The derivations of the optimal network throughputs in various scenarios for the paper:

Deep Reinforcement Learning Based MAC Protocol for Underwater Acoustic Networks

Xiaowen Ye, Yiding Yu, Liqun Fu

I. Detailed Derivation

This is a supplementary document to the paper: Deep Reinforcement Learning Based MAC Protocol for Underwater Acoustic Networks. In this document, we give the optimal network throughput when the DR-DLMA node (see our paper for definition) coexists with nodes using other protocols, then we use the optimal network throughput as the benchmark for our paper.

We consider a underwater acoustic network (UWAN) consisting of N nodes and a buoy as an access point (AP) in a 3D area (see our paper). The nodes transmit on a shared uplink channel. Different nodes in the network may adopt different transmission strategies. Specifically, we assume that at least one node uses the DR-DLMA protocol, and the other nodes use TDMA or ALOHA protocol. The aim of the DR-DLMA node is to make full use of the available time slots that are not used by other nodes or resulted from propagation delays, so as to maximize the throughput of the overall UWAN.

In order to get the optimal throughputs in various scenarios, we use the model-aware node that knows the propagation delays and the transmission strategies of other nodes to replace the DR-DLMA node.

A. Coexistence with TDMA Networks

We first consider the coexistence of one model-aware node and one TDMA node. Let D_1 and D_2 denote the propagation delay from the model-aware node and the TDMA node to AP, respectively. Suppose that the TDMA node transmits a data packet in time slot t, which will reach AP in time slot $t+D_2$. If the model-aware node transmits a data packet in time slot $t+D_2-D_1$, a collision will occur at AP. In order to maximize the total network throughput without collisions, the optimal transmission policy of the model-aware node is to transmit in all time slots except time slot $t+D_2-D_1$.

Furthermore, when multiple model-ware nodes coexist with multiple TDMA nodes, we assume that model-aware nodes are aware of each other and can cooperate to fully utilize the available time slots that are not used by TDMA nodes. Then the optimal total network throughput is also 1.

Xiaowen Ye and Liqun Fu are with the School of Informatics, and the Key laboratory of Underwater Acoustic Communication and Marine Information Technology Ministry of Education, Xiamen University, Xiamen, China. Email:{xiaowen@stu.xmu.edu.cn, liqun@xmu.edu.cn}

Yiding Yu is with the Department of Information Engineering, The Chinese University of Hong Kong, Shatin, Hong Kong Special Administrative Region, China. Email:yy016@ie.cuhk.edu.hk

B. Coexistence with ALOHA Networks

We first consider the coexistence of one model-aware node and one ALOHA node. Let D_1 and D_2 denote the propagation delay from the model-aware node and the ALOHA node to AP, respectively. For this coexistence scenario, in each time slot, the model-aware nodes determine whether to transmit or not according to the transmission probability of the ALOHA node. We assume that the ALOHA node transmits a data packet with probability q in time slot t and the transmission probability of the model-aware node is b in time slot $t+D_2-D_1$, then the total network throughput in time slot t can be calculated as follows:

$$f(b) = b(1-q)+(1-b)q,$$

Thus

$$\frac{\mathrm{d}f(b)}{\mathrm{d}b} = (1-q)-q = 1-2q,$$

$$\frac{\mathrm{d}^2 f(b)}{\mathrm{d}b^2} = 0,$$

indicating that $f(b)$ is convex in b. When $\mathrm{d}f(b)/\mathrm{d}b<0$, i.e., $q>0.5$, if $b=0$, $f(b)$ can get the maximum value; when $\mathrm{d}f(b)/\mathrm{d}b \geq 0$, i.e., $q\leq 0.5$, if $b=1$, $f(b)$ can get the maximum value. As a result, when the ALOHA node transmits in time slot t with probability q, the optimal transmission policy of the model-aware node is as follows:, if $q>0.5$, the model-aware node does not transmit in time slot $t+D_2-D_1$; otherwise, the model-aware node transmits in time slot $t+D_2-D_1$, i.e.,

$$b^* = \begin{cases} 0, & \text{if } q>0.5, \\ 1, & \text{otherwise.} \end{cases}$$

Then the optimal network throughput is

$$\begin{cases} q, & \text{if } q>0.5, \\ 1-q, & \text{otherwise.} \end{cases} \tag{1}$$

Furthermore, we now consider the coexistence of multiple model-aware nodes and N ALOHA nodes in the network. To derive the optimal results achieved by model-aware nodes, we assume that multiple model-aware nodes are aware of each other and can cooperate with each other in each time slot. For convenience of illustration, we regard the model-aware nodes as "a big model-aware node". We assume that the propagation delays from the ALOHA node i ($i=1,2,\cdots,N$) and the big model-aware node to AP are D_i and D_{N+1}, respectively.

Let q_i $(i = 1, 2, \cdots, N)$ denote the transmission probability of ALOHA node i in time slot $t - D_i$, and b denote the transmission probability of the big model-aware node in time slot $t - D_{N+1}$. Then the total network throughput in time slot t can be calculated as follows:

$$f(b) \ = b\prod_{i=1}^{N}(1-q_i) + (1-b)\sum_{i=1}^{N}\big(q_i \prod_{j=1_{j\neq i}}^{N}(1-q_j)\big),$$

Thus

$$\frac{\mathrm{d}f(b)}{\mathrm{d}b} = \prod_{i=1}^{N}(1-q_i) - \sum_{i=1}^{N}\big(q_i \prod_{j=1_{j\neq i}}^{N}(1-q_j)\big),$$

$$\frac{\mathrm{d}^2 f(b)}{\mathrm{d}b^2} = 0,$$

indicating that $f(b)$ is convex in b. When $\mathrm{d}f(b)/\mathrm{d}b < 0$, if $b = 0$, $f(b)$ can get the maximum value; when $\mathrm{d}f(b)/\mathrm{d}b \geq 0$, if $b = 1$, $f(b)$ can get the maximum value. As a result, the optimal transmission policy of the big model-aware node is as follows: the big model-aware node does not transmit in time slot $t - D_{N+1}$ when $\mathrm{d}f(b)/\mathrm{d}b < 0$; the big model-aware node transmits in time slot $t - D_{N+1}$ when $\mathrm{d}f(b)/\mathrm{d}b \geq 0$, i.e.,

$$b^* = \begin{cases} 0, & \text{if } \mathrm{d}f(b)/\mathrm{d}b < 0, \\ 1, & \text{otherwise.} \end{cases}$$

Then the optimal network throughput is

$$\begin{cases} \sum_{i=1}^{N}\big(q_i \prod_{j=1_{j\neq i}}^{N}(1-q_j)\big), & \text{if } \mathrm{d}f(b)/\mathrm{d}b < 0, \\ \prod_{i=1}^{N}(1-q_i), & \text{otherwise.} \end{cases} \tag{2}$$

C. Coexistence with TDMA and ALOHA Networks

We first consider the coexistence of one model-aware node, one TDMA node and N ALOHA nodes. Let D_i $(i = 1, 2, \cdots, N)$, D_{N+1}, and D_{N+2} denote the propagation delay from the ALOHA node i, the TDMA node, and the model-aware node to AP, respectively, and let q_i denote the transmission probability of ALOHA node i $(i = 1, 2, \cdots, N)$ in time slot $t - D_i$. Suppose that the TDMA node transmits a packet in time slot $t - D_{N+1}$, which will reach AP in time slot t. If the model-aware node transmits a data packet in time slot $t - D_{N+2}$, collisions will occur at AP. Therefore, in the time slot $t - D_{N+2}$, the model-aware node refrains from transmission, then the network throughput in time slot t is

$$\prod_{i=1}^{N}(1-q_i).$$

In the time slots except time slot $t - D_{N+2}$, the model-aware node decides whether to transmit or not according to the value of N and q_i (i.e., the model-aware nodes will coexist with ALOHA nodes). Specifically, let b denote the transmission probability of the model-aware node in each time slot except time slot $t - D_{N+2}$, then the network throughput in a certain time slot except time slot t is

$$b\prod_{i=1}^{N}(1-q_i) + (1-b)\sum_{i=1}^{N}\big(q_i \prod_{j=1_{j\neq i}}^{N}(1-q_j)\big).$$

Let p denote the ratio of the number of time slots used by the TDMA node in a frame, then the average network throughput in each time slot can be calculated as follows:

$$\begin{aligned} F(b) \ =& p\prod_{i=1}^{N}(1-q_i) + (1-p)\Big(b\prod_{i=1}^{N}(1-q_i) + \\ & (1-b)\sum_{i=1}^{N}\big(q_i \prod_{j=1_{j\neq i}}^{N}(1-q_j)\big)\Big). \end{aligned}$$

Thus,

$$\frac{\mathrm{d}F(b)}{\mathrm{d}b} = (1-p)\Big(\prod_{i=1}^{N}(1-q_i) - \sum_{i=1}^{N}\big(q_i \prod_{j=1_{j\neq i}}^{N}(1-q_j)\big)\Big),$$

$$\frac{\mathrm{d}^2 F(b)}{\mathrm{d}b^2} = 0,$$

indicating that $F(b)$ is convex in b. When $\mathrm{d}F(b)/\mathrm{d}b < 0$, if $b = 0$, $F(b)$ has the maximum value:

$$p\prod_{i=1}^{N}(1-q_i) + (1-p)\sum_{i=1}^{N}\big(q_i \prod_{j=1_{j\neq i}}^{N}(1-q_j)\big).$$

When $\mathrm{d}F(b)/\mathrm{d}b \geq 0$, if $b = 1$, $F(b)$ has the maximum value:

$$\prod_{i=1}^{N}(1-q_i).$$

As a result, if the TDMA node transmits in time slot $t - D_{N+1}$, then the optimal transmission policy of the model-aware node is as follows: in time slot $t - D_{N+2}$, the model-aware node refrains from transmission. In the time slots except time slot $t - D_{N+2}$, if $\mathrm{d}F(b)/\mathrm{d}b \geq 0$, the model-aware node transmits; otherwise, it does not transmit. Let $z = \mathrm{d}F(b)/\mathrm{d}b$, then the optimal network throughput can be summarized as follows:

$$\begin{cases} p\prod_{i=1}^{N}(1-q_i) + (1-p)\sum_{i=1}^{N}\big(q_i \prod_{j=1_{j\neq i}}^{N}(1-q_j)\big), & z<0, \\ \prod_{i=1}^{N}(1-q_i), & z \geq 0. \end{cases} \tag{3}$$

Furthermore, when multiple model-aware nodes coexist with M TDMA nodes and N ALOHA nodes, we assume that model-aware nodes are aware of each other and can cooperate with each other. For convenience of illustration, we regard multiple model-aware nodes as “a big model-aware node”. Let D_i, D_j, and D_{N+M+1} denote the propagation delays from ALOHA node i $(i = 1, 2, \cdots, N)$, TDMA node j $(j = N+1, N+2, \cdots, N+M)$, and the big model-aware node, respectively, and q_i denote the transmission probability of ALOHA node i in time slot $t - D_i$. Suppose that TDMA node j transmits a data packet in time slot $t - D_j$, then the optimal transmission policy for the big model-aware node is as follows: in the time slot $t - D_j$, the big model-aware node refrain from transmission. In the time slots except time slot $t - D_j$, if $z<0$, the big model-aware node does not transmit in time slot $t - D_{N+M+1}$; otherwise, the big model-aware node will cooperate to transmit in time slot $t - D_{N+M+1}$. Let p denote the total ratio of the number of time slots used by TDMA nodes in a frame, then the optimal average network throughput is the same as (3).