Reinforcement Learning Framework for Energy Efficient Wireless Sensor Networks

Mr. Ankit B. Patel[1], Dr. Hitesh B. Shah[2]

[1] ME in Communication System Engineering, E & C, L D College of Engineering, Gujarat, India
[2] PhD IIT Delhi, M. Tech IIT Roorkee, E & C, G H Patel College of Engineering and Technology, Gujarat, India

---***---

Abstract - *In recent years, there has been an increasing interest in Wireless Sensor Networks. Sensor nodes in Wireless Sensor Networks are energy constrained. Therefore, in WSNs one of the major design challenges is to minimize consumed energy at the sensor nodes. Hence, a number of routing schemes are designed that make efficient use of limited energy of the sensor nodes. Hierarchical routing protocols are best known in regard to energy efficiency. Our first aim is an implement of routing strategies using Q-routing algorithms and compared them for the energy efficient aspects. Our second step is to develop an energy efficient shortest path Q-routing algorithm using Reinforcement Learning property of ML techniques and compared it with an existing energy aware Q-routing algorithm. We have applied these algorithms on different topologies like heterogeneous, homogeneous and introduced a novel hybrid topology by combining heterogeneous and homogeneous topologies.*

Key Words: Wireless Sensor Networks, Quality of Service routing, Reinforcement Learning, Machine Learning

1. Introduction

A collection of various distributed sensors is widely known as Wireless Sensor Network (WSN).Wireless Sensor Networks (WSNs) have gained worldwide attention in recent years, particularly with the explosion in Micro-Electro-Mechanical Systems (MEMS) technology which has facilitated the development of smart sensors [1]. A Wireless Sensor Network is type of Wireless Network. It is small and infrastructure less. Basically Wireless Sensor Network consist a number of sensor node, called tiny device and these are working together to detect a region to take data about the environment [2].

Basically two types of WSNs: Structured and Unstructured. And these deployed in ad-hoc way into an area. Once deployed, the network is absent unattended achieve monitoring and reporting functions. In other structured WSNs, the all sensor nodes are deployed in predesigned manner. The sensor nodes of the WSNs are allows random deployment in inaccessible terrains, this means protocol of the wireless sensor is self-organized; another important feature of the WSNs is supportive effort of sensor nodes. Sensor nodes are gathering data about

environment, after gathering it they process it and then transmit to the base station. Base station offers an interface between user and internet. Basic characteristic of the WSNs are limited energy, dynamic network topology, lower power, node failure and mobility of the nodes, short-range broadcast communication and multi-hop routing and large scale of deployment.

To transmit data from source node to destination node or vice versa in WSNs is known as routing. In comparison with mobile ad hoc networks or cellular networks routing in WSNs is very challenging [3].

In general, *proactive, reactive and hybrid* are the three different strategies in routing protocol:
Proactive strategy referred to as table driven, relies on periodic dissemination of routing information to maintain consistent and accurate routing tables across all sensor nodes of the network. Some of the well-known proactive routing presented in literature Destination-Sequenced Distance Vector (DSDV) and Fisheye State Routing (FSR).
Reactive strategy is on-demand, in the sense that the routes are discovered, when a node needs to send a packet. This type of protocol takes into account only the minimum hop count to touch the destination. . Some of the well-known reactive routing presented in literature Dynamic Source Routing (DSR) and Ad Hoc On-Demand Distance Vector (AODV).
Hybrid strategy incorporates the best features of the proactive and reactive strategies. A hybrid routing strategy can be adopted whereby proactive routing is used within a cluster and reactive routing is used across clusters. Some of the well-known hybrid routing presented in literature Zone Routing Protocol (ZRP) and Sharp Hybrid Adaptive Routing Protocol (SHARP).

For designing the routing protocol, many challenging factors affecting like as minimal computational and memory requirement, automaticity and self-organization, energy efficiency, scalability, architecture matching the characteristics of traffic patterns, support for in network data aggregation.

1.1 Reinforcement Learning for WSNs

With the rapid growth in MEMS technology, it is possible to deploy large-scale sensors in the field in the near future. However, large-scale sensor network also unavoidably introduce large amount of data in WSNs to be

processed, transmitted and received. Transmitting all data back to a base station for processing and making inferences is merely impossible due to the sensor limited energy and bandwidth constraints. Thus, there is a need for applying Machine Learning (ML) methods in WSNs. This strategy could significantly reduce the amount of data communications and truly utilize the distributive characteristic of WSNs [4].

The purpose of this class of ML algorithms is to automatically learn the properties of the environment and to adjust their behavior rapidly and simply to them. Neural Network (NN), Fuzzy Logic (FL), Evolutionary Algorithm (EA), Reinforcement Learning (RL), and Swarm Intelligence (SI) are different ML algorithms are used in WSNs.

ML in WSN application from two perspectives, namely the *network associated issue* and *application associated issue* explained in [4]. We found from literature that, ML can effectively solve various challenges in WSNs and improve their performance significantly. ML algorithm is best suited for addressed challenges of WSNs [5].

With the use of Reinforcement Learning technique we have only knowledge about the environment i.e. existing algorithms, available protocols and information about topology which is used to implement Wireless Sensor Networks. We may receive some evaluation of action (Reinforcement), but is not told of which action is the correct one to achieve goal i.e. to design energy efficient network. The network lifetime can increase by choosing the optimal path for routing the data packets in the network.

2. Related Work

A lot of research has been done already in the field of energy-aware routing algorithms and Reinforcement Learning based routing algorithms. Now days, Machine Learning algorithms are widely used to solve the issues related to Wireless Sensor Networks. Reinforcement Learning property of Machine Learning is very useful to develop the routing algorithms. Several routing algorithms are discussed here.

2.1 Basic Q-routing algorithm

The basic Q-routing algorithm initially developed by Boyan and Littman [6] and the purpose of this algorithm was to provide dynamic load balancing in packet-switched networks based on a Reinforcement Learning algorithm. Consider Q_x (d, y) as the time that node x estimates for a packet to arrive at destination node d when it forwards the packet to its neighboring node y. This process includes any delay time that the packet may experience at node x. Every node in the network will store a local table with estimates for each pair of neighbors and destinations. When x sends a packet with destination d to node y, node y will immediately retaliate to node x with its own estimate T for the remaining travel time to destination d. The formula for obtaining a value for T is given by,

$$T = \min_{z \in neighbors of y} Q_y(d, z) \qquad 2.1$$

Node x will update its local table of estimates by taking this variable T into account. It will also use some other factors. These include the time that the sending node needed to process the message (the time between receiving a message and forwarding it), represented below as q as well as the time the message spent traveling to the next node (indicating link quality), and represented as s. After that, it compares the new estimate to the old **estimate and applies a learning factor η to the difference.** This gives a delta value that will be used to update the **local estimate for a node's neighbor.**

Update rule:

$$Q_x(d, y) = Q_x(d, y) + \Delta Q_x(d, y) \qquad 2.2$$

With

$$\Delta Q_x(d, y) = \eta(q + s + T - Q_x(d, y))$$

The main disadvantage using the original Q-routing algorithm is its exploration behavior in certain situations. For example, after the network experiences a high traffic load and returns to a lower traffic load the original Q-routing algorithm is unable to recover its routing policy to a shortest path policy and will continue with using the load balancing policy.

2.2 Shortest Path Q-routing Algorithm

There are various shortest path algorithms are available for routing data packets in Wireless Sensor **Networks. Dijkstra's algorithm is widely used** for routing data packets in the network. We have developed shortest path Q-routing algorithm for routing data packets. This algorithm is based on Reinforcement Learning technique and use to find the shortest efficient path from source to destination. The shortest path Q-routing algorithm is designed as below:

$$M(i,j) = \sqrt{[(x(j) - x(i)]^2 + [y(j) - y(i)]^2} \qquad 2.3$$

Where, (i, j) = Number of nodes for i≠j

This algorithm starts at source node and spreads over the network. At each and every node in the network this algorithm explores the shortest path to the destination node and selects the optimal path from source to destination. We have applied this algorithm on heterogeneous, homogeneous and hybrid topologies to compare the network lifetime performance results.

2.3 Energy Aware Q-routing Algorithm

This algorithm was recently proposed by Forster for balancing energy expenditure in Wireless Sensor Networks through Reinforcement Learning [7]. In this algorithm, packets flow from the sinks to the sources to announce their data requirements and from the sources to

sinks to deliver data. It investigates the optimal route from a source to multiple destinations. It reduces the number of hops to reach multiple sinks in a network and spreads the energy expenditure during the network lifetime. The most predominant property of algorithm is its adaptability to node failures and mobility. Because of dynamic nature of topology, the new route costs must be quickly re-learned by exploration and the Q-value must be updated. In this algorithm, the hop based cost estimate and the battery level information are combined. The cost function is based on two elements, hop count and battery so the feedback section of the DATA packet is enlarged to carry both components of the Q-value. The stored Q-value for each route is extended to store both its components. The function f that combines the two estimates into a single Q-value is given by,

$$f(E_{hops}, E_{battery}) = hcm(E_{battery}) * E_{hops} \qquad 2.4$$

The hcm is known as the hop-count-multiplier. It is a function that weights the hop count estimate based on the remaining energy. The purpose of hcm is to increase the cost of some route when the nodes on it have depleted their batteries making that route less appealing in comparison to routes composed of nodes with better battery levels. The initial estimate, E_{hops}, for a node to route to some subset of all destinations $D_i \subseteq D$ through a single neighbor, n_i, is given by,

$$E_{hops}(n_i) = \left(\sum_{d \in d_i} hops_d^{n_i}\right) - 2(|D_i| - 1) \qquad 2.5$$

Where, neighbor n_i has an estimate of $hops_d$ hops to each of the sinks which is shown as,

$$d \in d_i$$

To calculate the full cost to route to all required destinations through possibly multiple neighbors, the full estimated cost is given by,

$$E_{hops}(route) = \left(\sum_{i=1}^{k} E_{hops}(n_i)\right) - (k - 1) \qquad 2.6$$

Where, k is the number of neighbors selected to serve as the next hop to reach all destinations.

2.4 Lowest Energy on Path Feedback Algorithm

The approach used in this algorithm is referred as "Maximum Minimum Energy Available Node Route" or "Lowest Energy on Path". The aim of this approach is to identify the weakest node in terms of available energy on each path and use these nodes from different paths to compare the energy performance of these paths. Because Q-learning algorithms generate dynamic routing trees this energy level should be estimated and appropriately adapted [8]. This is realized as follows:

"Whenever a node needs to send feedback, it will first look at its routing tables to verify which node is its current parent. It will then look at the estimated energy level of this parent node and compare this with its own energy level. The node will then return the lower of the two energy levels; as such the minimum energy level on the path is propagated towards the source **of the path.**" This

often results in significant improvements of the network lifetime.

2.5 Q-RC Algorithm

Q-routing with compression attempts to aggregate messages as early as possible in the routing process and tries to compress them together before sending them to a single sink. Q-learning techniques are used to learn the best compression path towards the sink. The algorithm is fully distributed and its concepts can be applied to the field of Wireless Sensor Networks [9].

3. Simulation Environment

We created a simulation environment for Wireless Sensor Networks topologies in the MATLAB software. We developed a new shortest path Q-routing algorithm to increase the network lifetime. The simulator was designed to test different routing algorithms and supports multiple routing algorithms. In order to validate the simulator we tested multiple well known algorithms and verified that their behavior was in line with expectations on all topologies. All algorithms were tested in a multi-source single sink scenario.

3.1 Example Topologies Used in Experiment

We have fundamentally used two well known topologies: homogeneous topologies and heterogeneous topologies and developed new hybrid topologies by combining these two topologies for the comparison of network lifetime performance results.

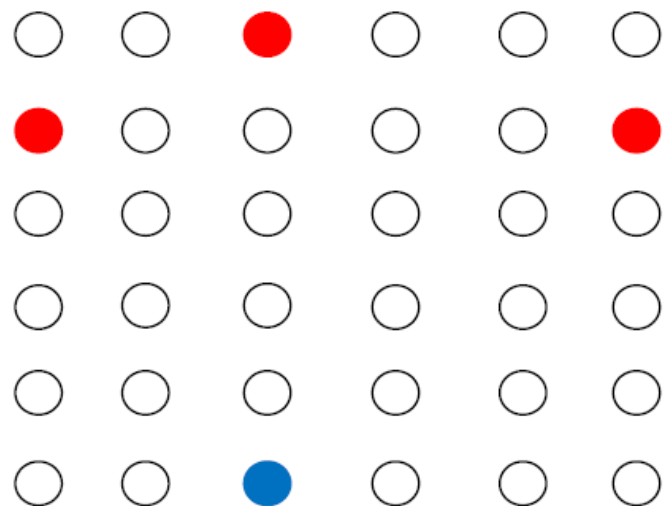## 3.1.1 Homogeneous Topologies

Figure 3.1 Homogeneous Topologies
Sources are shown in red. Sink is shown in blue.

When simulating homogeneous topologies we use the topologies shown in figure 3.1 where all nodes have the same probability of generating a data packet and all nodes have the same energy capacity.
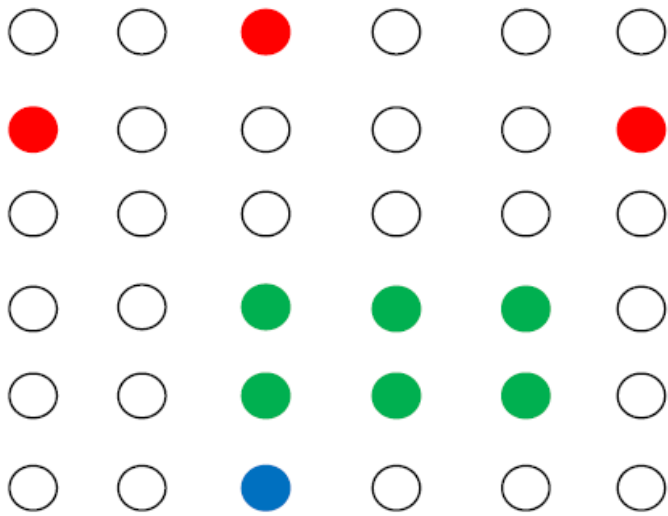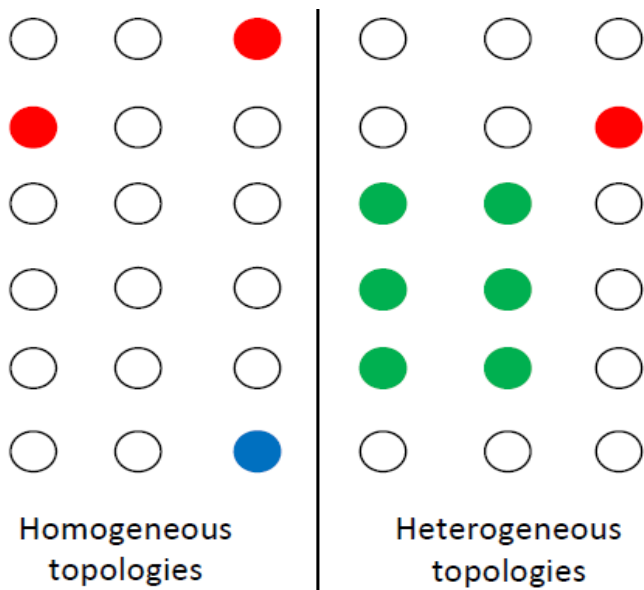
3.1.2 Heterogeneous Topologies

Figure 3.2 Heterogeneous Topologies
Sources are shown in red. Sink is shown in blue.
High probability nodes are shown in green.

For simulations of heterogeneous topologies, we use the topologies shown in figure 3.2. This time however, all nodes shown in green are three times more likely to generate a data packet when compared to other nodes. This results in an area that consumes significantly more energy than the rest of the network. When using the hop distance as routing metric, the lowest energy on path algorithm is able to better route data around this area of increased traffic.

3.1.3 Hybrid Topologies

Homogeneous topologies   Heterogeneous topologies

Figure 3.3 Hybrid Topologies
Sources are shown in red. Sink is shown in blue.
High probability nodes are shown in green.

For simulations of hybrid topologies, we use the topologies shown in figure 3.3. This time however, all nodes shown in green are three times more likely to generate a data packet when compared to other nodes. This results in an area that consumes significantly more energy than the rest of the network. When using the hop distance as routing metric, the lowest energy on path algorithm is able to better route data around this area of increased traffic.

4. Simulation Results

To evaluate and compare algorithms we ran multiple simulations in our simulation environment. We have applied both shortest path Q-routing algorithm and energy aware Q-routing algorithm on different topologies. All lifetime results are averaged over 15 independent runs by using shortest path Q-routing algorithm. We separate the results based on the type of topology the experiment was **performed on, as our results show that an algorithm's** performance can differ significantly on different topologies.

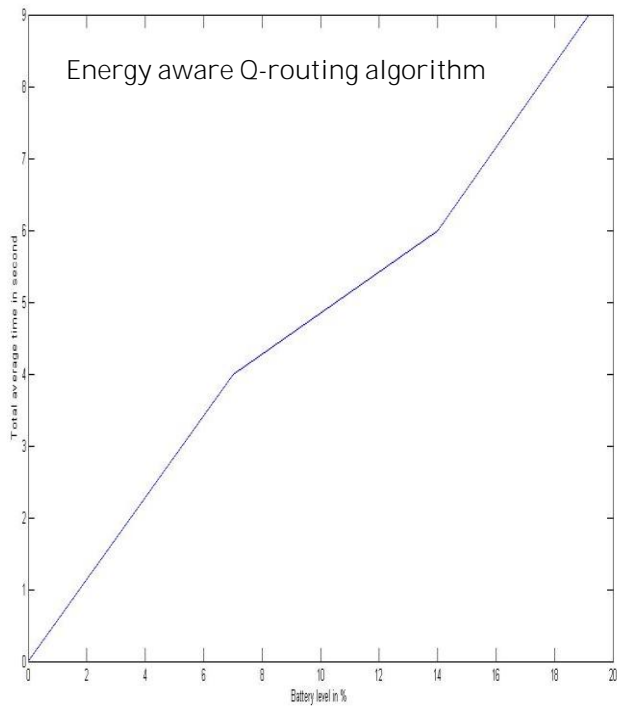## 4.1.1 Network Lifetime Measurements for Homogeneous Topologies
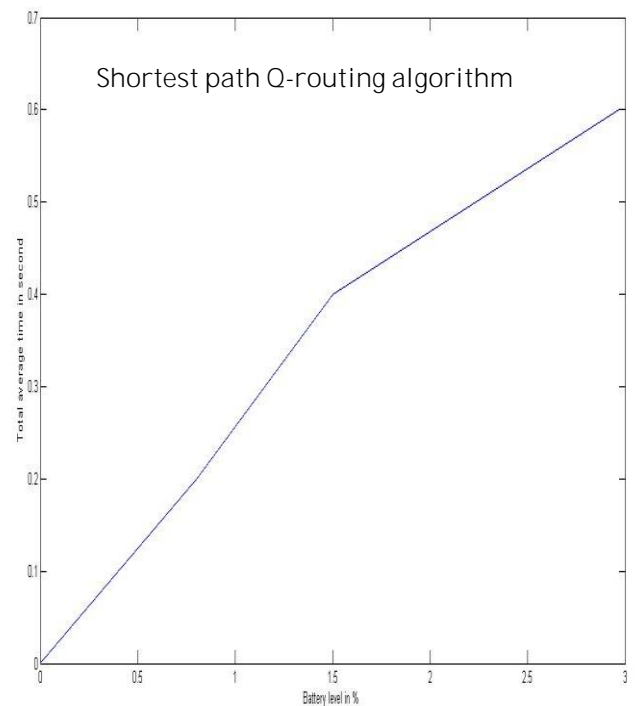
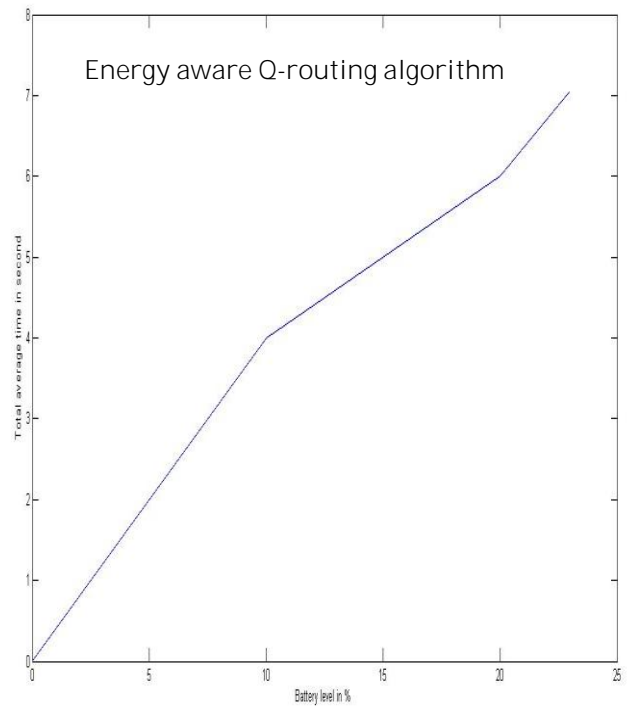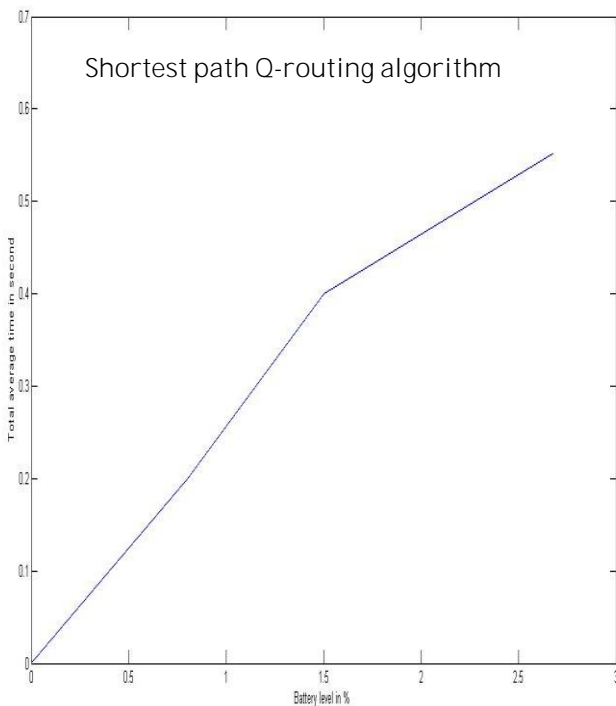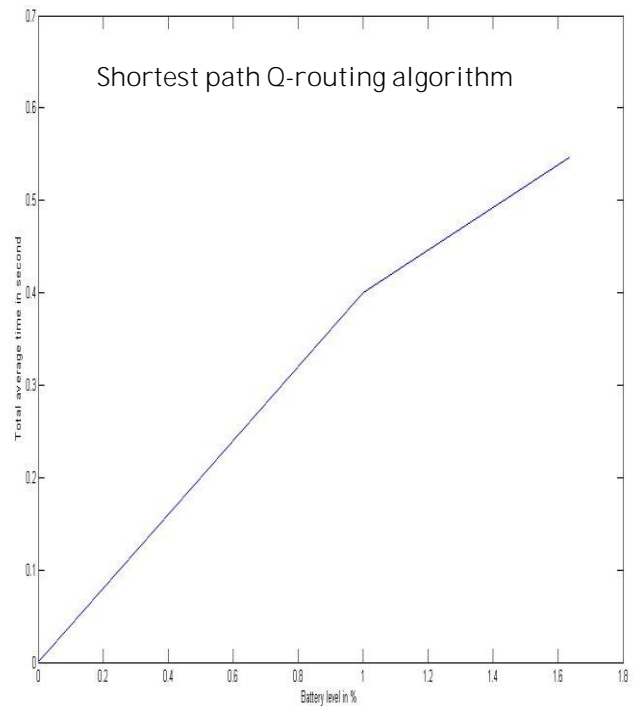Figure 4.1 Network Lifetime Measurements for Homogeneous Topologies

Figure 4.2 Network Lifetime Measurements for Heterogeneous Topologies

4.1.2 Network Lifetime Measurements for Heterogeneous Topologies

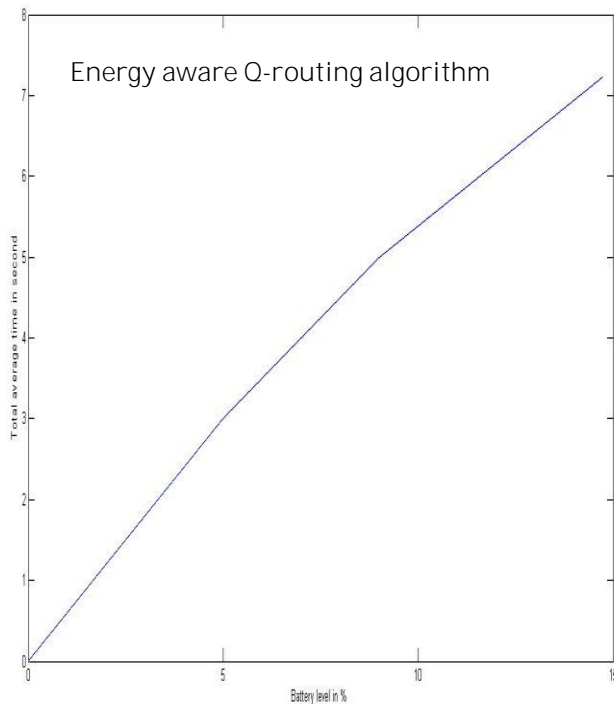## 4.1.3 Network Lifetime Measurements for Hybrid Topologies

Figure 4.3 Network Lifetime Measurements
for Hybrid Topologies

4.1.4 Comparison of Simulation Results in Tabular Form

Type of Topology	Type of Algorithm	Number of Nodes	Number of Paths	Consumed Battery Level in %	Total Average Time in Second
Homogeneous	Shortest path Q-routing algorithm	36	15	2.9703	0.6
	Energy aware Q-routing algorithm	36	15	19.1800	9
Heterogeneous	Shortest path Q-routing algorithm	36	15	2.6758	0.5523
	Energy aware Q-routing algorithm	36	15	22.9571	7.0358
Hybrid	Shortest path Q-routing algorithm	36	15	1.6349	0.5459
	Energy aware Q-routing algorithm	36	15	14.7256	7.2256

5. Conclusions and Future Scope

One of the challenges in WSNs is to develop an energy-efficient routing strategy in order to increase network lifetime. We have surveyed literature on energy efficient routing protocols, energy efficient routing strategies, topology control scheme and power management scheme in comparison with standard Quality of Service routing algorithms in Wireless Sensor Networks. Through the research papers we found that by sensor network life time may increase using energy efficient routing. Hence, our object towards developing energy efficient routing algorithm and energy efficient routing topologies.

We have developed a new shortest path Q-routing algorithm and hybrid topologies by combining homogeneous topologies and heterogeneous topologies to increase the network lifetime. We have applied both shortest path Q-routing algorithm and an existing energy aware Q-routing algorithm on different topologies like homogeneous topologies, heterogeneous topologies and hybrid topologies to compare the network lifetime performance. Using simulation results we conclude that our developed a new shortest path Q-routing algorithm gives optimal results on all topologies. Our developed hybrid topologies outputs better than homogeneous topologies and heterogeneous topologies by using both algorithms. The shortest path Q-routing algorithm gives better results on heterogeneous topologies and hybrid topologies than homogenous topologies. We concluded that energy aware Q-routing algorithm always gives good results on our developed hybrid topologies.

Finally, we conclude that our developed shortest path Q-routing algorithm always performs better on different topologies than an existing energy aware Q-routing algorithm. We noticed that our developed hybrid topologies increase the network lifetime compared to homogeneous topologies and heterogeneous topologies.

Given the highly satisfying results achieved in this thesis, we believe RL proves to be an efficient approach to solve many problems in WSNs. With the use of these two algorithms, a new algorithm, Q-RC can be used to increase the network lifetime. Q-routing with compression attempts to aggregate messages as early as possible in the routing process and tries to compress them together before sending them to a single sink. This thesis covers the way to further applications and energy efficient routing strategy, which will inherently improve the performance of WSNs, lower their design and deployment complication, and increase their application areas.

ACKNOWLEDGEMENT

period. His guidance and encouragement throughout my dissertation have been invaluable and unsparing. I learnt **from him deep and persistent analyses of one's own work** and his review comments put a great increase to my knowledge.

These acknowledgements would be incomplete without thanking all my fellow classmates for their continuous support and encouragement to work hard. Finally I want thank to my guru and my family members for their never ending love and encouragement.

REFERENCES

[1]   Jennifer Yick, Biswanath Mukherjee, Dipak Ghosal, **"Wireless sensor network survey,"** *Computer Networks Elsevier 52* , pp 2292–2330, 2008

[2]   Kiran Maraiya, Kamal Kant, Nitin Gupta, "Application based Study on Wireless Sensor Network", *International Journal of Computer Applications (0975 – 8887)* Vol-21, No-8, May 2011

[3]   **K. Akkaya, M. Younis, "A survey on routing protocols for wireless sensor networks,"** *Ad Hoc Networks 3*, 2005

[4]   **Ma Di, Er Meng Joo, "A Survey of Machine Learning in Wireless Sensor Networks,"** *ICICS,* 2007

[5]   Raghavendra V. Kulkarni, Anna Förster, Ganesh **Kumar Venayagamoorthy, "Computational** Intelligence in Wireless Sensor Networks: A **Survey,"** *IEEE communications surveys & tutorials,* Vol-13, No-1, pp 68- 96, 2011

[6]   Justin A. Boyan and Michael L. Littman. Packet routing in dynamically changing networks: A reinforcement learning approach. In Advances in Neural Information Processing Systems 6, pages 671–678. Morgan Kaufmann, 1994.

[7]   **Forster, A. and Murphy, A. L.; "Balancing Energy Expenditure in WSNs through Reinforcement Learning: A Study" In** Proceedings of the 1st International Workshop on Energy in Wireless Sensor Networks (WEWSN); Santorini Island, Greece (p. 7pp), 2008

[8]   **Deville, M.; Borgne, Y and Nowe, A;** "Reinforcement Learning for Energy Efficient Routing in Wireless Sensor Networks" *Proceedings of the 23rd Benelux Conference on Artificial Intelligence,* 2011

[9]   P Beyens, K Steenhaut, and A Nowe. Routing with compression in wireless sensor networks: A qlearning approach. In Proceedings of the 5th European Workshop on Adaptive Agents and Multi-Agent Systems. AAMAS, 2005.y

BIOGRAPHIES

Mr. Ankit B. Patel received the Bachelor of Engineering degree from G H Patel College of Engineering and Technology, Vallabh Vidhyanagar, Gujarat Technological University, Ahmedabad, Gujarat, India. He is currently pursuing towards the Master of Engineering degree in Communication System Engineering at the L D College of Engineering,Ahmedabad, Gujarat, India. His research interests lie in the Machine Learning and Reinforcement Learning for Wireless Sensor Networks.

Dr. Hitesh B. Shah received the Bachelor of Engineering degree from Birla Vishwakarma Mahavidyalaya, Vallabh Vidhyanagar, Sardar Patel, University, Gujarat, India and Master of Technology degree from IIT, Roorkee. He completed Ph.D. degree from IIT, Delhi. He is a Professor in the Department of Electronics and Communication at G H Patel College of Engineering and Technology, Vallabh Vidhyanagar, Gujarat Technological University, Gujarat, India. His research interests lie in the Artificial Intelligence and Machine Learning.