Indonesian Journal of Electrical Engineering and Computer Science
Vol. 7, No. 3, September 2017, pp. 718 ~ 723
DOI: 10.11591/ijeecs.v7.i3.pp718-723

Learning Based Route Management in Mobile Ad Hoc Networks

Rahul Desai[*1], B P Patil[2], Davinder Pal Sharma[3]
[1]Sinhad College of Engineering, Army Institute of Technology, Pune, INDIA
[2]E & TC Department, Army Institute of Technology, Pune, INDIA
[3]The University of the West Indies, Department of Physics St. Augustine, St. George, TT
*Corrsponding author, email-desaimrahul@yahoo.com

Abstract

Ad hoc networks are mobile wireless networks where each node is acting as a router. The existing routing protocols such as Destination sequences distance vector (DSDV), Optimized list state routing protocols (OLSR), Ad hoc on demand routing protocol (AODV), dynamic source routing (DSR) are optimized versions of distance vector or link state routing protocols. Reinforcement Learning is new method evolved recently which is learning from interaction with an environment. Q Learning which is based on reinforcement learning that learns from the delayed reinforcements and becomes more popular in areas of networking. Q Learning is applied to the routing algorithms where the routing tables in the distance vector algorithms are replaced by the estimation tables called as Q values. These Q values are based on the link delay. In this paper, various optimization techniques over Q routing are described in detail with their algorithms.

***Keywords**: Q Routing, Reinforcement, CQ routing, DRQ routing, CDRQ routing, DSR, AODV, DSDV*

1. Introduction

An ad Hoc network is a technology where no fixed infrastructure is required; all nodes are mobile, thus moving from one network to another [1, 2]. Ad hoc network is a temporary network where each node is also acting as a router. All nodes are self configured (addresses and routing features) nodes, multiple hops are required to transfer data from one node to another. Energy is also one of the most important parameter as all nodes have limited power supply. Ad hoc network characteristics includes Peer-to-peer, zero administration, low power, Multihop, dynamic and auto configured. 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.

There are two types of protocols–proactive routing protocols and on demand also known as reactive routing protocols are widely adopted for an ad hoc network. Proactive protocols always maintains routing paths between all pairs of nodes irrespective of their usage while reactive protocols finds out the path to reach to the node only when needed. Pro-active 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 behaviour. Destination sequenced Distance Vector (DSDV) is one of earliest protocols developed for ad hoc networks [3, 4]. It is based on distance vector algorithm and uses sequence numbers to avoid count to infinity problem. Every node communicates, finds out their neighbours 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.

Optimized link state routing protocol [5, 6] 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. There are two types of

Received May 29, 2017; Revised August 15, 2017; Accepted August 30, 2017

algorithms which are widely used for wires as well as wireless networks, first is distance vectors routing protocols, where the distances in terms of number of hops are communicated to the neighbours and builds up the routing table. Routing tables basically consists of three columns, first column for destination node, second column will be the next hop where the packet are to be delivered, third column stands for metric or cost.

In on demand routing protocols, 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. 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 [7-9].

2. Survey of Reinforcement Based Routing Methods

Reinforcement learning is the process of mapping the situations to the actions and tries to maximize a reward signal. There are various strategies such as positive or negative approaches as well as model based or model free approaches are used. Q Routing is new evolved concept arises in the modern world which is also based on reinforcement concepts. Each node in the network contains reinforcement learning module which tries to find out the optimum path to the destination. Direct or indirect training signal is required to improve the routing policy. As illustrated in Figure 1, Let $Q_X(Y, D)$ represents the time that a node X takes to deliver a packet P to the destination node D when the packet is transmitted to the next neighbour node Y. After sending the packet, node X will also get node Y’s estimate of the remaining time in the network [10-11].

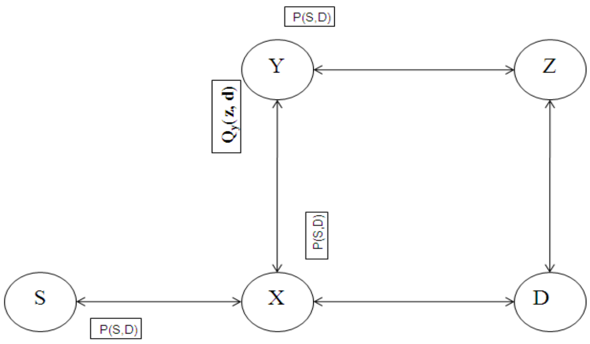

Figure 1. Q routing

In Reinforcement Learning (Q Routing) each node maintains database of Q values which represents delays for each of the next hops. For every incoming packet, nodes consult its Q table and decided the next hop based on the least delivery time required to reach the packet to the destination [10, 11]. At the same time the sending node receives the estimate of the remaining delivery time for the packet to the destination. Thus after every packet transmitted by the source node and all intermediate nodes[11] Q values are received by these nodes and updates their Q table to represents the steady state of the network. As soon as the node X sends a packet P to the destination node D to one of the neighbouring nodes Y, node Y send back to node X, its best estimate $Q_Y(Z, D)$ for the destination D. $Q_Y(Z, D)$ for the destination D shows its remaining time required to reach the packet to the destination node D [10,11].

PacketSend (X)
1 Receive the packet from Packet Queue
2 Find out the best neighbour Y = min(Q_x(Y, D))
3 Forward Packet to the neighbour Y
4 Receive Estimate (Q_y(Z, D) + q_y) from node Y.
5 Update Q value Q_x(Y, D).

PacketReceive (Y)
1 Receive a packet from neighbour X
2 Calculate best estimate for node D; Q_y(Z, D) and send back to node X.
3 Get ready for receiving next packet

By adding confidence measure, the quality of exploration is improved by learning faster thus Q values represent the current state of the network more closely. Each node in the network contains C tables consisting of confidence values, where each Q value is associated with C value. This value is the real number between 0-1 and essentially specifies the confidence in the corresponding Q value [10]

In standard Q routing, learning rate is always maintained to be constant, it means there is way to specify reliability of Q values but in Confidence based Q Routing, the learning rate depends on the confidence of the Q value being updated and its new estimate. In particular, when node X sends a packet to its neighbour Y, it also gets back the confidence value C_y(Z, D) associated with this Q value. When node X updates its Q_x(Y, D) value, it first computes the learning rate Π which depends on both C_x (Y, D) and C_y (Z, D). Simple and effective learning rate function is given by: Πf (C_{old}, C_{new}) = max (C_{new}, 1- C_{old}). The confidence value always represents the reliability of the corresponding Q value, and thus always changes with time. This confidence value decays with time if their Q values are not updated in the last time step [11].

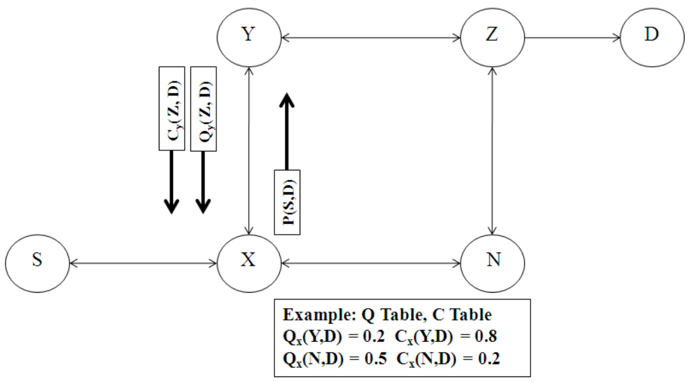

Figure 2. CQ routing

In confidence based Q Routing, algorithm for Packet Send and Packet Receive can be summarized as follows [10-12].

PacketSend(X)
1. Receive the packet from Packet Queue
2. Find out the best neighbour Y = min(Q_x(Y, D))
3. Forward Packet to the neighbour Y
4. Receive Estimate (Q_y(Z, D) + q_y) and C_y(Z, D) from node Y.
5. Update Q value Q_x(Y, D) and C_y(Z, D) value.

PacketReceive(Y)
1. Receive a packet from neighbour X
2. Calculate best estimate for node D; Q_y(Z, D) and send back to node X.
3. Find the corresponding confidence value C_y(Z, D) and send back to node X.
4. Get ready for receiving next packet

In Dual reinforcement Q Routing (DRQ) the learning process occurs in both ways and thus the learning performance of the Q Routing algorithm doubles. Instead of using a single reinforcement signal, an indirect reinforcement signal extracted from the incoming information is also used to update the state of the network. When a node X sends a packet to neighbour node Y, it will also send additional routing information which will be used to update node Y's decisions in opposite direction. Thus backward exploration is also added to standard Q Routing [11] Figure 3 illustrates the backward exploration in standard Q routing.

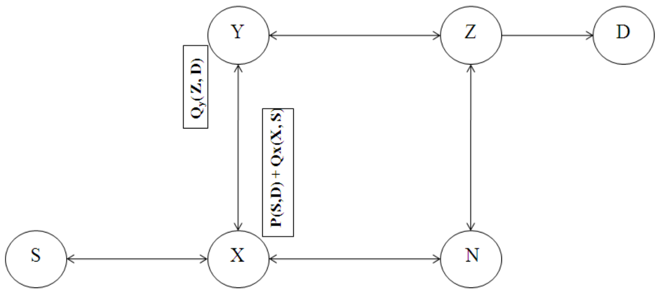

Figure 3. CQ routing

In dual reinforcement confidence based Q Routing, algorithm for Packet Send and Packet Receive can be summarized as follows [10-12].

PacketSend (X)

1. Receive the packet from Packet Queue
2. Find out the best estimate $Q_x(Z, d)$
3. Append $(Q_x(Z, S)+q_x)$ and $C_x(Z, S)$ to the packet P(S, D).
4. Find out the best neighbour Y = min($Q_x(Y, D)$)
5. Forward Packet to the neighbour Y
6. Receive Estimate $(Q_y(Z, D) + q_y)$ and $C_y(Z, D)$ from node Y.
7. Update Q value $Q_x(Y, D)$ and C value $C_x(Y, D)$.

PacketReceive (Y)

1. Receive a packet from neighbour X
2. Using the received estimate $Q_x (Z, D) + q_x$ and $C_x (Z, D)$ update Q value $Q_y(X, S)$ and $C_y(X,)$.
3. Calculate best estimate for node D; $Q_y (Z, D)$ and $C_y (Z, D)$, send back to node X.
4. Get ready for receiving next packet.

Thus confidence values are used not only for exploration but also in making routing decisions [12].

3. Analysis

The experiment is performed using the simulator NS2 which is open source software and used to do research on wired and wireless networks. The number of nodes varies from 10 to 100. The topology Size is 1000 m × 1000 m. The simulation time is 200 seconds. DSDV, DSR, AODV, Dual reinforcement Q routing protocols are analysed.

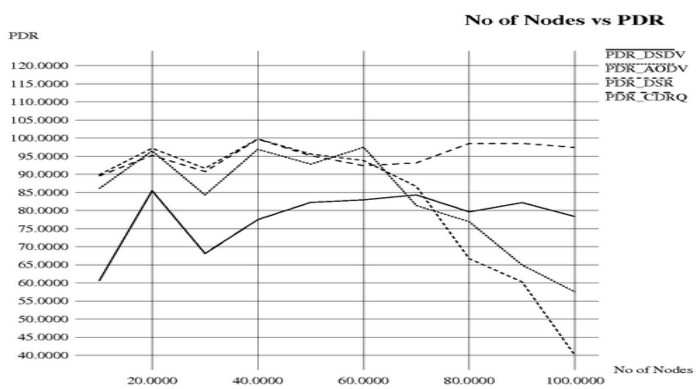

Figure 4. No of Nodes vs PDR

It is observed that when the network size increases beyond 60 nodes, AODV or DSR protocols starts dropping packets. But CDRQ protocols maintains consistent ratio throughout the network irrespective of the network size.

End-to-end Delay is the time taken by a data packet to reach to the destination. The result of end to end delay is illustrated in Figure 5. Here again dual reinforcement confidence based routing provides low delay compared with standard routing and other non optimized variants of Q routing.

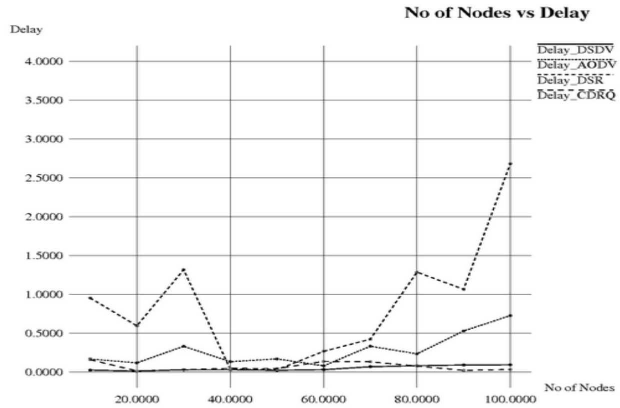

Figure 5. No of Nodes vs. Delay

4. Conclusion

This paper explains the comparative analysis of various optimized versions of existing routing protocols with dual reinforcement confidence based Q routing in NS2. This research study compares DSDV, AODV and DSR protocols with CDRQ routing protocols for an ad hoc network. PDR and delay are very important parameters when deciding how a reliable a protocols works. CDRQ variant based on reinforcement learning shows significant results as compared with existing routing protocols.

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] 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.

[4] 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.

[5] 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.

[6] 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.

[7] 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.

[8] 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.

[9] 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, 2014, pp. 1528~1535.
[10] S. Kumar. *Confidence based dual reinforcement Q-routing: an on-line adaptive network routing algorithm*. MS thesis. University of Texas at Austin, 1998.
[11] S. Kumar, R. Miikkulainen. Dual Reinforcement Q-Routing: An On-Line Adaptive Routing Algorithm. *Artificial Neural Networks in Engineering*, 1997.
[12] 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.