iJRASET
International Journal For Research in
Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 6 **Issue:** IV **Month of publication:** April 2018

DOI: http://doi.org/10.22214/ijraset.2018.4584

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 6.887
Volume 6 Issue IV, April 2018- Available at www.ijraset.com

Reinforcement Learning based Routing Protocols in WSNs: A Survey

Anju Arya[1], Nisha Nehra[2]
[1]*Computer Science and Engineering Department, Deenbandhu Chhotu Ram University of Science and Technology*
[2]*Computer Science and Engineering Department, Maharishi Markandeshwar University*

Abstract: Recent advances in the technology and reduction in processor size, memory, and wireless antenna size has enabled the construction of low cost, low power and multifunctional computing devices (Sensor nodes) that resulted in high demand for development of Wireless Sensor Networks. Much of the research has been done in the development of routing protocols for WSNs. This paper investigates the routing protocols based on Reinforcement Learning for WSNs.
Keywords: Reinforcement Learning (RL), Wireless Sensor Network (WSN), Routing Protocols, Hole detection, Adaptivety.

I. INTRODUCTION

The Micro-Electro Mechanical Systems (MEMS) technology has acted as a catalyst in the world-wide attention to the development of WSNs and the scope of its application areas. A wireless sensor network (WSN) [1] consists of a collection of autonomous sensors densely deployed in a structured or unstructured manner to monitor a physical environmental and collect the relevant information and cooperatively transfer it through the network to the sink via gateway or intermediate sensor nodes (actually smart sensors). This dense deployment is necessary to obtain accurate information and achieve speed, flexibility, and reliability, especially in a highly dynamic environment. The research on sensor networks started with DSN program at DARPA by about 1980 and its beginning is marked by the 1998 SmartDust project. The development of these low-cost, low-power, and multifunctional devices (or smart sensors) has attracted a special attention after the Second World War and is primarily motivated by the military applications. After then, they found a major application in the field of Optimal Control Systems. They are now being used in many other non-defence applications or monitoring and tracking activities, such as environment or area monitoring, greenhouse monitoring, agriculture, structural monitoring, passive localization and tracking, etc.

The wireless sensor networks exhibit properties such as 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. With regard to the previously mentioned challenges, the major research challenges include power, hardware cost, security, system architecture (no unified architecture), real-time, programming abstractions, real world protocols, and analytical and practical results. The researchers are still trying to improve sensor devices by miniaturizing them as much as possible, without affecting their performance. However, a trade-off still exists between the QoS and the physical requirements. The main focus is on the two aspects, namely Information Quality and Energy Consumption. Thus, the efficient routing protocols are designed by keeping two things in mind namely, Information Processing and Energy Conservation. Reinforcement learning approach has now been widely used and applied to the routing protocols, at the network layer, to handle resource constraints. It was first proposed to minimize delay in the static networking routing problem. RL addresses mainly 2 problems:-

A. Prediction (evaluate the future, given a policy)
B. Control problems (optimize the future, find the best policy).

Reinforcement learning [2 is suitable for distributed problems (e.g. routing). It has medium requirements for memory and need of low computation at the individual nodes. The most widely used reinforcement learning algorithm is Q-Learning, which assigns a Q-Value to each possible action to represent their goodness or quality. The main advantage of RL based routing is that each node does not need global network information, but still it can approximate global optimality. Most of the RL techniques focus on finding an optimal path and increasing residual energy (energy saver approach) of the network, thus prolonging network lifetime and efficiency. However, some try to balance the energy consumption to efficiently use the available resources. The approach being us depends upon the aspect being focused or the QoS required according to the application area. In this paper, we investigate the routing protocols for WSNs based on Reinforcement Learning.

II. REINFORCEMENT LEARNING AND WIRELESS SENSOR NETWORKS

Routing is a way of efficiently transmitting or distributing data over the network and at the same time considering the other factors too, such as energy consumption, cost, quality of service (i.e. throughput, delay or performance) and network lifetime (i.e. Failure tolerance or handling and reliability). It consider four important factors - energy cost, robustness, throughput and delay. Most of the RL based protocols or methods focus on finding an optimal path and prolonging network lifetime by being energy-aware or increasing residual energy uniformly across the network.

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

A. On the basis of Adaptivity

Adaptivity is the most desired and unavoidable feature or requirement. It can be to dynamic topologies, application requirements, power management, spectrum usage or changing traffic patterns. It can be achieved by using:

1) *Adaptive Routing Schemes*: Adaptive Routing [4], describes the capability of a network which characterize the routes by their destinations and change in the route paths in response to the dynamic network changes or conditions. Thus, as many routes as possible can be allowed to handle the changes through adaptation. Adaptive routing (AdaR) [5], Adaptive Tree Protocol (ATP) [6] or learning based adaptive routing tree, Distributed GAPS [7], and RL-MAC [8] [9] comes under such schemes.
2) *Cross-Layer protocols*: Cross-layer concept is used because the OSI model is not necessarily the correct approach for some modalities of wireless systems. The competing goals of energy-efficiency and flexibility can be achieved through specialization through cross-layer protocol design and through modularity in a layered protocol design, respectively. The idea behind cross- layer design (CLD) is to exploit the information from multiple layers to jointly optimize performance of those layers. The System designers implement the functionality through either dynamic or static methods. Dynamic CLDs are used to respond to changing network conditions. Static CLDs utilize known characteristics of the network and layers. However, these modifications destroy most of the benefit that the OSI framework provided. Thus, while implementing cross-layer designs the designers must make careful decisions. Cross-layer design proposals are generally specified in the form of Optimization objectives (e.g. minimum node lifetime, total power consumption or concave functions of node lifetimes) and System constraints (e.g. Flow conservation). Some of the protocols using cross-layer concept are namely XLM (unified), MAC-CROSS, EOA, SCL, CLEEP and CLAR. RL-MAC [46] [48], DReL, and DIRL [11] are the cross-layer protocols using reinforcement learning concept.
3) *Learning Automata schemes*: A Learning Automata can be used to model the learning systems [11].

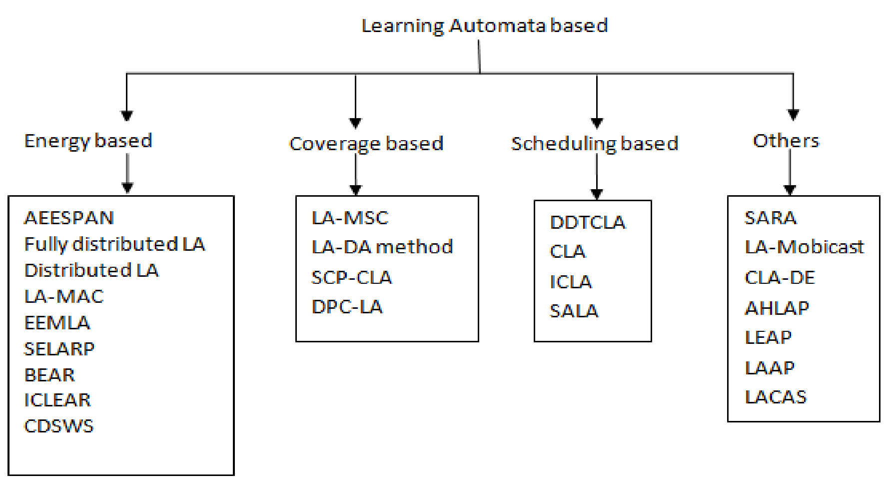


Fig 1. Classification on the basis of LA scheme used

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 6.887
Volume 6 Issue IV, April 2018- Available at www.ijraset.com

It doesn't require the knowledge of the environment it operates in. Its operation can be viewed in the framework of reinforcement learning. The sensor nodes equipped with an Intelligent Automata can learn or adapt to the environment as it changes and become intelligent as the time passes. It can solve various optimization problems. AEESPAN [12], SARA [13] [14], FEAR [15], etc. are some of the LA based protocols (or methods).

B. *On the basis of Routing Mechanisms*

The Routing protocols can be classified as structure-less and structure-based. The Structure based mechanisms use some data structure such as a Spanning tree or a Routing table to store information that can be used later on to take routing decisions or through multiple paths. These structures can be updated periodically or on-demand basis. They are useful for stable networks. The Structure-less mechanisms include real-time search approaches or methods (ant routing, TD methods), flooding based mechanisms and greedy approaches, suitable for dynamic networks. They have to take decisions at each hop. Adaptive Tree Protocol (ATP) [6] is both structured and structure-less.

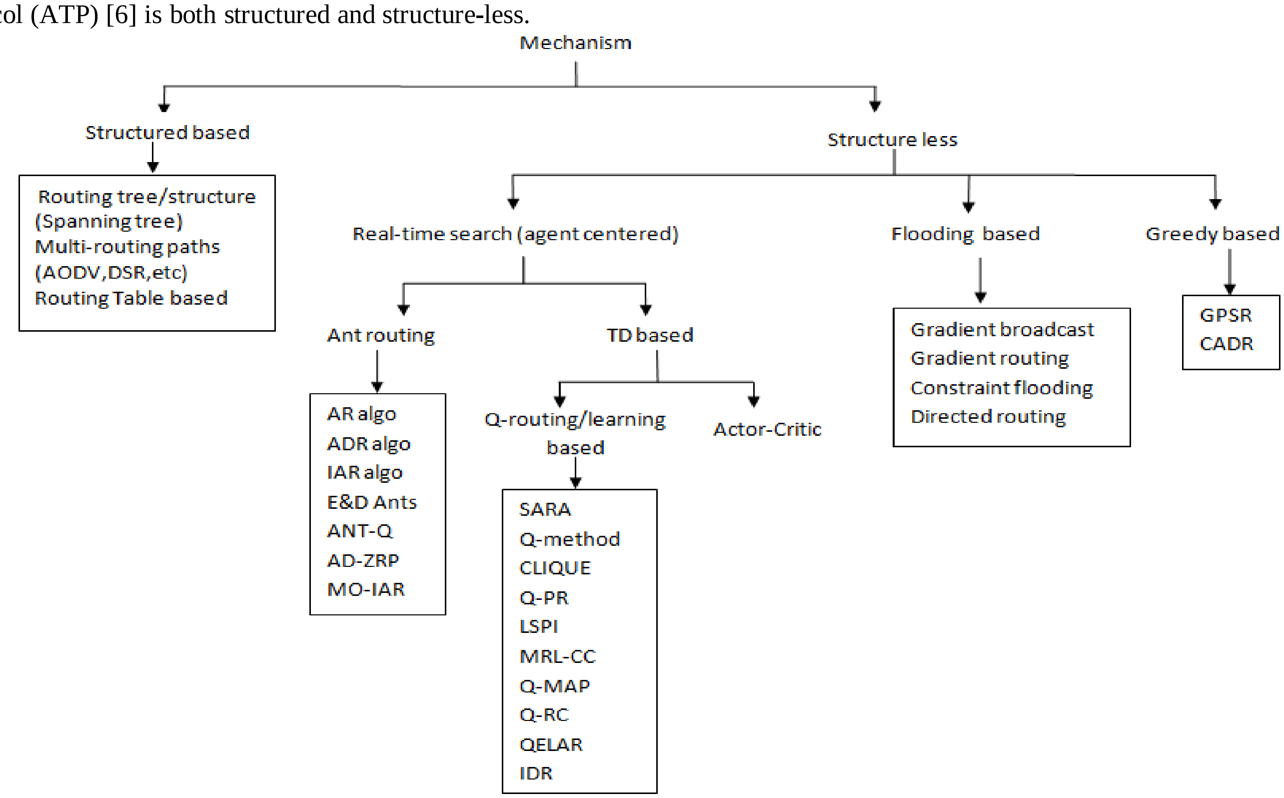


Fig 2. Classification on the basis of routing mechanism used

C. *On the basis of Routing Mechanisms*

A multi-agent system approach needs to be adopted for getting a more accurate model because of the large number of sensor nodes. The two types of coordination based RL approaches are:

1) Single-agent reinforcement learning (SARL) is suitable for centralized networks. When it is embedded in each host of the network, then it is referred to as Auction- based approach. Example: COORD for sensing coverage scheme to reduce power consumption.
2) Multi-agent reinforcement learning (MARL) is suitable for distributed networks where there are multiple agents. The two categories of MARL approaches in terms of the no. of hops involved in the payoff message propagation are Single-hop coordination-based MARL and Multiple-hop coordination- based MARL.

D. *On the basis of Aspect being focused*

1) *Energy-aware:* The energy-aware routing always tries to minimize energy waste and maximize network lifetime through uniform resource utilization and route packets such that energy consumption is distributed uniformly around a node (forwarding). It is being used in optimization problems. The factors affecting energy consumption have been discussed in [16],

such as routing path length, link reliability, aggregation, load balance (of nodes on shortest paths with low residual energy), or various sensor network design parameters. The Energy waste can be due to Idle listening (largest contribution due to lack of knowledge about packet transmissions), Packet collision (leads to retransmission), Message overhearing, Control packet overhead and Over-emitting (sending packet when destination is not ready). The energy-aware protocols can be categorized as Energy-saver and Energy- balancer. Examples include RLGR [17], DRLR [1], and FEAR [15].

2) *Feedback:* A Feedback is a scalar or real value that helps to take decisions related to routing and for learning and adapting to the network. It can be in the form of estimated route costs or link status, which in turn helps to select alternate paths, adjust links or start exploration process and helps to handle failure, recovery or node mobility situations. The idea of attaching routing feedback to data packets and learning is a powerful tool in the wireless network domain (sensors), since it requires only limited local knowledge and achieves significant results and also does not increase network costs. Example: FROMS[3] [18].

3) *Quality of Service:* In general, QoS refers to the quality as perceived by the end-user or application. From networking perspective, QoS is interpreted as a measure of service quality offered by the network to the end user or application. Achieving QoS provisioning in wireless sensor networks is not a straight forward task due to resource constraints (energy budgets, memory constraints), limited computational capabilities and limited BW associated with sensor technology. The meaning and elements defining QoS varies according to the application requirements or perspectives. The QoS requirement can be specified from two perspectives as application specific QoS and Network QoS. The different metrics or levels used to measure or define QoS can be latency, packet delivery within a defined deadline or transmission delays delay in real-time applications, reliability, accuracy, aggregation delay, coverage, fault tolerance and network lifetime, energy management , current system state (mobility, energy level, neighbors)and throughput(traffic priorities during varying loads), which are application specific. It can be achieved through Data Redundancy, which increases successful packet delivery rate & reliability, inspite of its routing and energy overhead. QoS requirements can be classified into two domains namely Timeliness and Reliability. Example: RL-MAC [8] [9] and QoS routing [19].

4) *Location Aware:* Location-based protocols use the position information to take decisions about next hop to forward the data packet. The nodes use the location and energy information of its neighbors to select the next hop. This information also helps to balance load distribution and distribute energy consumption uniformly over the network and thus avoids EISS and network partitioning problems. The main objectives are to reduce delay and maximize network lifetime. Example: RLGR [17].

5) *Fault-tolerant:* Fault tolerance is the ability of a system to provide a desired level of functionality even in the presence of faults. It is actually a critical issue in wireless sensor networks. It is highly desired in application such as clustering, time synchronization, gateway assignment, etc. The failure in a network can be due to energy depletion, hardware failure, communication or unpredictable link errors, node failure, routing failure, malicious attack, packet loss due to congestion and so on. The protocols with fault- tolerant and failure recovery capability are desired. Example: ATP [6] and FROMS [3] [18].

There is certainly no protocol that supports mobility, energy efficiency, scalability, data aggregation and fault tolerance at the same time. There would always be a trade-off between different quality factors or requirement.

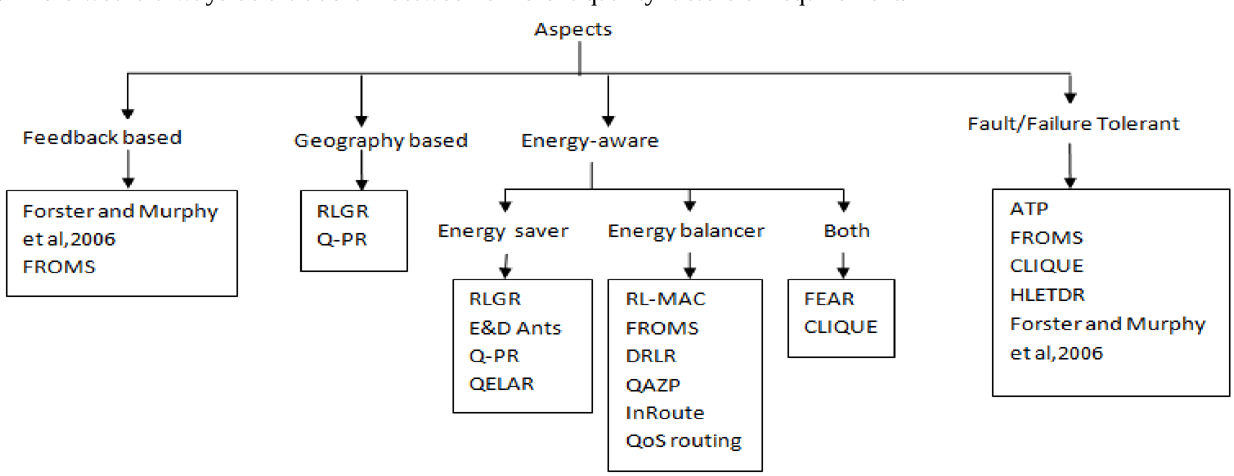


Fig 3. Classification on the basis of aspect being focused

TABLE I.
SUMMARY OF EXISTING ROUTING PROTOCOLS

Protocol/ method name				Aspect being Focused							
	Q-learning & variation	Ant based	Adaptive tree	QoS-aware	Fault Tolerance	Multi -agent	Sink mobility	Feed-back	Geo-graphy	Energy aware	Multiple sink
RL-MAC	✓			✓						✓	
QoS routing				✓						✓	
QELAR	✓									✓	
MRL-QRP				✓		✓					
MRL-CC	✓			✓		✓					
FROMS	✓				✓		✓	✓		✓	✓
AdaR			✓								
ATP			✓		✓		✓				
Distributed GAPS			✓								
ADR		✓									
E&D Ants		✓									
IAR, AR		✓									
MO-IAR		✓			✓						
ANT-Q	✓	✓									
GSOMs	✓										
Forster and Murphy et al.,2006			✓		✓			✓			✓
SARA	✓										
Q-method /routing	✓										
CLIQUE	✓				✓		✓			✓	✓

IDR	✓								✓		
Q-PR	✓								✓	✓	
Q-MAP	✓					✓					
RLGR									✓	✓	
FEAR										✓	
HLETDR					✓		✓				
DRLR				✓						✓	
QAZP	✓									✓	✓
InRoute	✓			✓						✓	✓
QwPAE	✓					✓					
Constant Share RL	✓					✓					
Max Share RL	✓					✓					
RL-LSSR				✓							

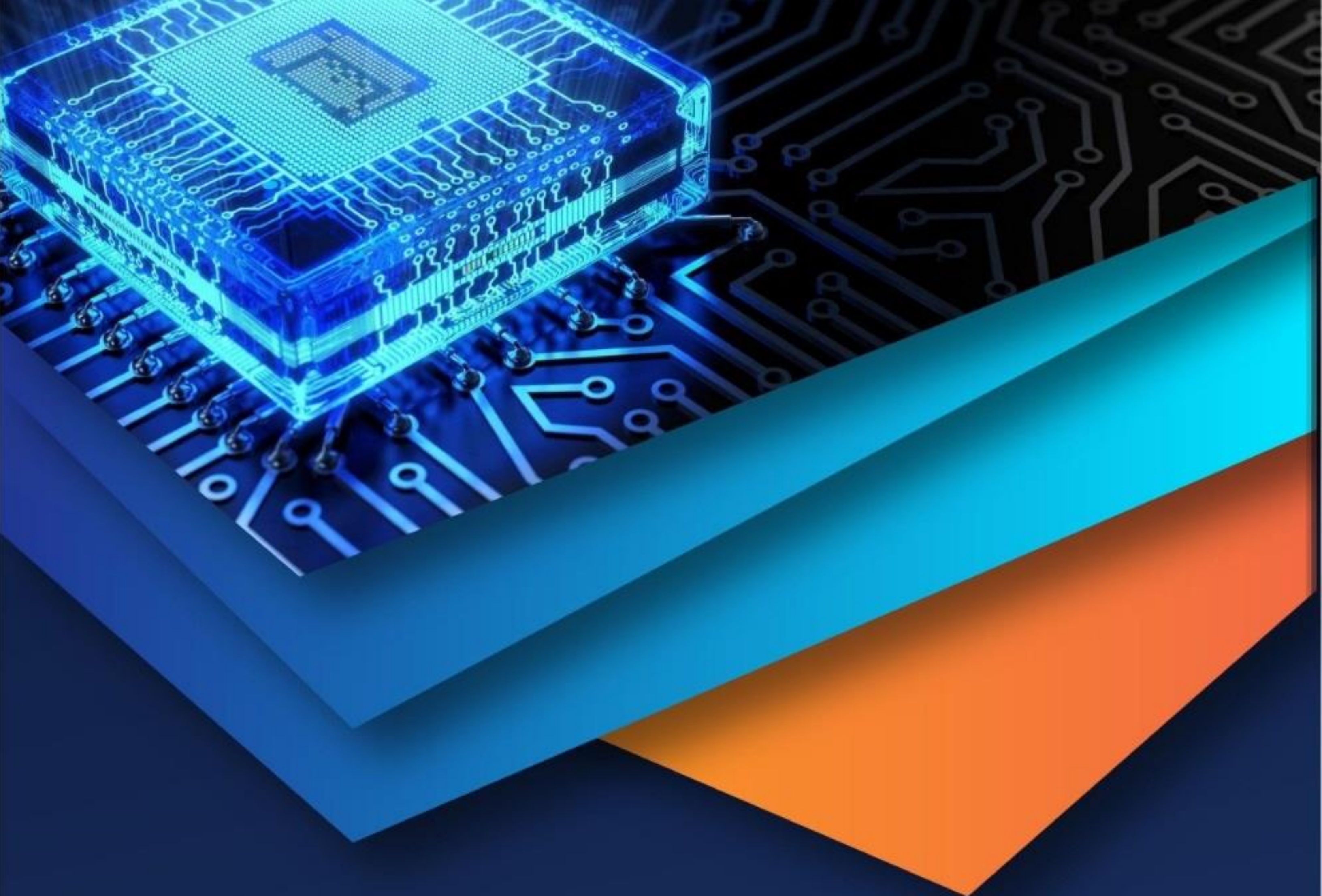
III. CONCLUSION

Most of the routing protocols or methods are using Q- learning concept. The most famous among them is FROMS. The protocol or method being used depends upon service requirements and the kind of environment operating in. This paper has just provided an overview to the techniques using reinforcement learning concept for achieving adaptivity and finding optimal routes in a resource-constrained environments of WSNs.

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] 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.
[6] 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.
[7] Leonid Peshkin and Virginia Savova, ―Reinforcement Learning for Adaptive Routing, ‖ in IEEE 2002.
[8] Zhenzhen Liu and Itamar Elhanany, "RL-MAC: a reinforcement based MAC protocol for wireless sensor networks," appeared in International Journal of Sensor Networks, Volume 1, Number 3-4/2006, pp. 117-124.
[9] 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.
[10] Kunal Shah and Mohan Kumar, "Distributed Independent Reinforcement Learning (DIRL) Approach to resource Management in Wireless Sensor Networks," in IEEE 2007.
[11] Petros Nicopolitidis, Georgios I. Papadimitiou, and Andreas S. Pomportsis, "Adaptive Wireless Networks Using Learning Automata," in IEEE Wireless Communications,2011.
[12] 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.

[13] 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.
[14] 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.
[15] [15] 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.
[16] K. Arulmozhi, V.R. Sarma Dhulipala and R. Prabakaran, "A Study on Optimized Power Consumption Routing in Wireless Sensor Networks," in IEEE 2011.
[17] 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.
[18] 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.
[19] 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.



10.22214/IJRASET

INDEX COPERNICUS 45.98

ISRA JIF IMPACT FACTOR: 7.129

TOGETHER WE REACH THE GOAL

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY