

Reinforcement Learning Based Scheduling in Underwater NDN

Muhammad Toaha Raza Khan, Muhammad Saad Malik, Muhammad Ashar Tariq*, Md. Mahmudul Islam, Junho Seo,

Ru Yang, Dongkyun Kim

School of Computer Science and Engineering, Kyungpook National University, South Korea

**Department of Artificial Intelligence, Kyungpook National University, South Korea*

{toaha,maliksaad,tariqashar,mislam,junhoseo,yr0818,dongkyun}@knu.ac.kr

Abstract—Day by day the demand for monitoring the marine environment and exploring the ocean is increasing. However, the limited bandwidth and long propagation delays and transmission delays make it challenging. In this paper, we propose the Reinforcement Learning (RL) based efficient scheduling for underwater nodes which communicates with the sink node using Named Data Networking (NDN) protocol. The RL based approach is applied at the sink node which provides assistance to the sensor nodes in scheduling for transmitting data. RL based approach assist in avoiding collision of packets that engenders efficient use of an acoustic channel.

Index Terms—Reinforcement Learning (RL), Named Data Network (NDN), Underwater, Surface Sink, Sensor Nodes

I. INTRODUCTION

Internet of underwater things (IoUT) is defined as the network of various smart underwater objects which can communicate to each other [1]. Different types of underwater entities including autonomous underwater vehicles (AUVs), underwater sensor nodes, surface sinks and so on for the network in IoUT. The underwater nodes in IoUT, however, have a limited processing capability and battery. Therefore, these nodes need to communicate using some lighter protocol for minimum resource consumption. Constrained application protocol (CoAP) is a lightweight internet of things (IoT) protocol developed for the constrained IoT devices. Since the underwater nodes are constrained, CoAP is a highly suitable candidate for communication among nodes, AUVs and surface sinks or ships. The CoAP has a minimum header size of 4 bytes and supports reliable communication with a built-in congestion control mechanism. A detailed overview of the CoAP can be found in [2].

The communication of multiple underwater objects like AUVs, nodes, etc. with the surface sink requires scheduling to communicate over same channel. Numerous scheduling schemes are available for communication between surface sink and underwater objects. Two scheduling schemes based on time domain multiple access (TDMA) are discussed in [3] for communication between underwater sensor nodes and surface sink, named as Transmit Delay Allocation MAC (TDA-MAC) and Accelerated TDA-MAC. The TDA-MAC uses ping messages to calculate the propagation delay between each node and surface sink and then sends a transmit delay instruction

(TDI) packet to each node, informing the amount of time it has to wait to start transmission after receiving the request (REQ) packet from the sink. The accelerated TDA-MAC caters the channel underutilization issues in TDA-MAC.

The limited bandwidth and slow propagation speed of acoustic signals leads to low data throughput for underwater networks. Machine learning techniques such as RL approach is applied to medium access control that engenders efficient use of an acoustic channel. In [4] RL based mechanism is applied for distributed scheduling in underwater networks to avoid collision. However, if two nodes sense the channel simultaneously before taking action, there is also a probability of collision. If the two nodes share the neighboring information between each other such as for neighbor discovery, the exchange of beacon messages can reduce the probability of collision. However exchange of beacon messages can increase the network overhead. In this paper we proposed the semi-centralized scheduling for transmission of data by the nodes using RL based mechanism.

The remainder of this paper is as follow. Section II gives the overview of ALOHA in underwater networks. Section III discuss about the proposed approach for scheduling in underwater NDN. Finally the conclusion is drawn in Section IV.

II. ALOHA IN UNDERWATER NETWORKS

Underwater sensor networks (UWSNs) is quite similar to terrestrial wireless networks, both shares a common channel for message propagation. Having a common channel for transmission and reception leads to collision when multiple devices access the shared medium. To reduce the data packet collision there should be a mechanism which will control the allocation of the common channel to various users. MAC protocols do the job for efficient channel access mechanism. The primary task of a MAC protocol is to avoid collision when allocating channel access to different nodes. Pure ALOHA (P-ALOHA) was the earliest contention-based MAC protocol invented in the 1970s. Slotted ALOHA (S-ALOHA) was proposed to enhance P-ALOHA. S-ALOHA divides the transmission time into multiple slots [5]. But still there is a probability of collision if two nodes sense the channel simultaneously.

III. REINFORCEMENT LEARNING BASED SCHEDULING IN UNDERWATER NDN

In underwater networks there is long propagation delays and transmission delays. In underwater NDN based network, the channel for transmission is divided into slotted ALOHA. Each node when have data to send, sense the transmission channel and sends the data on the respective free slot. The channel is divided into limited number of larger time span slots because of long propagation delay. When nodes sense the channel for data transmission there is a probability of collision when both nodes sense the channel simultaneously for transmission. In order to overcome this problem we proposed the centralized scheduling of data transmission for nodes using RL approach. This scheme still allows the nodes to behave in distributed manner rather than fully controlled by the surface sink as in centralized scheduling scheme. Wherein distributed scheduling scheme which is applied at the underwater nodes increase the network overhead, if nodes share the beacon message with the neighboring nodes which is not desired for underwater networks.

Machine learning based RL mechanism is applied at the surface sink. The surface sink gives the opinion regarding the scheduling of data transmission and broadcast the message for other nodes not to transmit at the respective time. The RL based machine learning is applied at the surface sink which is considered as the agent as shown in Fig. 1.

RL based mechanism consists of state and action taken by the agent and in return gets the reward. The surface sink observes the current state and based on optimal policy takes the action. The surface sink gets the message packet from nodes. From message packet it extracts the control information and the data information. The data information comprises of 3 dimensional hierarchical scheme such as $\backslash\backslash$ location\time\type. The receive data message gives information regarding the location of the sensing region, at what time data was sensed and the type of sensed data such as salinity and turbidity. The control information collected from the message packet gives information regarding the received interference strength in channel (I_t), channel gain between sending node and the surface sink (H_t), selected slot indices (N_t) for transmission and (L_t) represents the proportion of bits remaining to transmit. So the state comprises of number of observations which includes as given by equation (1).

$$S_t = [I_t, H_t, N_t, L_t] \quad (1)$$

The agent based on the policy takes the action gives reward to the decision made. The surface sink after receiving the message packet sends the scheduling intervals to the nodes and also sends the acknowledgement to the sender. The agent after taking the action rewards +1 if no collision happens, -1 if collision happens and -0.5 if the sink node receives the duplicate packet. The collision is detected if the sink does not receive the message packet at the given slot assigned to the sending node. The objective of the RL is to find the policy to

maximize the expected cumulative reward as given equation (2).

$$R_t = E[\sum_{n=0}^{\infty} \beta^n r_t + n] \quad (2)$$

where $\beta \in [0, 1]$ is the discount factor.

The state transition and reward are stochastic and modelled as a Markov decision process (MDP), where the state transition probabilities and rewards depend only on the state of the environment and the action taken by the agent. The transition from s_t to s_{t+1} with reward r_t when action a_t is taken can be characterized by the conditional transition probability, $p(s_{t+1}, r_t | s_t, a_t)$.

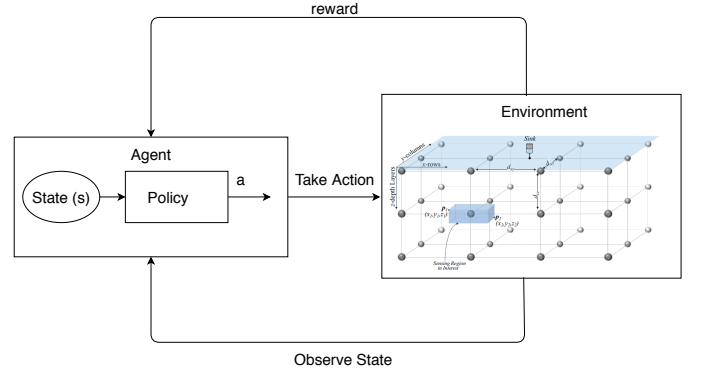


Fig. 1. Reinforcement Learning for Scheduling in Underwater NDN Network

IV. CONCLUSIONS

In this paper, we propose the RL based scheduling scheme for the transmission of data. This RL based approach is applied at the surface sink which performs the scheduling on the behalf sensor nodes to avoid the collision of data packets. RL based potentially offers opportunities for underwater NDN network design, due to its adaptive capability and its responsiveness to environmental changes.

V. ACKNOWLEDGEMENT

This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (2016R1D1A3B01015510).

REFERENCES

- [1] C.-C. Kao, Y.-S. Lin, G.-D. Wu, and C.-J. Huang, "A comprehensive study on the internet of underwater things: applications, challenges, and channel models," *Sensors*, vol. 17, no. 7, p. 1477, 2017.
- [2] M. A. Tariq, M. Khan, M. T. Raza Khan, and D. Kim, "Enhancements and challenges in coop—a survey," *Sensors*, vol. 20, no. 21, p. 6391, 2020.
- [3] N. Morozs, P. Mitchell, and Y. V. Zakharov, "Tda-mac: Tdma without clock synchronization in underwater acoustic networks," *IEEE Access*, vol. 6, pp. 1091–1108, 2017.
- [4] S. H. Park, P. D. Mitchell, and D. Grace, "Reinforcement learning based mac protocol (uw-aloha-q) for underwater acoustic sensor networks," *IEEE Access*, vol. 7, pp. 165531–165542, 2019.
- [5] K. Chen, M. Ma, E. Cheng, F. Yuan, and W. Su, "A survey on mac protocols for underwater wireless sensor networks," *IEEE Communications Surveys & Tutorials*, vol. 16, no. 3, pp. 1433–1447, 2014.