Reinforcement Learning based Routing Protocol for Underwater Acoustic Wireless Sensor Network

S. Ponlatha[#1], V.Pavithra[*2], D.Chitra[#3]

*[#1]Associate Professor, [*2]PG Scholar, [#3]Professor*
Department of Electronics and Communication Engineering, Mahendra Engineering College, Namakkal,Tamilnadu,India

Abstract: This paper focuses on Underwater wireless sensor networks (UWSNs) have emerged as a promising networking technology owing to their various underwater applications. Many applications require sensed data to be routed to a centralized location. However, the routing of sensor networks in underwater environments presents several challenges in terms of underwater infrastructure, including high energy consumption, narrow bandwidths, and longer propagation delays than other sensor networks. Efficient routing protocols play a vital role in this regard. Recently, reinforcement learning (RL)-based routing algorithms have been investigated by different researchers seeking to exploit the learning procedure via trial-and-error methods of RL. RL algorithms are capable of operating in underwater environments without prior knowledge of the infrastructure.

Keywords: Reinforcement learning,Underwater wireless sensor network.

I. Introduction

Underwater wireless sensor networks (UWSNs) represent an emerging field in wireless communication, owing to their significant advantages in various underwater applications. A typical UWSN consists of several self-configurable sensor nodes anchored to the ocean floor; these are interconnected by automatically adaptive wireless links featuring one or more underwater gateways.

Reinforcement learning

Reinforcement Learning (RL) is the science of decision making. It is about learning the optimal behavior in an environment to obtain maximum reward. This optimal behavior is learned through interactions with the environment and observations of how it responds, similar to children exploring the world around them and learning the actions that help them achieve a goal.ReinforcementLearning probleminvolves an agent exploring an unknown environment to achieve a goal. RL is based on the hypothesis that all goals can be described by the maximization of expected cumulative reward. The agent must learn to sense and perturb the state of the environment using its actions to derive maximal reward. The formal framework for RL borrows from the problem of optimal control of Markov Decision Processes (MDP). A useful abstraction of the reward signal is the value function, which faithfully captures the 'goodness' of a state. While the reward signal represents the immediate benefit of being in a certain state, the value function captures the cumulative reward that is expected to be collected from that state on, going into the future. The objective of an RL algorithm is to discover the action policy that maximizes the average value that it can extract from every state of the system RL algorithms can be broadly categorized as model-freeandmodel-based.Model-free algorithms .Do not build an explicit model of the environment, or more rigorously, the MDP. They are closer to trial-and-error algorithms that run experiments with the environment using actions and derive the optimal policy from it directly. Model-free algorithms are either value-based or policy-based.Value-based algorithms consider optimal policy to be a direct result of estimating the value function of every state accurately. Using a recursive relation described by the Bellman equation, the agent interacts with the environment to sample trajectories of states and rewards. Given enough trajectories, the value function of the MDP can be estimated.

Once the value function is known, discovering the optimal policy is simply a matter of acting greedily with respect to the value function at every state of the process. Some popular value-based algorithms are SARSA and Q-learning. Policy-based algorithms, on the other hand, directly estimate the optimal policy without modelling the value function. By parameterizing the policy directly using learnable weights, they render the learning problem into an explicit optimization problem. Like value-based algorithms, the agent samples trajectories of states and rewards; however, this information is used to explicitly improve the policy by maximizing the average value function across all states. Popular policy-based RL algorithms include Monte Carlo policy gradient (REINFORCE) and deterministic policy gradient (DPG)

II Problem Formulation

The agent observes the state of the environment during each decision step, and it selects actions randomly or by following a policy. Next, it receives an immediate reward based upon the selected action and goes on to the next state. The reward function is designed to provide feedback to the learning algorithm, reflecting the primary objective of the task.There, the agent observes state st from the environment. In that particular state, the agent chooses action at by exploration or exploitation. According to the taken action, the agent receives a reward rt and goes to the next state. To solve a problem with the help of RL, the problem should be designed as a Markov decision process (MDP) [36]. Therefore, MDP can be regarded as the theoretical basis of RL. The mathematical framework of MDP consists of a tuple of < S, A, P, R >, where S is a finite set of environment states, A is a set of actions available for the agent, P is the transition probability from the current state to the next state via a particular action, and R is the rewardreceived after transitioning to the next state with the taken action. The transition probability can be written as,

$$P_a(s, s') = P_r(s_{t+1} = s' | s_t = s, a_t = a), \qquad (1)$$

Where Pa is the probability of transitioning from state s at time t to state s 0 at time t + 1 by taking action a. After the transition from s to s 0, the agent receives an immediate reward, which can be denoted by Ra(s, s0). The reward represents an evaluation of the quality of an action in a particular state.

The goal of an RL agent is to identify a policy π that maximizes the cumulative rewards; typically, this is the expected discounted sum of rewards. The policy is a function that maps a given state to the probability of selecting each possible action from that state. Thus, following a policy π, the probability of taking action an in state s at time t can be denoted by π(a|s). The function that estimates how desirable it is for an agent to be in a given state, or how desirable it is to select a particular action in a given state, is called a value function. The value function can be a function of state or of state–action pairs.

III Routing protocols

Routing Protocols are the set of defined rules used by the routers to communicate between source & destination. They do not move the information to the source to a destination, but only update the routing table that contains the information. Network Router protocols helps you to specify way routers communicate with each other. It allows the network to select routes between any two nodes on a computer network

A.Static Routing Protocol

Static routing protocols are used when an administrator manually assigns the path from source to the destination network. It offers more security to the network**.**

B.Dynamic Routing Protocol

Dynamic routing protocols are another important type of routing protocol. It helps routers to add information to their routing tables from connected routers automatically. These types of protocols also send out topology updates whenever the network changes' topological structure.

Underwater wireless sensor architecture

Underwater wireless sensor networks comprise of nodes that are deployable on the surface and under the water. All nodes need to communicate and exchange information with other nodes in the same network and with the base station. Communication systems in the sensor network involve the transmission of data using acoustic, electromagnetic, or optical wave media. Among these types of media, acoustic communication is the most popular and widely used method due to its attenuation features in water. The factor of low transmission is derived from absorption and conversion of energy into heat in water. Meanwhile, acoustic signals operate at low frequencies, which enables them to be transmitted and received over long distances.

This study is conducted with the aim of investigating the domain of underwater wireless sensor networks and provide comprehensive insight into UWSNs requirements, platforms, recent advances, taxonomy and challenges of UWSNs. Additionally, this paper offers the newest evidence for various aspects that can satisfy the requirements for rapid development of UWSNs.

Longevity

Network lifetime is one of the key requirements of UWSNs. It has a significance impact on the cost, time, maintenance tasks and performing underwater sensor nodes. It is crucial for maximizing the network lifetime, especially for mobile sensor nodes operations. Therefore, firmware has a vital responsibility in ensuring an effective practice of hardware features such as sleep modes, allows interruptions to replace polling and easy to set up. In addition, routing protocol and deployment of nodes have a huge role in controlling the energy consumption. It leads to a significant amount of research works on the development and evaluation processes.

Accessibility

Each sensor node communicates to each other within a communication range located in the region. The communication range is another important requirement for UWSN which affect the density of nodes, deployment feasibility and the network cost of the targeted monitoring area. There are two communication modes for underwater sensor networks; Acoustic and Optical communication. Underwater acoustic wireless communications have been one of the most used technology as it is accessible and requires communication over great distances. However, acoustic waves still have many shortcomings including scattering, excessive delay because of the low propagation velocities, high attenuation, low bandwidth, and adverse effects on the underwater creatures.

Recently, orbital angular momentum has developed as an alternative multiplexing freedom to encrypt data onto vortex beams for enhancing the capacity of acoustic communication. Due to the limitations of acoustic communication, another approach is to use optical waves. According to [5], the current research on underwater optical communications focus on expanding the data rate and transmission range. Optical waves have the advantage of higher data rate, low latency, and energy efficiency at the expense of limited communication ranges.

Complexity

The specification of sensor node placement at the position is also crucial for UWSN. Thus, a complexity factor need to be considered before setting up the networking platform which incorporates physical aspect, firmware and network configuration of nodes placement. Additionally, routing protocol selection and computing complexity contributes in identifying routes dynamically with no added information or prior knowledge about other nodes. Apart from that, node algorithm complexity is another factor that need to be considered since it influences the energy optimization of the nodes. Local nodes' energy consumption is correspondingly depending on computational complexity and transmission power aspects. Underwater acoustic channel complexity such as multipath, Doppler shift, considerable attenuation and a high delay are also requirements that affect the performance of node localization methods.

Security and Privacy

UWSNs is correlated with security and privacy factors which related to sensor nodes connectivity, synchronization, and data transaction tasks. The dynamic features of underwater environment and its environment expose the network to various treats and malicious attacks. It is required for the networks to create trust before all the nodes can securely connect to the network to allow communication for information exchange. It is essential to study what level of security due to the increased computational load and the amount of transmitted data, yet consuming more energy within the network.

Environmental Sustainability

The deployment of communication technologies in UWSN is required to consider the impact to environment and wildlife. Reference reported that wildlife is influenced to ambient and boat noise which can lead to stress and the rise of extinction risk. Moreover, marine environment with increasing noise can generate behaviour changes, population distribution and hearing impairment of fish species.

Simulator: ns2

Ns (from network simulator) is a name for series of discrete event network simulators, specifically ns-1, ns-2 and ns-3. All of them are discrete-event network simulator, primarily used in research and teaching. ns-3 is free software, publicly available under the GNU GPLv2 license for research, development, and use. The goal of the ns-3 project is to create an open simulation environment for networking research that will be preferred inside the research community. It should be aligned with the simulation needs of modern networking research. It should encourage community contribution, peer review, and validation of the software. Since the process of creation of a network simulator that contains a sufficient number of high-quality validated, tested and actively maintained models requires a lot of work, ns-3 project spreads this workload over a large community of users and developers.

In 1996-97, ns version 2 (ns-2) was initiated based on a refactoring by Steve McCanne. Use of Tcl was replaced by MIT's Object Tcl (OTcl), an object-oriented dialect Tcl. The core of ns-2 is also written in C++, but the C++ simulation objects are linked to shadow objects in OTcl and variables can be linked between both language realms. Simulation scripts are written in the OTcl language, an extension of the Tcl scripting language. Presently, ns-2 consists of over 300,000 lines of source code, and there is probably a comparable amount of contributed code that is not integrated directly into the main distribution (many forks of ns-2 exist, both maintained and unmaintained). It runs on GNU/Linux, FreeBSD, Solaris, Mac OS X and Windows versions that support Cygwin. It is licensed for use under version 2 of the GNU General Public License.

NETWORK ANIMATOR (NAM)

Network animator (NAM) is an animator tool for viewing network simulation and real world packet teaches. It supports topology layout, packet level animation and various DATA inspection tool. Before to use NAM, trace file need to be created. This trace file is usually generated by NS.

It contains topology information, e.g nodes and links, as well as packet traces. During a simulation, the user can produce topology configuration, layout information and packet trace using tracing events in NS. Once the trace file is generated, NAM can be used to animate it.

ADVANTAGES OF NS2

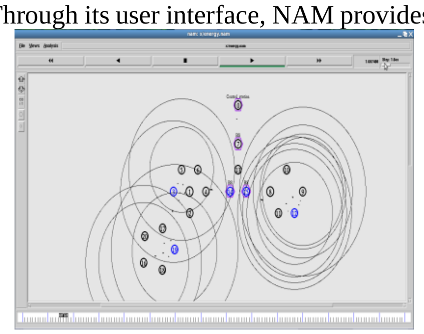
- NS is an ongoing effort of research and development.
- It does not require costly equipment.
- Simulators help in easy verification of protocols in less time, money
- Complex scenarios can be easily tested.
- NS offers support for simulating a variety of protocol suites and scenarios

OUTPUT:

- There are three types of outputs
  - Nam window
  - Tracing
  - Xgraph

NETWORK ANIMATOR

Network Animator (NAM) provides the packet level simulation output in a graphical manner. Network Animator is an animation tool for viewing network simulation traces and real world packet traces. It supports topology layout, packet level animation and various data inspection tools. Before starting to use NAM, a trace file needs to be created. This trace file is usually generated by NS. It contains topology information, e.g. nodes and links, as well as packet traces. During a simulation, the user can produce topology configurations, layout information and packet traces using tracing events in NS. Once the trace file is generated, NAM can be used to animate it as shown in figure 5.1. Upon startup, NAM will read the trace file, create topology, pop up a window, do layout if necessary and then pause at time 0. Through its user interface, NAM provides control over many aspects of animation.



- Fig : Simulated output of NAM

Tracing Trace packets on individual link Trace file format

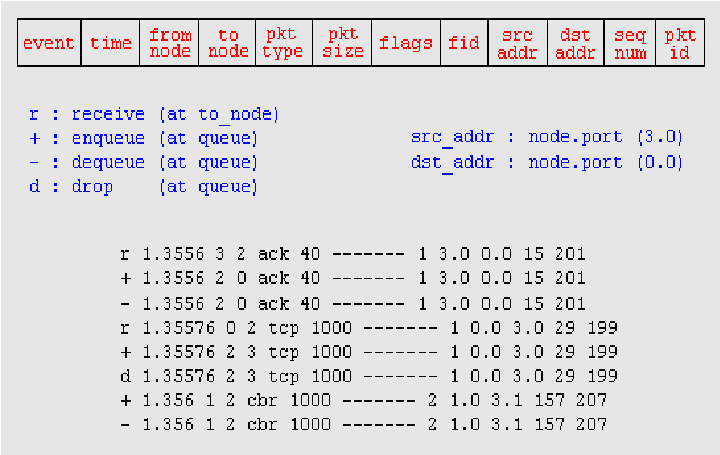


Figure : Model Xgraph

Provides the throughput comparison based on a graph which will be generated automatically based on the TCL coding

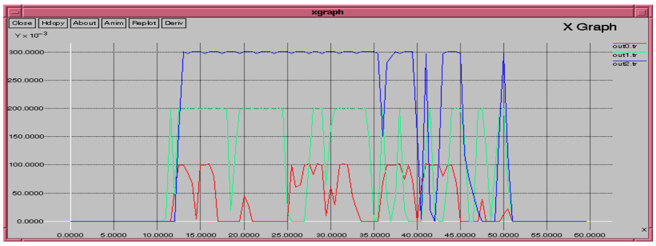


Fig : Xgraph for simulated output

SYSTEM MODELLING:

System modelling refers to an act of representing an actual system in a simply way. System modelling is extremely important in system design and development, since it gives an idea of how the system would perform if actually implemented.

Traditionally, there are two modelling approaches: analytical approach and simulation approach.

ANALYTICAL APPROACH:

The general concept of analytical modelling approach is to first come up with a way to describe a system mathematically with the help of applied mathematical tools such as queuing and probability theories, and then apply numerical methods to gain insight from the developed mathematical model. When the system is simple and relatively small, analytical modelling would be preferable (over simulation). In this case, the model tends to be mathematically tractable. The numerical solutions to this model in effect require lightweight computational efforts.

If properly employed, analytical modelling can be cost-effective and can provide an abstract view of the components interacting with one another in the system. However, if many simplifying assumptions on the system are made during the modelling process, analytical models may not give an accurate representation of the real system.

SIMULATION APPROACH:

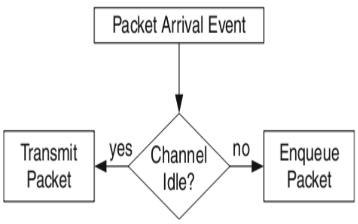
Simulation is a process of designing a model of a real system and conducting experiments with this model for the purpose of understanding the behavior of the system and/or evaluating various strategies for the operation of the system.

Simulation is widely-used in system modeling for applications ranging from engineering research, business analysis, manufacturing planning, and biological science experimentation, just to name a few.

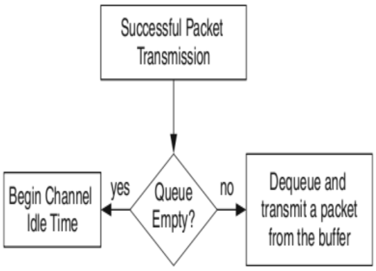
Compared to analytical modeling, simulation usually requires less abstraction in the model (i.e., fewer simplifying assumptions) since almost every possible detail of the specifications of the system can be put into the simulation model to best describe the actual system.

When the system is rather large and complex, a straightforward mathematical formulation may not be feasible. In this case, the simulation approach is usually preferred to the analytical approach.

In common with analytical modelling, simulation modelling may leave out some details, since too many details may result in an unmanageable simulation and substantial computation effort. It is important to carefully consider a measure under consideration and not to include irrelevant detail into the simulation.



Packet Arrival event



Successful Packet Transmission

CONCLUSION

In this paper I have introduced Routing for UWSNs is one of the most crucial issues in underwater applications. In RL, the efficiency of a system is increased with experience and time. This capability of RL algorithms has been widely considered in different wireless networking scenarios. RL has also been shown to significantly improve the design of routing protocols or UWSNs. In this article, present an extensive survey of RL-based underwater routing protocols. The methods are discussed, and their advantages, disadvantages, and suitable application environments are presented. The reviewed protocols are further compared in terms of their key ideas, RL mechanisms, optimization parameters, and evaluation techniques. The applications of RL are also separately compared for all protocols. For future researchers, the research gaps and areas requiring critical improvement are emphasized as open research issues. The analysis, discussion, comparison, and future research directions highlighted in this investigation will provide UWSN researchers with an in-depth overview of existing routing protocols. When compared to previously proposed methodologies, we obtained considerable gains on both actual and simulated data.

REFERENCE

1. M. Jouhari, K. Ibrahimi, H. Tembine, and J. Ben-Othman, "Underwater wireless sensor networks: A survey on enabling technologies, localization protocols, and internet of underwater things," IEEE Access, vol. 7, pp. 96 879–96 899, 2019.
2. I. F. Akyildiz, D. Pompili, and T. Melodia, "Underwater acoustic sensor networks: research challenges," Ad hoc networks, vol. 3, no. 3, pp. 257– 279, 2005.
3. H. Kaushal and G. Kaddoum, "Underwater optical wireless communication," IEEE access, vol. 4, pp. 1518–1547, 2016.
4. A. Palmeiro, M. Martin, I. Crowther, and M. Rhodes, "Underwater radio frequency communications," in OCEANS 2011 IEEE-Spain. IEEE, 2011, pp. 1–8.
5. C. Gabriel, M.-A. Khalighi, S. Bourennane, P. Leon, and V. Rigaud, "Channel modeling for underwater optical communication," in 2011 IEEE GLOBECOM Workshops (GC Wkshps). IEEE, 2011, pp. 833–837.
6. P. Lacovara, "High-bandwidth underwater communications," Marine Technology Society Journal, vol. 42, no. 1, pp. 93–102, 2008.
7. E. M. Sozer, M. Stojanovic, and J. G. Proakis, "Underwater acoustic networks," IEEE journal of oceanic engineering, vol. 25, no. 1, pp. 72– 83, 2000.
8. M. Stojanovic and J. Preisig, "Underwater acoustic communication channels: Propagation models and statistical characterization," IEEE communications magazine, vol. 47, no. 1, pp. 84–89, 2009.
9. M. Chitre, S. Shahabudeen, and M. Stojanovic, "Underwater acoustic communications and networking: Recent advances and future challenges," Marine technology society journal, vol. 42, no. 1, pp. 103–116, 2008.
10. Z. Zeng, S. Fu, H. Zhang, Y. Dong, and J. Cheng, "A survey of underwater optical wireless communications," IEEE communications surveys & tutorials, vol. 19, no. 1, pp. 204–238, 2016.
11. G. SchirripaSpagnolo, L. Cozzella, and F. Leccese, "Underwater optical wireless communications: Overview," Sensors, vol. 20, no. 8, p. 2261, 2020.
12. C. Gabriel, M.-A. Khalighi, S. Bourennane, P. Leon, and V. Rigaud, "Channel modeling for underwater optical communication," in 2011 IEEE GLOBECOM Workshops (GC Wkshps). IEEE, 2011, pp. 833–837.
13. S. Arnon, "Underwater optical wireless communication network," Optical Engineering, vol. 49, no. 1, p. 015001, 2010.
14. I. F. Akyildiz, P. Wang, and Z. Sun, "Realizing underwater communication through magnetic induction," IEEE Communications Magazine, vol. 53, no. 11, pp. 42–48, 2015.
15. M. C. Domingo, "Magnetic induction for underwater wireless communication networks," IEEE Transactions on Antennas and Propagation, vol. 6, no. 60, pp. 2929–2939, 2012.