2013a), we simulated the deployment plan for a large-scale ICSN with 1500 SNs, 16 RNs, 9 CNs and a sink over a square target area of side A = 1050m. The SNs were distributed randomly and uniformly over the target area. CNs and RNs are deployed at fixed, equidistant locations on a 2-dimensional square grid as described in the deployment plan, and shown in Figure 3-5. The RNs are deployed at the corners of each square grid and the CNs at the center. The Sink is deployed at the center of the deployment region. Node connections and interactions are based on a hierarchical ZigBee topology model [9]. The network is built over an IEEE 802.15.4 MAC-PHY simulator based on the work in [36]. Simulation parameters were set as listed in Table 3-III. The reader is referred to the work by Zayani, Gauthier and Zeghlace [37] for more details on the simulation parameters, and the setup of the analytical simulation model in Matlab for the IEEE 802.15.4 MAC-PHY joint simulation model.

The parameters that were varied in our simulation model were: (a) N_ active: the number of nodes attempting to simultaneously transmit data (b) Load: Application payload in terms of the size of the MAC frame payload in bytes, and (c) Per node offered load: per node frame arrival rate expressed as a fraction of the application payload, in bits per second. Impact of variation of these parameters on the QoI attributes of latency (OL), reliability (NR) and throughput (IT), and average throughput were studied. The maximum and minimum possible values for N_ active were identified from the node binding information available at the time of network deployment.

Table 3-III: Simulation Parameters.

Parameter Name	**Value**
Operational frequency	916MHz (ISM band)
Data rate	250kbps
Transmit Power	0dBm for RN; any value from the set {0, 3, 5, 7, 10} dBm for the CN
Modulation	Amplitude Shift Keying
Encoding	Non-Return to Zero
Path loss model	Log Normal Shadowing n=4, σ=4
SN transmission range	175m
Application payload size	0 – 127Bytes

From 10 sets of random deployment of sensor nodes, we found a lower bound of about 10 sensor nodes per CN, and an upper bound of close to 60 sensor nodes per CN, which we used in the simulations. The range of values chosen for the application payload size was 51Bytes – 121Bytes. This range for the MAC frame payload size was chosen based on the size of the data field supported by IEEE802.15.4 Physical layer packets: 0 to 127 bytes [42]. The range of values for per node offered load was 0 to 1400bits per second, such that the load could be expressed as a fraction of the application payload, ranging from 0.1 to 1.4 times the size of the application payload.

3.7.3 Simulation Results

From the simulation results shown in Figure 3-8, we analyze the impact of varying N_active on the QoI attributes for different application payload sizes. From Figure 3-8a, we see an overall trend of increase in latency as the number N_active increases. Latency increases with increase in application payload size, for any value of N_ active. For instance, we see that the latency for N_ active = 10, for a 51B payload, is less than 0.04s. Figure 3-8b indicates an overall trend of decrease in reliability with increase in N_active for any payload size. For a given value of N_active, lower reliability values were observed for higher payloads. For instance, for N_active=20, reliability improves from a value of about 0.4 at a payload of 121B to about 0.8 at 51B, which is almost a 50% improvement in reliability. At higher values of N_active, although the absolute value of reliability is small, the percentage difference between reliability values for different payload sizes remains almost same, especially when compared at 51B and 121B payloads.

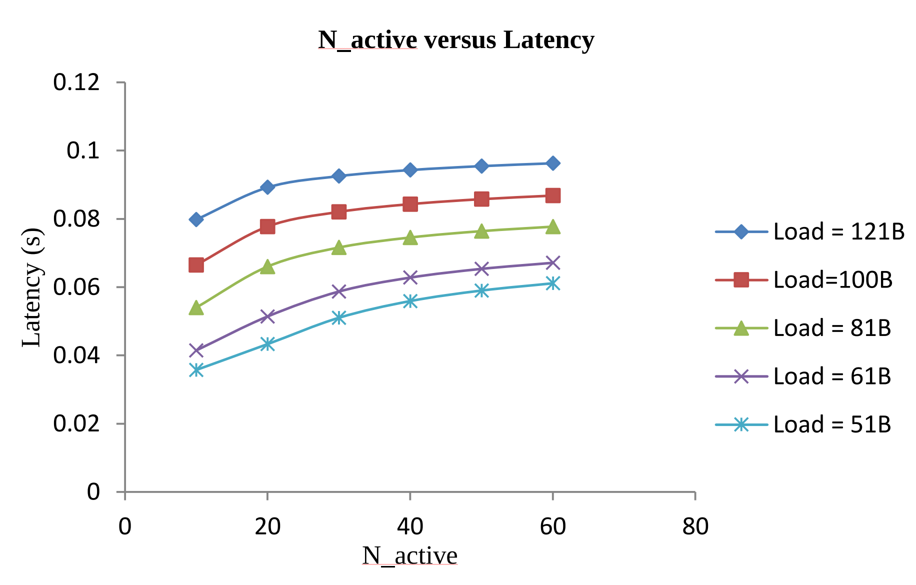


Figure 3-8a: Impact of N_active and Load on Latency.

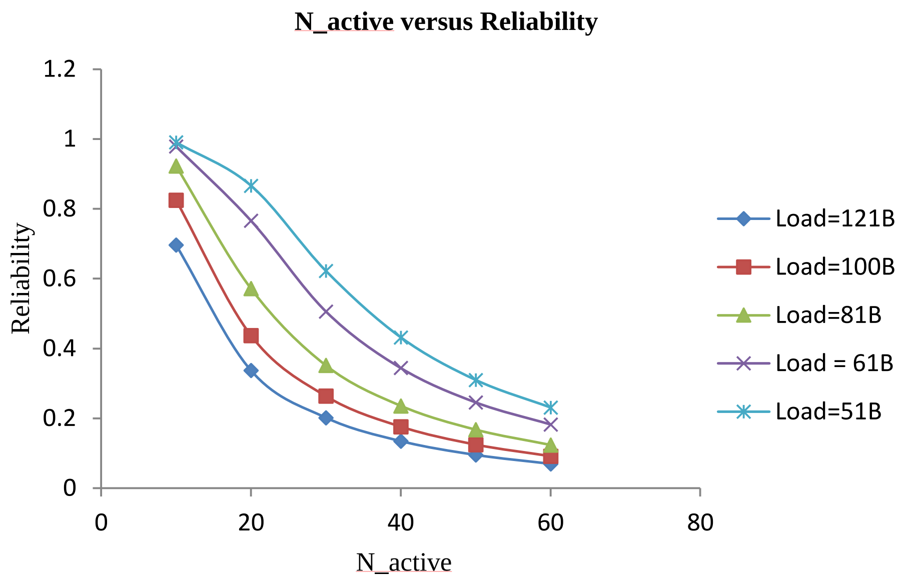


Figure 3-8b: Impact of N_active and Load on Reliability.

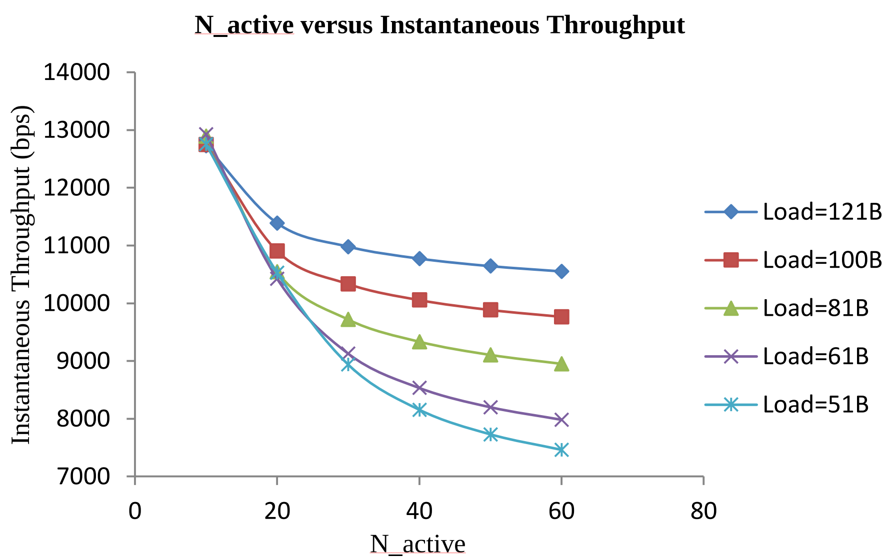
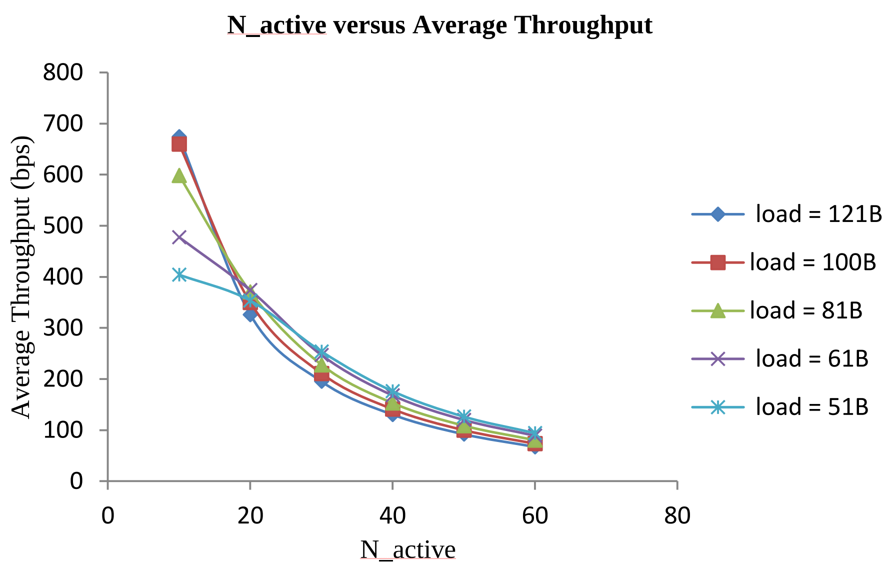


Figure 3-8c: Impact of N_active and Load on Instantaneous Throughput.



3-8d: Impact of N_active and Load on Average throughput.

Figure 3-8: Impact of varying N_active and payload size on QoI attributes.

Figures 3-8c and 3-8d indicate an overall trend of decrease in throughput as N_active increases. However, both average and instantaneous throughput values are higher for higher application payload sizes especially for values of N_active less than 20. Thus, controlling the number of nodes that are scheduled for simultaneous transmission to keep it between 10 and 20 nodes helps to improve the network performance in terms of the QoI attributes, even at high application payloads. Next, we go on to analyze the inter-dependence of the QoI attributes, and the variation of average wait time at a node as the frame arrival rates increase as shown in Figure 3-9 and Figure 3-10. From Figure 3-9 we can see that IT drops to almost half its value of about 5x10^4bps for a frame arrival rate of 6fps, from an original value of 9.8*10^4bps for a frame arrival rate of less than 1fps. The average wait time also sees a steeper increase beyond a 5fps frame arrival rate.

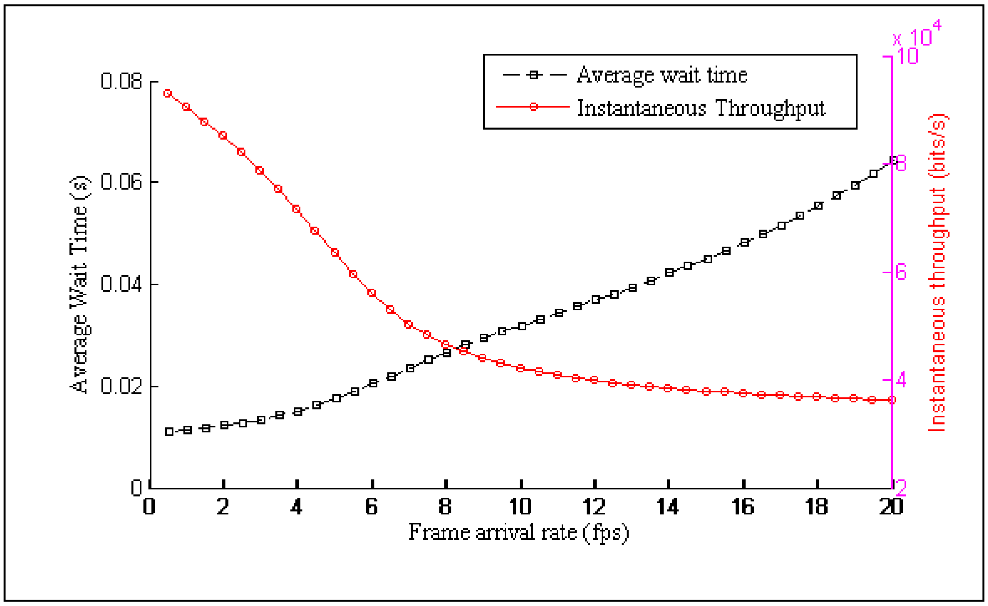


Figure 3-9: Average wait time and Instantaneous Throughput versus per node frame arrival rate.

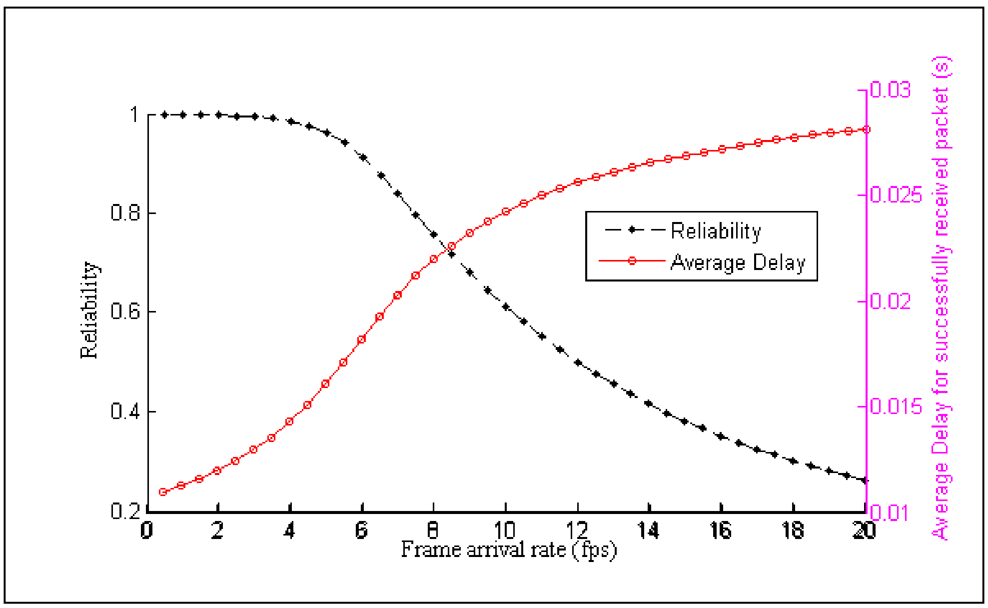


Figure 3-10: Reliability and Average Delay for successfully received packet versus per node frame arrival rate.

So, when increased traffic volumes in the network causes an increase in the frame arrival rate at the network nodes, the performance in terms of the per-node OL and IT degrades. NR follows a similar trend as observed from Figure 3-10. As the frame arrival rate increases beyond 6fps, NR drops below 0.9 and the average delay for successfully received packets is about 75% more; from about 0.01s at 1fps arrival rate to 0.018s at 6fps. In summary, we observe that the performance of each of the QoI attributes deteriorates with increased frame arrival rates at the nodes. But the link quality also affects the network's performance, apart from the per node performance.

From the observed performance of the QoI attributes in large-scale ICSN applications, we see that there is scope for improving the network performance in terms of the network traffic and resource management. In addition, we were also able to identify how the nodes can be scheduled to achieve better performance in terms of the QoI attributes. Thus, we suggest that cognitive nodes can be introduced to better manage the traffic flows and network resources, in a way that is cognizant of the end-user's service requirements. If an application requires high throughput guarantees, then cognitive nodes can exercise their learning mechanism to understand the user requirements and the reasoning mechanism can exercise its control to limit the number of nodes scheduled for simultaneous transmission, to provide the user-desired reliability for the application being serviced. These control instructions from cognitive nodes are passed on to RNs as well and the performance of the entire network can be tuned to meet the application's service requirements.

3.8 Conclusions

WSNs require additional capabilities in terms of being able to understand the network and users' dynamic behavior and store the large volume of generate data in the CICSN infrastructure. The network is dynamic due to changing application requirements, end-user behavior and network topology. To address these challenges, we proposed the introduction of cognitive nodes which implement learning, reasoning and knowledge representation as elements of cognition, in large-scale CICSN applications. We proposed the cognitive node architecture, and described the interactions among the cognitive elements and how they would impact the network performance. The reasoning mechanism responds to current network conditions that require immediate attention; while the learning mechanism uses the knowledge acquired during the network operation over a period of time, for planning and controlling the response to predicted network conditions.

Next, based on the energy consumption and hardware cost of the relay and cognitive nodes, the number and position of RNs and CNs to be used in the network was determined by the deployment plan. The proposed deployment plan was evaluated in terms of its performance of the QoI attributes of latency, reliability and throughput. We found that if the CNs could exercise control over the number of nodes that are scheduled for simultaneous transmission, and keep the number below 20 nodes, the QoI attributes' performance could be maintained at reasonable values for the proposed deployment plan and transmit power capabilities. In addition, the inter-dependence among the QoI attributes was also studied, so that the performance trade-offs can be adapted based on the end-user's requirements of the QoI attributes for the different traffic types.

In our future work, we will explore how the CN design and the proposed deployment plan impact the network longevity, and connectivity of the sensor nodes with the sink during network operation. Since the CICSN is expected to serve as a platform that supports different applications and diverse end-users, the success of the CICSN platform will be decided by its ability to adapt the data delivery decisions to changing network and traffic conditions, and user requirements. So, our next step is to evaluate the performance of the CICSN platform under different traffic conditions, and observe if the CNs are able to deliver data to the sink with user-desired QoI. In addition, the energy consumption pattern during network operation will be studied to see if a re-allocation of the initial energy at the CNs or RNs is required, and if the deployment plan can be improvised to better support larger areas and mobility in sensor and cognitive nodes.

3.9 References

[1] ITU-T technology watch briefing report series, No.4, “Ubiquitous Sensor Networks”, Feb. 2008.

[2] I. Stojmenović and S. Olariu”, Data-Centric Protocols for Wireless Sensor Networks," *Handbook of Sensor Networks,* I. Stojmenovic, ed., John Wiley &, Sons, pp. 417-456, 2005.

[3] A. Al-Fagih, F. Al-Turjman, W. Alsalih and H. Hassanein, “A priced public sensing framework for heterogeneous IoT architectures,” *IEEE Transactions on Emerging Topics in Computing,* vol. 1, no. 1., pp. 133-147, June 2013.

[4] F. Al-Turjman, A. Alfagih, W. Alsalih, and H. Hassanein, “A delay-tolerant framework for integrated RSNs in IoT”, *Elsevier: Computer Communications Journal*, vol. 36, no. 9, pp. 998–1010, May, 2013.

[5] A. Alfagih, F. Al-Turjman, and H. Hassanein, “Ubiquitous Robust Data Delivery for Integrated RSNs in IoT”, *In Proc. of the IEEE International Global Communications Conf. (GLOBECOM’12),* Anaheim, California, 2012, pp. 298-303.

[6] Z. Yun, X. Bai, D. Xuan, T. H. Lai, and W. Jia, “Optimal Deployment Patterns for Full Coverage and k-connectivity (k≤ 6) Wireless Sensor Networks”, *IEEE/ACM Transactions on Networking (TON)*, vol. 18, no. 3, pp. 934-947, 2010.

[7] P. Cheng, C. N. Chuah, and X. Liu, "Energy-aware Node Placement in Wireless Sensor Networks", *Global Telecommunications Conference, 2004. GLOBECOM'04. IEEE* (Vol. 5, pp. 3210-3214). IEEE. (2004, November).

[8] M. Cardei, M. T. Thai, Y. Li and W. Wu, "Energy-efficient Target Coverage in Wireless Sensor Networks", In *INFOCOM 2005. 24th Annual Joint Conference of the IEEE Computer and Communications Societies. Proceedings IEEE*, vol. 3, pp. 1976-1984). Mar. 2005.

[9] F. M. Al-Turjman, A. E. Al-Fagih, H. S. Hassanein, and M. Ibnkahla, "Deploying Fault-tolerant Grid-based Wireless Sensor Networks for Environmental Applications", *2010 IEEE 35th Conference on Local Computer Networks (LCN),* pp. 715-722, Oct. 2010.

[10] L. Tran-Thanh and J. Levendovszky, "A Novel Reliability Based Routing Protocol for Power Aware Communications in Wireless Sensor Networks", Proceedings of the 2009 IEEE conference on Wireless Communications & Networking Conference, pp. 2308-2313, April 05-08, 2009.

[11] A.Tufail, "Reliable Latency-Aware Routing for Clustered WSNs," International Journal of Distributed Sensor Networks, vol. 2012, Article ID. 681273, 6 pages, 2012.

[12] ZigBee Specifications. [Online]. Available: http://www.zigbee.org ZigBee Document 053474r17, Jan.2008.

[13] X. Cheng, D.Z. Du, L. Wang and B. Xu, "Relay sensor placement in wireless sensor networks", ACM/Springer Journal of Wireless Networks, vol. 14, no. 3, pp. 347-355, June 2008.

[14] X. Han, X. Cao, E. L. Lloyd and C. C. Shen, "Fault-tolerant Relay Node Placement in Heterogeneous Wireless Sensor Networks", *IEEE Transactions on Mobile Computing, vol. 9, no.* 5, pp. 643-656, 2010.

[15] E. L. Lloyd and G. Xue, "Relay Node Placement in Wireless Sensor Networks", *IEEE Transactions on Computers, vol. 56, no.*1, pp. 134-138, 2007.

[16] K. Xu, H. Hassanein, G. Takahara, and Q.Wang, "Relay Node Deployment Strategies in Heterogeneous Wireless Sensor Networks", *IEEE Transactions on Mobile Computing, vol. 9, no.* 2, pp. 145-159, 2010.

[17] F.M. Al-Turjman, H. S. Hassanein and M. Ibnkahla, "Optimized Relay Placement to Federate Wireless Sensor Networks in environmental applications", Wireless Communications and Mobile Computing Conference (IWCMC), 2011 7th International , pp. 2040-2045, 4-8 July 2011.

[18] M. Li and Y. Liu, "Underground Coal Mine Monitoring with Wireless Sensor Networks". *ACM Trans. Sen. Netw., vol.* 5, no. 2, Article 10, 29 pages, Apr. 2009.

[19] F. Wang, D. Wang, and J. Liu, "Traffic-aware Relay Node Deployment: Maximizing Lifetime for Data Collection Wireless Sensor Networks", *IEEE Transactions on Parallel and Distributed Systems, vol.* 22, no. 8, pp. 1415-1423, 2011.

[20] Y. B. Reddy and C. Bullmaster, "Application of game theory for cross- layer design in cognitive wireless networks," in Proc. 6th Int. Conf. Inform. Technology: New Generations, ITNG, 2009, pp. 510–515.

[21] L. Reznik and G. Von Pless, "Neural networks for cognitive sensor net- works," in Proc. IEEE Int. Joint Conf. Neural Network., IJCNN 2008, Jun. 2008, pp. 1235–1241.

[22] C. Bisdikian, L. M. Kaplan and M. B. Srivastava, "On the Quality of Information in Sensor Networks", ACM Trans. Sensor Netw, Vol. 9, no. 4, Article 48, July 2013.

[23] B. Ahlgren, C. Dannewitz, C. Imbrenda, D. Kutscher and B. Ohlman, “A Survey of Information-Centric Networking”, Communications Magazine, IEEE , vol.50, no.7, pp.26-36, July 2012.

[24] F. Al-Turjman and H. Hassanein, “Enhanced Data Delivery Framework for Dynamic Information-Centric Networks (ICNs)”, In Proc. of the IEEE Local Computer Networks (LCN), Sydney, Australia, 2013.(Accepted).

[25] B. Krishnamachari, D. Estrin, and S. Wicker, “Modelling DataCentric Routing in Wireless Sensor Networks”, IEEE Infocom, vol. 2, pp. 39-44, Jun. 2002.

[26] J. Heidemann, F. Silva, C.Intanagonwiwat, R. Govindan, D. Estrin and D. Ganesan, “Building Efficient Wireless Sensor Networks with Low-Level Naming”, 18th ACM Symposium on Operating Systems Principles, Oct. 21-24, 2001.

[27] W. R. Heinzelman, A. Chandrakasan, and H. Balakrishnan, “Energy-Efficient Communication Protocol for Wireless Microsensor Networks”, *IEEE Proc. Hawaii Int',l. Conf. Sys. Sci.*, pp. 1-10, 2000.

[28] W.B. Heinzelman, A.P. Chandrakasan and H. Balakrishnan, “An Application-specific Protocol Architecture for Wireless Microsensor Networks”, *Wireless Communications, IEEE Transactions on* , vol. 1, no. 4, pp. 660-670, Oct 2002.

[29] Q. Zhao, L. Tong and Y. Chen, “Energy-aware data-centric MAC for Application-specific Sensor Networks”, Statistical Signal Processing, 2005 IEEE/SP 13th Workshop on , pp.1238-1243, 17-20 July 2005.

[30] G. Vijay and M. Ibnkahla, “CCAWSN: A Cognitive Communication Architecture for Wireless Sensor Networks”, Communications (QBSC), 2012 26th Biennial Symposium on, pp.132-137, 28-29 May 2012.

[31] G. Vijay, E. B. Bdira and M. Ibnkahla, "Cognition in Wireless Sensor Networks: A Perspective", Sensors Journal, IEEE , vol.11, no.3, pp.582,592, March 2011.

[32] G. T. Singh, M. Abu-Elkheir, F. Al-Turjman, A-E.M. Taha, "Towards prolonged lifetime for large-scale Information-Centric Sensor Networks," *Communications (QBSC), 2014 27th Biennial Symposium on* , pp.87-91, 1-4 June 2014.

[33] Lithium Based Batteries. [Online]. Available: http://batteryuniversity.com/learn/article/lithium_based_batteries

[34] M. Zuniga and B. Krishnamachari, "Analyzing the Transitional Region in Low Power Wireless Links", *Sensor and Ad Hoc Communications and Networks, First Annual IEEE Communications Society Conference on , IEEE SECON 2004.* pp. 517-526. Oct. 2004.

[35] T. S. Rappapport. "Wireless Communications: Principles and Practice". Prentice Hall.

[36] M-H. Zayani, and V. Gauthier, "Usage of IEEE 802.15.4 MAC–PHY Model", Online: http://www-public.it-sudparis.eu/~gauthier/Tools/802_15_4_MAC_PHY_Usage.pdf

[37] M.-H. Zayani, V. Gauthier, and D. Zeghlache, "A Joint Model for IEEE 802.15.4 Physical and Medium Access Control Layers", In proc. of IEEE The 7th International Wireless Communications and Mobile Computing Conference (IWCMC 2011), 2011.

[38] P. Park, P. Di Marco, P. Soldati, C. Fischione, and K.H. Johansson, "A generalized Markov chain model for effective analysis of slotted IEEE 802.15.4", *Mobile Adhoc and Sensor Systems, 2009. MASS '09. IEEE 6th International Conference on* , vol.130, no.139, pp. 12-15 Oct.2009.

[39] C. Intanagonwiwat, R. Govindan and D. Estrin, "Directed Diffusion: A Scalable and Robust Communication Paradigm for Sensor Networks", *Proceedings of the 6th annual international conference on Mobile computing and networking*. ACM, 2000.

[40] Dinh, Ngoc-Thanh; Kim, Younghan, "Potential of information-centric wireless sensor and actor networking," Computing, Management and Telecommunications (ComManTel), International Conference on, pp.163-168, Jan. 2013.

[41] Anonymous, “Pythagorean Theorem”, [Online]. Available: http://mathworld.wolfram.com/PythagoreanTheorem.html

Chapter 4

A Data Delivery Framework for Cognitive Information-Centric Sensor Networks in Smart Outdoor Monitoring

Preface

This chapter has been accepted for publication in Elsevier's Computer Communication Journal. Details of the COGNICENSE framework used for introducing cognition in information centric sensor networks, the data delivery strategy in the multi-user smart environment, and the associated simulation results have been provided in this chapter. The simulation setup and parameters used are the same as used in Chapter 3, and the reader is referred to section 3.7.2 for complete details. Mathematical details about the Analytic Hierarchy Process are presented in Appendix I of this thesis.

4.1 Abstract

Cognitive Information-Centric Sensor Networks represent a paradigm of wireless sensor networks in which sensory information is identified from the network using named-data, and elements of cognition are used to deliver information to the sink with quality that satisfies the end-user requirements. Specialized nodes called Local Cognitive Nodes (LCNs) implement knowledge representation, reasoning and learning as elements of cognition in the network. These LCNs identify user-requested sensory information, and establish data delivery paths to the sink by prioritizing quality of information (QoI) attributes (e.g., latency, reliability, and throughput) at each hop based on the network traffic type. Analytic Hierarchy Processing (AHP) is the reasoning tool used to identify these paths based on QoI-attribute priorities set by the user. From extensive simulations, parameters that can be controlled to improve the values of QoI attributes along each hop were identified, and performance of the AHP based data delivery technique was compared with two traditional data-centric techniques in terms of the number of transmission rounds and QoI attribute performance. It was found that the use of cognition improves the number of successful transmissions to the GCN by close to 30%, while closely adapting the data delivery paths to the QoI requirements of the user.

4.2 Introduction

Wireless Sensor Network (WSN) applications have evolved from catering to application-specific requirements, to supporting large scale application platforms such as smart cities and Smart Outdoor Monitoring (SOM) in public sensing [1]. These applications typically require a large scale, dense deployment of the sensor network, which generates a large amount of data. However, end-users may be interested in accessing specific information from the network (such as temperature in the north-east region of deployment, or issue pollen alerts for people with allergies). These ‘smart’ application platforms require the underlying WSN to not only gather information from the relevant information sources, but also prioritize and efficiently manage the heterogeneous traffic flows generated by the requests, and deliver information with quality that satisfies the end-user’s requirements in terms of attributes such as reliability and latency. Providing a good quality of experience to end-users in such large-scale deployments requires a shift in focus from traditional address-centric communication abstractions to data-centric routing and storage, where information from multiple, concurrent information sources produced anywhere in the network can be coherently delivered to the end-user.

Information Centric Network (ICN) is one such paradigm that focuses on content delivery, rather than the point-to-point information flow in the network [2, 3]. It makes use of “named data objects” instead of IP addresses to gather data, thus decoupling information source from its location or node identification. ICN is touted as the future technology for content delivery over the internet because of its ability to bring information to the network layer to improve communication efficiency. Moreover, using the information-centric approach in such a resource rich, static environment, positively impacts data delivery to the end-user. Data-Centric Sensor Networks (DCSNs) [4-8] are a parallel paradigm in WSNs where attribute-value pairs are used for named identification of sensed data. Although DCSNs existed much before ICNs, the limited resource and energy capabilities of sensor nodes, and their inability to adapt data delivery decisions to the dynamic network conditions decreased the popularity of this approach in WSNs. Later, with the introduction of the ZigBee standard [9], most of the data processing and communication tasks were off-loaded to relay nodes. However, this also led to a shift to a more address-centric approach for WSNs. Then, with need to enhance the multi-objective optimization and dynamic decision making capabilities of the network, increased research activity in the field of applying cognition to sensor networks. These cognitive sensor networks were able to achieve various goals such as making the sensor network aware of user requirements, reduce network resource consumption, and make the network exhibit self-configuration, self-healing and self-optimization properties [10-12]. Despite these advances, it still remains a challenge for sensor networks to differentiate traffic flows in smart environments, where the user requirements change over time. Sensor networks still lack the ability to adapt data delivery techniques to different traffic flows generated by the network. In addition, it is desirable to have the sensor network functioning as an information gathering network, to make it easier for users to make name-based requests, and for ease of adaptability to the future ICN.

To cater to all these requirements, we put together the idea of an information-centric approach from ICNs/DCSNs, along with the concept of cognition in this paper, and propose a Cognitive Information Centric Sensor Network (ICSN) framework-COGNICENSE. The information centric strategy is used to identify relevant sensed information from the network, and the elements of cognition (i.e. knowledge representation, reasoning and learning) are implemented at special nodes called Local Cognitive Nodes (LCNs) and Global Cognitive Nodes (GCNs), to enhance their information processing and intuitive decision making capabilities. GCNs interpret the user request for the network, and the LCNs help to identify appropriate return paths for data delivery. Relay nodes participate in information transmission over multiple hops, thus maintaining the network's scalability. End-user satisfaction is based on the Quality of Information (QoI) delivered to the sink [13, 14], characterized by the attributes of latency, reliability, and throughput associated with the application specific traffic. Accordingly, we summarize our contributions in this paper as follows:

i. We propose a framework called COGNICENSE that makes use of elements of cognition and an information-centric approach for data delivery in WSN applications for Smart Outdoor Monitoring (SOM).

ii. We investigate three Quality of Information (QoI) attributes: latency, reliability and throughput. Based on simulations considering an IEEE 802.15.4 PHY-MAC model, we identify the parameters that affect these QoI attributes.

iii. Using a multi-criteria decision making (reasoning) technique called Analytic Hierarchy Process (AHP), we show how the values of the QoI attributes obtained from the simulations can be used to make decision choices about the data delivery path that provides the best value of information at the sink (end-user).

The rest of the chapter has been organized as follows: Section 4.3 reviews related work in literature. Section 4.4 provides the system models and problem description. Section 4.5 provides details about the proposed data delivery framework using elements of cognition, i.e. knowledge representation and inference. Section 4.6 provides simulation results and discussions, and we conclude the chapter in section 4.7.

4.3 Related work

The idea of focusing on information objects rather than the host of the information in communication networks is hardly new. Data-centric sensor networks in the wireless world and the TRIAD project [15] for the internet, described early forms of information centric networks, that aim to move away from the end-to-end communication paradigm and focus on the content being delivered to the end user. In this section, we review DCSNs, and ICNs with respect to their network and design components, and implementation challenges. We also explore the use of cognition in wireless networks with respect to their ability to enable networks to adapt to changing environment conditions, and cater to end-user requirements as they evolve with the applications.

4.3.1 Information Centric Networks

Information Centric Network is an information-oriented communication model proposed for the future internet, to help with managing the huge amount of IP traffic being exchanged globally. Unlike traditional host-centric networks where data routing requires the establishment of single end-to-end path to the host, ICNs decouple senders and receivers by leveraging in-network caching [16, 17] and replication of data. User requests for named data objects are addressed irrespective of the source of the publisher or the content's location. This is facilitated by the use of intermediate nodes, which are in-network devices that process and cache named data objects. Thus named data access, routing of requests and data, and information caching comprise the important features of ICNs, and the intermediate nodes play a very important role in implementing these features. These nodes will need to make smart decisions to coordinate their actions and decisions across the network, and also adapt to services and applications as they evolve. Despite the various ongoing research activities in ICNs, not much work is being done with regards to empowering the intermediate nodes to adapt dynamically to changes in the network and end-user behavior, to help them learn and evolve on their own.

4.3.2 Data-Centric Sensor networks

The DCSN approach is very similar to ICNs, in naming the sensed objects and in caching data as it if forwarded to the sink. One of the striking differences between DCSNs and ICNs in terms of the network components is that the DCSNs approaches consider only 2 types of devices in the network – sensor nodes and sink, whereas ICNs typically use 3 types of devices – publishers, subscribers and intermediate nodes. Some DCSNs do propose choosing sensor nodes as cluster heads and involve them in routing data to the sink [18], but this approach burdens the sensor node in terms of energy, data processing and memory capacities and affects the network lifetime and performance on the whole. What has not been explored much in DCSN is applying the ZigBee network model for DCSNs. ZigBee routers are a better choice in terms of conserving sensor's energy and making routers available for more functions such as information processing, routing and data caching. ZigBee topology is a big energy saver in terms of off-loading the burden from sensor nodes. Another aspect that has not been explored much in DCSNs is the ability to deal with heterogeneous traffic flows generated in the network as a result of the different request that the network receives. The request could be event-driven, time-driven, query-driven or a mix of any of these types [19]. Most DCSNs deal with one type of traffic, typically query-driven traffic. However, the challenge is in enabling the network to deal with all types of requests and provide satisfactory service to the end-user while adapting to changing network conditions and application requests at the same time [20]. But just as the case with intermediate nodes in ICNs, routers in DCSNs would be burdened with too many responsibilities, if they had to carry out all these function and are not empowered with techniques to deal with them effectively. Hence we look at the possibility of introducing cognition in the routers of the DCSNs.

4.3.3 Cognition in Communication Networks and Cognitive Sensor Networks

To understand the correlation between cognition and communication networks, we'll start with the way wired and wireless communication network architectures have been standardized: the layered protocol stacks of the OSI and TCP-IP models, and the 802 series specifications. As network sizes grew, it became challenging to correlate information from different parts of the network, and make decisions with incomplete or inconsistent information from different layers of the protocol stack. So the concept of a knowledge plane was proposed by Clark *et al*. [21] for the wired world, to break the barriers of the layered architecture and enable seamless communication across the layers of the protocol stack and across the network. This idea from the wired world was adopted into wireless networks by Thomas *et al*. [22], who proposed the idea of a Cognitive Network. This network would be aware of the application requirements as well as the network dynamics, and make use of learning, reasoning and feedback from past interactions to make decisions that improve both network performance and end-user satisfaction. The feedback in the network is based on an Observe-Analyze-Decide-Act loop [23], which when combined with learning and reasoning constituted the idea of cognition in the cognitive network. This concept of cognition has been extended to WSNs as well [24], which we will collectively refer to as cognitive sensor networks (CSNs) in this work. But these architectures and applications are address-centric, which cater to the end-to-end communication paradigm. To the authors' best knowledge, information-centric architectures (ICNs and DCSNs) have not leveraged the idea of cognition, in the way we have described above to handle diverse traffic flows and satisfy end-user requirements simultaneously. Specifically, cognition in data-centric sensor networks can provide the following benefits: (i). In-network information processing (aggregation) can save the energy expended on the huge amount of data exchanged within the network before being delivered to the sink, and (ii). Using intermediate nodes that incorporate cognition can reduce the burden on sensor nodes and make smart data delivery decisions based on evolving application requirements, and changing environment conditions. Table 4-I shows a comparison of the infrastructure and data-delivery techniques used in DCSNs, ICNs, and CSNs.

Table 4-I: Comparison of Infrastructure and Data Delivery techniques in DCSNs, ICNs and CSNs.

	DCSN	ICN	CSN
Network components	Sensor nodes (SNs) and Sink node(s). SNs participate in sensing, transmission, and even data aggregation when they function as cluster heads. Sink nodes disseminate request, store data returned from network, process stored data to respond to user queries, and manage network topology.	Publisher, Subscriber and Intermediate nodes. Publishers only publish the information. Intermediate nodes deliver published information to the Subscriber. Senders and receivers are decoupled.	Typically address-centric sensor networks with sensor nodes, relay nodes (RNs) and a Sink node or Base Station. In ZigBee based networks, SNs gather sensed data, transmit to RNs only. RNs participate in multi-hop transmission to Sink. Intelligent agents modelled as software agents within network nodes.
Node deployment and Control	Typically self-organizing. SNs randomly deployed. Dynamic network with Centralized control and decision making at Sink.	The ICN environment is a static, resource rich environment for wired communication networks.	Random, deterministic or mixed deployment for network nodes in a dynamic network environment. Distributed control through intelligent agents within the network.
Request dissemination	Requests are sent out in attribute-value pairs from the sink, which are disseminated in the network through flooding, multicasting or geocasting or some combination of multicasting and flooding	Name-resolution (content name is resolved into components to identify locators for request), or name- based routing (request forwarded based on identifier name)	Request dissemination is mostly address centric, containing node addresses or end-point ids from where data is to be fetched, for end-to-end communication.
Data gathering / aggregation	Typically along reverse paths of memorized links, established during request dissemination through broadcast trees; using chains of reporting sensor nodes or through token circulation among equally probable next hop nodes. Data may or may not be aggregated	ICNs explore prefix aggregation, request aggregation and aggregation of routing information for functions such as load balancing, and better routing scalability.	Most implementations of CSNs do not depend on or focus on data aggregation methods, or the benefits it can offer. However, data may be aggregated in dense deployments. The cognitive agents focus more on achieving various

	depending on correlation of observed data. Minimum spanning trees are constructed for aggregating data at specific nodes before forwarding them for reporting.		objectives such as reduced resource consumption, enabling self-organizing and self-healing capabilities of the network and QoS routing under diverse application scenarios.
Cache storage and replacement	Information sensed from a given region may vary over time. Hence stored data may become stale and provide inaccurate information to users demanding current information. Hence responding to query requires awareness of its type in order to generate useful responses from the network. This traffic classification, and cache replacement policies suitable for such environments do not currently exist.	Caching is inherent in the architecture. Published data doesn't vary over time. Hence cached information can be reused any number of times and improves network performance over time, as data becomes available from caches closer to the subscriber than the original publisher.	Data storage aspects have not been explored by intelligent agents of CSNs.
Scalability	Scalability and communication range are limited by the use of resource constrained sensor nodes in the network.	The information centric approach has been proposed to overcome the limitation imposed by IP addressing, for improved scalability.	Since CSNs are based mostly on ZigBee based communication, scalability is not an issue. RNs provide multi-hop communication over long distances.
Limitations/ Challenges	Energy consumption and delay involved in data processing, aggregation and delivery. Resource limitations at sensor nodes hinder implementation of advanced routing algorithms and limit caching.	Privacy issues, scalability in caching, cost efficiency.	Cognition has not been explored in a way that can be applied to sensor networks at an architectural level. Implementations are very application/goal specific.

To this end, the COGNICENSE framework we propose will be able to deal with changing application requirements, and make smart decisions to provide the requested information to end-users with quality that satisfies the SOM application requirements. SOM applications are challenging to handle in terms of the large amounts of data that needs to be handled in-network, and the network nodes are prone to disruptions caused by loss of nodes or poor link quality among communicating nodes [25-27]. Hence the ability to provide information with QoI attributes of high reliability, low latency and good hop-to-hop throughput are essential for improving the experience of an end-user receiving such data. We make use of an information-centric approach to deal the large amount of information available in the network. Sensed data is identified using attribute tags at sensor nodes. Request for sensory information issued at the sink is routed towards the location(s) in the network where the information has been published. As the request traverses through the network, intermediate nodes are checked for cached copies. As soon as an instance of the desired sensory information is found, it is returned to the sink using cognitive data delivery techniques based on the relative priorities of the QoI attributes that satisfy end-user requirements for a given traffic flow.

4.4 System models

In this section, we explain the COGNICENSE system models and its core components in details, in addition to listing our main assumptions.

4.4.1 Quality of Information (QoI)

QoI is defined as the level of satisfaction experienced/perceived by the end-user on the information received from the network [13]. Attributes such as reliability, latency and throughput are used to evaluate the QoI of data delivered to the sink. To differentiate QoI from Quality of Service (QoS) of WSNs [28], QoS takes care of the operational aspects of the network, while QoI is associated with the characteristics of the sensory information made available to the end-user. In our proposed approach, priorities are evaluated for these QoI attributes for each application traffic type at the sink, and the network tries to deliver the information with the desired QoI to the sink/end-user. For SOM applications in WSNs, QoI attributes that help us assess how well the network is able to gather and provide relevant sensory information is based on the following QoI attributes: reliability, latency and throughput. Their definitions are based on the work in [29], and are presented here briefly:

Latency (L): is defined in terms of the mean frame service time at the MAC layer and is estimated as the time interval from the instant a packet is at the head of its MAC queue and ready to transmit, till an ACK for such a packet is received. In other words, it is the average delay for a successfully received packet.

Reliability (R): is defined as the probability that a frame is not blocked, or lost due to channel access failure or discarded as a result of reaching the maximum number of retries limit.

Average throughput (AT): is a function of reliability and is defined as: λ * Reliability * Application load (bits), where λ is the average frame arrival rate at a node in bits/second.

Instantaneous throughput (IT): is a function of latency and is defined as: Application payload (bits)/Latency(s). We use the instantaneous throughput value for computations in our work, and refer to it simply as T.

4.4.2 Network Lifetime

In this work, we propose a novel definition for network lifetime based on the Quality of Information (QoI) perceived by the end-user. Network Lifetime is defined as: ***the time or number of transmission rounds beyond which the network can no longer deliver useful information to the end-user. This is reflected by the network's inability to find a data delivery path with***

satisfactory values for QoI attributes (latency, reliability and throughput), as determined by the end-user, or when there is insufficient energy in the network nodes to deliver such data to the sink for any of the application generated requests.

This definition not only caters to satisfying the application requirements, but also considers the status of the network and node resources (especially in terms of remaining energy at the nodes) in defining the network lifetime. If sensor nodes or LCNs were drained of energy, then at each hop, the QoI attribute values would be affected, and thus reflected in the overall value of information delivered at the sink. Thus it also justifies the fact that if the network doesn’t have sufficient resources to deliver data, it cannot satisfy the end-user, and hence it should be considered as the end-of-life of the network, as no useful information can be derived from it.

4.4.3 Application traffic profiles for smart outdoor monitoring applications

Application traffic is profiled into three categories [30] based on how often sensed information from the network needs to be delivered to the end-user, and the priorities associated with the QoI attributes for each traffic type. They traffic profiles are as follows:

Type I: periodic (application defined rate),

Type II: intermittent (application/external stimulus defined rate) or event driven/query driven traffic

Type III: low-latency data (emergency/alerting information)

We illustrate this traffic classification by making use of a sensor network deployed in the following SOM applications. The first one is a sensor network deployed for urban environment monitoring. In this application, traffic flow for an air-quality monitoring station is classified as Type I. Information flow generated in response to queries from an operator or end-user, requesting for specific information such as temperature or humidity at a specific time of the day is classified as Type II traffic. Finally, a service that issues alerts such as: High Ultra-Violet radiation warning, heat wave warning during extreme temperatures, reduced visibility warning, and pollen alerts, has traffic flow corresponding to Type III.

Another example of a SOM is a sensor network deployed for monitoring a forest environment [31]. When the network transmits information corresponding to periodically sensed data from the forest region, the flow corresponds to Type I traffic. Information flow corresponding to the assessment of factors that influence the type of flora and fauna found in the monitored region is classified as Type II traffic, and traffic flow associated with alerts issued in emergency situations such as forest fires is classified as Type III traffic.

4.4.4 Network architecture and components

Figure 4-1 represents the components of the COGNICENSE framework and their interactions. Sensor nodes (SNs), Relay Nodes (RNs), Local Cognitive Nodes (LCNs) and Global Cognitive Nodes (GCNs) constitute the nodes of the cognitive information–centric sensor network (CICSN). SNs constitute the leaf nodes that are deployed uniformly and randomly in the network. They communicate with LCNs and RNs lying within their communication range. Typically, SNs communicate with only one parent LCN or RN at a time. LCNs communicate with each other, with RNs, and a cognitive sink node called the GCN, which is located at the center of the deployment region. The GCN carries information to and from the sensor network to the end-user through a gateway and access-point. When hierarchically represented, the CICSN node interactions are as depicted in Figure 4-2a. LCNs and RNs are deployed at pre-determined locations on a grid as shown in Figure 4-2b, so as to ensure complete coverage of the target area and connectivity of SNs with the GCN.

4.4.4.1 Cognition in ICSNs

Haykin [32] and Mitola [33] have perhaps defined cognition in its most extensive form in the context of wireless communication systems. Going beyond simple adaptations, they make use of a feedback loop: the Observe-Analyze-Decide-Act (OADA) loop [20], to model cognition in a way that doesn't deal with imitating human-like behavior, but in making intuitive decisions based on learning from the environment to adapt to current network conditions, while inferring from past behavior and knowledge, to predict a course of action for the future that the network can benefit from.

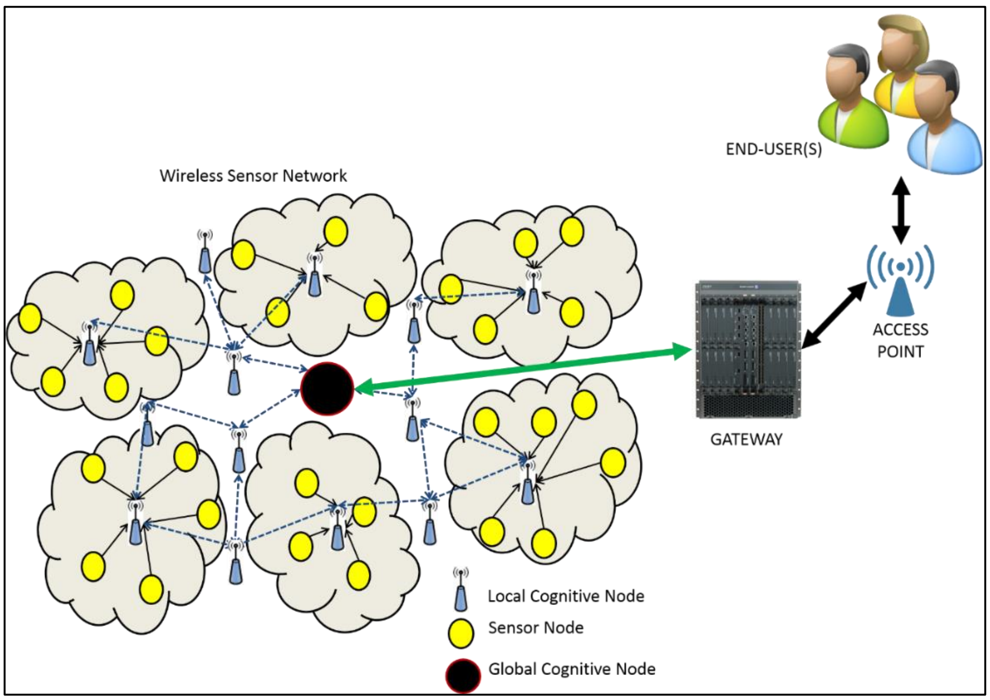


Figure 4-1: The Cognitive Information-Centric Sensor Network Architecture.

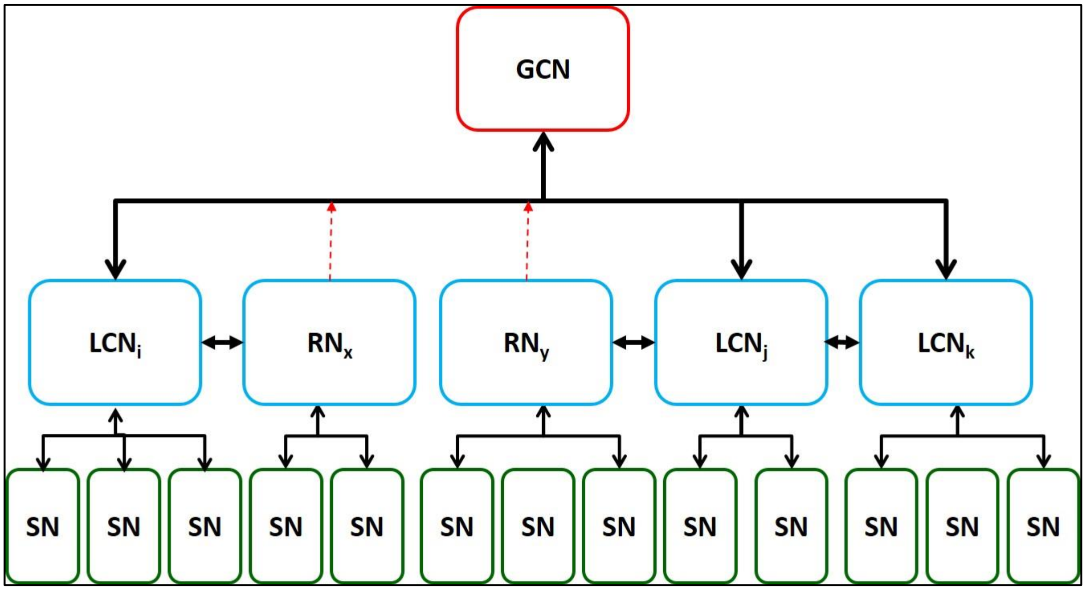


Figure 4-2a: Hierarchical organization of network nodes in the CICSN.

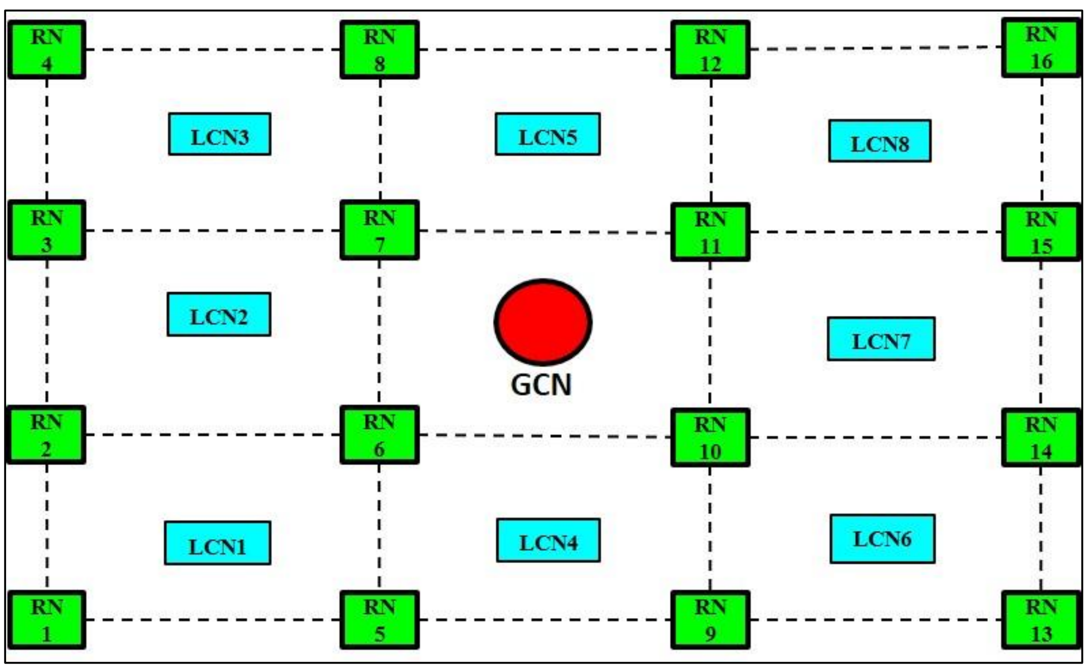


Figure 4-2b: Representation of LCNs and RNs in a 2-Dimensional grid structure.

Based on this idea, and drawing from the work on cognitive networks [34] and extending our work on cognitive information centric sensor networks [35, 36], we define *elements of cognition* to implement the functionality of the Observe-Analyze-Decide-Act (OADA) loop. Knowledge representation, reasoning, and learning constitute the elements of cognition, which when implemented in specialized nodes of the network, will help them make cognitive decisions, and make the sensor network, a cognitive one. In the CICSN, LCNs and GCNs are the specialized nodes that implement the elements of cognition.

4.4.4.2 Node Functions

In this section we describe the functions of the sensor, relay and cognitive nodes of the ICSN. We start with the sensor nodes.Sensor nodes host a multitude of sensors as required by the application platform. Raw sensed-data is stored in attribute-value pairs. This representation facilitates named-data identification to locate the user-requested information. Thus, the two main functions of the sensor nodes are: (i) Sensing raw-data, and (ii) Storing sensed information in attribute-value pairs. Details of the attribute-value pair representation follow in section 4.5.1, where we deal with Knowledge Representation. They communicate with relay and local cognitive nodes. Relay nodes communicate with SNs and LCNs to act as intermediate nodes that gather information from SNs, and forward them to their LCN neighbors. They deliver data over multi-hop paths to the GCN.

LCNs perform two main functions: (i) gathering sensory-data from sensor nodes, and forwarded information from relay nodes, (ii) data delivery based on QoI requirements of the traffic type. LCNs also function as caches to store the data as it travels through the network. LCNs make use of the sensor attributes to identify the relevant data, similar to the named data-object search in ICNs and DCSNs. The requirements on the QoI attributes are based on the type of traffic flow generated as a result of the end-user's request. As for dealing with the QoI attribute requirements, an Analytic hierarchy process (AHP) [14, 37] is implemented as the reasoning element of cognition to make the decision in the LCNs. We elaborate on this technique in section 4.5.4.

GCNs have the following main functions: They receive user requests and synthesize it to identify the following information: application traffic type, requested sensor attributes, and QoI attribute priorities. They broadcast the synthesized information to the LCNs, so that they may process it further to gather the requested information from the network. Once the network returns the requested information, GCNs process it to determine if the QoI provided by the network meets with the user requirements, and deliver information with acceptable QoI to end-user. They also determine when the network is no longer able to deliver useful information from the network, thus flagging the end-of-life of the network.

4.5 The COGNICENSE Framework

Elements of cognition in the network nodes and an Information-Centric data delivery approach are the two main constituents of the COGNICENSE framework. The elements that help in implementing cognition in the cognitive nodes are: knowledge representation, reasoning and learning. Knowledge representation helps in identifying data using attribute-value pairs, contributing towards identifying named-data objects for the information-centric approach. Reasoning helps in multi-criteria decision making to prioritize the QoI attributes for a given traffic flow, and decide on the number of sensor nodes chosen for data transmission to the LCN, or the next hop node chosen along the data delivery path to the GCN. While reasoning helps in achieving short-term objectives and making decisions that help the current situation, learning helps in achieving long-term goals of the network, such as improving its lifetime. Feedback obtained from the network's past behavior aids the learning process, and helps in planning proactive responses to changes in network behavior and user requests.

4.5.1 Knowledge Representation

A Frame structure based on attribute-value pairs is used in sensor nodes and the cognitive nodes for knowledge representation. In frame-based knowledge representation [38], the frame is defined as a hierarchical data-structure with inheritance [39]. It has slots which are function-specific cells for data. In sensor nodes, these function-specific cells store sensor attribute-value pairs. In LCNs, they store more information, such as the one-hop neighbor LCNs and the associated values of QoI attributes in the last communication round. Information accumulated over several rounds of information transmission leads to the formation of a Knowledge Base (KB), which can be looked up by the reasoning mechanism to make quick decisions on choosing the data delivery path which satisfies the QoI delivered to the end-user.

We make use of a semantic naming scheme using strings (sequence of characters) that provide information about the originator of the request, traffic type expected to be generated in response to the request, direction from which the data is requested, and the sensor data attribute(s) corresponding to which the data is to be gathered. The naming scheme has two main components: (i) Request Classifier, and (ii) Information Attributes. The Request Classifier (RQ) field has two sub-fields: the originator of the request, and the type of traffic expected. The Information Attribute (IA) component also has two sub-fields: Direction Attribute and Sensory data attribute. The two fields are separated by a colon ':' and the sub-fields within a field are separated by an underscore '_'. Here is the format of a request string: <Request_classifier> : <Information_Attribute>. Let us consider an example request string. *Sink_type1:N_temp.* Here, "Sink" indicates that the request has been originated by the sink. "type1" indicates that the expected response from the network is a periodic traffic flow. "N" indicates that the direction from which the data is expected to be gathered is North. "temp" indicates that temperature data is being requested. Thus the request string means: Sink initiated a request to collect periodic data from the Northern region of the deployment for the temperature attribute. Further, a combination of logical and relational operators can be used to add more details in the request. For example, the request string *Sink_type1-60:N_temp&&humd* specifies that temperature and humidity values are to be returned periodically, every 60 minutes. Once a complete match is found for the request string, the data is returned in attribute-value pairs to the sink by concatenating it to the original request string using a ":" operator, and changing "Sink" to "Source". For example, the response string: *Source_type1:N_temp:temp-25_temp-26_temp-24* indicates the temperature-value pairs recorded were 24C, 25C and 26C. The alphabets required for a complete representation of this language are represented in Table 4-II. For further digitizing the representation, each of the alphabet's values can be uniquely binary encoded.

Table 4-II: Alphabet used for representation of attributes as semantic information.

ALPHABET	**VALUES**	**REMARK**
α (Information Source)	{Sink, Source}	Indicates if this is a request or response
β (Traffic Type)	{type1, type2, type3}	Traffic flow type expected in the network in response to request.
γ (Direction attribute)	{N,E,W,S,NE,NW,SE,SW,ALL}	Direction(s) from which data may be requested. "All" indicates broadcast throughout the network.
δ (Attributes of Sensed data)	{temp, humd, uvi, co2,time}	Sensory attributes for which data can be provided by sensor nodes. "time" indicates the time stamp at which data was registered at the sensor node.
Logical and relational operators	&&, >, <. >=. <=	

The cognitive nodes (GCN and LCN) will be able to generate and parse these strings and arrange the information gathered from SNs / RNs in the desired format.

4.5.2 Learning

Learning is used in the COGNICENSE framework for identifying data delivery paths towards the GCN that satisfy the user's requirements in terms of QoI attributes. In this work, we make use of a direction-based heuristic to determine the data delivery path through RNs that lie in the direction of the GCN. This means that each time an LCN has to choose from among multiple RNs to decide the next hop, the direction-based heuristic eliminates RNs that increase the distance between the current LCN and GCN. Knowledge of the positions of the LCN and its one-hop RNs is used by the heuristic to determine the set of such RNs, which we call *forward-hop-RNs*. Thus the forward-hop-RNs of an LCN identified by the direction-heuristic is constituted by those RNs that reduce the distance between the LCN and the GCN. This information is stored in the LCN's knowledge base for use in the next transmission rounds. Feedback about QoI delivered along the forward-hop-RNs is used to identify the best forward- hop-RN for each traffic type. Thus the direction-based heuristic, along with feedback from the network about the QoI delivered along the chosen paths helps the LCNs to learn data delivery paths to the sink, as the network topology changes due to link variations and node deaths.

4.5.3 Reasoning

An Analytic Hierarchy Process (AHP) is used for implementing the reasoning element of cognition. AHP aids with multiple-criteria decision making while deciding on the data delivery path based on the Quality of Information requirements of the requested application. Example: For Type III traffic, requesting for low latency data, the QoI requirements are as follows: Highest priority: Latency, followed by reliability and finally throughput. This means that while choosing the next hop node for data delivery, the node which provides the lowest latency, will be chosen. Reliability is more important than throughput. Hence, if two next-hops guarantee the same latency then the next attribute to compare will be reliability, and lastly, throughput. AHP provides a method for pair-wise comparison of each of the QoI attributes and helps to choose the node that can provide the best value of information with respect to all three QoI attributes. Subsequent sections have more details with a running example on AHP. While these calculations help in deciding the next-hop, they also help in planning for future actions. The cognitive nodes are able to store the calculated priorities of the QoI attributes, which they can use to decide which type of traffic the LCNs can best provide for. Hence, these calculations need not be necessarily calculated for every transmission round.

4.5.4 The AHP framework for data delivery based on QoI attributes

There are three levels in the AHP hierarchy constituted of: Goal, Criteria and Alternatives as shown in Figure 4-3.

- *Goal*: Deliver application-requested sensory information to the GCN from LCN by identifying the next hop node.

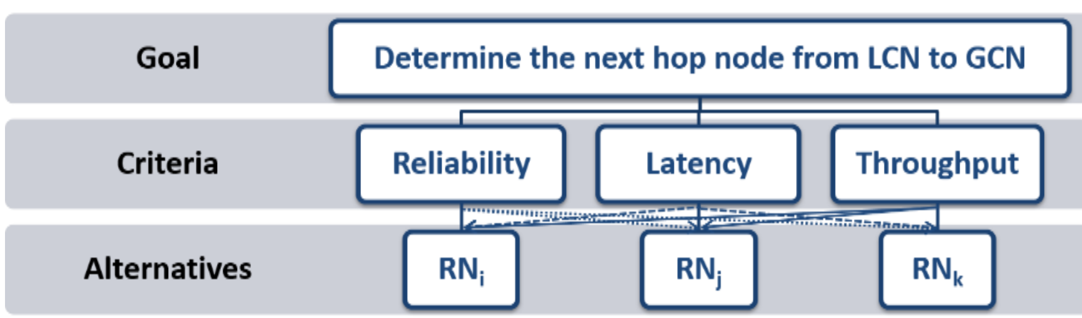


Figure 4-3: The AHP Hierarchy.

- *Criteria*: Data must be delivered with the appropriate priorities of QoI Attributes for each application type. The QoI attributes that are considered are: latency, reliability, and throughput.
- *Alternatives*: The RNs in the network available to forward the data over multiple-hops in the network.

A fundamental scale for pairwise comparisons is then used to set application-defined priorities for the QoI attributes [37]. Then the priorities of QoI attributes are established using pair-wise comparison. Let us consider an example where a SOM application wishes to transmit low-latency alerting information to its users. From the three QoI attributes of latency, reliability and throughput, we would assign the highest priority to latency, to ensure timely delivery of the alert, followed by reliability and then throughput. We tabulate the relative priorities of each the QoI attributes using pair-wise comparison and generate Table 4-III. Then, the AHP computation involves generating the Eigen vector for the values in Table IV, using the following steps:

i. Represent the values of Table 4-IV in matrix form {A = [1,4,6;1/4,1,3;1/6,1/3,1]}
ii. Compute the eigen vector of the matrix A { [v,d] = eig(A)}
iii. Isolate the absolute, real values of the eigen vector { q=abs(real(v(:,1)))}
iv. Compute the normalized, relative priority values as {Relative_Priorities=q/norm(q,1)}

This way, the AHP algorithm is implemented at LCNs to establish the relative priorities of QoI attributes and helps in multi-criteria decision making. The QoI attributes are the criteria and the goal is to find the next-hop RN during data delivery from LCN towards GCN, which provides the highest value for a specific QoI attribute, or provides the overall best value of information (VoI) as illustrated in Table V. VoI is based on the combined value of QoI attributes and energy consumed during the process of delivering information to the GCN.

Table 4-III: Pair-wise comparison of QoI attributes.

Latency	4	Reliability	1
Latency	6	Throughput	1
Reliability	3	Throughput	1

Table 4-IV: AHP for QoI Attributes v/s Goal.

Goal: Best QoI	**Latency**	**Reliability**	**Throughput**	**Relative Priorities of the QoI attributes**
Latency	1	4	6	**0.691**
Reliability	1/4	1	3	0.2176
Throughput	1/6	1/3	1	0.0914

Table 4-V: AHP to evaluate the overall priorities for all possible next-hop RNs.

Best candidate for next hop RNx	**Priority with respect to**			
	Latency	**Reliability**	**Throughput**	**Goal**
RNi	**0.252**	0.015	0.101	**0.375**
RNj	0.2	0.018	0.11	0.329
RNk	0.164	**0.019**	**0.116**	0.296

VoI delivered to the end user is said to be maximized when data is delivered over links that provide the best effective QoI for each traffic type, while minimizing the energy consumed in the network while doing so.

$$\textit{Value of Information (VoI)} = \sum_{n\text{-}hops} (\textit{Effective QoI}) - \sum_{n\text{-}hops} (\textit{Energy Cost}) \quad (1)$$

Equation (1) highlights that lower the energy cost of delivering data to the sink, higher is the VoI associated with that data/information object. The QoI must be maximized and energy cost minimized to achieve the best VoI. If energy consumption is measured as a function of the number of transactions taking place before data is delivered to the GCN, a simple metric - the hop count can be used to approximate the energy cost. If the information is transmitted from source to GCN over minimum number of hops, each link providing the best combined QoI for that traffic type, we can say that the information was delivered to the GCN with good VoI. The steps used in the AHP to establish priorities for the QoI attributes and identify the best next-hop path in delivering the application data to the GCN are illustrated in Algorithm 1.

Information about the relative priorities of the QoI attributes as desired by the user are received as input from GCN in steps 1 - 3. The output is a next hop RN that provides the best QoI as shown by steps 4-5. The simulations are set to run till no path can be found to GCN or till 50% of RNs and LCNs die. In steps 9-11, AHP analysis identifies the best next-hop RN that satisfies these requirements, and identifies the next-hop path for data transmission. Steps 12 – 17 define actions to be taken when data reaches the GCN and leads to a successful transmission, or reaches another LCN from where next hop has to be identified. Steps 18-21 indicate that if a path to GCN was not found along the chosen path, GCN issues a re-transmit request. These computations can be initially carried out for each next-hop node decision in the data delivery path. This technique helps to build the learning database at each LCN about its next-hop neighbor, and the priorities each of them offers with respect to the QoI attributes. This information can be stored and used for planning future rounds of data delivery for application traffic that may need to choose a different next hop for the same source LCN, based on the expected values of attribute priorities at the GCN. Thus we can see that this AHP process helps in adaptive multi-criteria decision making during data delivery, in considering the desired attribute priorities for each application-traffic type.

Algorithm1: AHP analysis to determine the data delivery path

1. **Function AHP (QoI.priorities)**
2. **Input**
3. QoI.priorities: End-user defined priorities on QoI attributes for requested data
4. **Output**
5. RN_x: Forward-hop $RN_x \in \{ RN_1 \ldots.. RN_n \}$ with best QoI
6. **Begin**
7. **Initialize**: QoI priority matrix for traffic type; Success=0;
8. **While** (*number of dead nodes<50% or network not disconnected)*
9.   *AHP_analysis*(Next-hop RNs v/s QoI attributes) //
10.   Next hop RN = RN_x //This is the RN with best QoI for chosen traffic type
11.   Transmit data to next-hop RN
12.   **If** (next hop = GCN)
13.   Success=1;
14.   **Else**
15.   *Choose next-hop LCN*
16.   goto step 8
17.   **End**
18.   **If** *(Success==0)*
19.   GCN Retransmits request
20.   **End**
21. **End**

4.5.5 Node mobility support in the COGNICENSE framework

The COGNICENSE framework allows for caching data at LCNs that act as intermediate nodes. This makes data readily available for users at nodes other than sensor nodes, thus offering two main advantages: (i) It prevents requests being sent out to sensor nodes, which may be in a sleep cycle, leading to a lost request and (ii) it helps to conserve valuable energy resources by reducing the number of transmissions occurring in the network; both from sensor nodes towards the sink, and over multiple relay nodes that transmit the sensory information from the sensor nodes to the sink. Furthermore, the named-data identification enhances the advantages offered by the data caching feature at the LCNs in terms of supporting node mobility. We discuss the issue of node mobility under two categories: (a) Sensor node mobility, and (b) LCN mobility.

4.5.5.1 Sensor Node mobility support

In the COGNICENSE framework, search for data is name-based, which means that the request is not associated with any specific address, location or an end-point. This is in contrast with the IP based approach, where an address is associated with each sensor node, and the request-response cycle involves the establishment and maintenance of an end-to-end connection between the sensor node and the Sink. This restricts the ability of the network to support node mobility, as the loss of connectivity with the source-sensor node or any intermediate node involved in the end-to-end connection, due to node death or lossy links, affects the data gathering and routing capability of the network.

However, in a cognitive ICSN, the requested information could be located anywhere in the network, and the user will be able to access it, since the request is not tied with any specific node address. Any node that can provide a match to the requested information can provide the data. Moreover, the routing path is not fixed, and can adapt to the changing network topology. This is made possible by the LCNs that make use of cognitive reasoning to dynamically identify data delivery paths based on the type of request, and how well a link had performed in a previous round. The data delivery paths are chosen based on the QoI attributes of latency, reliability and throughput. The LCNs offer another advantage of acting as a data cache. Information gathered from sensor nodes can be stored in these LCNs to make them available on-demand, without having to access the source sensor nodes. We assume that cooperative caching techniques designed for wireless sensor networks [40] that deal with large amounts of sensed data, can be applied at the LCNs to enable them to manage information storage. In addition, we assume that cache replacement algorithms such as Least Value First (LVF) replacement [16] can be used to maintain availability of relevant data while evicting stale and unused data from the cache, to make space for fresh data. Thus, even if a source sensor node was mobile, the sensed information is stored in LCNs whenever the node lies in close proximity with the LCN, and is made available to the user, irrespective of the mobility condition and/or pattern of the sensor node. Thus the COGNICENSE framework is capable of supporting sensor node mobility, without negatively affecting the network performance.

4.5.5.2 Local Cognitive Node mobility support

A further advancement that can be made to the COGNICENSE framework, is the ability to support LCN mobility. A combination of static and mobile data collector LCNs could be used in the information-centric sensor network to improve the data gathering capability of the network. The advantage offered by having mobile LCNs is that, when a part of the network starts to deteriorate in its energy capacity and link conditions, the mobile LCNs will still be able to gather information from that part of the network, and store it in their cache. Thus preventing a part of the network from getting completely disconnected from the rest of the network, as long as the sensor nodes remain functional. These mobile LCNs could then communicate amongst themselves and with the static LCNs, to decide on the best way to deliver the collected data to the Sink, and also to maintain information about the entire network to make informed decisions while responding to user requests.

4.5.6 Energy considerations in the COGNICENSE framework

Energy conservation is one of the most important aspects of WSN design. In ZigBee based address-centric WSNs, sensor nodes off-load the energy-draining communication tasks to relay nodes. SNs being leaf nodes do not have the network layer to forward data beyond their one-hop relay nodes, and they do not even communicate amongst each other. The multi-hop relaying between source and sink is done by RNs, which have higher battery and processing capacity. Let us denote the energy cost of the relay node using Equation (2):

$$C_{RN-E} = C(TE_{Tx} + RE_{Rx}) \tag{2}$$

Most of the energy consumption at the RN is due to data communication, represented by E_{Tx} for energy consumed during transmission and E_{Rx} for energy consumed during data reception. T and R represent the number of transmitted and received packets respectively. Let us compare this energy with that at the cognitive node (C_{CN-E}). Typical functions of CNs that consume additional energy compared to regular RNs are data aggregation and the cognitive decision process. Additional energy consumption is accounted for by two factors: (a) protocol overhead incurred during cognitive data delivery due to feedback from the network during the learning process and the exchange of values of QoI attributes such as latency, reliability and throughput while making routing decisions and (b) increased transmit power for increasing the communication range of CNs.

$$C_{CN-E} = C(TE_{Tx} + RE_{Rx}) + C(AE_{ag}) + C(PE_{cog-process}) \tag{3}$$

In Equation (3) T, R, A, and P, are the total number of packets that are transmitted, received, aggregated and processed by the cognitive elements respectively, in each transmission round. $C(TE_{Tx} + RE_{Rx})$ is the energy cost incurred during data transmission and reception, $C(AE_{ag})$

represents the energy cost incurred during data aggregation and $\mathrm{C(PE_{cog-process}})$ represents the energy cost due to protocol and processing overhead during the cognitive processes. Expressing Equation (3) in terms of the energy cost of RNs we get:

$$C_{CN-E} \geq C_{RN-E} + AE_{ag} + CE_{cog-process} \tag{4}$$

If the relay and cognitive nodes use the same transmit power, then the equality sign holds true in Equation (4). In any case, the energy cost of the cognitive node is higher than that of the relay node. In order to ensure that the energy cost of CNs does not offset the advantages offered by it in terms of adapting to the dynamic traffic flows and network topology changes, the cost can be optimized by maximizing the number of RNs and minimizing the LCNs in the deployment.

4.6 Simulations and results

A CICSN for a SOM application was implemented on top of an IEEE 802.15.4 MAC-PHY simulator [41, 42] in Matlab. The deployment and interconnection among the network nodes (SNs, RNs, LCNs, and the GCN) is as shown in Figure 4-4. The cyan and magenta lines indicate links between SNs and LCNs and SNs and RNs respectively. GCN in red is located at the center of the target area. Blue lines show inter-LCN communication links and the black lines indicate interactions between LCNs and RNs. Green lines indicate the links between the GCN and its one-hop RNs, and the red lines are the links between the GCN and nearest LCNs. The simulations were used to evaluate the impact of network and node parameters on QoI attributes.

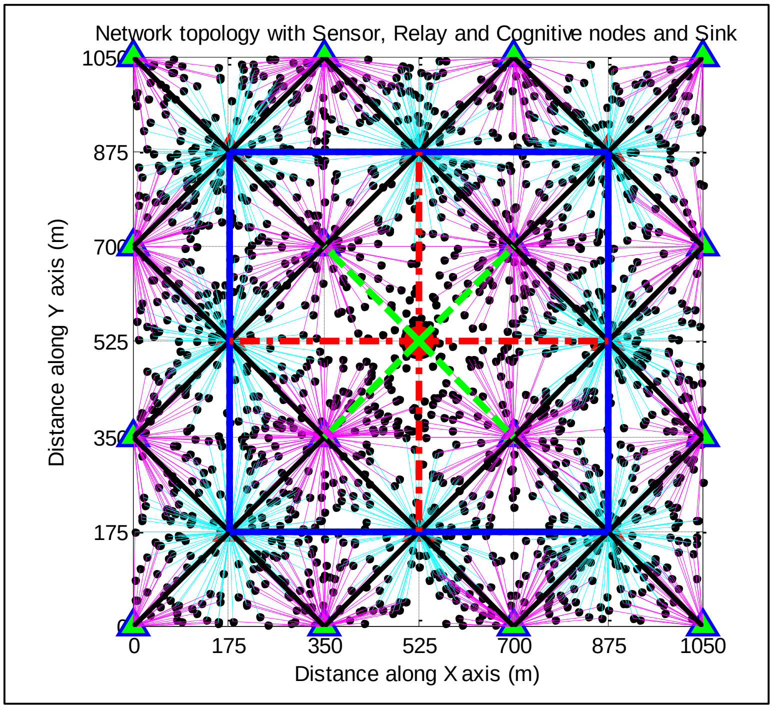


Figure 4-4: Interconnection among SNs, LCNs, and RNs.

Using parameters identified from this simulation, the AHP based data delivery technique (AHPDD) was implemented, and its performance was compared with two other techniques – a multipath data delivery technique (MDD), and a higher remaining battery based data delivery technique (HRBDD) in terms of the number of data transmissions to the GCN, and the QoI along the data delivery path.

4.6.1 Simulation setup and parameters

The first set of simulations was used to identify parameters that affect the QoI attributes of latency, reliability and throughput, for the application. Parameters chosen for observation were: (a) N_active: the number of nodes attempting to simultaneously transmit data, and (b) the offered load: the per node frame arrival rate expressed as a fraction of the application payload in bits per second. The simulation was setup to identify the impact of varying the offered load on the QoI attributes for different values of N_active. The maximum and minimum possible values for N_active were chosen based on the node binding information obtained from the deployed CICSN. From 10 sets of random deployment of sensor nodes, we found a lower bound of about 10 sensor nodes per LCN, and an upper bound of close to 60 sensor nodes per LCN. The range of values for per node offered load was 0 to 1400 bits per second, such that the load could be expressed as a fraction of the application payload, ranging from 0.1 to 1.4 times the size of the maximum application payload of 121 bytes. The remaining simulation parameters were set as shown in Table 4-VI.

Table 4-VI: Parameters of the simulated CICSN.

Parameter	Value
Target area	1050m x 1050m
Number of nodes	SNs: 1500 RNs: 16 LCNs: 8
Transmit power	SN: <3dB RN: 3dB LCN: {3dB, 5dB, 7dB}
Communication Range	SN: 175m RN: 250m LCN: 350m GCN: 500m
Application payload size	121Bytes
Per node offered load	0-1400 bits/s

4.6.2 Simulations showing the impact of network and node parameter variations on the QoI attributes

The impact of varying the offered load and N_active on the QoI attributes of latency, reliability, and throughput for the SOM application, is analyzed using the simulation results in Figure 4-5. Figure 4-5a indicates that latency increases almost linearly with increase in offered load for small values of N_active, upto 10 nodes. However, for higher values of N_active, latency saturated around 0.1s for loads greater than 1000bps. Figure 4-5b shows an overall trend of decrease in reliability as the offered load increases. However, there is a marked difference in the variation of reliability with increase in N_active. Reliability drops exponentially for values of N_active greater than 30, as offered load increases. For values of N_active around 20, reliability remains around 1for loads upto 500bps per node, after which it drops linearly with increase in offered load. Figure 4-5c indicates an overall decrease in throughput as offered load increases. For N_active =10, the decrease is linear, but for higher values of N_active, (20 nodes and above), the decrease in throughput with increase in offered load is exponential. Figure 4-5d indicates a very different trend compared with instantaneous throughput at N_active =10. There is an increase in throughput with increase in offered load, and stabilizes at around 700bps for offered load over 1250bps. However, as the value of N_active is increased, the absolute value of throughput decreases, and the increasing trend in throughput that was seen for N_active =10, starts reversing for loads greater than 500bps for N_active over 30. We made the following observations from analyzing the impact of varying the per-node offered load and N_active on the QoI attributes: (i). Values of each of the QoI attributes deteriorates as the offered load increases, and (ii). Restricting the number of nodes attempting to simultaneously transmit data (N_active) to around 10 nodes, helps to achieve good values for all the QoI attributes. We use these observations to setup the simulation parameters for our next set of simulations.

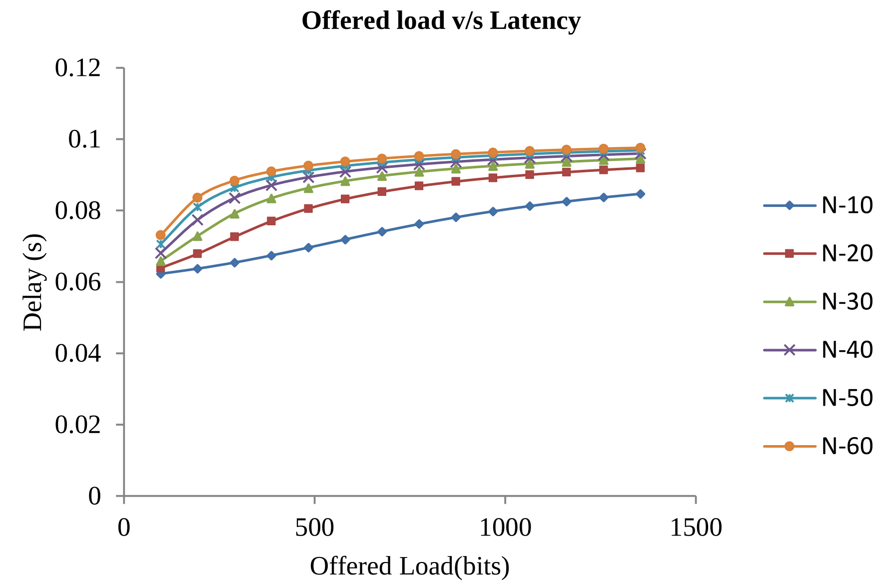


Figure 4-5a: Impact of N_active and offered load on Latency.

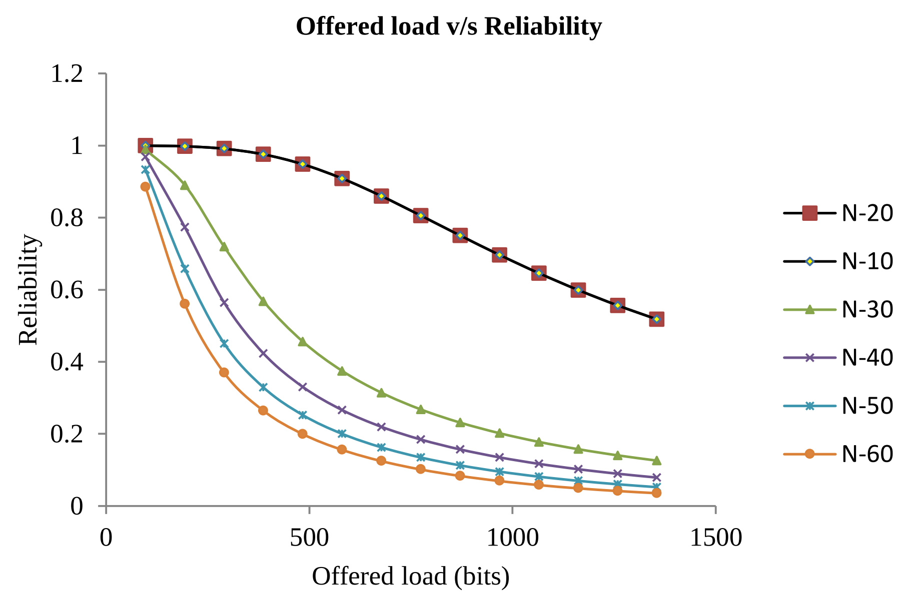


Figure 4-5b: Impact of N_active and offered load on Reliability.

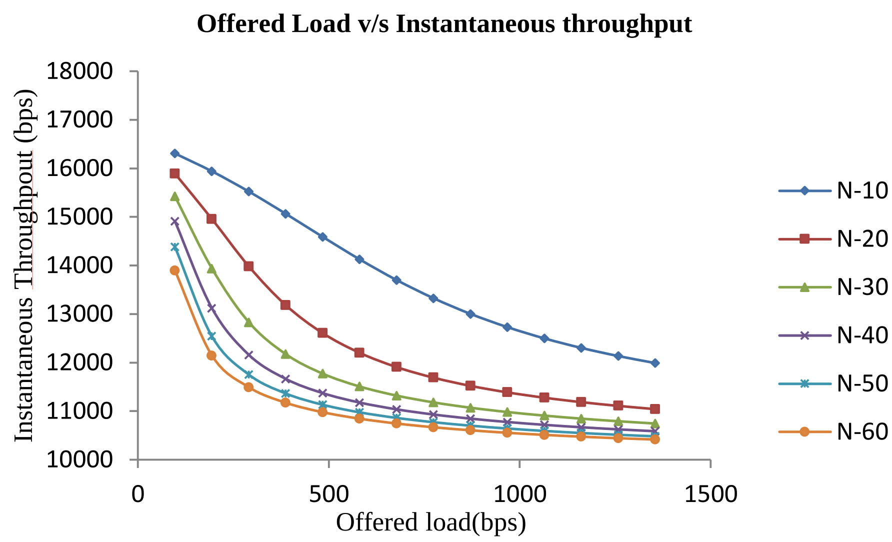


Figure 4-5c: Impact of N_active and offered load on Instantaneous Throughput.

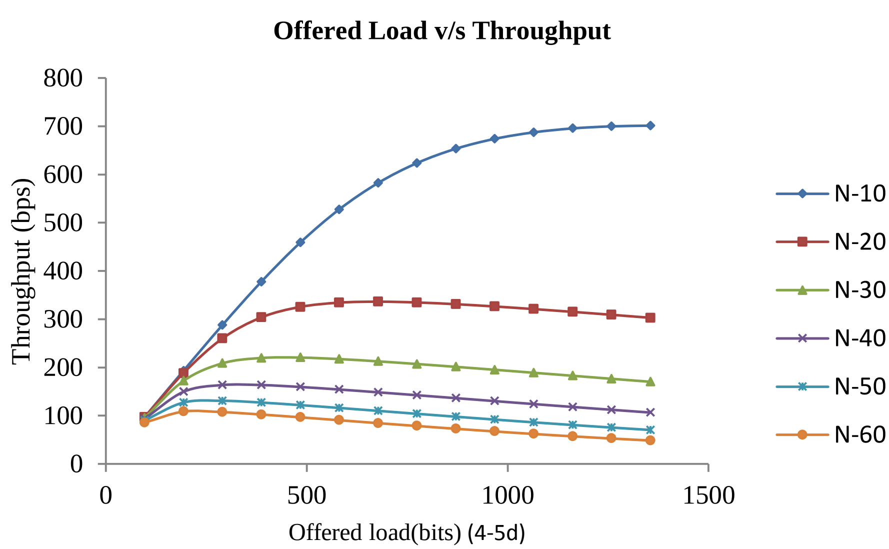


Figure 4-5d: Impact of N_active and offered load on Average Throughput.

Figure 4-5: Impact of varying offered load and N_active on QoI attributes.

4.6.3 Comparative evaluation of data delivery protocols: AHPDD, HRBDD, and MDD

Using the observations from Section 4.6.2, a network environment in which less than 10 nodes are scheduled at a time for simultaneous transmission, and the maximum transmission load is limited to 5 frames per second (fps) is setup. Channel conditions were varied by varying the application payload and N_active values. We assume that LCNs and RNs start with an initial energy of 25 units, and SNs have an initial energy of 15 units. Each transmission from a SN consumes 1unit of its energy; and transmissions from RN to LCN and vice-versa consumes 2units of energy at the transmitting node. Direct communication among LCNs or LCN to GCN consumes 3 units of power. These values are based on the transmit power and communication range capabilities of the nodes. At the start of the simulation, we identify a source LCN at which required data is available. Delivery of data from the identified source LCN to the GCN is considered as one successful transmission round. Using this setup, we analyze the performance of the AHP based data delivery protocol (AHPDD) based on the number of transmission rounds of delivering data from a source LCN to GCN, until one or both of the following simulation termination conditions are satisfied: (i) 50% of the total number of LCNs and RNs die out, or (ii) the network is no longer able to deliver information to the GCN as all the one-hop neighbor RNs and LCNs to the GCN are dead. At this point, the simulations are terminated. AHP analysis is implemented at LCNs to identify the best next hop RN. The priority matrix for AHP analysis is set to identify data delivery path for each of the three traffic types. The AHP based decision protocol is then compared with two other decision criteria in the same network setup, but without considering the cognitive reasoning capabilities at the LCN or GCN. These routing strategies are based on the ones described by Stojmenovic [7] for reporting via alternate paths in a broadcast tree in DCSNs.

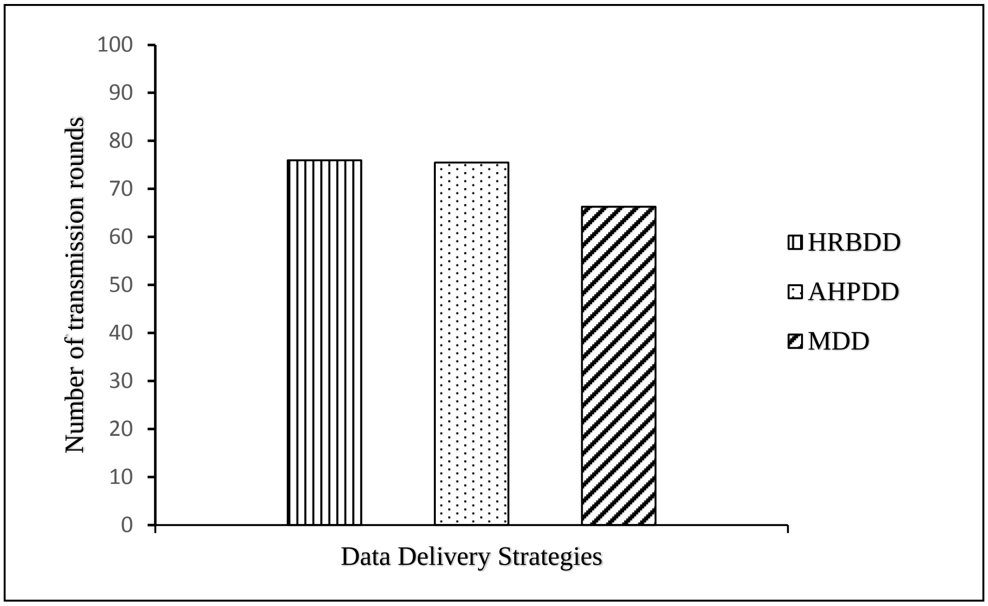


Figure 4-7: Comparison of the total number of transmission rounds.

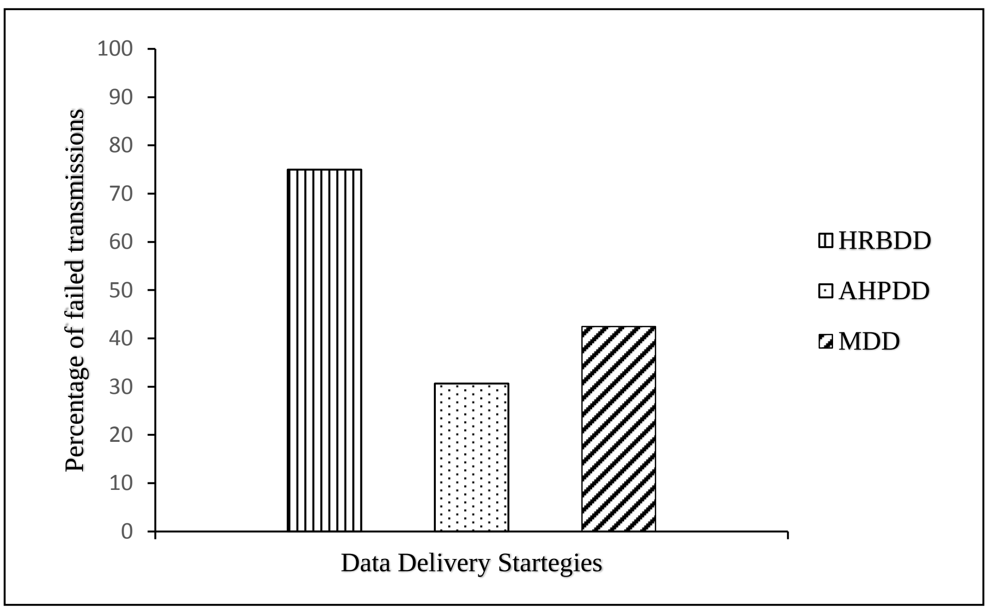


Figure 4-6: Comparison of failure rates.

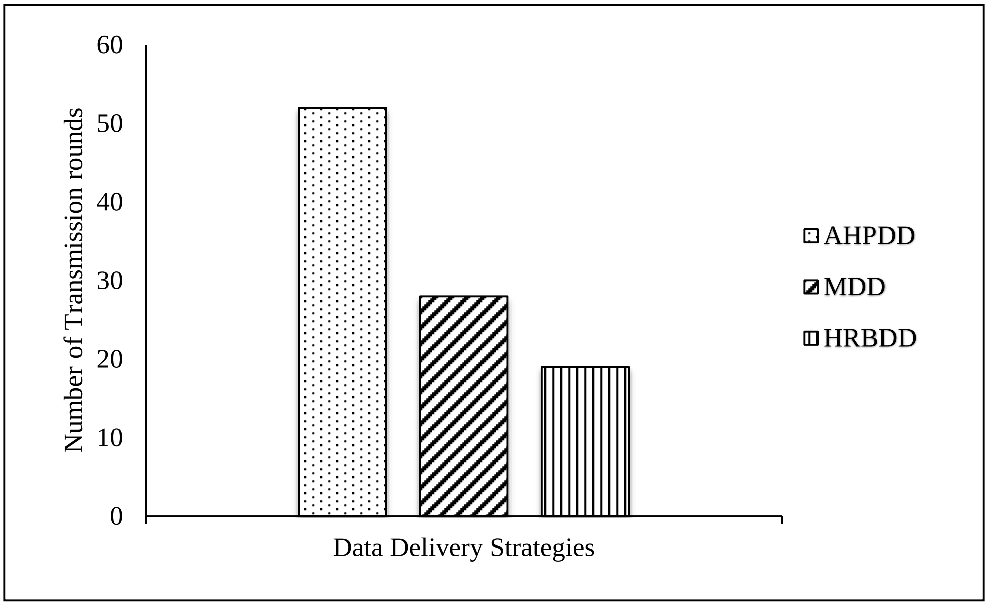


Figure 4-8: Comparison of the number of successful transmission rounds.

The first one is based on choosing an RN with the highest remaining energy from among the one-hop neighbor nodes, and is called highest remaining battery based data delivery technique (HBRDD). The second one is called multipath data delivery (MDD), where each node transmits through all its one-hop neighboring nodes with equal probability to improve the chances of successful data delivery to the sink. Data is delivered via multiple paths at each hop, until at least one of the paths leads to the Sink, which is the non-cognitive version of the GCN. The simulations were allowed to run till one or both the simulation termination conditions were met, and the average value of 25 such simulations was taken. The number of transmission rounds during which data was not delivered to the GCN was also recorded. The following criteria were used to determine unsuccessful transmissions to the GCN: (i) inability of the routing protocol to forward data to the GCN due to node deaths along the path chosen for data transmission, (ii) transmission failure due to insufficient remaining energy at the forwarding nodes. The difference between the total number of transmission rounds, and the number of failed transmissions gives a measure of the number of transmission rounds in which data was successfully transmitted to the GCN. Thus we define the failure rate of the routing protocols in Eq. 5 as follows:

$$Failure\ Rate = \left({}^{Number\ of\ failed\ transmissions}/_{Total\ number\ of\ transmission\ rounds} \right) * 100 \tag{5}$$

From the simulation results in Figure 4-6, we can see that AHPDD and HRBDD perform equally well, and better than MDD, in terms of the number of transmission rounds. However, from Figure 4-7, we see that the number of failed transmissions is very high for HRBDD (57 out of 76). On comparing the failure rates, we find that MDD in fact performs better than HRBDD by 12%. While only 31% of the transmissions using AHPDD fail to reach the GCN, the failure rate is as high as 75% with HRBDD, which is almost twice as worse when compared with the 42% failure rate of MDD. Figure 4-8 shows the number of successful transmission rounds for each of the data delivery techniques. We see that although MDD doesn't keep the network running for more number of transmission rounds compared to HRBDD, it is able to deliver data to the sink successfully for an average of 42% of the total transmission rounds, which is 17% higher than

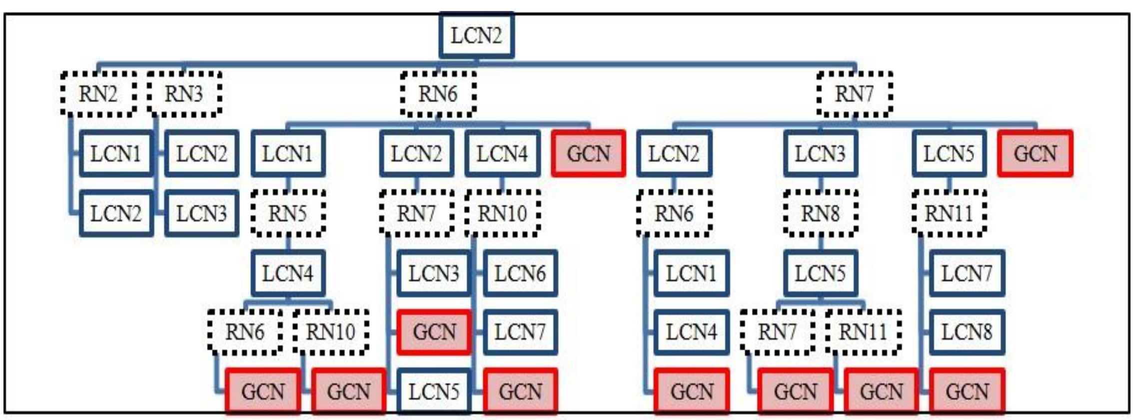


Figure 4-9: Tree-based illustration of a sub-set of paths from LCN2 to GCN through RNs 2, 3, 6, 7.

what is achieved by the HRBDD. However, AHPDD out performs both these protocols by adapting the data delivery decisions to user priorities, and successfully delivering data to the GCN for 70% of the total transmission rounds. From these simulations, we can say that AHPDD is better able to adapt to the changing network topology and deliver data to the GCN with a lower failure rate compared to the other two techniques.

4.6.4 Use-case analysis of the data delivery protocols based on QoI attribute performance

To analyze the performance of the three data delivery techniques in terms of the QoI attributes, we hereby adopt a use case based on the simulations in Section 4.6.3. The remainder of this use case will refer to Figure 4-9 and Table VII.

LCN2 is identified as the source node that has data to be delivered to the GCN, in response to periodic requests (Traffic Type 1) during each transmission round. The one-hop neighbor RNs of LCN2 are RN2, RN3, RN6, and RN7, and have battery levels of 11,9,7, and 5 units respectively at the start of the simulation instant. Values of the QoI attributes are recorded for each of the one-hop RNs. AHP analysis is performed to identify the best forward hop RN for AHPDD as marked in red under the column titled "Effective QoI". The theoretical best next hop RN for the other two protocols is found using AHP analysis (highlighted in green), to compare the QoI performance of the actual next-hop node chosen by the other two protocols. Comparing the QoI performance of the chosen next hop node, we make the following observations: AHPDD always chooses the best QoI providing node between RN6 and RN7, as long as they are available. Although RN2 or RN3 might provide better QoI values for the next hop in some cases, choosing the forward hop RNs reduces the number of hops to reach the GCN. This leads to lesser energy consumption in the network on the whole, and also reduces the cumulative latency along the data delivery path to the GCN. However, this also means that once the forward hop RNs die out, AHPDD has to make use of longer data delivery paths to the GCN. But again, the QoI attributes are still considered in choosing the best among the available next hop nodes. MDD on the other hand, is always able to deliver data through at least one next-hop node that provides the best effective QoI for each traffic type, even though it doesn't have a mechanism to identify the best next hop node.

It is also able to find the shortest route to the Sink because of the multipath approach at each next hop node. However, this performance comes at the cost of a higher overall energy consumption in the network. This can be seen from Table VII, where all the one-hop nodes run out of energy before the other two techniques. Comparing with the observations made from Figures 4-6 to 4-8, we see that although MDD lasts for lesser number of transmission rounds, not only does it provides a lower failure rate, it also performs well in terms of identifying at least one next hop node that provides the best QoI performance. As for HRBDD, what stands out from Table VII is the increased number of hops in delivering data to the Sink, causing an overall increase in energy consumption in the network. This is because HRBDD is always trying to find a node with higher remaining energy at each next hop, irrespective of its QoI performance. Although the chosen next hop node sometimes provides the best QoI, HRBDD's performance with respect to QoI attributes is not consistently good. Over a period of time, this leads to death of more intermediate nodes, causing a higher failure rate as indicated by Figure 4-7, as the sink cannot be reached along a chosen path. This leads to lesser number of successful transmissions to the sink, even though the network might be able to run for a little longer than the multipath routing technique, as shown by Figure 4-8 and Figure 4-6 respectively. Thus, HRBDD performs relatively poorly among the three data delivery strategies, both in terms of delivering data with user-desired QoI attributes, and in terms of the number of successful transmission rounds.

Table 4-VII: Comparative Analysis of data delivery paths in terms of QoI attributes.

	Remaining battery levels			QoI Attributes			Effective QoI	Chosen Next hop Node			Data delivery path		
RN#	**AHPDD**	**HRBDD**	**MDD**	**Latency**	**Reliability**	**Throughput**	**Type1 Traffic**	**AHPDD**	**HRBDD**	**MDD**	**AHPDD**	**HRBDD**	**MDD**
2	11	9	9	0.0219	0.7659	4.6606	0.1999	7	2	2,3,6,7	LCN2->RN7->GCN	LCN2->RN2->LCN1->RN5->LCN4->RN10->SINK	LCN2->RN6/7->SINK
3	9	9	7	0.0126	0.9958	8.1039	0.3013	Hop that offers best QoI=RN2					
6	7	7	5	0.0168	0.951	6.0936	0.2457						
7	3	5	3	0.0161	0.9619	6.3564	0.2531						
2	11	9	7	0.0126	0.9958	8.1023	0.2734	6	3	2,3,6,7	LCN2->RN6->GCN	LCN2->RN3->LCN3->RN8->LCN5->RN7->SINK	LCN2->RN6/7->SINK
3	9	7	5	0.0246	0.5878	4.1472	0.1659	Hop that offers best QoI=RN6					
6	5	7	3	0.0117	0.9976	8.6991	0.2873						
7	3	3	1	0.0126	0.9958	8.1023	0.2734						
2	11	7	5	0.0203	0.8455	5.0411	0.2752	6	2	2,3,6	LCN2->RN6->GCN	LCN2->RN2->LCN1->RN6->SINK	LCN2->RN6->SINK
3	9	7	3	0.0196	0.8691	5.2057	0.2704	Hop that offers best QoI=RN2					
6	3	5	1	0.0168	0.9496	6.0932	0.2478						
7	3	3	1	0.012	0.9972	8.5106	0.2067						
2	11	7	3	0.0224	0.7383	4.5636	0.1963	7	3	2,3	LCN2->RN7->GCN	LCN2->RN3->LCN3->RN7->SINK	LCN2->RN2->LCN1->RN5->LCN4->RN10->SINK
3	9	7	1	0.0117	0.9976	8.6991	0.3178	Hop that offers best QoI=RN3					
6	3	5	1	0.0224	0.7383	4.5636	0.1963						
7	1	1	1	0.0133	0.9926	7.6506	0.2895						
2	11	5	1	0.0117	0.9976	8.6991	0.2758	6	2	2	LCN2->RN6->GCN	LCN2->RN2->LCN1->RN6->SINK	LCN2->RN2->LCN1->RN5->LCN4->RN10->SINK
3	9	7	1	0.0196	0.8691	5.2057	0.1957	Hop that offers best QoI=RN6					
6	3	3	1	0.0117	0.9976	8.6991	0.2758						
7	1	1	1										
2	11	5	1	0.0117	0.9976	8.6991	0.2745	6	3	-	LCN2->RN6->GCN	LCN2->RN3->LCN3->RN8->LCN5->RN11	LCN2 disconnected from Sink (failed transmission)
3	9	7	1	0.0133	0.9926	7.6506	0.2514	Hop that offers best QoI=RN2				RN11 is one-hop from Sink; assumed dead	
6	1	3	1	0.0161	0.963	6.3564	0.2227						
7	1	1	1	0.0133	0.9926	7.6506	0.2514						

4.7 Conclusions

In this paper, we proposed a framework for Cognitive Information Centric Sensor Networks (ICSN) that can be used to implement information-centric data delivery using elements of cognition, i.e. knowledge representation, and inference to advance data-centric sensor networks to cognitive information-centric sensor networks. These ICSNs are able to handle heterogeneous traffic flows in the network generated as a result of requests coming from multiple clients in SOM applications, while considering the QoI attribute priorities for each traffic flow. From the simulations we were able to identify the number of sensor nodes that should be simultaneously scheduled while gathering data, to ensure good quality data from the sensor nodes. Optimally choosing the number of simultaneously transmitting sensor nodes improves the average throughput by about 85%, reliability by about 90% and reduces the latency by about 18% for a given value of offered load (1000bits). The simulation-generated values were used in the next set of simulations that implemented AHP analysis to decide the best next-hop node that should be used for data delivery to the GCN. It was found that the network lasted for significantly more number of transmission rounds, and performed well in responding to varying traffic types and changing network topology, when it implemented cognitive routing decisions, when compared with traditional decision techniques. In our future work, we will enhancing the learning strategy, and implement cache replacement at LCNs to further exploit the cognitive node's capabilities to improve network performance and prolong the network lifetime, while meeting the end-user's requirements.

4.8 References

[1] Al-Fagih, A., Al-Turjman, F., Alsalih, W., and Hassanein, H., "A priced public sensing framework for heterogeneous IoT architectures," *IEEE Transactions on Emerging Topics in Computing*, vol. 1, no. 1, pp. 135-147, Oct. 2013.

[2] Ahlgren, B.; Dannewitz, C.; Imbrenda, C.; Kutscher, D.; Ohlman, B., "A survey of information-centric networking", Communications Magazine, IEEE , vol.50, no.7, pp.26,36, July 2012.doi: 10.1109/MCOM.2012.6231276

[3] Al-Turjman, F. and Hassanein, H. , "Enhanced data delivery framework for dynamic Information-Centric Networks (ICNs)", *In Proc. of the IEEE Local Computer Networks (LCN),* Sydney, Australia, 2013, pp. 831-838.

[4] Ahmed, K., and Gregory, M.A., "Techniques and Challenges of Data Centric Storage Scheme in Wireless Sensor Network." J. Sens. Actuator Netw. Vol.1,no. 1, pp.59-85, 2012. doi: 10.3390/jsan1010059

[5] "Handbook of Sensor Networks: Algorithms and Architectures", Edited by I. Stojmenovic´, CH. 13, pp. 417-456, 2005, John Wiley & Sons, Inc., ISBN 0-471-68472-4

[6] Krishnamachari, B., Estrin, D., and Wicker, S.," Modelling data-centric routing in wireless sensor networks", IEEE Infocom, Vol. 2, pp. 39-44, Jun. 2002.

[7] Intanagonwiwat, C., Govindan, R., Estrin, D., Heidemann, J., and Silva, F., "Directed diffusion for wireless sensor networking", Networking, IEEE/ACM Transactions on, vol. 11, no. 1, pp. 2-16, 2003.

[8] Ratnasamy, S., Karp, B., Shenker, S., Estrin, D., Govindan, R., Yin, L., and Yu, F., "Data-Centric Storage in Sensornets with GHT, a Geographic Hash Table", Mobile Networks and Applications, Vol. 8,no.4, pp.427–442, 2003.

[9] Online: http://www.zigbee.org ZigBee Document 053474r17, "ZigBee specification", Jan.2008,

[10] Vijay, G., Ben Ali Bdira, E. and Ibnkahla, M. "Cognition in wireless sensor networks: A Perspective", Sensors Journal, IEEE, Vol.11, No. 3, pp.582-592, 2011.

[11] Shenai, K. and Mukhopadhyay, S. , "Cognitive sensor networks", Proc. IEEE 26th Int. Conf. Microelectronics (MIEL), pp. 315-320, 2008.

[12] Reznik, L. and Von Pless, G. ,"Neural networks for cognitive sensor networks", Proc. IEEE Int. Joint Conf. Neural Network., IJCNN, pp.1235-1241, 2008.

[13] Sachidananda, V., Khelil, A., and Suri, N., “Quality of information in wireless sensor networks: A survey”, 15th Int’l Conf. on Information Quality (ICIQ 2010). Little Rock, AK, USA, 193–207. 2010.

[14] Bisdikian, C., Kaplan, L. M., and Srivastava, M. B., “On the quality of information in sensor networks”, ACM Trans. Sensor Netw, Vol. 9, no. 4, Article 48 , July 2013. DOI: http://dx.doi.org/10.1145/2489253.2489265

[15] Cheriton, D., and Gritter, M., “TRIAD: A New Next-Generation Internet Architecture”, Jan 2000.

[16] Al-Turjman, F. , Alfagih, A. , and Hassanein, H. ,“A Value-Based Cache Replacement Approach for Information-Centric Networks”, *In Proc. of the IEEE Local Computer Networks (LCN),* Sydney, Australia, 2013, pp. 902-909.

[17] Ming, Z., Xu, M., Wang, D., "Age-based cooperative caching in Information-Centric Networks," *Computer Communications Workshops (INFOCOM WKSHPS), 2012 IEEE Conference on* , pp. 268 - 273, March 2012. doi: 10.1109/INFCOMW.2012.6193504.

[18] Heinzelman, W.B., Chandrakasan, A.P., and Balakrishnan, H., "An application-specific protocol architecture for wireless microsensor networks," Wireless Communications, IEEE Transactions on , vol.1, no.4, pp.660-670, Oct 2002. doi: 10.1109/TWC.2002.804190.

[19] Alsbou, T. A. A., Hammoudeh, M., Bandar, Z. and Nisbet, A. , "An overview and classification of approaches to information extraction in wireless sensor networks," in *Proc. of the 5th Intl. Conference on Sensor Technologies and Applications (SENSORCOMM '11)*, Nice, Saint Laurent du Var, France, IARIA, 2011.

[20] Akbas, M.I., and Turgut, D., "Lightweight Routing with Dynamic Interests in Wireless Sensor and Actor Networks." *Elsevier Ad Hoc Networks*, vol. 11, no. 8, pp.2313–2328, November 2013.

[21] Clark, D. D, Partrige, C., Ramming, J. C., and Wroclawski, J. T.,"A knowledge plane for the Internet," Proc. SIGCOMM 2003, pp. 3-10, 2003.

[22] Thomas, R. W., Friend, D. H., DaSilva, L. A., and MacKenzie, A. B. , "Cognitive networks: Adaptation and learning to achieve end-to-end performance objectives," IEEE Commun. Mag., vol. 44, no. 12, pp. 51-57, 2006.

[23] Boyd, J., "A discourse on winning and losing: Patterns of conflict," 1986.

[24] Vijay, G., and Ibnkahla, M., "CCAWSN: A Cognitive Communication Architecture for Wireless Sensor Networks", In Proc. 26th Biennial Symposium on Communications, QBSC 2012, pp. 132-137.

[25] Al-Turjman, F., Hassanein H., and Ibnkahla, M., "Towards prolonged lifetime for deployed WSNs in outdoor environment monitoring", *Elsevier Ad Hoc Networks Journal*, vol. 24, no. A, pp. 172 – 185, Jan., 2015. *DOI: 10.1016/j.adhoc.2014.08.017*.

[26] Al-Turjman, F., Hassanein H., and Ibnkahla, M., “Quantifying connectivity in wireless sensor networks with grid-based deployments”, *Elsevier: Journal of Network & Computer Applications*, vol. 36, no. 1, pp. 368-377, Jan, 2013.

[27] Al-Turjman, F., Hassanein H., and Ibnkahla, M., “Efficient deployment of wireless sensor networks targeting environment monitoring applications”, *Elsevier: Computer Communications Journal*, vol. 36, no. 2, pp. 135–148, Jan. 2013.

[28] Chen, D., and Varshney, P.K.,“QoS support in wireless sensor networks: A survey”, Proc. Intl. Conf. on Wireless Networks, (ICWN) 2004.

[29] Park, P., Di Marco, P., Soldati, P., Fischione, C., and Johansson, K.H., "A generalized Markov chain model for effective analysis of slotted IEEE 802.15.4," *Mobile Adhoc and Sensor Systems, 2009. MASS '09. IEEE 6th International Conference on* , vol.130, no.139, pp. 12-15 Oct.2009. doi: 10.1109/MOBHOC.2009.5337007

[30] Online :http://grouper.ieee.org/groups/802/15/pub/2003/Jan03/03036r0P802-15_WG-802-15-4-TG4-Tutorial.ppt

[31] ITU-T Series Y recommendation: ITU-T Y.2221; “Requirements for support of ubiquitous sensor network applications and services in the NGN environment”, Jan. 2010.

[32] Haykin, S., “Cognitive radio: Brain-empowered wireless communications”, IEEE J Sel Area Comm, no. 23, pp 201–220, 2005.

[33] Mitola, J. and Maguire, G. Q., “Cognitive radio: Making software radios more personal,” IEEE Personal Communications, vol. 6, no. 4, pp. 13–18, 1999.

[34] Friend, D. H. , Thomas, R. W. , MacKenzie, A. B., and L. A. DaSilva, “Distributed learning and reasoning in cognitive networks: Methods and design decisions," in Cognitive Networks

- Towards Self-Aware Networks (Q. H. Mahmoud, ed.), pp. 223-246, John Wiley & Sons, 2007.

[35] Singh, G., Abu-Elkheir, M., Al-Turjman**,** F., and Taha, A., “Towards Prolonged Lifetime for Large-scale Information-Centric Sensor Networks” *In Proc. of the IEEE Queen's Biennial Symposium on Communications (QBSC),* Kingston, ON., Canada, 2014, pp. 87-91.

[36] Singh, G., and Al-Turjman**,** F., “Cognitive Routing for Information-Centric Sensor Networks in Smart Cities” *In Proc. of the International Wireless Communications and Mobile Computing Conference (IWCMC)*, Nicosia, Cyprus, 2014, pp. 1124 - 1129.

[37] Online: http://en.wikipedia.org/wiki/Analytic_hierarchy_process.

[38] Steels, L., “Frame-based knowledge representation”, Working paper 170, Cambridge, MA, MIT AI Laboratory, 1978.

[39] Stengel, R., “Lecture slides on “Knowledge Representation,” Online: http://www.princeton.edu/~stengel/MAE345Lectures.html.

[40] Dimokas, N., Katsaros, D., Tassiulas, L., and Manolopoulos, Y., “High performance, low complexity cooperative caching for wireless sensor networks”, *J. Wireless Networks, vol.* 17, no. 3, pp. 717-737, April 2011.

[41] Zayani, M-H., and Gauthier, V., "Usage of IEEE 802.15.4 MAC–PHY Model", Online: http://www-public.it-sudparis.eu/~gauthier/Tools/802_15_4_MAC_PHY_Usage.pdf

[42] Zayani,M.-H., Gauthier, V., and Zeghlache, D., "A Joint Model for IEEE 802.15.4 Physical and Medium Access Control Layers," In proc. of IEEE The 7th International Wireless Communications and Mobile Computing Conference (IWCMC 2011), 2011.

Chapter 5

Learning Data Delivery Paths in QoI-Aware Information Centric Sensor Networks

Preface

This chapter has been submitted to IEEE Internet of Things Journal's Special Issue on Large-scale Internet of Things: Theory and Practice.

In Chapter 4, we had proposed the CICSN architecture with local and global cognitive nodes. We had focused mainly on data delivery using knowledge representation and reasoning at the cognitive nodes. In this chapter, we elaborate on the learning techniques that can be used with reasoning to improve the cognitive capabilities at the local cognitive nodes of the CICSN, operating in a large-scale deployment.

5.1 Abstract

In this paper, we envision future sensor networks to be operating as information gathering networks in large scale Internet of Things applications such as Smart Cities, which serve multiple users with diverse Quality of Information (QoI) requirements on the data delivered by the network. To learn data delivery paths that dynamically adapt to changing user requirements in this Information Centric Sensor Network (ICSN) environment, we make use of cognitive nodes that implement both learning and reasoning in the network. In this paper, we focus on the learning strategies, and propose two techniques, namely, Learning Data Delivery A* (LDDA*) and Cumulative-Heuristic Accelerated Learning (CHAL) that use heuristics to improve the success rate of data delivered to the sink in the cognitive ICSN. While LDDA* updates a single heuristic function to choose paths that can deliver data with good QoI to the sink, CHAL accumulates heuristic values from multiple observations from the environment to choose data delivery paths that are more resource aware and considerate towards the energy consumption of the network. From Matlab simulations we observed an improvement of about 40% in the average rate of successful data delivery to the sink with the use of heuristic learning, when compared with a network that didn't implement any learning. Further, LDDA* delivered information with consistently good QoI to the network's sink, while CHAL was the more energy efficient technique. However, this performance improvement came at the cost of a 28% reduction in network lifetime when compared with a non-learning network.

5.2 Introduction

Wireless Sensor Networks (WSNs) have evolved from simple, application-specific deployments, to being an integral and essential part of large-scale application platforms in the Internet of Things (IoT) domain. Specifically, Smart-city applications of the IoT, involving large-scale deployments of WSNs have gained a lot of research attention in recent times. The SmartSantander project in Spain is an example of one such city-scale research project involving the deployment of over 3000 sensor and relay nodes within the city, supporting multiple applications [1]. Environmental monitoring, Outdoor parking area management, parks and gardens irrigation are some of the many use cases being tested on the test-bed deployed in the city. The Santander project, and other such smart-city projects [2] have provided a platform for researchers to experiment with the routing protocols, network coding schemes, and data mining techniques in a large-scale, multi-user application platform for sensor networks. Here, the sensor network has to deal with large amounts of data, support requests from multiple users, and support information extraction from the network rather than serving as point-to-point communication network, and transmit data from multiple information sources to the sink. Figure. 5-1 represents such a framework that supports different types of users as the IoT user base, including individual users, private data centers and government agencies. These users interact with the sensor network through an internet-based interface access network, which is perceived to be an information centric network (ICN) in the future [3]. Data gathered from the sensor network is delivered to the interface access network through gateway nodes. These gateway nodes could either communicate directly with the sink node of the sensor network, or could be distributed throughout the network to provide multiple access points. Currently, the experiments are focused on enabling each application as a separate entity on the test bed. However, in the real world, many of the applications run simultaneously, and the network receives requests from multiple end-users in IoT user-base at the same time.

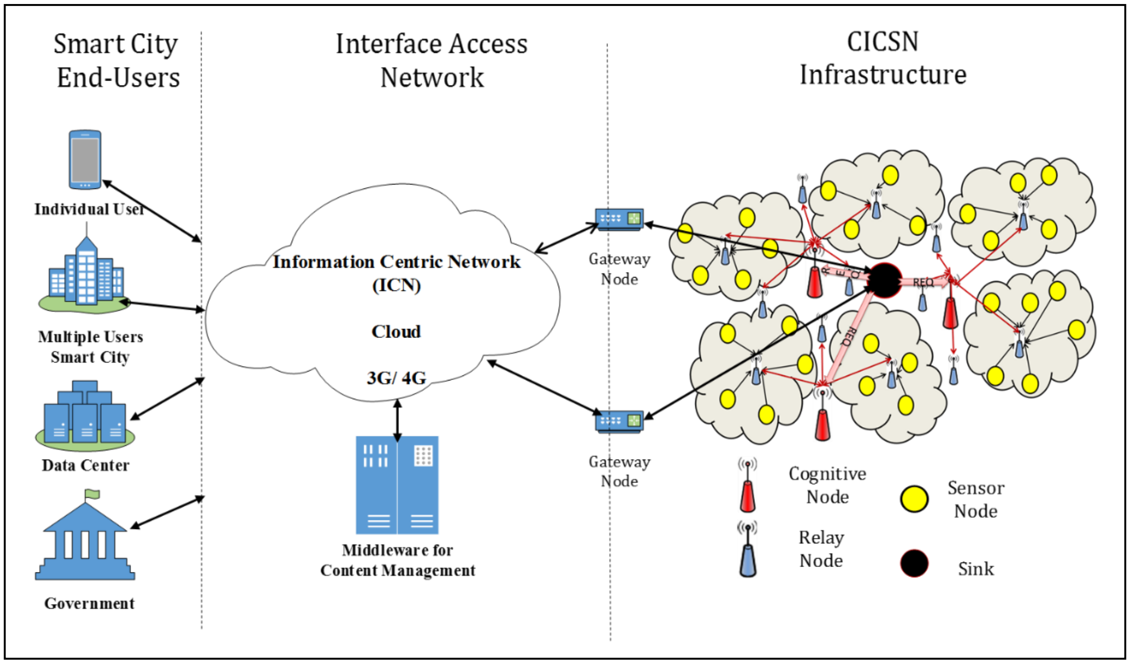


Figure 5-1: Conceptual design for Information Centric Sensor Network in an IoT application.

For example, in a smart city application, the same infrastructure (sensor and relay nodes) that is used in periodic monitoring of traffic intensity, is also used for outdoor parking management and to provide information about traffic – congested areas to users' on-demand. It can even be used to send out high-priority alerts about hazardous road conditions or accidents to users. The information generated by each of these request types has different attributes associated with it. While the periodically monitored information needs to be reliable, it does not have a strict upper bound on the time taken to gather and deliver the data, as long as it happens before the end of the stipulated time period. However, for on-demand requests generated by the user, such as a user requesting to know the availability of a free parking space in a region, the information has to be delivered quickly (low latency) to the user. In case of emergency alerts, the information must be transmitted reliably and as quickly as possible, to all users in the area. This shows that the sensor network must be able to segregate the requests and manage the heterogeneous traffic flows in a way that satisfactorily responds to the end-user in terms of the perceived quality of information for each request type. Latency, reliability, accuracy, relevance, and robustness are some of the attributes that collectively provide an estimate of the Quality of Information (QoI) perceived by the user [4]. To enable such QoI-aware data delivery, we had proposed the use of cognition in the underlying sensor network in our work [5-7]. Cognition refers to the ability to be aware of the environment, be able to learn from the past actions and use it to make future decisions that benefit the network [8]. We made use of knowledge representation using attribute-value pairs, reasoning using analytic hierarchy process (AHP), and learning as the cognitive elements that help in adapting the data delivery paths based on end-user requirements. AHP analysis[*1] was applied on the QoI attributes of latency, reliability and throughput to prioritize these attributes based on the traffic type generated by each request, and data was delivered accordingly to the end-user. The traffic type could be one of periodic, on-demand or emergency traffic, and the paths were chosen based on the priorities associated with the QoI attributes for each of the traffic types. In addition to using cognition, an information-centric approach was used for named-data identification and delivery, while keeping the QoI in focus. This cognitive information centric sensor network (CICSN) was able to achieve better performance when compared with non-cognitive data centric sensor networks [5]. Although we were able to observe the benefits of using cognition to make dynamic data delivery decisions, the CICSN framework did not have an effective learning mechanism to influence data delivery decisions that would be useful to the network over the long-term.

[*1] *AHP Analysis technique for making QoI-aware data delivery decisions has been explained in detail in Appendix I.*

In the work presented here, we propose the use of use of two heuristic learning techniques to improve the data delivery performance of Cognitive Information Centric Sensor Networks, namely, Learning Data Delivery A* (LDDA*) and Cumulative-Heuristic Accelerated Learning (CHAL). Both the learning techniques work with AHP based reasoning to achieve the networks' goals, which is to increase the success rate of data delivered to the sink while satisfying the end-user's requirements in terms of QoI or energy conservation, as the case may be. Their performance is compared with a non-learning network that uses only AHP based reasoning, and with a network that uses reinforcement learning. The metrics used for comparison of these algorithms are the success rate of data delivery to the sink, the network lifetime, the QoI delivered at the sink, and the energy consumed during the data delivery process.

Accordingly, we summarize the main contributions of this chapter as follows:

i. We propose two learning techniques to improve the average rate of successful data delivery at the network's sink.

ii. We identify the learning technique that is capable of providing better QoI of data delivered to the sink.

iii. We also compare the energy consumed by the learning techniques in terms of the energy remaining in the network at the end of the network's lifetime, to identify the more-energy efficient learning approach between the two proposed techniques.

Together, improving the QoI and successful data delivery rate at the sink contribute towards improving the quality of the end-user's experience in terms of guaranteeing a response to each of their queries in a multi-user IoT application environment. The remainder of the chapter has been organized as follows: We present related work in learning techniques used in sensor networks in Section 5.3, followed by the system models in Section 5.4. Details of the heuristic planning algorithms are presented in Section 5.5. In Section 5.6 simulations results are presented before concluding the paper in Section 5.7.

5.3 Related work

Learning and reasoning have been identified as the two important elements of cognition that can be applied to wireless networks to make them cognitive [8-10]. While reasoning helps to decide on the immediate actions to be taken by a network, learning is used to achieve the long-term goals of the network, such as maximizing the number of transmissions in which data is successfully delivered to the sink over the network's lifetime, and consistently providing data with good QoI to the end-user. Learning algorithms can be broadly classified into two categories, namely, online and offline algorithms [11-17]. This classification is based on whether the entire data required to solve a problem, or the entire sequence of requests for an algorithm are known from the beginning or not. On-line algorithms receive a sequence of requests and process them one at a time, consuming resources and incurring a cost for each action taken, without knowing the future requests or knowing the impact of their action on the future. Off-line algorithms on the other hand, know the possible sequences of requests in advance, and are able to choose their actions optimally [15]. The 8-Queen's problem, and Tower of Hanoi are examples of problems where off-line learning is used [12]. In networking tasks such as adaptive routing, identifying low cost and energy-balanced data delivery paths, and in information processing tasks such as data aggregation and inference, learning needs to be done on-line [16]. This is because neither the sequence of requests, nor the consequences of the actions taken in response to requests from the user can be known from the beginning. Particularly for wireless sensor networks operating in the IoT application environment, involving large-scale operations and active participation from end-users, on-line learning will help the network cope with varying end-user requirements in the dynamic network environment. We now briefly review some of the learning techniques that have been commonly used in wireless sensor networks.

5.3.1 Reinforcement learning

Reinforcement learning (RL) [11, 13] is a rewards-based, *on-line learning algorithm* that is commonly used in WSNs because of its ability to learn incremental changes, without the aid of any prior information about the environment in which it is operating. It emphasizes on learning while interacting with the environment, without relying on explicit supervision or requiring a complete model of the environment. Concepts of learning have been applied to cognitive sensor networks as well [10, 18], and software agents have been used to implement distributed intelligence [19, 20]. However, RL algorithms have a major drawback, which is very slow learning rates. These algorithms converge only after multiple iterations and extensive exploration of the state-action space, which can be very time consuming [21]. In large-scale dense sensor networks with IoT applications, where multiple-users are able to access the network, leading to heterogeneous traffic flows and dynamic resource consumption patterns [22], network conditions change before the learning from a previous set of observations can be applied usefully to the network, making these on-line RL algorithms too slow and expensive for the resource-limited sensor network.

In more recent research, heuristics have been considered to accelerate the RL algorithms in robotic applications such as robot navigation. Bianchi *et al.* [17, 23] have explored the use of heuristics to accelerate the convergence rate of RL algorithms, and enhance the learning rate by selecting promising actions during the learning process. Heuristics could be inferred from prior domain knowledge of the problem, by using a case-based approach, by using the solution of a previous sample problem as the initialization heuristic for a new problem, or by using the information from the learning process to infer a heuristic during execution time [24]. However, all the considered applications have been for the Artificial Intelligence domain and have been designed for specific applications with well-defined domain knowledge such as in robot navigation and robot soccer. This is very different from the dynamic environment that cognitive agents in WSNs have to support in an IoT application environment, where there are *multiple applications being simultaneously supported on the same hardware platform*, making it difficult to define the domain knowledge at the cognitive nodes that will be sufficient to handle the huge traffic volumes and diverse traffic flow patterns.

5.3.2 Informed search using heuristics

In another domain involving problem solving using informed search, heuristics are often used by Artificial Intelligence (AI) agents to improve the average-case performance on a problem-solving task [12]. Heuristics are rules of thumb that make use of problem-specific knowledge to find better estimates of the solution cost, when compared against uninformed search strategies such as breadth-first search, depth-first search, uniform-cost search and their variants. A-star, commonly represented as A* is a *heuristic offline search algorithm* based on best-first search, where the heuristic is an estimate of the remaining path length to reach the goal. This algorithm produces optimal path lengths to the goal state if the heuristic function *(h(n)*) is admissible, i.e. it never over estimates the remaining path length. The cost associated with reaching a goal state 'f' is denoted by $f(n)$, which is the sum of the path costs from initial state to intermediate state n (denoted by $g(n)$), and a heuristics estimate of the path cost from the intermediate state n to goal the state (denoted by $h(n)$). This way, $f(n)$ provides an estimate of the least cost path through the intermediate state n.

In problem domains such as data delivery in WSNs, this A* search can be applied to finding paths to the same Sink node, but with different start nodes based on the source LCN containing the requested information. This can be mapped to solving a problem repeatedly with the same goal state, but different initial states, and in such a situation, it is desirable that the algorithm is able to improve its performance over time. This is done by storing the best values of $f(n)$ at each node that is expanded, and making use of the heuristic values from a solved problem instance as the minimal value for the next problem instance. This involves *learning in real-time, and is an on-line version of the A* heuristic search algorithm*, called Learning Real-time A* (LRTA*). This on-line algorithm is more suitable for finding data delivery paths in WSNs compared to its off-line counterpart A*, as neither the sequence of requests, nor the impact of the actions taken in response to these requests are known *a priori*. In addition, the channel conditions and availability of network nodes also varies with time, making the network environment a dynamically changing one. So, an improvisation of LRTA*, called Learning in Real-Time Search (LRTS) was proposed to improve real-time, online routing decisions in WSNs [16]. This algorithm made use of a backtracking mechanism controlled by a learning quota, along with a weighted heuristic, to find data delivery paths from randomly deployed sensor nodes in a target area to a fixed sink. Although LRTS was only an extension of the distance vector routing (DVR) algorithm, the authors were able to show that with the heuristic weight and learning mechanism, LRTS was able to reduce the total network traffic, and find near-optimal routes faster, and at lower energy consumption. However, this was only applied to an application-specific sensor network, and the ability of this algorithm to scale to a larger network, with *multiple messages targeted to the same destination simultaneously through the network, with multi-user access*, such as in WSNs applications targeted for IoT applications, has not been explored. In addition, the ability to *deliver* *data with different QoI requirements between the same set of source and goal nodes* has not been explored either, which will often be the case for WSNs operating in the IoT application domain.

In this work, we explore the use of heuristics to accelerate the convergence of learning algorithms in large scale information centric sensor networks for IoT applications, supporting multiple users and multiple applications on a common platform. As described in the context of cognitive psychology, the heuristics will be used as strategies that ignore a part of the information to make decisions faster, and sometimes more accurately compared to more complex methods [25]. We experiment with the use of an online version of the A* heuristic search algorithm, which learns from the information available in the knowledge base of the cognitive nodes. We call this *Learning Data Delivery A* (LDDA*) algorithm.* The heuristics will be used to make approximate decision choices, as opposed to optimal decision choices. The heuristics can help the learning algorithm to converge faster, and arrive at general and scalable solutions for identifying the next hop path for delivering data towards the sink. We compare this with a *Cumulative-Heuristic Accelerated Learning (CHAL)* technique that accumulates the heuristic values at each state (relay and cognitive nodes), and makes use of as much information as possible from observations made in the network before making the data delivery path choices. It also uses negative heuristic weights to punish poor next-hop node choices, such as revisiting a node along a data delivery path. This way, LDDA* and CHAL will differ in the heuristic weights accumulated by the learning process. Since learning is typically used to improve the decisions made by reasoning engine in cognitive networks, we implement LDDA* and CHAL in a network that uses an Analytic Hierarch Process (AHP) based reasoning technique at the cognitive nodes to make data delivery decisions. (The details of AHP based data delivery (AHPDD) in CICSNs have been described in our previous work [5], and in Appendix I.) Performance of the heuristically accelerated learning techniques LDDA* and CHAL, are compared against the non-learning AHPDD in terms of the QoI observed at the Sink where data is delivered at the end of each transmission round. The algorithms will also be compared in terms of the rate of successful data delivery and the energy consumed during the data delivery process at the end of network's lifetime. The knowledge of the deterministic deployment of the RNs and CNs, and the knowledge accumulated in the knowledge base (KB) of the cognitive nodes, will be used to update the weight of the heuristics during network operation.

5.4 System Models

In this section we describe the CICSN network model [26], on which the heuristic search algorithm for identifying data delivery paths is implemented. We also briefly describe how heuristic functions should be defined.

5.4.1 CICSN Network model

The components comprising our CICSN network model are as follows:

1) **Sensor Nodes (SNs):** These nodes sense the environment, capture the sensed data, and transmit them to relay nodes or local cognitive nodes. They interact with the RNs and LCNs.
2) **Relay Nodes (RNs):** forward the data received from SNs to the Sink or neighboring local cognitive nodes in response to requests they receive from the cognitive nodes. They interact with SNs, LCNs and the Sink.
3) **Local cognitive Nodes (LCNs):** Cognitive nodes have elements of cognition, i.e. knowledge representation, reasoning, and learning, which help them interpret requests coming from the user and respond with data accordingly. They interact with RNs, SNs and the Sink.

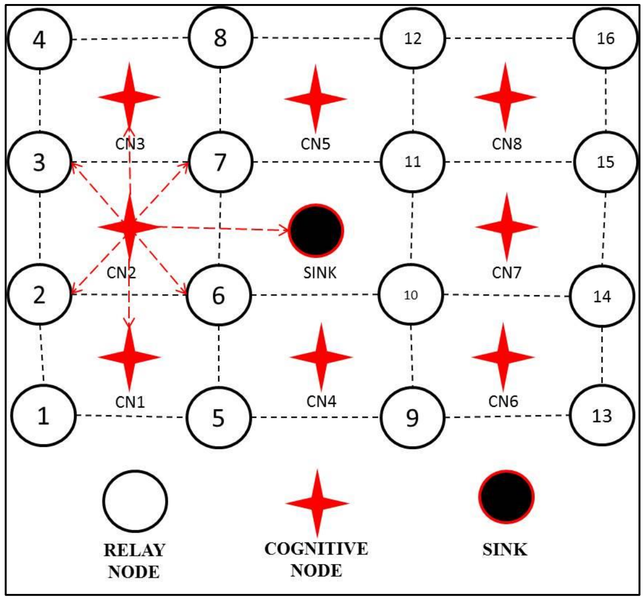


Figure 5-2: Node deployment in the CICSN.

4) **Sink:** This is where all the data gathered from the network is delivered. In our model, we also assume that the user requests are conveyed to the network through the Sink**.** A Sink enhanced with cognitive capabilities to interpret the traffic type and user requests is called a Global Cognitive Node.

SNs are randomly deployed in the network, and all other nodes are deployed at deterministic locations on a two-dimensional grid as shown in Figure 5-2 [26].

5.4.2 Heuristic Functions in CICSN

Three main features that define heuristic functions are as follows: i. Evaluation Criteria, ii. Selection Criteria, and iii. Termination Criteria. Each of these criteria must be defined to formulate a heuristic function, which can in turn be used to accelerate the learning process. We make use of the A-star (also written as A*) algorithm in the CICSN framework, as an example to illustrate the use of these criteria. The heuristic function will be used to accelerate the decision choice on the next hop RN from a give source LCN, to deliver the data towards the sink. Next-hop nodes are evaluated using two values: *g(n)*, which is the cost to reach an intermediate node n; and *h(n)* is the cost to get to the sink from node n. Using these values, the A* algorithm would estimate the cost of the cheapest solution through node n using equation (1).

$$g(n) + h(n) \leq f(n) \tag{1}$$

For the CICSN, let $g(n)$ represents the cost of reaching a next-hop node from the LCN. It is a function of the Quality of Information attributes (latency, reliability and throughput). $h(n)$ is the heuristic function, whose value depends on the energy cost and the traffic type. The energy cost depends on the estimated energy to reach the sink. The traffic type affects the QoI attribute that is prioritized while deciding the path. The heuristic function $h(n)$ can take on different values, as long as it satisfies the consistency condition. A heuristic is said to be consistent if the cost of the transition from current node n to next node n' caused by action a, plus the cost of reaching the sink node (goal state) from node n' is less than or equal to the cost of reaching the goal state from node n itself. In this work, node n represents an LCN, n' could be an LCN or RN and the goal state is the sink node. Since the heuristic functions can be evaluated only at the cognitive nodes, we assume that the current node n is always an LCN where data to be delivered has been collected/stored.

i. Evaluation criteria

$$g\ (n) + h(n) \leq f(n) \tag{2}$$

ii. Selection criteria

Chosen RN index = max $(g(n) + h(n))$

iii. Termination criteria

All one-hop RNs and LCNs to Sink are dead or less than 50% LCNs remain alive in the network

As observed by Geffner [27], this simple learning will guarantee that the goal is eventually reached when the next states are fully or partially observable, and actions have probabilistic effects. If this process is restarted many times while preserving the heuristic values in the knowledge base of the LCNs, the goal state can be eventually reached optimally, depending on the initial heuristic. In the subsequent section, we will describe the heuristics techniques used in this work to accelerate the learning process.

5.5 Learning approaches for the CICSN framework

In this section, we describe the two learning techniques that have been proposed for the CICSN framework, to improve the decision of choosing the best node for forwarding data to the sink. Both algorithms make use of heuristics based on the A* algorithm, which has been widely used in informed search [12]. We start with describing the LDDA* algorithm, which is a heuristic search strategy combined with online learning. This algorithm targets the IoT user-base comprising of individual users that directly interface with the underlying network, as shown in Figure 5-1, with the goal of having a positive impact on the end-user's experience with the underlying network. LDDA* accelerates the decisions made by the learning process, and ensures good QoI for data delivered to the sink. The CHAL technique on the other hand, has been developed for the IoT user-base comprised of Data Centers and Government establishments. In these applications, the users' focus is on conserving the network's energy, so that the network is able to provide data for longer. Learning in CHAL involves the accumulation of multiple heuristics based on observations made from current network conditions and past decision choices, so as to choose data delivery paths that more definitely lead towards the sink, rather than expending energy in exploring new paths that may or may not end at the sink. The algorithms for LDDA* and CHAL have been described in the following sub-sections.

5.5.1 Heuristic search for data delivery paths using the Learning data Delivery A* (LDDA*) algorithm

Data delivery decisions in the CICSN are based on observing the dynamically changing topology of the network. Since the learning process might be too slow to respond/converge before further changes take place in the network, we choose a heuristic search strategy to aid in making quicker decisions. With the help of a model of the original network topology combined with the currently observed changes, we make use of a modified version of the A* heuristic algorithm, to identify the nodes that can be used for data delivery to the sink. This approach is useful when a problem is to be solved repeatedly with the same goal state (sink), but with different initial states (LCNs), as we would like the algorithm to improve its performance over time. In the LDDA* algorithm, RNs and LCNs are initially assigned heuristic values at the time of deployment based on their proximity to the sink. Nodes that lie at 1-hop distance to the sink, have the highest probability of successful data delivery to the sink. Hence are given the highest heuristic weight (0.1 in our study). Nodes lying further away from the sink are given lower weights (0.05), so that they have lesser influence on the heuristic decision making. However, they are not assigned a zero weight, because these node will still be able to participate in multi-hop routing in case the nodes with direct access to the sink become unavailable due to poor link conditions, network congestion or node deaths. Higher the weight of the heuristic, higher will be the chance that the node will be chosen for data transmission to the sink. These values remain fixed at each RN and LCN until the nodes die, at which time the heuristic values are made ‘0’ as they do not influence the heuristic decision anymore. We now describe the steps of the LDDA* heuristic using Algorithm1. The function name, inputs and outputs are described in Steps 1-3. Step 5 is used to initialize the transmit success variable to ‘0’ during the start of each transmission round, indicating that data has not been delivered to the sink. Once a source LCN is identified in Step 6, all its available one-hop RNs are identified in step 7. If there is at least one RN connected with the source LCN, as checked in step 8, then steps 10 – 12 are repeated for each of the available one-hop RNs. First, information about the QoI that can be offered on the path from LCN to RN is gathered in step 10. This is called *g(n)*. Next, in step 11, a heuristic estimate *h(*n) of the QoI along the path from the RN to the sink is made, which is called as *hQo*I. In step 12, these two values are summed up as *f(n)* to give the *effective QoI* (eQoI), which is the QoI on the last hop that the data travels to reach the sink during each transmission round. In step 14, the RN that offers the highest value of *f(n)* is chosen as the next hop RN to deliver data from the LCN towards the sink. If the source LCN is not connected with any RNs, then steps 16 – 20 are followed to either transmit data directly from the LCN to sink when possible, or to another neighboring LCN to keep the data moving towards the sink. In case the LCN is disconnected from the network, the knowledge base of all the LCNs and RNs are updated to reflect this information and the heuristic value is dropped to '0' for disconnected nodes. The next hop neighbor information is updated in the LCNs and RNs in Step 22 to reflect the latest observations from the network environment for the next transmission round. These steps are repeated for each transmission round till the termination criteria described in step 25 is reached. This is the condition for detecting the network's end-of-life, when all one-hop nodes to the sink are dead or when less than 50% of the cognitive decision making LCNs remain alive in the network.

Algorithm 1: Learning Data Delivery A* (LDDA*) heuristic algorithm

1. **Function LDDA (hQoI)**
2. **Input: {hQoI}** //Heuristic estimate of QoI values from each RN to Sink
3. **Output: {$\mathbf{RN_x}$}** // RN indices chosen by learning function LDDA to deliver data towards sink
4. **Begin**
5. **Initialize: Transmit success = 0; //**for each RN and LCN at the start of a transmission round
6. Identify a source LCN as start node for current transmission round
7. List all the one-hop RN indices from source LCN
8. **If** there is at least one RN connected with this LCN
9. **For** each one-hop RN index ‘i’ do
10. $g(n_i) \leftarrow$ QoI (LCN to RN_i) // QoI from LCN to RN
11. $h(n_i) \leftarrow$ hQoI (RN_i to Sink) // Heuristic estimate of QoI from RN to Sink
12. eQoI $= g(n_i) + h(n_i)$ // effective QoI from source LCN to Sink
13. **End**
14. Chosen RN index = max(eQoI)
15. **Else** //there are no one-hop RNs linked with chosen LCN source
16. **If** source LCN is connected with Sink or any other LCN
17. transmit from LCN to Sink or to neighbor LCN
18. **Else**
19. Mark LCN as disconnected from network
20. **End**
21. **End**
22. Update next-hop neighbor information in LCNs and RNs
23. **Loop**
24. **Termination Criteria**:
25. All one-hop RNs and LCNs to Sink are dead or less than 50% LCNs remain alive in the network
26. **End**

At this time, the network will not be able to respond to the end-user with useful information from the network, and hence we call it the network end-of-life. The heuristic values at the LCNs are not accumulated over previous transmission round's values for any of the connected RNs, so as to allow more flexibility in the decision choice made by the LCN in choosing the next hop RN. We vary this feature in our next technique, which is the cumulative-heuristic accelerated learning.

5.5.2 Cumulative-Heuristic accelerated learning (CHAL)

In this method, heuristics accumulate at RNs over the transmission rounds to influence the choice made by LCNs for data delivery to the sink. Three features alter the heuristic weight assigned to the RN. They are as follows: (1) Is the RN connected to the Sink? (2) Has the RN not been visited previously in the current transmission round? (3) Is the RN connected with more than one LCN? For each of these conditions satisfied, the heuristic weight at each RN is incremented by 0.1. The only time an RN is assigned a negative weight (- 0.1), is when it is revisited. A variable 'visited' is used to indicate whether a node has already been visited in the current transmission round. This variable is reset to 0 at the start of each transmission round. A visited node is less preferred over a node that has not been previously used in the current data transmission round, and the negative weights help to avoid loops while choosing data delivery paths to the sink. When a node dies, its heuristic weight is reset to '0' and the connectivity updates are propagated in the network as data delivery paths are explored. The node connectivity information is updated in the knowledge base of the LCNs. In Algorithm 2 describing the Cumulative heuristic accelerated learning (CHAL) technique, we highlight the steps involved in updating the heuristic values, and show how the decision choice is made at each LCN along the data delivery path. Inputs assigned to the CHAL function in Step 2, are the heuristic weights at each LCN and RN, initialized to either 0.1 or 0.05 based on whether or not the nodes are connected to the sink.

Algorithm 2: Cumulative-Heuristic Accelerated Reinforcement Learning Algorithm

1. **Function CHAL ($h(RN_i)$, $h(LCN_j)$)**
2. **Input: {$h(RN_i)$, $h(LCN_j)$}** //Heuristic weight at each RN and LCN assigned at start of simulation
3. **Output: {RN_x, LCNy}** // RNs or LCNs chosen by learning function CHAL to deliver data towards sink
4. **Begin**
5. **Initialize:** LCN visited = 0; RN visited = 0;
6. Identify a source LCN as start node for current transmission round
7. List all the one-hop RN indices from source LCN
8. **If** there is at least one RN connected with this LCN
9. **For** each one-hop RN index ‘i’ do
10. $g(n_i) \leftarrow$ QoI (LCN to RN_i)
11. **If** RN_i is connected with Sink
12. $h(RN_i)= h(RN_i)+0.1$
13. **End**
14. **If** RN_i is not marked as ‘Visited’
15. $h(RN_i)= h(RN_i)+0.1$
16. **Else**
17. $h(RN_i)= h(RN_i)-0.1$
18. **End**
19. **If** RN_i is connected with more than 1 LCNs
20. $h(RN_i)= h(RN_i)+0.1$
21. **End**
22. eQoI = $g(n_i) + h(RN_i)$ // effective QoI from source LCN to Sink
23. **End//** end of for loop
24. Chosen RN index = max(eQoI)
25. **If** $f(n_i) \leq h(LCN_i)$)
26. Transmit through chosen RN
27. **Elseif** $f(n_i) > h(LCN_i)$)or LCN is not connected with RNs
28. Transmit from LCN to Sink or neighbor LCN
29. **End //** end of **if** condition starting on line 25
30. **Else**
31. Mark LCN as disconnected from network
32. **End//** end of **if** condition starting on line 8
33. Update next-hop neighbor information in LCNs and RNs
34. **Loop**
35. **Termination Criteria**:
36. All one-hop RNs and LCNs to Sink are dead or less than 50% LCNs remain alive in the network
37. **End**

Output of this function is the set of RNs and LCNs chosen along the data delivery path towards the sink as indicated in step 3, for each request from the user. The variable '*visited'* is initialized to '0' at the LCNs and RNs at the start of simulation, and is also reset at the start of each transmission round, as shown in Step 5. The source LCN for each request is identified in Step 6, and the one-hop neighbor RNs of the chosen LCN are identified in step 7. The existence of at least one RN connected with the source LCN is checked in Step 8. If this condition is satisfied, then for each RN connected with the source LCN, steps 9 to step 23 are repeated. In the loop starting from Step 9, the QoI from LCN to RN is stored in $g(n_i)$ in step 10. Then the heuristic weight at each RN is calculated based on the three conditions discussed above, and implemented in steps 12, 15, 17 and 20. The effective QoI (*eQoI)* from the source LCN to sink through each RN is calculated in Step 22. It is calculated as the sum of $g(n_i)$ and $h(RN_i)$ for each RN connected with the source LCN. Among the RNs for which the heuristic has been calculated, the one with the highest *eQoI* is chosen as the next hop for data transmission from the LCN towards sink in Step 24. To ensure admissibility of the chosen heuristic, a check is performed in Step 25 to make sure that the cost of transmission through the RN does not exceed the cost of transmission directly from LCN to Sink (given by $f(n_i)$). If this condition is satisfied, then the data is transmitted through the chosen RN as mentioned in step 26. Otherwise, if the condition is not satisfied or the LCN is not connected with any other RN, then the LCN attempts to transmit data through neighboring LCNs or directly to the Sink if possible as described by steps 27 and 28. Else, the LCN is marked as disconnected from the network in step 31, and this information is updated in the LCNs and RNs which communicated with this LCN in step 33. These steps are for each transmission round in the simulation till the termination condition in step 36 is reached, which is the network's end-of-life condition. The network is no longer able to respond to user requests with useful information from the network, at the end of the network's lifetime.

5.6 Performance Evaluation

In this section, we evaluate and compare the performance of the algorithms described in this section using Matlab simulations. We start with a brief description of the simulation setup and network parameters used, and follow it up with the description of the metrics that will be used in evaluating the performance of the algorithms. Simulations results and a detailed analysis of the results will also be presented in this section.

5.6.1 Simulation setup, network parameters, and evaluation metrics

The network is setup as described in Figure. 5-2, with randomly deployed sensor nodes, and fixed deployment of relay and cognitive nodes. Simulation parameters are as described in Table 5-I. Energy deductions at the LCNs and RNs during data transmission are as represented in Table 5-II, based on the transmit powers. The transmit power at RNs is fixed at 3 dBm, and it can be adapted at the LCNs to improve the probability of successful transmission as described in our work [6]. In the simulations, the energy deductions at RNs and LCNs do not include the energy consumed during data gathering from sensor nodes at these nodes. We do not deal with the details of sensor node scheduling either. Data delivery paths from source LCNs in the network are initially established based on AHP analysis of paths along next-hop neighboring RNs. Heuristic learning is introduced in this simulation to increase the average success rate of data delivery to the sink, irrespective of the randomness with which the requests for different traffic types are generated in the network.

Table 5-I : Simulation Parameters.

Parameter	Value
Target area	1050m x 1050m
Number of nodes	SNs: 1500
	RNs: 16
	LCNs: 8
Transmit power	SN: <3dB
	RN: 3dB
	LCN: {3dB, 5dB, 7dB}
Communication Range	SN: 175m
	RN: 250m
	LCN: 350m
	GCN: 500m
Application payload size	121Bytes
Per node offered load	0-1400 bits per second

Table 5-II: Transmit power consumption.

Ptx (dBm)	Cycle life reduction (units)
3 – 5	2
5 - 7	3
7 - 9	4
≥10	5

Following are the three performance evaluation metrics that will be used to compare the performance of the various algorithms described in section 5.5:

i. <u>Network Lifetime</u>: Number of transmission rounds till all one-hop nodes to GCN/Sink node are dead (includes RNs and LCNs), or till more than 50% of cognitive decision making LCNs are dead in the network.

ii. <u>Success Rate (ρ)</u>: It is defined as the ratio of the number of times a request was successfully responded to (*s*), with data delivered to the sink, over the total number of transmission rounds during the network's lifetime (T), expressed as a percentage. In other words, it is the ratio of the number of successful transmissions to the sink over the total number of transmission rounds during the network's lifetime. This is represented by Equation 3 as follows:

$$\rho = \frac{s}{T} * 100 \tag{3}$$

iii. Failure Rate (ϕ): It is defined as the ratio of the number of failed transmissions to the sink *(f)* due to non-availability of suitable paths towards the sink or the non-availability of LCNs in the region from where information has been requested, over the total number of transmission rounds during the network's lifetime (T), expressed as a percentage. In other words, it is the ratio of the number of failed transmissions to the sink over the total number of transmission rounds during the network's lifetime. This is represented by Equation 4 as follows:

$$\emptyset = \frac{f}{T} * 100 \tag{4}$$

iv. eQoI: *effective-QoI* or eQoI is the heuristics estimate of the QoI associated with data delivered to the sink at the end of successful transmission round. It is not the cumulative value of the QoI measured hop-over-hop from the source LCN to the sink, but a heuristic estimate of the value of the QoI at the last hop that delivered the information to the sink.

v. Hopcount: Number of hops over which data is carried, starting from the source LCN till data is delivered to the sink, or till transmission is abandoned for the current transmission round as no paths can be found to the sink. These hops include both RNs and LCNs.

vi. Remaining Energy: This metric represents the battery level at the end of network lifetime in the LCNs and RNs.

5.6.2 Simulation Results and analysis

Simulation results for the four techniques described in the methodology section (IV) are summarized in Table 5-III.

Table 5-III: Summary of Simulation Results.

Method	Lifetime (rounds)	Average Success Rate	Average Failure rate	Best-Case Success Rate	Worst-Case Failure Rate
AHPDD	78	63	37	79	52
RL	106	47	53	68	63
CHAL	59	84	16	90	19
LDDA*	56	88	12	92	22

Results from the simulation using AHP analysis (AHPDD) suggests that during an average lifetime of 78 transmission rounds, the average success rate is 63%, and the average failure rate is 37%. However, during the worst case, transmissions can fail for over 50% of the requests, as suggested by the worst case failure rate.

Heuristic values were accumulated at nodes that delivered data to the sink, and nodes that were re-visited were penalized to modify the decisions made by the AHP analysis. It was found that the average success rate had dropped to 47%, and the worst case failure rate had increased to over 60%. This decrease in success rate is attributed to the performance deterioration of the learning algorithm as simulation progresses, due to node deaths and the resulting poor network connectivity. The network degradation leaves fewer network nodes available to the learning algorithm to support exploration of the environment. Loss of network nodes (RNs and LCNs) leads to accumulation of more penalties than rewards as nodes get re-visited, without being able to deliver data to the sink. But the requests continue to be sent out to the network till end of the network lifetime, i.e., when either all one-hop neighbors of the sink are dead or all LCNs are dead. This leads to an increase in the number of failed transmissions compared to total number of requests accepted by the network, thus causing the average success rate to decrease. With the cumulative-heuristic accelerated learning (CHAL), it was found that the average success rate increased to 84%, and the worst-case failure rate was as low as 19%. The best case success rate was 90%, which was only 6% off from the average success rate. This shows that the heuristics performed consistently well under various traffic loads and request arrival patterns. The performance of CHAL was matched very closely by the LDDA* heuristic search algorithm, which provided an 88% data delivery success rate, but a slightly higher failure rate of 22% in the worst-case scenario when compared with CHAL. Figure 5-3 to Figure 5-6 provide a visual representation of the relative performance of the four techniques described in the methodology section.

For a CICSN deployed in a smart environment, the success of the network is evaluated based on the number of times the network is able respond successfully to the end-user's request. Although it is desirable to improve the network's longevity, longer network lifetimes are not a good indicator of the ability of the network to support better data delivery success rates. Figure 5-3 shows a comparison of the network lifetimes achievable with the four different techniques, and Figure 5-4 shows a comparison of their average success and failure rates. We observe that RL offers the best network lifetime, but its success rate is poor. The two other techniques LDDA* and CHAL that have lower network lifetimes of about 60 transmission rounds, perform very well in terms of the average success rates of data delivery, with over 80% success rate, while keeping the failure rates below 20%. LDDA* provides slightly higher average and maximum success rates, compared with CHAL. It is a simple, yet effective heuristic that is able to make good approximations about the best paths along which data can be successfully delivered to the sink.

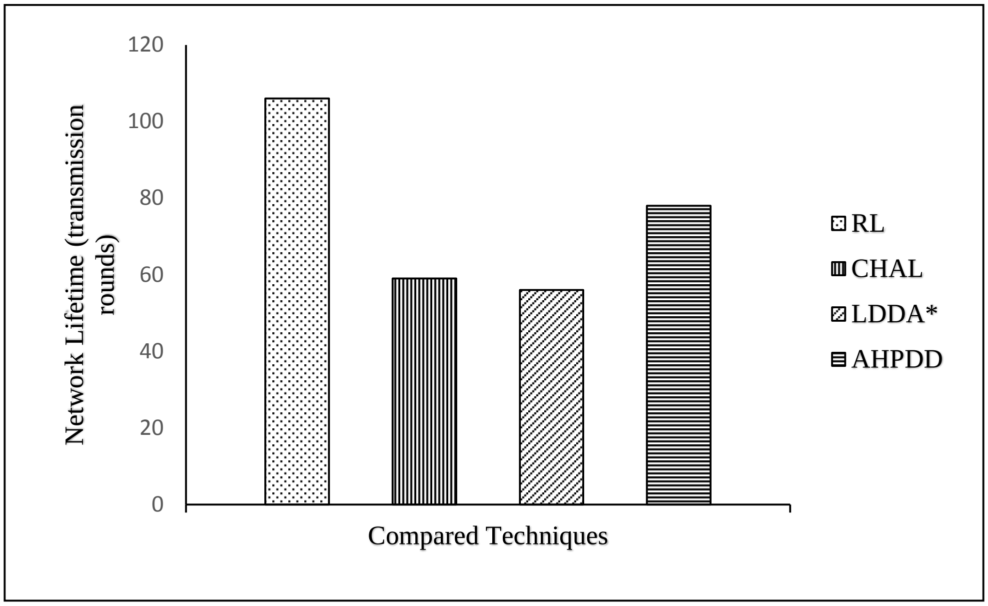


Figure 5-3: **Comparison of Network lifetimes.**

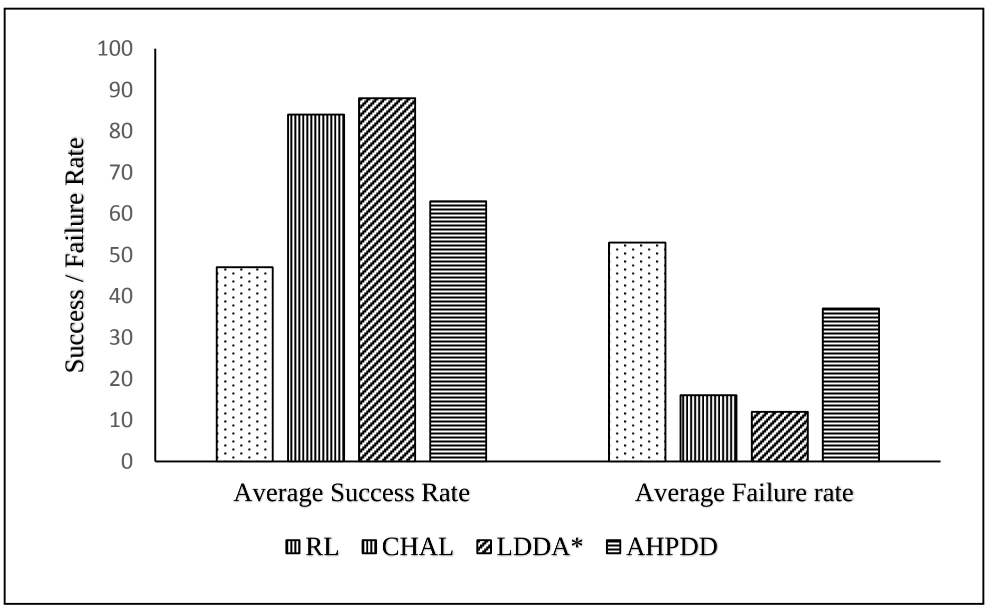


Figure 5-4: **Comparison of Average Success and Failure Rate.**

Together, Figures 5-3 and 5-4 represent the successful data delivery rate versus lifetime tradeoff in the network. Although there is a 50% decrease in lifetime, LDDA* and CHAL provide over 50% improvement in success rate of data delivery compared with the other two techniques. The extra energy consumption goes towards ensuring guaranteed data delivery, while satisfying user's QoI needs for each request.

Considering the fact that these techniques are being developed for use in an IoT environment, it is desirable to improve the success rate of data delivery to improve the user's experience of interaction with the network. Being more sure of getting a response back from the network for each of the queries sent out, improves the levels of satisfaction of the user's interaction with the network. As for the network lifetime, energy harvesting techniques or wire-line powered communications can be used at the network nodes when possible to increase the network's longevity. Figure 5-5 and Figure 5-6 compare the average, minimum and maximum values of success rates and failure rates respectively, for all the four techniques. We see that the LDDA* heuristic has the highest average success rate of 88%, and its worst-case failure rate is only 3% more than CHAL. Since it is more desirable to have a higher success rate in smart IoT applications, we further compare the performance of LDDA* and CHAL techniques in terms of their effective QoI (eQoI) as observed at the Sink to identify the best heuristic of the two.

Figure 5-7 shows the result of the comparison of the eQoI values for LDDA*, and CHAL, with AHPDD, which doesn't use any form of learning at the LCNs. In general, we observe that using some form of learning at the LCNs improves the eQoI of the data delivered to the Sink. Of the learning techniques, we observe that LDDA* performs the best in terms of consistently delivering data with higher eQoI at the sink, even towards the end of the network's lifetime. Now, this eQoI is the hop-over-hop value of QoI associated with the data delivered to the sink with respect to latency, reliability and throughput. Apart from the hop-over-hop latency, the cumulative delay in receiving a response from the network for a request is reflected by the number of hops taken along the path from source to sink.

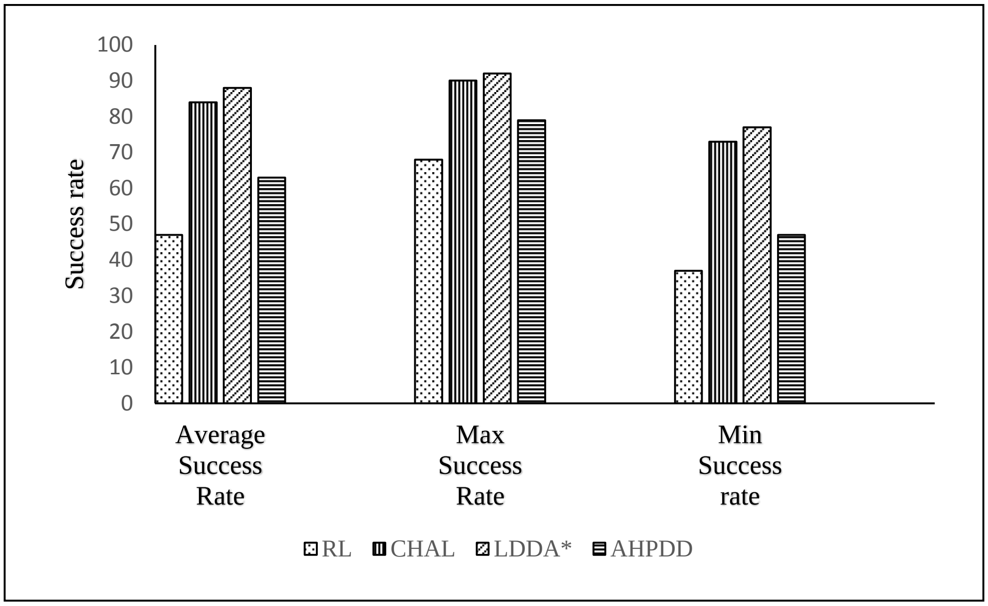


Figure 5-6: Comparison of Success rates.

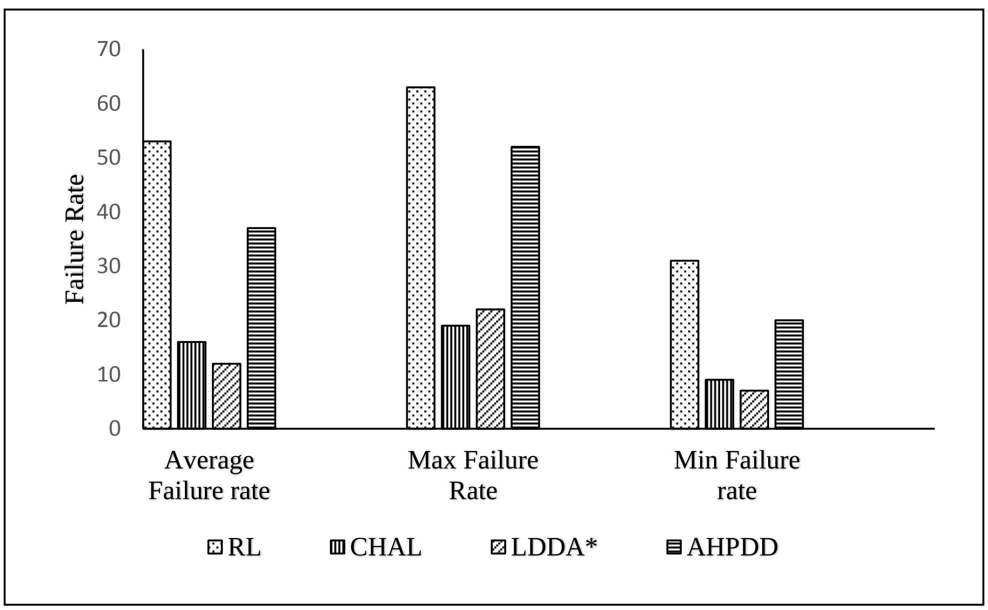


Figure 5-5: Comparison of Failure Rate.

Keeping this in view, we compare the hop count over the network lifetime for LDDA*, CHAL, and AHPDD, and the results are as shown in Figure 5-8. The observation made from the compared techniques is the spike in hop count seen towards the end of the network's lifetime. This is because more number of nodes are lost due to node deaths as the simulations progress, making it increasingly difficult for the remaining alive nodes to find paths to deliver data to the sink. This search for alternate paths leads to an increase in the hop count. However, one of the marked differences between LDDA* and CHAL is that CHAL has a constant hop-count of 2 till about 30 transmission rounds. This is because it starts out with exploiting its knowledge of paths through RNs that are one-hop away from the sink. But LDDA* starts with exploring paths and eventually learns the network connections that helps to reduce the worst case hop-count towards the end of the network's lifetime. This difference in strategies accounts for the slightly lower network lifetime of LDDA* (as observed from Figure 5-3) when compared with CHAL, due to higher energy consumed in exploring the paths. However, when averaged over the entire network lifetime, LDDA* is on average, 2 hops more expensive than CHAL, which it compensates very well with its higher average rate of successful data delivery and eQoI at the sink. The eQoI values and hop counts for Figures 5-7 and 5-8 have been taken from simulations for which the average success rate and network lifetime lie close to average values as reported in Table 5-II for each of the compared algorithms.

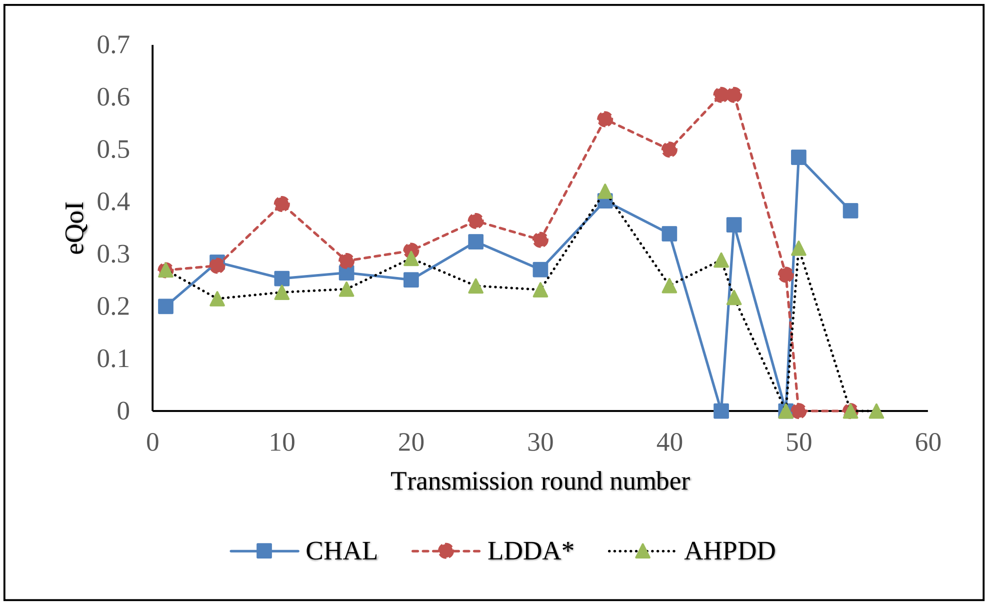


Figure 5-8: Comparison of eQoI as observed at the Sink over the network lifetime.

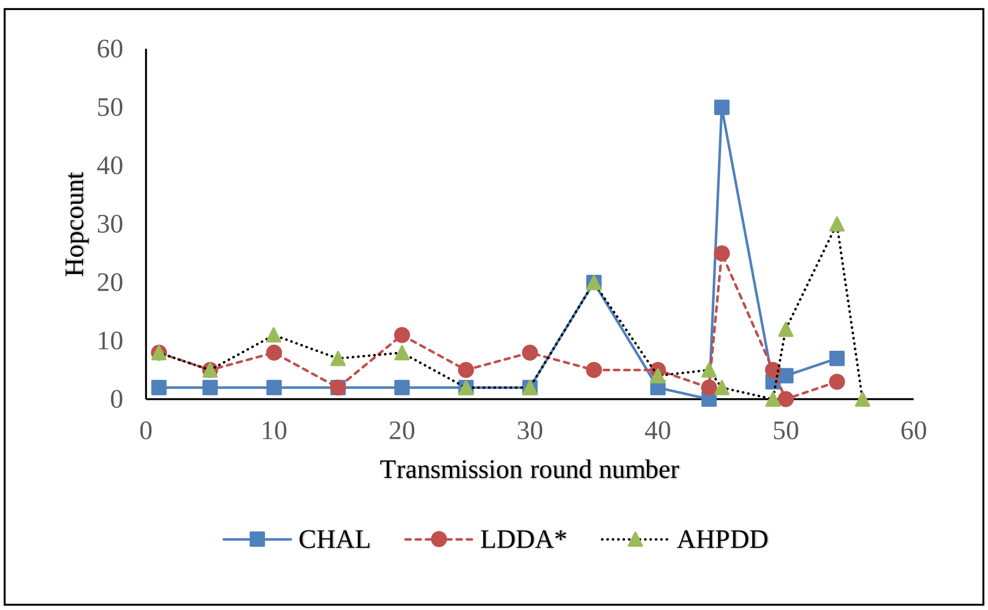


Figure 5-7: Comparison of hop counts in delivering data from source to sink.

Figure 5-9 compares the remaining energy at the LCNs and RNs for LDDA*, AHPDD, and CHAL at the end of the network lifetime. The dotted-lines are the linear trend lines for each of the compared techniques. They show an increasing trend for the remaining energy at the RNs and LCNs, as the distance from the sink increases, for all the compared techniques. This is because the one-hop neighbors of the sink are the nodes that get most frequently used in delivering data to the sink, and the nodes at the edge and corners of the network do not participate in data delivery towards the sink, as much as the nodes lying closer to the sink do. They mainly aid in gathering data from sensor nodes lying in those regions. On inspecting the differences in the trend lines, we observe that the trend line for CHAL lies above those for LDDA* and AHPDD, indicating that it has higher remaining energy at its nodes at the end of the simulation. So, CHAL may be considered more energy efficient compared to the other techniques. LDDA* has the next highest average amount of remaining energy at the end of simulations. LDDA*'s higher energy consumption may be attributed to its higher hop count compared to CHAL during the first half of the network's lifetime. Thus, we can conclude that of the two proposed techniques, LDDA* is capable of delivering data to the sink with a higher average success rate, with better eQoI, whereas CHAL is the more energy efficient technique. Either of these techniques may be used for data delivery in the CICSN for IoT applications, based on the application and end-user requirements on the eQoI, rate of successful data delivery, energy efficiency and cumulative delay from source to sink.

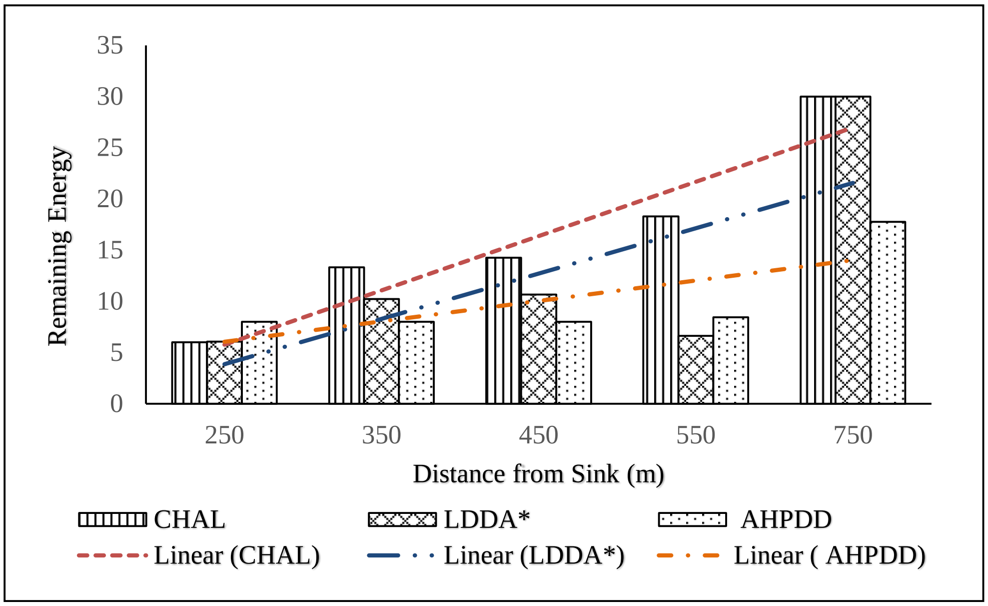


Figure 5-9: Remaining Battery Level at RNs and LCNs with Trend lines.

5.7 Conclusions

In this paper, we studied the use of heuristically accelerated learning techniques in improving the data delivery success rate of cognitive information centric sensor networks. We examined the performance in terms of impact on the network lifetime, average success and failure rates, and the energy consumption at the relay and cognitive nodes, and the QoI at the sink. We found that heuristics, when chosen correctly, can be simple, yet effective in achieving user-defined goals. The LDDA* heuristic was able to provide a 40% improvement in average data delivery success rate, when compared with the CICSN that used only AHP based reasoning and knowledge representation as its cognitive elements. The CHAL algorithm performed equally well in terms of the data delivery success rate, but, performed slightly better than LDDA* in terms of the energy consumed during the data delivery process, as reflected by the higher remaining energy at the end of the network lifetime. On the other hand, the LDDA* algorithm is a better choice when the application requires a higher eQoI at the sink, and higher best case success rate.

However, the tradeoff in using the heuristic learning techniques was a 28% reduction in the network lifetime, when compared with the non-learning CICSN that used AHPDD for data delivery. But this reduced lifetime was a more accurate estimate of the duration for which the CICSN could respond to end users requests, and deliver data for over 85% of the requests.

An interesting extension of this work is to consider the impact of change in density ratio of the RNs and LCNs around the sink, on the QoI delivered at the sink. As observed from the remaining battery level of the RNs' at the end of simulations, the nodes at the corners and edges of the network are most often not actively involved in data delivery. Changing the initial allocation of the energy at these nodes or considering a different deployment plan and re-evaluating the impact of the learning on the network lifetime and success rate will be interesting future work. In addition, it will be interesting to test these strategies on a simulator that supports sensor network simulations, and even on a hardware test bed.

5.8 References

[1] SmartSantander Project, Santander facility, [Online]. Available: http://www.smartsantander.eu/index.php/testbeds/item/132-santander-summary.

[2] European Smart Cities 3.0 (2014). [Online]. Available: http://www.smart-cities.eu/?cid=01&ver=3.

[3] B. Ahlgren, C. Dannewitz, C. Imbrenda, D. Kutscher, B. Ohlman, "A survey of information-centric networking", IEEE Communications Magazine, vol.50, no.7, pp.26-36, July 2012.

[4] C. Bisdikian, L.M. Kaplan and M.B. Srivastava, "On the Quality of Information in Sensor Networks", ACM Trans. Sensor Netw, Vol. 9, no. 4, Article 48 , July 2013.

[5] G.T. Singh and F.M. Al-Turjman, “A Data Delivery Framework for Cognitive Information-Centric Sensor Networks in Smart Outdoor Monitoring”, Elsevier, Comput. Commun. Jan. 2015. http://dx.doi.org/10.1016/j.comcom.2015.01.002.

[6] G. T. Singh, M. Abu-Elkheir, F. M. Al-Turjman, and A. Taha, “Towards Prolonged Lifetime for Large-scale Information-Centric Sensor Networks” *In Proc. of the IEEE Queen's Biennial Symposium on Communications (QBSC)*, Kingston, ON. Canada, 2014, pp. 87-91.

[7] G. T. Singh, and F. M. Al-Turjman, “Cognitive Routing for Information-Centric Sensor Networks in Smart Cities” *In Proc. of the International Wireless Communications and Mobile Computing Conference (IWCMC)*, Nicosia, Cyprus, 2014, pp. 1124 - 1129.

[8] D. H. Friend, “Cognitive Networks: Foundation to Applications”, Ph.D. Dissertation, Electrical and Comput. Eng., Virginia Polytechnic and State Univ., Blacksburg, VA, March 6, 2009.

[9] D. H. Friend, R. W. Thomas, A.B. MacKenzie, and L.A. DaSilva, “Distributed Learning and reasoning in cognitive networks: methods and design decisions”, Book chapter 9, Edited by Q. Mahmoud, 2012.

[10] L. Reznik and G. Von Pless, “Neural networks for cognitive sensor networks,” in *Proc. IEEE Int. Joint Conf. Neural Network., IJCNN 2008*, Jun. 2008, pp. 1235–1241.

[11] A. Förster, “Machine Learning Techniques Applied to Wireless Ad-Hoc Networks: Guide and Survey,” in *Proc. 3rd Intl. Conf. Intelligent Sensors, Sensor Networks and Information Processing* (ISSNIP), Melbourne, Australia, December 2007.

[12] S. Russell, and P. Norvig, “Artificial Intelligence A Modern Approach”, Third edition, Pearson Education Inc., 2010.

[13] A. Förster, “Machine Learning Techniques Applied to Wireless Ad-Hoc Networks: Guide and Survey,” in *Proc. 3rd Intl. Conf. Intelligent Sensors, Sensor Networks and Information Processing* (ISSNIP), 2007.

[14] R. S. Sutton, and A.G. Barto, “Reinforcement learning: An introduction”, Cambridge, MA: MIT Press, 1998.

[15] R. M. Karp, “On-line algorithms versus off-line algorithms: How much is it worth to know the future?” Technical report, TR-92-044 July 1992. Available online: http://www.icsi.berkeley.edu/pubs/techreports/TR-92-044.pdf

[16] V. Bulitko, G. Lee, “Learning in Real-time search: A unifying framework,” *J. Aritificial Intelligence Research,* no. 25, pp. 119-157, 2006.

[17] R. A. Bianchi, C. H. Ribeiro, and A.H. Costa, “Accelerating autonomous learning by using heuristic selection of actions,” *Journal of Heuristics*, Vol. 14, no. 2, pp. 135-168, 2008.

[18] W. Youssef, and M. Younis, “A cognitive scheme for gateway protection in wireless sensor network,” *Appl. Intell. J.*, vol. 29, no. 3, pp. 216–227, 2008.

[19] P. Boonma, and J. Suzuki, “Exploring self-star properties in cognitive sensor networking,” in *Proc. IEEE/SCS Int. Symp. Performance Evaluation Comput. Telecommun. Syst. (SPECTS)*, Edinburgh, U.K., Jun. 2008, pp. 36–43.

[20] K. Shenai, and S. Mukhopadhyay, “Cognitive sensor networks,” in *Proc. IEEE 26th Int. Conf. Microelectronics (MIEL)*, May 2008, pp. 315–320.

[21] R.A.C. Bianchi, C. H. C. Ribeiro, and A.H.R. Costa, “Heuristic Selection of Actions in Multiagent Reinforcement Learning”, *IJCAI'07*, Hyderabad, India, 2007.

[22] A. Al-Fagih, F. Al-Turjman, W. Alsalih and H. Hassanein, “A priced public sensing framework for heterogeneous IoT architectures,” *IEEE Transactions on Emerging Topics in Computing*, vol. 1, no. 1., pp. 133-147, June 2013.

[23] R. A. C. Bianchi, R. Ros, and R. L. De Mantaras, "Improving reinforcement learning by using case based heuristics." In *Case-Based Reasoning Research and Development*, pp. 75-89. Springer Berlin Heidelberg, 2009.

[24] R. A. C. Bianchi, C.H.C. Ribeiro, and A.H.R. Costa, "Heuristically Accelerated Reinforcement Learning: Theoretical and Experimental Results." In *ECAI*, pp. 169-174. 2012.

[25] G. Gigerenzer, and W. Gaissmaier, "Heuristic decision making." *Annual review of psychology* 62 , pp. 451-482, 2011.

[26] G. T. Singh, F. M. Al-Turjman, “Cognitive-Node Architecture and Deployment for Future Sensor Networks”, Wiley Publications, Wiley, Wireless Communications and Mobile Computing, Feb. 2015. [Submitted]

[27] H. Geffner, “Heuristics, Planning and Cognition”, In R. Dechter, H. Geffner, and J. Halpern, editors, Heuristics, Probability and Causality. A Tribute to Judea Pearl. College Publications, 2010.

Chapter 6

Conclusions and Future Directions

Smart City environments have emerged as one of the most promising applications of the Internet of Things era because of their human centric nature and ability to impact almost every activity in everyday life. Research in the IoT domain has spanned across a lot of areas including the future internet, object identification using RFID and Sensor networks, and various artificial intelligent techniques to improve the network performance and end-user experience. In contributing towards making the network more responsive towards user requirements, in this thesis, we proposed and evaluated the use of a cognitive information centric sensor network for large-scale IoT applications in smart cities.

6.1 Summary

We started the work in this thesis with hardware based experiments to explore the use of intelligent decision making in sensor networks; the results of which were discussed in Chapter 2. From the experimental results of the case study, we arrived at the conclusion that the ZigBee stack is capable of supporting intelligent behavior. We also suggested that it would be useful to develop a generic knowledge-based cognitive framework that can be applied to any sensor application platform, wherein the network decisions are based on learning and reasoning, and on information shared among the network nodes from the observations made. From there on, we focused on developing the cognitive techniques for sensor networks in the IoT application space; with applications in smart outdoor environments, such as smart cities and smart environment monitoring.

In Chapter 3 we proposed a Cognitive Information Centric Sensor Network (CICSN) architecture for introducing cognition in information centric sensor networks using cognitive nodes deployed at specific locations in the network. These cognitive nodes implemented the features of the OADA feedback loop and implemented knowledge representation, reasoning and learning as the elements of cognition. A cost comparison between relay and cognitive nodes showed that cognitive nodes were more expensive. Hence a 2-dimensional grid based deployment plan was proposed for the CICSN to control the number of relay nodes (RNs) and cognitive nodes (CNs) used, and to keep the number of CNs lower than the number of RNs in the network. The CNs were the decision makers in the network and RNs were used for forwarding the data. The sensor nodes were assumed to be uniformly, randomly deployed throughout the target area. The separation distance between the RNs and CNs was planned based on the probability of successful data reception at each node for a given transmit power and communication range. A Matlab simulation model for the IEEE 802.15.4 PHY and MAC was also used to evaluate the impact of varying load and number of active neighboring nodes on the network's performance in terms of latency, node reliability and instantaneous throughput.

In Chapter 4, we presented the details of the COGNICENSE framework, and elaborated on the reasoning and knowledge representation techniques used at the cognitive nodes. Knowledge was represented using attribute-value pairs, and Analytic Hierarchy Process (AHP) was used as the reasoning element of cognition to help with Quality of Information (QoI) aware decision making during data delivery. We showed how the CNs put these elements of cognition to use, by identifying data delivery paths to the sink from any source LCN in the CICSN. Latency, reliability and throughput were the attributes used to identify the QoI associated with data that was delivered to the end-user. User-set priorities assigned to these QoI attributes influenced the classification of traffic flows, and the choice of the data delivery paths.

From extensive simulations, we found that by restricting the number of nodes scheduled for simultaneous transmission to about 10, and the per node offered load to about 500bit/s, QoI attribute values could be maintained at acceptable values along each hop of the data delivery path. We also compared the AHP based QoI-aware data delivery strategy (AHPDD) with two other routing strategies used in traditional data-centric sensor networks, namely, highest remaining battery based data delivery (HRBDD), and multipath data delivery (MDD). From simulations, we found that the network was able to deliver data successfully to the GCN more often with the AHPDD technique, when compared with the other two techniques. In addition, the network performed well in responding to varying traffic types and changing network topology, thus establishing the suitability of the COGNICENSE framework for information centric sensor networks in smart environment applications of the IoT.

In chapter 5, we elaborated on the learning techniques that could be used with AHP-based reasoning to improve the cognitive decision making capabilities at the local cognitive nodes of the ICSN, operating in an IoT application environment. The challenge was to adapt the learning to (1) respond to the changing user requirements/requests from multiple users with good QoI on the data delivered by the network, and (2) to consider the energy consumption in the network while catering to the QoI requirements, and maintain a high average rate of successful data delivery to the sink while doing so. To address these challenges, we proposed two learning techniques, namely, Learning Data Delivery A* (LDDA*) and Cumulative-Heuristic Accelerated Learning (CHAL) that used heuristics to improve the success rate of data delivered to the sink in the cognitive ICSN. LDDA* was able to deliver data with good QoI at the sink, while CHAL was the more energy considerate technique, and both these learning strategies had comparable performance with respect to the success rate of data delivered to the sink.

Thus a new architecture for introducing cognition in wireless sensor networks, making use of an information centric approach was proposed and evaluated in this thesis. A considerable improvement was seen in the success rate of data delivered by the network to its sink, and this architecture can be used for different IoT applications, and supports multiple users simultaneously.

6.2 Future Directions

With the research done in this thesis, WSNs will be able to provide better infrastructure support for the Smarter Planet / Smart City initiatives across the globe. Not only this, several future research directions and open issues can also be derived from the work done in this thesis thus far. We outline some of these directions in the following section.

1. Exploiting the Caching capabilities of the cognitive nodes:

   One of the aspects of the information centric capabilities in the COGNICENSE framework, is its ability to cache information in cognitive nodes and use it collaboratively for information sharing across the network. While we acknowledged the advantage such caching would offer in a network with mobility-enabled nodes, we did not delve deeply into this topic. Thus, exploring the role of caching in information access and data delivery, and the study of cache replacement techniques that suit the cognitive nodes in the ICSN environment will be an interesting direction to explore from this work. In Appendix II we briefly describe the work we have initiated in this direction, called Value of Sensed Information (VoSI) based cache replacement.

2. Mobility enabled sensor and cognitive nodes:

   Effect of node mobility on the QoI delivered by the system, and its impact on the networks adaptability and longevity can be evaluated. While the CICSN architecture proposed here helps to overcome the limitations of the cross-layer design, we mostly dealt with a static environment. The impact of sensor and cognitive node mobility, and the network dynamics that it would introduce, along with the impact of node deaths, have not been explored in depth. Dynamicity is an important consideration in WSN protocols, and including it in optimization problem models is another challenge that can be addressed.

3. Enhancing the sensor network gateway for better knowledge representation:

   In terms of enhancing the GCNs ability to support diverse application platforms, this work could be enhanced to develop application or domain specific ontologies for better knowledge representation at a higher level. The creation of such domain ontologies contributes towards the development of an enterprise architecture framework that can be applied to different application domains using the same underlying cognitive sensor network platform.

4. Integration of CICSNs with next generation wireless network infrastructure:

   More functions could be incorporated at the GCN to integrate it with the Next Generation Network (NGN) constituted of cognitive radio enable nodes, working in cognitive network setup. The expansion of the GCN functions to a cognitive gateway node can be considered. The cognitive gateway node would then be able to take requests directly from different wireless users such as cell phone users, wireless access points and base stations, thus making the CICSN platform more accessible to end-users.

5. Cognition in the future internet – ICNs:

   The idea of cognition (cognitive elements) can be used in the intermediate routers of the future internet to provide on-demand content to users quickly. Data need not be requested from specific hosts, nor has to travel end-to-end in the network. Instead, it can be cached at the network's edge. Cognitive routers could be used to understand the user request patterns and manage the cache content intelligently.

6. Connectivity and Communication between cognitive nodes:

   a. We can develop an optimal learning policy that will identify from a set of CNs with heterogeneous capabilities, the minimum cover set required to maintain network connectivity while minimizing the number of CNs used. The minimum cover set approach with optimal learning can be used to find the minimum number of sensor nodes that should be connect to the CNs, to minimize energy consumption while maintaining accuracy of information.

   b. CNs can communicate amongst themselves and share their data bases to act altruistically towards a common network goal. This information sharing can improve what the network learns from feedback and observation, and can make better decisions for network resource management and data delivery.

7. Flexibility of the COGNICENSE framework in choosing QoI attributes:

   The advantage of using AHP analysis was that the technique could be applied to any measurable QoI attribute(s) in the WSN, thus making the COGNICENSE framework adaptable to any IoT application. The priorities assigned to the QoI attributes could also be varied to influence the classification of traffic flows, and the choice of the data delivery paths.

8. Security and Privacy issues in the IoT framework:

Data security and privacy are key areas to be addressed to encourage the widespread adoption of IoT applications and technologies, especially in large-scale deployments such as Smart Cities. In early IoT deployments, security solutions have not been incorporated in a planned way. The enabling technologies of RFID and sensor networks were simply integrated vertically with the internet infrastructure, without considering the security and privacy issues at the time of data gathering and delivery. These security considerations span not only along the data management, application and service levels, but also at the communication and networking level, and are important areas to consider for future research.

Appendix I

AHP Analysis

The Analytic Hierarchy Process (AHP) used for implementing the reasoning element of cognition at the Cognitive Nodes (CNs) was briefly described in Section 4.5.4 of Chapter 4. Here we provide additional material to explain the details of this technique, as applied to the context of identifying the best next-hop Relay Nodes (RNs) along the data delivery path to the network's sink. At each hop along the identified data delivery path, the nodes chosen are expected to satisfy the user's requirements in terms of the QoI attribute priorities. Latency, reliability, and throughput are the considered QoI attributes, and their values are measured at the PHY and MAC layers of the participating network nodes. Their definitions are as explained in section 4.4.1. Some parts of the presented material here may overlap with the material in section 4.5.4, as it helps to maintain the context.

The first step in AHP analysis is to use the fundamental scale for pairwise comparison shown in Table I to set application-defined priorities for the QoI attributes. Intermediate values such as 2, 4, 6, and 8, or equal values may also be assigned to QoI attributes to represent that they lie very close to each other in terms of importance. Relative priorities of the QoI attributes are then established using the pair-wise comparison scale, and tabulated as shown in Table II.

Table I: Fundamental scale for pairwise comparisons.

Intensity	1	3	5	7	9
Relative importance	Equal	Moderate	Strong	Very Strong	Extreme

Table II: Pair-wise comparison of QoI attributes.

Latency	4	Reliability	1
Latency	6	Throughput	1
Reliability	3	Throughput	1

The AHP analysis is then used to establish the relative priorities of QoI attributes for the values in Table II using the following steps:

- Represent the values of Table IV in matrix form

$$\boldsymbol{A} = \begin{matrix} 1 & 4 & 6 \\ 1/4 & 1 & 3 \\ 1/6 & 1/3 & 1 \end{matrix} \tag{1}$$

- Compute the eigen vector **v** of the matrix ***A***
- Isolate the absolute, real values of the eigen vector **v** as **q**

$$\mathbf{q} = |Re\ (\mathbf{v})| \tag{2}$$

- Compute the normalized values of **q** and tabulate them as the values of the Relative Priorities of the attributes as shown in the last column of Table III

Table III: Results of AHP analysis for QoI the attributes.

	Latency	**Reliability**	**Throughput**	**Relative Priorities**
Latency	1	4	6	0.691
Reliability	¼	1	3	0.2176
Throughput	1/6	1/3	1	0.0914

Table IV is then populated with actual values of Latency, Reliability and Instantaneous throughput observed at each of the next-hop RNs connected with a source LCN. The RNs are then pair-wise compared with respect to each of the QoI attributes and the priority values '**P**' are obtained by calculating the eigen vector of the matrix and normalizing it. Thus, priorities for each QoI attribute that the RNs offer is obtained. Sample values from actual simulations (based on the simulation setup used in Chapter 4) have been used in Table IV. The simulations assume that a minimum of 2 and maximum of 8 nodes are assumed to be transmitting simultaneously to a given LCN at a time. The application load ranges from 0-127Bytes, the frame arrival rate λ is about 1000bits/s per node, and the data rate is19.2Kbps.

Table IV: Table for the LCN alternatives v/s actual values of each QoI Atribute.

	Latency	Reliability	Throughput
RNi (121B, 8nodes)	0.083	0.8	0.135
RNj (121B, 5 nodes)	0.066	0.98	0.148
RNk (100B, 3nodes)	0.054	1	0.155

Values in Table IV can also be obtained for a different system size, application payload, frame arrival rate and number of stations trying to simultaneously transmit information. The values for throughput have been uniformly scaled down for this analysis. Table V is derived by using pair-wise comparisons of the alternatives against each criterion from Table IV as part of the AHP analysis.

Table V: AHP for the LCN alternatives v/s each QoI Attribute.

	Latency				**Reliability**				**Throughput**			
	RNi	**RNj**	**RNk**	**P**	**RNi**	**RNj**	**RNk**	**P**	**RNi**	**RNj**	**RNk**	**P**
RNi	1	1.257	1.537	**0.408**	1	0.816	0.8	**0.287**	1	0.912	0.87	**0.308**
RNj	0.795	1	1.222	**0.325**	1.225	1	0.98	**0.352**	1.09	1	0.954	**0.337**
RNk	0.65	0.818	1	**0.266**	1.25	1.02	1	**0.359**	1.148	1.047	1	**0.354**

The values obtained in Table V are further processed to obtain the overall attribute priorities for all possible next-hop RNs, as shown in Table VII. To obtain these values, we run the AHP analysis to calculate the actual value of priorities v/s goal as shown in Table VI. The results in Table VII suggests that RNi chosen as next hop performs best with respect to all three QoI attributes, thus providing the next hop RN for delivering the data. But if we wanted to consider only one of the QoI attributes, then a column-wise comparison for each attribute will provide the desired information.

Table VI: AHP to evaluate actual value of QoI Attributes v/s Goal.

Actual value: QoI v/s goal	Latency	Reliability	Throughput	Actual value of priorities
Latency	1	9.638	1.626	0.6174
Reliability	0.067	1	0.151	0.0535
Throughput	0.348	6.451	1	0.3291

Table VII: AHP to evaluate the overall priorities for all possible next-hop LCNs.

Best candidate for next hop LCNx	Priority with respect to			
	Latency	Reliability	Throughput	Goal
LCNi	**0.252**	0.015	0.101	**0.375**
LCNj	0.2	0.018	0.11	0.329
LCNk	0.164	**0.019**	**0.116**	0.296

However, if we want to consume as less node resources as possible (compute and memory in this case), we can just make use of the information already available from Table V, which points towards the same information as obtained from Table VII. These computations can be initially carried out for each next-hop node decision in the data-delivery path. This technique helps to build the learning database at each LCN about its next-hop neighbors, and the priorities each of them offers with respect to the QoI attributes. This information can be stored and used for planning future rounds of data-delivery for application traffic that may need to choose a different next hop for the same source LCN, based on the expected values of attribute priorities at the GCN. Thus we can see that this AHP process helps in adaptive multi-criteria decision making during data-delivery, in considering the desired attribute priorities for each application-traffic type.

In Chapter 4, Algorithm 1 described the use of the AHP analysis in the LCNs of the CICSN. This AHP analysis is run on LCNs as per the directive of the GCN, as long as there are over 50% of the LCNs remaining alive in the network. Although the LCNs do not have global knowledge of the network, the GCN is aware of the network status, and is capable of instructing the LCNs about when they should run this cognitive reasoning mechanism to identify data delivery paths. Once the number of LCNs falls to less than 50% of the originally deployed number, or the network is found to be disconnected by the GCN, then the GCN instructs the LCNs in its next broadcast that AHP analysis should not be run at the LCNs. This way, LCNs and GCN interact to co-ordinate the delivery of user-requested data by implementing AHP analysis as the reasoning mechanism.

Appendix II

A Value of Sensed Information Based Cache Replacement Strategy

1. Introduction

Cognitive nodes are used as data caches to save energy for the network and make data readily available to the user. We propose the use of a Value of Sensed Information (VoSI) based cache replacement strategy to retain the most valuable information in the cache for longer. Three parameters, namely, age of data based on periodic request, popularity of on-demand requests, and the duration for which the sensor node is required to operate in active mode to capture the sensed readings, are considered together to assign a value to the data. A least valuable information first approach is used to identify the least valuable sensed data, which can be replaced first in the cache of cognitive nodes to maintain availability of useful data for longer. This strategy for replacement of cached data helps to improve data availability in the network as shown by the initial simulation results.

2. Value of Sensed Information based Cache Replacement Policy

For cache replacement in CICSNs, we need to ensure that we choose data appropriately for storage based on the following criteria: Firstly, data that takes longer to sense should be stored for longer to conserve the sensor node's energy. Secondly, data storage must be a function of the periodicity of the requests based on the traffic type. This will help to store data till fresher data is available, and in servicing requests for different traffic types in a timely manner. Lastly, value of the data based on its age, i.e. if temperature in a region has changed considerably from the last time it was sensed, then the cached information is stale and does not provide correct information.

Hence the freshness of data is also an important criteria when servicing requests for data on demand. Since these criteria are known and fixed, the cache replacement plan can be programmed into the LCN. We propose a Value of Sensed Information (VoSI) based cache replacement strategy for the LCNs in an ICSN, which follows closely on the lines of the work by Al-Turjman *et al*[*1], with changes and adaptations for the CICSN model.

2.1 Delay model

Different sensors have different durations for which they need to be exposed to the environment, so that they can capture the sensed readings accurately. This affects the duration of the on-time of the sensor node, which in turn affects the lifetime of the sensor node. In order to prolong the lifetime of the sensor node, it is useful to store the sensed data for longer when the delay involved in acquiring the reading is more. This is called the sensing delay. In addition, if data has to be propagated from sensor nodes to LCNs every time data is requested, it would add to the propagation delay of the data, especially if the sensor nodes are located far away from the Sink. Thus the delay components we consider are the sensing delay δ and the propagation delay τ. We also limit the number of hops (*n*) within which the data has to be delivered to the sink to 6, so as to avoid unnecessary wastage of energy by involving multiple nodes in the data transmission. Accordingly, the constraints on τ and δ are given by Equations 1 and 2 respectively.

$$\tau \propto n, \qquad \text{for } n<6 \tag{1}$$

$$\delta \propto \max(d_1, d_2, d_3, ..d_k) \tag{2}$$

**1 F. M. Al-Turjman, A. E. Al-Fagih, H. Hassanein, " A Value based cache replacement approach for information centric networks", IEEE WLN ,13th Annual Workshop on Wireless Local Networks, 2013.*

k is the total number of sensors available on board the sensor node, and d_i represents the fixed sensing delay value of the sensor type *i*.

Thus the sensing delay is a function of the maximum delay from among the sensor types that have been activated to provide fresh data. Putting these two delays together, the total delay (Δ) involved in delivering freshly sensed data to the sink is a combination of the sensing and propagation delay, given by Equation 3.

$$\Delta = \tau + \delta \tag{3}$$

2.2 Age model

Our age model makes use of the following two conditions to decide what content should be dropped from the cache. The first is the based on the periodicity of the periodic request (Type1 traffic), and second, when the node's cache is full. We make use of the periodicity of the periodic request, because freshly sensed data has to be provided at the start of each periodic request cycle. Thus, when the cache is full at the end of one periodic request cycle, old data can be discarded from the cache. Thus the age of a sensed attribute-value pair is represented by its time-to-live (TTL) which is based on the periodicity of the request of each application type. This value is provided to the LCN by the GCN/Sink. Since we are not considering the use of historic data, our model implies that cached contents may be refreshed after every periodic time interval, as long as the data is being transmitted to the sink at the end of each cycle.

$$TTL_{SI} \propto T_{periodic} \tag{4}$$

Equation (4) represents that the TTL of the Sensed Information (SI) represented as attribute-value pair, is directly dependent on the periodicity of a request in Type1 traffic flow. In case the application requires that the periodic data is stored for a prolonged duration of time, for example

24hours, before making a single transmission to the sink, then the cache retention period becomes a function of the transmission cycle's periodicity.

2.3 Popularity of On-demand Requests

Traffic flow generated in response to on-demand requests have been classified as Type 2 traffic. More number of users may be interested in a particular type of sensed data, or a specific sensed data may be requested more number of times by one or more users. Such sensor data is said to be popular, and can be retained for longer in the LCN's cache. Thus the popularity of the sensed attribute-value pair is given by equation (5)

$$Popularity_{SI} \propto \mathrm{Re}\, q_{SI} / \mathrm{Re}\, q_{total} \qquad (5)$$

Where $\mathrm{Re}\, q_{SI}$ is the total number of requests for an attribute-value pair received at an LCN, and $\mathrm{Re}\, q_{total}$ is the total number of requests received by that LCN, within a particular operational cycle.

In addition, when sensor nodes start to die out in the network, LCNs should store the data for longer to maintain their availability. When the primary LCN storing such data itself starts to die out, storing the data in neighboring LCNs provides extra storage guarantees and ensure availability of data in the network for longer. This storage requirement based on non-availability of alive sensor nodes, is managed by the planning algorithm for data delivery based on the traffic flow in the network, and remaining energy at LCNs. Based on these models, we formulate the VoSI based cache replacement function in Equation 6 as follows:

$$VoI_{SI} = \alpha * \Delta + \beta * TTL_{SI} + \gamma * Popularity_{SI} \qquad (6)$$

Here α, β, and γ are the tuning parameters that are specified based on the traffic type and the user requests.

3. Performance Evaluation

In this section, we provide initial performance evaluation results for the VoSI based cache replacement technique, which we have compared with first-in-first-out (FIFO) and least-recently-used (LRU) techniques using ADEVS, a discrete event simulator. We make use of the Cache Hit Ratio to compare the performance of the different cache replacement strategies. Cache hit ratio is defined as the ratio of the number of times requested data was found in the cache divided by the total number of times data was requested from the cache. Simulation results are compared for VoSI, LRU and FIFO replacement techniques.

3.1 Simulation results

The VoSI based replacement technique provide a slightly better hit ratio for different cache sizes, when compared to the other two techniques. These simulations were run at a cache sizes ranging from 10 to 100, and the simulations end after serving 1000 packet requests. There are 100 different requests from which the packet requests are randomly generated.

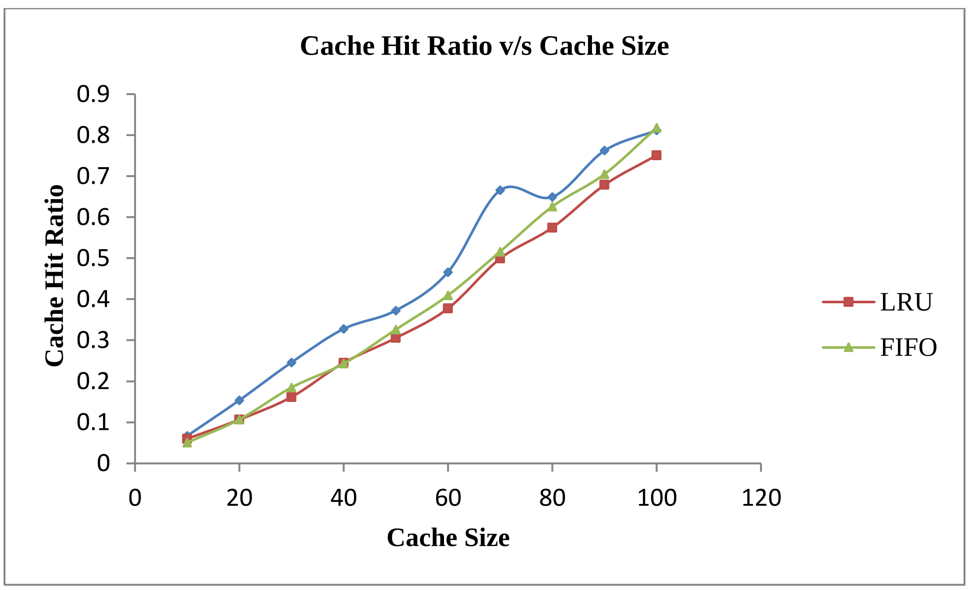


Figure 1: Plot of cache hit ratio versus cache size for VoSI, FIFO and LRU.

The advantage offered by the VoSI based cache replacement technique is that it can replace data based on the user-requirements and VoI of data. Other cache replacement techniques are only worried about the match of the data request packet numbers, irrespective of the age of data or its popularity or the delay involved in sensing and transmitting it to the Sink. But with VoSI based technique, factors like age of data, popularity and VoSI are also considered during replacement. Older data is replaced by fresher one. Unlike other techniques that look only for a number match irrespective of its age. Based on this, we suggest the use of 2-levels of caches, one at the CN, and one at RN. At CN we could use Value of Information (VoSI) based cache replacement strategy, and at RN, any of FIFO or LRU to make the computation less complex. Size of 1st level cache and 2nd level cache can be 100, as seen from Figure 1. The performance gains do not increase much when the cache size is increased beyond 100, when there are only100 different types of requests considered. Since the decision making is only at CNs, implementing the VoSI based cache replacement strategy at CNs can help the network save more resources if a cache hit is found at the first level of cache.

4. Conclusions

The VoSI based technique is suitable for use in CICSNs which use named data association for the sensed data, and are expected to support node mobility in the future. We can expect that the users are more satisfied with the response received from the LCNs, as they retain information in their cache based on both data popularity and various parameters that affect gathering sensed information and the energy involved in doing so, as the network scales up to larger sizes. Further, the VoSI cache replacement strategy will help in graceful degradation of the network, as cached data can be provided from LCNs even after sensor node deaths. However, this work requires further extensive simulations to evaluate the impact of varying network loads and inter-LCN communication on the effectiveness of the cache replacement strategy.