Flexible and Scalable Routing Approach for Mobile Ad Hoc Networks Based on Reinforcement Learning

Rahul Desai [1], B P Patil [2]

[1] *Research Scholar, Sinhgad College of Engg Asst Professor, Army Institute of Technology, Savitribai Phule Pune University, India*

[2] *Professor, E & TC Department, Army Institute of Technology, Savitribai Phule Pune University, India*

---***---

Abstract - *This paper describes and evaluates the performance of various reinforcement learning algorithms with shortest path algorithms that are widely used for routing packets through the network. In high traffic or high mobility conditions, the shortest path get flooded with huge number of packets and congestion occurs, so such shortest path does not provides the real shortest path and increases delay for reaching the packets to the destination. Reinforcement learning algorithms are adaptive algorithms where the path is selected based on the traffic present on the network at real time. Thus they guarantee the least delivery time to reach the packets to the destination. Analysis done on a 6 by 6 irregular grid and sample ad hoc network shows that performance parameters used for judging the network such as packet delivery ratio, delay etc provides optimum results using reinforcement learning algorithms.*

Key Words: Packet Delivery Ratio, CQ Routing, Delay, DRQ Routing, Q Routing

1. INTRODUCTION

Routing is the process of transmitting packets from one network to another. The most simplest and effective policy used is the shortest path routing. This policy is not always good as there are some intermediate nodes present in the network that are always get flooded with huge number of packets. In such cases, it is always better to select alternate path for transmitting the packets. Such routes are dynamically selected in real time based on the actual traffic present on the network. Hence when the more traffic is present on some popular routes, some un-popular routes must be selected for delivering the packets. For example, in order to demonstrate limitation of shortest path algorithms (fig 1), consider that Node 0, Node 9 and Node 15 are simultaneously transferring data to Node 20. Route Node 15-16-17-18-19-20 get flooded with huge number of packets and then it starts dropping the packets. Thus shortest path routing is non-adaptive routing algorithm that does not take care of traffic present on some popular routes of the network.

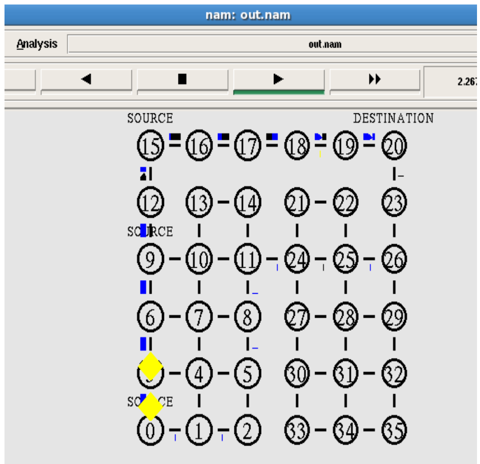

Fig 1: Limitation of Shortest Path Algorithms

The main goal is to optimize the delivery time for the packets to reach to the destination and preventing the network to go into the congestion. There is no training signal available for deciding optimum policy at run time, instead decision must be taken when the packets are routed and packets reaches to the destination on popular routes.

2. LITERATURE SURVEY OF VARIOUS REINFORCEMENT LEARNING ALGORITHMS

Reinforcement learning is learning where the mapping between situations to actions is carried out so as to maximize rewards. Q Routing is reinforcement based learning algorithm. It is based on the Q learning principle in order to learn the optimal path to the destination. Each node in the network has a reinforcement learning module to dynamically determine the optimum path to the destination [1,2].

In Q Routing, each node maintains information about Q values for each of the possible next hops[3]. These Q values represents the delivery time for the packets to reach to the destination. For every packet, the node makes a choice of the next hop that has the least estimate of the time it takes to reach the destination. Also, an update is also sent to the previous node regarding the present Q value. In order to keep the Q value as close to the actual values as possible and to reflect the changes in the state of the network, the Q value estimates need to be updated with minimum possible overhead. Fig 2 shows Q routing forward exploration. As soon as the node X sends a packet P(S, D) destined for node D to one of the neighboring nodes Y, node Y send back to node X, its best estimate Qy(Z, D) for the destination D. This value essentially estimates the remaining time in the journey of packet P(S, D). Upon receiving Qy(Z, D), node X computes the new estimate[4-5].

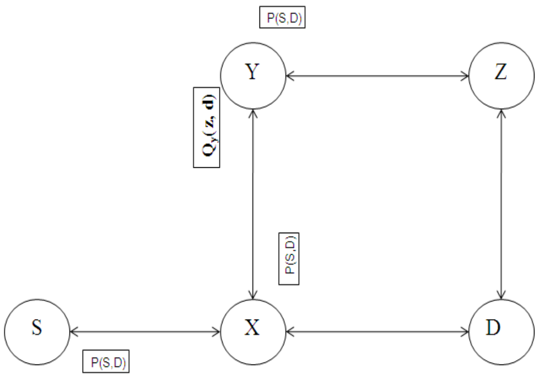

Fig 2: Q routing Forward Exploration

In another optimized form, Confidence Based Q Routing, each Q value Qx(Y, D) in the network is associated with a measure of confidence Cx(Y, D), which is a real number between 0 and 1. A value of 1 means that there is full confidence in the corresponding Q value and that this Q value reflects the current state of the network. In other words, this Q value has recently been updated. A value of 0, on the other hand, means that the corresponding Q value is random and does not necessarily reflect anything about the current state of the network. In other words, this Q value has never been updated. All Intermediate nodes along with Q value, also transmits Confidence value which will updated in confidence table [4-6]. Fig 3 shows Confidence based Q routing forward exploration.

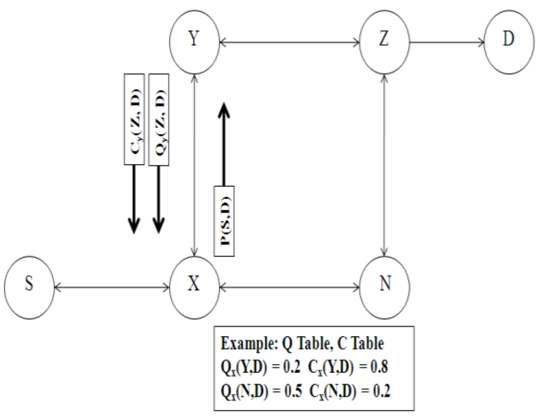

Fig 3: Confidence Q routing

Dual reinforcement Q Routing (DRQ) is a modified version of the Q-Routing algorithm, where learning occurs in both ways. Since, the learning process occurs in both ways the learning performance of the Q-Routing algorithm doubles. Instead of trying to use the single reinforcement signal, an indirect reinforcement signal is extracted from the incoming information and is used to update the local decision maker. When a node X sends a packet to neighboring node Y, some additional routing information can be sent along with the packet. This information can be used to update node Y's decisions in the direction opposite to the direction of the packet. This update adds backward exploration to Q Routing [4-6]. Fig 4 shows Dual reinforcement Q routing which involves backward exploration.

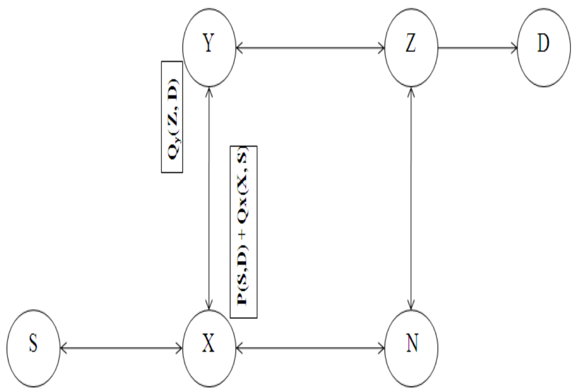

Fig 4: Dual Reinforcement Based Q routing

3. ANALYSIS AND PERFORMANCE COMPARISON

Network Simulator NS2 is used for experimentation. NS2 is the standard network simulator used for analysis of wired and wireless networks. Two different experiments are performed to judge the quality of reinforcement learning algorithms using different performance parameters, in first experiment 6×6 irregular grid is used to test the performance of reinforcement learning for random traffic. Second experiment is performed on an ad hoc network consisting of 10 to 100 nodes with random mobility of nodes and random traffic generated on the network.

The network topology used is the 6×6 irregular grid shown in the fig 5. In this network there are two possible ways of routing packets between the left cluster (nodes 1 through 18) and the right cluster (nodes 19 through 36): the route including nodes 12 and 25 (R1) and the route including nodes 18 and 19 (R2). For every pair of source and destination nodes in different clusters, either of the two routes, R1 or R2 can be chosen.

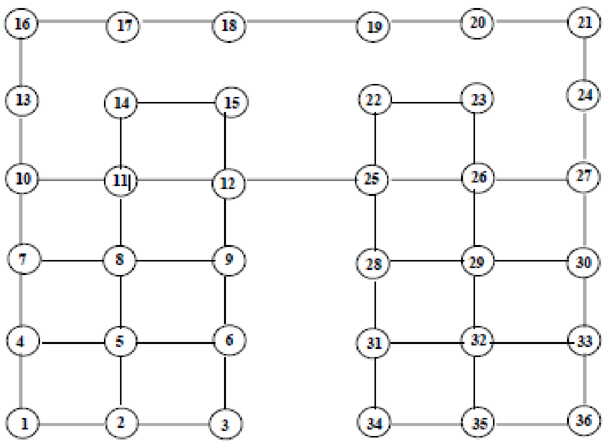

Fig 5: The 6×6 Irregular Grid

Route R1 is always selected by shortest path routing. For low loads (1 packet/simulation time/connection), shortest path routing is giving best results. Average packet delivery time is very less as compared with reinforcement learning methods (fig 6)

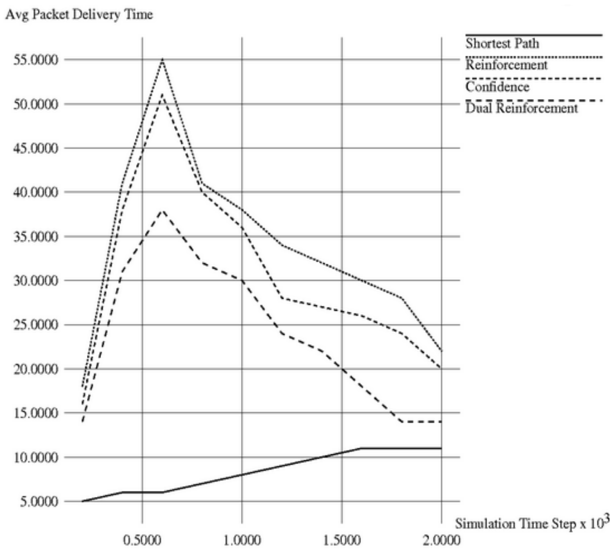

Fig 6: Average Packet Delivery Time vs. Simulation Time step for low loads.

At medium (2 packets/simulation time/connection) (fig 7) , or high load conditions (3 packets/simulation time/connection) (fig 8), imposed over the network, it is found that the shortest path routing breaks down and the average packet delivery time grows linearly as the simulation time progresses. This is because the packet queues at particular nodes 12 and 25 increases without bound. A lot of queue waiting time is incurred by packets going through these nodes. In reinforcement routing, simulation time steps 1500 to 2000 are required to find out the optimum paths, and there after they settle down to most stable routing policy.

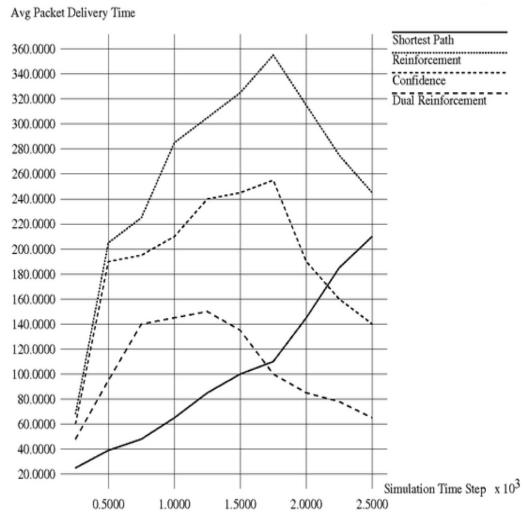

Fig 7: Average Packet Delivery Time vs. Simulation Time step for Medium loads

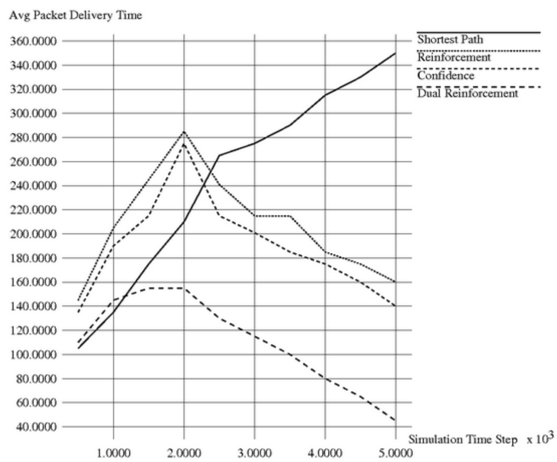

Fig 8: Average Packet Delivery Time vs. Simulation Time step for High loads

3. CONCLUSIONS

In this paper, various reinforcement learning algorithms were presented. Confidence based reinforcement and Dual reinforcement routing are showing prominent results as compared with shortest path routing for medium and high load conditions. At high loads, dual reinforcement Q routing performs more than twice as fast as Q-Routing. In dual reinforcement routing, as backward exploration is involved including confidence measure, less time is required in order to settle down the Q values thus they more accurately predict the state of network at run time. It is found that, though mobility rate changes at high rate as well as high traffic, dual reinforcement routing obtains more accurate result as compared with Q routing.

REFERENCES

[1] Richard S. Sutton and Andrew G. Barto "Reinforcement Learning: An Introduction" A Bradford Book, The MIT Press Cambridge, Massachusetts London, England

[2] Ouzecki, D.; Jevtic, D., "Reinforcement learning as adaptive network routing of mobile agents," MIPRO, 2010 Proceedings of the 33rd International Convention , pp.479,484, 24-28 May 2010

[3] Ramzi A. Haraty and Badieh Traboulsi "MANET with the Q-Routing Protocol" ICN 2012 : The Eleventh International Conference on Networks

[4] S Kumar, Confidence based Dual Reinforcement Q Routing : An on line Adaptive Network Routing Algorithm. Technical Report, University of Texas, Austin 1998.

[5] Kumar, S., 1998, "Confidence based Dual Reinforcement Q-Routing: An On-line Adaptive Network Routing Algorithm," Master's thesis, Department of Computer Sciences, The University of Texas at Austin, Austin, TX-78712, USA Tech. Report AI98-267.

[6] Kumar, S., Miikkulainen, R., 1997, "Dual Reinforcement Q-Routing: An On-line Adaptive Routing Algorithm," Proc. Proceedings of the Artificial Neural Networks in Engineering Conference.