Reinforcement Learning based Routing Protocols in WSNs: A Survey

Anju Arya[1]
Computer Science Department
Deenbandhu Chhotu Ram University of Science and Technology, Murthal
Haryana, India

Amita Malik[2]
Computer Science Department
Deenbandhu Chhotu Ram University of Science and Technology, Murthal
Haryana, India

Ritu Garg[3]
Computer Science Department
Deenbandhu Chhotu Ram University of Science and Technology, Murthal
Haryana, India

Abstract— *Advances in the technology along with reduction in processor size, its memory, and wireless antenna size has facilitated the construction of low cost, low powered and multifunctional Sensor nodes which in turn led to high demand for development of Wireless Sensor Networks. A lot of research work has been done regarding the development of routing protocols for WSNs. This paper provides a brief overview of the routing protocols using Reinforcement learning approach for WSNs.*

Keywords— Reinforcement Learning (RL), Wireless Sensor Network (WSN), Routing Protocols.

I. INTRODUCTION

The Micro-Electro Mechanical Systems technology has attracted the world-wide attention towards the development of WSNs and the various application areas it can cover. A wireless sensor network [1] consists of a collection of autonomous sensors densely deployed in a structured or unstructured manner to monitor a physical environment and gather the relevant information and cooperatively and coordinately transfer it through the network to the sink through gateway nodes. The dense deployment helps to obtain accurate data and at the samw time achieve speed, flexibility, and reliability in the given environment.

The research on sensor networks started at DARPA with DSN program by about 1980 and its beginning is marked by the 1998 SmartDust project. These networks have found a major use and application in the field of Optimal Control Systems. They are also now used in other applications related to monitoring and tracking activities, such as area monitoring, greenhouse monitoring, structural monitoring, passive localization and tracking activities, etc.

The main features of wireless sensor networks are Embedded routers, Dense connectivity, Resource constrained nodes, Asymmetric links, Dynamic topology, Broadcast communication paradigm, Heterogeneity of nodes, Withstand harsh environments, Autonomous in nature, Infrastructure less and self operable, and Multi-hop routing. The main challenges faced by wireless sensor networks are frequent topology changes, limited battery, limited capacity, limited memory, prone to failures, no global ID, and adaptation.

Reinforcement learning approach [2] addresses mainly 2 problems namely, Prediction and Control problems. It is now being widely used by the routing protocols, at the network layer, to handle resource constraints. The most widely used reinforcement learning algorithm is Q-Learning approach. The major advantage of RL based routing is that each node does not need global information of the network, but still it can approximate global optimality. The focus of RL techniques is mainly to find an optimal path and increase residual energy of the network, which helps to prolong network's lifetime and its efficiency. The routing approach used in the network depends upon the aspect being given importance or the QoS required for a particular application area. The remaining sections provides a brief overview about the routing protocols for WSNs based on Reinforcement Learning approach.

II. REINFORCEMENT LEARNING AND WIRELESS SENSOR NETWORKS

Routing can be defined as efficiently transmitting data over the network, simultaneously considering the other factors such as energy consumption, cost, quality of service, network lifetime, etc. The four important factors to be considered are energy cost, robustness, throughput and delay. Most of the RL based protocols finds optimal paths and prolonging network lifetime by being energy-aware (balancer or saver approach) or by increasing residual energy uniformly across the network.

The routing strategies can grouped into being structured and Structure-less. Some greedy based approaches are GPSR, CADR, etc and some search based approaches are GEAR, Q-Routing [3], etc. Adaptive Tree Protocol (ATP) is between structured and structure-less. The routing strategies can also be grouped on the basis of aspect being focused such as location-based, feedback-based, Energy-awareness, fault-tolerance, or cost effectiveness.

Thus, on the basis of adaptiveness, routing mechanisms, coordinating agent, and aspect being focused, the routing protocols or approaches can be categorized as follows:

A. On the basis of Adaptivity

Adaptivity is the most desired feature for a dynamic environment. The dynamic behavior can be due to dynamic topologies, application requirements, power requirements, and traffic patterns. It can be obtained through the use of:

1) *Adaptive Routing Schemes:* It includes Adaptive Routing [4, 5] which characterizes the route paths by their sinks or destinations and change in the route paths due to dynamic network conditions. Adaptive routing (AdaR) [6], Adaptive Tree Protocol (ATP) [7], etc. comes under such schemes.
2) *Cross-Layer protocols: These protocols help as to achieve* competing goals of energy-efficiency and flexibility through specialization through cross-layer protocol design and through modularity in a layered protocol design, respectively. The main logic behind cross-layer design (CLD) is to use the information from multiple layers to jointly optimize performance of those layers. XLM (unified), MAC-CROSS [8], CLEEP [9], RL-MAC [10] [11], DReL, DIRL [12], etc. are some of the cross-layer protocols using reinforcement learning concept.
3) *Learning automata schemes: This scheme* can be used to model the learning systems [13] and also it doesn't require the information of the environment it operates in. It helps the intelligent automata equipped sensor nodes to learn or adapt to the environment as it changes and become intelligent with time. Some of the LA based protocols are AEESPAN [14], SARA [15] [16], FEAR [17], etc.

B. On the basis of routing mechanisms

On the basis of routing mechanisms, protocols can be classified as structure-less and structure-based. The Structure based mechanisms use some data structure such as a Routing table to store information, updated periodically or on-demand, that can be used to take routing decisions, later on. These mechanisms are suitable for stable networks. Real-time search approaches such as ant routing, TD methods etc, flooding based mechanisms and greedy approaches are some structure-less mechanisms for dynamic environment. Adaptive Tree Protocol (ATP) [7] comes under both structured and structure-less mechanisms.

C. On the basis of coordinating agent

To get a more accurate model for a large number of sensor nodes, a multi-agent system approach can to be adopted. The two kinds of coordination based Reinforcement Learning-based approaches are:

1) Single-agent reinforcement learning (SARL): It is suitable for centralized networks. Example: COORD.
2) Multi-agent reinforcement learning (MARL): It is suitable for distributed networks.

The MARL approaches can further be categorized in terms of the no. of hops involved in the payoff message propagation as Single-hop coordination-based MARL and Multiple-hop coordination-based MARL.

D. On the basis of Aspect being focused

Energy-aware

The energy-aware routing is used to minimize energy waste and maximize network lifetime by utilizing resource uniformly. The packets are routed such that energy consumption is distributed uniformly around a forwarding node. It is used for handling optimization problems. The factors affecting energy consumption are discussed in [18], such as routing path length, link reliability, aggregation, load balance, etc. The energy-aware protocols can further be categorized as Energy-saver and Energy-balancer. RLGR [19], DRLR [1], etc are some examples of energy-aware protocols.

Feedback

A Feedback is a scalar value that helps to take routing decisions, learning and adapting to the network. This scalar value can be in the form of estimated route costs or link status. The attachment of routing feedback to data or route packets along with learning is a powerful tool in the wireless network domain because it needs limited local knowledge and achieves significant results. It also does not increase network costs. FROMS [3] [20] is one such protocol, which uses feedback for learning network.

Quality of Service

QoS can be interpreted as a measure of service quality offered by the network to the end user from networking perspective. The various factors or elements defining QoS varies according to the application requirements. The two perspectives from which QoS can be specified are application specific QoS and Network QoS. The different metrics, which can be used to measure QoS are latency, packet delivery time, reliability, accuracy, aggregation delay, coverage, fault tolerance and network lifetime, energy management, etc. QoS can be achieved through Data Redundancy, even if it increases routing and energy overhead. QoS requirements can further be categorized as Timeliness and Reliability. RL-MAC [10] [11] and QoS routing [21] are some protocols that provides QoS.

Location aware

Location-aware protocols use the location or geographic information to take routing decisions to forward the data packets. The sensor nodes use both location and energy information of its neighbors to select the next hop for routing packets. This data can also be used to balance load distribution and distribute energy consumption uniformly over the network. It also helps to avoid EISS and problems of network partitioning. The focus is on reducing delay and maximize network lifetime. RLGR [19] is a one such location-aware protocol using RL approach.

Fault-tolerant

Fault tolerance is the ability of a system to provide a good or desired level of performance, even in the presence of faults or bugs. This property is highly desired in application such as clustering, time synchronization,, etc. The network failure can be due to energy depletion, hardware or system failure, communication or unpredictable link errors, routing failure, malicious attacks, packet loss, etc. ATP [7] and FROMS [3] [20] are some of the protocols with fault-tolerant capability.

The following table classifies some reinforcement learning based protocols on the basis of the concept or approach used by then for routing.

Protocol/ method name / **Concept**	RL-MAC	QoS routing	FROMS	AdaR	ATP	Distributed GAPS	ADR	E&D Ants	IAR, AR	Forster and Murphy et al.,2006	SARA	Q-method /routing	CLIQUE	IDR
Q-learning & variation	√		√								√	√	√	√
Ant based							√	√	√					
Adaptive tree				√	√	√								

TABLE I. SOME PROTOCOLS USING RL APPROACH

III. CONCLUSION

Reinforcement learning approaches are now widely used for adapting to dynamic networks and performing well in such highly changeable environment. Most of the reinforcement learning based routing protocols uses Q-learning concept. The approach being used by a protocol depends upon service requirements and the kind of environment it is operating in. This paper has provided an overview of the reinforcement learning techniques or methods available for achieving adaptivity and finding optimal routes in resource-constrained environments of wireless sensor networks.

IV. REFERENCES

[1] Malik Tubaishat and Sanjay Madria, -Sensor Networks:An Overview, ‖ in IEEE 2003.

[2] R.S. Sutton and A.G. Barto, ―Reinforcement Learning,‖ in Journal of Cognitive Neuroscience,1999.

[3] A. Forster and A. Murphy, ―FROMS: Feedback Routing for Optimizing Multiple Sinks in WSN with Reinforcement Learning,‖ appeared in Intelligent Sensors, Sensor Networks and Information, 2007. ISSNIP 2007. 3rd International Conference on 3-6 Dec. 2007, pp. 371-376.

[4] Brajendra kumar singh, ―Adaptive routing for wireless sensor networks: a survey,‖ in Winter 2007.

[5] Leonid Peshkin and Virginia Savova, ―Reinforcement Learning for Adaptive Routing,‖ in IEEE 2002.

[6] Ping Wang and Ting Wang, ―Adaptive Routing for Sensor Networks using Reinforcement Learning,‖ appeared in Computer and Information Technology, 2006. CIT '06. The Sixth IEEE International Conference in Sept. 2006, Page(s) 219.

[7] Ying Zhang and Qingfeng Huang, ―A Learning-based Adaptive Routing Tree for Wireless Sensor Networks,‖ appeared in Journal of Communications, Volume 1, No 2 (2006), pp. 12-21, May 2006.

[8] Suh, Changsu, Young-Bae Ko, and Dong-Min Son. "An energy efficient cross-layer MAC protocol for wireless sensor networks." *Advanced Web and Network Technologies, and Applications*. Springer Berlin Heidelberg, 2006. 410-419.

[9] Liu, Shujuan, Yuebin Bai, Mo Sha, Qingyuan Deng, and Depei Qian. "CLEEP: A novel cross-layer energy-efficient protocol for wireless sensor networks." In *Wireless Communications, Networking and Mobile Computing, 2008. WiCOM'08. 4th International Conference on*, pp. 1-4. IEEE, 2008.

[10] Zhenzhen Liu and Itamar Elhanany, "RL-MAC:a reinforcement baed MAC protocol for wireless sensor networks," appeared in International Journal of Sensor Networks, Volume 1, Number 3-4/2006, pp. 117-124.

[11] Zhenzhen Liu and Itamar Elhanany, "RL-MAC: A QoS-Aware Reinforcement Learning based MAC Protocol for Wireless Sensor Networks," appeared in Networking, Sensing and Control, 2006. ICNSC '06. Proceedings of the 2006 IEEE International Conference, pp. 768-773.

[12] Kunal Shah and Mohan Kumar, "Distributed Independent Reinforcement Learning (DIRL) Approach to resource Management in Wireless Sensor Networks," in IEEE 2007.

[13] Petros Nicopolitidis, Georgios I. Papadimitiou, and Andreas S. Pomportsis, "Adaptive Wireless Networks Using Learning Automata," in IEEE Wireless Communications, 2011.

[14] Zahra Eskandari and Mohammad Hossien Yaghmaee, "AEESPAN: Automata Based Energy Efficient Spanning Tree for Data Aggregation in Wireless Sensor Networks," appeared in Wireless Sensor Networks, Volume 1, Issue 4, pp. 316-323, 2009 or Communication Systems, 2008. ICCS 2008. 11th IEEE Singapore International Conference on 19-21 Nov. 2008, pp. 943-947.

[15] T. Joshi, D. Ahuja, D. Singh, and D.P. Agrawal, "SARA: Stochastic Automata Rate Adaptation for IEEE 802.11 Networks," appeared in Parallel and Distributed Systems, IEEE Transactions in Nov. 2008, Volume: 19, Issue: 11, pp. 1579 – 1590.

[16] Y. Xing and R. Chandramouli, ―Stochastic Learning Solution for Distributed Discrete Power Control Game in Wireless Data Networks,‖ appeared in IEEE/ACM Trans. Net., Volume 16, Number 4, Aug.2008, pp. 932-44.

[17] Ehsan Ahvar, Alireza Pourmoslemi, and Md. Jalil Piran, "FEAR: A Fuzzy-based Energy-Aware Routing Protocol for Wireless Sensor Networks," appeared in arXiv.org on 13 August 2011.

[18] K. Arulmozhi, V.R. Sarma Dhulipala and R. Prabakaran, "A Study on Optimized Power Consumption Routing in Wireless Sensor Networks," in IEEE 2011.

[19] Shaoqiang dong, P. Agrawal, and K. Shiva, "Reinforcement learning based geographic routing protocol for UWB wireless sensor network," appeared in Global Telecommunications Conference, 2007. GLOBECOM '07. IEEE on 26-30 Nov. 2007, pp. 652-656.

[20] A. Forster and A. Murphy, "A Feedback-Enhanced Learning Approach for Routing in WSN," appeared in Communication in Distributed Systems (KiVS), 2007 ITG-GI Conference on Feb. 26 2007-March 2 2007, pp. 1-12.

[21] Abinash Mahapatra, Kumar Anand, and Dharma P. Agrawal, "QoS and energy aware routing for real-time traffic in wireless sensor networks," appeared in Computer Communications 29, no. 4 (2006), pp. 437-445.