ORIGINAL ARTICLE

A novel algorithm for wireless sensor network routing protocols based on reinforcement learning

Anil Kumar Yadav[1] · Purushottam Sharma[2] · Rakesh Kumar Yadav[3]

Abstract Major challenging problems for wireless sensor networks are the utilization of energy and lifetime routing maximization in the network layer. In wireless sensor network protocols are more critical over data routing in the network. Energy awareness has been described in the context of data-centric, spatial based and categorized protocols. This research paper presents energy consumption analytical analysis based on adoptable routing algorithms based on reinforcement learning using Q-Learning algorithms. Performance comparisons with distributed routing algorithms in the context of the rate of delivery, energy consumption, flow rate, number of packets lost and lifetime of the system were evaluated.

Keywords Q-learning · Reinforcement learning · Wireless sensor networks · Routing Protocol

1 Introduction

End-to-end latency, delivery rate, network lifespan, and energy usage are measures of successful network management and providing quality of service in Wireless Sensor Networks (WSNs) (Giordano 2002). WSN nodes must deal with routing effectively and adaptively to meet the stringent requirements of these parameters.

Indeed, in mobile environments, the routing protocol must perform well; it must be able to adapt automatically to high mobility, complex network topology, and connection changes. Simple guidelines are insufficient to prolong the network's lifespan.

As a result, packet routing decisions can be regulated using Reinforcement learning (RL) (Sutton and Barto 2014) In general, there are two types of reinforcement learning-based routing protocols for Wireless Sensor Networks (WSNs): global and local. This paper uses reinforcing learning to compare the performance of different Global Routing protocols. The small battery-powered device is made up collection of sensor networks and wireless infrastructure that track and record environmental conditions.

The wireless network is a self-governing device that works over a distributed network, which used protocol to distribute data over the network. It also used for real monitoring or working over for weather forecasting (Fig. 1) (Devika et al. 2013).

Sensor Networks use sensors to track temperature, pressure, humanity, applied over the direction of wind and light of speed, able to vibration, enhance intensity, voltage, detection and monitoring of chemical compound, to check pollution level and monitoring all function along with sensor network (Perkins et al. 2003).

1.1 Types of wireless sensor networks

Environment formed by a sensor network, it applied to overcrop monitoring in water management and coordinate

✉ Purushottam Sharma
  puru.mit2002@gmail.com

1 Department of Computer Science and Engineering, IES College of Technology, Bhopal, Madhya Pradesh, India

2 ASET, Amity University Uttar Pradesh, Noida, India

3 Department of Computer Science and Engineering, IFTM University, Moradabad, Uttar Pradesh, India
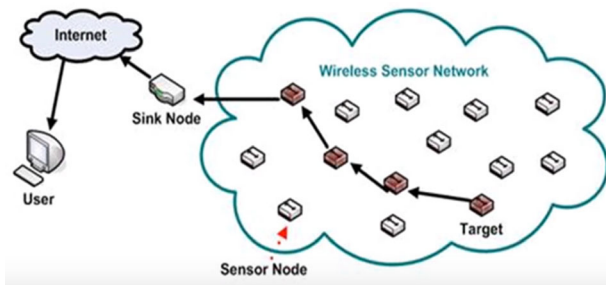
Fig. 1 Wireless sensor network

detection, etc. global sensor network and secretive wireless sensor network handsets network connected though varies set lights or other devices.

It combined thousands of wireless sensor network acts as a node to restructure in ad-hoc (unstructured) to communicate with the base station. The distribution of the sensor nodes is unsystematically over the given goal position in the field of agriculture. What are uses unstructured phenomenon for dripping along with static mode over it? This network is planned as a 2D and 3D model in the context of the grid and optimal placement accountability.

For minimal power utilization along with a wireless sensor network model with a solar cell that provides more efficient power to the network. To evaluation of the minimum delay and cycle along with shortest routing nodes between various nodes as the given network. That help saves energy towards wireless sensor network. In the case of global sensor networks more costly in the context of providing services, the cost of equipment and the assembly of underground implementation.

There are several sensor nodes recombined to be made up of high qualitative devices to produces efficient conditions.

Additional sensing nodes are situated above the normal level that effective for relay data from source to destination to the base station. There are many problems to recharge sensing devices to work effectively and also recharge the solar battery to given backup when required charge over the network.

Due to a larger degree of attention, low signal and loss of signal have created more difficulty for communication for the underground. In survey report b national organization total consumption of water occupied more than 70 % of the earth. These are constructed by several sensor nodes to apply over the targeted problem.

Sensor nodes are used for gathering information from all nodes and after evaluation, they will control the vehicle and other devices. Due to propagation delay sensor connectivity and transfer of data failure from point-to-point connectivity that is a major challenge of sensor networks.

Wireless sensor network containing high-capacity storage battery backup to sensing device to establish communication between different devices (Sutton and Barto 1998; Kim et al. 2002; NS 2004; Rabiner et al. 1999; Heinzelman 2000; OssamaYounis and Fahmy 2002, 2004).

Wireless networks along with media networks developed so that events such as images, video and au are tracked and monitored by multimedia.

These networks are made up of low-cost, cameraman-equipped sensor nodes. These nodes are interconnected via a wireless data compression, data collection and correlation connection.

In wireless networks for improving bandwidth and data processing used data compression methods form part of the challenges with WSN multimedia. Furthermore, multimedia content requires the correct and easy delivery of content with high bandwidth (Watkins and Dayan 1992; Chettibi and Chikhi 2011; Vassileva and Barcelo-Arroyo 2008).

Several sensing devices communicate with the real environment of the distributed environment. Sensitivity and communication are possible for mobile nodes (Manjeshwar and Agrawal 2001; Lindsey and Raghavendra 2002; Taheri et al. 2012; Devika et al. 2013; Lin and Chen 2014).

Static sensor networks are less effective than mobile sensor networks in a distributed environment. Including improving energy utilization and effectiveness, improved channel capacity, etc., are the benefits for static sensing network huge networks over a distributed environment.

2 Related work

There are better perform conscious routes along with maximum lifetime protocols for energy in terms of local routing and global routing. In both routing evaluate modeling behavior of the routing path of the nodes. That is based on intelligent routing, combining prior search routing algorithms to search shortest and effective nodes from sources to destination.

2.1 Universal (global) routing protocol

Every sensing mobile node takes part along with the process of road innovation in global routing through transferring packets for the Route Request (RREQ). Paths that are then discovered are assessed by source or destination knots toward metric energy for all distributed nodes.

In Cho and Kim (2002), each node sends the concept to the request for the time delay route. RREQ data packet holds various nodes it depends on the relationship as inverse relation with battery energy and residual this time.

Therefore, pathways with energy-poor nodes have little chance of being selected.

In Naruephiphat and Usaha (2008), the author seeks to maximize the lifetime of nodes while reducing the amount of energy. Each spring sensor node acting over reinforcement learning algorithms as Q-learning on the first visit, so that three main parameters are used to choose the best path: minimum energy path, This path counted as residual battery backup and minimum cost of the given node in routing network.

In Nurmi (2007), described combined reinforcement learning algorithm along with the routing protocol. This modeled based on a sequential decision-making system based on trial-and-error interaction using Marko decision process that is fully observable.

Ravi and Kashwan (2015) The report of the Council of Europe a new reducing excess energy approach, like fidelity energy conservation algorithm, (AFECA). This algorithm is known as the "Energy-Aware Spain Routing tool." The use of a physical circuit to reawaken napping nodes further optimizes vitality consumption. These methodologies to energy saving to the already re-active routing protocols in a sensor network.

2.2 Local (native) routing protocol

In Local Routing, each intermediate node takes its own decision by its energy profile, either to take part in the discovery of routes or not to delay the transmission of the EQ, or, to adjust its forwarding rate eventually. The model of routing in (Xu et al. 2001; Srinivasan et al. 2003) provides nodes with two possible behavior patterns: cooperation (forward packages) or defects (drop packets).

Every j node in Altman et al. (2004) transmits μj-probably packets. Each node calculates the current balancing strategy and uses this as the delivery test capabilities if a packet is sent.

Furthermore, there is a penalty mechanism where nodes decrease the probabilities when someone deviates from the strategy of balance.

Although these works are proven to be effective, authors at Chettibi and Chikhi (2012, 2014) offer additional efficient methods of routing, using reinforcement learnings, to enable each node to learn appropriate shipping rates reflecting their willingness to participate in the process of discovery.

In Chettibi et al. (2016) was described a dynamic AODV power using routing protocol in ad-hoc networks where every node takes a Fuzzy Logic System to identify and takes the decision to select Requests during discovery route (RREQs).

In Chettibi and Chikhi (2013), a Fuzzy Logic System (FLS), in contrast to Chettibi et al. (2016), was used for the OLSR protocol to adjust its readiness parameter. The FLS's residual energy and its expected residual life take into account decisions at every mobile node.

In Das and Tripathi (2018) was described intelligent routing protocol based on serval intelligentsia technique of Decision Making, namely intuitional fuzzy soft set (IFSS).

3 Proposed method and algorithm

This section describes environment generalization along with state action (s, a), which are responsible for sensory input during training sessions. For this purpose, we need to model various nodes as a state over a given environment. It performed random state action and start searching available action within given four available actions left, right, up and down to archive a new state. If the current state is similar to the last stat action then repeat from start step one. Then checking goal status id achieved then exit otherwise updates lookup table and proceeds next steps.

We selected Q-learns (which have a model-free enhancement study characteristic, which means that we don't have an initial system knowledge to train the network's various nodes in the given environment under optimal policy rewards and penalty. Through this agent learn to acts optimally toward Marko process via the number of trial and error interaction. Every iterative process store in given lookup table, during testing it will select ransom 9.

Value from available action.

Applying Q-learning algorithm all nodes in given environment train using objective function approximation (Q function) like of a satellite state-action input pair. The physical activity during the train of the network process involves sensory input and targeted output in the given problem. At the beginning of the training process, based on state action pair values Q-learning algorithm as a reward value.

Classifications of events, repeated knowledge processes within each occurrence in Q learning occur in nth time:

We propose new RL routing algorithms in this work as follows.

(i)   Randomly generates starter state (sn) in a given network

(ii)   Search for actions available (an)

(iii)   Randomly selects an action

(iv)   if the goal is not achieved

(v)   go to step (i)

(vi)   When the goal is reached

(vii)   Iterative value store (sn, an) in temp array

(viii)   Table Update of with $R_1 = R\gamma^{(x1-i)}$

(ix)   Now uses the State-action pair to generate a next state (sn + 1)

(x)   If then the next goal is reached.

(xi)   Repeat step xi again until the stop criterion is met

3.1 Implementation of reinforcement learning routing algorithms

I/P=Start search, define state(s, a) and move (L R U D) possible at different nodes

O/P= Search objective (malicious node in the given network) Realized / completed search

\# Begin for (every state si: s) algorithm start / / Generates transitional action \

Start

Step 1: fetch the initial node if,

Step2: Find and execute possible network or search movement (L R U D),

Step 3: Target Status Check (find as per given goal to find malicious node)

Step 4: if (yes: goal=start as targeted node)

Update the d1=dμN1-i and i=i+1 discount rate.

Step5: else

{

Goal ≠ the targeted node states (malicious node)

Store (s,a) in the search table value

}

Step6: Take step2 //Last but not least.

End

End.   // Search finished \

4 Results analysis and comparisons

Performance of the reinforcement learning routing algorithm using Q-learning measures can be used to assess learning accuracy. E_Q is a measure of goal tracking efficiency in the given node of the networks and overall percentage of correctly evaluated training rate.

$$E_Q = \left[1 - \frac{\text{Minimum count step}(i) - \text{Total count step}(x1)}{\text{Total program output}(T)}\right] \times 100$$

where T is the total number of states, the minimum count step is I and x_1 is the total count step.

As can be seen from the above Tables 1, 2, the accuracy of reinforcement learning routing algorithm with different episodes, so compare these data and show the accuracy in the form of percentage in Figs. 2, 3.

As can be seen from Table 1, the learning accuracy of the Reinforcement learning routing algorithm is over 95 % most of which are higher than Distributed Routing Algorithm. After completion of all the experimental operations, we got the result that the Reinforcement learning routing algorithm is better than Distributed Routing Algorithm. Here we have done the comparison of the Reinforcement learning routing algorithm and Distributed Routing Algorithm and we found that the Reinforcement learning routing algorithm is better in the context of efficient packets delivery, learning time, memory usage. Finally, we compare and show the comparison graph is as follows.

Table 1 Comparison table between different episodes for Reinforcement Learning Routing Algorithm

Rate of delivery	For Episode E = 1000	For Episode E = 3000	For Episode E = 5000	For Episode E = 7000	For Episode E = 9000	For Episode E = 11,000
d1	69	97.80	81	99.89	81	99.70
d2	70.50	69	66	84	96.71	97.82
d3	74.20	69.80	98.77	97.77	84.90	85.60
d4	87	76.60	97.66	84.80	97.90	99.70
d5	98.80	97.92	61.11	89.95	99.80	99.50
Average	79.98 %	81.40 %	79.69 %	91.69 %	92.70 %	96.20 %

Table 2 Comparison table between different episodes for Distributed Routing Algorithm

| Rate of delivery	For Episode E = 1000	For Episode E = 3000	For Episode E = 5000	For Episode E = 7000	For Episode E = 9000	For Episode E = 11,000
d1	96.80	72	75.60	97.80	98.20	99.78
d2	72	70.12	63	78.80	81.50	83.90
d3	68.90	73	92.54	75.60	84.54	96.77
d4	76	81.90	96	81	83.90	81.90
d5	98.72	97.80	54	96	95.40	99.90
Average	81.70 %	79.94 %	75.78 %	85.50 %	88.70 %	93.47 %

Fig. 2 Comparison View of different Episodes
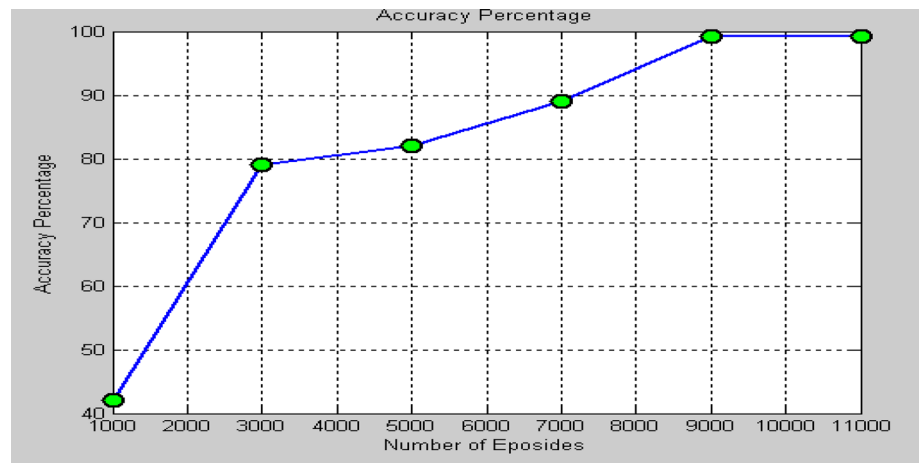
5 Conclusions

We evaluate the reinforcement learning routing algorithm and various intelligent routing algorithms such as distributed routing algorithms, in particular, and compare our performance. The knowledge that the agent gains from various sensing environments and the respective decision-making path have also been demonstrated. We noted that the reinforcement learning routing algorithm is efficient to calculate various sensing nodes with prior system knowledge in connection with the delivery rate of packets.

In the future, we can improve learning through the combination of multiple learners and other decision-making techniques like transfer learning.

Fig. 3 Performance Comparison graph between Reinforcement Learning Routing Algorithm and Distributed Routing Algorithm
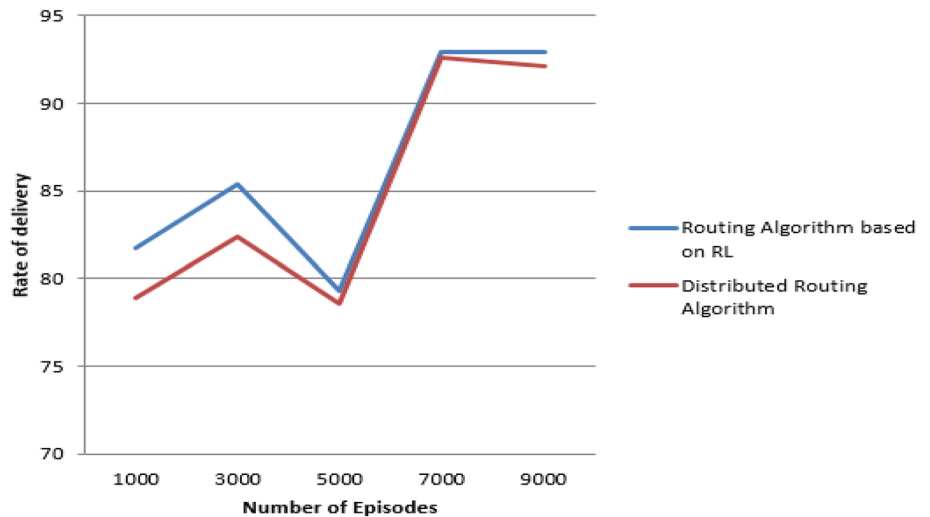
References

Altman E, Kherani AA, Michiardi P, Molva R (2004) Non-cooperative forwarding in ad- hoc networks. In: Proceedings of the 15th IEEE international symposium on personal, indoor and mobile radio communications

Arjunan S, Pothula S (2016) A survey on unequal clustering protocols in wireless sensor networks. J King Saud Univ Comput Inf Sci

Chand S, Kumar R, Kumar B, Singh S, Malik A (2016) NEECP: novel energy-efficient clustering protocol for prolonging lifetime of WSNs IET Wirel. Sens Syst 6(5):151–157

Chettibi S, Chikhi S (2011) A survey of reinforcement learning based routing proto- cols for mobile Ad-Hoc networks, recent trends in wireless and mobile networks, 162. Communications in Computer and information Science, Springer, pp. 1–13

Chettibi S, Chikhi S (2012) An adaptive energy-aware routing protocol for MANETs Us- ing the SARSA reinforcement lea ing algorithm. Evolving and Adaptive Intelligent Systems (EAIS), IEEE Conference on, 84–89

Chettibi S, Chikhi S (2013) FEA-OLSR: an adaptive energy aware routing protocol for manets using zero-order sugeno fuzzy system. Int J Comput Sci Issues 10(2):136–141

Chettibi S, Chikhi S (2014) Adaptive maximum-lifetime routing in mobile ad-hoc networks using temporal difference reinforcement learning. Evol Syst 5, Springer: Berlin, Heidelberg

Chettibi S, Chikhi S (2016) Dynamic fuzzy (local routing) logic and reinforcement learning for adaptive energy efficient routing in mobile ad-hoc networks. Applied Soft Compu- ting 38:321–328

Cho W, Kim SL (2002) A fully distributed routing algorithm for maximizing life time of a wireless ad hoc network. In: Proc IEEE 4 th Int workshop-mobile & wireless commun. Network, pp. 670–674

Das SK, Tripathi S (2018) Intelligent energy-aware efficient routing for MANET. Wirel Netw 24(4):1139–1159

Devika R, Santhi B, Sivasubramanian T (2013) Survey on routing protocol in wireless sensor network I jet, 5(1):350–356

Heinzelman WR (2000) Application-Specific Protocol Architectures for Wireless Networks, Ph.D. thesis, Massachusetts Institute of Technology

Giordano S (2002) Mobile ad hoc networks. Handbook of wireless networks and mobile computing. pp. 325–346

Kim D, Aceves JJGL, Obraczka, Cano JC, Manzoni P (2002) Power-aware routing based on the energy drain rate for mobile ad-hoc networks. In: 11th international confer-ence on computer communications and networks

Lin H, Chen P, Wang L (2014) Fan-shaped clustering for large-scale sensor networks Proc– 2014 Int Conf Cyber-Enabled Distrib Comput Knowl Discov CyberC. pp. 361–364

Lindsey S, Raghavendra CS (2002) PEGASIS: power-efficient gathering in sensor Information system. In: Proceedings IEEE aerospace conference, 3, Big Sky, MT, pp. 1125–1130

Manjeshwar A, Agrawal DP (2001) TEEN: a protocol for enhanced efficiency in wireless sensor networks. In: The proceedings of the 1st international workshop on parallel and distributed computing issues in wireless networks and mobile computing, San Francisco, CA

Naruephiphat W, Usaha W (2008) Balancing tradeoffs for energy-efficient routing in MA- NETs based on reinforcement learning. In: The IEEE 67th vehicular technology con-ference

Nurmi P (2007) Reinforcement learning for routing in ad-hoc networks. In: Proceedings of the fifth international symposium on modeling and optimization in mobile, Ad-Hoc, and Wireless Networks (WiOpt)

NS (2004) The UCB/LBNL/VINT Network Simulator (NS), http://www.isi.edu/nsnam/ns/

Perkins C, Belding-Royer E, Das S (2003) Ad hoc on-demand distance vector (AODV) Routing. Network Working Group

Rabiner W, Kulik J, Balakrishnan H (1999) Adaptive protocols for information dissemination in wireless sensor networks. In: Proceedings of the fifth annual international conference on mobile computing and networking (MOBICOM), Seattle, WA, USA, pp. 74–185

Ravi G, Kashwan KR (2015) A new routing protocol for energy efficient mobile applica- tions for ad hoc networks. Comput Electr Eng 48:77–85

Srinivasan V, Nuggehalli P, Chiasserini CF, Rao RR (2003) Cooperation in wireless ad hoc networks. In: Proceedings of the 22nd annual joint conference of the IEEE computer and communications societies (INFOCOM). IEEE Computer Society, pp. 808– 817

Sutton R, Barto A (1998) Reinforcement learning. MIT Press: Cambridge, MA

Sutton RS, Barto AG (2014) Reinforcement Learning, Second edition, in progress, MIT Press

Taheri H, Neamatollahi P, Younis OM, Naghibzadeh S, Yaghmaee MH (2012) An energy-aware distributed clustering protocol in wireless sensor networks using fuzzy logic Ad Hoc Netw., 10(7):1469–1481

Vassileva N, Barcelo-Arroyo F (2008) A survey of routing protocols for maximizing the lifetime of Ad Hoc wireless networks. Int J Softw Eng Appl 2(3):77–79

Watkins CJ, Dayan P (1992) Qlearnin. Mach Learn 8: 279–292

Xu Y, Heidemann J, Estrin DD (2001) Geography informed energy conservation for Ad- Hoc routing. In: Proceedings of 7th annual international conference on mobile computing and networking, pp. 70–84

Younis O, Fahmy S (2002) Distributed custering in Ad-hoc sensor networks: a hybrid, energy-efficient approach

Younis O, Fahmy S (2004) Heed: a hybrid, energy-efficient, distributed clustering approach for Ad-hoc networks. IEEE Trans Mob Comput 3(4):366–369