Design of S-Aloha Protocol based on Q-Learning for Underwater Acoustic Sensor Network

Yangxia Lu[1,a], Dongling Zou[1]

[1]College of Information Engineering, Shanghai Maritime University, Shanghai 200000, China.

[a]584609134@qq.com

Abstract

To solve the problem of collisions among data packets in underwater acoustic sensor networks with variable topologies, a slotted Aloha protocol based on Q-Learning algorithm (QS-Aloha) is proposed. Each sensor node in the protocol can adaptively learn and select the optimal slot while transmitting data, so as to reduce the collision among packets and improve the transmission success rate. After the system converges, the proportional fair scheduling algorithm is applied to select the better node for transmission scheduling, so as to obtain the communication balance state. Simulation results show that QS-Aloha protocol has a higher success rate than traditional slotted Aloha protocol, and the performance of the protocol system after adding proportional fair scheduling algorithm has been better optimized.

Keywords

Underwater Acoustic Communication; MAC Protocol; Q-Learning; Proportional Fair Scheduling.

1. Introduction

The nodes in underwater acoustic sensor networks (UASN) use sensors to monitor, collect and fuse the data, and then send the information to the receiver through the nodes with transmission capacity [1]. The Medium Access Control (MAC) protocol in UASN coordinates network nodes to access channels, so that can share a common wireless channel with multiple neighboring nodes within the transmission range. When data packets collide, the protocol uses corresponding control mechanisms to ensure that limited channel resources are allocated to terminals reasonably and fairly, so as to reduce and avoid collisions among packets. Compared with terrestrial wireless channel, underwater acoustic channel resources are limited and complex, so designing effective MAC protocol is an important direction in underwater acoustic sensor network research [2-3].

According to the channel allocation method, the MAC protocol can be divided into a scheduling-based MAC protocol and a contention-based MAC protocol [4]. The contention-based MAC protocol is mainly for data transmission by competing for channel resources, and is divided into random Aloha access mode and handshake-based mode. As the earliest competitive MAC protocol, Aloha protocol can be directly applied to underwater acoustic networks [5]. They will send packets immediately when nodes have data to send. While packets sent by multiple nodes overlap at the receiver, they need to wait for the next random time to be transmitted. Obviously, this mechanism makes Aloha protocol further aggravate the collision of data packets after retransmission. In terrestrial wireless networks, the concept of slot Aloha(S-Aloha) [6] is put forward, which divides the transmission time into slots of the same length and transmits synchronously at the beginning of the slots. Compared with the traditional Aloha protocol, this protocol effectively reduces the collision of data packets. However, spatial temporal uncertainty exists in underwater acoustic network [7]. Nodes that send packets earlier may conflict with other nodes at the sink due to large transmission delay, and the system performance is not improved better in underwater acoustic network. Therefore, Zou et al. [8] put forward a scheme to improve the S-Aloha protocol, which ensures that packets arrive at the receiver within the slot by adjusting the sending time of nodes, but does not adopt collision avoidance mechanism, so the collision probability of packets is still very high.

Therefore, in order to solve the problem of packet collision in underwater acoustic dynamic network, this paper proposes a slot selection strategy based on reinforcement learning Q-Learning algorithm on S-Aloha protocol. By selecting the optimal slot to control the sending and receiving of nodes, the probability of successful packet transmission is improved and the collision is minimized. When the network tends to converge, the proportional fair scheduling algorithm is further adopted to schedule the better nodes for transmission. Simulation results show that the successful packet transmission rate of S-Aloha protocol with reinforcement learning is higher than that of traditional Aloha protocol, and the network performance is optimized after adding proportional fair scheduling algorithm.

2. Protocol description

This section introduces the QS-Aloha protocol proposed in this paper. Firstly, it introduces the conflict problem of slotted Aloha protocol in the process of data transmission in underwater acoustic sensor network, and then proposes an improved transmission scheme, which applies Q-learning to S-Aloha protocol to realize conflict minimization, and adds proportional fair scheduling algorithm after the system is stable to improve the transmission success rate of data packets.

2.1 The problem of S-Aloha conflict

In underwater acoustic sensor networks, S-Aloha protocol divides time into the same slots, each node sends data at the beginning of the slot, and whether the data packet is successfully received at the receiving end is not only related to the data packet sending time, but also related to the propagation delay. The slot model is shown in Figure 1. Due to the influence of spatial location, the data packet sent by node A in the first slot and the data packet sent by node B in the first slot collide at the sink end, and the packet transmission fails.

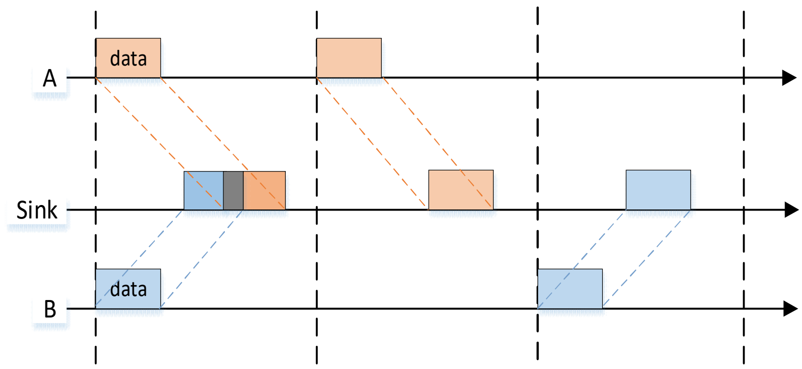

Fig.1 S-Aloha protocol model

Therefore, in order to reduce the collision probability of the system and obtain a higher success rate of data transmission, this paper adopts the slot Aloha model shown in Figure 1, and combines the Q-Learning algorithm to realize the intelligent selection of slots by sensor nodes, so that the nodes can adaptively learn the optimal slots while transmitting data, thus improving the system performance.

2.2 Design of S-Aloha protocol based on Q-Learning

In reinforcement learning (RL) [9], agents interact with the external environment through state perception and action selection. Under the current system state, agents select actions from the action space by means of feedback rewards, so as to maximize the expected benefits and decide their own behaviors.

Q-Learning algorithm, as the most widely used algorithm in reinforcement learning method, is an off-line strategy control algorithm, which is expressed by state action value function $Q(s,a)$, and its formula is as follows:

$$Q'(s_t,a_t) = Q(s_t,a_t) + \alpha[r_t + \gamma \underbrace{max}_{a} Q(s_{t+1},a_{t+1}) - Q(s_t,a_t)] \tag{1}$$

Where α is the learning rate (usually set to less than 1), which affects the results of previous training, and the discount factor $\gamma \in [0,1]$, which is used to weigh the immediate return and the future long-term return. The greater γ, the greater the impact on previous learning experience. r_t is a return function, which generally has two types: continuous change and constant according to the actual scene. The value of $Q(s_t, a_t)$ is the cumulative value obtained by the agent selecting action a according to the policy and executing it circularly in the state s at a certain moment. It can be seen from the formula that the maximum value of $Q(s_{t+1}, a_{t+1})$is adopted in each step, which is equivalent to the action with the maximum value of $Q(s_{t+1}, a_{t+1})$. Therefore, the main idea of the algorithm is to build the state and action into a Q table to store the Q value, and select the action that can get the maximum benefit according to the Q value.

In this paper, time is divided into frame structure, each frame contains multiple slots, and each node can only select one slot for transmission in the same frame. At the same time, each sensor node has its own Q value in the slot, and its Q value is updated according to the transmission result. When the network is in a stable state, the node only needs to be active in the slot with the highest Q value, and does not need to listen idle.

The state s of the sensor node is active in the slot allocation problem, which means it is still in the process of learning, it is converted into node serial number i, and the value of action a is converted into that the node selects a certain slot k for data transmission. It can be expressed by $Q_t(i,k)$, that is, the action value function obtained after node i sends packets to slot k at time t. Because the state of the sensor network channel is unpredictable, and the future return of randomness is meaningless to update the whole Q value, γ is set to 0, so the iterative update equation of Q-Learning according to formula (2) is:

$$Q_{t+1}(i,k) = Q_t(i,k) + \alpha[r_t - Q_t(i,k)] \tag{2}$$

Set the return value r_t={1, −1}, if the transmission is successful, get the positive reward value, and if fails, feedback the negative value. Q-Learning determines the strategy π according to the optimal action value function, and the final output strategy is:

$$\pi = arg\ maxQ(i,k) \tag{3}$$

Taking figure 2 as an example, assume that there are two sending nodes and one receiving node in the network, each frame contains four slots, and the length of each data packet is less than one slot.

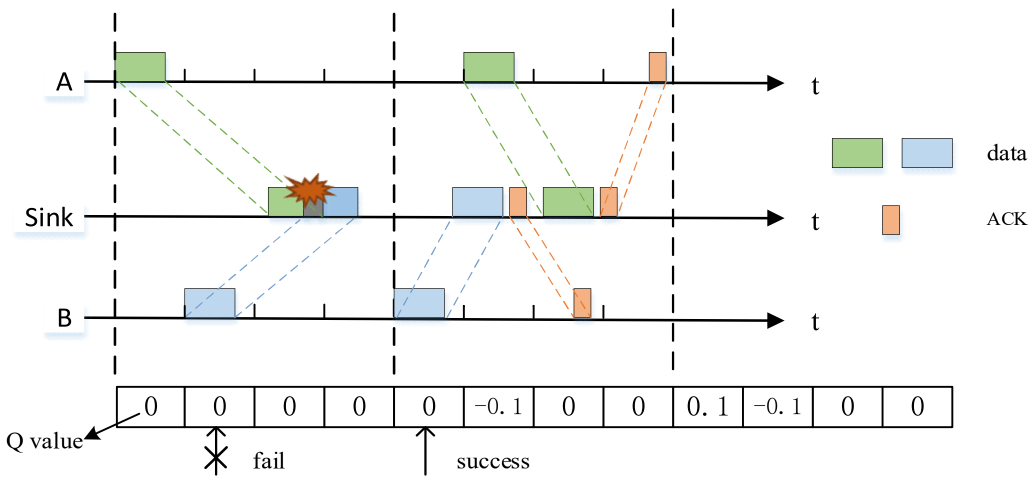

Fig.2 Transmission process diagram of QS-aloha protocol

The following describes the specific implementation process of QS-Aloha protocol:

(1) Slots in a frame are given unique weights, and the weights are initialized to 0. After each transmission attempt, the Q value table is updated according to the feedback of the receiver.

(2) During the first transmission, the node randomly selects a slot in the frame to send a data packet, and if the acknowledgement packet is received, the report value is +1. In figure 3, the node B selecting slot 2 collides with the packet of node A at the receiver, and returns the slot report value -1, so that the Q table is updated.

(3) From the second transmission attempt, the node selects the highest value as the current preferred slot and sends data. If multiple slots have highest weight, the node will randomly select one of them.

From the example of slot selection based on Q-Learning in Figure 2, it can be seen that a single Q value reflects the transmission history of a specific slot to a certain extent. With the increase of transmission attempts, all active nodes can find their own contention-free slots, that is, the system reaches a stable state. Therefore, compared with the traditional S-Aloha protocol, the improved transmission scheme is easier to improve the success rate of data packet transmission. At the same time, QS-Aloha protocol does not need to exchange information in advance to pre-allocate slots for each node, and has strong adaptability to network topology.

For the design of MAC protocol in underwater network, we should not only consider the performance of collision avoidance, but also consider the fairness factor. The fairness of the system refers to the reasonable scheduling among multiple nodes and the allocation of limited resources. Therefore, this paper uses the proportional fair scheduling algorithm to further optimize the system performance.

2.3 Proportional fair scheduling algorithm

Proportional Fair Scheduling (PF) algorithm [10] is a widely used strategy in the field of radio frequency at present, which is a compromise between fairness and system throughput. For example, it is set as the default scheduling strategy in the downlink of CDMA system, so it is applied to underwater communication system as the benchmark of system performance analysis.

With the reduction of packet collisions, the system channel is unbalanced in resource utilization, which creates conditions for the design of scheduling algorithm. Therefore, this paper selects one of the users who share the same channel to send data, so as to achieve a better data rate.In the k slot, the available channel capacity of each node in this slot is normalized to the actual average channel capacity of this node so far, and the node with the largest ratio is selected as the active node of this slot, with the formula as follows:

$$i_k = \arg\max_i \frac{R_{i,k}}{\bar{R}_{i,k}} \tag{4}$$

where $\bar{R}_{i,k}$ represents the average channel capacity of node i in k slots, and the expression is:

$$R_{i+1,k} = \left(1 - \frac{1}{T_s}\right) \bar{R}_{i,k} + \frac{1}{T_s} R_{i,k} \delta(k - i) \tag{5}$$

T_s is called time frame, which is measured by the number of slots. When the number of nodes is equal to the number of slots, $\delta(k - i) = 1$, otherwise, it is 0.

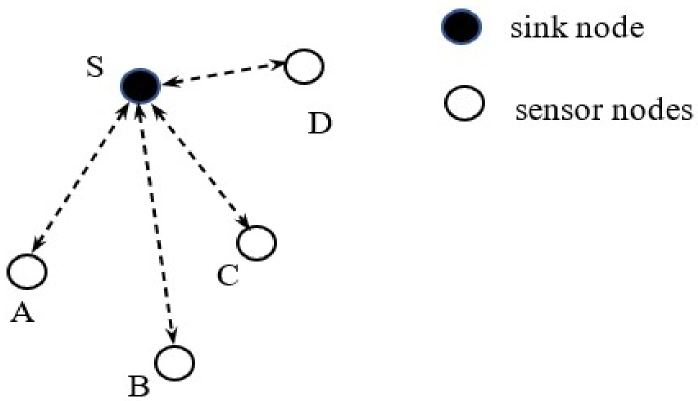

Fig.3 Topology diagram of underwater acoustic sensor network

3. Simulation and analysis

3.1 Simulation model

The topology diagram of underwater acoustic sensor network is shown in Figure 3. Sink node is in the center of the network, which is responsible for collecting information and has the function of data storage and forwarding. Multiple sensor nodes are randomly distributed around the sink node to generate packets of the same length, and send data to the sink node according to the MAC protocol mechanism. All nodes communicate with each other in a single hop manner.

3.2 Simulation parameter setting

In this experiment, simulation analysis is carried out on MATLAB, and the performance of packet transmission success rate is compared with that of S-Aloha protocol. Monte Carlo method is adopted, and the nodes are distributed in a square area of 1km×1km, the transmission speed of underwater communication signal is 1.5km/s, the data packet size sent by each node is 512 bits, the number of slots per frame is four, the length of slots is 0.8s, and the learning rate in Q-Learning is α=0.1.

3.3 Simulation results

In the fixed-length slot frame of this experiment, when the number of nodes N is less than the number of slots, that is, the packet transmission rate is low, the collision probability of adopting QS-Aloha protocol is very small, so the packet transmission success rate will increase. In about 20 frames, the system converges to a steady state, and the success rate reaches 1, as shown in figure 4. However, in this case, the channel utilization rate decreases, so when N is greater than the number of slots, the probability of successful data packet transmission decreases with the increase of the number of nodes, because in a fixed length of time, the increased nodes need a period of time to learn before they can access the slots effectively, but this situation has less impact on the protocol than the collision avoidance of the whole network.

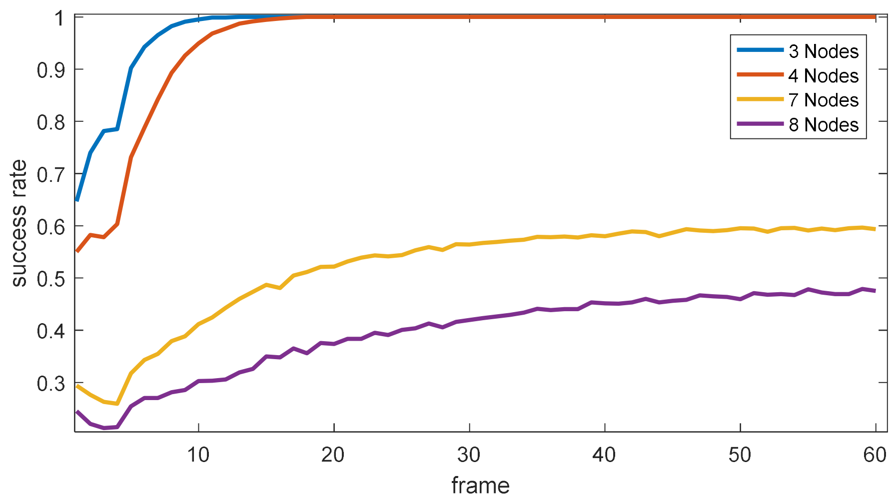

Fig.4 Data packet transmission success rate of different sensor nodes

Fig. 5 shows the comparison between slotted Aloha protocol and protocol with Q-Learning algorithm in terms of packet transmission success rate. When N=5, the packet success rate of QS-Aloha protocol is about 43.7% higher than that of S-Aloha, when N=6, the success rate is about 35.8% higher, and when N=7, the success rate is about 28.6% higher. Because in the single-hop network, the optimized slot scheduling can be realized by the Q-Learning algorithm, the only overhead required is the acknowledgement (ACK) packet. Compared with the traditional protocol, when the number of successfully transmitted bits increases, the overall throughput of the improved protocol also increases, so the system performance is improved.

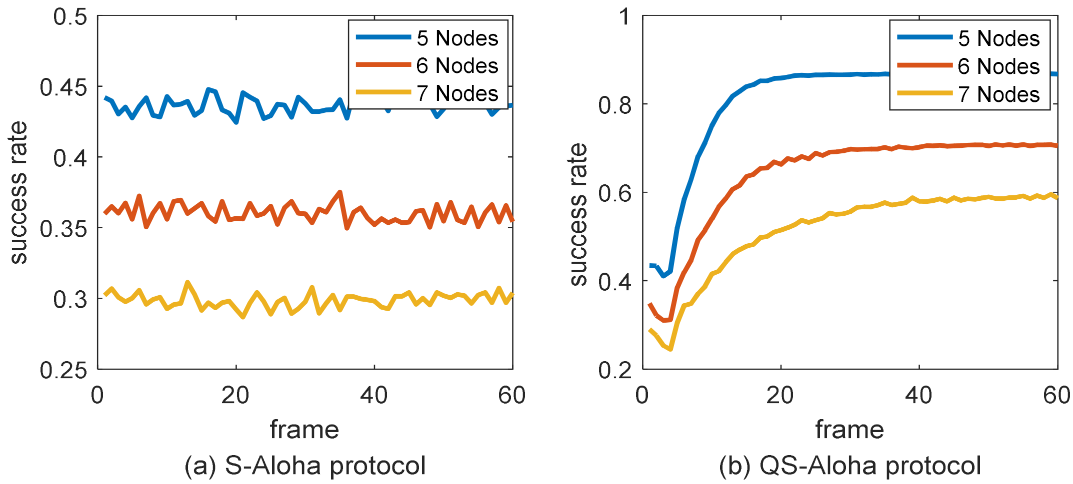

Fig.5 The success rate of sending data packets by s-aloha protocol and QS-Aloha protocol

Q-Learning algorithm avoids the situation that multiple nodes compete for one slot to a certain extent. Now, the system performance is evaluated from another angle. The proportional fair scheduling algorithm is adopted for nodes in figure 6. For S-Aloha protocol, after adding PF algorithm, the system calls a better node. From the 40th frame, the success rate of packet transmission increases by about 21.35%. For QS-Aloha protocol, the success rate of packet transmission increased by about 24.51% after adding PF algorithm. It also proves that this strategy can improve system performance and fairness of node scheduling.

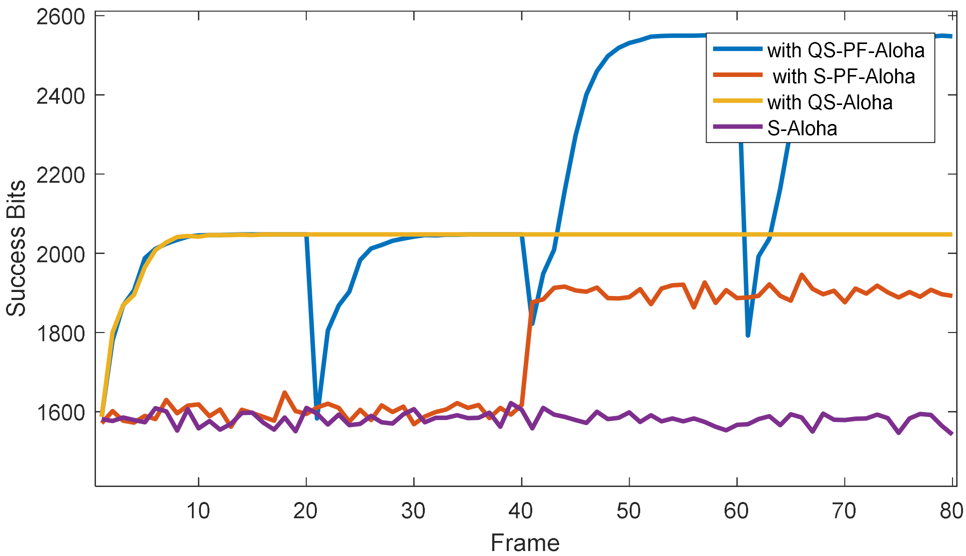

Fig.6 Comparison diagram of data packet bits with proportional fair scheduling algorithm

4. Summary

In this paper, the underwater acoustic MAC protocol is considered in an intelligent way, and the reinforcement learning strategy is incorporated into the slotted Aloha protocol. For this single-hop network, the Q-Learning algorithm minimizes the collision between packets with the minimum overhead in steady state, and improves the channel utilization rate. At the same time, the proportional fair scheduling algorithm is used to schedule and transmit nodes, and the effectiveness of the design is verified by simulation.

In underwater environment, the system efficiency depends on the number of slots per frame, which is difficult to determine in unpredictable environment. Therefore, how to make the sensor nodes achieve better system performance by evaluating the channel contention degree and dynamically adjusting the frame size according to their own needs remains to be further studied.

References

[1] Z.W. Guo, H.J. Luo, F. Hong, et al. Research progress of underwater wireless sensor networks[J]. Computer Research and Development, 2010, 47(1):377-389.

[2] Shahabudeen S, Motani M, Chitre M. Analysis of a High-Performance MAC Protocol for Underwater Acoustic Networks[J]. IEEE Journal of Oceanic Engineering, 2013.

[3] NOH Y, SHIN S. Survey on MAC protocols in underwater acoustic sensor networks[C].

[4] Communications and Information Technologies (ISCIT), 2014 14th International Symposium on IEEE, 2015:80-84.

[5] I. Demirkol, C. Ersoy and F. Alagoz, "MAC protocols for wireless sensor networks: A survey", Communications Magazine, vol.44, pp. 115-121, 2006.

[6] Chirdchoo N, Soh W S, Chua K C. Aloha-Based MAC Protocols with Collision Avoidance for Underwater Acoustic Networks[J]. Proceedings - IEEE INFOCOM, 2007:2271-2275.

[7] Roberts, Lawrence G. ALOHA packet system with and without slots and capture[J]. Acm Sigcomm Computer Communication Review, 1975, 5(2):28-42.

[8] Syed, A., Ye, W., Heidemann, J.S., & Krishnamachari, B. (2007). Understanding spatio-temporal uncertainty in medium access with ALOHA protocols. WuWNet '07.

[9] Zhou Y, Chen K, He J, et al. Enhanced Slotted Aloha Protocols for Underwater Sensor Networks with Large Propagation Delay[C]//Proceedings of the 73rd IEEE Vehicular Technology Conference, VTC Spring 2011, 15-18 May 2011, Budapest, Hungary. IEEE, 2011.

[10] Johnson J D, Li J, Chen Z. Reinforcement Learning: An Introduction: R.S. Sutton, A.G. Barto, MIT Press, Cambridge, MA 1998, 322 pp. ISBN 0-262-19398-1[J]. 2000, 35(1-4):205-206.

[11] J.Y. Wei, X.Y. Rui. Improved Proportional Fair Scheduling for Two-Hop Relay Networks [J]. Communications Technology, 2018, 51(4):829-834.