

Dual Reinforcement Q Routing for Ad Hoc Networks

Rahul Desai^{*1}, B P Patil²

¹Sinhad College of Engineering, Army Institute of Technology, Pune, India

²E & TC Department, Army Institute of Technology, Pune, India

Corresponding author, email-desaimrahul@yahoo.com*

Abstract

Ad Hoc Networks are infrastructure less network in which nodes are connected by Multi-hop wireless links. Each node is acting as a router as it supports distributed routing. Routing challenges occurs as there are frequent path breaks due to the mobility. Various application domains include military applications, emergency search and rescue operations and collaborative computing. The existing protocols used are divided into proactive and on demand routing protocols. The various new routing algorithms are also designed to optimize the performance of a network in terms of various performance parameters. Dual reinforcement routing is learning based approach used for routing. This paper describes the implementation, mathematical evaluation and judging the performance of a network and analyze it to find the performance of a network.

Keywords: Ad Hoc Network, MANET, Q Routing, CQ Routing, CDRQ Routing

Copyright © 2017 Institute of Advanced Engineering and Science. All rights reserved.

1. Introduction

A wireless ad hoc network is a decentralized type of wireless network [1, 2] This network is ad hoc network because it does not rely on a pre existing infrastructure, such as routers in wired networks or access points in Infrastructured wireless networks. Each node participates in routing by forwarding data for other nodes, so the determination of which nodes forward data is made dynamically on the basis of network connectivity.

Routing consists of two steps; forwarding packets to the next hop and to decide how the forwarding process to reach the packets to the destination in minimum number of hops. To judge the merit of a routing protocol, qualitative and quantitative metrics are used to measure its suitability and performance. Various performance parameters such as packet delivery ratio, delay, jitter, control overhead etc are used judge the performance of routing protocols.

Ad hoc networks, due to their quick and economically less demanding deployment, find applications in many areas. Ad hoc networks can be very useful in establishing communication among a group of soldiers for tactical operations. Setting up a fixed infrastructure for communication among a group of soldiers in enemy territories may not be possible. In such cases, ad hoc wireless network provides the required communications quickly. It also includes the coordination of military objects moving at high speeds. Such applications require quick and reliable communication. Another domain in which the ad hoc wireless networks find applications is collaborative computing. The requirement of a temporary communication infrastructure for quick communication with minimal configuration among a group of people in a conference. Ad hoc wireless networks are also very useful in emergency operations such as search and rescue, crowd control and commando operations. Figure 1 shows an example of mobile ad hoc network which is an infrastructure less network.

The responsibilities of a routing protocol includes exchanging the route information, finding a feasible path to a destination based on criteria such as hop length, minimum power control and lifetime of the wireless link; gathering information about path breaks; mending the broken paths expending minimum processing power and bandwidth. The major challenges that a routing protocols for ad hoc network faces are mobility, bandwidth constrained, error-prone and shared channel, location dependent contention and other resource constraints such as power and buffer storage and link capabilities etc. The major requirements of a routing protocol in ad hoc wireless networks are minimum route acquisition delay, quick route reconfiguration,

loop free routing, distributed routing approach, minimum control overhead, scalability, provision of QoS, support for time sensitive traffic and security with privacy [3].



Figure 1. Example of Ad Hoc Wireless Network

Routing protocols for ad hoc networks can be classified into several types based on different criteria. The routing protocols are broadly classified into four categories based on routing information update mechanism, use of temporal information for routing, routing topology and Utilization of specific resources.

2. Classification of Routing Protocols

Based on the routing information update mechanism, they are classified into proactive or table driven routing protocols, reactive or on demand routing protocols and hybrid routing protocols. In table driven routing protocols, every node maintains the network topology information in the form of routing tables by periodically exchanging routing information. In on demand routing protocols, nodes obtain paths when it is required, by using a connection establishment process. Hybrid protocols combine the best features of above two categories.

Proactive routing protocols always find the optimum routes to reach to every destination nodes. But these types of protocols are not suitable for large network because of high overheads and their poor convergence behavior. Destination sequenced Distance Vector (DSDV) is one of earliest protocols developed for ad hoc networks [4, 5]. It is based on distance vector algorithm and uses sequence numbers to avoid count to infinity problem. Every node communicates, finds out their neighbors by sending hello messages and exchanges their routing tables with them. Periodic full updates and small updates are also transmitted to maintain routing tables up to date. Wireless Routing Protocol (WRP) is another distance vector protocol optimized for ad hoc networks. WRP belongs to a class of distance vector routing protocols called path finding algorithms. The algorithm of this class uses the next hop and second-to-last hop information to overcome the count-to-infinity problem.

Optimized link state routing protocol [6,7] is another proactive routing protocol based on link state algorithm. Here, every node broadcasts link state updates to every other node present in the network and thus creates link tables from which routing tables are designed. In order to reduce the overheads, multipoint relay concept is widely used. Figure 2 shows working of MPR. Node j chooses i , k , l and m as MPR nodes, since they are sufficient to reach all its two-hop neighbors.

In on demand routing protocols [8], route to the destination is obtained only when there is a need. When source nodes want to transmit data packets to the destination nodes, it initiates route discovery process. Route request (RREQ) messages float over the network and finally the packet reaches to the destination, Destination nodes replies with route reply message (RREP) and unicast towards the source node. All nodes including the source node keeps this route information in caches for future purpose. Dynamic Source Routing Protocol (DSR) is thus characterized by the use of source routing. The data packets carry the source route in the packet header. When the link or node goes down, existing route is no longer available; source node again initiates route discovery process to find out the optimum route. Route Error packets

and acknowledgement packets are also used. Ad Hoc on Demand Distance Vector Routing (AODV) is also on-demand routing protocol. It uses traditional routing tables, one entry per destination [9, 10]. In AODV, only one route path is available in routing table, if this path fails, it again initiates route discovery process to find out another optimum path. Route Request Message (RREQ) from the source to the destination and route reply message (RREP) from the destination to the source is shown in Figure 3 and Figure 4 respectively. To overcome this limitation, Ad Hoc Multipath Distance Vector (AOMDV) comes in picture.

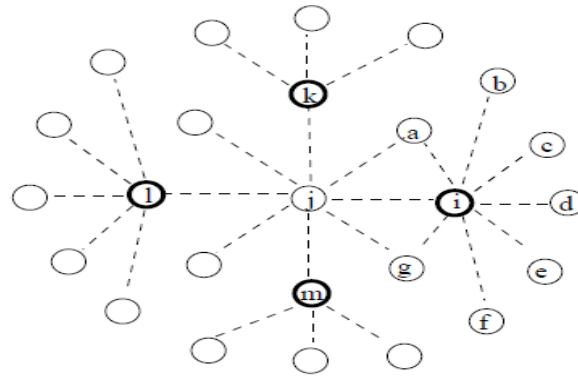


Figure 2. The Working of OLSR

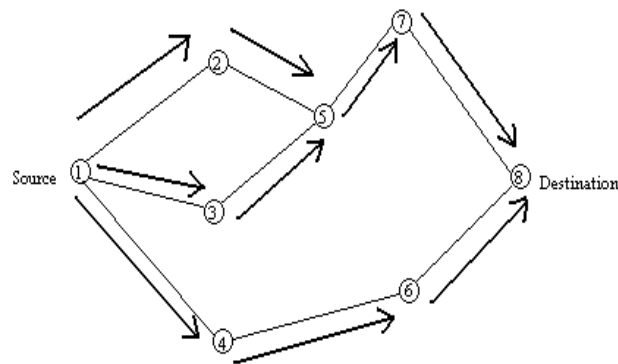


Figure 3. Route Request (RREQ) from source to destination

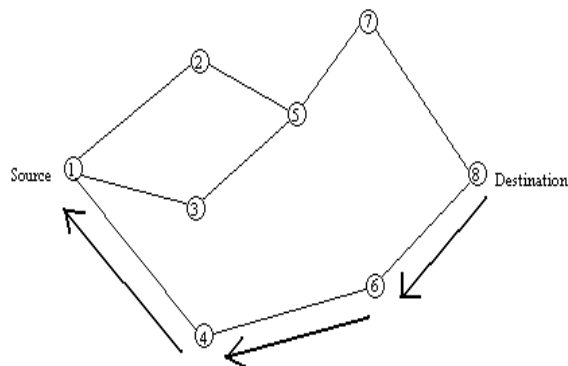


Figure 4. Route Reply (RREP) Reply from destination to source

3. Reinforcement Learning Based Routing Protocols

In section II, various routing protocols are introduced used for ad hoc wireless networks. In this section, new routing protocol based on reinforcement learning is introduced. Reinforcement learning is an example of the model-based approach where model of the system is learned in terms of Q values. These Q values are used to make decisions and these estimates are also updated in order to reflect the changes in the network. Thus entire routing tables are expressed in terms of Q values. Each Q value in the routing table is in the form of $Q(S, A)$ which represents the expected reinforcement of taking action 'A' in state 'S'. Thus Each node X in the network represents its own view of the state of the network through its Q table Q_x [11-12].

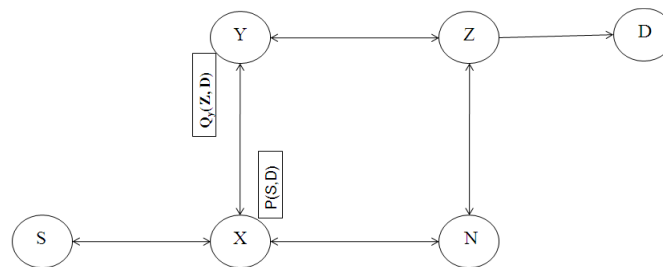


Figure 5. Reinforcement Based Learning Method – Q Routing

Assume that $P(S, D)$ is the packet that is to be transmitted from sender S node to Destination D node. $Send_Packet(X)$ function describes the action performed by C node while sending the packet while $Receive_Packet(Y)$ function described the action taken by node Y after receiving the packet.

Send_Packet(X) (assuming that queue is not full)

1. Receive the packet $P(S, D)$ and keep it in the queue.
2. Receive the packet from the queue to process it when its turn arises.
3. Find out the best neighbor by consulting its routing table.
4. Forward packet to best neighbor node obtained in step 3.
5. Receive estimate from best neighbor node and update its corresponding Q value.
6. Get ready to send next packet.

Receive_Packet(Y)

1. Receive a packet $P(S, D)$ from neighbor X.
2. Calculate best estimate for destination node and send back to node X.
3. If $(D = Y)$ then Consume Packet ($P(S, D)$) else append packet to packet Queue ($P(S, D)$)
4. Get ready for receiving next packet.

In step 3 of $Send_Packet(X)$, the best neighbor is obtained by Equation 1 and step 5 of $Send_Packet$ function, corresponding Q values is updated using Equation 2.

$$Q_y(z', d) = \min_{z \in N(y)} Q_y(z, d) \quad (1)$$

$$\Delta Q_x(y, d) = \eta_f (Q_y(z', d) + q_y - Q_x(y, d)) \quad (2)$$

$\Delta Q_x(y, d)$ is the new estimate value for node x to the destination d via the neighboring node d. This new estimation is calculated by subtracting old estimation value $Q_x(y, d)$ from the sum of the estimation time for packet travelling from node y to destination d via neighbor z ($Q_y(z, d)$) and current queue delay for the packet in node x (q). η_f is the learning rate parameter defined by the programmer [11].

Backward exploration together with the forward exploration is applied in DRQ algorithm in order to improve the learning rate of the Q-Routing algorithm. In DRQ, exact delay values learnt from the backward learning have also been used in the routing tables in addition to the estimation values learnt from forward learning in Q-Routing. This has doubled the learning information available in the algorithm thus improved the learning rate of the algorithm [11-14].

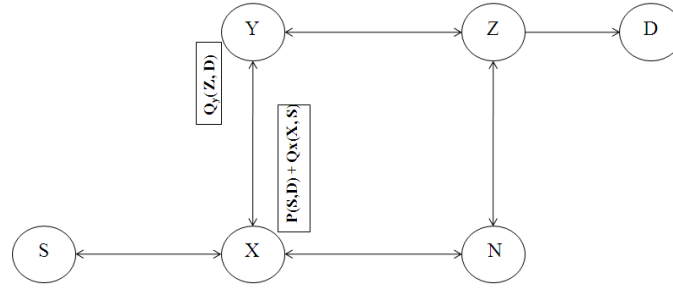


Figure 6. Reinforcement Based Learning Method-Dual Reinforcement Q Routing

Assume that $P(S, D)$ is the packet that is to be transmitted from sender S node to Destination D node. $Send_Packet(X)$ function describes the action performed by C node while sending the packet while $Receive_Packet(Y)$ function described the action taken by node Y after receiving the packet [12, 14].

Send_Packet(X) (assuming that queue is not full)

1. Receive the packet $P(S, D)$ and keep it in the queue.
2. Receive the packet from the queue to process it when its turn arises.
3. Find out the best neighbor by consulting its routing table.
4. Compute best estimate and append this estimate to the packet $P(S, D)$
5. Forward packet and best estimate to best neighbor node obtained in step 3.
6. Receive estimate from best neighbor node and update its corresponding Q value.
7. Get ready to send next packet.

Receive_Packet(Y)

1. Receive a packet $P(S, D)$ from neighbor X .
2. Extract estimate from packet $P(S, D)$ and update Q value.
3. Calculate best estimate for destination node and send back to node X .
4. If $(D = Y)$ then Consume Packet ($P(S, D)$) else append packet to packet Queue ($P(S, D)$)
5. Get ready for receiving next packet.

In DRQ algorithm, when X sends a packet to node Y to get its estimated remaining trip times, Y also gets X 's estimated trip times for its link with S .

$$Q_x(z', s) = \min_{z \in N(y)} Q_y(z, s) \quad (3)$$

$$\Delta Q_y(y, s) = \eta_b (Q_x(z', s) + q_y - Q_y(x, s)) \quad (4)$$

In Figure 6, packet at node x arriving from source node S is sent to node Y , also carries the estimated time that it takes from node X to s , $Q_x(Z, S)$ (Equation 3). With this information node Y updates its own estimate $Q_y(X, S)$ for the entry node X associated with the destination S (Equation 4). Therefore, in DRQ both backward and forward exploration can be used to update the Q entries.

In Q routing, some of the Q values (the Q values, which are just updated) are reliable and others may not be reliable. In Q routing, Q values are updated only when the packet is transmitted by the node in the network. If the packet is not transmitted for a longer time, Q values become less reliable. Thus the decision taken on such un-reliable Q values turns to be wrong and an optimum path may not be achieved. Hence to represent reliability of Q values, another value called as confidence value is also included. Thus every node contains two-tables- Q table that stores Q value and C table, which contains C values, which represents the reliability of Q values. If C value is one, this indicates, that corresponding Q value is 100% reliable (as packet is just transmitted by the network through this node) and confidence value of zero indicates that Q value is not trustable. The decision taken on such Q value may turn to be wrong and thus an optimum path may not be achieved. This confidence value should also decay after certain time representing that reliability of Q value is less.

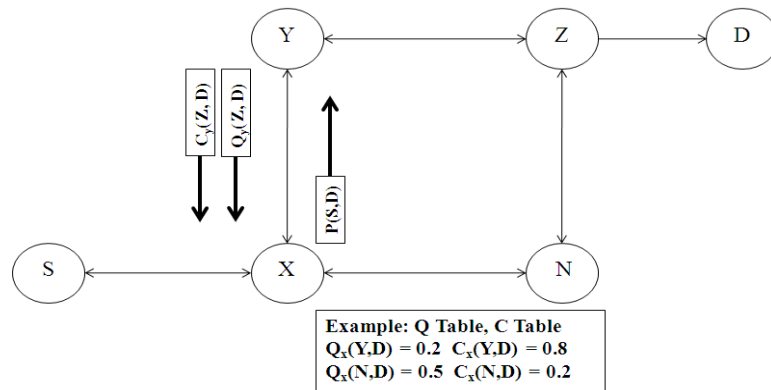


Figure 7. Confidence based Q routing

Thus every node transmitting packets will also receive C value along with their Q value, which is used to update old Q values and old C values (Figure 7). In standard Q routing, learning rate is fixed but in confidence based Q routing learning rate is a function of confidence values. When Q value with low confidence need to be updated, high learning rate should be used. If estimated C value is high, then the learning rate should be high. Learning rate should also be high if either: confidence value in the old Q value, C^{old} is low or confidence value in the new estimate Q value, C^{est} is high.

3. Results and Discussion

The experiment is performed using the simulator NS2 which is open source software and used to do research on wired and wireless networks. Experiment is performed on 6 by 6 irregular grid (Figure 8). In 6 by 6 irregular grid, there is left cluster and right cluster. Left cluster consist of nodes 1 through 10 while right cluster consists of nodes 25 through 36. There are two possible routes, route 1 consisting of nodes 12 and 25 and route 2 consisting of nodes 18 and 19. The shortest path routing algorithm always selects route 1 as shortest path routing algorithm select the path having minimum number of hops.

Figure 9 shows average packet delivery time (APDT) for low load, Figure 10 and 11 shows average packet delivery time for medium and high load respectively.

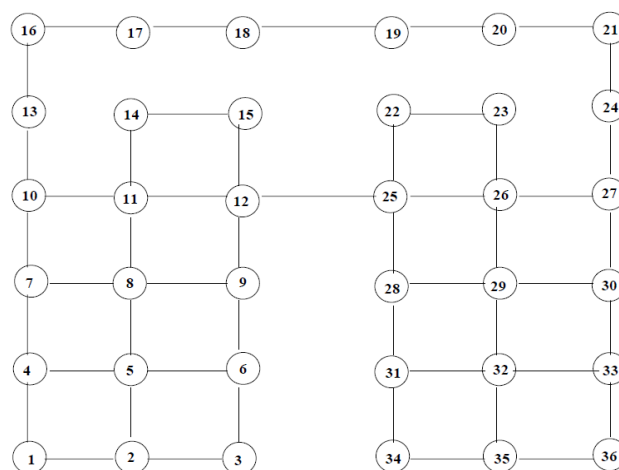


Figure 8. Network of 6 by 6-irregular grid

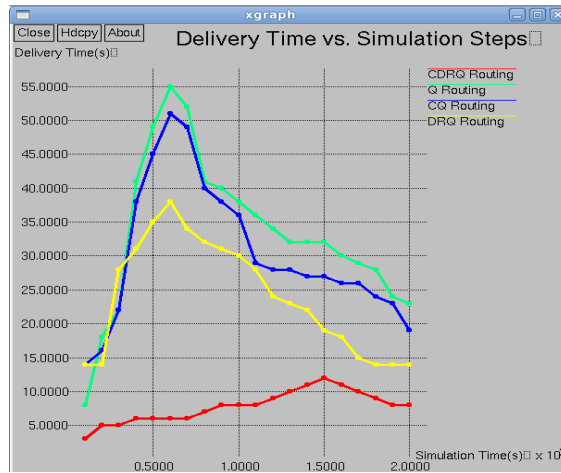


Figure 9. Average Packet Delivery Time for Low Loads

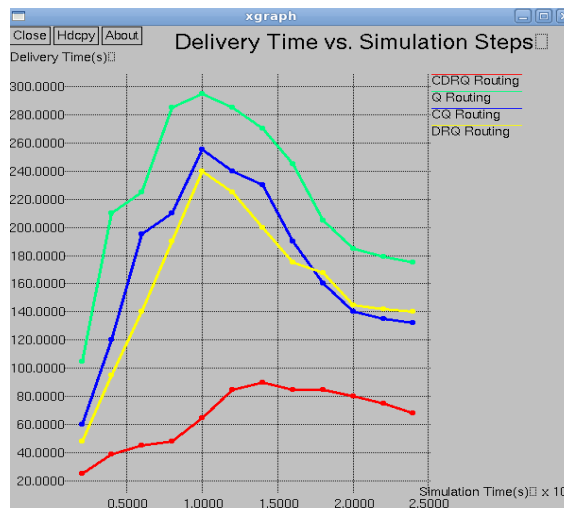


Figure 10. Average Packet Delivery Time for Medium Loads

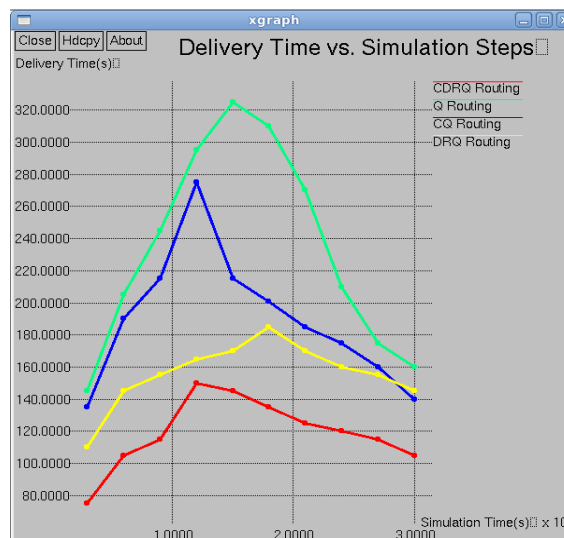


Figure 11. Average Packet Delivery Time for High Loads

At low loads, shortest path always gives best performance as shortest path always selects the path having minimum number of hops and there is low traffic in the network. Initially Q values in Q table are zero, so packets are transmitted randomly. Some amount of time, it takes to settle down Q table in the network to represent real state of the network. There is initial learning phase till Q values settle down to their optimum values. At initial phase average packet delivery time is large and once these Q values settle down to their optimum values, average packet delivery time decreases rapidly. CQ routing gives better performance as compared with Q routing, as Q values are made more reliable as compared with Q routing. DRQ routing is faster as compared Q routing, as Q values settle down as it involves exploration in both direction. CDRQ gives better performance as it includes confidence values and dual reinforcement. It is also observed, that at medium and high loads, Q values converges very slowly to their optimum values as compared with CDRQ routing at low loads.

Experiment is carried out on 50 nodes MANET by changing the interval from 0.010 to 0.014. The simulation is carried out for 200 seconds. The size of packet is 512 bytes. The results obtained are shown in Figure 12 to 13. PDR obtained in Q routing is in range of 12% to 19% and it is improved in CDRQ routing from 34% to 44%.

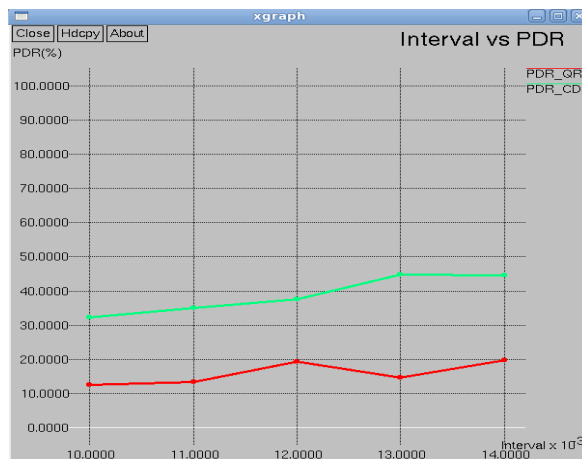


Figure 12. Interval vs. PDR

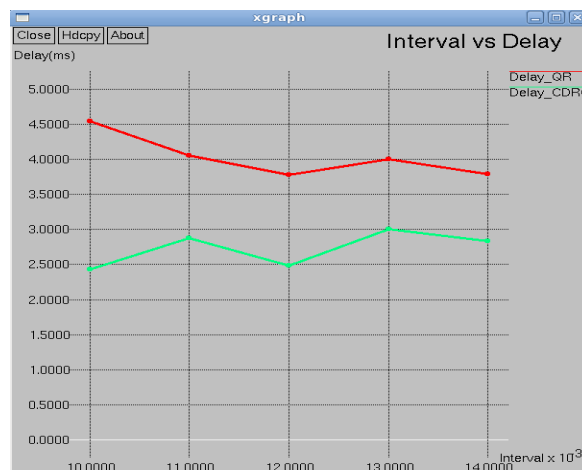


Figure 13. Interval vs. Delay

Comparatively, CDRQ routing protocol gives less delay as compared with Q routing, this is because the adaption time required for Q tables in CDRQ routing is less. Table 1 represents the comparative values of PDR and end-to-end delay for Q routing and CDRQ routing.

Table 1. Evaluation of CDRQ routing Protocol on 50 Nodes MANET

Interval (s)	Packet Delivery Ratio (%)		End to End Delay (ms)	
	Q Routing	CDRQ	Q Routing	CDRQ
0.010	12.66	32.37	4.544	2.427
0.011	13.37	35.07	4.048	2.871
0.012	19.45	37.56	3.773	2.489
0.013	14.73	44.70	4.002	3.002
0.014	19.84	44.60	3.785	2.836

5. Conclusion

This paper explains the existing routing protocols such as DSDV, AODV, DSR and OLSR which are based on shortest path. Also comparative analysis of Q routing and CDRQ routing is done on 6 by 6 irregular grid and 50 nodes mobile ad hoc network with random mobility. PDR and delay are very important parameters when deciding how a reliable a protocols works. CDRQ provides very good results as compared with Q routing because increased exploration and exploitation. CDRQ is much suitable for medium and high traffic where shortest path routing fails to work.

References

- [1] Mukhtiar Ahmed, Mazleena Salleh, M.Ibrahim Channa, Mohd Foad Rohani, "Review on localization based Routing Protocols for Underwater Wireless Sensor Network", *International Journal of Electrical and Computer Engineering*, Vol 7, No 1: February 2017
- [2] Dilip Singh Sisodia, Riya Singhal, Vijay Khandal, "A Performance Review of Intra and Inter-Group MANET Routing Protocols under Varying Speed of Nodes " *International Journal of Electrical and Computer Engineering*, Vol 7, No 5: October 2017
- [3] S. Murthy and J. J. Garcia-Luna-Aceves, "An Efficient Routing Protocol for Wireless Networks," *ACM J. Mobile Networks and Applications, special issue on Routing in Mobile Communication Networks*, vol. 1, no. 2, 1996, pp. 183–97.
- [4] Justin Sophia I, N. Rama, "Improving the Proactive Routing Protocol using Depth First Iterative Deepening Spanning Tree in Mobile Ad Hoc Network", *International Journal of Electrical and Computer Engineering*, Vol 7, No 1: February 2017
- [5] Rahul Desai, B P Patil, "Analysis of Reinforcement Based Adaptive Routing in MANET", *Indonesian Journal of Electrical Engineering and Computer Science*, Vol 2, No 3: June 2016
- [6] P. Jacquet et al., "Optimized Link State Routing Protocol for Ad Hoc Networks," Proc. IEEE Int'l. Multi Topic Conf. (INMIC '01), 2001, pp. 62–68.
- [7] T. Clausen and P. Jacquet, Optimized Link State Routing Protocol (OLSR), document RFC 3626, IETF, Oct. 2003. [Online]. Available: <http://www.ietf.org/rfc/rfc3626.txt>.
- [8] Reji Mano, P.C. Kishore Raja, Christeena Joseph, Radhika Baskar, "Hardware Implementation of Intrusion Detection System for Ad-Hoc Network", *International Journal of Reconfigurable and Embedded Systems (IJRES)*, Vol 5, No 3: November 2016
- [9] Shalini Singh, Rajeev Tripathi, "Performance Analysis of Extended AODV with IEEE802.11e HCCA to support QoS in Hybrid Network", *Indonesian Journal of Electrical Engineering and Computer Science*, Vol 12, No 9: September 2014
- [10] AL-Gabri Malek, Chunlin LI, Layuan Li, "Improving ZigBee AODV Mesh Routing Algorithm Topology and Simulation Analysis", *TELKOMNIKA Indonesian Journal of Electrical Engineering* Vol.12, No.2, February 2014, pp. 1528 ~ 1535
- [11] S. Kumar, "Confidence based dual reinforcement Q-routing: an on-line adaptive network routing algorithm," MS thesis. University of Texas at Austin, 1998.
- [12] S. Kumar and R. Miikkulainen, "Dual Reinforcement Q-Routing: An On-Line Adaptive Routing Algorithm," *Artificial neural networks in engineering*, 1997
- [13] N. Coutinho, R. Matos, C. Marques, A. Reis, S. Sargento, J. Chakareski, and A. Kessler, "Dynamic dual-reinforcement-learning routing strategies for quality of experience-aware wireless mesh networking," *Computer Networks*, vol. 88, pp. 269–285, 2015.
- [14] Rahul Desai, B P Patil, "Prioritized Sweeping Reinforcement Learning Based Routing for MANETs", *Indonesian Journal of Electrical Engineering and Computer Science* Vol. 5, No. 2, Feb 2017, pp. 684 ~ 694 DOI: 10.11591/ijeecs.v2.i3.pp648-694.