International Journal of Business Data Communications and Networking
Volume 17 • Issue 1 • January-June 2021

Disaster Relief Management Using Reinforcement Learning-Based Routing

Gajanan Madhavrao Walunjkar, Vel Tech Rangarajan Dr. Sagunthala R&D Institute of Science and Technology, Chennai, India & AIT, Pune, India

https://orcid.org/0000-0003-0177-5206

Anne Koteswara Rao, Vel Tech Rangarajan Dr. Sagunthala R&D Institute of Science and Technology, Chennai, India

V. Srinivasa Rao, Vel Tech Rangarajan Dr. Sagunthala R&D Institute of Science and Technology, Chennai, India

ABSTRACT

Effective disaster management is required for the peoples who are trapped in the disaster scenario but unfortunately when disaster situation occurs the infrastructure support is no longer available to the rescue team. Ad hoc networks which are infrastructure-less networks can easily deploy in such situation. In disaster area mobility model, disaster area is divided into different zones such as incident zone, casualty treatment zones, transport areas, hospital zones, etc. Also, in order to tackle high mobility of nodes and frequent failure of links in a network, there is a need of adaptive routing protocol. Reinforcement learning is used to design such adaptive routing protocol which shows good improvement in packet delivery ratio, delay and average energy consumed.

KEYWORDS

AODV, AOMDV, CDRQ, CQ, DM, DRQ, DSDV, DSR, Q Routing

1. INTRODUCTION

Effective disaster management plays very vital role in saving lives of the people who are trapped in the disaster scenario. To carry out complete disaster operation successfully the lifetime of the communication network must be enough. But unfortunately when disaster occurs the infrastructure support is no longer available to the rescue team. Ad hoc networks are more preferred in such applications where the laying down the infrastructure network is not possible either due to short period of time or any emergency situations such as battlefield, military and disaster scenario. The various advantages gained from mobile ad hoc networks over wireless communication are cost factor, resource sharing, good QoS, security and reliability.

The various characteristic of mobile ad hoc network are:

1. **Multi-Hopping:** Packets moves through intermediate nodes to reach to the destination.
2. **Mobility:** Nodes can move randomly which frequently changes the topology.
3. **Self-Organization:** Nodes are intelligent and configures without the support of external entity.

DOI: 10.4018/IJBDCN.2021010102

4. **Limited Energy:** Nodes are operated through battery which has limited energy. Energy of the node should be conserved for increasing the overall lifetime of network.
5. **Bandwidth Constrained:** Frequent breakages of communication link changes the topology and again reestablishment of communication link is required. This increases the bandwidth consumption.

Various existing routing protocols are divided into two classes – proactive and reactive (Nossenson & Schwartz, 2013). Proactive protocols always maintain paths and thus consistent information is always available in routing database. Destination-Sequenced Distance Vector (DSDV) comes under this category. Reactive routing protocols (Recent Trends in Networks and Communications, 2010) which are on-demand protocols where the actual communication path is obtained only when never it is required. Dynamic source routing (DSR), ad hoc on-demand distance vector (AODV), ad hoc on demand multipath (AOMDV) comes under this category.

In DSDV, routes are updated periodically by triggering to their neighbors. But it is observed, that routing and caching overhead is high and throughput is low in DSDV. AODV is more suitable and adaptable to large highly dynamic topologies. DSR reduces route discovery overhead but creates more delay to the packets reaching to the destinations. It is also observed, in high mobility, on-demand routing protocols gives good performance. The DSDV and AOMDV generate high number of control packets and thus consume more bandwidth.

There are several mobility models present. Random waypoint model includes select random direction, speed and variable pause time. The node moves towards random destination with a velocity chosen randomly from [0, V_{max}]. In Random walk mobility model, node moves to a new location by randomly choosing a direction and speed. This is similar to RWP but nodes change their speed/direction every time slot, New direction θ is chosen randomly between (0, 2π], New speed chosen from uniform (or Gaussian) distribution (Handbook of Mobile Ad Hoc Networks for Mobility Models, 2011). When node reaches boundary it bounces back with (π-θ). In Manhattan Grid Model nodes move only on predefined paths. In disaster mobility model, disaster area is divided into different zones such as incident zone, casualties treatment zone, transport area and hospital zone (Walunjkar & Rao, 2019a) etc. Figure 1 shows the division of zones in disaster area scenario.

This paper is organized as follows: Section 2 introduces the problem and the issue about non adaptive routing and various methods of designs of adaptive routing. Section 2 also describes the need of further optimization required on adaptive routing protocols in case of higher mobility and higher data rate. The research methodology used in proposed method in detail is given in section 3. Results obtained by performance comparison with existing routing protocol are given in section 4.

2. BACKGROUND AND PROBLEM STATEMENT

Incident location is location where disaster happens. All affected and injured peoples are in IL and need to carry to CCS area. For TOC and CCS areas, all nodes are stationary while in IL, all nodes are mobile nodes. These nodes initially have a random location inside incident location, start moving from exit point of IL and enter through entry point of PWT. All patients in PWT area are carrying from exit point of PWT to CCS through entry point of CCS. From CCS, all patients are transferred to hospital zone through APP. More detail is shown in figure 2.

The shortest path algorithm always selects the shortest path, which is in terms of number of hops (Li, 1991). The shortest path may not be called as optimum path as shortest path routing algorithm never considers the traffic present on the network. The shortest path routing may be good when there are less number of packets present on the network in other words; there is less traffic on the network. Whenever the traffic increases, it is always better to select alternate route for the destination which

Figure 1. Disaster area scenario

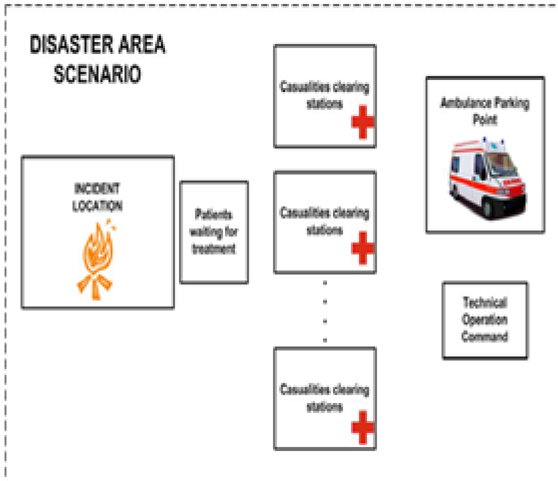
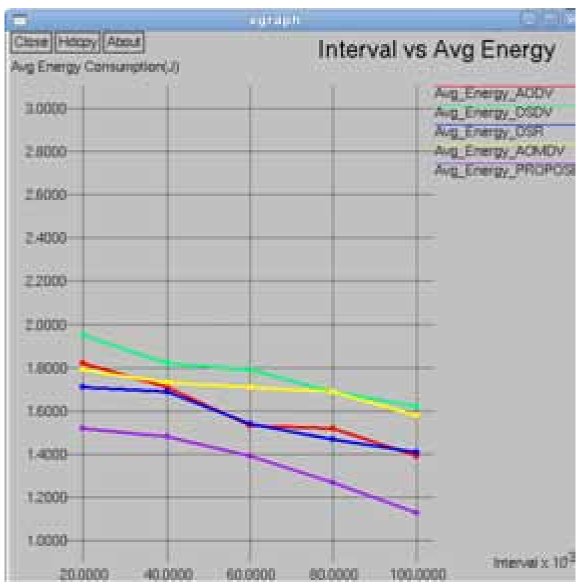


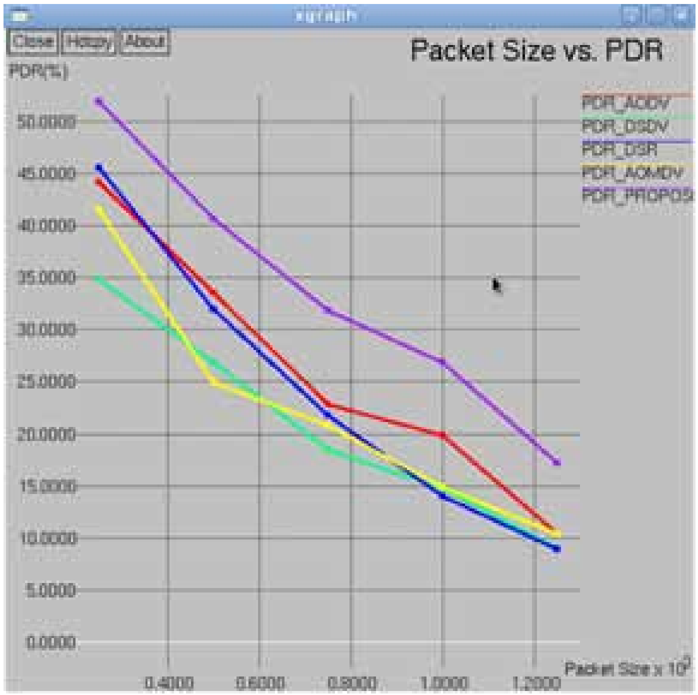
Figure 2. Disaster area scenario having more than one incident location



may not be shortest in terms of number of hops, but this alternate path would result in the shortest delivery time (Jiang et al., 2004).

Machine learning could be incorporated in designing such adaptive routing algorithms. Machine learning (ML) algorithms fall into the categories of supervised, unsupervised learning (Anvik & Murphy, 2011). Reinforcement learning is also a form of machine learning. Reinforcement learning based routing is model-based approach used for adaptive routing. In this, the complete system is modeled in terms of Q values. Q tables contain Q values, which are the estimates of delivery times of the packets to reach to the destinations. It is absolute necessary that these Q values should be

Figure 3. Q routing table at Node 2



learned more accurately to represent real state of the network. In addition, these Q values should also update frequently (rather after every packet send from source or intermediate nodes) for bringing more accuracy in estimating the delivery times of the packets to reach to the destinations. For every destination node, these Q values are zero as packets are already reached to the destination. In addition, if the destination node is the next-hop, Q value is ∂ which represents the transmission delay over the link from node X to node Y.

After every packet transmission by the source node, it also receives learning update, which is used to update its Q values to represent real state of the network. Whenever there are two possible routes, the nodes select the route having minimum Q value and transmit the packets hoping to reach the packet to the destination in optimum time. Optimum path is thus selected provided Q values represent real state of the network. Sample Q routing table at node 2 is shown in figure 3.

At very high load or with high mobility, Q routing does not guarantee the shortest path. It is necessary to optimize this protocol further to stabilize routing table in shortest amount of time to indicate the real picture of the network (Walunjkar & Rao, 2019b).

The Q routing is also classified as CQ, DRQ and CDRQ routing. Confidence values are also generated and stored in confidence table to increase the quality of exploration. Sample Q routing table and C table at node 2 is shown in figure 4.

These confidence values are updated by using same manner as Q values are updated. It is also possible to increase the quantity of exploration which is achieved in DRQ routing where the Q values of sender and intermediate nodes are updated simultaneously. Both confidence values and dual reinforcement is implemented in CDRQ protocol (figure 5).

Figure 4. Q table and C table at Node 2

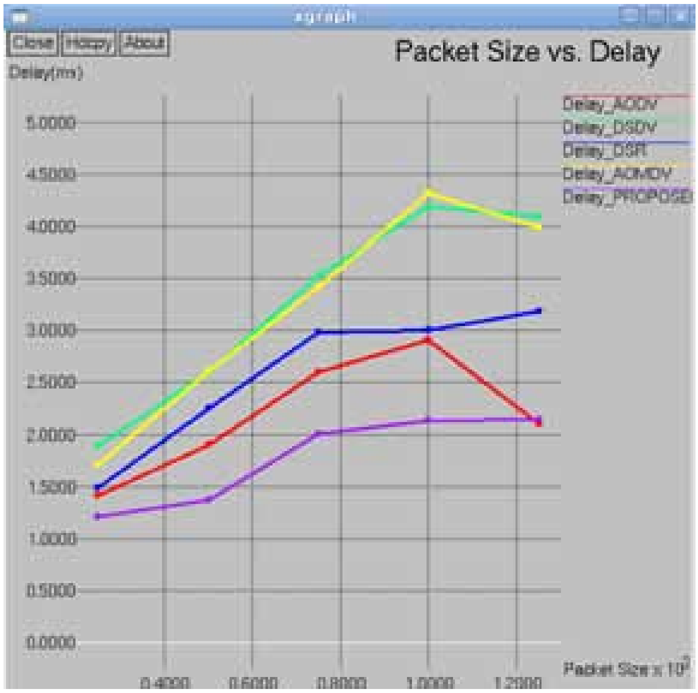
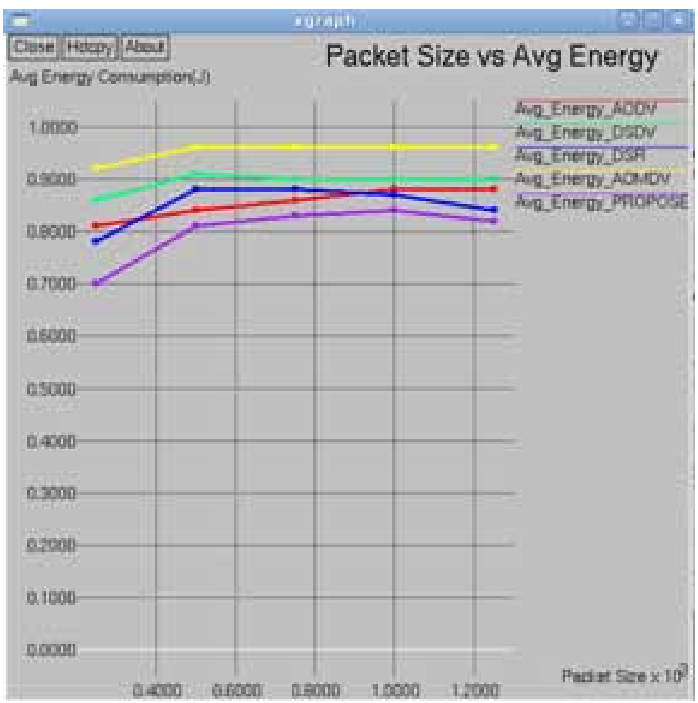


Figure 5. Forward and backward exploration in CDRQ routing



3. RESEARCH METHODOLOGY

3.1 Algorithm for Disaster Area Mobility Model

Algorithm 1

```
Require: VG ← compute visibility graph with obstacles H and areas R
for each area r in R do
while nodes in area r < Nr^trans+ Nr^stat
assign new node n to area r
y  ← 1
if nodes in area r < |Nr^trans| then
n = Nr^trans
else
n = Nr^trans
t:= 0
end if
while t <= d do
if n = Nr^trans then
W_n path based on movement cycle Z_r and VG using Dijkstra
else
n = Nr^trans
v:= rand([v_min; v_max])
p:= rand([t_min; t_max])
add waypoints for W_n to the trace based on r and p
t:= t + time-needed(W_n)
end if
end while
```

3.2 Proposed Method for Improved Routing

When a new node enters inside a network, it initiates initialization process. When new node wants to transmit data, it initiates route discovery process to find out the optimum path. When new node does not want to transmit data, it just monitors HELLO messages.

When node wants to transmit the packets, it finds the best path by consulting Q tables. Estimated Q value for backward exploration embedded in the packet and transmitted towards the best neighbor. Node also receives the estimates from neighbor node and updates Q and C tables. It also decays Confidence values of non-selected nodes and sends control packets to preceding nodes. Q and C tables are shared among neighbors. A neighbor node updates their Q and C tables. When node receives control packet, it extracts reward and puts the destination node information in the queue. Maintenance occurs when there is a change in topology which updates their Q and C values. When the link is restored Q value is made as 0 while if nodes dies, Q value is made as 1.

In proposed method, learning occurs in both directions, quantity of exploration doubles. Backward exploration is also involved which are more accurate values as compared with forward exploration. All Q values are made reliable by using confidence values. Thus Q values of all nodes are updated. Learning rate depends on reliability. Whenever the traffic pattern changes, less adaption time is required as routing policy depends upon Q values and recovery rate. Backward propogation [Figure 7] of Q values from the receiver nodes happens till the packet reaches to the sender node.

This algorithm is briefly described as follows:

Figure 6. Optimization using reinforcement learning

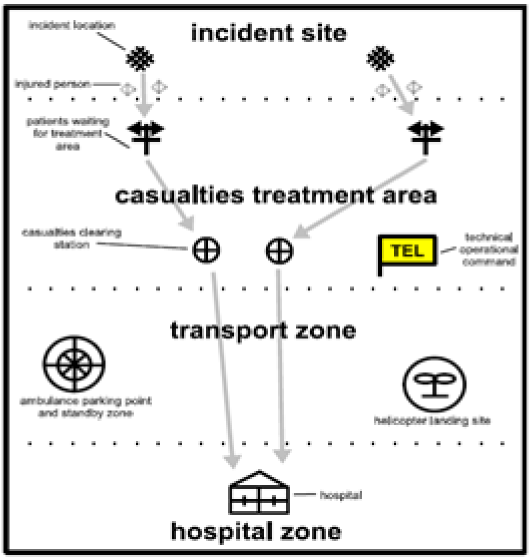
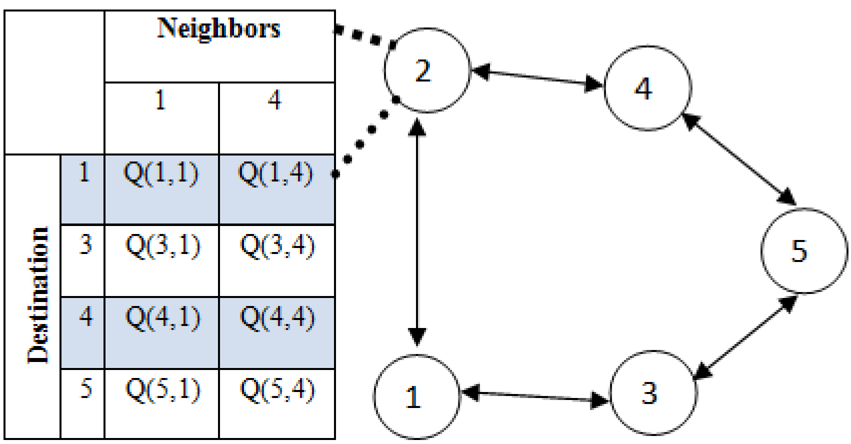


Figure 7. Backward Propagation of Q values from the receiver nodes



Algorithm 2

```
Initialization: When a new node enter inside a network
1. When new node wants to transmit data – Initiates route
discovery process
2. When new node does not want to transmit data – Just monitor
HELLO messages
Learning: When node want to transmit the packets
1. Find the best path by consulting Q tables
2. Estimated Q value for backward exploration embedded in the
packet
3. Transmit packet towards the best neighbor
```

```
4. Receives the estimates from neighbor node
5. Updates Q and C tables
6. Decays Confidence values of non-selected nodes
7. Sends control packets to preceding nodes
Sharing: Q and C tables are shared among neighbors
1. Neighbor nodes update their Q and C tables
2. When node receives control packet, it extracts reward and puts
the destination node information in the queue.
Maintenance: Occurs when there is a change in topology
1. Updates their Q and C values
2. Link restore - Q(*, y) = 0
3. Node dies - Q(*, y) = 1
```

4. RESULTS AND DISCUSSION

The proposed method need to compare with routing protocols against some performance metrics in disaster scenario. The performance evaluation is carried out using standard simulator tool – Network Simulator 2.34. Many network simulator tools are available such OPNET, QualNet, packet tracer, OMNeT++ and NS3. NS2 uses discrete simulator and programming is done using C++ and OTCL. Traffic is created using CBRGEN utility and disaster area scenario is generated using bonnMotion software. Following parameters are used for the network [Table 1].

The PDR obtained by changing the interval is shown in figure 8. The values are shown in table 2. For 0.02 intervals of packets, proposed method provides packet delivery fraction as 21.74% where AODV provides PDR of 16.18%. Higher traffic on the network makes adaptive routing to take fast decisions which avoid packet loss in few percentages. It is observed, that proposed method provides more PDR as compared with other non-adaptive routing protocols.

The delay obtained by changing the interval is shown in figure 9. The values are shown in table 3. As we can see in the above scenario the delay is getting less in proposed method because less delay path is selected for data transmission by calculating estimated delay at real time. For 0.04 seconds of interval, proposed method provides delay of 2.43 seconds which is less as compared with other routing protocols. It is possible to minimize delay by increasing the quality and quantity of exploration.

The average energy utilized / consumed obtained by changing the interval is shown in figure 10. The values are shown in table 4. We are able to decrease the average energy utilization as compared with other routing protocols. For 0.02 seconds of interval, proposed method provides average energy utilization 1.52 Joules as compared with 1.71 Joules using DSR protocol.

Table 1. Simulation parameters for Experiment 1

Parameter	Value	Observations
Interval	0.02 sec to 0.10 Sec	Interval varies from 0.02 sec to 0.10 Sec
Size of Packet	1024 bytes	Packet size is 1024 bytes
Simulation Time	400 Seconds	Simulation carried for 400 seconds
Number of Nodes	250	Total nodes are 250.
Mobility Model	DM	Disaster Area Mobility Model is used.
Topology Size	500 by 500	Size of Topography
Energy	100 Joules	100 Joules per node
No of Connections	50	No of senders and receivers are 50

Figure 8. Interval vs. PDR

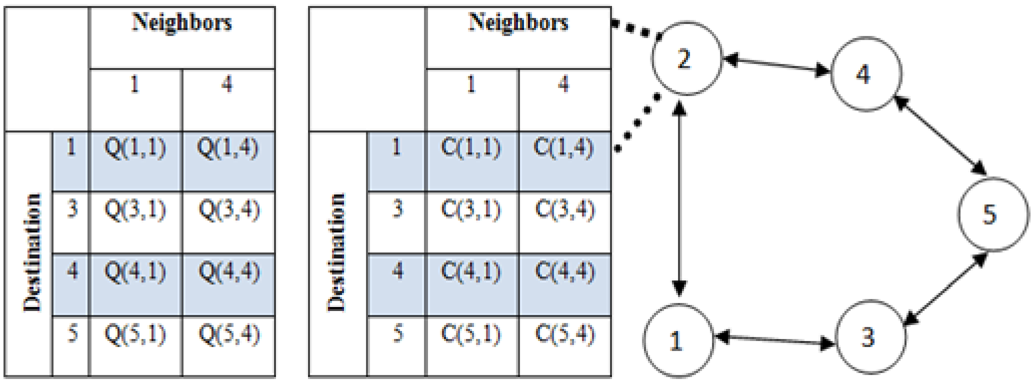


Table 2. Interval vs. PDR for Experiment 1

Interval	0.02	0.04	0.06	0.08	0.10
DSDV	12.85	23.78	30.15	37.65	48.96
DSR	10.80	19.74	21.89	40.48	61.52
AODV	16.18	34.18	48.07	62.19	74.59
AOMDV	14.85	28.75	41.45	54.85	64.80
PROPOSED	21.74	39.14	56.36	67.84	81.45

Figure 9. Interval vs. Delay

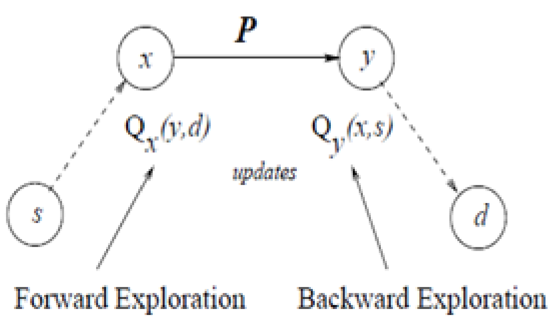


Table 3. Interval vs. Delay for Experiment 1

Interval	0.02	0.04	0.06	0.08	0.10
DSDV	4.52	4.32	4.28	4.25	3.90
DSR	3.12	3.28	2.85	1.75	1.70
AODV	2.81	2.70	2.55	1.97	1.20
AOMDV	4.52	4.40	3.81	3.35	2.76
PROPOSED	2.50	2.43	2.25	1.80	1.05

Figure 10. Interval vs. Average energy utilization

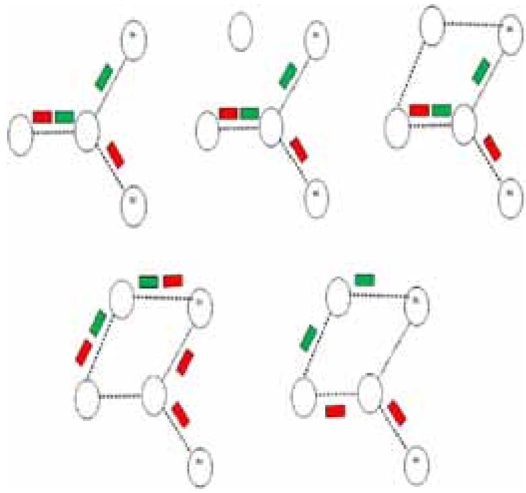


Table 4. Interval vs. Average energy utilization for Experiment 1

Interval	0.02	0.04	0.06	0.08	0.10
DSDV	1.95	1.82	1.79	1.69	1.62
DSR	1.71	1.69	1.54	1.47	1.41
AODV	1.82	1.71	1.53	1.52	1.39
AOMDV	1.79	1.73	1.71	1.69	1.58
PROPOSED	1.52	1.48	1.39	1.27	1.13

Experiment is also carried out by changing the size of packet. The size of packet varies from 250 bytes to 1250 bytes while the interval between successive packets is 0.02 seconds. The parameters are specified in table 5.

The PDR obtained by changing the size of packet is shown in figure 11. The values are shown in table 6. For small packet of size 250 bytes, proposed method provides packet delivery fraction as

Table 5. Simulation parameters for Experiment 2

Parameter	Value	Observations
Interval	0.02 Sec	Interval between packets is 0.02 sec
Size of Packet	250 bytes to 1250 bytes	Packet size varies from 250 bytes to 1250 bytes.
Simulation Time	400 Seconds	Simulation carried for 400 seconds
Number of Nodes	250	Total nodes are 250.
Mobility Model	DM	Disaster Area Mobility Model is used.
Topology Size	500 by 500	Size of Topography
Energy	100 Joules	100 Joules per node
No of Connections	50	No of senders and receivers are 50

Figure 11. Size of packet vs. PDR

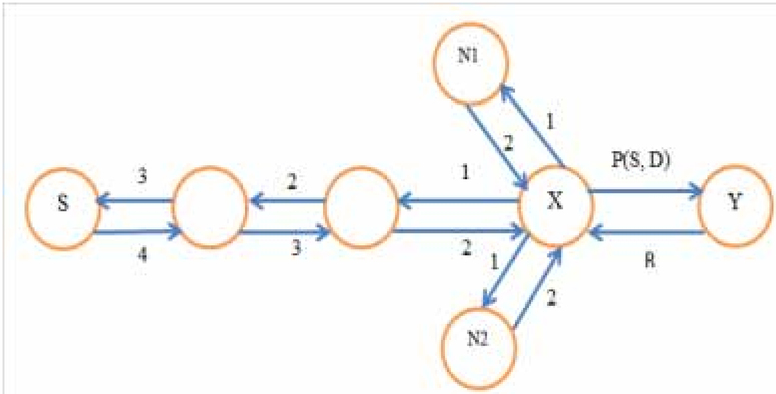


Table 6. Size of Packet vs. PDR

Size of Packet	250	500	750	1000	1250
DSDV	34.85	26.90	18.52	14.80	9.08
DSR	45.52	31.89	21.74	14.06	8.90
AODV	44.15	33.50	22.86	19.85	10.45
AOMDV	41.45	24.96	20.85	14.89	10.40
PROPOSED	51.89	40.65	31.85	26.82	17.25

51.89% while DSR provides PDR of 45.52%. It is observed, that proposed method provides more PDR as compared with other non-adaptive routing protocols.

The delay obtained by changing the interval is shown in figure 12. The values are shown in table 7. As we can see in the above scenario, the delay is less in proposed method because less delay path is selected for data transmission by calculating estimated delay at real time. For packet size of 250 bytes, proposed method provides delay of 1.24 which is less as compared with other routing protocols. It is possible to minimize delay by increasing the quality and quantity of exploration.

The average energy utilized / consumed obtained by changing the interval is shown in figure 13. The values are shown in table 8. We are able to decrease the average energy utilization as compared with other routing protocols. For packet size of 250 bytes, proposed method provides average energy utilization 0.70 Joules as compared with 0.78 Joules using DSR protocol.

We compare the proposed method with existing routing protocols such as DSDV, AODV, DSR and AOMDV protocol. Disaster area mobility model with 250 nodes is considered. PDR, delay along with average energy are compared with varying parameters such as Interval and packet size. It is observed, that the proposed method not only increases the utilization of bandwidth and also reduces the cost required of delivering the packets from source to destinations.

5. CONCLUSION

This paper presents a new approach to improve routing using reinforcement learning. Simulation shows that there is a significant increase in the speed of adaption of routing policy and higher

Figure 12. Size of packet vs. Average delay

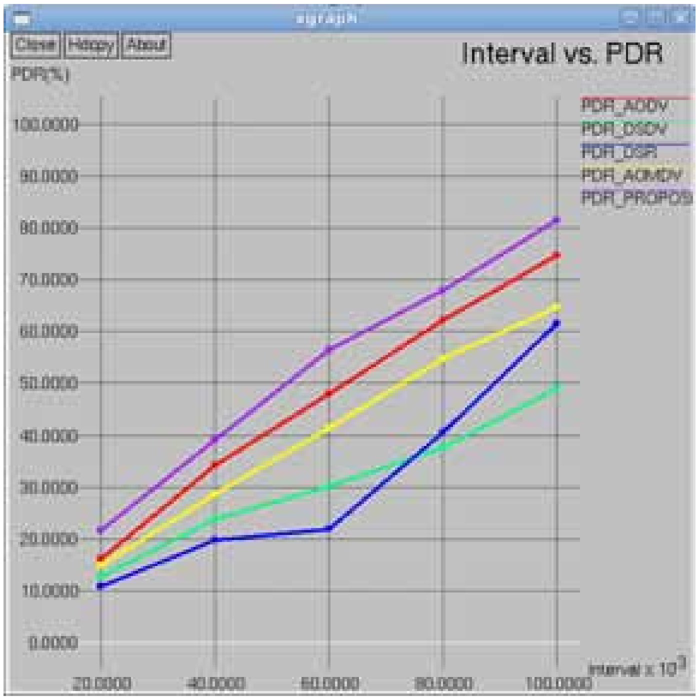


Table 7. Size of packet vs. Average delay

Size of Packet	250	500	750	1000	1250
DSDV	1.89	2.60	3.52	4.19	4.10
DSR	1.49	2.25	2.98	3.00	3.18
AODV	1.41	1.90	2.60	2.91	2.10
AOMDV	1.71	2.61	3.42	4.32	3.99
PROPOSED	1.21	1.37	2.01	2.13	2.14

exploration-exploitation activity, there is a significant increase in PDR, significant decrease in the end-to-end delay and average energy utilization. From our experiments, we summarizes that the proposed method out-performs any distance vector or link state algorithms which are based on shortest path and non-adaptive algorithm in high traffic or/and high mobility or whenever there are frequent changes in the topology. The adaptive routing algorithms are designed to revise the routing policy when the network operates in non-stationary environment and to minimize the actual delivery time to reach the packets to the destination. Average packet delivery time is lowest in proposed method and highest packet delivery ratio could be obtained. Proposed method is better than conventional routing protocols under both high load and varying network conditions.

Figure 13. Size of packet vs. Average energy utilization

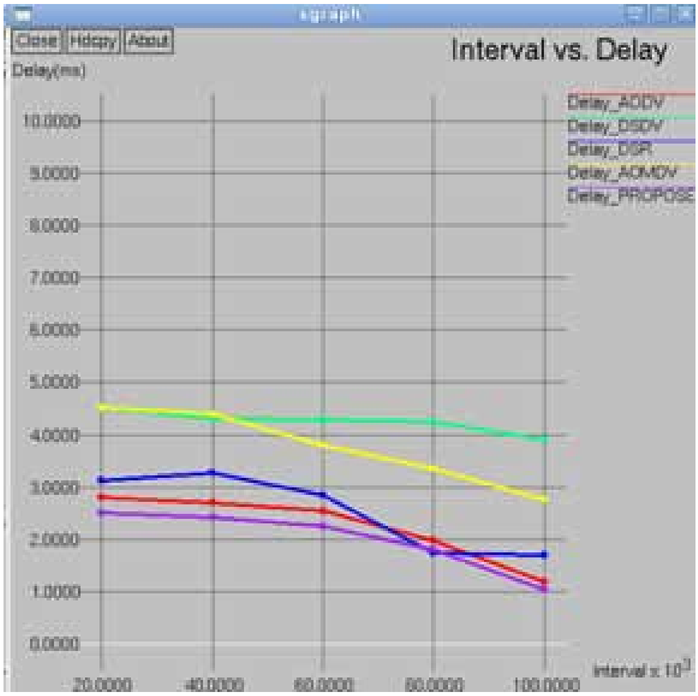


Table 8. Size of packet vs. Average energy utilization

Size of Packet	250	500	750	1000	1250
DSDV	0.86	0.91	0.90	0.90	0.90
DSR	0.78	0.88	0.88	0.87	0.84
AODV	0.81	0.84	0.86	0.88	0.88
AOMDV	0.92	0.96	0.96	0.96	0.96
PROPOSED	0.70	0.81	0.83	0.84	0.82

CONFLICTS OF INTEREST

We wish to confirm that there are no known conflicts of interest associated with this publication and there has been no significant financial support for this work that could have influenced its outcome.

Funding Statement

No funding was received for this work.

Process Dates:

Received: March 25, 2020, Accepted: September 15, 2020

Corresponding Author:

Correspondence should be addressed to Gajanan Walunjkar, gwalunjkar@aitpune.edu.in

REFERENCES

Anvik, J., & Murphy, G. C. (2011). Reducing the effort of bug report triage. *ACM Transactions on Software Engineering and Methodology*, *20*(3), 1–35. doi:10.1145/2000791.2000794

Handbook of Mobile Ad Hoc Networks for Mobility Models. (2011). Academic Press.

Jiang, S., Liu, Y., Jiang, Y., & Yin, Q. (2004). Provisioning of Adaptability to Variable Topologies for Routing Schemes in MANETs. *IEEE Journal on Selected Areas in Communications*, *22*(7), 1347–1356. doi:10.1109/JSAC.2004.829352

Li. (1991). A survey of research and standards in high-speed networks. *International Journal of Digital & Analog Communication Systems, 10*.

Nossenson, R., & Schwartz, A. (2013). *On the Initial Network Topology Factor in Mobile Ad-Hoc Network*. Springer Science an-d Business Media LLC. doi:10.1007/978-94-007-6190-2_29

Recent Trends in Networks and Communications. (2010). Springer Science and Business Media LLC.

Walunjkar, G., & Rao, A. K. (2019a). Simulation and Evaluation of Different Mobility Models in Disaster Scenarios. *2019 4th International Conference on Recent Trends on Electronics, Information, Communication & Technology (RTEICT)*. doi:10.1109/RTEICT46194.2019.9016893

Walunjkar, G., & Rao, A. K. (2019b). A Reinforcement Learning Approach for Evaluation of Disaster Relief. *2019 International Conference on Cutting-edge Technologies in Engineering (ICon-CuTE)*. doi:10.1109/ICon-CuTE47290.2019.8991541

Gajanan Walunjkar received his Masters in Computer Engineering degree from Pune University in 2009 and currently pursuing Ph.D. in Computer Science from Vel Tech Rangarajan Dr. Sagunthala R & D Institute of Science and Technology, Chennai, India. He is presently working as Asst Professor, Dept. of Information Technology in Army Institute of Technology, Pune, India. He is having 15 years of experience and has published 20 plus research papers in various international and national referred journals and conferences. His area of interest includes wireless communication and ad hoc networks.

Anne Koteswara Rao is Professor and Director Academics in computer science and Engineering department at Vel Tech Rangarajan Dr.Sagunthala R&D Institute of Science and Technology. Computer Networking, Engineering Education are his area of interest. He has also published quality research papers, articles, books in indexed journals. His experience also includes Head of the department and Dean Academics in different institutes such as University of Klagenfurt, Austria, University of Hanover, Germany, and VR Siddhartha Engineering College, India. He has also played important role in National Board of Accreditation (NBA), representing at International Engineering Alliance Meetings, member of Washington Accord signatory by developing accreditation criteria.

V. Srinivasa Rao is the Dean of School of Computing at Vel Tech Rangarajan Dr.Sagunthala R & D Institute of Science and Technology, Chennai and he is also a Professor and Head of the Computer Science and Engineering department. His research interests include image processing, ad hoc networks, machine learning. He published high quality research articles/books in indexed journals and completed/doing research projects in these areas. Before becoming Dean SOC at Vel Tech Rangarajan Dr.Sagunthala R & D Institute of Science and Technology Chennai, Dr V Srinivasa Rao has varied experience in academic roles such as Professor, Head of department in institute such as V. R. Siddhartha Engineering College, Vijayawada, Andhra Pradesh, India.