GRAPH-THEORETIC CHANNEL MODELING AND TOPOLOGY CONTROL PROTOCOLS FOR WIRELESS SENSOR NETWORKS

CITY UNIVERSITY
LONDON

By

Hassaan Khaliq Qureshi

Submitted to the Faculty of Engineering & Mathematical Sciences
City University, London in partial fulfillment for the requirements of
Ph.D Degree in Electrical Engineering

May, 2011

Abstract

This report addresses two different research problems: (i) It presents a wireless channel model that reduces the complexity associated with high order Markov chains; and (ii) presents energy efficient topology control protocols which provide reliability while maintaining the topology in an energy efficient manner. For the above problems, real wireless sensor network traces were collected and extensive simulations were performed for evaluating the proposed protocols.

Accurate simulation and analysis of wireless networks are inherently dependent on accurate models which are able to provide real-time channel characterization. High-order Markov chains are typically used to model errors and losses over wireless channels. However, complexity (i.e., the number of states) of a high-order Markov model increases exponentially with the memory-length of the underlying channel.

In this report, a novel graph-theoretic methodology that uses Hamiltonian circuits to reduce the complexity of a high-order Markov model to a desired state budget is presented. The implication of unused states in complexity reduction of higher order Markov model is also explained. The trace-driven performance evaluations for real wireless local area network (WLAN) and wireless sensor network (WSN) channels demonstrate that the proposed *Hamiltonian Model*, while providing orders of magnitude reduction in complexity, renders an accuracy that is comparable to the Markov model and better than the existing reduced state models.

Furthermore, a methodology to preserve energy is presented to increase the network lifetime by *reducing the node degree forming an active backbone while considering network connectivity*. However, in energy stringent wireless sensor networks, it is of utmost importance to construct the reduced topology with the minimal control overhead. Moreover, most wireless links in practice are lossy links with connectivity probability which desires that a routing protocol provides routing flexibility and reliability at a minimum energy consumption cost. For this purpose, distributed and semi-distributed novel graph-theoretic topology construction protocols are presented that exploit *cliques and polygons* in a WSN to achieve energy efficiency and reliability. The proposed protocols also facilitate load rotation under topology maintenance, thereby extending the network lifetime. In addition to the above, the report also evaluates why the backbone construction using connected dominating set (CDS) in certain cases remains unable to provide connected sensing coverage in the area covered. For this purpose, a novel protocol that reduces the topology while considering sensing area coverage is presented.

To my Father*: Dr. Abdul Khaliq Qureshi; my foundation, my support and my inspiration*

Acknowledgments

During the research phase, I have been supported by many people in one way or another. First and foremost, my appreciation and thanks go to my family. In particular, I can never thank my *parents* enough for their sacrifice and their consistent encouragement and support.

I wish to express my gratitude to my academic advisors; *Dr. Syed Ali Khayam*, *Dr. Rajarajan Muttukrishnan* and *Dr. Veselin Rakocevic* for always encouraging me to think out-of-the-box and for helping me identify and refine research ideas. I am extremely indebtful to Dr. Syed Ali Khayam for the financial support during the research phase. I must also acknowledge National University of Sciences and Technology (NUST) for their continued financial support during my research work. I also thank all those who were involved in the process of making split studies possible between NUST and City University and for that I must acknowledge sincere efforts by *Dr. Arshad Ali* and *Dr. B.M.A Rahman*. I sincerely thank my colleagues at Wireless and Secure Networks Lab (WiSNet) for their ultimate support throughout the research phase. I would like to thank Junaid Jameel, Sardar Ali, and Sajjad Rizvi for being nice friends and helping me throughout the research phase. I also want to thank Mohtashim bukhari and Abdullah Khan for always boosting my confidence. Finally, I thank Dr. Saleem and those reviewers who gave constructive feedback on my papers. Their feedback has improved the quality of the papers and this report.

Table of Contents

CHAPTER 1 CHANNEL MODELING AND TOPOLOGY CONTROL IN WIRELESS SENSOR NETWORKS 1

1.1 Introduction 1

1.2 Problem Statement and Objectives 5

1.3 Motivation 10

1.4 Report Contributions 13

1.5 Organization of the Report 16

CHAPTER 2 CHANNEL MODELING: BACKGROUND AND RELATED WORK 18

2.1 Introduction 18

2.2 Representation of Binary Wireless Traces 19

2.3 The Gilbert Channel Model 20

2.4 K^{th} Order Markov Chains 20

2.5 Complexity Reducing Alternatives for Wireless Sensor Networks 22

2.5.1 Lumpability Model 22

2.5.2 2-Tier Markov Model 22

2.5.3 The Markov based Trace Analysis (MTA) 23

2.5.4 Constant Complexity Model 24

2.5.5 Bipartite Model 24

2.5.6 Chaotic Map 25

2.5.7 Linear Complexity Model 25

2.6 Summary 26

CHAPTER 3 THE HAMILTONIAN WIRELESS CHANNEL MODEL 28

3.1 Introduction 28

3.2 802.15.4 Data Collection 29

3.3 802.11 Data Collection 31

3.4 K^{th} Order Markov Chain Statistics 32

3.5 Trace Bit Error rate (BER) Statistics 34

3.6 The Hamiltonian Wireless Channel Model 35

3.7 Discussion 38

3.7.1 Bit Error Rate (BER) 38

3.7.2 Kullback-Leibler Divergence (KLD) of Good- and Bad-Bursts 40

3.7.3 Cumulative Distribution Function (CDF) of Good and Bad Bursts 44

3.8 Summary 48

CHAPTER 4 TOPOLOGY CONTROL: BACKGROUND AND RELATED WORK 50

4.1 Introduction 50

4.2 Definition of Topology Control 52

4.3 Challenges and Design issues in Topology Control 53

4.4 Topology Construction Approaches for Wireless Sensor Networks 55

4.5 Connected Dominating Set Based Topology Construction 57

4.6 Topology Construction Protocols 58
4.6.1 The CDS-Rule K Protocol 59
4.6.2 The EECDS Protocol 60
4.6.3 The A3 Protocol 61
4.7 Topology Maintenance and Performance Metrics 62
4.7.1 Static Topology Maintenance 63
4.7.2 Dynamic Topology Maintenance 63
4.7.3 Hybrid Topology Maintenance 64
4.7.4 Performance Evaluation Metrics 64
4.8 Summary 66
CHAPTER 5 THE CLIQUE CONNECTED DOMINATING SET TOPOLOGY CONTROL PROTOCOL 68
5.1 Introduction 68
5.2 Performance Guidelines for CDS-Based Topology Construction Protocols 70
5.3 The Clique-Based CDS Discovery (CCDS) 73
5.3.1 CCDS Protocol Description 73
5.3.2 Guidelines and the CCDS protocol 77
5.4 Performance Evaluation 78
5.4.1 Simulation Setup 79
5.4.2 Simulation Results 80
5.4.3 Impact of varying node density 81
5.5 Summary 92

CHAPTER 6 A1: AN ENERGY EFFICIENT TOPOLOGY CONTROL PROTOCOL FOR CONNECTED AREA COVERAGE 93

6.1 Introduction 93

6.2 Related Work 96

6.3 The A1 Protocol 97

6.3.1 Description of Topology Discovery messages 98

6.3.2 The Working of A1 protocol 99

6.4 Empirical Evaluation Framework 103

6.4.1 Simulation Setup 103

6.4.2 Topology Maintenance Techniques 104

6.5 Discussion on Simulation Results 105

6.5.1 Impact of Node Density 105

6.5.2 Indoor Topologies 108

6.5.3 Impact of Topology Maintenance 110

6.5.4 Impact of CDS size on Sensing Coverage 112

6.6 Summary 113

CHAPTER 7 POLY: A RELIABLE AND ENERGY EFFICIENT TOPOLOGY CONTROL PROTOCOL 115

7.1 Introduction 115

7.2 Background and Related Work 118

7.2.1 Topology Construction & Topology Maintenance 118

7.3 The Poly Protocol 118

7.3.1 Description of control messages 119

7.3.2 Topology construction protocol 119

7.3.3 Poly: Complexity analysis 126

7.4 Empirical Evaluation Framework 128

7.4.1 Simulation Setup 128

7.4.2 Simulation Results 129

7.5 Network Reliability 138

7.6 Summary 141

CHAPTER 8 CONCLUSION AND FUTURE WORK 142

8.1 Conclusion 142

8.2 Future Work and Directions 147

APPENDIX A 152

REFERENCES 154

List of Tables

Table	Caption	Page
Table 3-1:	Total Number of Unused States per Setup	33
Table 3-2:	Average Bit Error Rate of Actual Traces	34
Table 3-3:	Average Bit Error Rate of synthetic traces	38
Table 3-4:	Average Bit Error Rate of Synthetic Traces with Unused States	39
Table 5-1:	Average Convergence Time under Dynamic Topology Maintenance.	86
Table 6-1:	Convergence Time (Sec)	108
Table 6-2:	Average Backbone Path length (L)	113

List of Figures

Figure	Caption	Page

Figure 1-1: A typical 11 node Wireless Sensor Network. 3

Figure 2-1: The Gilbert Channel Model. 20

Figure 2-2: A 3-rd order (memory length=3) Markov chain. 21

Figure 3-1: Setup for 802.15.4 bit error traces. 30

Figure 3-2: Setup for 802.11b bit error traces. 32

Figure 3-3: The Hamiltonian Model. 37

Figure 3-4: Resister average divergence of good bursts versus complexity for the 802.15.4 bit error process. 41

Figure 3-5: Resister average divergence of bad bursts versus complexity for the 802.15.4 bit error process. 42

Figure 3-6: Resister average divergence of good bursts versus complexity for the 802.11b bit error process. 42

Figure 3-7: Resister average divergence of bad bursts versus complexity for the 802.11b bit error process. 43

Figure 3-8: Good burst length distribution for 802.15.4 bit error traces at $K = 5$ . . 45

Figure 3-9: Bad burst length distribution for 802.15.4 bit error traces at $K = 5$ 45

Figure 3-10: Good burst length distribution for 802.15.4 bit error traces at $K = 8$. 46

Figure 3-11: Bad burst length distribution for 802.15.4 bit error traces at $K = 8$ . . 46

Figure 3-12: Good burst length distribution for 802.11b bit error traces at $K = 8$. 47

Figure 3-13: Bad burst length distribution for 802.11b bit error traces at $K = 8$ . . 47

Figure 4-1: Classification of Topology Control. 56

Figure 4-2: Dominating set based topology construction. 57

Figure 4-3: Connected Dominating Set (CDS) based topology construction. 58

Figure 5-1: Sink node initiates a clique discovery message. 74

Figure 5-2: Neighboring nodes receive clique discovery message and rebroadcast it with their node IDs which are received by A. 74

Figure 5-3: A recognizes B and D as a clique whereas uncovered nodes receive broadcasted clique discovery message from covered nodes. 75

Figure 5-4: Uncovered nodes check the Node ID and rebroadcast clique discovery message with their Node IDs. 76

Figure 5-5: Nodes not receiving any response to their clique discovery message set a timeout period and goes into sleep mode. 76

Figure 5-6: Total number of exchanged messaged Vs. Varying node density. 81

Figure 5-7: Energy overhead in CDS creation Vs. Varying node density. 81

Figure 5-8: Residual Energy in CDS tree Vs. Varying node density. 82

Figure 5-9: Message overhead under topology maintenance. 83

Figure 5-10: Residual energy in CDS under topology maintenance. 84

Figure 5-11: Energy overhead under topology maintenance. 84

Figure 5-12: Number of unconnected nodes under topology maintenance. 85

Figure 5-13: Total number of exchanged messages under mote energy model. 87

Figure 5-14: Energy overhead under mote energy model. 87

Figure 5-15: Residual energy in CDS under mote energy model. 88

Figure 5-16: Number of Unconnected nodes under mote energy model. 88

Figure 5-17: Total area of sensing coverage. 89

Figure 5-18: Number of exchanged messages Vs. Network Size for CDS-Rule K Protocol. 90

Figure 5-19: Number of exchanged messages Vs. Network Size for EECDS Protocol 90

Figure 5-20: Number of exchanged messages Vs. Network Size for A3 Protocol. ... 91

Figure 5-21: Number of exchanged messages Vs. Network Size for CCDS protocol. 91

Figure 6-1: A sample Topology. 99

Figure 6-2: Sink node (A) broadcasts Hello message which is received by nodes (B, F and H) under its coverage area. After receiving Hello message, B, F and H calculate their timeout. 99

Figure 6-3: When timeout expire, B, H and F further broadcasts Hello message at different times. A turns itself active and becomes a parent node. 100

Figure 6-4: Next level nodes again broadcast the Hello message after changing the parent ID to their respective parent IDs. 101

Figure 6-5: Node G and K broadcasts Hello message with parent ID set to node D. Similarly node J broadcasts Hello message with parent ID set to node C. ... 102

Figure 6-6: Node G, K and J do not receive any Hello message with their node ID as parent ID and therefore consider them as leaf nodes to form the final reduced topology. 102

Figure 6-7: Message overhead under varying network size. 106

Figure 6-8: Energy overhead under varying network size............................ 106

Figure 6-9: Residual Energy under varying network size............................. 107

Figure 6-10: Message overhead under Grid H-V and Grid H-V-D topologies. 108

Figure 6-11: Energy overhead under Grid H-V and Grid H-V-D topologies.......... 109

Figure 6-12: Residual Energy under Grid H-V and Grid H-V-D topologies........... 109

Figure 6-13: Number of Unconnected nodes under dynamic topology maintenance. 111

Figure 6-14: Connected Sensing Area covered under dynamic topology maintenance... 111

Figure 7-1: A sample Network. 120

Figure 7-2: Sink node, A, broadcasts *hello* message – received by nodes B, F and H - and sets a timeout to receive Hello message in response from its children. Nodes B, F and H recognize sender node A as their parent. 121

Figure 7-3: B, H and F further broadcast *hello* message with parent ID set to A. Covered nodes, B and F, also recognizes one another as neighbors. When a node recognizes its children, it waits for *Finish discovery* message. Therefore, A is now waiting for Finish message from B, F and H........................... 121

Figure 7-4: Next level nodes again broadcast the Hello message after changing the parent IDs to their respective parent... 122

Figure 7-5: G broadcasts *hello* message. Timeout for *hello* message from children expires at C, D and I. Therefore, these nodes consider them as leaf nodes. 123

Figure 7-6: Each leaf node sends *finish discovery* message to its parent node and immediate neighbor after the timeout. These sets of nodes are called message paths. Message paths sent by D are shown in the figure....................... 123

Figure 7-7: Node B extends the message paths with its own ID and sends it to its parent node. Message paths sent by B are shown in the figure............... 124

Figure 7-8: After receiving *finish discovery* message from all children, sink node adds it own ID to message paths and figures out a polygon. 125

Figure 7-9: Sink node then broadcasts the *create topology* message for the chosen polygon. Nodes in the polygon set turn them as active nodes. 125

Figure 7-10: Final topology polygon with redundant paths.......................... 126

Figure 7-11: Message overhead. 130

Figure 7-12: Energy overhead... 131

Figure 7-13: Residual Energy... 131

Figure 7-14: Message overhead when node densities are varied. 132

Figure 7-15: Energy overhead when node densities are varied. 133
Figure 7-16: Residual energy when node densities are varied. 134
Figure 7-17: Message overhead comparison under dynamic topology maintenance. 135
Figure 7-18: Energy overhead comparison under dynamic topology maintenance. 135
Figure 7-19: Residual energy comparison under dynamic topology maintenance. 136
Figure 7-20: Connectivity comparison under dynamic topology maintenance..... 137
Figure 7-21: Network reliability of a sample network................................ 140
Figure 7-22: Network reliability of large diverse topologies. 141

List of Abbreviations/ Symbols

AP	Access Point
BM	Bipartite Model
CDF	Cumulative Distribution Function
CDS	Connected Dominating Set
CCDS	Clique Connected Dominating Set
DCF	Distributed Coordination Function
EECDS	Energy Efficient Connected Dominating Set
FSM	Full State Markov
GPS	Global Positioning System
IEEE	Institute of Electrical and Electronics Engineers
IP	Internet Protocol
ISM	International Scientific Medical
KLD	Kullback Leibler Divergence
LOS	Line of Sight
HM	Hamiltonian Model
MAC	Medium Access Control
MCDS	Minimum Connected Dominating Set
MIS	Maximal Independent Set
MPDU	MAC Protocol Data Unit
RF	Radio Frequency
TC	Topology Control
TM	Topology Maintenance
UDG	Unit Disk Graph
VLSI	Very Large Scale Integration
WSN	Wireless Sensor Network
WLAN	Wireless Local Area Network
WPAN	Wireless Personal Area Network
WCDS	Weakly Connected Dominating Set
6LOWPAN	IPv6 Low Power Personal Area Network
V	Vertices
E	Edges
G	Graph

E_i	Initial Energy Level
E_d	Remaining Energy Level
RSS_c	Minimum Required Signal Strength
RSS_s	Signal Strength of Parent Node

Chapter 1

Channel Modeling and Topology Control in Wireless Sensor Networks

1.1 Introduction

Wireless Sensor Networks (WSN) can be defined as a wireless network consisting of spatially distributed autonomous devices, denoted as nodes, which can sense the environment and communicate the information gathered from the monitored field. Nodes in a WSN forwards data to other nodes in the network and can also forward data packets to the base station directly depending on the communication range which varies from few meters to hundreds of meters. In other words, WSN's can be configured to operate in a single hop and multiple hop scenarios.

WSN's have several common aspects with wireless ad hoc network and in many cases they are simply considered as a special case of them. Therefore, a sensor network normally constitutes a wireless ad-hoc network, meaning that each sensor supports a multi-hop routing algorithm. The data is forwarded, possibly via multiple hops, to a sink (also denoted as controller or base/receiving station) that can use it locally or is connected to other networks (e.g., the Internet) through a gateway node. WSN's have generated an increasing interest from industrial and research perspectives and have enabled new applications and new markets [1]-[10]. Due to this fact, the variety of possible applications of WSN's in the real world is practically unlimited, and varies from environmental monitoring, health care, positioning and tracking, to logistic, localization, and so on [11]-[14]. This possible classification distinguishes applications according to the type of data that must be gathered in the network.

On the other hand, the standardization process in the field of WSNs remained very active during past few years and an important outcome is represented by IEEE 802.15.4 which specifies a short-range communication system intended to provide applications with relaxed throughput and latency requirements in Wireless Personal Area Networks (WPAN). In addition, due to the increase in the popularity, several other standards are currently under development such as 6LOWPAN (IPv6 low power personal area networks).

Sensor nodes are typically equipped with a radio transceiver, a small microcontroller, and have small memory and a processing unit with limited computational power. In addition, an energy source, usually a battery, is also attached with the nodes. The WSN nodes normally operate in 2.4GHz ISM (International scientific and medical) frequency band with data rate varying from 30Kbps to 1Mbps. A WSN works as a network of nodes that cooperatively sense and may control the environment enabling interaction between persons or computers and the surrounding environment. The nodes in a WSN can be stationary or moving and may or may not be aware of their location. The two major participants in a WSN are the source and the sink. The source is a sensor that senses the data in its deployment scenario and reports the measurement to the sink. On the other hand, sink is interested in receiving data from the sensor nodes and can be either part of a WSN or an external device such as a laptop. In general, there is one sink however; multiple sink deployments are also possible. A typical WSN is shown in Figure 1-1 in which node A represents a sink node while rest of the nodes represents the source nodes from which sink node is interested in receiving the information. This form of representing a WSN is sometimes also called as a graph theoretic representation and allows researchers to prove mathematical formulations in terms of graph theory. As shown in Figure 1-1, the links are formed among different nodes with which communication is possible or those nodes lying in the communication radius of each other.

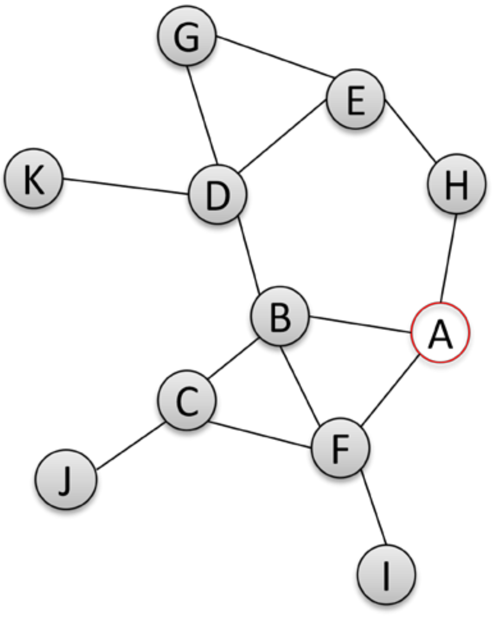


Figure 1-1: A typical 11 node Wireless Sensor Network.

Depending on the application, three types of communication models are commonly used. In the event driven reporting model, the sensor node reports to the sink node in case if certain event occurs. On the other hand, in periodic reporting model, sensor nodes periodically report events to the sink node. In the third reporting model which is a query reporting model, sensor nodes satisfying the query broadcasted by the sink node responds to the sink node. In all three communication models, the nodes can be deployed in a centralized and a decentralized manner. However, large scale deployment of WSN requires classical decentralized deployment, and therefore greater challenges exist in a WSN. In the context of decentralized detection, cooperation allows exchange of information among sensor nodes to continuously update their local decisions until consensus is reached across the nodes. Therefore, there are stringent power constraints for each node, as communication channels between nodes and the fusion center are severely bandwidth-constrained and are dependent on limited power supply. Therefore, one of the most important issues in the design of WSN's is energy efficiency. High energy efficiency means long network lifetime and limited network deployment and maintenance costs. Moreover, due to the wide variety of possible applications of WSN's, system requirements could change significantly. For instance, in case of environmental monitoring applications, energy efficiency is desired as nodes have a limited power supply. However, the amount of data to be sensed can be limited in forest fire detection and therefore communication must be robust to failure.

The main features of WSN's, as could be deduced are: scalability with respect to the number of nodes in the network, self-organization (can also be termed as decentralization), reliability (link level or hop by hop), energy efficiency, a sufficient degree of connectivity among nodes, low-complexity, low cost and size of nodes. The protocol architectures and technical solutions providing such features can be considered as a potential framework for the creation of these networks, but, unfortunately, the definition of such a protocol architecture and technical solution is not simple, and lots of research attention is still required. Due to this reason, accurate system simulation and analysis is necessary to validate emerging WSN protocols and applications. Therefore, before their deployment on actual sensor devices, an accurate channel model is a fundamentally important requirement for reliable theoretical or simulation-based study of wireless protocols and systems.

1.2 Problem Statement and Objectives

In this section, importance of channel modeling and topology control within the context of WSN is explained. Moreover, the desirable objectives in terms of accurate channel model and efficient topology control are also highlighted.

The channel modeling translates into two approaches which includes physical layer channel modeling and Medium Access Control (MAC) layer channel modeling. In physical layer channel modeling, constructive and destructive fading components are taken into account for characterizing the lossless environment. However, an alternative approach of modeling a wireless channel is to include physical-layer errors as a part of the underlying channel model and then model the residual channel (i.e., the channel observed after physical-layer processing) at the MAC layer.

The MAC layer channel modeling has been used to improve design of communication channels and systems for many decades [15]-[28]. Modeling the channel helps gaining insights into characteristics of the underlying error random process. This insight is essential for design and performance evaluation of a wide range of wireless protocols, applications and services. For instance, accurate channel models can facilitate design, parameter tuning and verification of the following wireless protocols:

- If the channel model is able to provide real-time error characterization at different hops of the network, the routing protocols for WSN can use MAC-to-MAC channel models to differentiate reliable versus shortest routes to different destinations [29], [30].
- MAC layer channel model can help characterizing the loss due to congestion or a loss due to vagaries of wireless communication in a WSN. Knowledge of losses due to channel errors, which is assumed in many wireless congestion

control solutions, can be provided by a real-time MAC to MAC channel model which helps tuning the congestion control protocol [31], [32].

- When the physical and MAC layer knowledge of the wireless medium is shared with the higher layers, in order to provide efficient methods of allocating network resources and applications, the real time MAC layer channel model can be used to choose reliable (e.g., using MAC layer retransmissions) versus cross-layer (e.g., ignoring data payload errors [33], [34]) protocols.
- An accurate channel model can predict future error characteristics, thereby saving the MAC layer protocol, and the overhead of switching to an inaccurate lower/higher data rate based on short-term observations.

Similarly, with the help of an accurate channel model, the design of many wireless applications can be improved. For instance:

- Real-time channel estimation provided by an accurate model can be employed by rate adaptive applications to perform channel and source-coding rate adaptation for efficient bandwidth utilization.
- Many wireless applications require different error control coding schemes. Therefore, a thorough understanding of errors above the physical layer can help in the design of effective error-control scheme.
- Error-resilience features of contemporary multimedia codec's can be effectively designed and verified with knowledge of MAC layer error characteristics.

It is worth noting that most benefits of a wireless MAC layer channel model can be realized if the model is able to provide real-time channel characterization and prediction.

It has been shown that the bit errors over a residual wireless channel exhibit bursty behavior, where each successful and unsuccessful packet transmission is dependent upon whether or not the previous K packet transmissions were successful. Such a channel is adequately modeled using a K^{th}-order Markov channel model in which the states of the model correspond to all possible combinations of correctly- and erroneously-received K previous packets. In complexity and power- constrained WSN environments, real time channel characterization is only possible with a low-complexity model or a model capable of reducing the complexity associated with the markov chains. On the other hand, the complexity of the markov model increases with the increase in the memory length and the current models do not cater for the complexity associated with the markov chain. Therefore, the probability distributions of good- and bad-bursts are analyzed for a FSM (Full State Markov) chain of arbitrary order. Moreover, the probability distributions are derived in terms of FSM chain transition and steady-state probabilities. These distributions render useful insights into important FSM characteristics, which are used to develop guidelines for defining FSM state space partitions in a low complexity model. Due to this reason, the objective of the present analysis is to ascertain partitions of FSM state space. With the help of guidelines, FSM states in a particular partition are then grouped together to form an aggregate state in the low-complexity approximating model. Moreover, it is also desired that the FSM state space partitions are performed such that the resulting aggregate process, while being less complex, closely matches the FSM chain characteristics. For characterizing the FSM state space partitions, they are represented as the order of nodes in a graph theoretic way e.g. if FSM has total of 8 states, then they can be represented as 8 nodes among which different transitions are possible.

While the size of sensor devices is rapidly decreasing, the slow improvement of the energy density in batteries aggravates the problem of energy limitations in WSNs. These constraints increase the importance of energy-efficient designs in all aspects of wireless communication. Therefore, energy efficiency is desired at different levels starting from the technology level (e.g., by adopting hardware that consumes less energy), energy efficient MAC, routing protocols up to the application level. As an example, at physical and MAC layers, nodes could operate with low duty cycle by spending most of their time in sleeping mode to save energy. However, new problems arise as nodes may not wake up at the same time, due to the drifts of their local clocks, thus making the communication impossible. Therefore, suitable network synchronization schemes and energy efficient and robust protocols are required in WSN's. Due to the need of energy efficient operation, the importance of topology control increases since it allows to increase the network lifetime while ensuring network wide connectivity with other nodes. Topology control is a well known strategy to preserve energy and increase the network lifetime [35]-[45]. Topology control protocols reduces the transmitting power of network nodes by considering network connectivity and the most famous topology control technique is to form a backbone using a Connected Dominating Set (CDS). As an example, nodes in Figure 1-1 can reduce their transmission power so that they communicate with only few neighbors. However, it is assured that in such a way that network wide connectivity is achieved while reducing the transmission power. On the other side, forming a reduced topology requires several considerations to be taken into account and in order to allow topology control protocols to have a widespread use; the energy constraint problem must be resolved. For instance:

- The control overhead must be taken into account to save bandwidth which in addition will save energy as well. The control overhead is related with the number of exchanged messages which increases with the increase in the number of nodes. Therefore, data reduction in terms of protocol overhead is a key to the energy efficiency of the system and many techniques and algorithms are proposed for this purpose but a need to decrease the overhead in terms of exchanged messages still exists.

- The reduced topology must not only save energy but also fulfills the primary task of a WSN which is sensing i.e. the nodes in the network must be capable of providing better sensing coverage with connected set of nodes in order to communicate sensed information to the sink node.
- Critical applications desire not only connectivity and an increase in the network lifetime, but also demand a reliable flow of information towards the sink node. Due to this reason, the importance of network wide reliability increases. The reliability in the network can be provided by maintaining redundant paths by which if link in one path fails then alternative path is available for the flow of information.

However, increasing the redundancy in the network impacts the energy efficiency among nodes in the network. Therefore, the prime objective should be finding a compromise between energy efficiency and network reliability.

Most of the studies describe methods to increase the network reliability but do not explain the impact of having the network reliability on energy consumption [46]-[52]. Most of the time, it is assumed that the pair of nodes are either connected or disconnected. In practice, most wireless links are lossy links with link success probability which desires that a routing protocol provides routing flexibility and reliability at a minimum energy consumption cost. For instance, topology control protocols like A3 [53], EECDS [55], CDS-Rule K [54] explain the topology construction approximation but do not explain the vulnerability associated with 1-connected CDS [56], [57]. However, algorithms for constructing K- connected dominating sets are proposed in the literature [56], [57], [58], but unfortunately they do not explain the network reliability impact on energy consumption.

Therefore, if the design objective is to have a reliable protocol, then energy efficiency should also be considered.

In addition to the above constraint, a protocol or a method can be sustainably viable only if it constructs topology while considering topology maintenance. Therefore, the importance of topology maintenance cannot be neglected. In topology maintenance, nodes shift their role to other nodes in a predefined manner after the construction of the topology; therefore it is of utmost importance to have such a topology construction protocol which is able to shift the nodes role to those nodes which have better energy resource among other network nodes. Otherwise, topology maintenance would not work and increase in the network lifetime becomes impossible. Most of the studies focus towards reducing the topology or constructing the topology in an energy efficient manner but do not highlight the importance of topology maintenenance [53], [54], [55]. However, topology maintenance is a key factor in determining an energy efficient protocol and should be considered in the design process of the protocol.

1.3 Motivation

Basic challenges in wireless sensor networks have sparked a number of different research themes. In fact, the activity of sensing, processing and communication under limited amount of energy, ignites a cross-layer design approach typically requiring the joint consideration of distributed signal/data processing, medium access control and topology control protocols. In addition, the predicted cost of WSN equipment is variable, ranging from hundreds of dollars to a few, depending on the complexity of the individual sensor nodes. Therefore, cost constraints on sensor nodes result in corresponding constraints on resources due to which most researchers rely on system level simulation.

Due to a lack of available infrastructure to perform realistic WSN experiments, system-level simulations are used to evaluate the performance of emerging wireless protocols and services. An accurate model of the wireless channel is an important component of such simulation-based performance evaluation. In the past three decades, channel error modeling techniques have been used extensively to improve the design of communication channels and the protocols that operate on these channels [15]-[18]. Using an accurate channel model, one can simulate the channel and can gain insights into the channel's underlying behavior. More importantly, an accurate and low-complexity channel model can be used to tune critical parameters of network protocols and applications at design time and in real-time. Lastly, a low-complexity channel model also allows real-time channel characterization and prediction which is required by rate adaptive protocols and applications.

In the channel modeling context, stochastic models have gained significant research attention [19]-[24]. In particular, high-order Markov channel models have been shown to be quite accurate in modeling link layer bit-errors and packet losses [19], [20], [23], [24], [25]. Unfortunately, the complexity of Markov models increases with their memory length and consequently the viability of using Markov model in resource-constrained wireless environment is very limited. Thus, accurate approximations of high-order Markov channel models are needed for wireless environments.

Many models have been proposed in the literature to reduce the complexity of high-order Markov chains [23]-[26]. While there exists a clear tradeoff between complexity and accuracy (lower the model complexity, lower the accuracy), existing low-complexity channel models [66] (with the exception of the bipartite model [25]) reduce the channel model's complexity to a fixed level and therefore do not cater for the emerging heterogeneous communication devices. For instance, on a given channel, high-end wireless devices (e.g., desktop and laptop computers) can afford higher complexity channel models than low-end devices (e.g., PDAs and smart phones.) To cater for such device heterogeneity, channel models are needed that can adapt their complexity to an arbitrary level in accordance with the resources available at a wireless device.

As explained before, in sensor networks, energy is valuable as nodes get energy supply from a battery source and due to deployment of nodes in uncertain locations it is always impossible to change the energy source. Therefore, the design of energy efficient communication protocols is a very peculiar issue of WSN's, without significant precedent in wireless network history. Generally, when a node is in transmit mode, the transceiver drains much more current from the battery than the microprocessor in active state or the sensors and the memory chip. The ratio between the energy needed for transmitting and for processing a bit of information is usually assumed to be much larger than one (more than one hundred or one thousand in most commercial platforms). For this reason, the communication protocols need to be designed according to paradigms of energy efficiency. Due to this reason, topology control is a very popular technique in a WSN to save energy and extend the lifetime of the network [35]-[45]. The main goal is to reduce the number of active nodes and active links, preserving the saved resources for future maintenance with the help of two phases namely: topology construction, which forms the reduced topology, and topology maintenance, which allows the maintenance of the reduced topology, so that characteristics like connectivity and coverage are preserved.

Most of the literature on WSNs deals with the design of energy efficient protocols which causes nodes to go into sleep/wake mode. However, over time, nodes will fail, they may run out of energy, overheat in the sun or even in fixed positions. Moreover, quality of RF communication links (and, thus, node's topologies) can change dramatically due to the vagaries of RF propagation. Due to this reason, coverage area of nodes increases/decreases depending on the transmission power. Therefore, topology control protocols suffer from volatility in the network when nodes move around and frequently form or sever many links. For this purpose, nodes should use minimum energy and network topology in such a manner that every node is accessed by any other node or can communicate with the sink node. In other words, the topology of the network while constructing a reduced topology should remain connected in case of any mishap or in other words nodes should adapt to the rapid change in their environment. Due to this reason, network reliability becomes an important aspect in topology control protocols. Network reliability is the probability that there is at least one spanning tree or the probability that the node or vertices can communicate with each other in case of random edge failures or the network is connected. Hence, reliability is another critical parameter that gives information about the connectivity of the network. Therefore, due to the desire of robust operation, this aspect has received interest in the research community [46]-[52]. However, increase in network reliability increases the energy consumption as redundant links are added to increase the network reliability. Therefore, reliable and energy efficient topology control protocols are much needed in WSN's. Therefore, in this report, topology control protocols are presented for energy efficient transmission in a WSN. In addition, the compromise between the energy efficiency and reliability is also explained by presenting a protocol for reliable communication.

1.4 Report Contributions

In this report, a new variable-complexity wireless channel model referred to as the Hamiltonian Model (HM) is presented. The HM reduces the complexity of high-order Markov channel models by identifying and aggregating Hamiltonian circuits present in the states of the Markov chain. Given a desired complexity budget in terms of the total number of Markov states, the proposed model identifies a Hamiltonian circuit in the Markov chain, finds cycles of the needed complexity, and then aggregates these cycles into odd and even states based on the number of total states present in the cycles. In addition, the transition probability matrices are further analyzed and reported for bit patterns that never occur in the collected traces and are referred to as the unused states. Moreover, when the unused state parameter is available, this report demonstrates that the complexity associated with higher order Markov channel models can be further reduced.

The performance of the proposed model is compared with the Bipartite Model (BM) of [25] which is a graph theoretic channel model. Therefore, bipartite model is selected to compare the performance characteristics of graph theoretic Hamiltonian Model using a comprehensive data-set of actual traces collected in two different environments: (i) 802.11 MAC layer bit errors at 5.5 Mbps; (ii) 802.15.4 MAC layer bit errors at 250 Kbps. The accuracies of the models are compared by measuring their closeness to the actual wireless channel traces using an information-theoretic Kullback–Leibler-based divergence measure and by comparing the Cumulative Distribution Functions (CDF) of bit errors. The results presented in this report demonstrate that HM has significantly higher modeling accuracy than BM.

The report also contributes towards efficient topology control protocols for WSNs. For this purpose, the clique connected dominating set (CCDS) protocol is presented which models the network as a connected graph and finds the number of 2-cliques present in the network. To reduce the control overhead, CCDS forms a backbone by using the inherent broadcast nature of WSN's. The CCDS protocol is also used with topology maintenance techniques which in turn increase the network lifetime. The CCDS protocol is completely distributed and easily scalable. The CCDS does not need the position or orientation information of the nodes. Moreover, the results demonstrate that the CCDS attains better energy efficiency which in turn provides longer network lifetime. Similarly, the impact of reducing the topology on sensing coverage region is also explained in proposed A1 protocol. The A1 protocol is also distributed and uses only one message to reduce the message overhead and to achieve energy efficiency. In addition, it provides better connected sensing area when compared with the other protocols.

It has been demonstrated that backbone construction using CDS cannot provide network reliability. Therefore, to achieve network reliability, a protocol named Poly is presented which exploits polygons in a network providing polygenic backbone, while keeping the network connected and covered. To achieve energy efficiency, the protocol forms a CDS like polygenic network which in turn provides reliability in the case of random link failures. Moreover, it adapts to topological changes in the network based on the remaining energy of the nodes. This allows topology maintenance among different set of nodes to increase the network lifetime. The Poly protocol is scalable and does not need any synchronization scheme. The Poly protocol can be applied to different data reporting models which includes the rendezvous design problem for WSN's with a mobile BS, which aims to find rendezvous point's (RPs) and provides polygenic redundancy to RPs. To validate the findings, extensive simulations are performed to demonstrate that the proposed protocols can achieve satisfactory performance under range of settings and different underlying topologies. In addition, the reliability of Poly is mathematically modeled and compared with other CDS protocols to show that it achieves better connectivity under highly dynamic network topologies. Similarly, the objective as part of the future work is also explained which is to calculate a desired reliability budget, if numbers of edges are given and where these edges should be induced to achieve the maximum increase in reliability.

1.5 Organization of the Report

The rest of this report is organized as follows. Chapter II describes related work in the channel modeling area and provides the background that is required to understand the contribution of this report within the context of variable complexity channel model. Chapter III describes experimental setup and the error trace collection on Wireless Sensor and Local Area Networks. Moreover, the proposed Hamiltonian model is explained in detail and the performance of the proposed model is compared to that of a bipartite model on various metrics. In addition, Chapter III summarizes the key conclusions obtained from the Hamiltonian model and other graph theoretic channel models.

The rest of this report is organized as follows. Chapter IV presents introduction to topology control and explains the key feature and design requirements for an efficient topology control protocol. It also explains related work required to understand the working of topology control protocols for WSN's. The related work also includes those protocols which are later used in the performance evaluation of the presented protocols in this report. In addition, performance metrics are explained with reasons for their selection in the evaluation process. Chapter V describes the detailed description of the Clique Connected Dominating set (CCDS) protocol while showing the experimental setup and performance evaluation with existing protocols. In Chapter VI, A1 protocol is presented which is fascinated from the A3 protocol. Moreover, performance evaluation of the A1 protocol with other protocols is also presented with focus on the sensing coverage issues of topology control protocols. Chapter VII describes the working of the Poly protocol. In this chapter, the mathematical model is explained which allows to evaluate network reliability of the protocols. Moreover, the experimental setup used to evaluate the performance of the protocol with existing protocols is also explained. For the future work in progress and the future directions for interested readers, Chapter

VIII explains the packet delivery and link reliability of CDS based topology control with sensing ranges issues in WSNs. Moreover, the chapter also summarizes the key conclusions obtained from the report.

Chapter 2

Channel Modeling: Background and Related Work

2.1 Introduction

Wireless bit-error processes are generally bursty and have a memory-length of greater than one bit, and therefore these processes cannot be modeled using a two state model. Therefore, to model a bursty behavior, a Markov chain is defined such that at each time instance, the process is characterized by as many bits as the memory-length. At each time instance, a new bit is added to the memory-window and the oldest bit is dropped from the memory-window. Therefore, this chapter first provides the background literature review to the nature of binary wireless traces and explain the K^{th} order Markov chains in detail. In addition, it also explains the proposed methods to reduce the complexity of higher order markov chains used for wireless channel modeling.

From this chapter, many intuitive observations regarding FSM chains can be stated. These observations are used in subsequent sections to derive important characteristics of FSM chains. It is important to outline the approximation of FSM chains. For this purpose, the approximate models in this report are developed by creating partitions of the FSM chain state space. All FSM states in a partition are then simply aggregated / grouped into an *aggregate* state of the approximate process with the help of graph theory. Hence, this chapter mainly addresses the following question: *How should one define partitions on the FSM state space such that the resulting aggregate process accurately approximates the underlying FSM chain?* In other words, this chapter finds out which FSM states can be aggregated together without compromising the FSM model's performance.

2.2 Representation of Binary Wireless Traces

A wireless sensor network consists of low-cost and low-energy sensors and a base station. These sensors are deployed in a vicinity of interest to monitor a phenomenon and send their observations to a base station which makes global decision. Traces collected over a wireless medium generally represent two states. One state is the good state and the other state is the bad state or the lossy state. Therefore, wireless traces can generally be characterized as a binary time series $\{x(n)\}_{n=1}^{\ell}$, where $x(n) \in \{0,1\}$ and ℓ is the length of the error trace.

Without loss of generality -- throughout this report -- zeros are used to represent an error-free bit and ones are used for a bit in error. The sequence of these bits forms alternating bursts of zeros and ones. If the burst consists of number of zeros then it is referred as a *good burst* and if the burst consists of number of one's then it is called a *bad burst*. The trace can hence be represented as pair of good and bad bursts: $(N_1, G_1), (N_2, G_2), \ldots, (N_n, G_n)$, where N_n and G_n represent the length of the n^{th} good and the bad bursts, respectively [1]. Many channel modeling studies have shown that this binary representation is suitable for representing channel traces [20], [25]-[28].

2.3 The Gilbert Channel Model

The Gilbert channel model shown in Figure 2-1 is used to model channels with 1st order memory. It has been used to model many wireless channels at bit, byte and packet levels [22], [27], [28]. The Gilbert model captures channel memory through a two-state Markov chain having a good and a bad state. The probability of the next (good or bad) symbol is dependent on the whether the last received symbol was good or bad. Since, wireless bit error processes are generally bursty and have a memory length of greater than one bit; therefore these processes cannot be modeled using the Gilbert model. For accurately characterizing the bit error processes, a Markov chain is defined such that at each time instance the process is characterized by as many bits as the memory-length. At each time instance, a new bit is added to the memory-window and the oldest bit is dropped from the memory-window. The K^{th} order markov chains are explained in the next section.

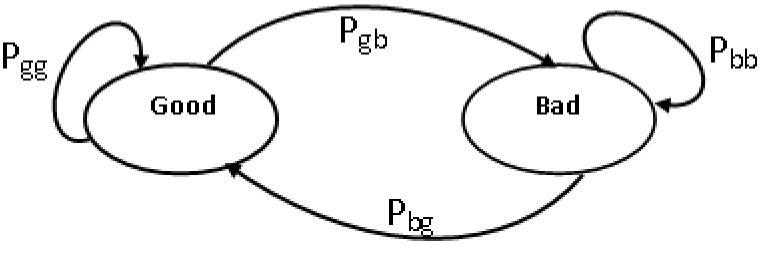


Figure 2-1: The Gilbert Channel Model.

2.4 K^{th} Order Markov Chains

A Markov chain of memory K is a discrete time random process whose probabilities for going to future states at a given present state are independent of the past states. For a memory length of K, the Markov chain comprises of 2^K possible combinations of K consecutive bits. Therefore, if a set of states K consists of $S = \{S_1, S_2, S_3, \ldots, S_n\}$, then the process starts in one of these states and moves successively from one state to another. If S_i is defined as a current state, then for moving to next state S_j, the probability will be denoted by P_{ij}. The probabilities are called transition probabilities and are computed by sliding bit by bit a K bit memory window over the data [59], [60]. In this report, Markov chain corresponding to even (odd) decimal numbers is referred as even (odd) states. Using the above notation, an example 3-rd order Markov chain is shown in Figure 2-2; only transitions to even states are shown.

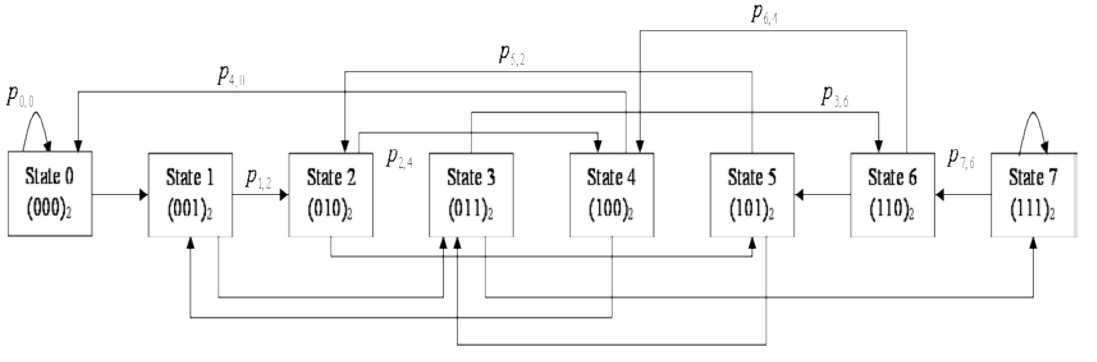


Figure 2-2: A 3-rd order (memory length=3) Markov chain.

If the Markov chain is in even state, the last received bit (i.e., the least significant bit position in the memory window) must be error-free. Similarly, a Markov chain in the odd states implies that the last bit was corrupted. Due to the binary nature of the underlying wireless bit-error process, each Markov chain state can transit to only two other states [59], [60], [61]. This is due to the Markov chain definition in which the memory-window at each time instance is left-shifted by one bit and a one or a zero bit is added to the least-significant bit position. Thus, from state S, a Markov chain can transit either to even state $(2S) \bmod 2^K$ or to odd state $(2S+1) \bmod 2^K$. Since the sum of all transitions from a Markov state must sum to one, therefore for any state S probability becomes $P_{(2S+1)\bmod 2^K} = 1 - P_{(2S)\bmod 2^K}$.

It should also be emphasized that once a non-corrupted bit is received, a K^{th} order Markov chain will return to state 0 (i.e., the no error state) only from state 2^{K-1}

after K transitions; see in Figure 2-2 that at state $2^{3-1} = 4$, the Markov chain wraps around to state 0.

2.5 Complexity Reducing Alternatives for Wireless Sensor Networks

In this section, some previous studies are presented which propose methods to reduce the complexity of high-order Markov models [20]-[27].

2.5.1 Lumpability Model

Chen and Rao [20], [21] used the lumpability framework to reduce the order of a Markov channel model. They described techniques for reducing FSM channel model to a two state model. They elaborated the Markov chain to be either strongly or weakly lumpable to a particular partition and stated that the original channel model can be exactly described by only 2-states. When the channel model fails to be strongly or weakly lumpable, a hazard rate ordering technique can be applied to stochastically upper and lower bound the original channel model to 2-state model. They showed that the upper bound is very tight and the goodness of the bounding process is based on higher-layer error control protocols. However, the lumpability conditions place very stringent constraints on the transition probabilities of a Markov chain. The presented ON-OFF model stochastically bounds the sojourn time distributions of the lumped good and bad states. However, an ON-OFF model assumes geometric (memory-less) distributions for good and bad periods. These constraints are generally not satisfied by real-life wireless channel models. Therefore, performance of the proposed technique with lumped Markov chains is not compared.

2.5.2 2-Tier Markov Model

For accurate system simulation and analysis, the authors in [22] present two different classes of WSN channel models for residual MAC layer bit-errors. In the first class of channel models, referred to as 1-Tier models, residual bit-errors are modeled as a stand-alone error process. This 1-Tier model is commonly known as the Binary Symmetric Channel (BSC) model. In this model, behavior of every bit is independent of any previous bit and the probability that a bit is in error does not change with any information about the previously received bits. Similarly, in the second model, frame-error model and a bit-error model is combined to form a 2-Tier model. The frame-error model at tier 1 is used to excite the bit-error model at tier 2. That is, if the tier 1 frame error model predicts an error-free frame then there is no need to invoke the bit-error model because it is known that all the bits in the frame are error-free. On the other hand, if the tier 1 frame-error model predicts a corrupted frame then the bit-level model at tier 2 is used to generate bit-errors in the corrupted frame.

To employ the model for both the tiers, markov models of different orders is used. At tier 1, a Gilbert model for frame errors is used and at tier 2, a 3-rd order Markov model for bit-errors is used. Thus, this model incorporates memory at both frame and bit levels.

2.5.3 The Markov based Trace Analysis (MTA)

In [23], authors present the *Markov based Trace Analysis* (MTA) algorithm, for the design of channel error models. The MTA algorithm assumes that a trace with non-stationary properties can be decomposed into a set of piecewise stationary traces consisting of “lossy” and “error-free” states. The MTA algorithm defines these states, and parameterizes transitions between them as a function of a preset parameter, which is the change-of-state constant. Therefore, a statistical constant from the wireless traces is derived and used to divide the previously non-stationary trace into stationary sub-traces representing lossy and error-free segments of transmission. By analyzing the length distributions of these segments, they effectively characterize the transitions between them, and create a model that more accurately represents the original trace.

2.5.4 Constant Complexity Model

In [24], guidelines were proposed to accurately model Markov based wireless channels and a constant complexity probabilistic model is proposed. Given a complexity budget in the form of the total number of states, the authors explain that some of the states should not be aggregated to develop an aggregate model with total number of states satisfying the complexity budget. Due to the high probability, they explain that FSM state 0 should not be aggregated with other states. Similarly, they also propose that FSM states 2^{K-1} and 1 should be aggregated with a minimal number of other states. Based on the analysis, they demonstrate that CCM approximates the behavior of the underlying 2^{K} state FSM quite closely.

2.5.5 Bipartite Model

Willig [25] proposed a scalable-complexity Bipartite Model (BM) for wireless channels. The bipartite model uses the notion of a binary indicator sequence which is a finite sequence of one's and zero's and divides the sequence according to its burst order. Binary indicator sequences are subdivided into error bursts and error free bursts according to a burst order K_o. They define an error-free burst of order K_o to be a maximum length contiguous all-zero subsequence with a length of at least $K_o + 1$. In contrast, an error burst of order K_o is a subsequence of at least one bit length and with ones at its fringes. Based on the burst order, the image of good and bad burst Probability Mass Functions (PMFs) are divided into burst intervals. A transition matrix is then computed for transiting between the burst intervals which describes the state transition from bad states to the good states and vice versa. In order to build a model from the traces one needs to choose the number of states, matrices, probability distributions and the bit error rates. The accuracy of bipartite model depends on a selected value of complexity. However, model accuracy should not be optional and even a low-complexity model should provide the requisite accuracy. Moreover, bipartite model require a large number of parameters to achieve a certain level of accuracy. In the next chapter, performance of the proposed model is compared with the BM, and is explained in detail.

2.5.6 Chaotic Map

In [26], authors use chaotic maps to model 802.11b bit-errors at low data rates (1 Mbps and 2 Mbps). Similar to the Gilbert-Elliot model, the chaotic map model uses two states “good” and “bad”. In the good state all bits are correct, while in the bad states, all bits are erroneous. Switching between states depends on the value of an auxiliary variable χ_t that is updated for each bit.

Due to the focus on low data rates, it was observed that: (a) probability of bit-error bursts of more than two bits is very low, and (b) there is almost no correlation in error traces. In addition, the chaotic map model ignores the correlation and captures only the heavy-tail behavior of bit-errors. While this assumption of “no autocorrelation in data” might be suitable for the particular experimental setup used in for the analysis, it is not generically applicable to network error and lossy data.

2.5.7 Linear Complexity Model

In [27], authors show empirically that low complexity hierarchal and hidden Markov models cannot characterize the bit error processes and proposes to employ high-order Markov chains for accurate channel characterization. They use an autocorrelation analysis to establish that the correlation drops to an in-significant value after a certain lag. Moreover, the maximum order of the Markov chain required to accurately represent the 802.11b MAC layer bit-error process is between 14 and 16 which results in 2^{14} to 2^{16} states. They further show that all FSM chains of order-9 and above represent the 5.5 Mbps bit-error process accurately. Similarly, the 2 Mbps bit-error process is adequately modeled by an order-10 FSM. This illustrates the feasibility of effective models with orders lower than the underlying process correlation. However, the complexity of the full-state model is still quite high and renders them impractical for most real-life simulation scenarios or real-time adaptive applications. Therefore, these studies, results in models with fixed non-scalable complexities. Since, complexity and accuracy of a model generally exhibit a direct proportionality relationship, therefore, in view of the heterogeneity of contemporary wireless devices, a model should be able to scale its complexity in accordance with the complexity that can be afforded at a wireless device. More specifically, given a complexity budget (for instance, in terms of the number of model states,) a scalable channel modeling algorithm should be able to produce a channel model to satisfy that budget. Only few studies ([20], [22], and [25]) approach the wireless channel modeling problem in this way, among which [25] is the graph theoretic model. Therefore, due to the similar nature of the model with the presented model, in the next chapter, only study in [25] is used in the evaluation of the proposed model.

2.6 Summary

This chapter presented a review of the concepts used in wireless channel modeling. Firstly, the binary nature of wireless traces is explained and later it was demonstrated that the Gilbert model cannot accurately characterize the wireless bit error random process. Therefore, the markov chains of order K are highlighted and properties of the markov chains were also explained. In addition, some related work and different types of low complexity models were explained -- that can be employed for wireless channels-- which however do not cater for the FSM chain performance accuracy and need large number of parameters.

Chapter 3

The Hamiltonian Wireless Channel Model

3.1 Introduction

In this chapter, collected data and some preliminary trace statistics used in the performance evaluation of Hamiltonian Model are discussed. To perform realistic performance evaluation over operational channels, a comprehensive dataset of wireless error traces was collected over two different channels: 1) an 802.15.4 WSN channel, and 2) an 802.11b WLAN channel. All traces were collected at the MAC layer after physical layer processing; MAC layer channels are referred to as *residual channels* in prior literature [22], [24]. Moreover, this chapter also analyzes how FSM chains capture good and bad-burst behavior of wireless channels and the implication of unused states on the bit error traces. In addition, the bit error rate (BER) statistics of collected traces is also explained.

Most benefits of a wireless MAC layer channel model can be realized if the model is able to provide real-time and online channel characterization and prediction. In complexity- and power-constrained wireless and mobile environments, such channel characterization is only possible with a low-complexity model. Despite some interest in reducing the complexity of wireless models [21], [25], development of accurate, pragmatic and low-complexity wireless channel models is still an open problem. Since, low-complexity models have not been thoroughly explored and verified for contemporary wireless and mobile networks, many of the protocols applications and systems have not been realized in practical wireless systems. The number of states of an FSM chain is an exponential function of the random process' memory-length - 2^k states for a process with a memory-length of k bits. This phenomenon is commonly referred to as state explosion. Due to state explosion, although FSM chains can provide accurate models of wireless bit-errors, their high complexity renders them impractical for realistic wireless environments.

To reduce FSM chains' complexity, in this chapter, *Hamiltonian Model* (HM) is also explained in detail. Later, the focus remains on directly approximating FSM chain behavior for which insightful observations about underlying characteristics of an FSM model are used. Using this analysis, important guidelines are used for the realization of accurate, effective and low-complexity models. These guidelines lead to a *Hamiltonian model* (HM) that always aggregates odd and even states without compromising the good and the bad states irrespective of the memory length.

In the end, the results are shown which demonstrate that the performance of the HM in modeling the real 802.15.4 and 802.11 MAC layer bit-error channels is comparable to exponential-complexity FSM chains and better than Bipartite model [25]. In the next section, discussion starts with the explanation on the collected traces.

3.2 802.15.4 Data Collection

Crossbow's Micaz motes were used [62] to collect residual bit-error traces over wireless sensor networks. These motes operate on the ISM frequency band of 2.4

GHz and support a peak data rate of 250 Kbps. Sensor motes were running the open-source TinyOS operating system [63]. The source code of TinyOS applications was modified to disable the MAC layer checksum feature at the receiver. Hence, corrupted packets were not dropped in a receiver's kernel, and were passed to a data logging application. The application logged all packets on an attached computer through the serial port. The traces at four different locations or setups were collected. These setups are named according to their geographical location as shown in Figure 3-1.

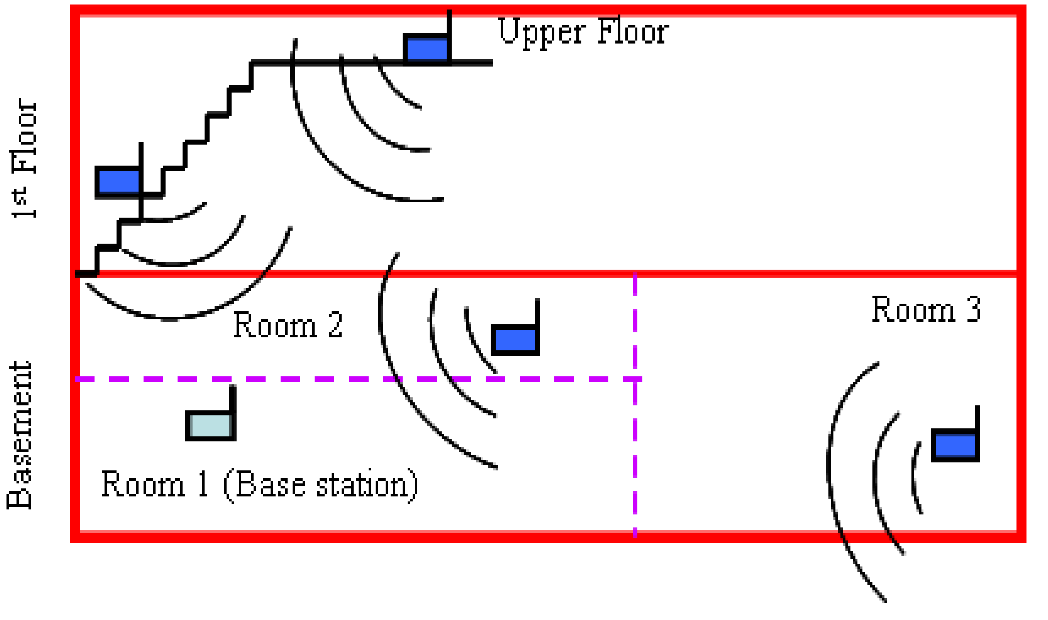


Figure 3-1: Setup for 802.15.4 bit error traces.

The light-shaded mote in Figure 3-1 is the base station which received and logged data, while the remaining motes are sending motes which transmitted packets with predefined contents to the base station. In each experiment, one sender transmitted uni-cast data to the base station and the other senders were inactive; i.e., in each trace collection, there was no channel contention and collisions as there was only a single sender and a single receiver. While performing experiments, motes were kept stationary.

In order to collect traces with varying error behaviors, the distance between the motes and the base station was varied from 5 to 12 meters. The senders transmitted 20-byte fixed-sized frames at a rate of 10 frames per second. At first, experiments were performed by having a direct line of sight (LoS) between the sender and the base station but the error rates observed in those experiments were too low to warrant further analysis. Therefore, in the experiments, non-LoS traces are focused.

The average number of frames per trace was approximately 31,000 frames. Thus, the average length of each trace was approximately 5 million bits [22]. For evaluation of HM and BM, five traces for each setup were used which counts to a total of 20 traces for all setups. While more actual wireless sensor network traces were collected, but the trends observed in 20 traces (5 per setup) are representative of the trends observed in other traces.

3.3 802.11 Data Collection

For 802.11b traces, the same process is repeated -- as it was in 802.15.4 traces -- using the topology shown in Figure 3-2. The AP (access point) was operating in Distributed Coordination Function (DCF) mode and three wireless stations communicating in the infrastructure network configuration. One of the stations was operating as the server and the remaining two as multicast clients. All wireless stations were Linux boxes using Dlink wireless cards with Prism2 chipset device drivers [64]. The server was stationary and transmitted a continuous stream of predetermined patterns to the multicast clients. Traces were generated for each bit rates at different stationary client positions with and without LoS. It was observed that, with clear LoS, the error rate at all bitrates was extremely low. Such excellent performance deemed further LoS study inconsequential. Hence, both clients were positioned in a separate room across two walls in order to simulate a more realistic business/classroom/home-network wireless setup and forced to receive non Line of Sight frames of 512-bytes at a physical layer data rate of 5.5 Mbps. The average length of each trace was approximately 6 million bits. At different locations, 3 traces for 5.5 Mbps were collected. Moreover, the traces at 11 Mbps were also calculated, but due to their non-Markovian behavior those traces are not used in this report.

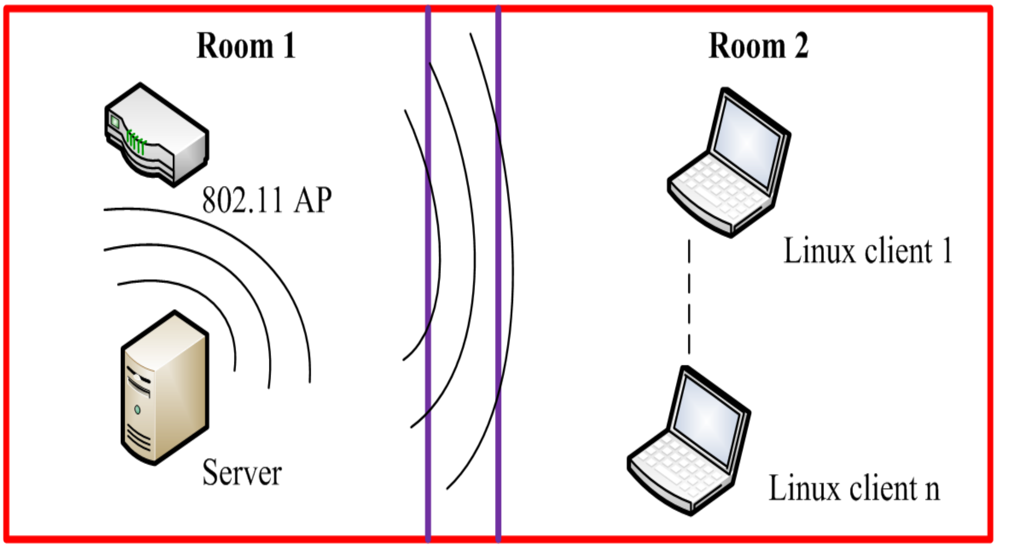


Figure 3-2: Setup for 802.11b bit error traces.

3.4 K^{th} Order Markov Chain Statistics

In order to accurately and efficiently represent transition probability data and to reduce the complexity associated with higher order Markov chains, the transition probability matrices for the 802.15.4 WSN and 802.11b WLAN collected traces were investigated. During investigation, it reveals that some bit patterns never occur in transition probability matrices and are referred to as unused states. In other words, such states result in all-zero columns of the transition probability matrix. An all-zero column implies that the probability of jumping to that state from any state is zero which is observed because that state was not present in the training data. The number of unused states is larger for the traces with lower BER because all the error states are not observed in a finite length trace. Conversely, as the error rate increases, the number of observing unused states decreases proportionally. Many unused states in 802.15.4 WSN traces were observed and their number grew as the order of Markov chain [28]. The total number of unused states per setup for 802.15.4 WSN traces is shown in Table 3-1. It shows that the number of unused states increases with the increase in the order of Markov chain and maximum number of unused states are observed at $K = 9$. The unused states in Room 3 are fewer as compared with other setups since the sender and receiver in Room 3 were separated by a concrete wall. Therefore, this setup has a high BER when compared with other setups. More unused states in case of Upper floor were observed, in which sender and receiver were at the farthest distance from each other. However in 802.11b WLAN traces, there is no unused state up to $K = 9$ i.e. the order by which HM is tested for complexity reduction.

Table 3-1: Total Number of Unused States per Setup.

Setup	**Number of States**					
	16	**32**	**64**	**128**	**256**	**512**
Room2	0	0	6	41	209	772
Room3	0	0	5	36	138	425
Stairs	0	0	0	13	129	595
Upper floor	0	4	46	205	610	1512

3.5 Trace Bit Error rate (BER) Statistics

Table 3-2 shows the average bit error rates of 802.15.4 and 802.11b original traces, respectively. It can be seen that highest bit error rate was observed for Room 3 as in this case the sender and receiver were separated by a concrete wall which was in between them. Room 2 has the lowest bit error rate as in this case the sender and receiver were at the closest distance from each other and there was a glass window in between them. The bit error rate of Upper floor and Stairs are similar because of being in the same vicinity but the distance of Upper floor was more as compared to Stairs setup. In case of 802.11b where the receiver's were kept stationary, the average bit error rate observed at different locations was 0.003. In the next section, the graph theoretic Hamiltonian model is explained in detail followed by the discussion on BER statistics of the artificial traces.

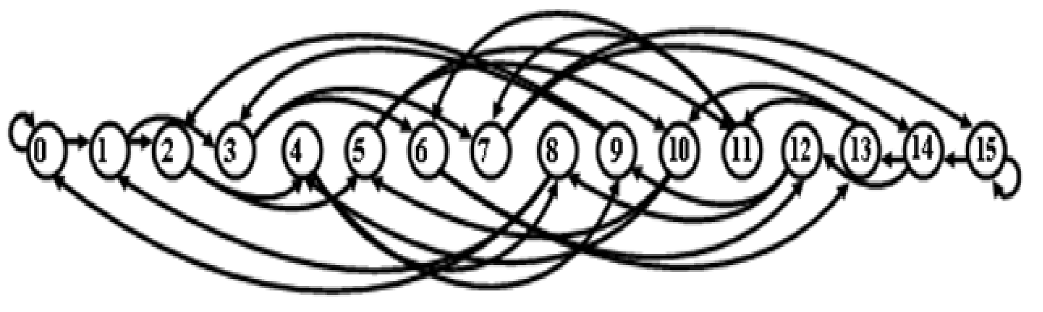
Table 3-2: Average Bit Error Rate of Actual Traces.

	Setup	**BER**
802.15.4	Room2	0.00085689
	Room3	0.01519663
	Stairs	0.00737452
	Upper floor	0.00735198
802.11b	Location 1,2,3	0.003067968

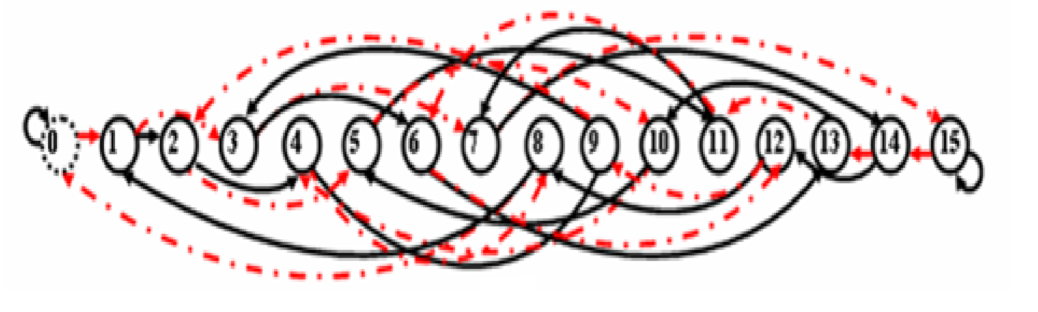
3.6 The Hamiltonian Wireless Channel Model

Let a K^{th} order Markov chain be represented as a K vertex connected digraph $G=(V,E)$ with positive edge weights. Markov chain of any order forms a connected graph as long as both state transitions probability in each Markov state are greater than zero. In addition, Markov chains exhibit many interesting graph-theoretic properties which can be used to reduce the complexity associated with higher order Markov chains and those properties can be used to develop a scalable model. A graph in which each vertex or node (referred to as a state in case of markov chain) is traversed exactly once forms a Hamiltonian circuit. A Hamiltonian circuit can be identified in Markov chains of any arbitrary order. By using this property, the states can be further arranged according to the node traversed, which gives an easy method for aggregating states that comprise the Hamiltonian circuit. Since a cycle of arbitrary length can be identified in the Markov digraph, the states of the circuit can be aggregated to a desired state budget. These characteristics are a consequence of the Markov chain construction and are therefore present at all orders of the Markov chains. Moreover, the Hamiltonian circuit formed in the Markov digraph clearly identifies the good and bad nodes which remain separated during state aggregation. This is a very important property because it is generally undesirable to merge good (even) and bad (odd) states together [25]. After Hamiltonian state aggregation, the merged states probabilities are aggregated and normalized into one aggregate state of the low-complexity model. This graph-theoretic realization helps us in reducing the complexity in a finite time and in a simple and easy manner. In the following discussion, an example of applying the proposed Hamiltonian Model based state aggregation on a Markov chain of order $K = 4$ is explained.

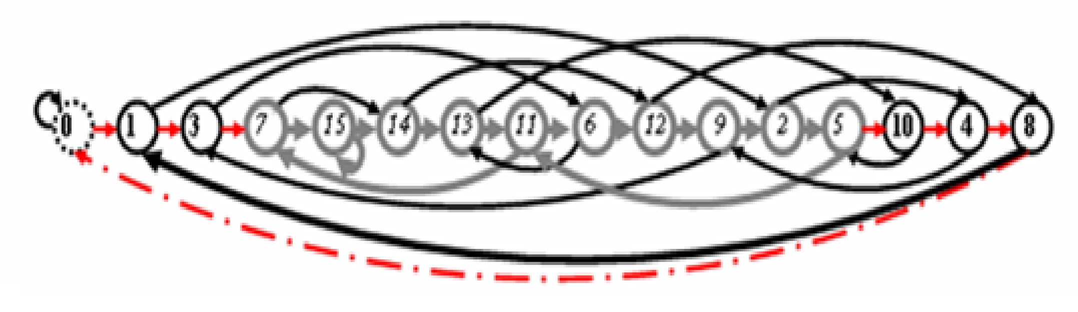
A Markov chain of order $K = 4$ is shown in Figure 3-3 [a]. The state zero is chosen as a start to traverse all the other vertices and by that the order of traversed vertices in the consequent Hamiltonian circuit will be: 0-1-3-7-15-14-13-11-6-12-9-2-5-10-4-8-0, respectively. This is depicted with dotted line or edges in Figure 3-3 [b]. After finding a Hamiltonian cycle, it was observed that there are n cycles in the graph. Based on the cycles still left in the graph, the algorithm is applied to



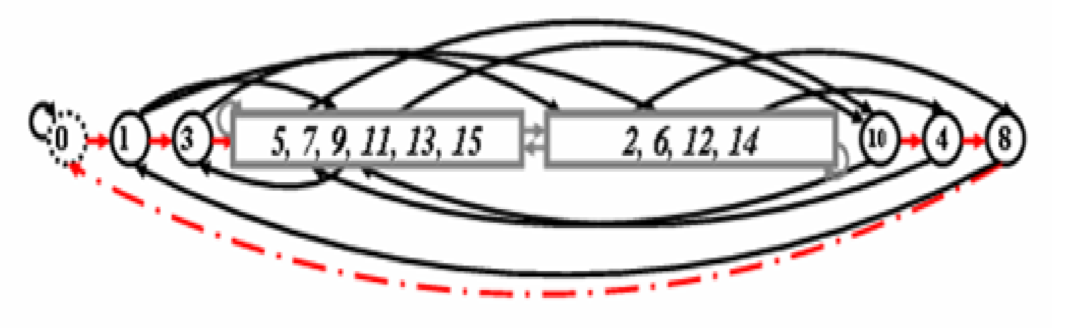
a) Markov chain of order $K = 4$.



b) Hamiltonian circuit.



c) Hamiltonian circuit arranged by vertices traversed.



d) Hamiltonian circuit reduced to $K = 3$.

Figure 3-3: The Hamiltonian Model.

work in such a way that it finds a cycle exactly aggregating the graph into $K = 3$ states. This procedure is shown in Figure 3-3 [c]. At first, a cycle consisting of 5 vertices (7, 5, 14, 13, and 11) is chosen and then again another cycle coming from vertex 5 towards vertex 11 is chosen. Hence, the cycles consisting of ten nodes is chosen in the start which is separated into odd and consecutive even nodes further.

Figure 3-3 [d] shows the final transformation from $K = 4$ to $K = 3$. The probabilities of the aggregated states in odd and even parts are normalized. The HM model is compressed from $2^9 = 512$ to $2^8 = 256$ and then to $2^7 = 128$ and so on up to $2^3 = 8$. The algorithm is applied on the graph as depicted in the figures to generate artificial traces at each state by the model. It is also possible to find the Hamiltonian circuit by traversing vertices different than the traversed vertices shown in Figure 3-3. For that the model remains the same and is still applicable in reducing the needed complexity.

3.7 Discussion

In this section, the error traces are used to compare the performance of the Hamiltonian Model with the Bipartite Model using Bit Error Rate (BER), Kullback-Leibler Divergence (KLD) and bit error distributions.

3.7.1 Bit Error Rate (BER)

The BER is calculated from 802.15.4 and 802.11b synthetic traces generated by Hamiltonian Model and the Bipartite Model. Table 3-3 shows the Average BER per setup of Hamiltonian Model traces parameterized from actual traces at $K = 6$, $K = 7$ and $K = 8$. It also shows the BER of Bipartite Model for varying number of K-states.

Table 3-3: Average Bit Error Rate of synthetic traces.

Model	Setup	Number of States		
		64	**128**	**256**
Hamiltonian_802.15.4	Room2	0.004119	0.001383	0.000416
	Room3	0.016309	0.016735	0.021816
	Stairs	0.021501	0.009547	0.011524
	Upper Floor	0.01853	0.008242	0.014039
Hamiltonian_802.11b	Location 1,2,3	0.007559	0.002649	0.005284
Bipartite_802.15.4	Room2	0.0259419	0.0388292	0.0540207
	Room3	0.0203247	0.0324104	0.0725722
	Stairs	0.0066553	0.0099122	0.0453318
	Upper Floor	0.0373605	0.0800648	0.1279515
Bipartite_802.11b	Location 1,2,3	0.038968	0.03524	0.030718

Comparing with Table 3-2**,** it can be observed that BER of Hamiltonian Model (HM) traces are closer to the actual traces' and better than the Bipartite Model (BM). Overall, the inaccuracy in BER estimates shows the opposite trends in HM and BM. For the HM, the inaccuracy decreases with an increase in the number of states except for Room 3 which has a larger number of unused states when compared with other setups. The BM, on the other hand, incurs more inaccuracy for higher number of states. Thus, for the BM introducing more state does not necessarily increase the accuracy of the model. Typically, an increase in the complexity of a model (e.g., with an increase in the order of a Markov chain) causes the accuracy of the model to improve. However, for the HM model, this proportionality trend between complexity and accuracy was not observed.

On the contrary, after empirically evaluating the HM model's accuracy for varying state merging orders, it was observed that a particular HM order provides better accuracy than orders above and below it when unused states are not used. When the unused state parameter is available, the complexity associated can be reduced with the model by tuning the unused state parameter. This is also shown in Table 3-4 for $K = 6$, $K = 7$ and $K = 8$.

Table 3-4: Average Bit Error Rate of Synthetic Traces with Unused States.

Model	**Setup**	**Number of States**		
		64	**128**	**256**
Hamiltonain_802.15.4	Room2	0.000898	0.000841	0.000835
	Room3	0.01614	0.015131	0.015937
	Stairs	0.007777	0.007411	0.006951
	Upper Floor	0.00782	0.007335	0.007236

For generating HM synthetic traces, model is tuned in such a way that when unused states occur in transition probability matrix the model should always jump to a good state. The notion of jumping to good state when zero columns occur in transition probability matrix is taken from the overall behavior of the WSN traces which shows that 99% of the times the state transitions are in good state. Hence, when compared with Table 3-2, the model shows very similar bit error rate for varying number of K-states in case of 802.15.4 WSN traces and is shown in Table 3-4.

While the BER estimates give us an overall picture of the accuracy of a model, for more elaborate performance comparison, the burst distributions of the HM and BM are further compared. For this purpose, the good- and bad-bursts' distributions of the two models are compared in the next two subsections.

3.7.2 Kullback-Leibler Divergence (KLD) of Good- and Bad-Bursts

Entropy is a measure of the average number of bits required to represent all outcomes of a probability distribution. The Kullback-Leibler divergence quantifies the difference in the entropies of two probability distributions [65]. The KLD divergence quantifies the source-coding-like overhead incurred by employing a model instead of the actual source. For two probabilities distributions p and q defined over a common alphabet Ψ, the KL divergence is defined as:

$$D(p \| q) = \sum_{x} p(x) \log_2 \frac{p(x)}{q(x)}. \quad (3.1)$$

The KLD has two shortcomings: 1) non symmetry, 2) it requires the two distributions to be continuous with respect to each other. Therefore, instead of Kullback-Leibler, the KL based Resistor-Average (R) divergence measure defined in

[65] is used. For accuracy evaluation of HM and BM, the R divergence of good- and bad-bursts distributions derived from actual traces and the models is compared. The R divergence observed in Figure 3-4, Figure 3-5, Figure 3-6, and Figure 3-7 at different states demonstrate that the HM shows results demonstrating very small R values for good bursts and outperforms the BM.

The results for R divergence of 802.15.4 HM traces and BM are demonstrated in Figure 3-4. It shows consistent performance with increasing order of K and renders good behavior with increasing state compression. Figure 3-5 shows the R divergence of bad-bursts for 802.15.4 traces generated from HM and BM. For bad bursts, it demonstrates very small R values elaborating similarity with actual traces when compared with BM and showing lower R divergence at $K = 5$ because of occurrence of unused states at $K = 6$.

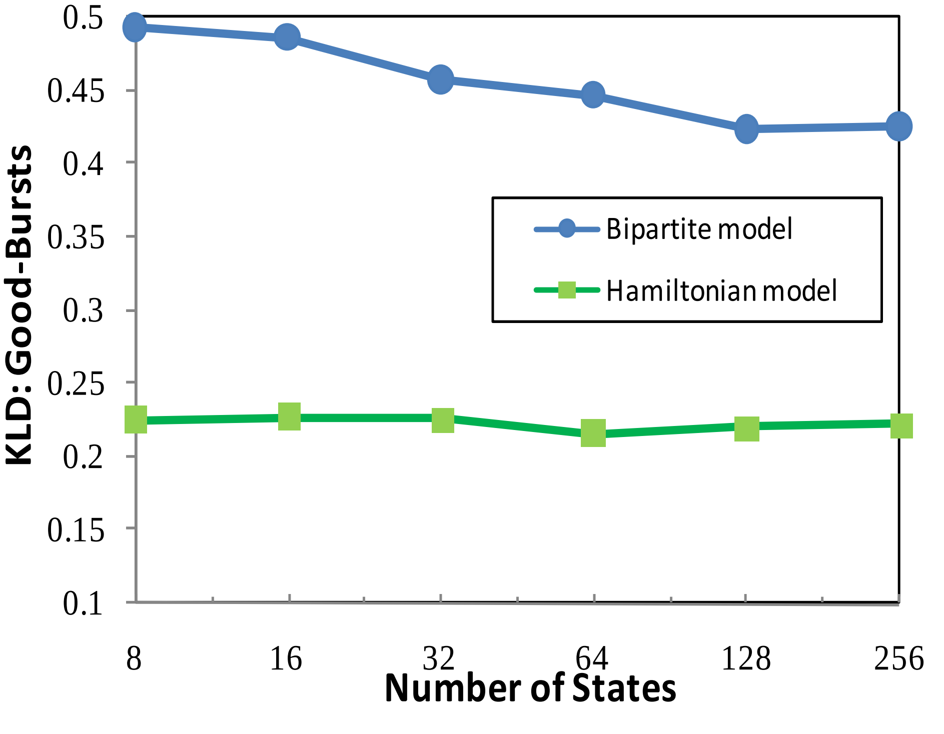


Figure 3-4: Resister average divergence of good bursts versus complexity for the 802.15.4 bit error process.

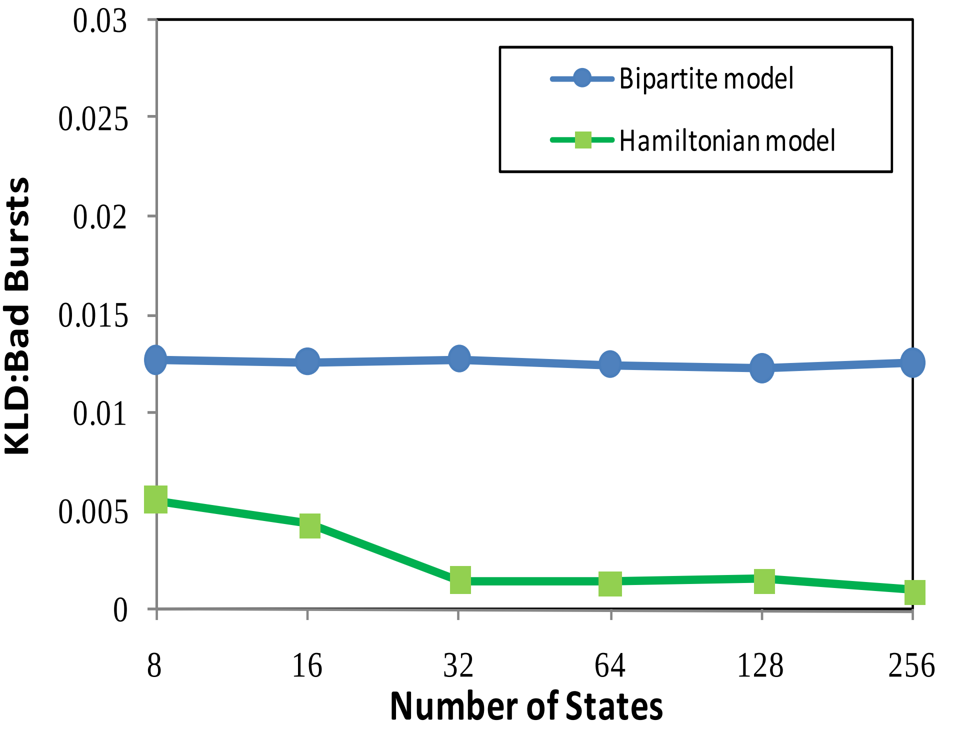


Figure 3-5: Resister average divergence of bad bursts versus complexity for the 802.15.4 bit error process.

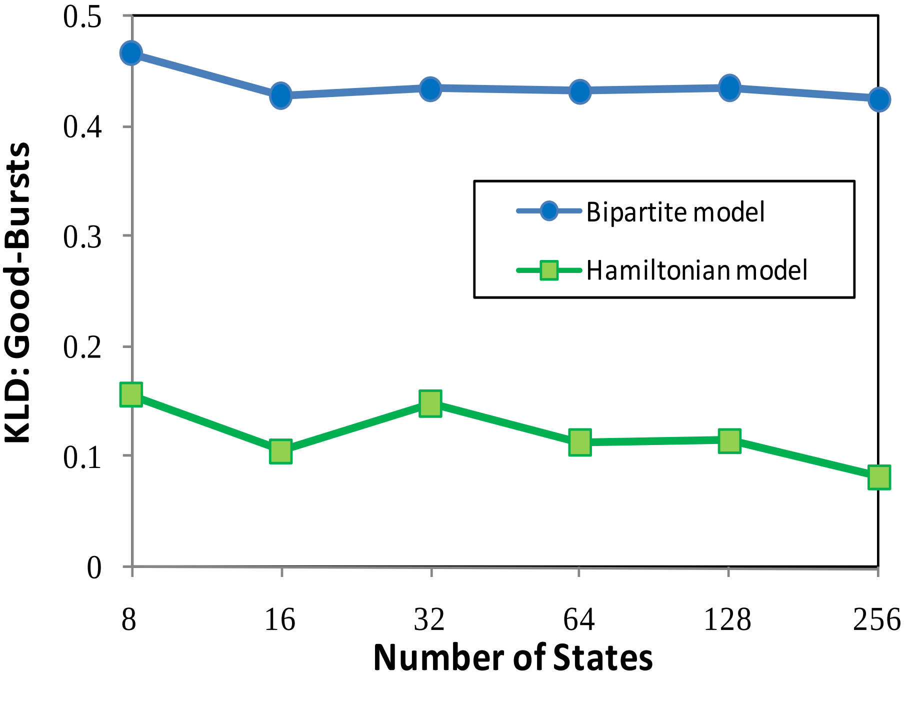


Figure 3-6: Resister average divergence of good bursts versus complexity for the 802.11b bit error process.

The increase in the order of Markov chain causes occurrence of more unused states and overall in our 802.15.4 WSN traces these states start at $K=6$ and which is also shown in Table 3-1. It can be clearly observed that at increasing order of K the HM has very small R divergence in case of bad bursts. On the other hand, the good-bursts and bad-bursts distributions of the BM diverge quite significantly from actual traces. R divergences of good and bad bursts for the 802.11b traces are shown in Figure 3-6 and Figure 3-7. It can be clearly observed that at increasing order of K the HM has very small R divergence in the good bursts case. On the other hand, the good-bursts distributions of the BM diverge quite significantly from actual traces. For bad-bursts, the R divergence of the HM is slightly higher than the BM at $K=9$. The slight variations in bad burst R divergence values can be removed by averaging over more traces. Nevertheless, both models are able to capture the bad-bursts behavior quite accurately.

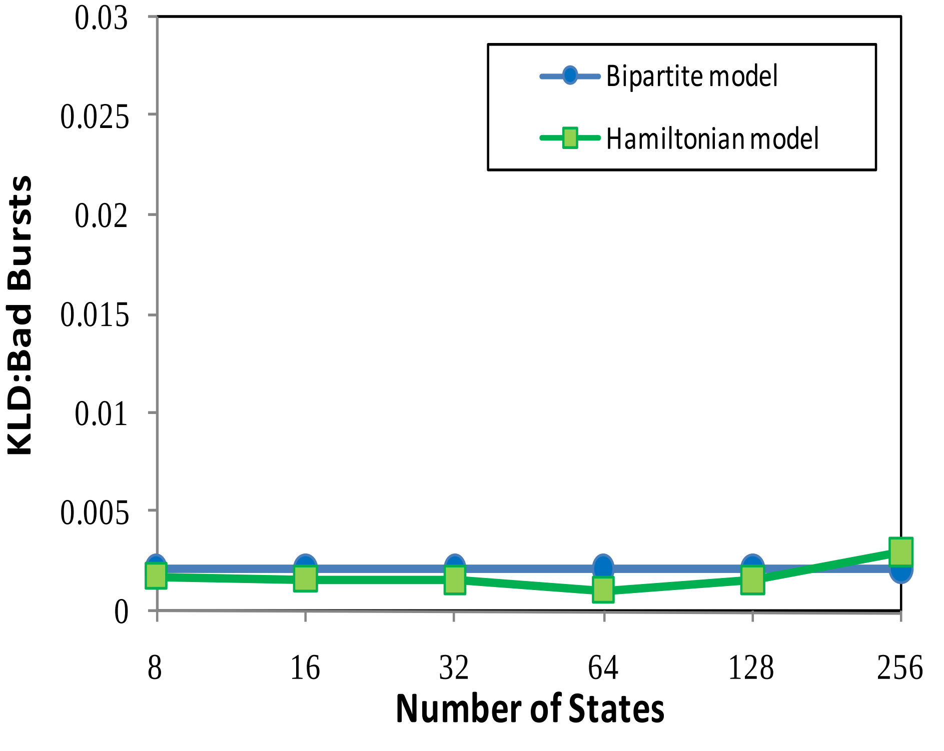


Figure 3-7: Resister average divergence of bad bursts versus complexity for the 802.11b bit error process.

3.7.3 Cumulative Distribution Function (CDF) of Good and Bad Bursts

Let p(x) and q(x) be two probability mass functions (PMFs) of a random variable x defined over an alphabet set Y. The random variable can be either the good bursts or the bad bursts random variable. The cumulative distribution function (CDF) of this random variable is defined as

$$F_X(t) = P\{X \leq t\} = \sum_{x \leq t} p_x(x). \qquad (3.2)$$

Let P(x) and Q(x) be the CDFs of p(x) and q(x) then, the standard error between P(x) and Q(x) is defined as

$$S(P,Q) = \sqrt{\frac{1}{l(l-2)}\left(l\sum_{Y}^{n} P^2 - \left(\sum_{Y} P\right)^2 - \frac{\left(l\sum PQ - \left(\sum_{Y} P\right)\left(\sum_{Y} Q\right)\right)^2}{l\sum_{Y} Q^2 - \left(\sum_{Y} Q\right)^2} \right)} \qquad (3.3)$$

Where l is the length of the CDF. Moreover, the standard error is used as a measure of error incurred by a model in approximating the actual CDF. The P(x) and Q(x) are represented as the CDFs derived from actual sensor network traces and traces artificially synthesized by HM and the BM [27]. Furthermore, the CDF for good and bad bursts of the exponential distribution with x varying for different burst lengths is evaluated. For evaluation, one trace per setup is taken from each of the model at different memory length and is compared with the original trace at a particular order.

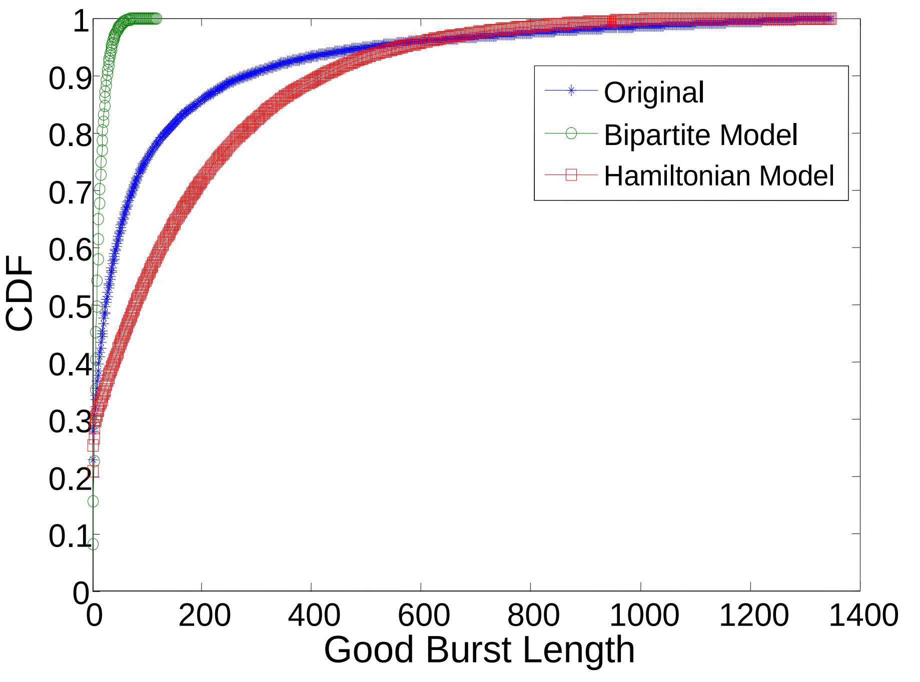


Figure 3-8: Good burst length distribution for 802.15.4 bit error traces at $K = 5$.

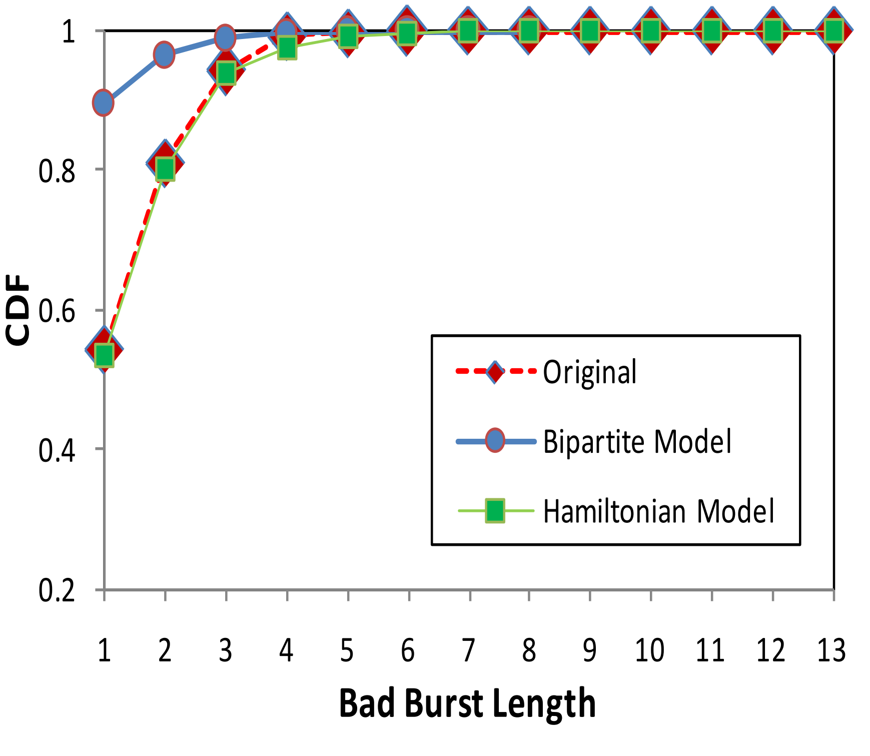


Figure 3-9: Bad burst length distribution for 802.15.4 bit error traces at $K = 5$.

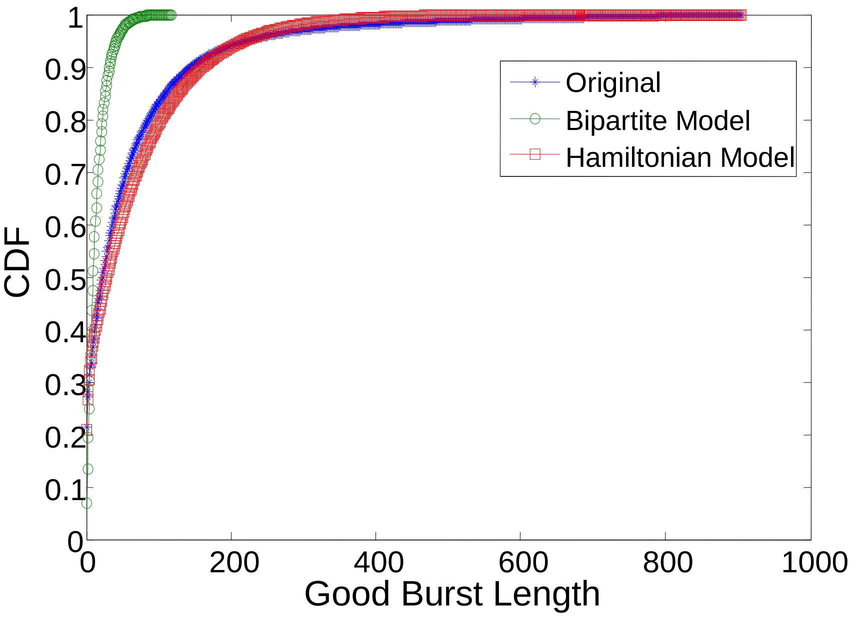


Figure 3-10: Good burst length distribution for 802.15.4 bit error traces at $K = 8$.

The cumulative distribution function for good and bad burst length at $K = 5$ and $K = 8$ is shown in Figure 3-8, Figure 3-9, Figure 3-10, and Figure 3-11 respectively. The CDF of the HM clearly follows the CDF of actual 802.15.4 traces for both the good and bad burst distributions. This further demonstrates that the standard error incurred by a model in approximating the actual CDF is very low.

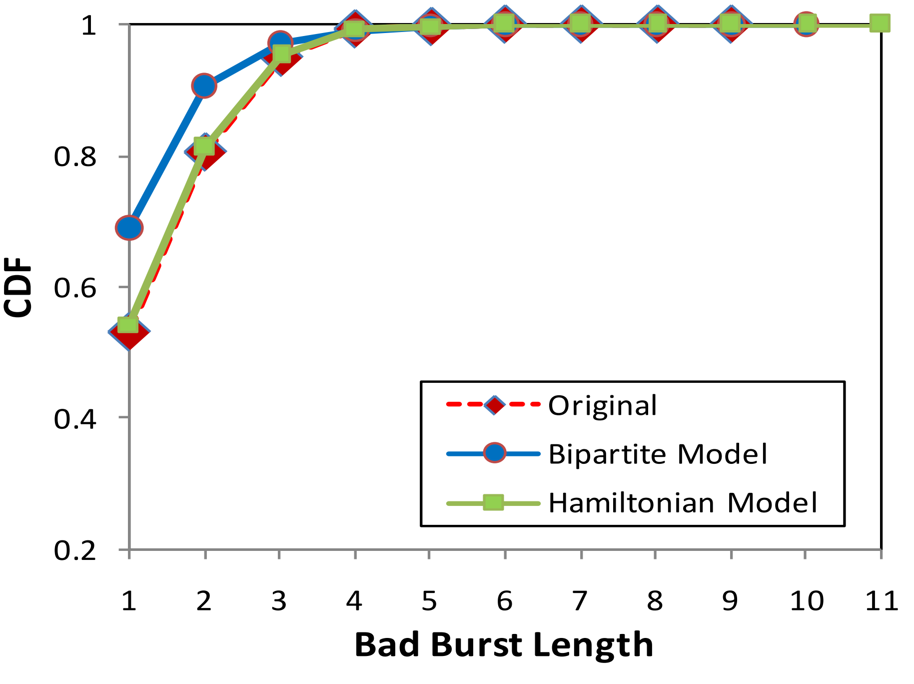


Figure 3-11: Bad burst length distribution for 802.15.4 bit error traces at $K = 8$.

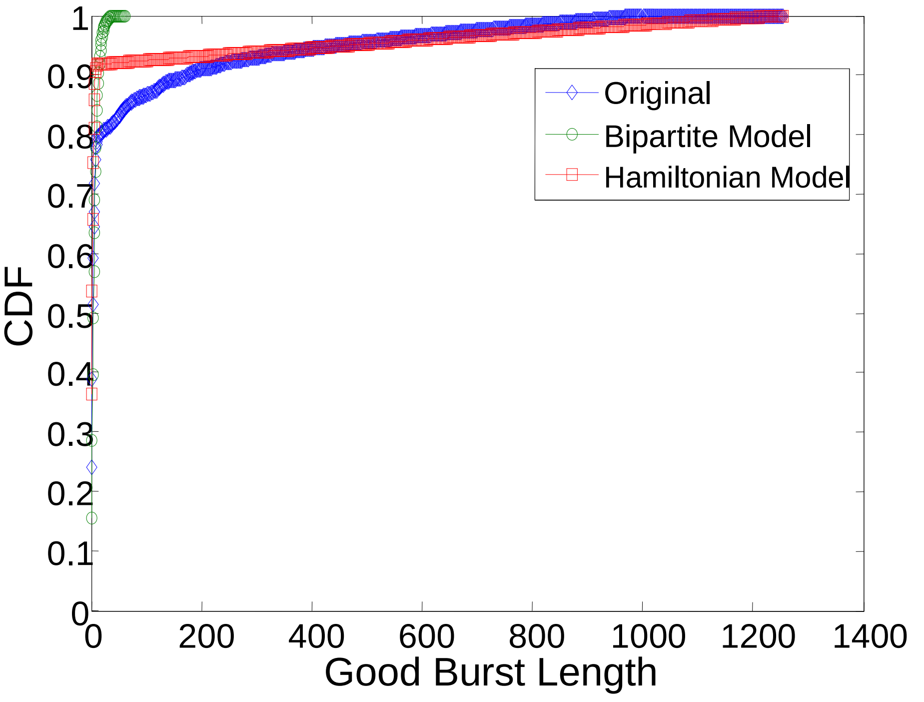


Figure 3-12: Good burst length distribution for 802.11b bit error traces at $K = 8$.

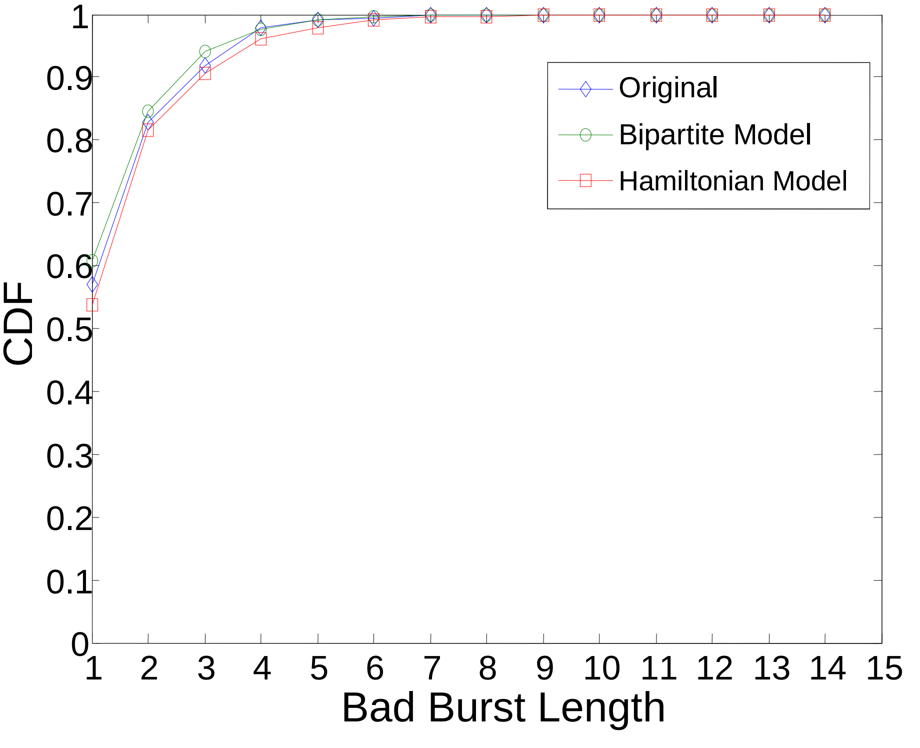


Figure 3-13: Bad burst length distribution for 802.11b bit error traces at $K = 8$.

The CDF for $K = 8$ is also plotted for good and bad burst lengths and is shown in Figure 3-12 and Figure 3-13. The CDF of the HM clearly follows the CDF of actual

802.11b bit error traces in both the good and bad burst length distributions, respectively. While the BM model captures the bad burst CDF reasonably accurately, the HM outperforms BM in capturing good bursts as shown in Figure 3-10 and Figure 3-12, respectively.

3.8 Summary

This chapter established the basic framework for the experiments conducted for presenting a wireless channel model. For data collection, two types of traces were used, the traces collected over 802.15.4 wireless sensor network and 802.11b wireless local area network. For 802.15.4, total of 20 traces were collected at different locations and for 802.11b, 3 traces were collected. After the insight into the collected wireless traces, initial bit error rate statistics were shown. The K^{th} order markov chain statistics give insight into the number of unused states that were never used in different order of markov chains. The number of unused states increases with the increase in the order which shows that the increase in the order of markov chains increases the number of unused states and markov chains only jumps into particular states.

In addition, a novel algorithm (Hamiltonian Model) is presented to reduce the complexity of high-order Markov channel models to a desired state budget. The state aggregation by identifying Hamiltonian circuits in the Markov digraph is performed and the circuit was then aggregated to a given and arbitrary state budget. Moreover, it has been demonstrated that the HM provides orders of magnitude reduction in complexity and renders very accurate performance when compared with BM. In addition, it is also shown that by using the unused state parameter, the model complexity can be reduced further and can provide results very similar to actual traces at varying number of K-states. When this parameter is not available, the model complexity renders behavior similar to actual traces at particular order of K-states. However, the HM model still performs better than the BM in capturing good and bad bursts.

Chapter 4

Topology Control: Background and Related Work

4.1 Introduction

Sensor networks have already been researched and deployed for decades. In recent years, their wireless extension has witnessed a tremendous upsurge in interest, applications, and protocol development [1]-[14], [67]-[74]. The ease of deploying wireless devices in unprecedented operating conditions and due to its vast variety of applications and its potential for generating large revenues have attracted many researchers and for this reason wireless sensor networks (WSN) are under keen consideration by academia and industry [75]. Pico-Radio [76] and Smart Dust [77] projects at UC Berkeley and WINS [78] project at UCLA were some initial projects that attempted to build wireless sensors.

The sensor nodes in a Wireless Sensor Networks (WSN) are resource-constrained in terms of computational, storage, communication and energy resources and capabilities. Moreover, lack of physical and the large-scale nature of most practical WSN deployment scenarios prohibit battery recharging. The two main sources of energy drain in a sensor mote are communication (bit transmission/reception) and local computation. Of all these constraints it is well known that communication consumes approximately two orders of magnitude higher energy than computation, storage and sensing for a majority of common applications [79], [80]. Therefore, communication overhead must be minimized for energy-efficient operation of a sensor network.

A WSN can be used in many application scenarios such as environmental and habitat monitoring, healthcare and surveillance, logistics and transportation, security and automation to monitor physical and environmental changes, such as pressure, temperature, vibration etc... Sensor node consists of a wireless transceiver, microcontroller and an energy source assembled as one unit of very small size. The size and energy source imposes unique constraints of energy, memory, connectivity, bandwidth and computational speed [81]-[88]. These special characteristics offers many research challenges in adaptation mechanism, packet size format, network wide reliability, energy constraints and scarce memory capacity, therefore highlight the need of a unique topology control protocol. Due to this reason, a viable option to reduce communication energy depletion is to use energy efficient network topology i.e. if the degrees of the nodes are controlled in such a way that a target property (e.g. strong connectivity) of the resulting network topology is guaranteed, while the network energy consumption is minimal.

The most famous way of forming a reduced subset of nodes is to form a hierarchal topology. In the hierarchal approach, reduced subsets of nodes are given more responsibility on behalf of the majority of nodes in the network. The reduced subset of nodes can be formed with well known connected dominating set (CDS) with nodes in a CDS forms a backbone which are responsible for relaying information for rest of the nodes in the network. Many Topology Control (TC) protocols have been proposed in the literature [35]-[44], [89], [90], [91], [92] with emphasis on either energy efficiency or network reliability (defined in terms of network connectivity.) While both of these metrics (energy efficiency and reliability) are equally important for mission-critical WSN deployments, existing CDS based topology construction protocols have not been designed to collectively cater for both metrics [56]-[58].

Before going into the proposed protocols details, it is necessary to understand the design guidelines, challenges and general techniques used for WSN topology control. Therefore, background required to understand the contribution of the report (related to topology control) is presented in this Chapter.

4.2 Definition of Topology Control

WSN's in a formal way are represented as a geometric random graph $G = (V, E, r)$, in which V is the set of vertices, E is the set of edges, and r is the radius of the transmission range of the nodes. Every vertex on V represents a wireless sensor device, and has a geometric coordinate associated to it and an open ball with radius r.

The communication links are established with nodes that are close enough for the radio signal to arrive with adequate signal strength. In other sense, links are established between nodes that are within a communication radius of each other. However, in order to improve energy efficiency, topology control process helps in reducing the communication links with other neighbors of the node in the network.

Topology control (TC) is an iterative process in which there is an initialization phase which is common to all WSN deployments. In the initialization phase, nodes make use of the discovery process by using maximum transmission power to build the initial topology. The initial network topology comprises of links and nodes that allow direct communication and every node communicates with a subset of the nodes according to the distance between them. In other words, TC is the reconstruction and organization of node parameters and modes of operation from time to time to modify the topology of the network. A TC protocol consists of two separate components: In the *topology construction* mechanism, nodes control their transmission/degree in order to achieve a desired topological property while preserving important network properties such as connectivity and coverage, and in the *topology maintenance* mechanism nodes further shift their role when they are unable to provide the requested service to increase the network lifetime.

The topology control protocols in the communications protocol stack normally exploit two layers during the topology control operation. The topology control protocols use the data link layer and network layer. The data link layer allows nodes to find their neighbor and establish the corresponding links. It also helps building the reduced topology from the original topology according to the topology construction protocol. Once the reduced topology is build, the network starts the route discovery process in order to route packets in the reduced topology. Therefore, these layers are called many times as part of the topology maintenance procedure.

4.3 Challenges and Design issues in Topology Control

The main task of the topology control is to build a reduced topology in such a manner that network wide connectivity and coverage is maintained in the whole network. Moreover, the algorithms and protocols designed to perform the topology construction and maintenance operation should be energy efficient as they may end up consuming equal or more energy than the energy saved as a result of the reduced topology. In addition to the above, most of the researchers only consider the topology construction phase. However, to extend the network lifetime it is always desirable to change the reduced topology in order for the active nodes to also save their energy resources. Therefore, the overhead and computations of the topology maintenance phase should also be included as part of the performance evaluation of topology control protocol. Hence, following design guidelines should always be kept in mind for making an energy efficient topology control protocol. Some of the design guidelines are as under:

- **Distributed protocol:** The Topology control protocol must discover the reduced topology in a distributed fashion as centralized protocols need complete information of the network and therefore are very expensive to be implemented in practice. The nodes in the network must also use the local information which they gather from their neighbors to discover the topology. This reduces the energy consumption and makes the protocol scalable. The node must also not use any extra hardware or support mechanisms which add to the cost and control overhead. One example is the need of location information, which might be provided by global positioning system (GPS) devices or localization protocols which are difficult to realize in practical WSN deployments.
- **Simplicity and Message Complexity:** To achieve energy efficiency, the topology control protocol must use as few messages as possible to construct the reduced topology. The message complexity should also be reduced while maintaining the topology. Moreover, the protocol should also be simple in operation in order for it to work with different topology maintenance mechanisms.
- **Connectivity and Coverage:** The reduced topology must be connected in such a manner that active nodes can exchange information among themselves and also with the sink node. The reduced topology must also cover the area of interest.
- **Node degree:** The small node degree is desirable in topology control protocols which allow nodes to save more energy as small node degree means a small number of neighbors, which at the same time reduces into a

lower number of collisions and retransmissions. This factor also allows mitigating the hidden and exposed terminal problems.

- **Spanner:** The reduced topology must be a spanner of the Unit Disk Graph in terms of both length and the number of hops. A sub-graph G' is a spanner of a graph G for length (number of hops) if there is a positive real constant X such that for any two nodes $U, V,$ the length (number of hops) of the shortest path in G' is at most X times of the length (number of hops) of the shortest path in G. The constant X is called the length (number of hops) stretch factor.

$$\text{Hop Stretch factor} = \underbrace{\max}_{u,v\in V} \frac{(u,v)_{G'}}{(u,v)_{G}} \quad \textbf{(4.4)}$$

Hence, the hop stretch factor is defined as the worst increase in path length for any pair of nodes $U, V,$ between the original graph G and the topology controlled path G'. The topology control protocols with small stretch factor are desirable.

In the next section, some prominent topology control approaches are explained in detail.

4.4 Topology Construction Approaches for Wireless Sensor Networks

Topology Control protocols have been proposed by many researchers based on different techniques. When a flat network topology (all nodes are considered equal) is desired, the set of neighbors of a node can be reduced by simply not communicating with some neighbors. There are several possible approaches to choose neighbors, but the one that is always promising and simple for WSN is to limit the reach of node's transmission. However, in this report, the main interest is focused towards relevant hierarchal graph-theoretic approaches which are used in topology construction. In the hierarchal topology construction approach, a communication hierarchy is created in which a reduced subset of the nodes is selected and given more responsibilities on behalf of a simplified and reduced functionality for the majority of the nodes. One disadvantage of the hierarchal approach is that the selected subset of nodes will work more than their unselected neighbors, and will consume their batteries. Hence, they must be used with efficient topology maintenance mechanism. The general classification of the topology control protocol is shown in Figure 4-1. The hierarchal backbone construction using dominating sets is explained in the subsequent section. The maintenance mechanisms and the protocols which construct CDS are also discussed in detail.

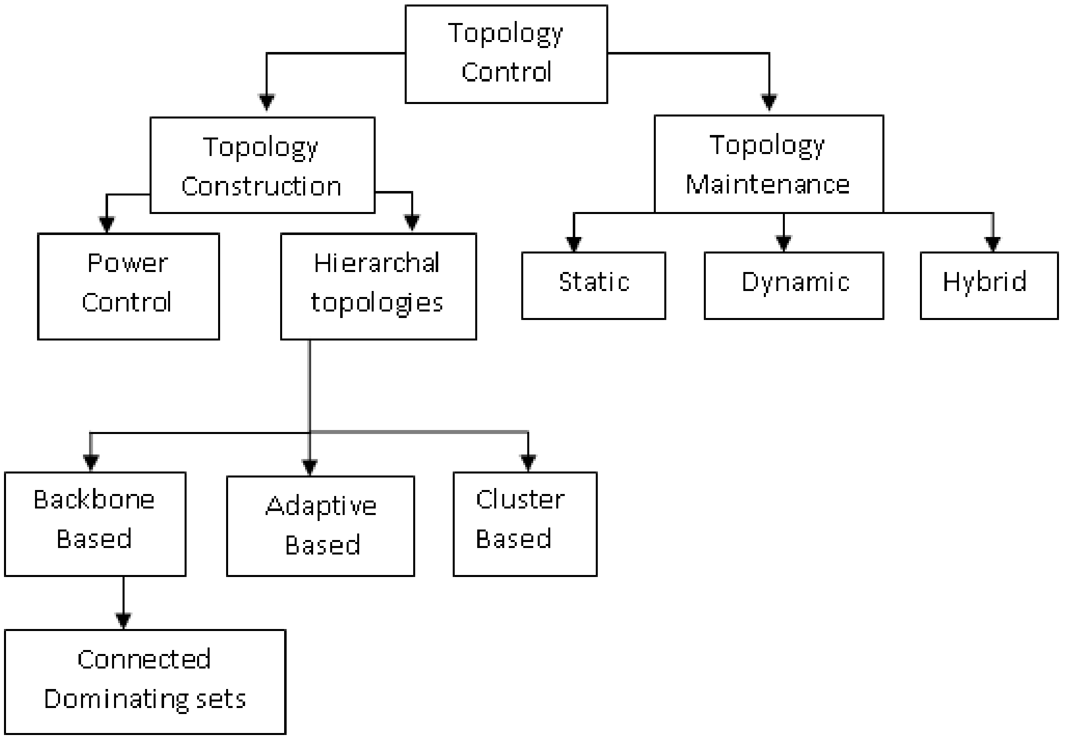


Figure 4-1: Classification of Topology Control.

4.5 Connected Dominating Set Based Topology Construction

The main task of the backbone based approach is to find a subset of nodes which guarantee connectivity and coverage while allowing nodes to communicate with the backbone nodes in a direct way. A communication backbone can be created by solving a widely known mathematical problem: the Connected Dominating Set (CDS) problem. A Dominating Set (DS) is a set of nodes $D \subset V$ in a graph, in which all other nodes that do not belong to the subset have a link (connection) to at least one node in the set. A topology control algorithm based on dominating sets is proposed in [89]. Dominating sets for a graph $G = (V, E)$ consists of set of vertices S whose neighbors along with themselves consist of all the vertices in the graph. It is also shown in Figure 4-2 in which nodes are marked with black color vertices. In CDS, these dominating nodes are connected with each other through a common neighbor. The topology construction based on dominating sets is shown in Figure 4-2.

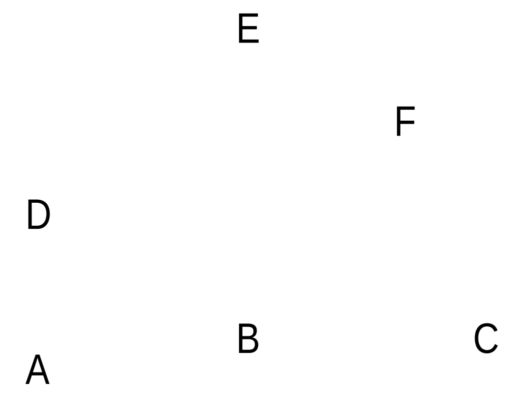


Figure 4-2: Dominating set based topology construction.

A CDS can be built by growing a tree and can be illustrated by comparing it with Prim's algorithm. The Prim's algorithm finds the minimum spanning tree of a graph. The construction process starts from the set of nodes which are part of the tree at time t. As the time passes, the nodes in the set evaluate all adjacent nodes in order to extend the tree. This process continues until all nodes on the graph are traversed and some nodes declare them as active while other enters into dormant mode in the tree. This procedure guarantees connectivity as the new selected nodes are always neighbors from at least one node of the tree set. However, every edge serves as a bridge edge in the tree which on the other hand does not provide reliability. The topology construction based on connected dominating sets is shown in Figure 4-3.

The CDS construction starts at a single node, the sink node, but in case of multiple sink nodes, several trees may be built in parallel. The CDS construction technique has characteristics that are desirable for routing protocols as it forms an organized structure which allows implementing routing based on hierarchal addressing. In the next section, prominent topology construction protocols are explained which are later utilized in the performance evaluation of the proposed protocols.

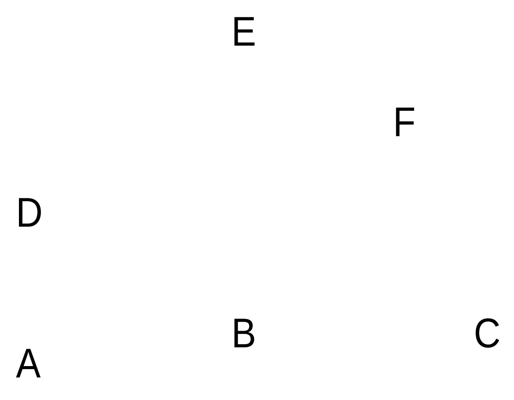


Figure 4-3: Connected Dominating Set (CDS) based topology construction.

4.6 Topology Construction Protocols

In this section, three existing topology construction protocols -- CDS-Rule K, A3 and EECDS -- used for the performance evaluation with proposed protocols are explained.

Historically, energy conserving topologies for WSN's were constructed by adjusting the transmission powers of sensor nodes [92], [94]. Some other techniques preserve energy using the nodes' geographical location information [95]. However, power control and location awareness are expensive propositions (in terms of energy consumption, footprint and cost) and hence are difficult to realize in practical WSN deployments.

As an alternative to power/location awareness, a CDS-based graph-theoretic solution was proposed in [89], in which a vertex dominating it-self and all the adjacent vertices form a cluster in the graph. Energy is then conserved by routing data through these cluster-heads. A similar solution was proposed in [90] which use independent dominating sets with lower IDs node acting as the cluster heads. Since these seminal papers, CDS based WSN clustering has received significant attention and is now widely considered as the predominant method for energy efficient topology construction in WSN's.

In the next subsection, focus remains on the three protocols which are used for the performance evaluation in this report. Other protocols are derived from these three base protocols.

4.6.1 The CDS-Rule K Protocol

In [54], the authors proposed a CDS-Rule K protocol by utilizing the marking and pruning rules. In the marking rule, a node is temporarily marked if it has two neighbors that are not directly connected. The authors further demonstrate that the temporarily marked node set is a CDS. However, after a set of temporarily marked nodes is derived by applying the marking process, it can be further reduced via the pruning rule. In the pruning rule, a temporarily marked node u can be unmarked if all its 1-hop neighbors $(N_1(u))$ are also neighbors of any one of K coverage nodes that are connected and have higher priorities. This means that when each node in $(N_1(u))$ is a neighbor of node set C, u is said to be covered by C. Due to the marking and pruning process, CDS-Rule K constructs the topology in two phases.

In the CDS-Rule K protocol, the construction process starts with a big set of nodes which allow nodes to exchange their neighbor's lists. A node remains marked if there is at least one pair of unconnected neighbors and un-marks itself if it determines that all of its neighbors are covered with higher priority. This is indicated by a level of the node in the tree. The final tree is a pruned version of the initial one with all redundant nodes with higher or equal priority removed.

4.6.2 The EECDS Protocol

In an undirected graph, a Maximal Independent Set (MIS) is also a Dominating Set (DS). Most of the distributed algorithms find's an MIS and connect this set to form a CDS. The authors in [92], [110], and [111] first proposed distributed algorithms for constructing CDS's in unit disk graphs (UDG's), which consists of two phases to form the CDS. They form a spanning tree and then utilize nodes in the tree to find an MIS. At start, all the nodes in an MIS are colored black. In the second phase, more nodes are added which have a blue color to connect the black nodes to form a CDS. Later, the authors in [55] proposed an Energy-Efficient CDS (EECDS) protocol that computes a sub-optimal CDS in an arbitrary connected graph. They also use two phase strategy to form a CDS. The EECDS also uses a coloring approach to build the MIS. The EECDS algorithm begins with all nodes being white. An initiator node elects itself as part of the MIS coloring itself black and sending a Black message to announce its neighbors that it is part of the MIS. Upon receiving this message, each white neighbor colors itself as gray and sends a Gray message to notify its own White neighbors that it has been converted to gray. Therefore, all white nodes receiving a Gray message are neighbors of a node that does not belong to the MIS. These nodes need to compete to become Black nodes. For this, a node sends an

Inquiry message to its neighbors to know about their state. If it does not receive any Black message in response and it has the highest weight, it becomes a Black node, and the process starts again. In EECDS, the second part of the protocol is to form a CDS using nodes that do not belong to the MIS. These nodes, called connectors, are selected in a greedy manner by MIS nodes using three types of messages namely Blue, Update, and Invite messages.

4.6.3 The A3 Protocol

Mostly, all the approximation ratios mentioned above are concerned with CDS size. While for efficient routing, not only a smaller CDS size is desired, but extra requirements like energy efficiency also needs to be considered. For this purpose, the authors of [53] have proposed a topology construction protocol that produces an approximate solution to form a sub-optimal CDS. The A3 protocol used four types of messages for topology discovery. The CDS building process is started by a pre-defined node that might be the sink, right after the nodes are deployed. The sink initiates the topology construction process by sending a hello message in its communication region. As the nodes receive a hello message from the parent which is also a sink node, they send back the information regarding remaining energy and signal strength in parent recognition message while calculating it with the help of selection metric. The selection metric sets low value of timeout or gives priority to those nodes with higher energy and which are farther away from the parent node. This helps in constructing the reduced topology with fewer nodes.

The parent node on receiving the parent recognition message sends back the sorted list to all its children. This sorted list contains the timeout information of all the children belonging to the parent. As the timeout expires, nodes further explore their neighbors by repeating the same steps to form the reduced topology. Nodes not receiving any message in reply to their hello message enter into sleep mode by broadcasting a sleeping message.

The protocols described in previous subsections build the reduced topology but do not maintain it. Consequently, the optimal topology depletes the battery of CDS nodes rapidly which directly reduces the network lifetime. There are other protocols [80], [99]-[102], which improve the overall lifetime of the network through different topology maintenance techniques. Therefore, in the next section, prominent topology maintenenance techniques are presented which are later used in the maintenance phases of the proposed topology control protocols.

4.7 Topology Maintenance and Performance Metrics

Topology maintenance techniques are classified as static, dynamic and hybrid techniques. Moreover, they are further classified as global or local. The global techniques consider all the nodes in the network in order to take a global-optimal decision, while local techniques only consider a small subset of the nodes in the network. In other words, global techniques switch the entire network topology, while local techniques switch only a portion of the network, such as a branch of a CDS tree, or a cluster. Moreover, in any of the above topology maintenance technique, the constructed topology is rotated on the basis of time and energy triggering mechanisms.

In time triggered topology maintenance techniques, pre-constructed topologies are rotated based on the timeout criteria. The amount of time is usually fixed and pre-defined, and is a very critical variable. Therefore, a very short time may cause unwanted extra overhead as a consequence of switching the topology more often than necessary. Similarly, a very long time may use a suboptimal network longer than necessary, with the possibility of losing important properties such as connectivity and poor coverage.

In this chapter, the prime focus remains only on global maintenance of the protocols based on energy thresholds i.e. that is the topology is rotated when the network nodes energy falls below a certain energy threshold. In the next subsection, topology maintenance techniques are presented which are subsequently used in the evaluation of the CDS-based protocols.

4.7.1 Static Topology Maintenance

Static topology maintenance techniques calculate all the possible set of CDS trees during the initial topology construction time and then rotate the possible set of CDS trees. However, the set of CDS trees can be restricted based on the nodes density. Static techniques can be triggered based on energy and time thresholds which allow the rotation between the a priori constructed CDS trees. If the network is sparse, static techniques do not build completely disjoint CDS trees. Static techniques make use of notification and reset messages sent to a sink node which then selects a new CDS tree.

The performance of static techniques mainly depends on the topology construction protocol. If a topology construction protocol is energy efficient then it is more likely that it will perform better with static maintenance techniques.

4.7.2 Dynamic Topology Maintenance

Dynamic topology maintenance methods do not make a priori calculations to determine the next possible set of nodes that form a CDS tree after the current set of nodes is no longer optimal, e.g. when the energy threshold is reached. Dynamic techniques are generally more time consuming because they switch the topology based on the actual status of the CDS tree. On the other hand, they are better in terms of selecting the most capable set of nodes in the network. Dynamic techniques can also be time/energy triggered. An example of dynamic topology maintenance can be found in [102], where network topology is changed in rounds based on a time-based triggering criterion.

In the case of dynamic global techniques, the sink broadcast a reset message to other nodes in the network and informs that the new topology is now being rotated. Similarly, a “notification message” is used by the nodes in the network to notify the sink node when the nodes energies are running out. In this manner, the sink node gets aware and sends the reset message to initiate the recreation of the topology.

4.7.3 Hybrid Topology Maintenance

In a hybrid mechanism, a set of potential CDS sets are pre-constructed and rotated based on time and energy thresholds. If existing CDS sets degrade performance, a new topology is built by invoking a dynamic technique. This process allows new CDS set in the possible combination that was built at the start using static maintenance. Hybrid topology maintenance works best with energy-based triggering because time-based methods invoke CDS tree updates too frequently, thereby resulting in large message overhead.

4.7.4 Performance Evaluation Metrics

In this section, we now provide formal definitions of the key concepts/metrics used in the evaluation process of the protocols.

The major function of a topology construction protocol is to build a routing backbone so that the data can be collected from each individual node. The function needs to be carried out under a set of constraints posed by the ad hoc nature of wireless networks and resource-constrained sensor nodes. Therefore, topology must be built with minimum of the signaling overhead in order to keep the network operational for an extended period of time. In other words, higher message overhead consumes higher energy and, in general, also needs significant processing overhead. Therefore, any protocol designed for WSN’s must try to minimize this metric. Secondly, maximum network connectivity under topology maintenance is also desired. Keeping in view of all these considerations, set of metrics were carefully selected such that they cover almost every aspect of topology control protocols [53], [54], [55], [109]. The definition of performance metrics are given below.

- **Message overhead**: Message overhead is defined as the total number of sent and received packets in the whole network during construction of the topology.
- **Energy overhead**: Energy overhead is defined as the fraction (or percentage) of the total network energy consumed during one run of an experiment.
- **Residual energy**: Residual energy is defined as the remaining energy in the active set of nodes at the end of an experiment.
- **Convergence time**: Convergence time is defined as the time taken by a protocol to construct the topology until the finishing criteria.
- **Unconnected nodes**: Unconnected nodes are defined as the number of nodes which are disconnected from the sink/initiator node at the end of topology maintenance operation.
- **Average backbone path length (L):** is defined as the number of edges along the shortest paths for all possible pairs of network nodes. It is given by

$$L = 1/n(n-1).\left[\sum_{i,j} d(v_i, v_j)\right] \qquad \textbf{(4.5)}$$

  where n is the number of vertices/nodes in the network and d is the distance between nodes i and j for all pair of active nodes in the network.
- **Connected sensing area:** is defined as the area covered by the connected nodes at the end of topology maintenance operation.

Most of the studies (explained in previous section) consider topology construction as the major process thereby ignoring the importance of topology maintenance.

Our choice of parameters considers both procedures as integral parts of a topology control protocol. Our choice of message overhead is an extremely important metric as it directly affects the energy consumed in the network. Many authors only consider the number of sent messages as the message overhead. However, we believe that message reception is also critical and, therefore, our definition of message overhead is set accordingly. Similarly, residual energy is a measure of network lifetime. As the residual energy falls below a certain threshold value, the probability of network partitioning increases.

Under topology maintenance, it is important to consider the protocols performance in terms of network connectivity. To analyze this, unconnected nodes parameter is selected to elaborate the performance of the protocols. This parameter also measures the effectiveness of a topology construction protocol. Similarly, convergence time is also an important measure because new topologies are frequently constructed. If a protocol has higher convergence time, it can affect the overall network performance. Finally, covered sensing area at the end of topology maintenance operation is also another important metric. This metric allows judging the capability of a protocol in terms of connected nodes covering the area. A protocol is better if it covers more area. Therefore, any protocol designed for WSNs must try to maximize this metric. In the end, an average backbone path length differentiates an easily negotiable network from one which is complicated and inefficient, with a shorter one stated being more desirable in many studies.

4.8 Summary

In this chapter, prominent features and definitions of topology control are presented with challenges which must be addressed in a topology control protocol. The relevant and pertinent graph theoretic approaches for topology construction in wireless sensor networks were explained. These approaches are further used in the evaluation process of the proposed protocols. Moreover, vulnerabilities associated with different studies are also highlighted. It has also been shown that the current techniques do not cater for the unique WSN requirements and the issues and problems related with the protocols were discussed. In the end, definitions of the performance evaluation metrics were explained on which the protocols are evaluated later in the next chapters.

Chapter 5

The Clique Connected Dominating Set Topology Control Protocol

5.1 Introduction

Due to their inherent energy, cost and footprint constraints, Wireless Sensor Networks (WSNs) generally have limited computation, storage and communication resources. Moreover, lack of physical access and the anticipated large-scale deployments of sensor nodes prohibit frequent battery recharging/ replacement. Communication (bit transmissions and receptions) and local computations are the two main sources of energy consumption in a sensor mote. Since it is well-known that energy consumed during communication is much higher than that consumed in other computational tasks, therefore communication overhead must be minimized for energy efficient operation of a sensor network. Due to this fact, topology control is used to reduce WSN's communication energy by deploying sensor nodes in such a way that a target property (e.g., strong connectivity) of the resulting network is guaranteed with low energy expenditure.

Most of the power and location-unaware WSN topology control protocols rely on the Connecting Dominating Set (CDS) principle [89], [90], [91], [92], and [93]. In CDS-based routing, a set of rich connectivity nodes -- called the CDS -- acts as a virtual backbone for relaying packets in the network. Non-CDS nodes then conserve energy by turning off their transceivers. Clearly, the size of a CDS presents a critical tradeoff. A small CDS size allows high energy efficiency as more nodes can be put to sleep. On the other hand, a small-sized CDS compromises the network reliability (ability of a network to remain functional for an extended period of time) because few nodes handle the bulk of the network traffic and consequently deplete their batteries quickly.

In this chapter, a distributed protocol that discovers a CDS by finding cliques in a WSN is presented. The protocol, referred as Clique-based Connected Dominating Set (CCDS), models the network as a connected graph and finds the number of 2-cliques present in the network; a 2-clique refers to a vertex set between pair-wise adjacent vertices. The CCDS considers each node to be a clique of size 1 and then merges 1- cliques into 2-cliques until no more merges are possible. Based on the energy threshold of different nodes, CCDS adaptively finds the backbone nodes by forming an active clique set thereby allowing other nodes to turn off their transceivers while keeping the network connected and covered. CCDS attains better energy efficiency as CCDS discovery/maintenance protocol exploits the broadcast nature of the wireless transmissions to reduce the message complexity.

The CCDS is compared through simulations with exiting CDS-Rule K [54], Energy Efficient CDS (EECDS) [55] and A3 [53] protocols. Simulation results show that CCDS protocol has a low and linearly-bounded message complexity which allows the CCDS to run several times as a part of the topology construction procedure. The simulations are performed to demonstrate that CCDS performs considerably and consistently better than EECDS, A3 and CDS-Rule K protocols in most operational scenarios. Moreover, it is also shown that CCDS is a scalable protocol that does not require any location information and synchronization mechanism.

The CCDS is compared with other topology construction protocols on the basis of five relevant metrics: message overhead, residual energy, energy overhead, number of unconnected nodes and convergence time. Our performance evaluation reveals several interesting insights. For instance, some of the protocols provide reasonable performance during topology construction, but fail to maintain the same performance in the topology maintenance phase. Moreover, it was also observed that the performances of these protocols differ considerably in static and dynamic scenarios. Therefore, lack of consistent performance metrics can be very misleading because a protocol performing well on one metric might incur serious performance degradations for another metric. Finally, it was also observed that protocols using link/physical layer side-information (e.g., use of signal strength in A3) lead to a non-uniform distribution of energy resources during the topology maintenance phase.

The rest of this chapter is organized as follows. At first, the performance guidelines for CDS based topology construction protocols are explained. Later, the working of the CCDS protocol followed by the performance evaluation of the protocol with prominent topology construction protocols is explained. In addition, the relation between the guidelines and the CCDS is also given. In the end, the salient findings of the CCDS protocol are summarized.

5.2 Performance Guidelines for CDS-Based Topology Construction Protocols

In this section, interesting insights gained while implementing and evaluating the three topology construction protocols are explained. For the topology construction protocol to perform well, these design guidelines should be followed by a CDS-based topology control protocol.

Guideline 1: CDS must be formed with a large set of nodes–preferably proportional to the network size–in order to extend the network lifetime.

It is generally assumed that small CDS size is beneficial to increase the network lifetime since it allows nodes to have more residual energy. However, the simulation results show that even though EECDS protocol does not form a small CDS, it still provides better network lifetime. On the other hand, A3 protocol uses a selection metric to form a small backbone but fails to consume energy resources evenly. This observation reveals the fact that if the backbone nodes are in proportion to the network size, it distributes the energy resources uniformly in the network.

Guideline 2: Instead of relying on the connectivity properties of nodes, it is important that nodes remaining energy level is taken into consideration during the formation of CDS.

In the EECDS protocol, nodes form a MIS when a new topology is constructed. Selection of nodes is based only on battery power and its effective degree. Consequently, EECDS forms a CDS comprising of high energy dominating set nodes. On the other hand, A3 protocol makes use of selection metric based on the node’s distance from the parent node to form a CDS directly. The distance metric is used to optimize coverage and to provide better connectivity. However, the distance metric does not allow high energy nodes to become part of the CDS in each case. Therefore, the residual energy is worse than the EECDS. Hence, it is inferred that node’s selection criteria based only on energy is a good metric.

Guideline 3: Network connectivity can be improved by choosing diverse nodes to form the CDS.

During topology maintenance, network connectivity and residual energy can be improved if the CDS is formed from diverse selection of nodes i.e. nodes chosen from larger set without any strict selection metric criterion. It is inferred from the guideline by noticing that CDS-Rule K protocol forms a new CDS by selecting the nodes which were part of the previous constructed topology. This causes residual energy of nodes to fall sharply.

Guideline 4: Dynamic topology maintenance must be considered during the design of a topology control protocol.

This design guideline was observed during the inconsistent performance of protocols under different topology maintenance techniques. The static maintenance techniques calculate the overhead of pre-constructed topologies which, in most cases, do not represent a realistic scenario as the backbone nodes chosen at the start can behave differently as the time progresses. The energy consumption pattern and nodes connectivity properties change consequently. Therefore, dynamic topology maintenance technique must be used in the topology control process.

Guideline 5: Message overhead is a critical parameter which must be reduced by an efficient topology control protocol.

The CDS-Rule K and the EECDS protocol construct the topology in two steps leading to higher message overhead. On the other hand, A3 protocol reduces the message overhead by constructing the CDS in a single phase leading to low energy overhead. However, A3 does not provide high residual energy resources since it selects nodes using a selection metric which causes non-uniform distribution of energy (see Guideline 3). To elaborate the significance of afore-mentioned guidelines, in the following section, a new CDS-based topology construction protocol is presented by utilizing the design guidelines. The use of the guidelines allows the protocol to outperform existing protocols under different evaluation metrics.

5.3 The Clique-Based CDS Discovery (CCDS)

In this section, a Clique-based CDS Discovery (CCDS) protocol which is inspired from the performance guidelines learnt from the existing protocols is explained. The pruning process used by CDS-Rule K increases with increase in the number of neighbors thereby increases the number of exchanged messages. A similar problem exists in EECDS due to its two phase strategy to form a CDS. The Clique-based CDS Discovery (CCDS) protocol addresses the shortcomings of existing solutions by exploiting the inherent broadcast nature of wireless medium. Use of broadcast messages reduces the need of multiple explicit messages that need to be transmitted to form a CDS. Consequently, message complexity of the protocol is reduced significantly making it more energy efficient.

5.3.1 CCDS Protocol Description

Cliques comprise parts of a graph in which all nodes are connected with each other. A simple arrangement for a node is to form a clique with its one hop neighbors by message broadcast which reduces the number of messages. A CCDS assumes no prior knowledge about the position or orientation of the nodes. However, nodes are aware of other nodes ID's contained in the received network messages. This information is used to select a clique on first-come-first-serve basis; i.e., nodes in the selected clique receive and process messages in the order of delivery. The CCDS protocol is executed by selecting a node called an initiator node to be a clique of size 1. As the nodes get aware of the total number of nodes in the network, the initiator node then covers the clique of size 2, as discussed below, which is ultimately transformed into an active clique set hence-forth referred to as the backbone nodes or CDS.

The CCDS protocol starts with an initiator node which can, for instance, be the sink node. CDS discovery then propagates by message rebroadcasting which is illustrated by the example shown in Figure 5-1. In this example, Node A acts as an initiator node and broadcasts a clique discovery message to all its neighbors to announce itself as a clique node. The message will be received by node B and node D located in the transmission range of Node A (see Figure 5-2).

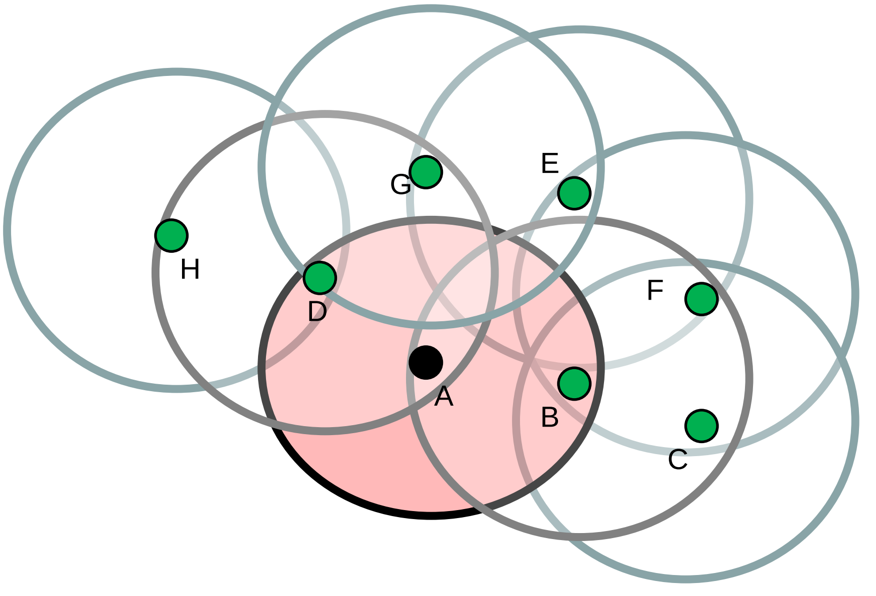


Figure 5-1: Sink node initiates a clique discovery message.

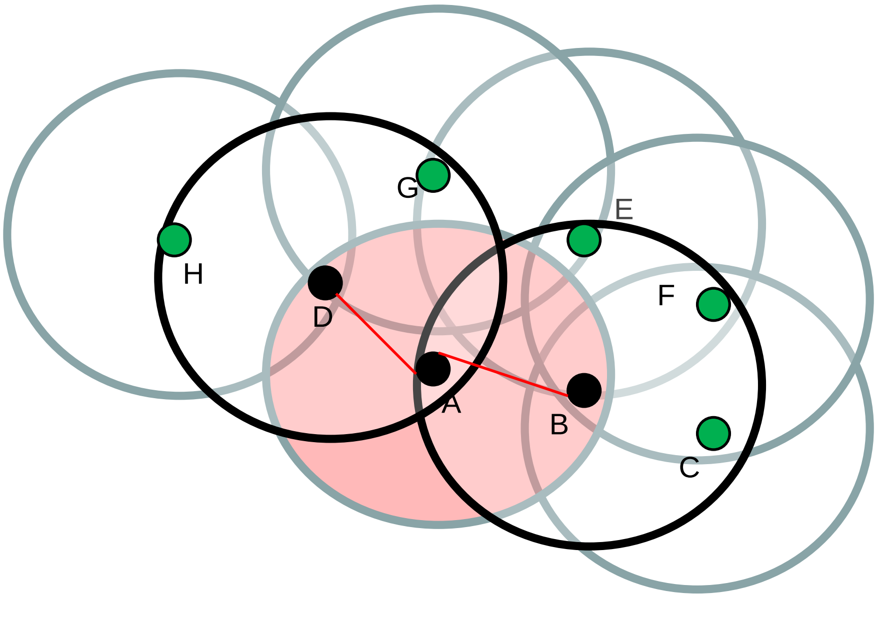


Figure 5-2: Neighboring nodes receive clique discovery message and rebroadcast it with their node IDs which are received by A.

The nodes on reception of the discovery message append their IDs to the message and broadcast it further. Note that neighbors of node A do not send any explicit response to node A. Instead, they simply repeat the broadcast of clique discovery message which allows node A to be aware of its neighbors. In this way, CCDS takes advantage of the broadcast mechanism for reducing the total number of messages exchanged during CDS discovery.

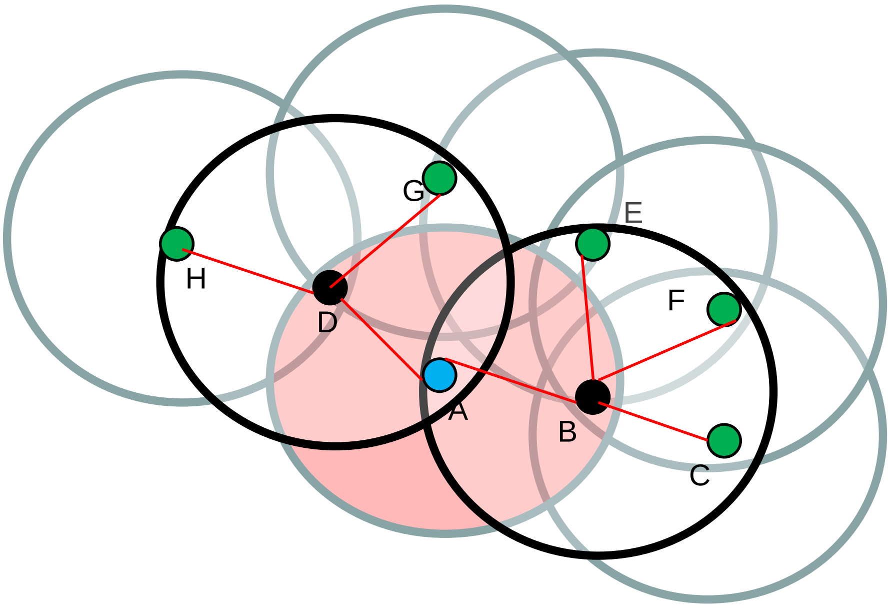


Figure 5-3: A recognizes B and D as a clique whereas uncovered nodes receive broadcasted clique discovery message from covered nodes.

The broadcasts of nodes B and D are received by their neighbors enabling node A to recognize the clique of size 2 with node B and node D as shown in Figure 5-3. Neighbors of node B and node D repeat the broadcast in a similar way and the process continues till the network is completely covered. Moreover, the node sending a clique discovery message sets a timeout period to be aware of their neighbors. If no discovery message is received during this time interval, the node assumes itself as a leaf node.

As the messages are exchanged, nodes send a recognition message back to the nodes from which they first received the message; in Figure 5-4, uncovered nodes (e.g., nodes E, F and C in the present example) send a clique discovery message back to node B. Similarly, nodes G and H form a clique of size 2 after the exchange of messages with node D. As the nodes get aware of different cliques, they go into sleep mode by setting up a wakeup timer. This is shown in Figure 5-5 in which nodes H, G, E, F and C send a sleep message to backbone nodes by setting up a wakeup timer.

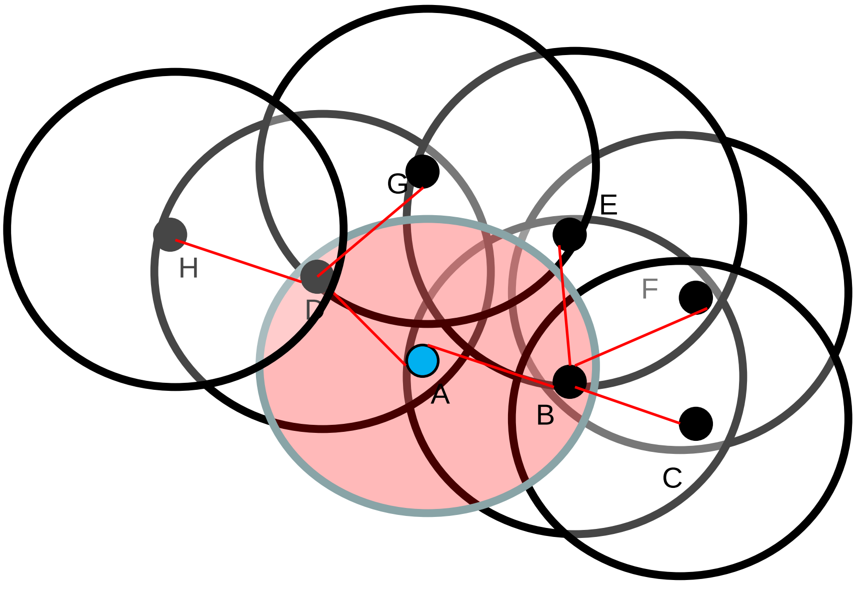


Figure 5-4: Uncovered nodes check the Node ID and rebroadcast clique discovery message with their Node IDs.

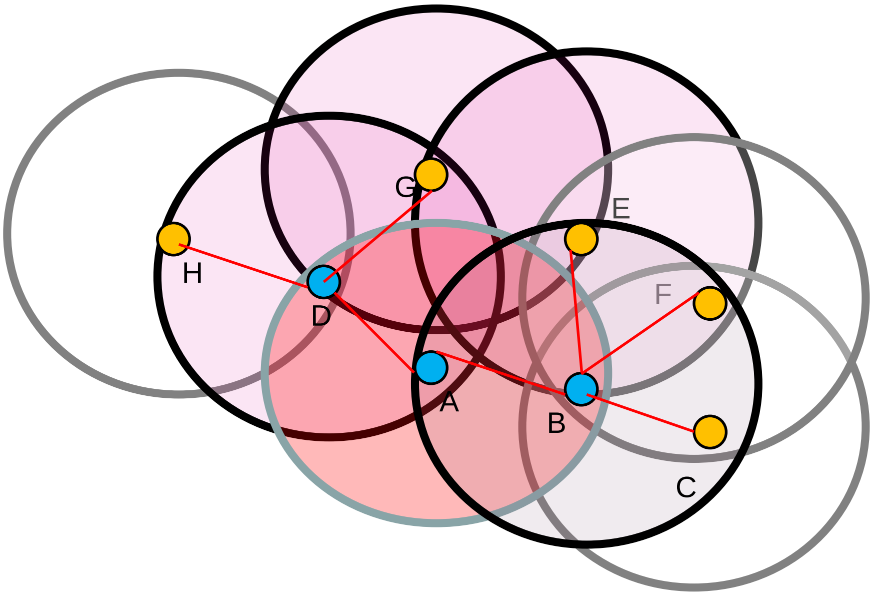


Figure 5-5: Nodes not receiving any response to their clique discovery message set a timeout period and goes into sleep mode.

This completes the description of the protocol. In the next subsection, the CCDS relevance with the guidelines is explained which is then followed by the performance evaluation of the CCDS protocol.

5.3.2 Guidelines and the CCDS protocol

In this section, the connection between the guidelines given in section 5.2 and the CCDS protocol is presented.

To follow the guideline 1, the CCDS protocol forms cliques of size 2 based on first come first serve basis. Since the CCDS does not select any node based on the selection metric, the clique sets form a CDS backbone which covers the whole network. Moreover, each node becomes part of the CDS in many cliques thereby forming a CDS proportional to the size of the network.

The protocols which follow the guideline 2, forms the MIS based on battery power and effective degree to provide better energy efficiency. However, in CCDS protocol this guideline is not used, since CCDS uses a single phase topology construction process to construct the CDS directly from the nodes in the network. In order to achieve energy efficiency, CCDS uses inherent wireless broadcast medium for parent discovery in the topology construction process. Node selection metrics or node marking restricts diverse nodes to become part of the CDS. The CCDS protocol constructs the CDS by forming cliques and does not use any marking rule or selection metric by following the 3rd guideline. The nodes become active based on the present state of the network. Under maintenance operation, this property allows diverse nodes to become part of the CDS. The static techniques compute the topology construction overhead at the start of maintenance operation which does not represent a realistic scenario. This is due to the reason that active nodes selected at the start may not have enough energy for active operation at the later stage. Hence, it is important to construct the topology on the fly based on the present state of the nodes in the network. Therefore, the proposed CCDS protocol uses dynamic topology maintenance by following the fourth guideline.

The CDS-Rule K and EECDS protocols use two phase topology construction process which increases the message overhead. Similarly, the A3 protocol sends explicit children recognition messages in reply to the parent recognition message which further increases the message overhead. To follow the 5th guideline, the CCDS is constructed in one phase with 2-cliques in the network. Moreover, nodes avoid the use of any explicit response message and use the inherent wireless broadcast medium for parent discovery. This leads to a significant reduction in the message overhead for the CDS discovery process as compared to the A3 protocol.

5.4 Performance Evaluation

Simulation-based performance evaluation is the most widely-used method to ascertain properties of ad-hoc protocols. While topology construction protocols have been an active research topic for the past few years, several shortcomings exist in the simulation environment which is used to evaluate these protocols. A summary of few of these shortcomings is presented which is as follow.

1) Lack of consistent evaluation metrics: Existing studies evaluate the performance of their protocols on the basis of a few metrics which are not even used consistently across different studies [53], [54], [55], and [91]. This impacts the credibility of the evaluation process in two ways. First, biases in the performance analysis go unnoticed. Second, the proposed protocols may not be performing well in terms of other metrics thereby limiting their scope.

2) Lack of evaluation on diverse network topologies: The reported simulation studies of topology construction protocols assume -- in general -- a fixed network topology and report the selected metric values. On one hand, it represents a biased evaluation affecting the credibility of the whole process. On the other hand, it does not unveil the performance of a protocol on diverse network topologies under topology maintenance. Consequently, its overall standings remain in the shadow.

3) Lack of one-to-one comparative evaluation: Mutual comparison of the prominent protocols over a range of relevant metrics has not been done before for different topology construction protocols using maintenance techniques. However, this should be the prime objective as such comparative analysis helps to identify and learn from the strengths and weaknesses of each protocol.

In the following subsections, a comprehensive performance analysis of the CCDS protocol is provided. In this context, the evaluation framework is explained which is followed by the discussion on the results obtained from simulation of each protocol in different network topologies.

5.4.1 Simulation Setup

The Atarraya Simulator [96] designed specifically for the evaluation of WSN topology control protocols was used for simulating the topology control environment. The simulator allows the scalability of the underlying network with the ease of selecting different network parameters, such as area, transmission range, etc. The results are reported for experiments with varying node densities in which the transmission range of nodes were fixed to a value of 42m. In addition, experiments were also performed with varying transmission ranges, but are skipped since those results did not provide any new insight. It was also assumed that the nodes are deployed randomly on a two dimensional plane of $600\text{m} \times 600\text{m}$. The network size was varied from 50–250 nodes to verify the scalability of the protocols. The nodes can communicate with each other using full duplex wireless radios that conform to 802.15.4 wireless standard.

In order to allow performance evaluation under practical settings, it was also assumed that the topology control protocols are operating over a standardcomplaint 802.15.4 WSN MAC/PHY layers. The IEEE 802.15.4 packets have a maximum MAC Protocol Data Unit (MPDU) size of 127 bytes. To use the MPDU in the experiments, it is assumed that the A3 protocol uses two types of messages for topology construction. The short messages are assumed to be of size 25 bytes and long messages are 100 bytes long. The EECDS protocol uses broadcast packet size of 25 bytes with 6 types of messages for topology construction which does not exceed broadcast packet size. Similarly, the CDS-Rule K protocol also uses 25 byte broadcast packet. For CCDS protocol, all messages have a size of 25 bytes. During topology maintenance, it is assumed that a sensed data packet equals 100 bytes for all the four protocols.

Each reported result is averaged over 50 simulation runs and the data packet size equals 25 bytes with no packet loss at the data link layer. Each node was assigned initial energy level of 1 Joule. The energy consumed during actuation equals $50\text{nJ}/\text{bit}$ while the energy consumed during communication is $100\text{PJ}/\text{m}^2$. The performances of the protocols were also evaluated under mote energy model recently added in new version of Attaraya (see subsection 5.4.3.2).

For topology maintenance, an energy threshold of 10% for energy triggered technique was assumed. The results of time-based triggering are not reported due to its very minor impact on the performance criteria. In the following subsection, the performances of the protocols based on the performance metrics are presented.

5.4.2 Simulation Results

The discussion on simulation results is divided into two subsections. First, the results for all the four protocols while constructing the topology are discussed. Subsequently, the performance of the protocols under the dynamic topology maintenance technique is elaborated.

5.4.3 Impact of varying node density

In this section, the impact of network size on the metrics for all the four protocols is explained. An increase in the network size provides an insight into the ability of the protocol keeping in view of the fact that the utilization of network resources increases considerably. For the CDS-Rule K, EECDS protocol, message complexity increase more rapidly than CCDS as shown in Figure 5-6.

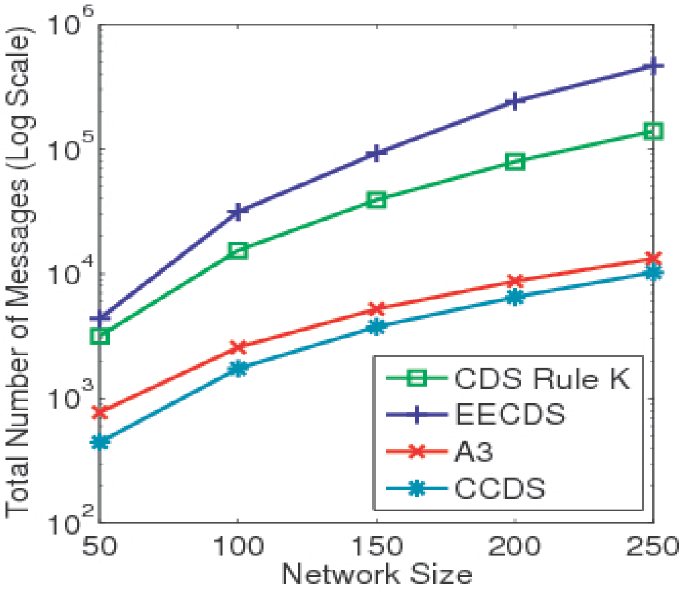


Figure 5-6: Total number of exchanged messaged Vs. Varying node density.

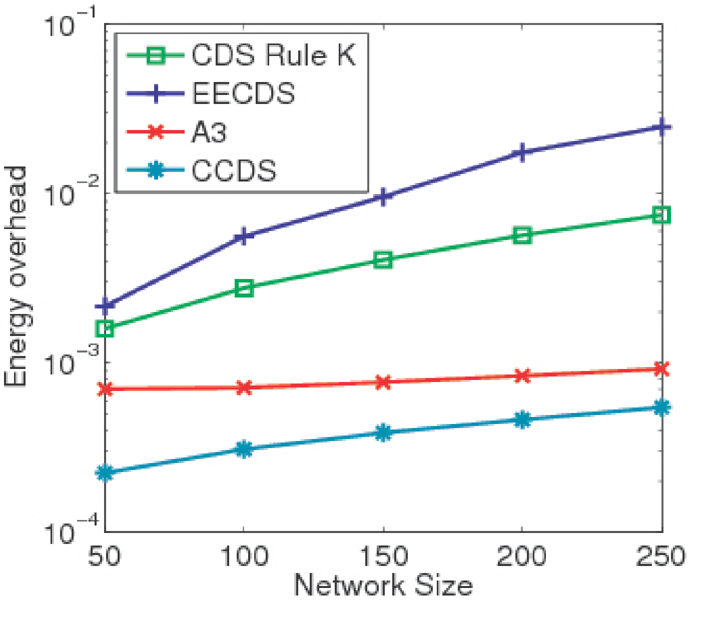


Figure 5-7: Energy overhead in CDS creation Vs. Varying node density.

The CDS-Rule K and EECDS protocol use two phase topology construction process which causes an increase in the number of exchanged messages. Moreover, it also leads towards high energy overhead for both the protocols as shown in Figure 5-7.

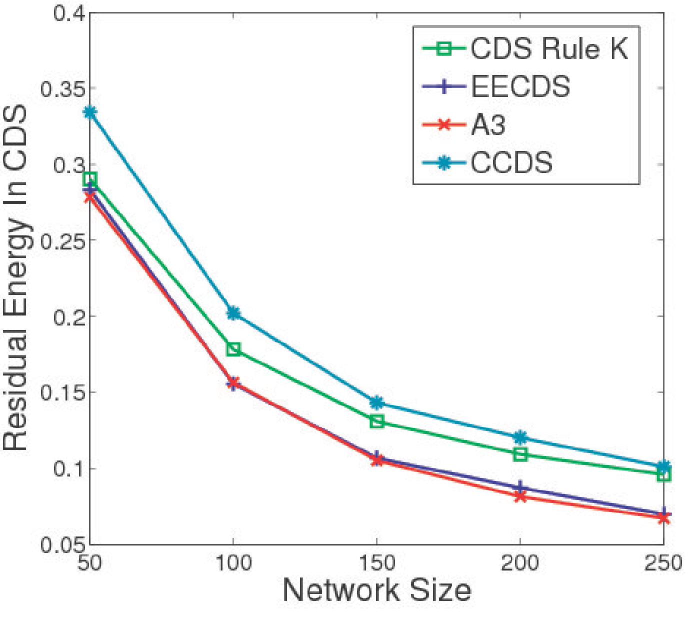


Figure 5-8: Residual Energy in CDS tree Vs. Varying node density.

The A3 protocol has low message and energy overhead due to the single phase construction of the CDS. The CCDS performs much better than the other three protocols in terms of message complexity and energy overhead. It is due to the reason that CCDS does not use separate recognition messages for each node but uses clique discovery message broadcast by the neighboring nodes. Figure 5-8 shows the residual energy present in CDS tree as the network size is increased. The CCDS protocol achieves higher residual energy and hence is more scalable than the existing protocols.

5.4.3.1 Impact of topology maintenance on simple energy model

As the network size and node density grow, message overhead of all the four protocols rises exponentially under dynamic maintenance as shown in Figure 5-9.

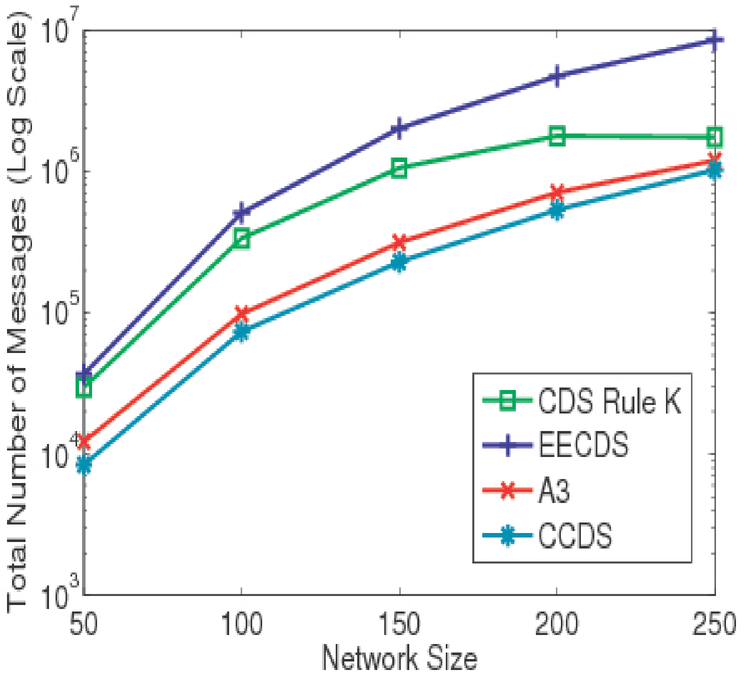


Figure 5-9: Message overhead under topology maintenance.

Message overhead of EECDS and CDS-Rule K is significantly higher than A3. This is caused by the two phase topology construction process utilized by these protocols which causes similar impact under topology maintenance. In comparison, A3 generates fewer messages because it chooses the distant nodes which consequently lead to quick convergence of the protocol. For CDS-Rule K, the number of messages starts decreasing with increase in the number of nodes. This is because of only a few nodes remain in the network or CDS tree (see Figure 5-12). The CCDS protocol achieves energy efficiency with low message overhead.

Figure 5-10 and Figure 5-11 show residual energy and energy overhead for the four protocols under dynamic topology maintenance with energy-based triggering. The CDS-Rule K protocol uses marking and pruning rules while EECDS uses a two-phase process for topology construction which leads to higher energy overhead under topology maintenance. This trend is visible in Figure 5-11. An interesting observation is that, although EECDS consumes higher total energy, it has significantly better residual energy (Figure 5-10).

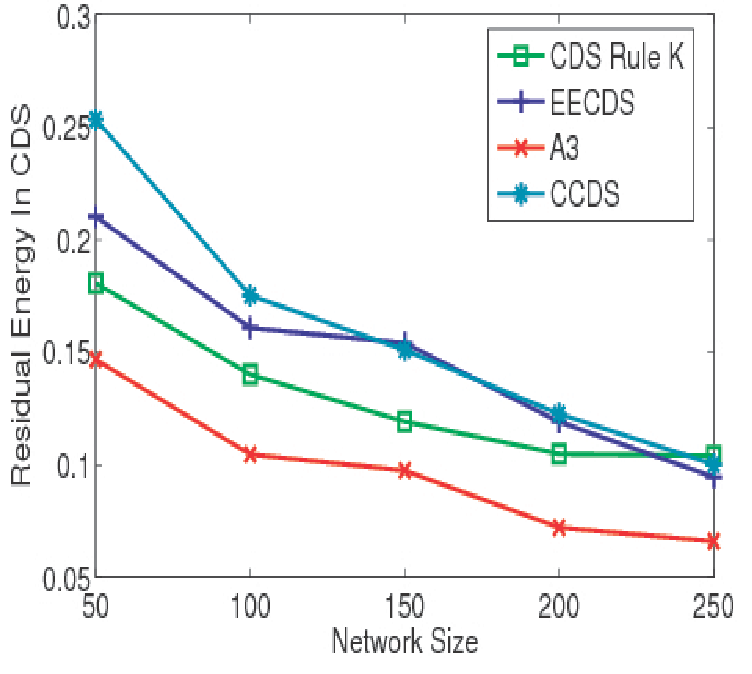


Figure 5-10: Residual energy in CDS under topology maintenance.

On the other hand, A3 consumes lesser total energy but its residual energy is lower than EECDS and CDS-Rule K. This is due to non-uniform distribution of communication overhead which drains the battery of fewer nodes resulting in lower residual energy level.

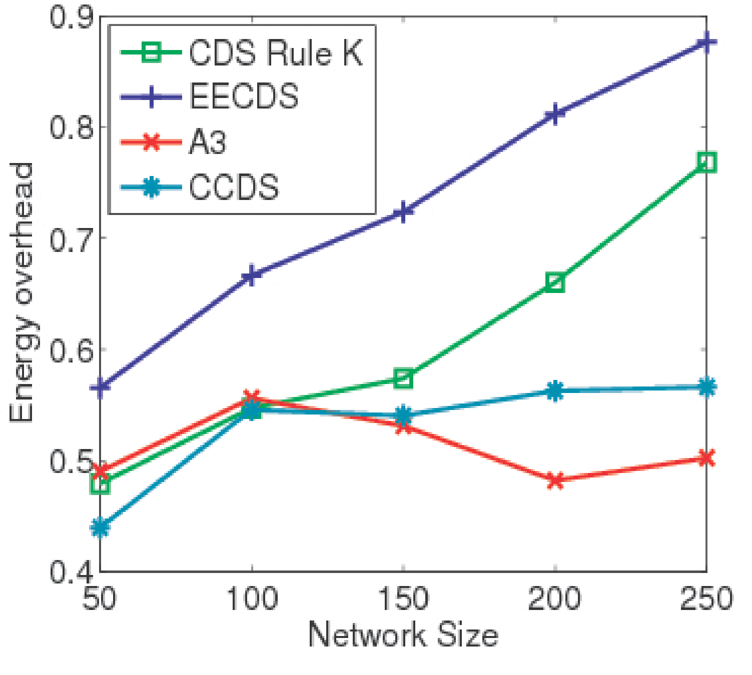


Figure 5-11: Energy overhead under topology maintenance.

The energy overhead of EECDS and CDS-Rule K increases with an increase in the number of nodes, while for A3 it starts decreasing with an increase in the number of nodes. As the node density rises, energy overhead of A3 decreases because it has less message overhead in dynamic case. In comparison, CDS-Rule K and EECDS protocols generated higher message overhead. Similarly, the residual energy decreases with increase in the number of nodes for all the three protocols. On the other hand, the use of the broadcast nature of wireless channels allows the CCDS protocol to cover the end nodes, providing better information in constructing a new CDS tree. It also leads to uniform distribution of energy resources [see Figure 5-10 and Figure 5-11].

The numbers of unconnected nodes with energy-based triggering for all the three protocols are shown in Figure 5-12. The number of unconnected nodes increases exponentially for CDS-Rule K and EECDS protocols. However, A3 protocol has less number of unconnected nodes under dynamic topology maintenance scheme. In CDS-Rule K, nodes remained marked if there is at least one pair of unconnected neighbor.

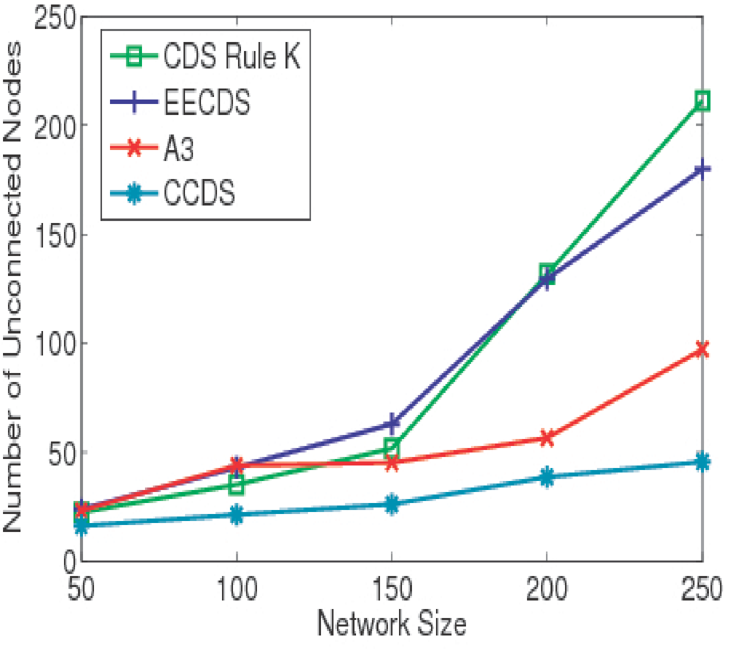


Figure 5-12: Number of unconnected nodes under topology maintenance.

The energy depletion of the marked node causes more number of unconnected nodes. The CCDS has less number of unconnected nodes when compared with the other three protocols which show that CCDS has better network lifetime and it also provides better network connectivity.

The convergence time for all the four protocols is tabulated in Table 5-1. A3 takes less execution time due to its nodes selection procedure which is based on signal strength. Since A3 converges quickly, its message overhead in the topology construction process is low as well. Consequently, it incurs less energy overhead leading to longer network lifetime. The CCDS also has less convergence time as it utilizes the inherent broadcast nature of wireless medium for parent discovery which allows the quick convergence of the protocol.

Table 5-1: Average Convergence Time under Dynamic Topology Maintenance.

	CDS- Rule K	**EECDS**	**A3**	**CCDS**
Dynamic Topology Maintenance	110.874124	144.440978	35.771068	3.35235081

5.4.3.2 Impact of topology maintenance on mote energy model

To increase the scope of the simulation and to get better insight into the protocols' performances, all the four protocols were evaluated under 500 simulation runs/ random generated topologies. Moreover, the realistic mica2 mote energy model recently added in the new version of Attaraya was used.

It is also now considered that the nodes energy distribution follows a Poisson process. In addition, the sensing coverage performance of the protocols in terms of deployment area is also analyzed. Figure 5-13 and Figure 5-14 shows the message overhead, energy overhead under dynamic topology maintenance. The results demonstrate that CCDS protocol has low message overhead.

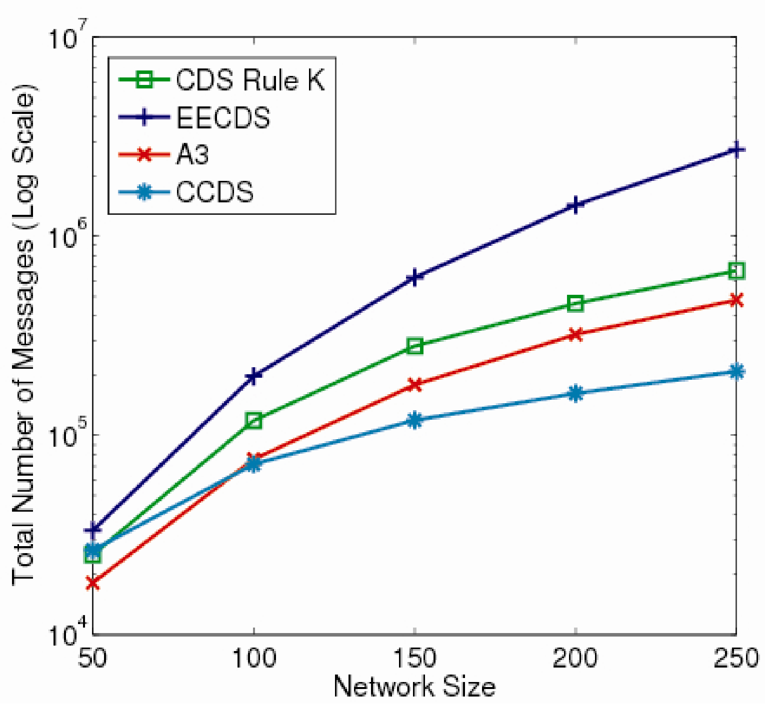


Figure 5-13: Total number of exchanged messages under mote energy model.

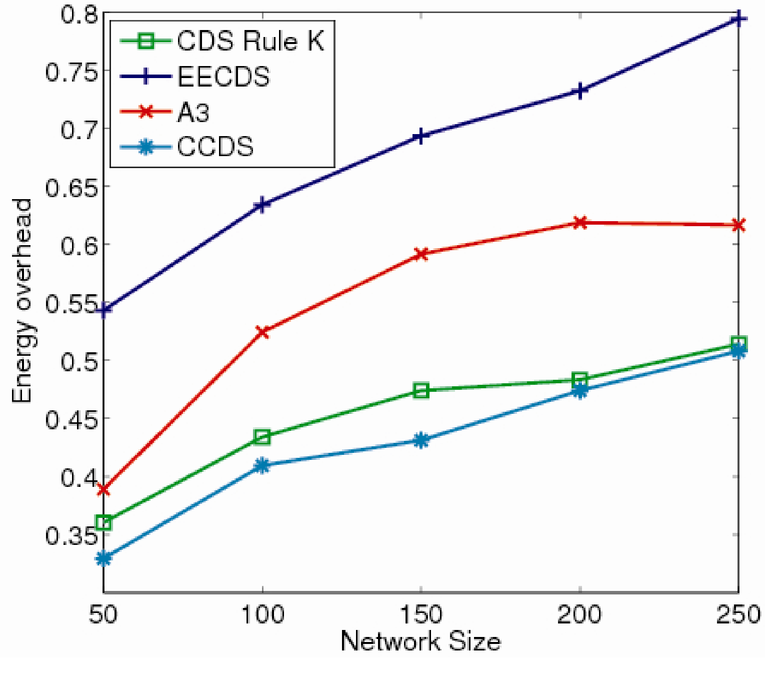


Figure 5-14: Energy overhead under mote energy model.

Figure 5-15 and Figure 5-16 shows the residual energy, and number of unconnected nodes under dynamic topology maintenance. The results demonstrate that the CCDS protocol has better energy resources when compared with the other protocols. Similarly, the CCDS protocol has more number of connected neighbors when used with dynamic topology maintenance as shown in Figure 5-16.

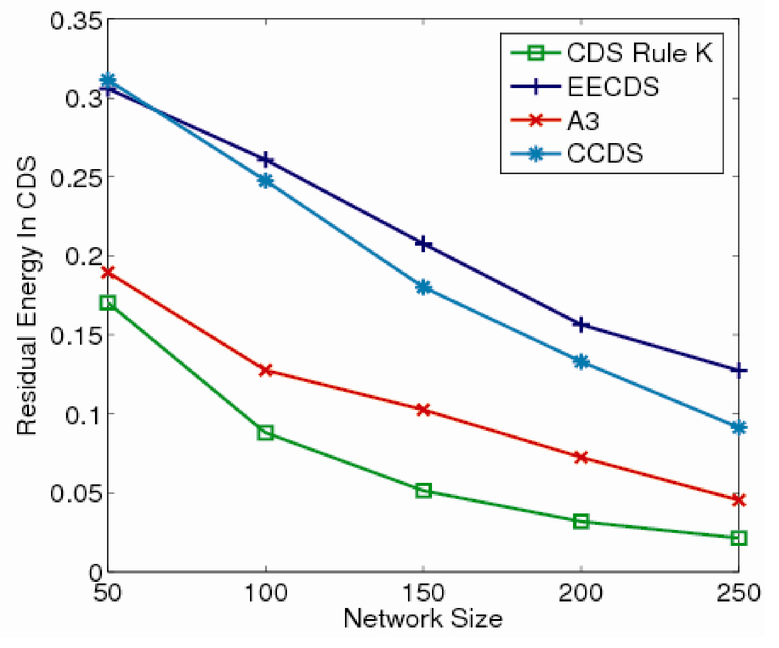


Figure 5-15: Residual energy in CDS under mote energy model.

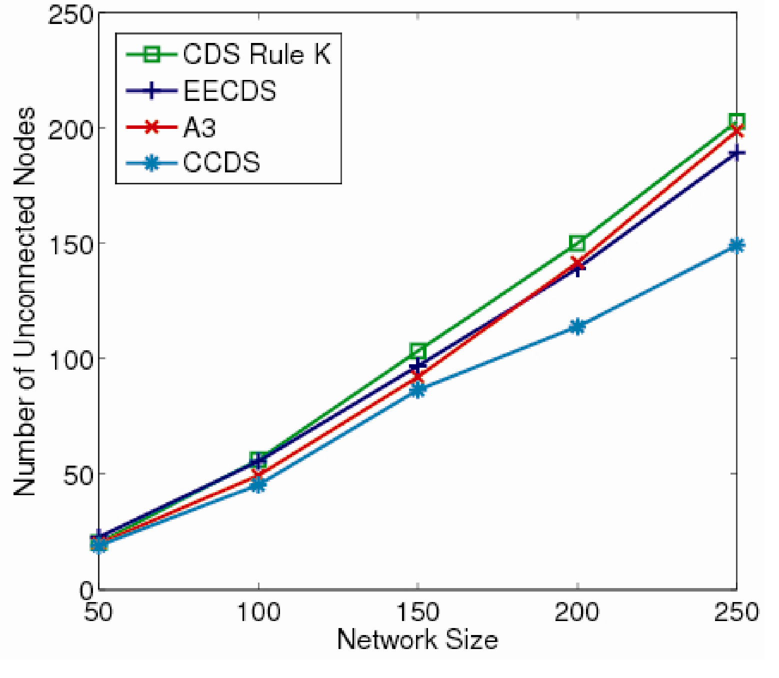


Figure 5-16: Number of Unconnected nodes under mote energy model.

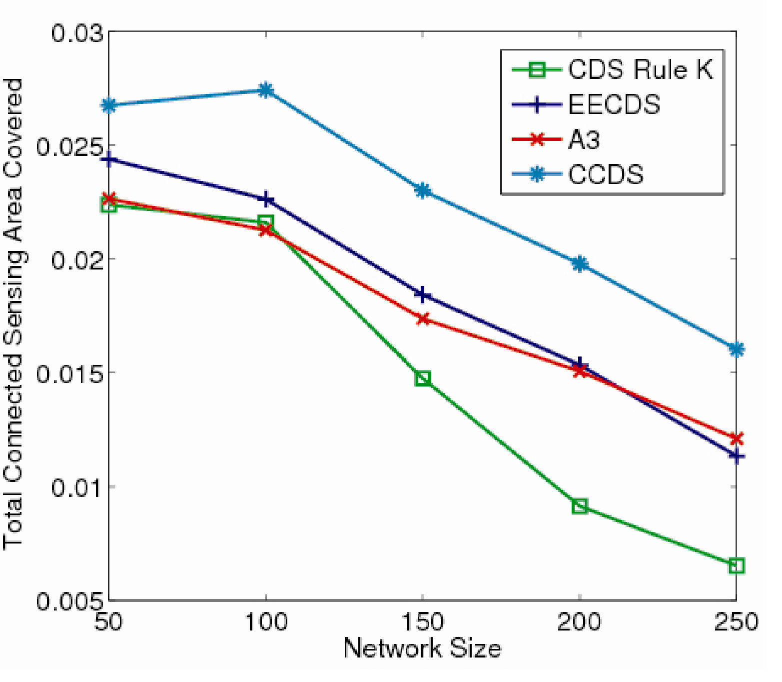


Figure 5-17: Total area of sensing coverage.

The average sensing area covered by all the four protocols under dynamic topology maintenance operation is shown in Figure 5-17. Under topology maintenance, sensing coverage reduces as nodes deplete the energy and lose connectivity with the sink node. This phenomenon is similar for all the four protocols as shown in Figure 5-17. However, these results demonstrate that CCDS covers more area while using fewer resources as compared to other protocols.

5.4.3.3 Impact of Channel Errors:

Figure 5-18, Figure 5-19, Figure 5-20, and Figure 5-21 shows the impact of channels errors on the number of exchanged messages during the topology construction for all the four protocols. As the bit-error increases, the number of exchanged messages also rises. Furthermore, it is also important to note that these results are qualitatively similar to the ones that were obtained earlier. More specifically, CCDS protocol still generates the least message overhead among all the candidate protocols which support the previous study.

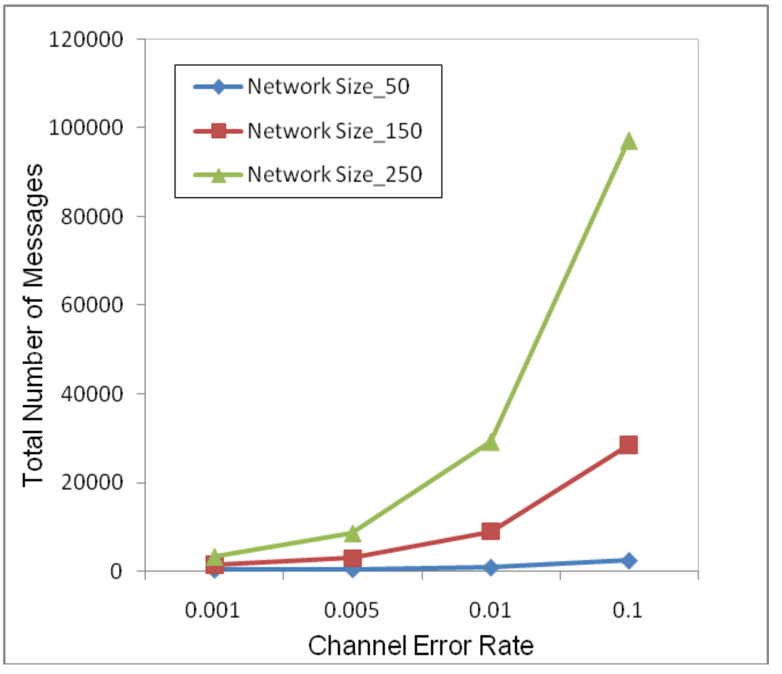


Figure 5-18: Number of exchanged messages Vs. Network Size for CDS-Rule K Protocol.

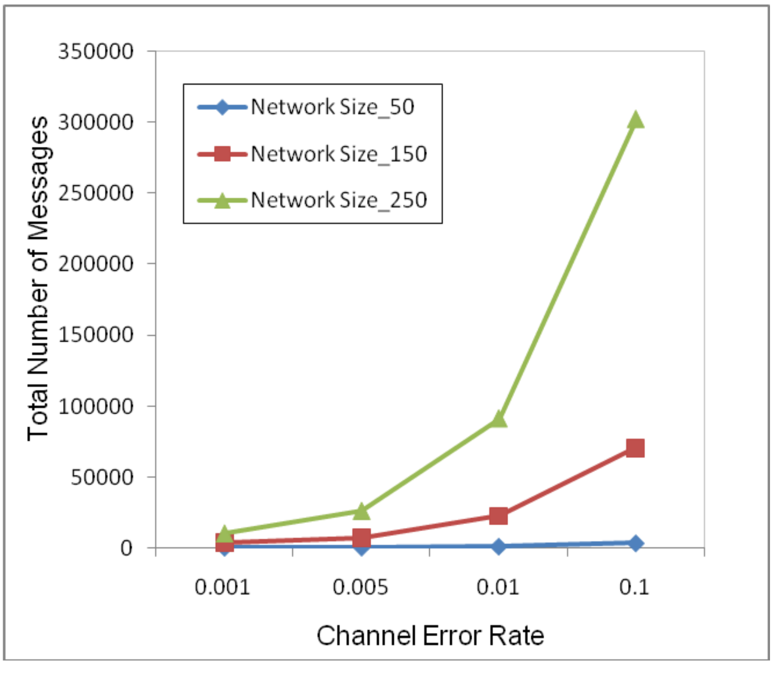


Figure 5-19: Number of exchanged messages Vs. Network Size for EECDS Protocol

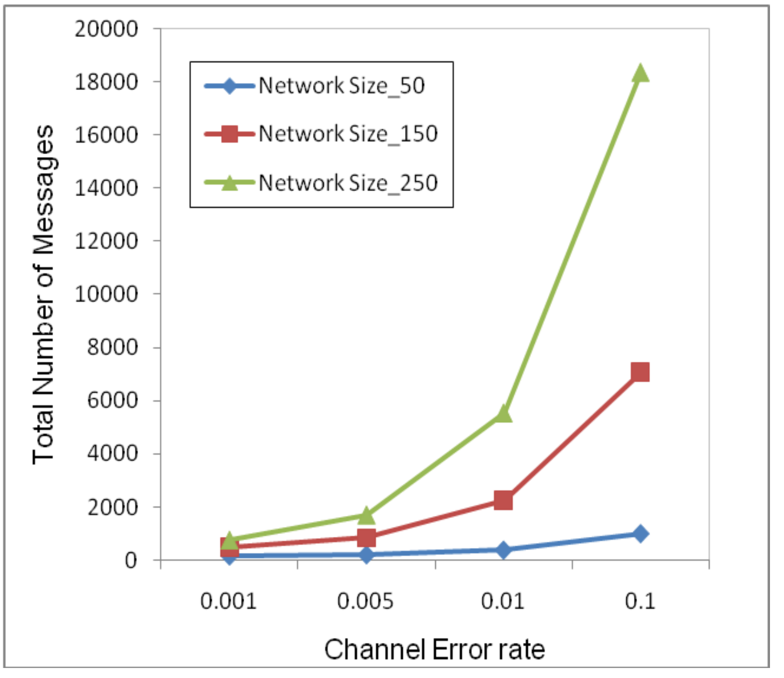


Figure 5-20: Number of exchanged messages Vs. Network Size for A3 Protocol.

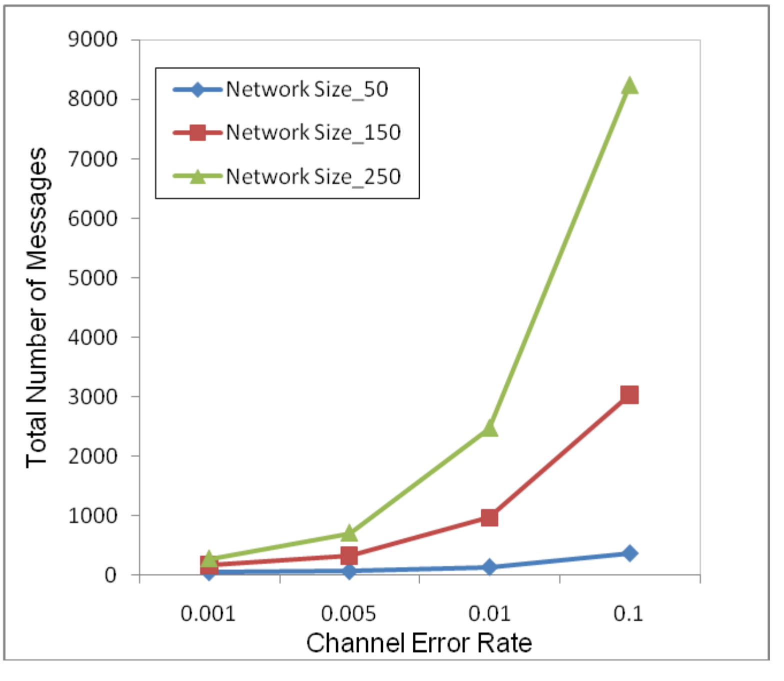


Figure 5-21: Number of exchanged messages Vs. Network Size for CCDS protocol.

5.5 Summary

In this chapter, prominent protocols and maintenance techniques used by topology construction protocols are explained. The performance evaluation is one of the critical components of a protocol engineering cycle. For this purpose, extensive simulations were performed to compare the performance of three prominent topology construction protocols; CDS-Rule K, A3 and EECDS over a large operational landscape. The empirical results demonstrate that A3 consumes less amount of energy due to its low message overhead. On the other hand, EECDS and CDS-Rule K, although consume higher energy than A3, achieve better residual levels that extend the overall network lifetime. It has also been shown that A3 converges quickly than EECDS and CDS-Rule K protocols. Based on the analysis of existing topology construction protocols, a set of guidelines that can be used to design efficient - in terms of energy and performance - CDS-based clustered WSN topologies are derived. As a proof of concept, these guidelines are utilized to propose a novel CDS-based protocol using 2-cliques in the network. In the end, it is shown through simulations results that CCDS - the proposed protocol - performs efficiently and has less associated message overhead.

Chapter 6

A1: An Energy Efficient Topology Control Protocol for Connected Area Coverage

6.1 Introduction

Wireless sensor networks continue to be a very popular technology to monitor and act upon events in dangerous or risky places for humans. WSN's are easy to deploy in an application field and the cost is relatively low by the continuing improvements in embedded sensor, (Very-large-scale integration) VLSI, and wireless radio technologies [105].

Although WSN's have evolved in many aspects, they continue to be networks with constrained resources in terms of energy, computing power, and memory. In addition, nodes have limited communication capabilities due to which a source node can cover only within its maximum transmission range. On the other hand, it causes nodes to relay messages through intermediate nodes to reach their destinations. Due to this reason, routing related tasks become much more complicated in WSN's since there is no pre-defined physical backbone infrastructure for topology control. This drawback motivates a virtual backbone to be employed in a WSN. Conceptually, a virtual backbone is a set of active nodes which can send message to the destination by forwarding the message to other neighboring active nodes. These set of active nodes provides many advantages to network routing and management. This is due to the reason that routing path gets reduced to the set of active nodes only which provides an efficient fault-tolerant routing. Moreover, the reduced topology reacts quickly to topological changes and is less vulnerable in terms of collision problems caused due to flooding based routing protocols [106].

The authors in [107], [108] introduced the first approximation algorithms to compute a virtual backbone using a Connected Dominating Set (CDS). Since then, CDS based topology control (TC) has emerged as the most popular method for energy efficient (TC) in WSN's. TC has two phases namely: topology construction and topology maintenance. In the topology construction phase, a desired topological property is established in the network while ensuring connectivity. Once the topology is constructed, topology maintenance phase starts in which nodes switch their roles to cater for topological changes.

In CDS-based TC schemes, some nodes are a part of the virtual backbone which is responsible for relaying packets in the WSN. These nodes are also called dominator nodes or active nodes. Non-CDS nodes or dominatees relay information through the active nodes. Hence, a CDS works as a virtual backbone in the reduced constructed topology. The CDS size remains the primary concern for measuring the quality of a CDS. The authors in [91], [109] prove that a smaller virtual backbone suffers less from the interference problem and performs more efficiently in routing and reducing the number of control messages. Moreover, this allows the maintenance of the CDS much easier and provides better reliability for a fixed probability of success. Due to these reasons, most research studies in this area focus on reducing the size of a CDS [53]-[58], [111], and [112]. However, most studies do not consider the impact of topology maintenance under which many nodes get disconnected from sink node. This is due to the reason that for small virtual backbones, fewer nodes handle the bulk of the network traffic and consequently deplete their batteries quickly. On the other side, this causes the reduction in the virtual backbone size, which affects the coverage region of WSN.

In this chapter, a distributed topology control protocol for wireless sensor networks is proposed. The protocol, referred to as the A1 protocol, models the topology as a connected network and finds the set of active nodes to form a CDS. The A1 protocol uses node IDs of different nodes and a node selection criteria for nodes to calculate their timeout. In this way, nodes turn-off themselves and later repeat the process -- after the timeout expires -- to discover neighbors desiring them to work as an active node. In this way, a reduced topology is formed while keeping the network connected and covered. To achieve energy efficiency, the protocol forms the CDS comprising of high energy nodes in a single phase construction process. In addition, it also forms a proportionate set of active nodes in order to provide better sensing coverage. Moreover, it adapts to the topological changes in the network based on the remaining energy of the nodes. This allows better topology maintenance among different set of nodes which increases the network lifetime.

The performance of the protocol is compared with Energy Efficient CDS (EECDS) [55], CDS-Rule K [54] and A3 [53] protocols. For this purpose, extensive simulations are performed under varying network sizes to analyze the message complexity and energy overhead in terms of spent energy and remaining energy in the CDS. The performance of the protocols under topology maintenance to verify the nodes connectivity in terms of number of unconnected nodes is also analyzed. As the primary task of a WSN network is to provide sensing coverage of the area, the performance of the protocols on connected sensing area covered at the end of topology maintenance is also verified. The results show the proposed A1 protocol has low message complexity. Moreover, it also provides better residual energy resources while having less number of unconnected nodes under topology maintenance. In addition, the A1 protocol has better connected sensing area and it covers 35% more area when compared with the other three protocols.

The rest of this chapter is organized as follows. Section 6.2 summarizes the related work in this area. The A1 protocol is explained in Section 6.3. In section 6.4, the empirical evaluation framework utilized for the performance analysis of A1 is explained. Section 6.5 shows the discussion on simulation results with sensing coverage analysis of the protocols. The salient findings during implementation and analysis of A1 are explained in Section 6.6.

6.2 Related Work

The CDS based topology construction in WSN's has been studied extensively. Some of the existing protocols [94] consider using the transmission power of WSN nodes to achieve energy efficiency while some used geographical location of the nodes [95]. However, power control and location awareness are difficult to realize in practical WSN deployments. Similarly, constructing CDS for heterogeneous networks by using directional antennas is proposed in [113]. In directional antenna models, the transmission/reception range is divided into several sectors and one or more sectors can be switched on for transmission. However, it is difficult to realize these schemes in case of WSN's because of limited energy source. We now explain some of the relevant CDS based research efforts in the area.

There exist some work in [56], [57], and [58] that describe the construction of k-connected m-dominating sets for fault tolerance. To this end, they have proposed two approximation algorithms – Connecting Dominating Set Augmentation (CDSA) and k-connected m-dominating set (k, mCDS) – to construct a k-connected virtual backbone which can accommodate the failure of one wireless node. However, most of the work focuses on reliability in their network models. In practice, reliability should be achieved while considering sensing coverage region of the nodes. Moreover, the studies do not analyze the impact of exchanged messages for (k, mCDS) on energy efficiency. Similarly, some other protocols in this context were earlier presented and explained in detail in section 4.6 and therefore are skipped here in this section.

We now explain the working of the A1 protocol in the next section.

6.3 The A1 Protocol

As the primary focus is on energy efficient reduced topology, the fundamental design application that is used to reduce the size of the backbone nodes is with the help of signal strength and energy based timeout criteria. The nodes selection criteria for timeout is given by

$$T_{d,s} = (E_d / E_i) + (RSS_s / RSS_c), \qquad (6.6)$$

where d and s represents the children node and parent node, E_d is the remaining energy level of the children node and E_i is the initial energy level. Similarly, RSS_S is the signal strength of parent node received by the children node and RSS_C is the minimum required signal strength to ensure connectivity. The selection criteria allow high energy nodes with better signal strength to be selected. This is due to the reason that the neighbors of the node select a low value for timeout if they calculate a high value for selection criteria. The selected nodes serve as a virtual backbone for all the nodes in the network and hence forming a CDS.

In the following two subsections, the CDS formation process in A1 protocol is explained. In the first subsection, the types of discovery messages that are used during the topology construction are explained. Subsequently, the mechanism that leads to the formation of CDS in the network is illustrated.

6.3.1 Description of Topology Discovery messages

There are several factors which impact energy efficiency. However, energy efficiency is mainly dependent on packet size and continuous listening in promiscuous mode [112]. Energy consumption increases with the increase in size of packets and affects both sending and receiving nodes in the network. In A3 protocol, children recognition messages contain ordered list of all the children of sender. This list is used by children to set a timer to compete for an active node. When the network is dense, this list increases with the increase in the message size and hence consumes more energy. The more the children, the more the length of the message and it will result in more energy consumption per children recognition message. Due to this reason, the A3 protocol uses a 100 bytes size for children recognition message apart from other messages of size 25 bytes. On the other hand, the EECDS protocol uses broadcast packet size of 25 bytes with 6 types of messages for topology construction which does not exceed broadcast packet size. Similarly, the CDS-Rule K protocol also uses 25 byte broadcast packet.

In order to improve the energy efficiency, the A1 protocol uses only one type of message for CDS formation. A hello message of size 25 bytes contains the parent ID of the sender discovers the reduced CDS topology. The parent node do not decide the timer value for its children by sending an explicit children recognition message. Instead, children nodes calculate and set a timeout period on their own after the reception of a hello message. This calculated timeout is independent of timeout of other nodes due to different energy and distance characteristics of the nodes. In this way, energy efficiency is achieved during topology construction and life of the network is prolonged.

6.3.2 The Working of A1 protocol

The A1 protocol constructs the topology in one phase. At start, the initiator node first discovers its neighbor. Similarly, the neighbors of the initiator node discover their neighbors as their timeout expires in the second phase. This process continues until the complete topology is formed with nodes acting as the virtual backbone (CDS) for rest of the nodes in the network.

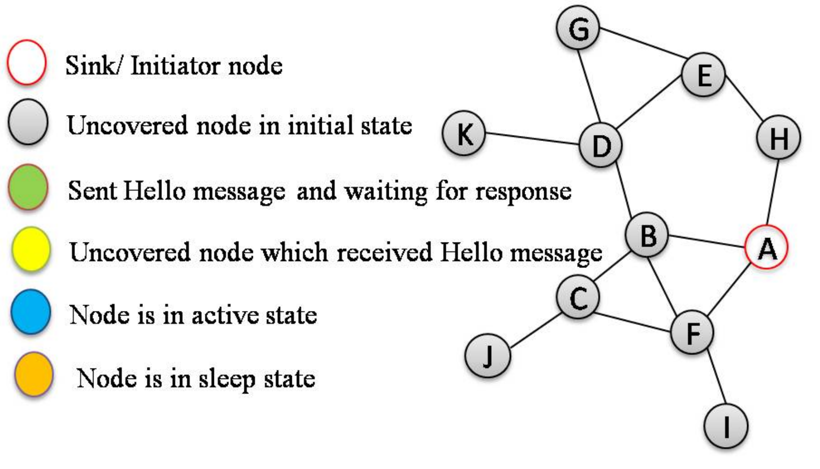


Figure 6-1: A sample Topology.

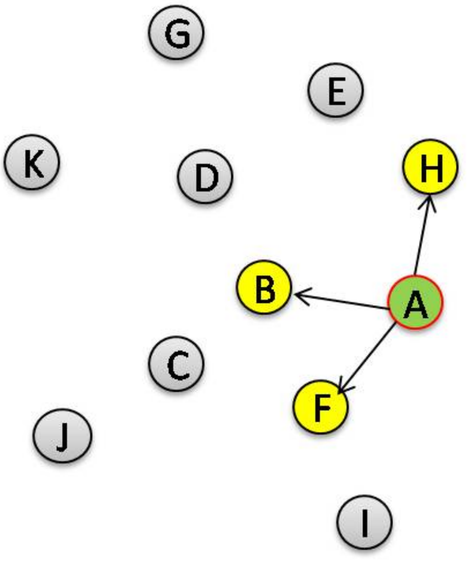


Figure 6-2: Sink node (A) broadcasts Hello message which is received by nodes (B, F and H) under its coverage area. After receiving Hello message, B, F and H calculate their timeout.

We describe the construction of the reduced topology -- formed with the A1 protocol-- with the help of an example network shown in Figure 6-1. The topology construction starts in A1 by a node called an initiator node. For protocol implementation, we selected a random node as an initiator node and if more than one node initiates the process, the node with the largest ID is chosen. In Figure 6-1, the initiator node A broadcasts a hello message to start the topology construction process. The parent node then waits to hear a message with parent ID set to its own ID. It is important to point out that the parent ID field is empty in case of the initiator node. The nodes B, F and H which are located within the transmission range of A receive the hello message (see Figure 6-2).

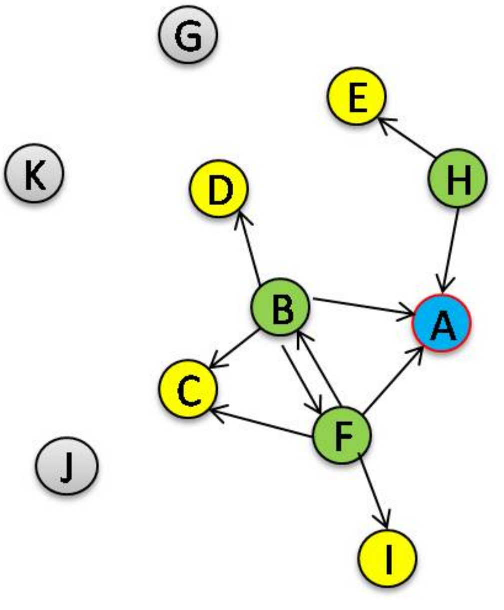


Figure 6-3: When timeout expire, B, H and F further broadcasts Hello message at different times. A turns itself active and becomes a parent node.

The nodes after the reception of the hello message, calculate the timeout according to equation (6.6) and enters into sleep mode according to the value of the calculated timeout. As the timeout expires, these nodes discover their neighbors further at different times and send another hello message with parent ID field now set to node A. This allows node A to become an active node. Nodes B and

F are located within each other's transmission range also receive the broadcasted message by both of them. Since in both messages, the parent ID is the same, both nodes recognize them as the children of the same parent node. Similarly, node C also receives the message from nodes B and node F. In addition, node E and node I receive the message from node H and node F respectively as shown in Figure 6-3.

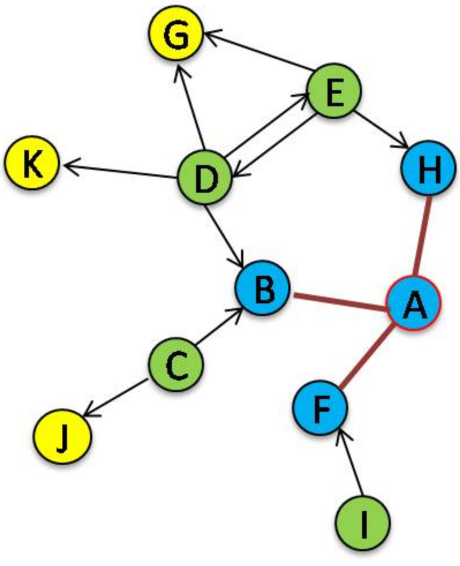


Figure 6-4: Next level nodes again broadcast the Hello message after changing the parent ID to their respective parent IDs.

Node E and node I changes the parent ID field to node H and node F respectively and broadcasts the message after the timeout expires. In this way, node H and node F becomes dominators/active nodes. Similarly, node C and D chooses node B as an active node by sending a message with parent ID field set to node B. It is worth noting that node C and node D selected node B as their parent since they received the message firstly from node B due to low value of timeout (see Figure 6-4). This message from node D is also received at node E which also sent the same message with different parent ID to node D. Since node E do not receive any message with its own parent ID, it discovers itself as a non active node. Similarly, node I also perform in the same manner (see Figure 6-5).

The nodes G and K broadcasts the message with parent ID set to node D which allows node D to work as an active node. On the other hand, node C gets aware due to the message reception from node J as shown in Figure 6-5.

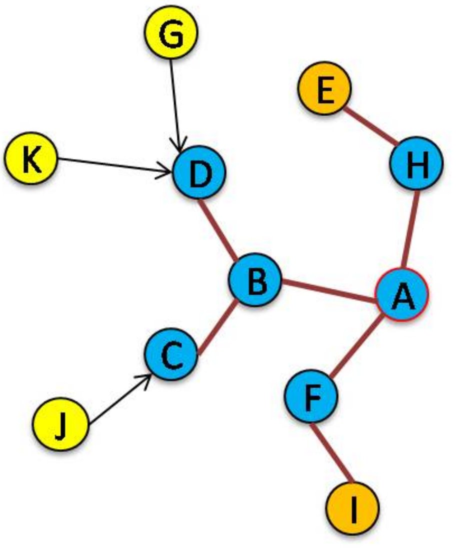


Figure 6-5: Node G and K broadcasts Hello message with parent ID set to node D. Similarly node J broadcasts Hello message with parent ID set to node C.

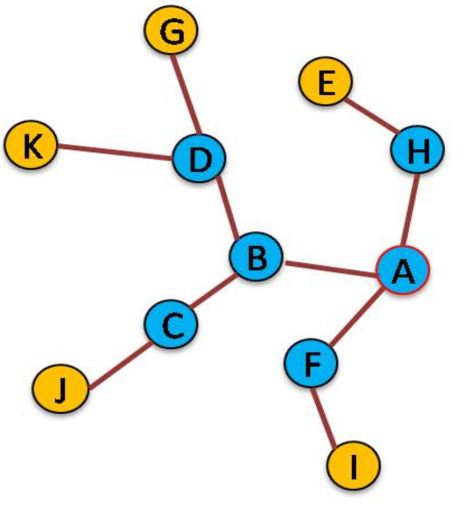


Figure 6-6: Node G, K and J do not receive any Hello message with their node ID as parent ID and therefore consider them as leaf nodes to form the final reduced topology.

In the end, nodes G, K and node I do not receive any message with parent ID set to their own ID and therefore enter into sleep mode after the expiration of calculated timeout. In this manner, a reduced and covered topology is formed in which some nodes work as a virtual backbone for rest of the nodes in the network as shown in Figure 6-6. This completes the description of the A1 protocol. We now provide our experimental setup which is used for the evaluation of the A1 protocol. It is then followed by a detailed discussion on simulation results.

6.4 Empirical Evaluation Framework

This section explains the empirical evaluation framework used for the evaluation of the A1 protocol and other CDS protocols, namely EECDS, CDS-Rule K, and A3. The discussion starts with the empirical setup that explains the simulation settings and underlying network topologies. In the subsequent section, the topology maintenance techniques are explained followed by the discussion on results.

6.4.1 Simulation Setup

The protocols were evaluated on a specifically designed simulator for WSN topology control protocols [96]. The simulator -- Atarraya -- allows the scalability of the underlying network with the ease of selecting different network parameters, such as deployment area, transmission ranges and network size.

For the simulations, a 600m × 600m virtual space was assumed in which nodes are randomly deployed. Two system parameters, the number of nodes in the space and the common transmission range of nodes were used. The number of nodes is increased from 50 to 250 nodes. We also performed experiments for the node density beyond 250 nodes; however the trend remains the same for all the four protocols. Similarly, the maximum transmission range was set to 42m in order to have a connected topology. In addition, nodes sensing range was set to 10m. In case of indoor topologies, a network of 169 nodes was assumed for Grid H-V while restricting the transmission range to 28m. For Grid H-V-D, the network size was increased to 324 nodes. The network size for indoor topologies was selected due to the deployment scenario possible as nodes communicate with their horizontal and vertical neighbors in Grid H-V, while in the Grid H-V-D topology; nodes also communicate with their diagonal neighbors.

For the same system parameter settings, 100 connected graph instances were randomly created to compute a CDS for each instance for all the four protocols. The initial energy level of each node was set to 1J with actuation energy equals 50nJ/bit, while the communication energy was set to 100PJ/bit/m^2. The nodes communicate with each other using full duplex wireless radios. In addition, to use the MAC Protocol Data Unit (MPDU) in the experiments, the message sizes of all the four protocols were used as explained in earlier section.

6.4.2 Topology Maintenance Techniques

Topology maintenance is a process in which a certain desired topological property is maintained to increase the network lifetime. As explained before, in static maintenance, a possible set of disjoint topologies are build at the start of the maintenance operation. The pre-constructed topologies are then rotated based on the time or energy based triggering mechanism. However, static techniques calculate the overhead of pre-constructed topologies at the start, which in most case; do not represent a realistic scenario as the backbone nodes chosen at the start can behave differently at the later stage. On the other hand, dynamic topology maintenance techniques form a new topology based on the present condition of the network, e.g. as the threshold is reached and therefore are used in the evaluation process.

In the next section, the results for dynamic topology maintenance techniques based on energy-threshold are reported. For this purpose, the energy threshold of 10% is defined i.e. topology maintenance process is triggered when the network energy falls by 10%. During topology maintenance, it was assumed that a sensed data packet equals 100 bytes for all the four protocols.

6.5 Discussion on Simulation Results

The discussion on simulation results is divided into four subsections. The discussion starts with the performance of the protocols under varying node densities. All the four protocols are then evaluated in indoor deployment environments: the Grid H-V and the Grid H-V-D topologies. Subsequently, the performance of the protocols under dynamic topology maintenance is explained. In the last subsection, the impact of CDS size on coverage area of WSN's is analyzed.

6.5.1 Impact of Node Density

The message overhead, energy overhead and residual energy results for varying node densities are shown in Figure 6-7, Figure 6-8, and Figure 6-9.

The number of exchanged messages increases with the increase in the network size. This is due to the reason that increase in the number of nodes also leads to an increase in node degree which also increases the number of exchanged messages. This trend is same for all the four protocols as shown in Figure 6-7. However, two phase topology construction leads to high message overhead for EECDS and CDS-Rule K protocols. On the other hand, A3 incurs fewer messages overhead due to single phase topology construction. Moreover, it uses less number of messages for topology construction when compared with EECDS and CDS-Rule K protocols. In comparison, A1 constructs the topology using one message and has less message overhead than EECDS and CDS-Rule K protocols. As can be intuitively argued, an increasing node density leads to higher energy overhead due to an increase in the number of received packets. This trend is visible in Figure 6-8 for all the four protocols. However, A1 protocol consumes less energy for the construction of the topology.

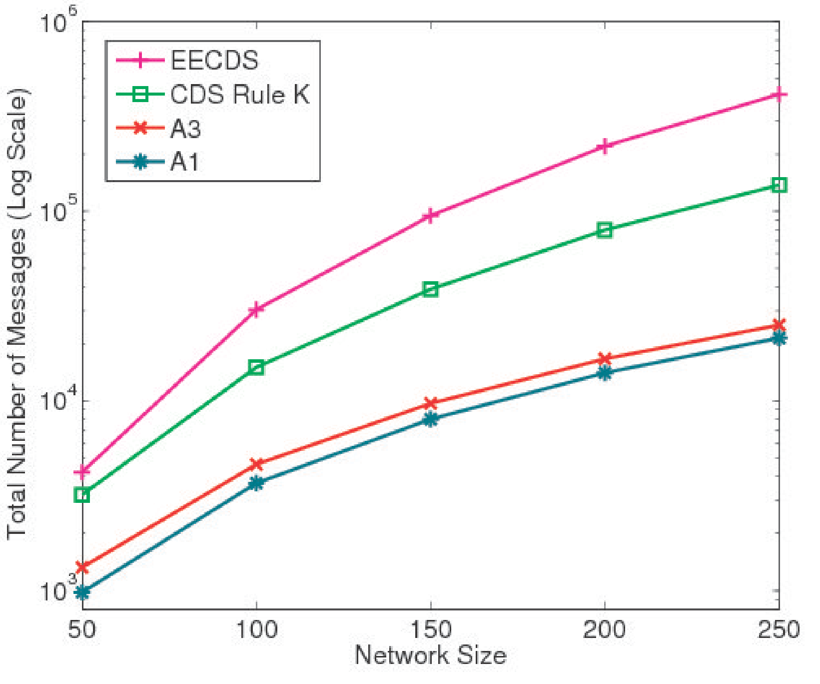


Figure 6-7: Message overhead under varying network size.

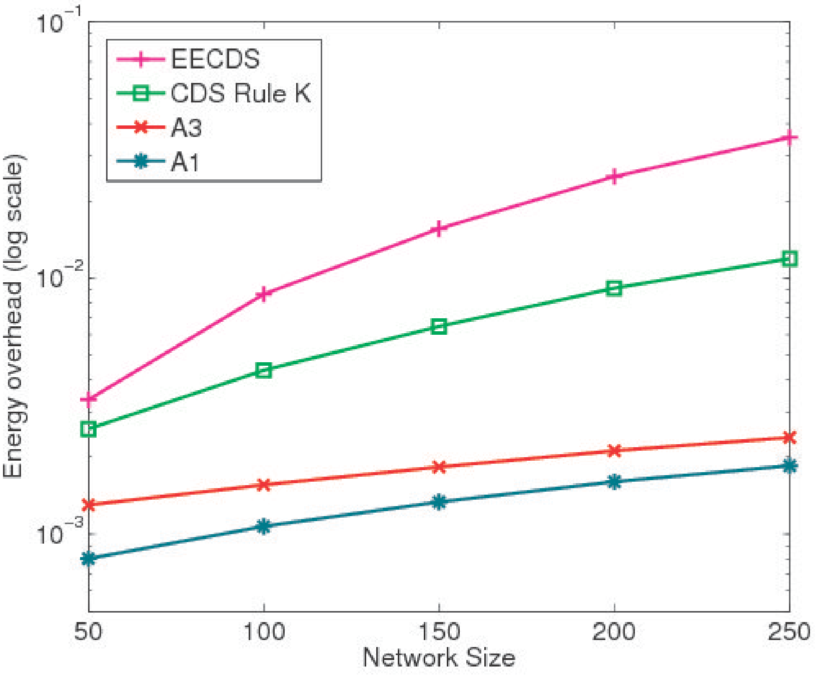


Figure 6-8: Energy overhead under varying network size.

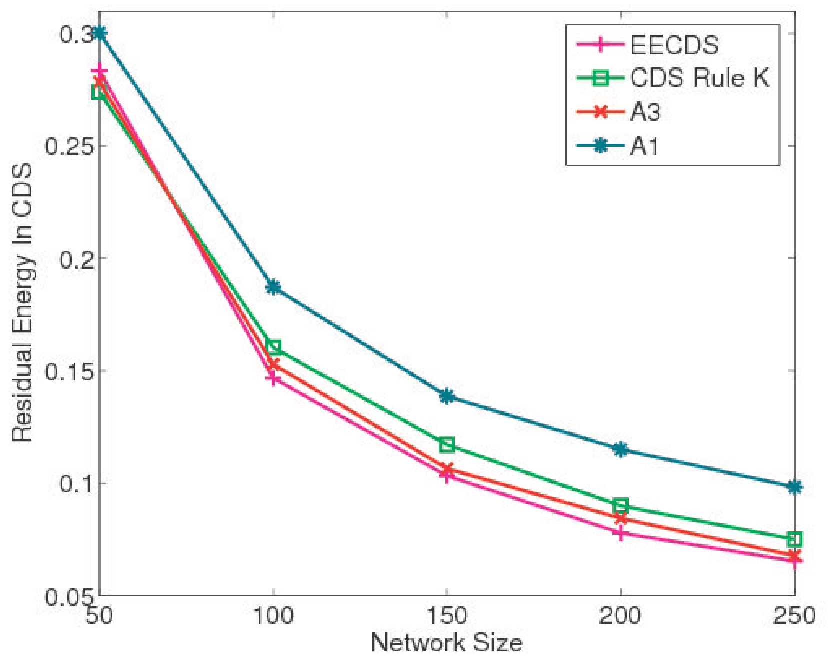


Figure 6-9: Residual Energy under varying network size.

Figure 6-9 shows the residual energy among active set of nodes for all the four protocols. Usually, high energy overhead leads to lower residual energy. But, it was observed that CDS-Rule K ends up with better residual energy resources. This is due to the reason that A3 protocol tries to reduce the virtual backbone by selecting far nodes from the parent node. This results in non-uniform distribution of communication overhead which drains the battery of fewer nodes resulting in lower residual energy levels among nodes in the network. On the other hand, A1 provides better residual energy when compared with all the three protocols. This is because the nodes calculate the timeout with selection criteria which results in balanced virtual backbone.

Table 6-1 shows the convergence time for all the four protocols. The convergence time is higher for EECDS and CDS-Rule K due to two phase topology construction. On the other hand, A3 and A1 protocol has less convergence time due to a single phase construction of the topology.

Table 6-1: Convergence Time (Sec)

Network Size	EECDS	CDS-Rule K	A3	A1
50	145.50	89.19	39.34	41.77
100	145.13	102.17	34.31	39.81
150	144.77	114.73	34.04	39.57
200	144.94	127.36	33.90	39.40
250	144.71	140.29	34.21	39.33

6.5.2 Indoor Topologies

Figure 6-10 shows the message overhead for all the four protocols under indoor deployment environments. The A1 protocol incurs fewer message overhead by constructing the topology with less energy overhead (see Figure 6-11). As the nodes are at equal distances in case of grid environments, nodes only calculate the timeout according to remaining energies of the nodes which results in better residual energy resources for A1 protocol (see Figure 6-12).

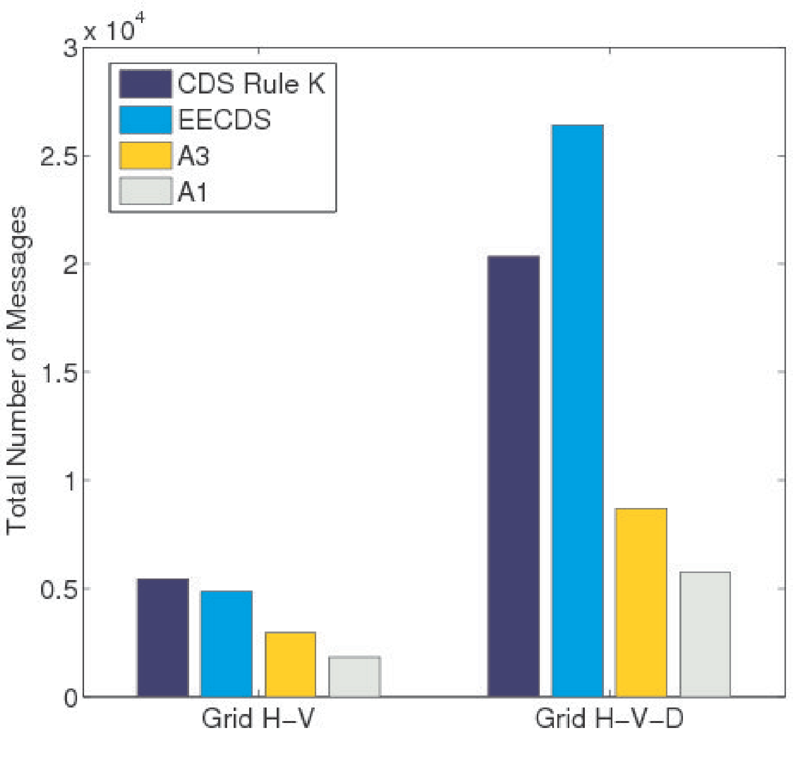


Figure 6-10: Message overhead under Grid H-V and Grid H-V-D topologies.

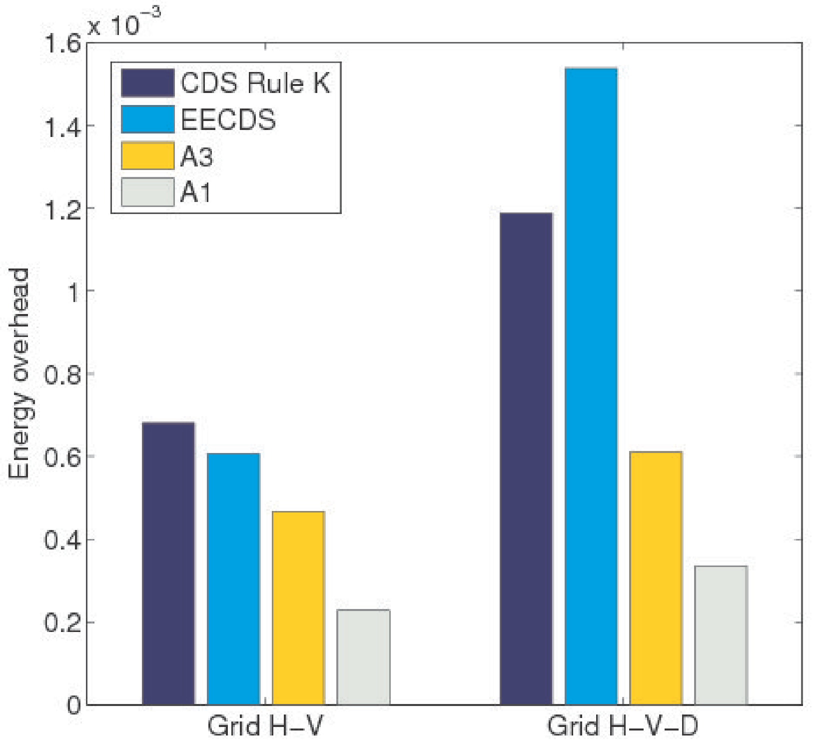


Figure 6-11: Energy overhead under Grid H-V and Grid H-V-D topologies.

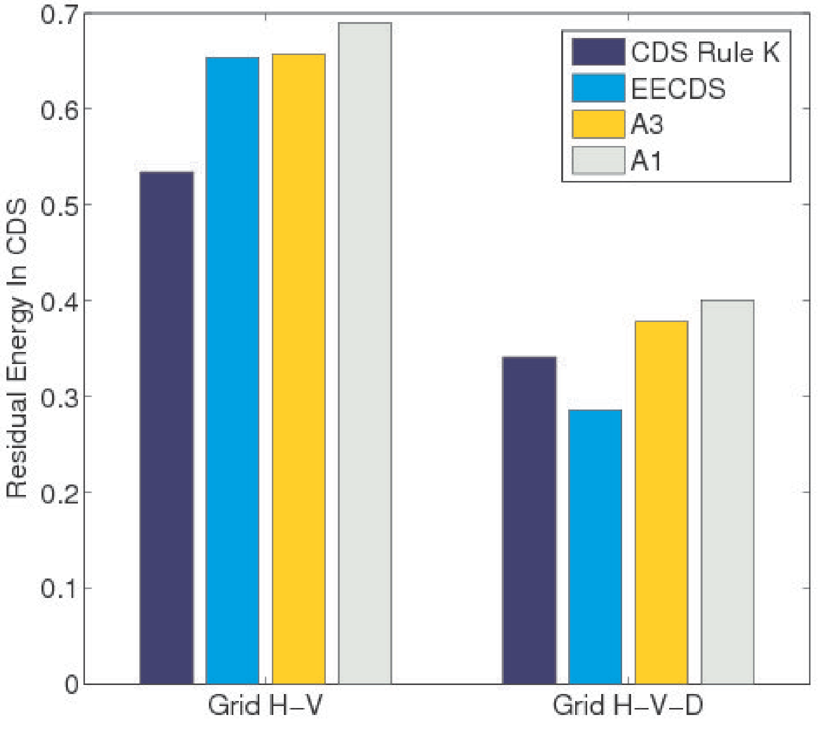


Figure 6-12: Residual Energy under Grid H-V and Grid H-V-D topologies.

This is also true for A3 protocol but it fails to perform better due to three way message handshakes for the construction of the topology which results in high message and energy overhead. Similarly, formation of Maximal Independent Set (MIS) and the formation of CDS in EECDS contribute to large number of exchanged messages as the network size is increased. Moreover, CDS-Rule K uses a pruning process in which every node updates its two hop neighbors when it is not marked and the process gradually increases as the node density is changed. This result in high energy overhead and less residual energy as shown in Figure 6-11 and Figure 6-12 respectively.

6.5.3 Impact of Topology Maintenance

Figure 6-13 and Figure 6-14 shows the metric values of all the four protocols under dynamic topology maintenance.

The number of unconnected nodes increases with increase in the network size for all the four protocols. However, CDS-Rule K protocol results in large number of unconnected nodes as shown in Figure 6-13. In CDS-Rule K, nodes remained marked if there is at least one pair of unconnected neighbors. The energy depletion of the marked node leads to higher number of unconnected nodes as compared with the other three protocols. Moreover, it fails to provide better sensing coverage which decreases with the increase in the number of unconnected nodes (Figure 6-14). On the other hand, A3 has less number of unconnected nodes due to its node selection process based on signal strength metric and provides better sensing coverage. In comparison, A1 results in very less number of unconnected nodes which on the other hand provides better sensing coverage when compared with all the three protocols.

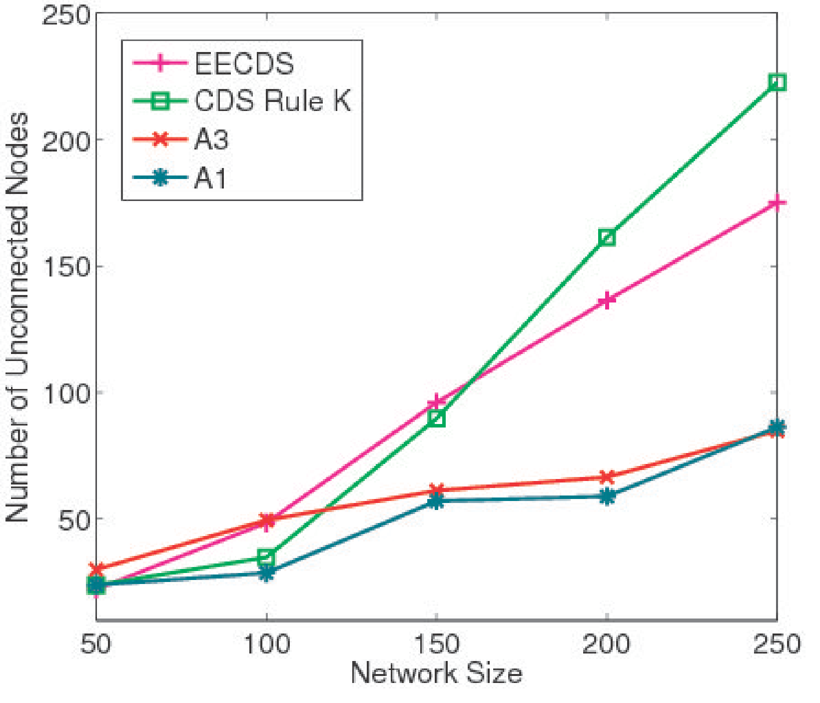


Figure 6-13: Number of Unconnected nodes under dynamic topology maintenance.

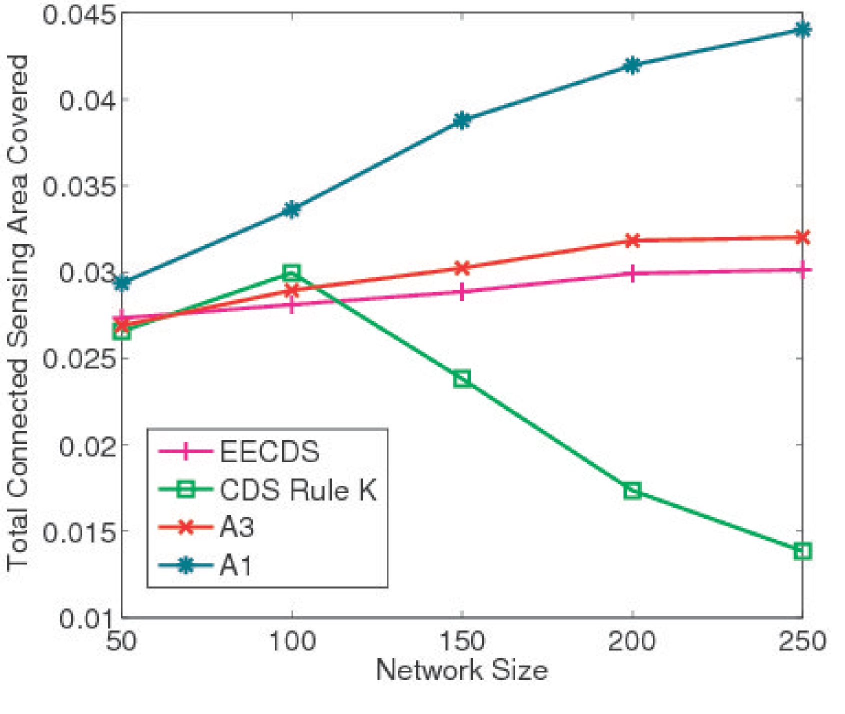


Figure 6-14: Connected Sensing Area covered under dynamic topology maintenance.

It is interesting to note that though the number of unconnected nodes increases in EECDS, it results in providing better sensing coverage as shown in Figure 6-14. This is due the reason that its two phase topology construction results in forming a proportionate CDS topology with more connected nodes covering the virtual area much better than CDS-Rule K protocol.

6.5.4 Impact of CDS size on Sensing Coverage

Network's delivery reliability is a critical parameter that measures the performance of the protocol. It is defined as the probability that the sensor nodes can communicate with each other with the increase in the node density. It is given by:

$$R(P_S, L) = (P_S)^L, \quad (6.7)$$

where P_S is the probability of success and L is average backbone path length (virtual backbone) in a CDS (defined as equation (4.5)). Hence, as L increases, the reliability that a packet will be successfully delivered decreases [114]. However, it was noticed that under topology maintenance operation, many nodes get disconnected from the network. Due to this reason, only few nodes remain connected with the sink node at the end of the topology maintenance (see Figure 6-13). This, on the other hand, computes a smaller average backbone path length. Now, if the reliability is modeled under fixed probability of success for such a topology, the network reliability appears to be high. On the other side, such a topology fails to provide better coverage.

To elaborate our findings, random topologies of network size were generated varying from 50 to 250 nodes for CDS (CDS-Rule K, EECDS and A3) protocols and compared them with the proposed A1 protocol. The average backbone path length computed for all the four protocols is shown in Table 6-2.

Table 6-2: Average Backbone Path length (L)

Network Size	EECDS	CDS-Rule K	A3	A1
50	3.00	3.00	3.00	3.33
100	3.00	3.33	3.00	3.66
150	3.33	2.66	3.33	4.33
200	3.33	2.00	3.33	4.66
250	3.33	1.66	3.66	5.33

The results reveal that CDS-Rule K protocol provides very small values for L under varying network size, which on the other hand should provide better reliability according to equation (6.7). However, Figure 6-14 shows that the protocol fails to provide better sensing coverage and has more number of unconnected nodes. This observation is also true for EECDS and A3 protocols. On the other hand, L is greater for A1 protocol but it provides better sensing coverage under varying network sizes. Hence, reducing the average backbone path length compromises the coverage region of the protocols. Therefore, size of a CDS should be accounted under topology maintenance while considering coverage area in order to have a better sensing coverage. The A1 protocol forms the reduced topology without any metric desired for the reduction in the size of the CDS. Due to this reason, the number of unconnected nodes as shown in Figure 6-13 increases in much slower proportion which on the other hand provides better sensing coverage.

6.6 Summary

In this chapter, the problem of constructing a CDS which provides better sensing coverage in an energy efficient manner is investigated. The observations reveal that single phase topology construction with fewer number of messages lead towards an efficient protocol. Due to this reason, A1 outperforms other protocols by using far less messages for topology construction. To validate the results, simulations are performed over a large operational spectrum to compare with

EECDS, CDS-Rule K, and A3 protocols. The results show that A1 has low message complexity and incurs less energy consumption. Moreover, it covers more sensing area under its coverage region and has better connectivity characteristics when tested under topology maintenance operation. Therefore, topology maintenance should also be considered for topology construction protocols.

Chapter 7

Poly: A Reliable and Energy Efficient Topology Control Protocol

7.1 Introduction

In the context of sensor topology control for routing and dissemination, Connected Dominating Set (CDS) based techniques proposed in prior literature provide the most promising efficiency as explained in previous chapters. In a CDS-based topology control technique, a backbone -- comprising a set of highly connected nodes -- is formed which allows communication between any arbitrary pair of nodes in the network. For this purpose, the distributed clique connected dominating set (CCDS) protocol and the A1 protocol were presented in the previous chapters. However, CDS protocols do not provide any reliability as the CDS based protocols forms a CDS tree. Therefore, in this chapter, the important requirements for mission-critical wireless sensor networks are explained.

Wireless Sensor Networks (WSNs) are envisioned as an enabling technology for a broad class of mission-critical applications. It is generally assumed that nodes in a WSN are connected to their neighbors with a certain probability of packet loss. Since wireless links are inherently unreliable, these packet losses are not acceptable for many mission critical WSN applications (e.g., forest fire detection, battle field monitoring) which require the network topology to provide a certain desired level of reliability. This reliability should however be achieved while keeping in mind the fundamental energy consumption constraint of a WSN. In this context, the graph-theoretic Connected Dominating Set (CDS) principle has emerged as the most popular method for energy-efficient topology control (TC) in WSNs.

In CDS based TC schemes, some nodes are a part of the virtual backbone which is responsible for relaying packets in the WSN. Non-CDS nodes conserve energy by turning off their transceivers. CDS size is a critical parameter which controls the compromise between reliability and energy efficiency. For instance, for small CDSs, fewer nodes handle the bulk of the network traffic and consequently deplete their batteries quickly [53]. The positive side of a small CDS is that more nodes can go to sleep mode. While both of these metrics – energy efficiency and reliability – are equally important for mission critical WSNs, existing CDS-based routing protocols cannot simultaneously cater both metrics [53] - [55].

In our earlier work, the performance of maintaining a cycle in a 10 node network was analyzed [104]. However, to understand practical limitations, it is important to analyze the performance on larger networks against other widely available protocols. Therefore, in this chapter, a semi distributed graph-theoretic topology control protocol for wireless sensor networks is proposed. The protocol, referred to as the Poly protocol, models the network as a connected graph and finds the number of polygons present in the network. Based on the duplicate node IDs of different nodes, Poly adaptively finds a polygenic backbone to turn-off the unnecessary nodes while keeping the network connected and covered.

To achieve energy efficiency, the protocol forms a CDS like polygenic network which in turn provides reliability in the case of random link failures. Moreover, it adapts to topological changes in the network based on the remaining energy of the nodes. This allows topology maintenance among different set of nodes to increase the network lifetime.

The Poly protocol has a low message complexity which allows the protocol to run multiple times during topology construction and maintenance phases. It can also be applied to different data reporting models which aim to find rendezvous point's (RPs) and can provide polygenic redundancy to RPs [103]. In addition, Poly achieves energy efficiency while considering network reliability.

The protocol is compared through simulations with A3, Energy Efficient CDS (EECDS) and CDS-Rule K protocols. Simulations are performed under different underlying topologies, varying node densities to analyze message overhead, energy overhead, residual energy and network connectivity. Simulation results show that the proposed Poly protocol has low energy overhead and it has 19% better residual energy when compared with CDS-Rule K protocol. Similarly, it has 32% and 34% better residual energy versus EECDS and A3 protocol while performing better under topology maintenance techniques. In addition, the results also demonstrate that Poly is more reliable under varying link probabilities than rest of the competitor protocols.

The rest of this chapter is organized as follows. At first, the background and related work in this area is explained which is then followed by the description of the Poly protocol. Later, the empirical evaluation framework utilized for the performance analysis of Poly is explained followed by the discussion on simulation results. In the end, the reliability analysis of Poly is presented and the salient findings of Poly are summarized.

7.2 Background and Related Work

In the next section, some related topology construction protocols are explained with the topology maintenance technique which is later utilized to evaluate the performances of different topology construction protocols.

7.2.1 Topology Construction & Topology Maintenance

As explained previously that DS based protocols have led to the concept of Connected Dominating Set (CDS) based topology construction protocols for the generation of energy efficient topology in WSN's. However, studies demonstrate [56]-[58] that CDS backbones are more vulnerable to node and link failures in WSN's. To this end, two approximation algorithms are proposed -- Connecting Dominating Set Augmentation (CDSA) and k-connected m-dominating set (k, mCDS) -- to construct a k-connected virtual backbone which can accommodate the failure of one wireless node. However, the studies do not analyze the impact of having k-connected virtual backbone on the energy efficiency of the network which is considered in the proposed protocol presented in this chapter.

Topology maintenance procedures may also be classified on the basis of time and energy triggering mechanisms. In time triggered methods, topology is rebuilt after a specific period of time. However, these mechanisms are generally expensive in terms of message and energy overhead. Therefore, in this chapter, focus remains on topology maintenance based on energy thresholds with dynamic topology maintenance.

The working of the Poly protocol is explained in the next section which is then followed by the complexity analysis of Poly.

7.3 The Poly Protocol

Due to the focus on mission-critical applications, two fundamental design constraints that are imposed on a topology construction protocol are: 1) its resultant topology should provide a desired level of packet delivery reliability, and 2) its energy efficiency should be comparable to or more than existing CDS-based topology construction protocols. To satisfy these constraints, the Poly protocol arranges the nodes in such a way that they form a closed path among a set of nodes. The closed path provides a reliable and energy efficient topology because: 1) the sink node gets polygenic redundancy with its neighbors which allows the nodes to use an alternative path in case of random link failures, and 2) it forms an active node set – nodes comprising a polygon – allowing leaf nodes to enter into the dormant / sleep mode. An additional advantage of polygenic is that the topology construction protocol does not need the position or orientation information of the nodes.

7.3.1 Description of control messages

The Poly protocol uses three types of messages which are involved in the polygon formation process. A hello message which contains the parent ID of the sender. A finish *discovery* message which is used by the parent node to announce the end of the topology discovery process. In the finish discovery message, each node sends a list of its discovered neighbors. Finally, a create topology message containing the IDs of active node set is propagated in the network.

7.3.2 Topology construction protocol

Topology construction phase of the Poly protocol is divided into three phases. In the first phase, a CDS is created during which the nodes discover their neighbors. The neighbor discovery process is initiated by a pre-defined node (e.g., the sink node) and terminates at the leaf nodes. In the second phase, each leaf node sends its neighbor list through the upstream neighbor -- the so-called parent node -- to the sink node. In the third phase, the sink node discovers polygons in the graph.

Subsequently, the polygon nodes are informed that they are part of the active node set. In this way, a closed path is formed with connecting paths to the branches.

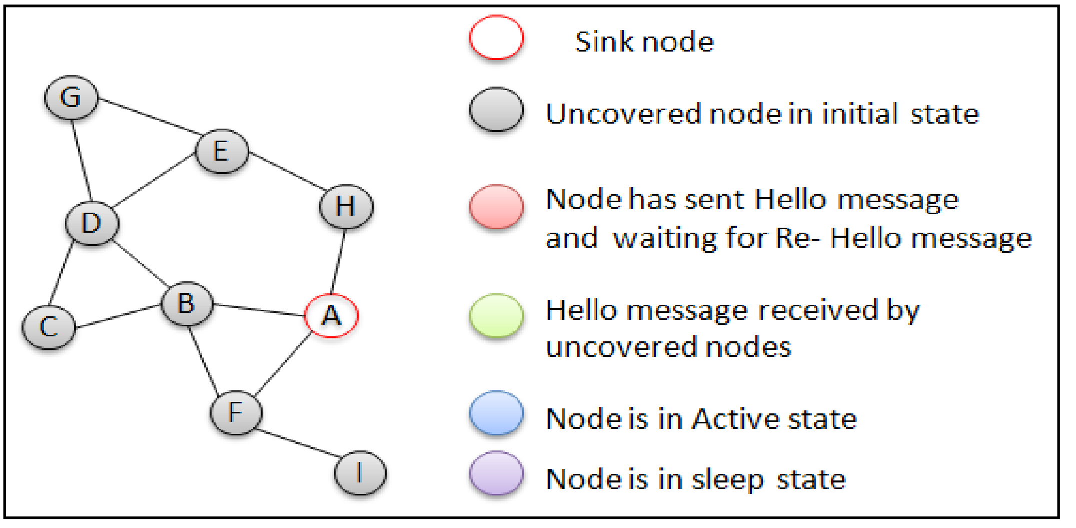


Figure 7-1: A sample Network.

The formation of polygon is described with the help of an example network shown in Figure 7-1. The topology construction process in Poly is initiated by a node called an initiator node. For protocol implementation, a random node is selected as an initiator node with the criterion that, if more than one node initiates this process, the node with the largest ID will persist. The initiator node A broadcasts a *hello* message and starts a timer to receive a *hello* response from its children (see Figure 7-2). As described earlier, the *hello* message contains the parent ID of the sending node. In the case of the initiator node, this field is empty. The *hello* of node A is received by B, F and H nodes located within its transmission radius. These nodes are *uncovered* nodes which mean that they are in the initial state and have not yet chosen any parent node. Therefore, nodes B, F and H – after receiving the message – choose A as their parent node. The uncovered nodes further rebroadcast the *hello* message to discover their children, and also start their respective timers to receive their children nodes' responses.

Figure 7-2: Sink node, A, broadcasts *hello* message – received by nodes B, F and H - and sets a timeout to receive Hello message in response from its children. Nodes B, F and H recognize sender node A as their parent.

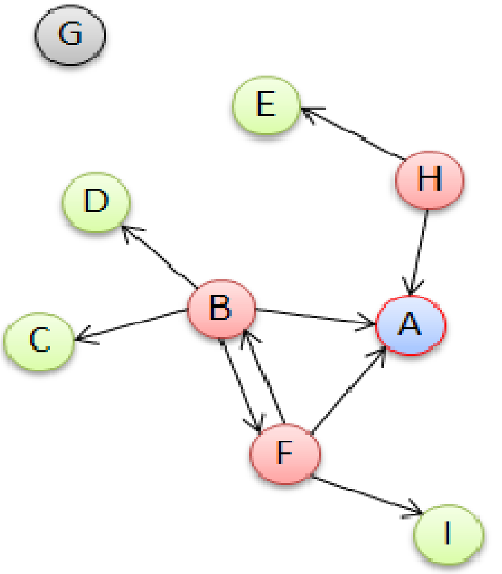


Figure 7-3: B, H and F further broadcast *hello* message with parent ID set to A. Covered nodes, B and F, also recognizes one another as neighbors. When a node recognizes its children, it waits for *Finish discovery* message. Therefore, A is now waiting for Finish message from B, F and H.

Every rebroadcasting node, before forwarding the message, updates the parent ID field by replacing it with its own parent ID; for instance, nodes B, F and H in Figure

7-3 update the parent field to the ID of node A. The rebroadcast *hello* message is also received by the parent node A. Consequently, node A identifies the sender of the *hello* message as one of its children. Once identified, nodes B, F and H are considered as *covered* nodes. Furthermore, when a node identifies a child node, it switches to an active state and starts to wait for *finish discovery* message from the children. When the *hello* message is received by a covered non-parent node, the receiving node identifies the sender as one of its neighbors. For instance, in the given example, the *hello* message from B is also received by node F – a nonparent node – leading to the identification of node B as a neighbor of node F. In this manner, the nodes discover their neighbors during CDS creation and the process is repeated until the network is completely covered (see Figure 7-4).

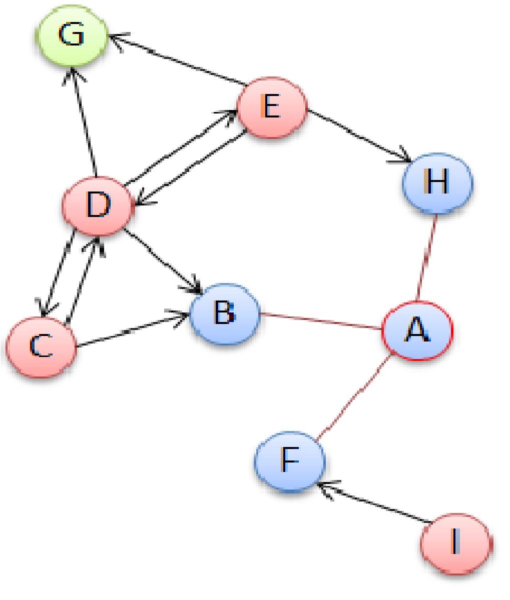


Figure 7-4: Next level nodes again broadcast the Hello message after changing the parent IDs to their respective parent.

The rebroadcast of *hello* messages continues until they reach the *leaf nodes* e.g. nodes C, D and I in Figure 7-5. The leaf nodes follow the same process but their timeout expires as these nodes do not have any child node. When timeout expires at leaf nodes, they send *finish discovery* messages to their parent nodes thereby initiating the second phase of the protocol as shown in Figure 7-5. After sending a *finish discovery* message, a leaf node enters the sleep mode and turns off its transceivers to conserve energy.

Figure 7-5: G broadcasts *hello* message. Timeout for *hello* message from children expires at C, D and I. Therefore, these nodes consider them as leaf nodes.

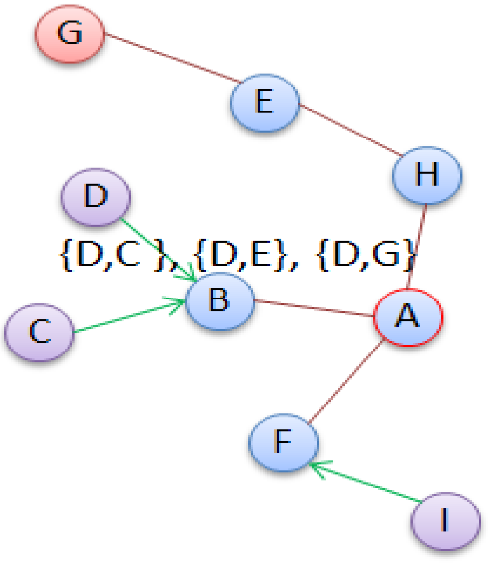


Figure 7-6: Each leaf node sends *finish discovery* message to its parent node and immediate neighbor after the timeout. These sets of nodes are called message paths. Message paths sent by D are shown in the figure.

Note that the neighbors of node A do not send any explicit response. Instead, they simply rebroadcast the *hello* message which also functions as a response message for node A. Consequently, nodes avoid the use of any explicit response message and reduce the number of control messages exchanged during the topology discovery.

In the *finish discovery* message, each node sends the list of its neighbors to its parent node. In the example, node C and node E are neighbors of node D. Therefore, node D sends {D, C}, {D, E}, {D, G} to its parent (node B) in the *finish discovery* message (see Figure 7-6).

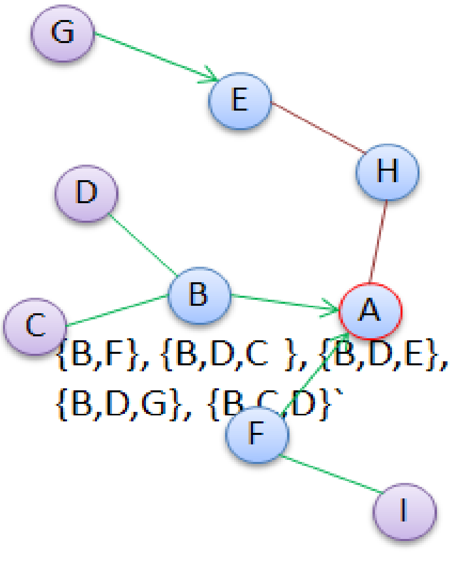


Figure 7-7: Node B extends the message paths with its own ID and sends it to its parent node. Message paths sent by B are shown in the figure.

These sets of nodes create message paths which can then be used for polygon formation among a set of nodes. A node waits for the *finish discovery* messages from all its children. When it receives all the expected messages, it creates its own *finish discovery* message and forwards it to its parent node. Node B in Figure 7-7 extends the received set of message paths with its own message ID. In this way, *finish discovery* message converges towards the sink node.

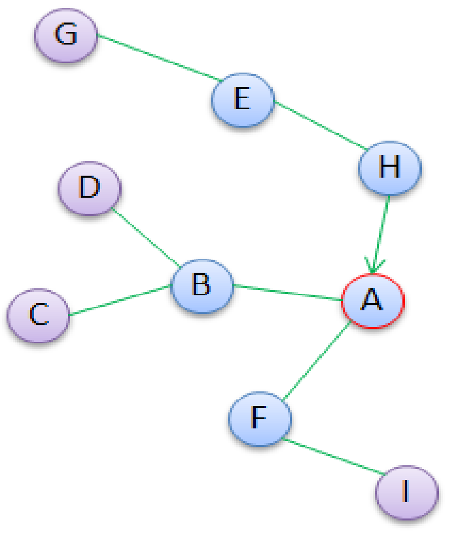


Figure 7-8: After receiving *finish discovery* message from all children, sink node adds it own ID to message paths and figures out a polygon.

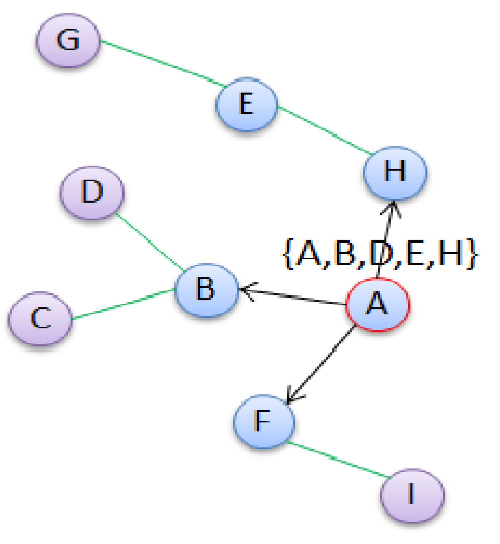


Figure 7-9: Sink node then broadcasts the *create topology* message for the chosen polygon. Nodes in the polygon set turn them as active nodes.

When the sink node has received *finish discovery* message from all of its children, it combines different message paths to discover polygons in the network. If there are at least two common nodes in the two message paths, then a polygon exists in the network. For instance, {A, B, D, G} and {A, H, E, D} are two message paths and have two common nodes A and D. Hence, it can be inferred that there exists a polygon in the network comprising nodes {A, B, D, E, H, A} (see Figure 7-8). Figure 7-9 shows the selected polygon by a sink node. Once selected, sink node broadcasts the *create topology* message which contains the list of nodes that are part of the polygon.

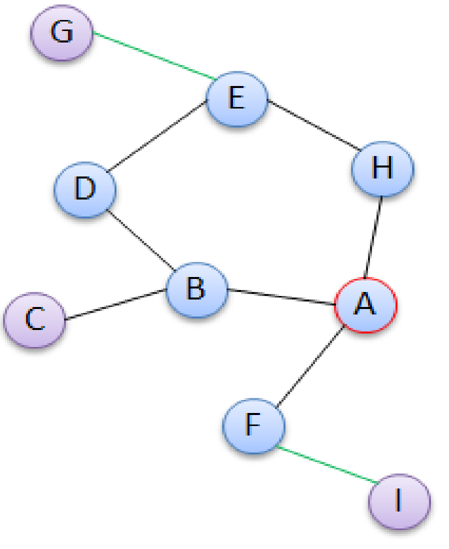


Figure 7-10: Final topology polygon with redundant paths.

The size of polygon is dependent on the needed level of reliability as a large size allows more nodes to be connected with the sink node. On the other hand, a small size provides minimum reliability among nodes forming a polygon. When a node receives the *create topology* message from a sink node, it searches its ID in the polygon. If it finds its ID in the polygon, it marks itself as an active node. At the end of this process, each node is either in *active* state or in *sleep* state as shown in Figure 7-10. The active nodes form a backbone responsible for data communication in the whole network.

7.3.3 Poly: Complexity analysis

All the protocols analyzed and proposed in this report start by undertaking a CDS discovery process. The complexity of the CDS discovery process is the same for the three existing protocols considered in this chapter. However, the Poly protocol has lower CDS discovery message complexity because, instead of exchanging explicit messages, it uses wireless broadcasts for parent discovery. After CDS discovery, existing protocols simply use the discovered CDS for topology construction and therefore do not incur any additional overhead. On the other hand, to achieve a higher level of reliability, the proposed Poly protocol introduces some additional complexity after CDS discovery. To reduce this additional complexity, all the cycles in a network are not found, but instead subsets of cycles are discovered. The sink node, therefore, processes a reduced subset of message paths and fewer cycles. The concept of cycle merging is also utilized in which smaller cycles are combined to form a larger cycle. This further reduces the protocol's complexity as explained below.

Assume that there are total of N message paths after the CDS discovery phase. When sink node receives *finish discovery* message from all child nodes, it starts the process of discovering a cycle. In order to find the protocol complexity, in the worse case, N^2 message paths of length L are needed to be compared, in order to find at least two common nodes in each iteration. Consequently, the worst case complexity of the searching algorithm is $O((NL)^2)$. However, all the N message paths are not handled. Rather, the search space is reduced significantly by handling a subset of message paths. If d is the divisor which is used to reduce the number of message paths, the complexity will be $O(NL/d)^2)$.

As the cycles are discovered, they are compared and merged to find larger cycle based on the needed level of reliability. So, in essence, the complexity of Poly is far less than $O(N^2)$. For storage purposes, among N branches of length L, NL/d nodes are needed to be saved at the sink node which reduces the complexity to the order $O(N)$. If the reliability constraint is discounted, then the Poly protocol's complexity is equal to or lower than the existing protocols. Thus, the additional complexity of the Poly protocol represents a tradeoff between reliability and energy efficiency.

This completes the description and analysis of the Poly protocol. The experimental setup which is used for the evaluation of the Poly protocol is now presented in the next section. It is then followed by a detailed discussion on simulation results.

7.4 Empirical Evaluation Framework

In this section, the empirical evaluation framework is presented which is utilized for the evaluation of the Poly protocol and three other CDS protocols, namely A3, EECDS, and CDS-Rule K. The empirical setup explains the description of various network topologies and simulation parameters. Moreover, the simulation results are also discussed.

7.4.1 Simulation Setup

To evaluate the protocols under consideration, the Atarraya simulator is used which has been designed specifically for WSN topology control protocols [96]. In our experiments, it is assumed that the sensor nodes are randomly deployed in an area of 600m × 600m. The experiments are performed in different network topologies ranging from 50 to 250 nodes. The protocols performance behavior remains same as the node density is further increased; therefore we only report results with node density varying up to 250 nodes. The transmission radius and initial energy level of each node are set to 42m and 1J, respectively. The nodes communicate with each other using full duplex wireless radios. The actuation energy equals $50\text{nJ}/\text{bit}$ while the communication energy is $100\text{PJ}/\text{bit}/\text{m}^2$. As described before, energy-based topology maintenance technique is only considered. To this end, the energy threshold is set to 10% i.e. topology maintenance process is triggered when the network energy falls by 10%. The message sizes in most of the cases do not exceed 25 bytes therefore data packet size of 25 bytes is used in the experiments. However, the A3 protocol also use children recognition message which contains the sorted list of the children. Therefore, it has a size of 100 bytes in the experiments. Moreover, an ideal Medium Access Control (MAC) layer is assumed; i.e. there is no packet loss due to channel contention / collisions.

For the experiments, different connected topologies were generated; however, the results do not deviate largely therefore the reported values of the selected metrics are averaged over 50 simulation runs. In static techniques, performance is mainly dependent on efficient topology construction. Therefore, the results for dynamic topology maintenance techniques based on energy-threshold are reported. Finally, it is reemphasized that size of the polygon in the Poly protocol is a critical parameter, characterizing the tradeoff between reliability and energy efficiency. Therefore, in all our experiments, the size of polygon is varied approximately between 10 to 50 nodes with the increase in the network size.

7.4.2 Simulation Results

Simulation results are described in three subsections. First, all the four protocols are evaluated in two ideal grid environments observed in controlled indoor deployments: the Grid H-V and the Grid H-V-D topologies. In the Grid HV topology, nodes can communicate with their horizontal and vertical neighbors, while in the Grid H-V-D topology; nodes can communicate with their diagonal neighbors as well. Subsequently, the protocols' performances under varying node densities are compared while assuming that: (1) the nodes are randomly deployed, and (2) the protocols only construct the topology. The performance of the protocols is then discussed under a dynamic topology maintenance technique triggered by energy thresholds. In the next section, the reliability of these protocols is modeled and compared.

7.4.2.1 Grid Topology

In the case of Grid H-V topology, a network of 169 nodes was assumed while restricting the transmission range to 28m. For Grid H-V-D, the network size was increased to 324 nodes. Due to the fact that nodes in grid topologies are at some predefined positions therefore in order to always have a connected deployment scenario, the network size of 169 and 324 was chosen. The message overhead, energy overhead and residual energy results are shown in Figure 7-11, Figure 7-12, and Figure 7-13 respectively.

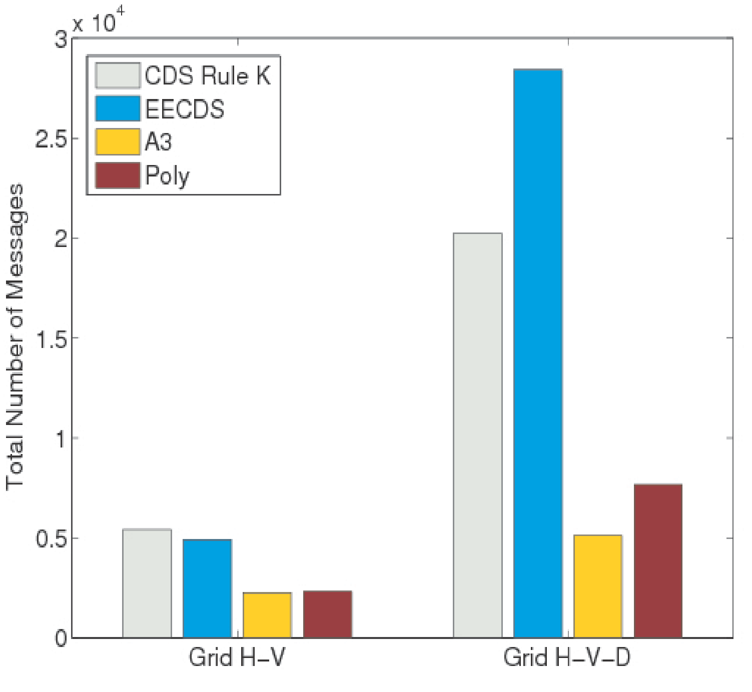


Figure 7-11: Message overhead.

The message and energy overhead of EECDS and CDS-Rule K protocol increases due to the two phase topology creation mechanism used by both protocols. On the other hand, A3 has low message overhead due to its three-way handshake process which allows nodes to have less energy overhead to form a reduced topology. The proposed Poly protocol has low energy overhead despite the fact that its message overhead is greater than A3 protocol. This is because A3 uses a selection metric based on signal strength which allows distant nodes to be selected in the CDS.

However, in grid topologies the neighbors of the sink node are at equal distances which introduces more energy overhead for the A3 protocol.

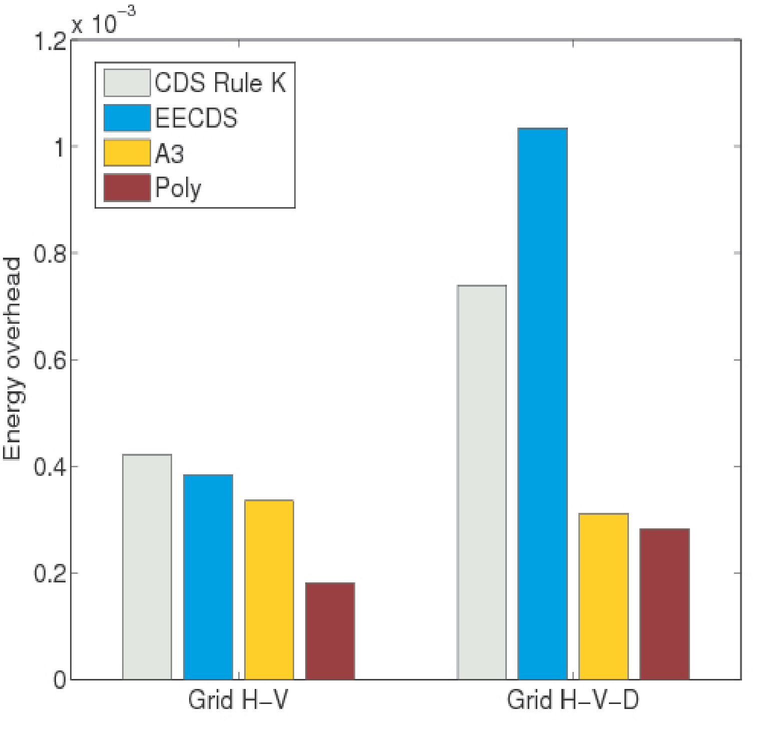


Figure 7-12: Energy overhead.

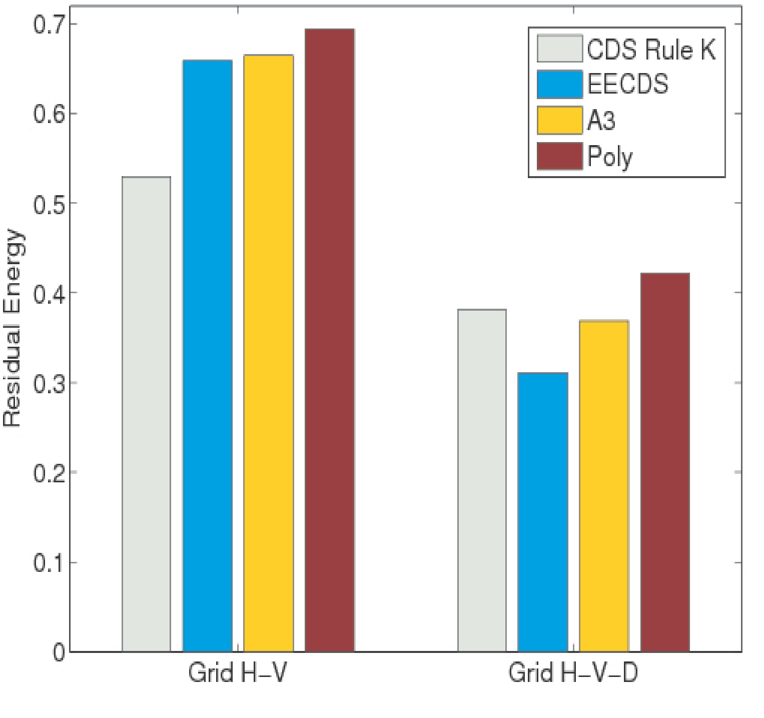


Figure 7-13: Residual Energy.

On the other hand, the Poly protocol uses a broadcast mechanism to select nodes in proportion with the size of the network, hence yielding better residual energy as compared to CDS-Rule K, EECDS, and A3 Protocols.

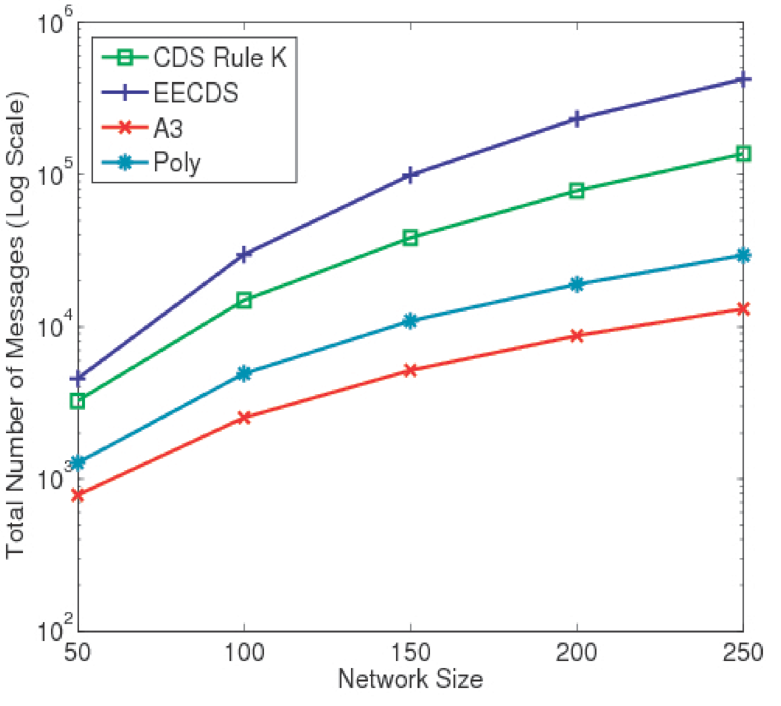


Figure 7-14: Message overhead when node densities are varied.

7.4.2.2 Impact of Node Density

Figure 7-14, Figure 7-15, and Figure 7-16 shows the message overhead, energy overhead and residual energy of all the four protocols respectively. As the network size grows, the number of exchanged messages rises exponentially for all the four protocols. The increase in the node density results in a proportional increase in the node degree which ultimately leads to an increase in the number of messages exchanged. This trend is noticeable in the results shown in Figure 7-14. Number of exchanged messages for EECDS and CDS-Rule K is significantly higher than the Poly protocol. This is caused by the two-phase topology construction process utilized by EECDS and CDS-Rule K protocols. In comparison, A3 generates fewer messages because it chooses the distant nodes using signal strength. This allows fewer nodes to become part of the CDS, thus leading to quick convergence of the protocol. On the other hand, the Poly protocol forms a polygon in which all the nodes send their IDs back to the parent node. This process incurs higher message overhead than the A3 protocol.

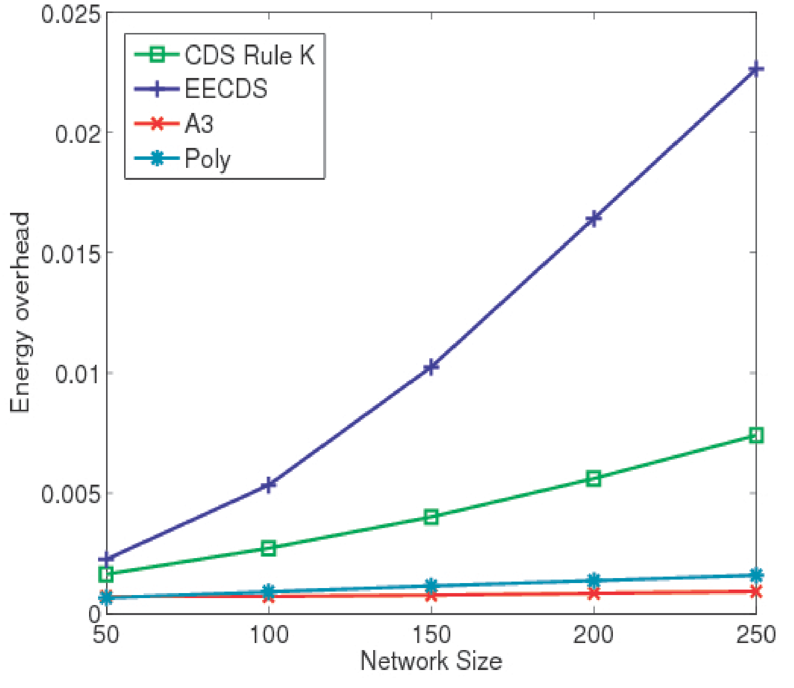


Figure 7-15: Energy overhead when node densities are varied.

Energy overheads of EECDS and CDS-Rule K are significantly higher than A3 and Poly, as shown in Figure 7-15. As can be intuitively argued, an increasing node density leads to higher energy consumption due to an increase in the number of received packets. However, Poly still has lower energy consumption due to its rebroadcast strategy for topology discovery. As mentioned earlier, Poly does not use any messages explicitly sent to a parent node by its children. Instead, it overhears the broadcast at the parent node to get aware of its children. A3 protocol has low message overhead when compared with all the other three protocols. The energy overhead curve flattens for A3 and Poly protocols because both protocols do not use a two-phase strategy like EECDS and CDS-Rule K. Figure 7-16 shows the residual energy of all the four protocols. Usually, high energy overheads lead to lower residual energies. However, it was observed that A3 which has low message and energy overhead has significantly less residual energy. This is due to non-uniform distribution of communication overhead which drains the battery of fewer nodes resulting in lower residual energy levels among nodes in the network. On the other hand, EECDS and CDS-Rule K protocols have less residual energy due to high energy overhead. Poly provides better residual energy when compared with all the three protocols. This is because: 1) It forms the active node set in proportion to the network size, and 2) it uses a rebroadcast mechanism which inherently consumes battery of nodes at an equal rate.

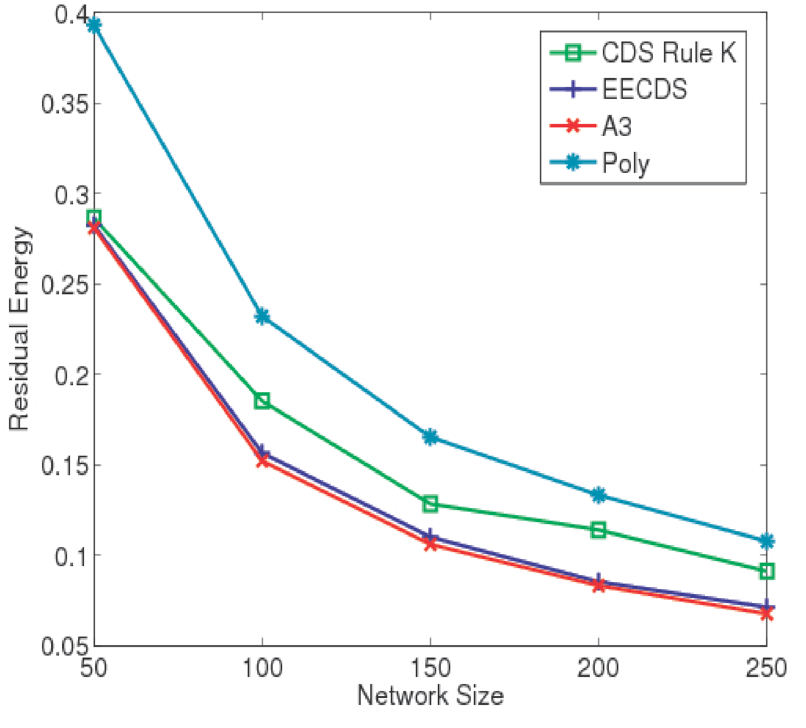


Figure 7-16: Residual energy when node densities are varied.

7.4.2.3 Dynamic Topology Maintenance

Figure 7-17, Figure 7-18, Figure 7-19, and Figure 7-20 shows the metric values of all the four protocols under dynamic topology maintenance. Formation of Maximal Independent Set (MIS) and the formation of CDS in EECDS contribute to large number of exchanged messages as the network size is increased. This is shown in Figure 7-17. However, the number of exchanged messages decreases slightly in case of CDS-Rule K protocol. This is due to less number of connected nodes as the node density is increased (see Figure 7-20).

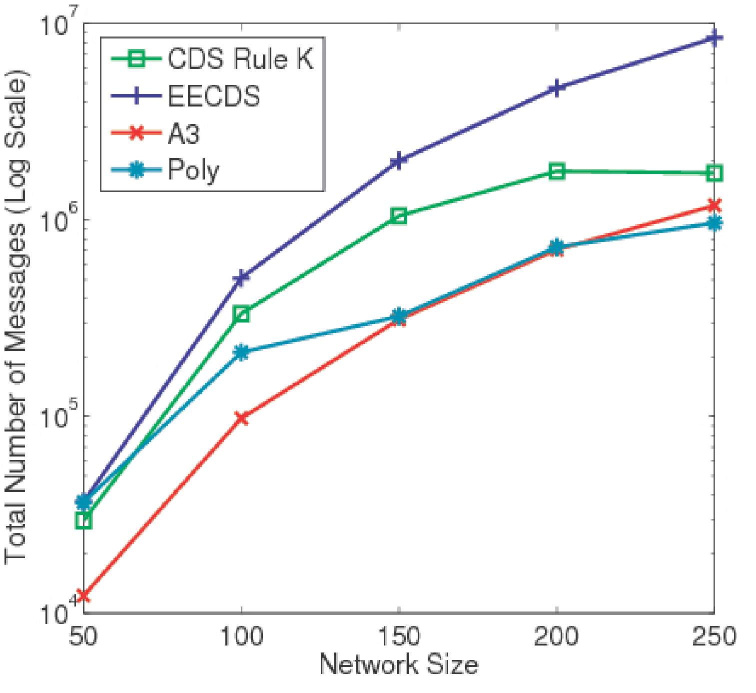


Figure 7-17: Message overhead comparison under dynamic topology maintenance.

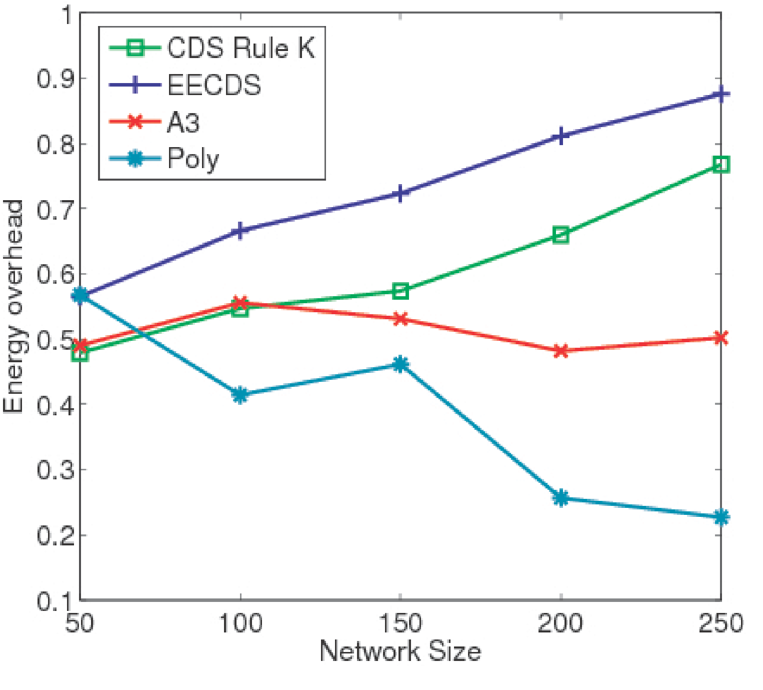


Figure 7-18: Energy overhead comparison under dynamic topology maintenance.

For A3 and Poly protocols, the number of exchanged messages increases exponentially due to higher number of connected nodes. Similarly, consumed energy also increases linearly in case of EECDS and CDS-Rule K protocols (Figure 7-18). However, EECDS allows uniform distribution of energy resources which results in better residual energy (Figure 7-19).

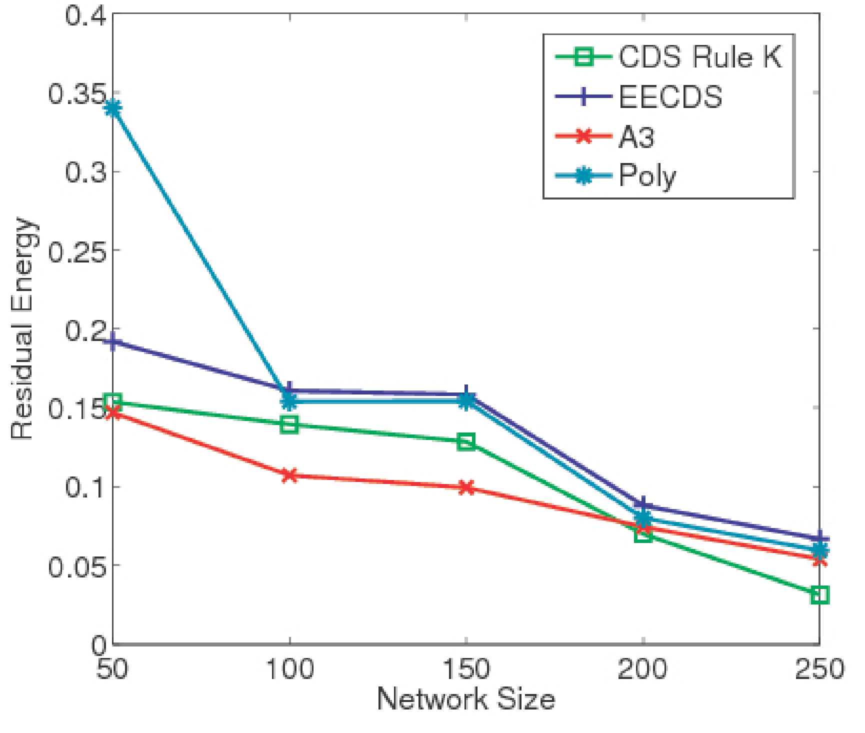


Figure 7-19: Residual energy comparison under dynamic topology maintenance.

On the other hand, CDS-Rule K uses a pruning process in which every node updates its two hop neighbors when it is not marked and the process gradually increases as the node density is changed. Therefore, CDS-Rule K has less residual energy as shown in Figure 7-19. A3 protocol shows consistent behavior in terms of energy overhead. However, it has less residual energy due to its three way message exchange and distant node selection metric (see Figure 7-19).

The energy overhead of Poly decreases with the increase in the number of connected nodes. However, it has better residual energy when compared with all the three protocols as shown in Figure 7-19 for the reasons mentioned in the previous subsection.

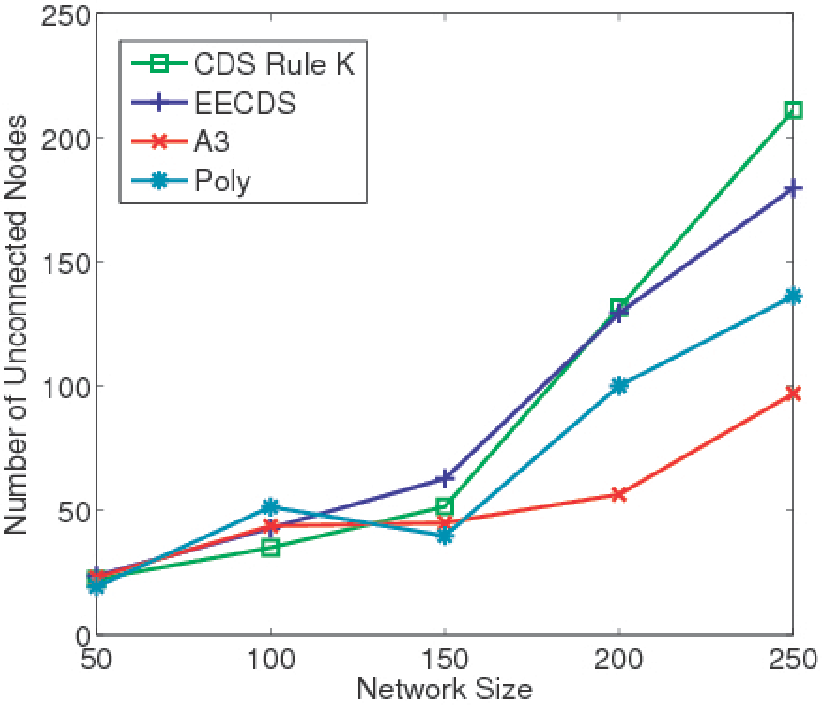


Figure 7-20: Connectivity comparison under dynamic topology maintenance.

Figure 7-20 shows the number of unconnected nodes under dynamic topology maintenance. In CDS-Rule K, nodes remained marked if there is at least one pair of unconnected neighbors. The energy depletion of the marked node leads to higher number of unconnected nodes as compared with the other three protocols. On the other hand, A3 has less number of unconnected nodes due to its node selection process based on signal strength metric. Performance of Poly in larger networks is better than EECDS and CDS-Rule K. However, it is alarming to note that the number of unconnected nodes increases as the network size gets bigger.

Before the chapter is concluded, in the following section, the reliability of all the four protocols is analyzed.

7.5 Network Reliability

In graph theory, redundancy is defined as *the expected number of functional spanning trees in a graph*. Removal of an edge from a spanning tree partitions the graph. Therefore, every edge may be considered as a bridge in a spanning tree. The number of spanning trees measures the network performance under highly dynamic conditions, e.g. frequent link failures. Reliability, on the other hand, is the probability that there is at least one spanning tree or the probability that the sensor nodes can communicate with each other in case of random link failures. Hence, reliability is another critical parameter that measures the redundancy of the protocol [98]. The linear algebra (linalg) package available in Maple [97] was used to analyze the performance of Poly protocol by assuming the network shown in Figure 7-1. The random topologies of different sizes were generated to calculate the reliability results which were later averaged. However, the polygon of any size is bi-connected and every edge is a bridge in a CDS. These topological properties lead to better network reliability on diverse topologies when compared with CDS protocols.

Let $A = (a_{ij})_{n.n}$ denote the adjacency matrix of graph G, then

$$a_{ij} = \begin{cases} 1 & \text{if vertex } v_i \text{ is connected to} \\ & \text{vertex } v_j \text{ by an edge,} \\ 0 & \text{otherwise.} \end{cases} \quad (7.8)$$

The degrees of the vertices are represented by a diagonal matrix. If $D = (d_{ij})_{n.n}$ denote the diagonal matrix of graph G, then

$$d_{ij} = \begin{cases} \deg(v_i), & \text{for } i = j, \\ 0 & \text{for } i \neq j. \end{cases} \quad (7.9)$$

The matrix tree theorem [98] was used to find the number of non-identical spanning trees for the network shown in Figure 7-1. According to the theorem, the spanning tree of graph G is the value of any cofactor of the matrix, i.e. $T = D - A$. Therefore, by using equation (7.8) and (7.9), the matrix T for the assumed network equals:

$$T = \begin{bmatrix} 3 & -1 & 0 & 0 & 0 & -1 & 0 & -1 & 0 \\ -1 & 4 & -1 & -1 & 0 & -1 & 0 & 0 & 0 \\ 0 & -1 & 2 & -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & -1 & -1 & 4 & -1 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & -1 & 3 & 0 & -1 & -1 & 0 \\ -1 & -1 & 0 & 0 & 0 & 3 & 0 & 0 & -1 \\ 0 & 0 & 0 & -1 & -1 & 0 & 2 & 0 & 0 \\ -1 & 0 & 0 & 0 & -1 & 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 1 \end{bmatrix} \qquad \textbf{(7.10)}$$

The cofactor of matrix T with any combination equals 108. Hence, there are 108 non-identical spanning trees which represent the total redundancy of the network. However, the interest is in measuring the probability that at least one of the spanning trees is working or the reliability that the network will be functional in events of random edge failures. To compute this, all spanning trees must be represented as a disjoint product as given below:

$$P(t_1 \cup t_2 \cup t_3 \cup \ldots\ldots \cup t_{108}) = P(t_1) + P(t_2\overline{t_1}) + P(t_3\overline{t_2 t_1}) + \ldots\ldots + P(t_{108}\overline{t_{107} t_{106}}\ldots\overline{t_1}), \qquad \textbf{(7.11)}$$

Where t is a spanning tree in the network.

The reliability for CDS (CDS-Rule K, EECDS and A3) protocols was computed and compared with the proposed Poly protocol. The adjacency matrix for CDS-based protocols remains the same. Therefore, all the three protocols have the same reliability. Consequently, these existing protocols maintain a CDS tree in which every edge serves as a bridge edge. If it's supposed that all edges have the same reliability $P_1 = P_2 = ... = P_n = P$ then, the reliability of the sample network shown in Figure 7-1 becomes $108P^8 - 315P^9 + 348P^{10} - 172P^{11} + 32P^{12}$, by using equation (7.11).

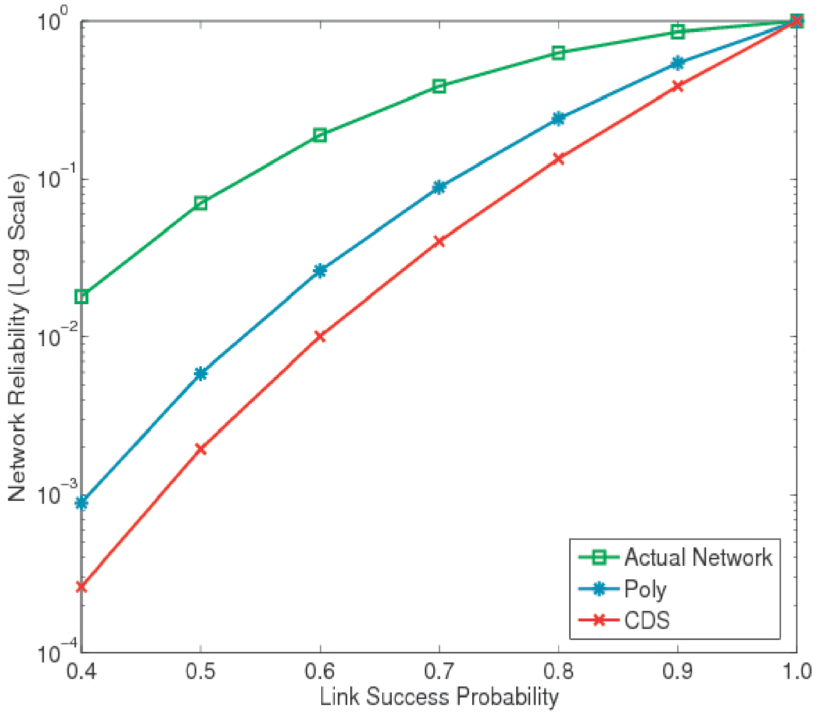


Figure 7-21: Network reliability of a sample network.

Figure 7-21 compares the reliability of poly protocol with CDS based protocols. The decrease in the link probability causes a proportional decrease in the network reliability. However, the Poly protocol provides better network reliability as the link probability is decreased. Existing CDS-based protocols have considerably lower network reliability because each edge (link) in these topologies serves as a bridge edge, and therefore does not provide any redundancy in the network.

The protocols performance behavior remains same as the node density is further increased. To validate this, the random topologies of network size varying from 50 to 250 were generated and averaged to get the reliability results. Figure 7-22 shows the averaged results on large diverse topologies. The results demonstrate that Poly provides better results when compared with existing CDS protocols.

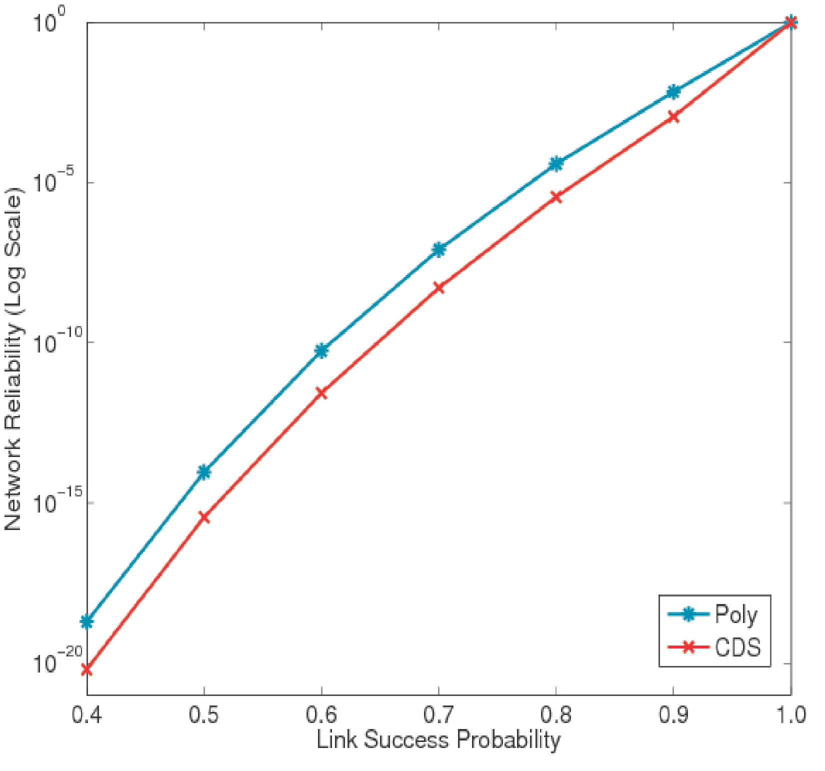


Figure 7-22: Network reliability of large diverse topologies.

7.6 Summary

In this chapter, a topology control protocol – Poly – which forms a CDS by finding polygons present in a WSN is proposed. The simulations were performed to compare the performance of Poly with other prominent topology construction protocols– CDS-Rule K, A3 and EECDS–over a large operational spectrum. Simulation results demonstrated that Poly has low message overhead and consumes less energy while providing higher network reliability. The Poly protocol also works well with dynamic topology maintenance techniques.

Chapter 8

Conclusion and Future Work

8.1 Conclusion

There has been a tremendous amount of research on self organizing Wireless Sensor Networks (WSNs) during the last few years. Due to this reason, in this report, two important yet different research problems for WSNs were addressed. First, a thorough review of the concepts used in wireless channel modeling was presented to understand the nature of the problem. Due to the cost constraint, it was observed that K^{th} order markov chains are used by most researchers to model a wireless channel and though many solutions exist for reducing the complexity associated with the markov chains, none of the solution accurately models a wireless channel and reduces the complexity to a desired state budget. In addition, most of the low complexity models need large number of parameters and calculate steady state and transition probabilities by an artificial mechanism which further questions the reliability of the models. Therefore, to understand and implement a low complexity wireless channel model which is parameterized from actual wireless traces, the binary nature of wireless traces was examined by collecting

802.15.4 and 802.11 traces. Traces collected over a wireless medium generally represent two states. One state is the good state and the other state is the bad state or the lossy state. Therefore, if the burst contains a number of zeros then it is referred as a *good burst* and if the burst contains a number of one's then it's called a *bad burst*. The trace can hence be represented as a combination of good and bad bursts.

In order to analyze behavior of actual traces, Crossbow's Micaz motes were used to collect residual bit-error traces over wireless sensor networks. Similarly, in case of 802.11, wireless stations with Linux boxes using Dlink wireless cards with Prism2 chipset device drivers were used. At first, experiments were performed by having a direct line of sight (LoS) between the sender and the base station but the error rates observed in those experiments were too low to warrant further analysis. Therefore, non-LoS traces were used. After analyzing the behavior of wireless traces, steady state and transition probabilities were calculated to model a markov chain. Similarly during investigation, it reveals that some bit patterns never occur in transition probability matrices and are referred to as unused states. In other words, such states result in all-zero columns of the transition probability matrix. The numbers of unused states are larger for the traces with lower BER because all the error states are not observed in a finite length trace. Many unused states in 802.15.4 WSN traces were observed and their number grew as the order of Markov chain and these states were further used to improve the accuracy of the model. While studying Markov chains, it was observed that FSM chains exhibit many interesting graph-theoretic properties which can be used to reduce the complexity associated with higher order Markov chains. Therefore, it was observed that with the help of those properties, a scalable model can be developed.

A graph in which each vertex or node is traversed exactly once forms a Hamiltonian circuit. A Hamiltonian circuit can be identified in Markov chains of any arbitrary order. Therefore, by using this property, the states can be further arranged according to the node traversed, which gives an easy method for aggregating states that comprise the Hamiltonian circuit. Since a cycle of arbitrary length can be identified in the Markov digraph, the states of the circuit can be aggregated to a desired state budget. Due to the reasons above, a Hamiltonian Model (HM) based on Hamiltonian circuits was presented to reduce the complexity. To analyze the performance of the model, a graph theoretic Bipartite Model (BM) was chosen and artificial traces were generated from both the models to quantify the similarity with the actual traces. In order to analyze the accuracy of both the models, Kullback-Leibler divergence measure was chosen, which quantifies the difference in the entropies of two probability distributions. The KLD divergence quantifies the source-coding-like overhead incurred by employing a model instead of the actual source. In addition, cumulative distribution function was used to measure the standard error which measures the error incurred by a model in approximating the actual CDF (calculated from actual traces). It was observed that HM outperforms BM. The traces generated from the HM closely follow the pattern in the actual traces and therefore it's concluded that HM can be used to model a wireless channel. To warrant further analysis, bit error rate statistics of the artificial traces were calculated and compared with the actual bit error statistics. It was observed that bit error statistics of HM lies in close range with the actual bit error statistics.

After highlighting the channel modeling problem, the report focuses on other existing challenges in Wireless Sensor Networks i.e. energy and bandwidth constraint which poses unique challenges for new protocols and applications to be used in these networks. It is well known that the two main causes of energy consumption in WSNs are communication (bit transmission/reception) and local computation. Moreover, communication consumes approximately two orders of magnitude higher energy than computation, storage and sensing for a majority of common applications. Therefore, communication overhead must be minimized for energy-efficient operation of a sensor network. In order to reduce the energy consumption constraint, importance of topology control protocols was highlighted as TC protocols are commonly used in the reconstruction and organization of node parameters and modes of operation to modify the topology of the network. In addition, design considerations for topology control protocols were also explained with existing challenges that need consideration for an efficient topology control protocol. It was observed that message overhead must be considered in order to have an efficient protocol. Therefore, the message exchange process in the formation of the reduced topology must be reduced.

To address the message overhead problem, the methods and techniques used by most researchers to reduce the network topology were highlighted. Of all these methods, the emergence of familiar Connected Dominating Set (CDS) based technique has reincarnated the topology control protocols for WSNs. To analyze the performance of the proposed protocols, many CDS based protocols were implemented and analyzed to derive the guidelines for having an efficient topology control protocol. Similarly, performance comparison parameters were also explained by highlighting the importance of them in performance evaluation process.

Later in the report, efficient CDS based topology control protocol is presented which uses cliques to construct a CDS. The novelty of the protocol lies in the fact that the protocol uses a single phase operation in order to construct a CDS. Whereas, most of the techniques and protocols use two phase construction mechanism i.e. construct an MIS and then connect the nodes in an MIS to form a CDS. Moreover, the cliques connected dominating set protocol uses the inherent broadcast mechanism in order to reduce the message exchange and do not use any explicit response messages back to the nodes from which they receive the message. This allows the protocol to have less consumed energy and have more residual energy among set of nodes present in the network. The evaluation of the CCDS protocol with other protocols demonstrates that CCDS performs better than other protocols in terms of message overhead and residual energy. However, it was also explained that it is not necessary that a protocol constructing a topology efficiently will also provide better network lifetime. Therefore, in order to check the robustness of the protocol, topology maintenance must also be taken into account. For this purpose, prominent methods for topology maintenenance were also explained and it was observed that dynamic topology maintenenance provides more realistic results since it recreates the topology while considering the present state of the nodes in the network. Therefore, the CCDS protocol was evaluated using dynamic topology maintenance and it was observed that it provides better network lifetime and connectivity among a set of nodes in the network.

While implementing the CCDS, it was observed that reducing the topology compromises the sensing area covered by the nodes in the network. It is due to the reason that the CDS size is taken as a critical parameter to control the reliability and energy efficiency. For instance, for small CDS's, fewer nodes handle the bulk of the network traffic and consequently deplete their batteries quickly. The positive side of a small CDS is that it reduces average path length to a certain extent as the path length is dependent on the network size. Hence, as the network size grows, the average path length of CDS based protocols increases. Therefore, most of the protocols reduce the topology while having a small CDS to achieve packet delivery reliability as average backbone path length is decreased. However, it was observed that the reduced topology must also consider the covered sensing area, since it's the main task of the nodes working in a WSN. For this purpose, A1 protocol was presented which considers the sensing coverage by forming a proportionate backbone and tries to achieve better sensing coverage while also achieving energy efficiency. The performance evaluation of the A1 with other protocols reveals that the A1 has better sensing coverage with more connected nodes under topology maintenance operation.

It is well known that wireless links are inherently unreliable, but on the other hand, the packet losses are not acceptable for many mission critical WSN applications (e.g., forest fire detection, battle field monitoring) which require the network topology to provide a certain desired level of reliability. Therefore, network reliability is needed in most of the application scenarios. However, it should be achieved while keeping in mind the fundamental energy consumption constraint of a WSN.

The number of spanning trees measures the network performance under highly dynamic conditions, e.g. frequent link failures. Therefore, in CDS protocols, the probability that there is at least one spanning tree or the probability that the sensor nodes can communicate with each other in case of random link failures is very low due to the formation of a tree. In this context, poly protocol was presented for reliable and energy efficient operation in a WSN. The poly protocol exploits the polygons inherently present in the network and forms a reduced topology comprising of two paths present near the sink node. To reduce the message complexity associated with polygon formation, the poly does not find all the polygons in the network. On the other hand, the performance evaluation of the protocol shows that Poly achieves network reliability and outperforms other CDS based protocols. This work will be taken further and extended as part of the future work and is explained in the next section.

8.2 Future Work and Directions

In this section, future work and directions which can be drawn out from the presented work are presented.

Mission critical applications in WSN's are loss and delay sensitive and therefore require transmission reliability to maintain detection and response capabilities in a timely manner. The transmission reliability in CDS techniques is generally achieved by using hop by hop retransmission. However, retransmission adversely affect the performance for lossy and delay sensitive applications. Moreover, it was earlier demonstrated that CDS based protocols are unable to provide desired level of reliability as they maintain a single disjoint path between source nodes and the sink node. Therefore, the removal of an edge in a CDS tree partitions the graph.

As part of our future work, the packet delivery reliability of multiple node disjoint paths is the main area of interest. While analyzing CDS protocols, the interest is in knowing that what level of link redundancy can reduce the average path length which on the other hand can provide higher level of packet delivery reliability. This reliability should however be achieved while keeping in mind the fundamental energy consumption constraint of a WSN.

To demonstrate the impact of maintaining more number of links, the energy consumption for maintaining a bidirectional link was modeled by 'assuming' two nodes namely node A and node B. The traffic t received and transmitted by node A is

$$t = t_{AB} + t_{BA}. \qquad (8.12)$$

Let $p_T(d)$ be the total power consumed while transmitting at distance d, then

$$p_T(d) = p_{To} + p_A(d), \qquad (8.13)$$

The p_{To} is the power consumed by the circuitry and $p_A(d)$ is the power consumed by the power amplifier to transmit over a distance d and is also equal to the DC power p_{DC}. Let p_R is the power consumed while receiving, then

$$p_R = p_{Ro} , \tag{8.14}$$

The p_{Ro} is the total power consumed while receiving a packet from node B . The ratio of RF output power which is power delivered to antenna to DC input power is called the drain efficiency which is about 50% of the actual transmitted for class A power amplifiers and is given by:

$$h = \frac{p_{Tx}}{p_{DC}}. \tag{8.15}$$

Since, DC power is equal to the power consumed by the power amplifier, the total power consumed while transmitting using equation (8.15) becomes:

$$p_T(d) = p_{To} + \frac{p_{Tx}(d)}{h}. \tag{8.16}$$

But, the interest in knowing the power consumption at node A, so by adding the transmitted and received power in equation (8.14) and (8.16), we get:

$$p_A = p_{To} + \frac{p_{Tx}(d)}{h} + p_{Ro}. \tag{8.17}$$

The total transmission and reception power consumption at node A by using equation (8.17) is then given by:

$$p_A(t) = p_T(d).t_{AB} + p_R.t_{BA} = (p_{To} + \frac{p_{Tx}(d)}{h}).t_{AB} + (p_{Ro}).t_{BA} \tag{8.18}$$

The above equation states the energy consumed for maintaining a bidirectional link. Hence, as the backbone nodes transmit/receive information from more nodes, energy consumption increases accordingly. On the other side, it is well known that routing a packet through a multiple node disjoint path can achieve higher reliability. Moreover, a multiple node disjoint path can reduce the number of retransmissions which on the other hand provide better energy efficiency.

Therefore, in the future work, answers to the two fundamental questions in terms of an efficient and reliable topology control protocol are desired.

- Given K new edges/links, where these edges should be induced to achieve the maximum increase in reliability and this maximum can be achieved with how many number of edges.
- Given a desired reliability, what is the minimum number of edges required to achieve the desired reliability.

As every pair of nodes in a CDS is connected through at most one path, the average path length (l) between any pair of nodes increases with the increase in the node density. Therefore, if these fundamental questions are answered then a desirable reliability can be achieved with a desired level of energy consumption.

In addition to the above work, some new directions can also be extracted from the work presented in this report. As the sensing operation consumes much energy, due to different sensing ranges of the protocols, therefore in this domain, the energy efficiency of topology control protocols can be further improved by having a mechanism to reduce the overlapping sensing areas. Due to this reason, a mechanism can be developed in order to detect the overlapping ranges which then are reduced to improve the energy efficiency. Moreover, in such cases, erasure coding can be used on top of the protocol to improve the data redundancy and achieve network reliability.

The protocols presented in this work can also be used to address some of the pressing questions in mobile networks. The distributed topology control protocols presented in this report are capable of quick modifications, based on the response updates and timeouts of hello messages. Therefore, with some more modifications, they can also be suitable for networks of moving objects – vehicular networks or networks of mobile relays [71], [72], [73]. In these networks energy efficiency is not the prime challenge – impact on network service is very important. Therefore, the location of the receiving station and its position in relation to the nearest relay is very important in the delivery of high-speed services and quick topology control is required.

With this in mind, it may be better to apply the protocols in a different environment by trying to solve not the power-constrained topology control problems in WSNs, but a service delivery problem in ‘moving objects’ networks.

Appendix A

Publications

[1] Hassaan Khaliq Qureshi, Khurram Shahzad, Syed Ali Khayam, Muttukrishnan Rajarajan, and Veselin Rakocevic, "Complexity Reduction of Markov Channel Models for Wireless Networks using Graph Theory," *IEEE Milcom*, November, 2008.

[2] Hassaan Khaliq Qureshi, Adnan Iqbal, Muhammad Saleem, Syed Ali Khayam, Muttukrishnan Rajarajan, and Veselin Rakocevic, "A Cycle-based Clustering Algorithm for Wireless Sensor Networks," *IEEE NGMAST*, September, 2009.

[3] Adnan Iqbal, Hassaan Khaliq Qureshi, Muhammad Saleem, and Syed Ali Khayam, "On the Tradeoff between Reliability and Frame Length in Wireless Networks," *IEEE NGMAST*, September, 2009.

[4] H. K. Qureshi, S. Rizvi, M. Saleem, S. A. Khayam, V. Rakocevic, and M. Rajarajan, "An Energy Efficient Clique-based CDS Discovery Protocol for Wireless Sensor Networks," *44th Annual conference on Information Sciences and Systems (CISS)*, March, 2010.

[5] Hassaan Khaliq Qureshi, Junaid Jameel Ahmad, Syed Ali Khayam, Muttukrishnan Rajarajan, and Veselin Rakocevic, "GRAPH-THEORETIC COMPLEXITY REDUCTION FOR MARKOVIAN WIRELESS CHANNEL MODELS, *Springer's Journal of Wireless Personal Communications (WPCJ),* vol. 58, no. 4, pp.831-849, June, 2011.

[6] Hassaan Khaliq Qureshi, Sajjad Rizvi, Muhammad Saleem, Syed Ali Khayam, Veselin Rakocevic, and Muttukrishnan Rajarajan, "Poly: A Reliable and

Energy Efficient Topology Control Protocol for Wireless Sensor Networks," accepted in *Elsevier Computer Communications Journal (Comcom)*, vol. 34, no. 10, pp. 1235-1242, May, 2011.

[7] Hassaan Khaliq Qureshi, Sajjad Rizvi, Muhammad Saleem, Syed Ali Khayam, Veselin Rakocevic, and Muttukrishnan Rajarajan, "Evaluation and Improvement of CDS-based Topology Control for Wireless Sensor Networks," conditionally accepted in *Springer Wireless Networks Journal (WINET)*, January, 2011.

[8] Hassaan Khaliq Qureshi, Sajjad Rizvi, Syed Ali Khayam, Veselin Rakocevic, and Muttukrishnan Rajarajan, " A1: An Energy Efficient Topology Control Protocol for Connected Area Coverage in Wireless Sensor Networks," under review process in *Elsevier Journal of Networks and Computer Applications (JNCA),* May, 2011.

References

[1] F. Salvadori, M. de Campos, P. S. Sausen, R. F. de Camargo, C. Gehrke, C. Rech, M. A. Spohn, A. C. Oliveira, “Monitoring in Industrial Systems Using Wireless Sensor Network With Dynamic Power Management”, *IEEE Transactions on Instrumentation and Measurement,* vol. 58, no. 9, pp. 3104-3111, September, 2009.

[2] T. H. Huang, T. Sang, Y.M. Yu, “An architecture for integrating wireless sensor networks into IP network”, *International Conference on Intelligent Computing and Integrated Systems (ICISS),* pp. 910-912, October, 2010.

[3] Z. Chi, Z. Yanchao, F. Yuguang, “A Coverage Inference Protocol for Wireless Sensor Networks”, *IEEE Transactions on Mobile Computing*, vol. 9, no. 6, pp. 850-864, June, 2010.

[4] W-Y. Chung, Y-D Lee, S-J. Jung, “A wireless sensor network compatible wearable u-healthcare monitoring system using integrated ECG, accelerometer and SpO2”, *IEEE 30th Annual International Conference of Engineering in Medicine and Biology Society, (EMBS),* pp.1529-1532, August, 2008.

[5] J. Wu, F. Dai, S. Yang, “Iterative Local Solutions for Connected Dominating Set in Ad Hoc Wireless Networks”, *IEEE Transactions on Computers,* vol. 57, no. 5, pp. 702-715, May, 2008.

[6] R. Misra, C. Mandal, “Minimum Connected Dominating Set Using a Collaborative Cover Heuristic for Ad Hoc Sensor Networks”, *IEEE Transactions on Parallel and Distributed Systems,* vol. 21, no. 3, pp. 292-302, March, 2010.

[7] E. Felemban, S. Vural, R. Murawski, E. Ekici, L. Kangwoo, M. Youngbag, P. Sangjoon, “SAMAC: A Cross-Layer Communication Protocol for Sensor Networks with Sectored Antennas”, *IEEE Transactions on Mobile Computing*, vol. 9, no. 8, pp. 1072-1088, August, 2010.

[8] X. Wang, X. Zhang, Q. Zhang, G. Chen, "An Energy-Efficient Integrated MAC and Routing Protocol for Wireless Sensor Networks", *IEEE ICC,* pp. 1-6, June, 2009.

[9] F. Hu, Y. Xiao, Q. Hao, "Congestion-aware loss-resilient bio-monitoring sensor networking for mobile health applications", *IEEE Journal on Selected Areas in Communications,* vol. 27, no. 4, pp. 450-465, May, 2009.

[10] X. Zhang, H, Jiang, L. Zhang, C. Zhang, Z. Wang, X. Chen, "An Energy-Efficient ASIC for Wireless Body Sensor Networks in Medical Applications", *IEEE Transactions on Biomedical Circuits and Systems,* vol. 4, no. 1, pp. 11-18, February, 2010.

[11] ALERT Systems, http://www.alertsystems.org.

[12] A. Mainwaring, J. Polastre, R. Szewczyk, and D. Culler, "Wireless Sensor Networks for Habitat Monitoring," *ACM Workshop on Wireless Workshop in Wireless Sensor Networks and Applications*, June, 2002.

[13] C. Chiasserini, E. Magli, and D. di. Elettronica, "Energy Consumption and Image Quality in Wireless Video Surveillance Networks," IEEE PIMRC, vol. 5, pp. 2357-2367, December, 2002.

[14] M. Galetzka, J. Haufe, M. Lindig, U. Eichler, P. Schneider, "Challenges of simulating robust wireless sensor network applications in building automation environments," *IEEE EFTA*, pp. 1-8, November, 2010.

[15] H. S. Wang, "Finite-state Markov channel – a useful model for radio communication channels," *IEEE Transactions on Vehicular Technology.,* vol. 44, no. 1, pp. 163-171, February, 1995.

[16] M. Zorzi, R. Rao, and L. Milstein, "On the accuracy of a first-order Markov model for data transmission on fading channels," in *Proceedings of Universal Personal Communications conference,* pp. 211-215, November, 1995.

[17] W. Turin and R. Van Nobelen, "Hidden Markov Modeling of Flat Fading Channels," *IEEE JSAC*, vol. 16, no.7, pp. 1809-1817, December, 1998.

[18] P. Ji, B. Liu, D. Towsley, and J. Kurose, "Modeling Frame-level Errors in GSM Wireless Channels," *IEEE Globecom*, vol. 3, pp.2483-2487, November, 2002.

[19] S. A. Khayam, H. Radha, S. Aviyente, and J. R. Deller, Jr., "Markov and multifractal wavelet models for wireless MAC-to-MAC channels," *Elsevier Performance Evaluation*, vol. 64, no. 4, pp. 298–314, May, 2007.

[20] A. M. Chen and R.R. Rao, "Wireless Channel Models – Coping with Complexity," *Wireless Multimedia Network Technologies*, Kluwer Academic Publishers, vol. 524, no. 2, pp. 271-288, 2002.

[21] A. M. Chen and R.R. Rao, "On tractable wireless channel models," *IEEE PIMRC*, vol. 2, pp. 825-831, September, 1998.

[22] A. Iqbal and S.A. Khayam, "Improving WSN Simulation and Analysis Accuracy Using Two-Tier Channel Models," IEEE *ICC*, pp. 349-353, May, 2008.

[23] A. Konrad, B.Y. Zhao, A.D. Joseph, and R. Ludwig, "A Markov based channel model algorithm for wireless networks," *ACM WINET*, vol. 9, no. 3, pp. 189-199, May, 2003.

[24] S. A. Khayam and H. Radha "Constant-Complexity Models for Wireless Channels," *IEEE Infocom*, pp. 1-11, April, 2006.

[25] A. Willig. "A new class of Packet- and Bit-Level Models for Wireless Channels," *IEEE PIMRC*, vol. 5, pp. 2434-2440, September, 2002.

[26] A. Köpke, A. Willig, and H. Carl, "Chaotic maps as parsimonious bit error models of wireless channels," *IEEE Infocom*, vol.1, pp. 513-523, March, 2003.

[27] S. A. Khayam and H. Radha, "Linear-Complexity models for wireless MAC-to-MAC channels,"*ACM WINET*, vol. 11, no. 5, pp. 543-555, September, 2005.

[28] S.A. Khayam and H. Radha, "Markov-based modeling of wireless local area networks," *ACM Mobicom International Workshop on Modeling, Analysis and Simulation of Wireless and Mobile Systems (MSWiM)*, September, 2003.

[29] J. Tang, G. Xue, and W. Zhang, "Reliable Routing in Mobile Ad Hoc Networks Based on Mobility Prediction", *IEEE MASS*, pp. 466-474. October, 2004.

[30] S. Mueller, R. P. Tsang, and D. Ghosal, "Multipath Routing in Mobile Ad Hoc Networks: Issues and Challenges", *Lecture Notes in Computer Science (LNCS)*, vol. 2965, pp. 209-234, April, 2004.

[31] G. Holland and N. Vaidya, "Analysis of TCP Performance over Mobile Ad Hoc Networks," *ACM Wireless Networks (WINET)*, vol. 8, pp. 275–288, May, 2002.

[32] Z. Fu, X. Meng, and S. Lu, "How Bad TCP Can Perform In Mobile Ad Hoc Networks," *IEEE ISCC*, pp. 298-303, November, 2002.

[33] C. H. Shih, Y. M. Tou, and C. K. Shieh, "A self-regulated redundancy control scheme for wireless video transmission," *IEEE WirelessCom*, vol. 1, pp. 439-444, June, 2005.

[34] E. Masala, M. Bottero, and J. C. De Martin, "MAC-level partial checksum for H.264 video transmission over 802.11 ad hoc wireless networks," *IEEE VTC*, vol. 5, pp. 2864-2868, June, 2005.

[35] P. Santi, "Topology Control in Wireless Ad Hoc and Sensor Networks", *ACM Computing Surveys (CSUR)*, vol. 37, no. 2, June, 2005.

[36] N. Li, J. C. Hou, "Topology control in heterogeneous wireless networks: problems and solutions", *IEEE Infocom*, March, 2004.

[37] L. Li, J. Y. Halpern, P. Bahl, Y. M. Wang, R. Wattenhofer, "A cone-based distributed topology-control algorithm for wireless multi-hop networks", *IEEE/ACM Transactions on Networking,* vol. 13, no. 1, pp. 147–159, February, 2005.

[38] N. Li, J. C. Hou, "FLSS: A fault-tolerant topology control algorithm for wireless networks", *ACM Mobicom*, pp. 275–286, September, 2004.

[39] X. Li, W. Song, Y. Wang, "Localized topology control for heterogeneous wireless sensor networks", *ACM Transactions on Sensor Networks (TOSN),* vol. 2, no. 1, pp. 129–153, February, 2006.

[40] Y. Jaewook, E. Noel, K. W. Tang, "Degree Constrained Topology Control for Very Dense Wireless Sensor Networks", *IEEE Globecom,* pp. 1-6, December, 2010.

[41] A. Jeng, R. Jan, "Adaptive Topology Control for Mobile Ad Hoc Networks", *IEEE transactions on parallel and distributed systems*, pp. 1-8, February, 2011.

[42] M. Arguedas, C. Perez, M. Carvalho, K. Hoback, W. Kraus, "A directed expansion topology control algorithm for mobile Ad-Hoc networks", *IEEE Milcom*, pp. 1417-1422, November, 2010.

[43] N. Li, J. C. Hou, L. Sha, "Design and analysis of an MST-based topology control algorithm", *IEEE Transactions on Wireless Communications,* vol. 4, no. 3, pp. 1195–1206, May, 2005.

[44] R. B. Matthias, H. C. R. Carlos, D. Turgut, S. Rothkugel, "LSWTC: A local small-world topology control algorithm for backbone-assisted mobile ad hoc networks", *IEEE 35th conference on Local computer Networks (LCN),* pp. 144-151, October, 2010.

[45] L. Ding, W. Wu, J. Willson, H. Du, W. Lee, D. Du, "Efficient Algorithms for Topology Control Problem with Routing Cost Constraints in Wireless Networks", *IEEE Transactions on Parallel and Distributed Systems*, pp. 1-9, January, 2011.

[46] S. Kim, R. Fonseca, and D. Culler, "Reliable Transfer on Wireless Sensor Networks," *IEEE SECON*, pp. 449-459, October, 2004.

[47] D. Kim, W. Wang, X. Li, Z. Zhang, W. Wu, "A New Constant Factor Approximation for Computing 3-Connected m-Dominating Sets in Homogeneous Wireless Networks", *IEEE Infocom*, pp. 1-9, March, 2010.

[48] H. Zhang, A. Arora, Y. R. I. Choi, and M. G. Gouda. "Reliable Bursty Convergecast in Wireless Sensor Networks," *ACM Mobihoc*, May, 2005.

[49] N. D. Thuan, H. Nishiyama, N. Ansari, N. Kato, "On Performance Evaluation of Reliable Topology Control Algorithms in Mobile Ad Hoc Networks", *IEEE vehicular technology conference*, pp. 1-5, September, 2010.

[50] K. Miyao, H. Nakayama, N. Ansari, N. Kato, "LTRT: An efficient and reliable topology control algorithm for ad-hoc networks", *IEEE Transactions on Wireless Communications*, vol. 8, no. 12, pp. 6050-6058, December, 2009.

[51] J. Wang, S. Medidi, "Topology Control for Reliable Sensor-to-Sink Data Transport in Sensor Networks", *IEEE ICC*, pp. 3215-3219, May, 2008.

[52] H. Jiang, L. Chen, J. Wu, S. Chen, H. Leung, "A Reliable and High-Bandwidth Multihop Wireless Sensor Network for Mine Tunnel Monitoring", *IEEE Sensors Journal*, vol. 9, no. 11, pp. 1511-1517, November, 2009.

[53] Wightman, P.M. Labrador, M.A, "A3: A Topology Construction Algorithm for Wireless Sensor Network", *IEEE Globecom*, pp. 1-6, December, 2008.

[54] J. Wu, M. Cardei, F. Dai, and S. Yang. "Extended dominating set and its applications in ad hoc networks using cooperative communication", *IEEE Transactions on Parallel and Distributed Systems*, vol. 17, no. 8, pp. 851-864, August, 2006.

[55] Z. Yuanyuan, X. Jia, and H. Yanxiang, "Energy efficient distributed connected dominating sets construction in wireless sensor networks", *ACM International Conference on Wireless Communications and Mobile Computing*, pp. 797–802, July, 2006.

[56] F. Wang, M. T. Thai, D. Z. Du, "On the construction of 2-connected Virtual Backbone in Wireless Network", *IEEE Transactions on Wireless Communications*, vol. 8, no. 3, pp. 1230-1237, March, 2009.

[57] Y. Wu, F. Wang, M. T. Thai and Y. Li, "Constructing k-Connected m-Dominating Sets in Wireless Sensor Networks", *IEEE Military Communications Conference*, pp. 1-7, October, 2007.

[58] Y. Wu, Y. Li, "Construction algorithms for k-connected m-dominating sets in wireless sensor networks", ACM international symposium on Mobile ad hoc networking and computing, pp. 26-30, May, 2008.

[59] J. G. Kemeny and J. L. Snell, "Finite Markov Chains," *Springer-Verlag*: New York, 1976.

[60] C.M. Grinstead and J.L. Snell, "Introduction to Probability: Second Revised Edition," *AMS publishers,* 1997.

[61] S.A Khayam and H. Radha, "On the Impact of Ignoring Markovian channel Memory on the Analysis of Wireless Systems," *IEEE* ICC, pp. 199-204, June, 2007.

[62] Crossbow homepage, http://www.xbow.com.

[63] TinyOS homepage, http://www.tinyos.net/.

[64] Linux-wlan homepage, http://www.linux-wlan.org/.

[65] T. Cover and J. A. Thomas, *Elements of Information Theory*, Second Edition. *Wiley publishers*, ISBN: 0-471-24195-4, July, 2006.

[66] J. M. Mcdougall. "Low complexity channel models for approximating flat Rayleigh fading in network simulations," *PhD thesis from Texas A&M University*, August, 2003.

[67] K. A. M. Almahorg, S. Naik, X. Shen, "Efficient Localized Protocols to Compute Connected Dominating Sets for Ad Hoc Networks", *IEEE Globecom*, pp. 1-5, December, 2010.

[68] S. Parikh, V. Vokkarane, L. Xing, D. Kasilingam, "An energy efficient routing mechanism for wireless sensor networks", *20th International Conference on Advanced Information Networking and Applications (AINA)*, vol. 2, pp. 308–312, 2006.

[69] N. Zhao, M. Wu, Y. Zhang, "A Localized MESH-CDS Structure and Maintenance for Wireless Sensor Networks", *6th International Conference on Wireless and Mobile Communications (ICWMC),* pp. 262-267, September, 2010.

[70] A. H. Shuaib, A. H. Aghvami, "A Routing Scheme for the IEEE-802.15.4-Enabled Wireless Sensor Networks", *IEEE Transactions on Vehicular Technology,* vol. 58, no. 9, pp. 5135-5151, November, 2009.

[71] F. Farnoud, S. Valaee, "Reliable Broadcast of Safety Messages in Vehicular Ad Hoc Networks", *IEEE Infocom*, pp. 226-234, April, 2009.

[72] B. Lorenzo, S. Glisic, "Traffic adaptive relaying topology control", *IEEE Transactions on Wireless Communications*, vol. 8, no. 11, pp. 5612-5620, November, 2009.

[73] M. Luis, R. Oliveira, L. Bernardo, A. Garrido, P. Pinto, "Joint topology control and routing in ad hoc vehicular networks", *European Wireless Conference (EW),* pp. 528-535, April, 2010.

[74] A. Durán, M. Toril, F. Ruiz, M. Solera, R. Navarro, "Analysis of topology control algorithms in multi-hop cellular networks", *Fifth International Conference on Broadband and Biomedical Communications (IB2Com),* pp. 1-6, December, 2010.

[75] M. Dohler, *et al.* "The ARESA Project: Facilitating Research, Development and Commercialization of WSNs," *IEEE SECON 2007*, pp. 590- 599, June 2007.

[76] A. Burnstein, K. Bult, D. Chang, and F. Chang, "Wireless Integrated Microsensors," *Proceedings Sensors EXPO*, Anaheim, CA, April, 1996.

[77] J. M. Kahn, R. H. Katz, and K. S. J. Pister, "Next Century Challenges: Mobile Networking for Smart Dust," *ACM/IEEE Mobicom*, August, 1999.

[78] J. Rabaey, J. Ammer, J. L. da Silva and D. Patel, "PicoRadio: Adhoc Wireless Networking of Ubiquitous Low-Energy Sensor/Monitor Nodes," *IEEE Computer Society Workshop on VLSI*, pp. 9-12, August, 2000.

[79] V. Raghunathan, C. Schurgers, S. Park, and M. B. Srivastava, "Energy-Aware Wireless Microsensor Networks", IEEE *Signal Processing Magazine*, vol. 19, no. 2, pp. 40-50, March, 2002.

[80] B. Chen, K. Jamieson, H. Balakrishnan, and R. Morris. "Span: An energy-efficient coordination algorithm for topology maintenance in ad hoc wireless networks", *ACM Wireless Networks Journal*, vol. 8, no. 5, pp. 481-494, September, 2002.

[81] S. Liqi, A. O. Fapojuwo, "TDMA Scheduling with Optimized Energy Efficiency and Minimum Delay in Clustered Wireless Sensor Networks", *IEEE Transactions on Mobile Computing*, vol. 9, no. 7, pp. 927-940, July, 2010.

[82] M. Nabi, M. Blagojevic, M. Geilen, T. Basten, T. Hendriks, "MCMAC: An Optimized Medium Access Control Protocol for Mobile Clusters in Wireless Sensor Networks", *IEEE SECON*, pp. 1-9, June, 2010.

[83] V. Kawadia, P. Kumar, "Power control and clustering in ad hoc networks", *IEEE Infocom*, pp. 459–469, April, 2003.

[84] H. M. Ammari, S. K. Das, "A Study of k-Coverage and Measures of Connectivity in 3D Wireless Sensor Networks", *IEEE Transactions on Computers*, vol. 59, no. 2, pp. 243-257, February, 2010.

[85] R. Khanna, H. Liu, H. H. Chen, "Self-Organization of Wireless Sensor Network for Autonomous Control in an IT Server Platform", *IEEE ICC*, pp. 1-5, May, 2010.

[86] Y.Sankarasubramaniam, I.F. Akyildiz and S.W. McLaughlin, "Energy efficiency based packet size optimization in wireless sensor networks," *IEEE Workshop on Sensor Network Protocols and Applications (SNPA)*, June, 2004.

[87] Z. Tian, D. Yuan, Q. Liang, "Energy Efficiency Analysis of Error Control Schemes in Wireless Sensor Networks", *International Wireless*

Communications and Mobile Computing Conference (IWCMC), pp. 401-405. August, 2008.

[88] M. C. Vuran, I. F. Akyildiz, “Error Control in Wireless Sensor Networks: A Cross Layer Analysis”, *IEEE/ACM Transactions on Networking*, vol. 17, no. 4, pp. 1186-1199, August, 2009.

[89] D. J. Baker and A. Ephremides, “The architectural organization of a mobile radio network via a distributed algorithm”, *IEEE Transactions on Communications*, vol.29, pp. 1694-1701. November, 1981.

[90] M. Gerla, J.T. Tsai, “Multicluster mobile multimedia radio network”, *Springer’s Wireless Networks Journal*, vol. 1, no. 3, pp. 255-265, 1995.

[91] D. Kim, Y. Wu, Y. Li, F. Zou and D. Zhu. Du, “Constructing Minimum Connected Dominating Sets with Bounded Diameters in Wireless Networks”, *IEEE Transactions on Parallel and Distributed computing*, vol. 2, no. 2, pp. 147-157, February, 2009.

[92] P. J. Wan, K. M. Alzoubi and O. Frieder, “Distributed Construction of Connected Dominating Sets in Wireless Ad Hoc Networks”, *IEEE Infocom*, vol. 3, pp. 1597-1604, November, 2002.

[93] B. Han, W. Jia, “Efficient Construction of Weakly-Connected Dominating Set for Clustering Wireless Ad Hoc Networks”, *IEEE Globecom*, pp. 1-5, November, 2006.

[94] R. Ramanathan, R. Rosales-Hain, “Topology control of Multihop Wireless Networks Using Transmit Power adjustment”, *IEEE Infocom*, vol. 2, pp. 404-413, March, 2000.

[95] V. Rodoplu, T. H.Meng, “Minimum Energy Mobile Wireless Networks”, *IEEE Journal of Selected Areas in Communications*, vol.17, no.8, pp.1333-1344, August, 1999.

[96] P. M. Wightman and M. A. Labrador, ``Atarraya: A Simulation Tool to Teach and Research Topology Control Algorithms for Wireless Sensor Networks'', *Create-Net 2nd International Conference on Simulation Tools and Techniques*, SIMUTools, pp. 1-10, March, 2009.

[97] Maple homepage, http://www.maplesoft.com.

[98] P.G. Ralph, “Discrete and Combinatorial Mathematics: An applied Introduction”, 3rd edition, *Addison-Wesley publishers*, Reading, MA, 1994.

[99] Y. Xu, J. Heidemann, and D. Estrin, “Adaptive Energy-Conserving Routing for Multi-hop Ad Hoc Networks'', *Research Report no. 527, USC/Information Sciences Institute*, October, 2000.

[100] Y. Xu, J. Heidemann, and D. Estrin, "Geography-informed Energy Conservation for Ad Hoc Routing," *7th Annual International Conference on Mobile Computing and Networking (Mobicom)*, pp. 70-84, July, 2001.

[101] R. Iyengar, K. Kar, and S. Banerjee, "Low-coordination Topologies for Redundancy in Sensor Networks," *6th ACM International Symposium on Mobile Ad Hoc Networking and Computing(Mobihoc)*, pp. 332-342, May, 2005.

[102] W. B. Heinzelman, A. P. Chandrakasan, and H. Balakrishnan, "Energy-Efficient Communication Protocols for Wireless Micro-sensor Networks," *33rd Hawaii International Conference on System Sciences (HICSS)*, vol. 8, January, 2000.

[103] G. Xing, T. Wang, W. Jia and M. Li, "Rendezvous Design Algorithms for Wireless Sensor Networks with a Mobile Base Station," *9th ACM International Symposium on Mobile Ad Hoc Networking and Computing (MobiHoc)*, May, 2008.

[104] H. K. Qureshi, M. Rajarajan, V. Rakocevic, A. Iqbal, M. Saleem, and S. A. Khayam, "A Cycle-based Clustering Algorithm for Wireless Sensor Networks," *IEEE Annual International Conference on Next Generation Mobile Applications, Services and Technologies (NGMAST)*, pp. 441-445, September, 2009.

[105] C. R. Dow, P. J. Lin, S. C. Chen, J. H. Lin, and S. F. Hwang, "A Study of Recent Research Trends and Experimental Guidelines in Mobile Ad Hoc Networks," *19th International Conference on Advanced Information Networking and Applications (AINA05)*, pp. 72-77, March, 2005.

[106] S. Ni, Y. Tseng, Y. Chen, and J. Sheu, "The Broadcast Storm Problem in a Mobile Ad Hoc Network," *ACM International Conference on Mobile Computing and Networking (Mobicom)*, pp. 152-162, August, 1999.

[107] A. Ephremides, J. Wieselthier, and D. Baker, "A Design Concept for Reliable Mobile Radio Networks with Frequency Hopping Signaling," *Proc. IEEE,* vol. 75, no. 1, pp. 56-73, January, 1987.

[108] S. Guha and S. Khuller, "Approximation Algorithms for Connected Dominating Sets," *Algorithmica*, vol. 20, pp. 374-387, April, 1998.

[109] K. Mohammed, L. Gewali, and V. Muthukumar, "Generating Quality Dominating Sets for Sensor Network," *Sixth International Conference on Computational Intelligence and Multimedia Applications (ICCIMA 05)*, pp. 204-211, August, 2005.

[110] K. M. Alzoubi, P. J. Wan, and O. Frieder, "New Distributed Algorithm for Connected Dominating Set in Wireless Ad Hoc Networks," *35th Hawaii International Conference on System Sciences (HICSS 02)*, vol. 9, pp. 3849-3855, January, 2002.

[111] K. M. Alzoubi, P. J. Wan, and O. Frieder, "Distributed Heuristics for Connected Dominating Sets in Wireless Ad Hoc Networks," *Journal of Communications and Networks*, vol. 4, no. 1, pp. 1-8, March, 2002.

[112] L. Feeney and M. Nilsson, "Investigating the energy consumption of a wireless network interface in an ad hoc networking environment," *IEEE Infocom*, vol. 3, pp. 1548-1557, April, 2001.

[113] S. Yang, J. Wu, and F. Dai, "Efficient Backbone Construction Methods in MANETs Using Directional Antennas, "*27th International Conference on Distributed Computing Systems (ICDCS)*, pp. 45-52, June, 2007.

[114] M. Saleem, I. Ullah, S. A. Khayam, and M. Farooq, "On the reliability of Ad hoc routing protocols for loss and delay sensitive applications," *Elsevier Ad hoc Networks*, vol. 9, no. 3, pp. 285-299, May, 2011.