Jamming-resilient algorithm for underwater cognitive acoustic networks

Zixiang Wang[1,2], Fan Zhen[1], Senlin Zhang[1], Meiqin Liu[1] and Qunfei Zhang[3]

Abstract
Due to the limit spectrum resource in the underwater acoustic networks, underwater cognitive acoustic communication is a promising technique. The channel sharing mechanism in cognitive networks can improve the communication capacity efficiently. Jamming attack is a common deny of service attack in cognitive networks. In the underwater cognitive acoustic networks, the anti-jamming problem is quite different from cognitive radio networks. It calls for an effective anti-jamming strategy in the cognitive acoustic channel access. In this article, we propose an online learning anti-jamming algorithm called multi-armed bandit–based acoustic channel access algorithm to achieve the jamming-resilient cognitive acoustic communication. The imperfect channel sensing and the constraints of underwater acoustic communication are considered in the anti-jamming game. Under different kinds of jamming attacks, the channel utilization can be improved with our jamming-resilient approach.

Keywords
Underwater cognitive acoustic network, anti-jamming, hidden Markov model, multi-armed bandit, channel access

Introduction

During the past decade, underwater wireless sensor networks (UWSNs) have attracted significant interests due to a variety of applications such as ocean environment monitoring, underwater target tracking, oceanography data collection. Underwater sensor nodes need to transmit the collected data in the underwater environment. Because the electronic wave decays quickly in the water, acoustic communication is the mainstream communication method in the UWSNs at current time. It calls for reliable and high data rate data transmission over an underwater acoustic channel. The attenuation of underwater acoustic signal is related to the frequency. In the underwater acoustic networks, the available communication frequencies are limited in a small range, which usually range from tens of hertz to hundreds of kilohertz.[1] Today, most underwater acoustic systems utilize the frequency band from 1 to 100 kHz, making the acoustic channel crowded.[2–4] In the scenario that multiple underwater systems such as autonomous underwater vehicles (AUVs), underwater fleet, and UWSNs are deployed in the same area, the scarce acoustic spectrum should be shared among them. The limited underwater acoustic spectrum resource points out the requirement for efficient spectrum utilization

[1]College of Electrical Engineering, Zhejiang University, Hangzhou, China
[2]State Grid Zhejiang Electric Power Research Institute, Hangzhou, China
[3]School of Marine Science and Technology, Northwestern Polytechnical University, Xi'an, China

Corresponding author:
Meiqin Liu, College of Electrical Engineering, Zhejiang University, Hangzhou 310027, China.
Email: liumeiqin@zju.edu.cn

schemes. Underwater cognitive acoustic network (UCAN) is an emerging technique for the underwater acoustic networks.[5] The idea of cognitive acoustic is brought from the cognitive radio networks (CRNs).[6,7] In cognitive networks, there are primary users (PUs) and secondary users (SUs). SUs are permitted to sense which portion of the PUs' spectrum is available and select the best available channel to access. Since there exists some states when the PUs are idle, some frequency bands might be vacant for considerable periods. SUs can utilize the resource of PUs at their vacant periods. By this way, the communication capacity can be increased. In the works,[8,9] it is verified that UCAN can improve acoustic communication capacity. The UCAN will be a promising technique in underwater acoustic communications.

Although some proposed approaches[8–10] have been shown to be able to bring gains for the users, the attackers are not considered in the channel allocation process. If the SUs are deployed in a hostile environment, there exist malicious attackers who try to prevent the efficient spectrum utilization of legitimate SUs. An attacker can interfere or jam the acoustic signal at the receiver by simply emitting a random signal in the water.[11] The interfered or jammed signal becomes unidentifiable at the receiver. Therefore, the channel allocation schemes which aim at maximizing the spectrum utilization may not perform well due to the jamming caused by attackers. Hence, it calls for an anti-jamming channel allocation scheme for UCAN to efficiently utilize the spectrum. Jamming attack is a denial-of-service (DoS) attack. There have been some traditional anti-jamming techniques such as direct-sequence spread spectrum (DSSS) and frequency-hopping spread spectrum (FHSS).[12] But these solutions are not suitable for cognitive networks. SUs can only access the free spectrum of PUs, and the availability of the spectrum is time varying. So the traditional spread spectrum communication techniques cannot be applied in cognitive networks.

Today, most of jamming-resilient methods for CRNs employ opportunistic channel sensing and utilization techniques.[13–20] In the works,[14–19] the opportunistic channel access strategy is designed through online learning algorithms. These algorithms need to know the all channels' status and perfect channel sensing is assumed. The channel access strategy is made after the channel sensing. However, this requirement is difficult to be satisfied in the underwater acoustic channel sensing. Because of the long propagation delay, it takes a long time to sense all the channels. It is not efficient to sense all the channels in UCAN. The perfect channel sensing is also not practical in underwater acoustic communication. In Zhao et al.[21] and Ahmad et al.,[22] SUs determine a sensing strategy and access the channels that are sensed idle. But the jamming attacks

are not considered in Zhao et al.[21] and Ahmad et al.[22] The authors in Wang et al.[13] modeled the anti-jamming problem in cognitive networks as a multi-armed bandit (MAB) problem. The opportunistic channel sensing strategy is made through solving the MAB problem. But the false channel state detection, which is quite common in UCANs, is not considered seriously. The existing anti-jamming approaches for CRNs are not suitable in the underwater environment. It calls for a novel jamming-resilient approach for UCANs.

A multi-armed bandit–based acoustic channel access (MAB-ACA) algorithm is proposed in this article. The algorithm consists of two phases. In the first phase, we build the hidden Markov model (HMM) model of each channel, and the potential reward of each channel can be estimated based on the HMM model. In the second phase, SUs make the channel sensing and access strategy through online learning. Each underwater acoustic channel will be sensed with a given probability. For the channels prone to be jammed, small probability will be assigned. The probability is determined through online learning based on the current reward. Although this approach shows some similarities with the anti-jamming algorithms in the CRNs, the properties of UCANs are considered in our approach. The major contributions of this article can be concluded as follows:

- We build the model of cognitive acoustic channel. Considering the channel state transition and imperfect channel sensing in the underwater environment, an HMM is employed to describe the underwater cognitive acoustic channel.
- We combine the Baum–Welch algorithm and genetic algorithm in the HMM estimation process, which solves the sensitive dependence on initial values of the traditional Baum–Welch algorithm.
- The posterior probability of a cognitive channel is deduced based on the HMM model.
- We formulate the anti-jamming problem in UCANs as an MAB problem. A heuristic algorithm to solve such MAB problem is proposed. The algorithm can improve the channel utilization while considering the latent underwater jammers.

The article is organized as follows: in section "Related works," we briefly describe the related works; in section "System and adversary model," we introduce the system and adversary model; in section "Jamming-resilient opportunistic acoustic channel access," a jamming-resilient opportunistic acoustic channel access is proposed; in section "MAB-based acoustic channel access algorithm," a heuristic approach to solve the jamming-resilient acoustic channel access problem is proposed; the simulation results are shown in section

"Performance evaluation and discussions"; in section "Conclusion," we conclude the article.

Related works

The UCAN provides a solution to improve the spectrum utilization for the underwater acoustic networks. The authors in Baldo et al.[8] and Bicen et al.[9] proposed cognitive spectrum access approaches for underwater acoustic communications, and the communication efficiency can be improved with the opportunistic spectrum access protocol. However, the works[8,9] do not consider the possible jammers in the underwater environment.

To defense the jamming attack, anti-jamming mechanisms in wireless communication have been extensively studied.[13–20,23] In Wood and Stankovic,[23] FHSS and DSSS techniques are used in the underwater acoustic communications. These techniques are resistant to interference from jammers, but they are not directly suitable for UCANs. First, such schemes do not consider the influence of PUs, who are important participants in cognitive networks. Second, such schemes require a wide range of frequencies, which are not available in UCANs. In the works,[13,16,17,20] the anti-jamming problem in cognitive networks is described as an MAB problem, and heuristic algorithms are proposed to solve the MAB problem. In Wang et al.,[13] authors formulated the anti-jamming channel access problem in cognitive networks as a non-stochastic MAB problem. Authors also proposed a heuristic algorithm to solve the MAB problem. In Wang et al.,[20] the same group of authors improved the theoretical performance analysis and experiments for the MAB-based channel access algorithm. Su and colleagues[16,17] applied the MAB-based anti-jamming approach in the smart gird scenario. In the works,[14,15,18,19] authors modeled the anti-jamming problem as a stochastic game. In Wang and colleagues,[14,15] authors used the stochastic zero-sum game to describe the anti-jamming scenario in cognitive networks. Minimax-Q learning is employed to deal with anti-jamming game. In Singh and Trivedi,[18] authors replaced the minimax-Q learning with QV learning and state-action-reward-state-action learning for the anti-jamming stochastic game. In Yang et al.,[19] authors studied the power control problem in wireless networks. They modeled the jamming defense problem as a Stackelberg game and derived the optimal transmission power.

Although the existing works can deal with the jamming attacks in cognitive networks, we cannot directly use these methods in the UCANs. The MAB-ACA algorithm, which is proposed in this article, is designed for the UCANs. The characteristics of underwater acoustic channel are considered in our approach. We use HMMs, instead of Markov chains, to describe the channel state transition. The channel state can be estimated through HMM. We formulate MAB problem for the jamming-resilient acoustic channel access game, and we solve the MAB problem based on the Exp3 algorithm.[24]

System and adversary model

In this section, we introduce the model of UCAN and the possible adversaries existing in the underwater system. Some licensed channels of PUs may be available for SUs, and SUs utilize these channels to increase communication capacity. The jammers, on the other hand, jam the idle channels of PUs to block the communication of SUs.

System model

In UCANs, PUs may share their spectrum with SUs. The available acoustic spectrum of PUs will be partitioned into channels with equal bandwidth, and an underwater SU will be assigned to an idle channel. Although SUs may have their own spectrum, we focus on the utilization of PUs' spectrum in UCANs. SUs are permitted to access the channels only when not interfering with PUs, and their spectrum utilization is not legally protected. Assume PUs own a spectrum consisting of n channels, each with bandwidth D_i ($i = 1, 2,..., n$). The UCAN system proceeds in discrete time slots. The occupancy or vacation of these n channels follows a discrete Markov process. A channel can be either busy or idle in each time slot, and the state of the channel can be described as a two-state Markov Chain as shown in Figure 1(a). We denote the busy state by "0" and the idle state by "1." The state of the
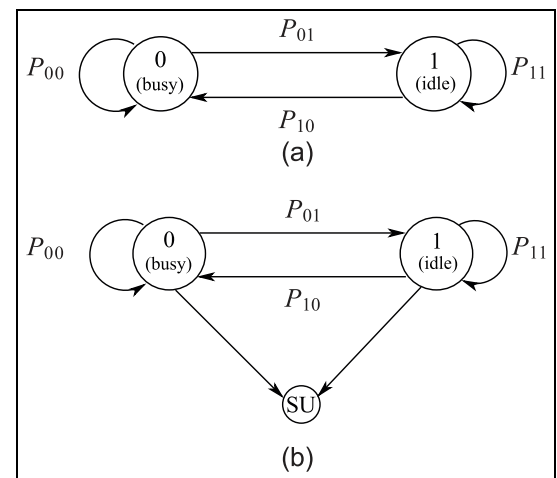

Figure 1. Markov model of channel state: (a) channel state transition and (b) HMM of cognitive acoustic channel.

kth channel in each slot is given by s_k where $s_k \in \{0, 1\}$. The channel may transit from busy state to idle state with the probability P_{01}, stay in busy state with the probability $P_{00} = 1 - P_{01}$, transit from idle state to busy state with the probability P_{10}, and stay in idle state with the probability $P_{11} = 1 - P_{10}$.

The SUs seek spectrum opportunities in PUs' N channels. They will sense availability of PUs' channels at first and access the idle channels to transmit data without interfering the PUs. If the channel is sensed busy, the SU will choose another channels to sense until an available is discovered. Once the data are received successfully at the receiver side, an Ack will be sent to the sender. In underwater environment, ad hoc networks are more appropriate. Therefore, we focus on SU networks without a central controller for coordinating the SUs. Each SU will sense and access the channels independently.

The spectrum sensing can be performed in time, frequency, and code domains.[10,25–28] The spectrum sensing is out of range of this article, but we should note that the false alarm and miss detection are common in UCANs. It is mainly caused by the imperfect acoustic spectrum sensing on SUs. Moreover, the long propagation delay may also lead to false sensing results. Since the signal of PUs suffers long propagation delay, the sensing results at SU side are always out-of-date. The delayed channel usage information increases the probability of false sensing. Although the channel sensing results are not reliable, they can still reflect the actual channel status. Such channel state transition and channel sensing process can be modeled as an HMM as shown in Figure 1(b). The channel state transition is a hidden Markov chain, whose states can be observed by SUs. The observed state of the kth channel is denoted by $O_k \in \{0, 1\}$. Thus, the false alarm probability and miss detection probability can be described as $P(O_k = 0 | s_k = 1)$ and $P(O_k = 1 | s_k = 0)$, respectively.

Different to CRNs, underwater nodes are usually equipped with single modem. An underwater acoustic SU can sense and access one channel each slot. The job of an SU is to choose which channel to sense and access the available channel based on the sensing result.

Adversary model

Jamming is a DoS attack which targets at disrupting communications at the physical and link layers. The jammers will sense the activity of the PUs and jam the idle channels. By injecting packets to the licensed channels, a jammer can prevent the legitimate users from accessing the available spectrum of PUs. In UCANs, PUs are protected by law and usually well physically protected. Due to the heavy penalty of being located

and captured, the jammers will not jam the PUs' communications.[7,13,17] Instead, the SUs' access to the PUs' spectrum is not legally protected. A jammer can attack the SUs and prevent them from using the spectrum for communication.

Similar to legitimate users in UCANs, jammers are also equipped with single modem. So one jammer can only jam one channel each slot. In a UCAN, multiple jammers may be deployed to disrupt the communications. Due to the limitation of cost, we assume $N_J (N_J < n)$ jammers exist in the network.

The jammers will adopt an attacking strategy to prevent SUs from accessing the idle channels. In this article, a strategy means a channel sensing and access decision. Both SUs and jammers should choose their strategies in UCANs. A jammer or an SU will choose a channel to sense and access the idle channel. Such channel sensing and access action is called a strategy. The jammers can be classified into four types:[13]

> *Static jammer*. A static jammer always senses the same channel of PUs and jams it if it is idle.
> *Random jammer*. A random jammer will randomly jam one of the idle channels of PUs.
> *Myopic jammer*. A myopic jammer is an intelligent jammer that runs the myopic algorithm.[21,22] It senses the occupancy of channels and makes the optimal jamming strategy that costs most damage.
> *Adaptive jammer*. Through listening the Ack packet in the Ack transmission interval, a jammer can know whether its attack is successful on the channels it jams. An adaptive jammer senses the parts of channels and jams idle channels based on its jamming history.

Under the deterministic channel sensing and access policy, an SU senses the channels in a fixed order. Once a channel is sensed idle, the SU will access the channel. Such policy is vulnerable to jamming attack, especially for adaptive jamming. The SUs need to choose an accessing strategy to alleviate the potential damage. By hopping among idle channels, the access pattern of SUs is not easy to be predicted, and the UCAN will be more resistant to the jamming attack.

Channel model

The attenuation of underwater acoustic channel can be modeled as a function of distance and frequency[1]

$$L(d, f) = d^l a^d(f) \tag{1}$$

where f is frequency, d is the distance between sender and receiver, l is the path loss coefficient, and $a(f)$ is the absorption coefficient. $a(f)$ can be approximated empirically using Thorp's formula[29]

$$\log a(f) = 0.011 \frac{f^2}{1 + f^2} + 4.4 \frac{f^2}{4100 + f^2} \quad (2)$$
$$+ 2.75 \times 10^{-5} f^2 + 0.0003$$

In the underwater environment, there exist multiple kinds of noise. The ambient noise major consists of four components: turbulence, shipping, waves, and thermal noise. The power spectrum density (p.s.d.) of these noise components can be expressed as[1]

$$10 \log N_t(f) = 17 - 30 \log(f) \quad (3)$$

$$10 \log N_s(f) = 40 + 20(v_s - 0.5)$$
$$+ 26 \log(f) - 60 \log(f + 0.03) \quad (4)$$

$$10 \log N_w(f) = 50 + 7.5\sqrt{v_w}$$
$$+ 20 \log(f) - 40 \log(f + 0.4) \quad (5)$$

$$10 \log N_{th}(f) = -15 + 20 \log(f) \quad (6)$$

where v_s is shipping factor and v_w is wind speed. Different kinds of noise influence different bands. For example, turbulence noise has impact only at very low frequency region, $f < 10$ Hz.

We denote the p.s.d. of transmitted signal for the distance d by $S_d(f)$. The channel capacity of a tone transmitted at the kth channel is

$$C = \int_{D_k} \log_2 \left(1 + \frac{S_d(f)}{L(d,f)N(f)}\right) df \quad (7)$$

where $N(f)$ is the p.s.d. of ambient noise and D_k is the bandwidth of the kth channel. According to equation (7), the channel capacity is related to distance and frequency. In UCANs, channel capacity will be increased if SUs can properly utilize the channels.

Jamming-resilient opportunistic acoustic channel access

Channel state transition

At the beginning of a time slot, some channels are reserved by PUs for their communication in the current time slot. Then SUs can temporarily access the idle channels that belong to PUs. A certain gain can be achieved by an SU through utilizing the channels of PUs. The gain is a proper quality of service (QoS) measure in underwater acoustic networks. In this article, we use channel capacity to evaluate the gain achieved. The gain of channel k at time t is denoted by $g_k(t)$. The underwater nodes are located at different positions, and the corresponding sender/receiver pairs will have different communication distances. Since attenuations vary with distances and frequency, different channel

assignments will get different channel capacities. When a channel is occupied by the PUs or jammers, SUs cannot access the channel. Otherwise, the gain is

$$g_k = \int_{D_k} \log_2 \left(1 + \frac{S_d(f)}{L(d,f)N(f)}\right) df$$

The availability of a licensed channel is only determined by the PUs. In cognitive acoustic networks, PUs are licensed users and their rights are protected by law. Only when the channel is idle can SUs or jammers access the channel. The state of the channels at time t can be described as a triple $\{s_k^p(t), s_k^u(t), s_k^J(t)\}$. $s_k^p(t)$ means whether the kth channel is occupied by the PUs, $s_k^u(t)$ means whether the kth channel is accessed by the SUs, and $s_k^J(t)$ means whether the kth channel is jammed. One state will jump to the other state in the next time slot.

In each slot, SUs and jammers will select some channels to sense at first. After observing the state of the licensed channels, both SUs and jammers will access the available channels. SUs try to utilize the channel to send data, and jammers try to block the communication of SUs. We consider a vector space $\{0, 1\}^n$ and number the licensed channels from 1 to n. Mathematically, the actions of the SUs and jammers can be denoted by $a_k(t) \subseteq \{0, 1\}^n$ and $a_k^J(t) \subseteq \{0, 1\}^n$, respectively. $a_k(t) = 1$ means the SU will sense the kth channel at t. The size of SUs' action space is n, that is, an SU can choose one of the n channels to sense. Without jammers, the state of channels is determined by the PUs. However, the jammers will jam part of channels so that the state of channels will also be affected by jammers. The action of a jammer is correlated with the SUs' actions. Therefore, the state of channels in the next slot $s(t + 1)$ is determined by $s(t)$, $a(t)$, $a^J(t)$. The dynamics of the channel state can be described as a Markov process with the transition probability $P(s(t + 1)|s(t), a(t), a^J(t))$. Since the dynamics of the PUs' activity is independent of the SUs and jammers, the transition probability $P(s(t + 1)|s(t), a(t), a^J(t))$ can be further expressed as

$$P(s(t + 1)|s(t), a(t), a^J(t)) = P(s^p(t + 1)|s^p(t))$$
$$P(s^u(t + 1), s^J(t + 1)|s(t), a(t), a^J(t)) \quad (8)$$

$P(s^p(t + 1)|s^p(t))$ represents the transition probability of the PUs' status and $P(s(t + 1)|s(t), a(t), a^J(t))$ represents the transition probability of the licensed channels. When $s^p(t) = 0$, the band is occupied by the PUs. Both SUs and jammers will not access the licensed channels. The action of SUs will be $a(t) = \mathbf{0}$ and the action of jammers will be $a^J(t) = \mathbf{0}$. Here $\mathbf{0}$ means a zero vector. When $s^p(t) = 0$, we have

$$P(s(t+1)|s^p(t)=0, s(t), a(t), a^J(t))$$
$$= P(s^p(t+1)|s^p(t)=0) \qquad (9)$$

When $s^p(t)=1$, the band is available to the underwater SUs. The jammers will jam N_J channels under their jamming strategy.

Opportunistic channel access

In each slot, three parties make their own actions. Since an underwater SUs want to utilize the PUs' channels to transmit data, it needs to sense these channels and access them if available. In the channel sensing process, SUs cannot distinguish between PUs and jammers. As long as a channel is occupied, the SUs will skip it and choose another channel to sense. Current researches about anti-jamming CRN assume the perfect channel sensing,[7,15–18] but it cannot be achieved in the underwater acoustic communication. Due to the long propagation delay, false alarm and miss detection cannot be avoided in CANs. We denote the false alarm probability by p_f and miss detection probability by p_m. The conditional probability that channel k is idle is

$$P(s_k=1|O_k) = \begin{cases} 1-p_m, & O_k=1 \\ p_f, & O_k=0 \end{cases} \qquad (10)$$

where O_k is the channel sensing result. The false alarm probability and miss detection probability are equivalent to the observation symbol probability in HMM. Once a packet is received by an SU receiver, a reward can be obtained. For an SU, the potential reward on each channel is the joint effect of the PUs' activity, jammers' actions, and its own action. We denote this by

$$r_k = s_k^p(t)\bar{s}_k^J(t)s_k^u(t)g_k(t) \qquad (11)$$

where r_k is the reward of channel k, s_k^J and s_k^u are determined by sensing results, and $\bar{s}_k^J(t)$ is the logic not of $s_k^J(t)$. Thus, the total reward is $\sum_{k=1}^{N} r_k$. We may say the cognitive acoustic communication is efficient if greater reward can be achieved on each channel. Without jammers, the optimal opportunistic channel access is to maximize $\sum_{k=1}^{N} s_k^p(t)s_k^u(t)g_k(t)$. A jamming-resilient acoustic channel access problem can be described as how to maximize the mean reward while considering the jammers' actions, that is

$$\arg\max_{a(t-1)} E\left[\sum_{k=1}^{N} s_k^p(t)\bar{s}_k^J(t)s_k^u(t)g_k(t)\right] \qquad (12)$$

However, due to the existence of the intelligent jammers, the actions of an SU at t may affect the jamming results in the future. A good channel access strategy also needs to consider the potential reward in the future. We define the reward function as

$$V = E\left[\sum_{t=1}^{\infty} \gamma^t \sum_{i=1}^{N} s_k^p(t)\bar{s}_k^J(t)s_k^u(t)g_k(t)\right] \qquad (13)$$

where $0 < \gamma < 1$ is the discount factor. Thus, the jamming-resilient acoustic channel access problem can be further described as

$$\arg\max_{a(t-1)} E\left[\sum_{t=1}^{\infty} \gamma^t \sum_{k=1}^{N} s_k^p(t)\bar{s}_k^J(t)s_k^u(t)g_k(t)\right] \qquad (14)$$

The problem equation (14) is a stochastic game, and it is known that such a problem has a nonempty set of optimal policies, and at least one of them is stationary.[30] The problem (14) may be solved using its respective dynamic programming (DP) representations,[22] which can be described as

$$V_T = \max_{a(T-1)} E\left[\sum_{k=1}^{N} r_k(T)\right]$$
$$V_t = \max_{a(t-1)} E\left[\sum_{k=1}^{N} r_k(t) + \gamma V_{t+1}\right], t = 1, 2, \ldots, T-1$$
$$(15)$$

Here $r_k(t)$ is the immediate total reward of all licensed channels. It can be obtained through the Ack packets received by the senders.

Channel state estimation

While solving the DP problem (15), we should get the mean reward in each iteration. According to equation (11), three parameters $s_k^p(t)$, s_k^J, $g_k(t)$ should be known at first. Given the distance between sender and receiver, we can determine the gain of each channel for the sender/receiver pair. However, it is difficult to know $s_k^p(t)$ and $s_k^J(t)$. Since a channel may be occupied by PUs or jammers, the channel state $s(t)$ is the combined effect of PUs and possible jammers. We may use a two-state Markov Chain to describe the channel state. Because the cognitive acoustic channel sensing process suffers a long delay, the sensing result does not reflect the actual channel state. Therefore, we can just estimate the actual channel state based on the past sensing result. As mentioned in section "System and adversary model," the channel state transition and channel sensing process in UCANs can be modeled as an HMM. The channel state estimation problem is equivalent to a parameter estimation problem of HMM. We can use a triple $\lambda = (A, B, \pi)$ to describe an underwater cognitive acoustic channel. A is the state transition probability distribution, B is the observation symbol probability distribution, and π is the initial state distribution. If an SU senses the kth channel for T rounds, a sensing result

sequence $O(1), O(2),..., O(T)$ will be obtained. We hope to estimate λ based on $O(1), O(2),..., O(T)$.

The Baum–Welch algorithm[31] is a common method to estimate λ. The Baum–Welch algorithm can provide the maximum a posteriori (MAP) estimation of the HMM. We denote the state transition probability from $i \in \{0, 1\}$ to $j \in \{0, 1\}$ at time t by

$$\xi_{ij}(t) = \frac{P(s(t) = i, s(t + 1) = j, O|\lambda)}{P(O|\lambda)} \qquad (16)$$

$P(s(t) = i, s(t + 1) = j, O|\lambda)$ and $P(O|\lambda)$ can be calculated through forward-backward algorithm[32,33]

$$P(s(t) = i, s(t + 1) = j, O|\lambda) = \alpha_i(t)A_{ij}$$
$$P(s(t + 1) = j|O(t + 1))\beta_j(t + 1) \qquad (17)$$

$$P(O|\lambda) = \sum_{i \in \{0,1\}} \sum_{j \in \{0,1\}} \alpha_i(t)A_{ij}$$
$$P(s(t + 1) = j|O(t + 1))\beta_j(t + 1) \qquad (18)$$

where

$$\alpha_i(t) = P(O(1), O(2), ..., O(t), s(t) = i|\lambda) \qquad (19)$$
$$\beta_j(t) = P(O(t + 1), O(t + 2), ..., O(T)|\lambda) \qquad (20)$$

With the channel sensing sequence $O(1), O(2),..., O(T)$, $\sum_{t=1}^{T-1} \xi_{ij}(t)$ represents the transition times from i to j, and $\sum_{t=1}^{T-1} \sum_{j=\{0,1\}} \xi_{ij}(t)$ represents the transition times from i. Based on $\sum_{i \in \{0,1\}} \xi_{ij}(t)$, $\sum_{t=1}^{T-1} \xi_{ij}(t)$, $\sum_{t=1}^{T-1} \sum_{j \in \{0,1\}} \xi_{ij}(t)$, the parameters of the two-state HMM λ can be estimated

$$\bar{\pi} = \sum_{j \in \{0,1\}} \xi_{ij}(1) \qquad (21)$$

$$\bar{A}_{ij} = \frac{\sum_{t=1}^{T-1} \xi_{ij}(t)}{\sum_{t=1}^{T-1} \sum_{j \in \{0,1\}} \xi_{ij}(t)} \qquad (22)$$

$$\bar{B}_{mi} = \frac{\sum_{O=m} \xi_{ij}(t)}{\sum_{t=1}^{T} \sum_{j \in \{0,1\}} \xi_{ij}(t)} \qquad (23)$$

Theorem 1. Building the HMM of a channel has time complexity $O(40T^2 - 38T + 6)$ and space complexity $O(T + 8)$.

Proof. In the forward-backward algorithm,[32,33] the time complexity to calculate $P(s(t) = i, s(t + 1) = j, O|\lambda)$ is $O(2N^2(T - 1) + N)$, where N is the number of possible states. Since the channel state is either 0 (busy) or 1 (idle), $N = 2$ in our scenario. Therefore, calculation of

$P(s(t) = i, s(t + 1) = j, O|\lambda)$ has time complexity $O(8T - 6)$. The time complexity for equation (16) is $O(8T - 6)$. The time complexity to calculate equations (21)–(23) is $O(40T^2 - 38T + 6)$. Once equations (21)–(23) are solved, we can build the HMM of a channel. So the time complexity for building the HMM is $O(40T^2 - 38T + 6)$.

In the HMM build process, we needs to collect T samples. From equations (16)–(23), there are eight variables. So the space complexity is $O(T + 8)$.

However, the Baum–Welch algorithm cannot guarantee the global convergence. In the two-state HMM estimation process, there are a lot of local extremums. The final result of the Baum–Welch will always converge near the initial value. Without any priori knowledge, it is impossible to determine a proper initial value. To improve the convergence of HMM estimation, genetic algorithm can be employed for the global search.[34,35] In this article, we combine the Baum–Welch algorithm with genetic algorithm for the two-state HMM estimation. Genetic algorithm implements the global model matching at first, and the Baum–Welch algorithm will use the result of genetic algorithm as the initial value.

Before applying genetic algorithm to solve the optimization problem, we should find a proper encoding method. For the HMM of cognitive acoustic channel $\lambda = (A, B, \pi)$, there are five independent parameters $A_{11}, A_{21}, B_{11}, B_{21}, \pi$. Thus, the chromosomes can be encoded as a vector of real numbers: $c = (A_{11}, A_{21}, B_{11}, B_{21}, \pi)$. The selection, crossover, mutation process will be implemented iteratively. Since the Baum–Welch algorithm is a kind of MAP estimation, the fitness function of genetic algorithm is chosen as

$$F_i = P(O|c_k) / \sum_{i=1}^{M} P(O|c_i) \qquad (24)$$

where c_i is the ith chromosome and M is the population. Such fitness function can guarantee the consistence of HMM training. Chromosomes with higher fitness values are more likely to be reproduced. Then the arithmetic crossover will be applied between each couple of chromosomes

$$c_i' = \mu c_i + (1 - \mu)c_j \qquad (25)$$
$$c_j' = \mu c_j + (1 - \mu)c_i \qquad (26)$$

where $0 < \mu < 1$. Finally, a few chromosomes will alter some genes to produce new mutated chromosomes

$$c_i' = c_i + \delta(1 - c_i)(1 - t_c/t_m)^b \qquad (27)$$

where $0 < \delta < 1$ is a random number, t_c is the current generation number, t_m is the maximum generation

number, and $b>0$ is a constant. The result of genetic algorithm will be the initial value of the Baum–Welch algorithm.

Once the parameters of the HMM are estimated, the probability that the ith channel is idle can be represented as

$$P(s_i^p = 1) = \frac{\bar{A}_{12}}{1 - \bar{A}_{22} + \bar{A}_{12}} \tag{28}$$

The posterior probability with channel sensing result is

$$P(s_i(t)|O(t)) = \frac{P(s_i(t), O(t))}{P(O(t))} = \frac{B_{s_i(t), O(t)}P(s_i)}{P(O(t))} \tag{29}$$

MAB-based acoustic channel access algorithm

In this section, we will give the channel accessing strategy while considering the possible jammers. Recall the jamming-resilient acoustic channel access problem (14), finding the optimal solution needs to know the channel state and jammer's actions at first. The channel state can be estimated through the Baum–Welch algorithm, but the jammer's action remains unknown. For the static jammers and random jammers, we can further express the problem (14) as

$$\arg\max_{a(t-1)} E\left[\sum_{t=1}^{\infty} \gamma^t \sum_{i=1}^{N} s_i^p(t)s_i^u(t)g_i(t)\right] E\left[\bar{s}_i^J(t)\right] \tag{30}$$

because these jammers are independent to legitimate users' actions. The myopic jammers and adaptive jammers, on the other hand, will change their strategies based on legitimate users' actions. It is difficult for us to estimate the mean reward, even though the channel state is known. Therefore, the DP-based method (15) cannot be solved directly in jamming-resilient acoustic channel access problem.

An anti-jamming strategy for UCANs is to access idle channels to maximize the reward. In this article, we use an online learning approach to approximate the solution of (15). The algorithm solves an MAB problem to achieve jamming-resilient acoustic channel access. Each channel can be considered as an arm. Communication through each channel is associated with a sequence of rewards. The PUs' actions and jammers' actions will affect the reward. During the slot T, the reward of an SU under the strategy Ψ is

$$R^{\Psi} = \sum_{t=1}^{T} \sum_{i=1}^{n} r_i^{\Psi}(t) \tag{31}$$

where r_i^{Ψ} is the corresponding reward of strategy Ψ and R^{Ψ} is the total reward within T. Assume there exists an optimal strategy Ψ^*, the regret of current strategy Ψ is

$$\rho = R^{\Psi^*} - R^{\Psi} \tag{32}$$

where ρ is the regret. Regret ρ represents the performance gap between the solution of (15) and our MAB-ACA algorithm. Obviously, we hope to find a strategy Ψ that can achieve the minimum regret.

In MAB problem, each strategy will be assigned a probability. The SU will choose a strategy based on the probability. In an UCAN with n channels, there are n accessing strategies Ψ_1, Ψ_2,..., Ψ_n. Each strategy is associated with a probability P_{Ψ_i}, $i = 1, 2, ..., n$, and the probability is deduced from the strategy weight. The weight of a strategy in slot t is the product of all rewards of the selected channels. During the system running, the SU will dynamically adjust the strategy weight based on the rewards.

In the existing researches, the reward will be counted only when the Ack is received through the channel. SUs can learn the properties of PUs and jammers through the reward. With the knowledge of the jammers' actions, jamming-resilient approach can be designed for the jamming avoidance and channel capacity optimization. For the channel sensed, the conditional probability that the channel is idle can be described as $P(s_i^p(t + 1) = 1|O(t))$. And the corresponding reward is

$$r_i = P(s_i^p(t + 1) = 1|O(t)) \int_{D_i} \log_2\left(1 + \frac{S_d(f)}{L(d,f)N(f)}\right)df \tag{33}$$

$$r_i = (P(s_i^p(t) = 0|O(t))\bar{A}_{12} + P(s_i^p(t) = 1|O(t))\bar{A}_{22}) \int_{D_i} \log_2\left(1 + \frac{S_d(f)}{L(d,f)N(f)}\right)df \tag{34}$$

However, the state of channels that are not sensed is not evaluated due to lack of information. As a result, even though some channels are quite good for the SUs, they may not be covered by the algorithm during several iterations. Moreover, because of the imperfect spectrum sensing in the underwater acoustic communication, the sensing results are not so accurate. Therefore, direct using the Ack information to formulate the reward may not correctly reflect the channels' availability. The idea of virtual channel reward can be employed to ensure that each channel is sensed sufficiently. We formulate the virtual channel reward as

$$\hat{r}_i = \frac{r_i}{P_{\Psi_i}} \tag{35}$$

Algorithm 1. MAB-based acoustic channel access algorithm.

---

1. Generate the initial value of the Baum–Welch algorithm based on genetic algorithm;
2. Estimate the parameters of HMM based on the Baum–Welch algorithm;
3. $\gamma = \sqrt{\frac{n \ln n}{T(e-1)}}$, $w_i(t) = 1$;
4. $P(\Psi_j) = (1 - \gamma)\frac{w_j(t)}{\sum\limits_{i=n} w_i(t)} + \frac{\gamma}{n}$;
5. Access the channel sensed idle based on $P(\Psi_1), P(\Psi_2), ..., P(\Psi_n)$;
6. If Ack packet is received, $r_k^{\Psi_k} = 1$; otherwise, $r_k^{\Psi_k} = 0$;
7. Update weights based on equations (35) and (36);

The channels with low probability to be sensed will get higher reward. Such operation can improve the sampling sufficiency in the channel sensing process.

We denote the weight of the ith channel at time t by $w_i(t)$. Based on the virtual reward, $w_i(t)$ can be calculated as

$$w_i(t + 1) = w_i(t)e^{\frac{\gamma \tilde{r}_i(t)}{n}} \qquad (36)$$

The probability of strategy Ψ_j is

$$P_{\Psi_j}(t) = (1 - \gamma)\frac{w_j^s(t)}{\sum\limits_{i=1}^{n} w_i^s(t)} + \frac{\gamma}{n} \qquad (37)$$

where $\gamma \in (0, 1]$. In each time slot, SU will access a channel based on strategy probability.

The detail steps of our MAB-based acoustic channel access (MAB-ACA) algorithm are shown in Algorithm 1. Since MAB-ACA algorithm just gives a heuristic solution to the jamming-resilient acoustic channel access problem, it is expected that our strategy can track the optimal strategy. Theorem 2 shows the effectiveness of our MAB-ACA algorithm.

Lemma 1. For any $n > 0$, $\gamma \in (0, 1]$

$$R^{\Psi^*} - R^{\Psi} \le (e - 1)\gamma R^{\Psi^*} + \frac{n \ln n}{\gamma} \qquad (38)$$

holds for any $w_i(t)$ and T.[24]

Theorem 2. The normalized regret ρ/T converges to 0 at rate $O(1/\sqrt{T})$.

Proof. We denote the upper bound of channel gain by $\bar{g}$. Because an SU can access one channel at most, the reward of the optimal channel accessing strategy satisfies

$$R^{\Psi^*} \le T\bar{g} \qquad (39)$$

within T. Let $\gamma = \sqrt{n \ln n/(T(e - 1))}$, we have

$$\rho = R^{\Psi^*} - R^{\Psi} \le (1 + \bar{g})\sqrt{T(e - 1)n \ln n} \qquad (40)$$

based on Lemma 1. Thus, the normalized regret satisfies

$$\frac{\rho}{T} \le \frac{(1 + \bar{g})\sqrt{(e - 1)n \ln n}}{\sqrt{T}} \qquad (41)$$

It converges to 0 at rate $O(1/\sqrt{T})$.

Theorem 3. Algorithm 1 has time complexity $O(n(40T^2 - 38T + 6) + nT)$ and space complexity $O(1 \ln + nT)$.

Proof. In Algorithm 1, we should build the HMM of each channel at first. As shown in Theorem 1, building the HMM of a channel has time complexity $O(40T^2 - 38T + 6)$. Consider there are n channels, the time complexity of HMM building is $O(n(40T^2 - 38T + 6))$. Steps 3–7 implement the channel sensing probability assignment through solving the MAB problem. There are n channels to sense, and an underwater node can sense one channel in each time slot. The size of strategy set is n, and the time complexity of channel sensing probability assignment is $O(nT)$. Therefore, the time complexity of Algorithm 1 is $O(n(40T^2 - 38T + 6) + nT)$.

For the kth strategy, there are three variables: w_k, r_k, and $P(\Psi_k)$. Since the size of strategy set is n, the space complexity to solve the MAB problem is $O(3n)$. According to Theorem 1, building the HMM of a channel has space complexity $O(T + 8)$. So Algorithm 1 has space complexity $O(11n + nT)$.

Performance evaluation and discussions

The channel utilization transition of PU can be described in Figure 1. The false alarm probability and miss detection probability in the channel sensing process are set to random value from 0.05 to 0.25. Four
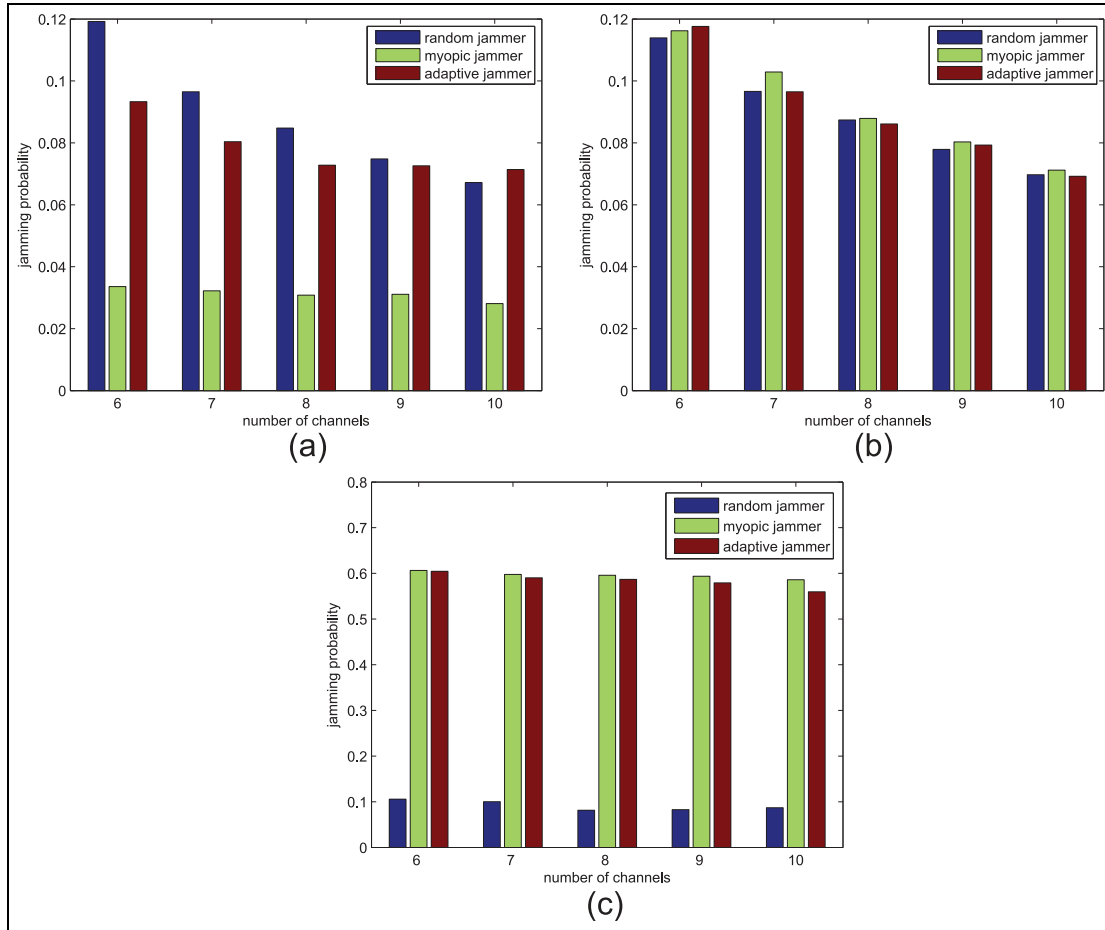
Figure 2. Jamming probability under different approaches: (a) jamming probability under MAB-ACA algorithm, (b) jamming probability under random access, and (c) jamming probability under myopic algorithm.

types of jammers are introduced in section "System and adversary model." The static jammer is not considered in the simulations because it is too easy to be avoided by the SUs. Myopic jammer employs the myopic algorithm in Ahmad et al.,[22] and adaptive jammer utilizes the algorithm in Wu et al.[15]

The performance of jamming-resilient underwater cognitive communication is evaluated in terms of the normalized average throughput, cumulative distribution function (CDF) of expected time to achieve message delivery and jamming probability. The normalized average throughput is used to evaluate the spectrum utilization. Normalized throughput is the ratio of actual throughput to saturated throughput. We repeat the simulations for 1000 times and obtain the normalized average throughput. The CDF of expected time to achieve message delivery describes the probability when a message is successfully received.

Jamming probability is one metric to evaluate the anti-jamming performance in UCANs. Large jamming probability means the SUs are vulnerable to jamming attack. The jamming probability under three kinds of

jammers is shown in Figure 2. Compared to random access, the jamming probability of MAB-ACA algorithm under myopic jammer is much smaller. Compared to myopic access, the jamming probability of MAB-ACA algorithm under myopic jammer and adaptive jammer is much smaller. It shows our approach has better anti-jamming performance in different scenarios. From Figure 2, we can also find the jamming probability will be smaller if there are more channels. For the random jammer, the jamming probability drops quickly when the channel number increases. For the myopic jammer and adaptive jammer, the influence of channel number is not so obvious.

Since the goal of cognitive acoustic network is to improve utilization of underwater channels, we use the normalized average throughput to evaluate the channel utilization in this article. High throughput means more data are transmitted through these channels. A good jamming-resilient algorithm in UCANs needs to achieve higher throughput. The normalized average throughput of the MAB-ACA algorithm is shown in Figure 3(a). As comparisons, the throughput of
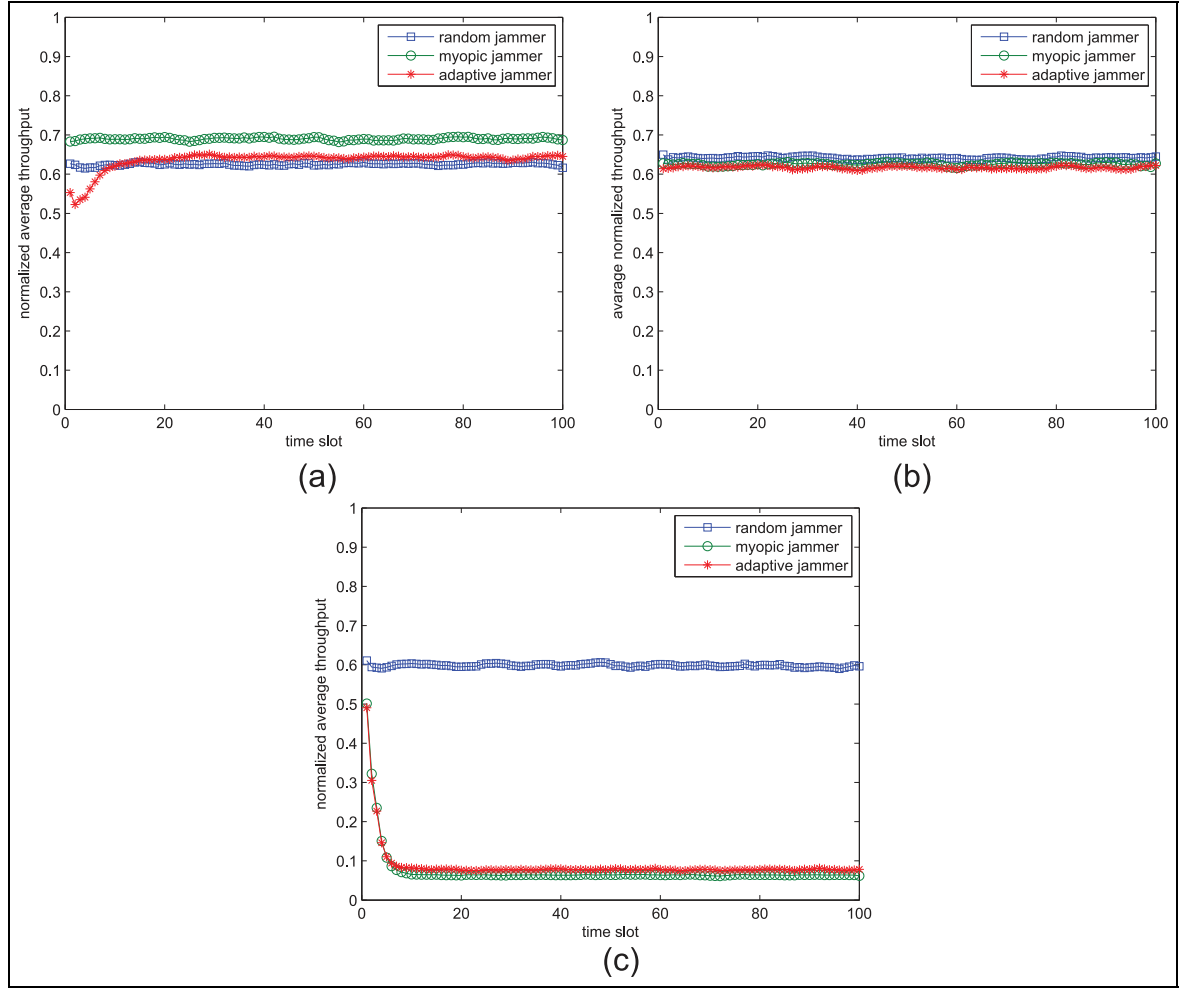
Figure 3. Normalized average throughput under different approaches: (a) the normalized average throughput under MAB-ACA algorithm, (b) the normalized average throughput under random access, and (c) the normalized average throughput under myopic algorithm.

random channel access and myopic channel access is also plotted in Figure 3(b) and (c). For the random channel access strategy, the throughput of three kinds of jammers is similar. Such simulation result is obvious because random access actions will not be influenced by the jammers' strategies. Under random jamming attack, myopic channel access gives a better performance than random jammer in terms of throughput. However, the throughput drops fast under myopic jamming and adaptive jamming. Compared to random channel access, the MAB-ACA algorithm can achieve higher throughput among all three kinds of jammers. The enhancement is quite obvious under myopic jammer. Compared to myopic channel access, the MAB-ACA algorithm achieves similar throughput under random jammer, but the throughput under myopic jammer and adaptive jammer is much higher. For the myopic jammer, the throughput of MAB-ACA algorithm is 7.83 times higher than myopic channel

access. For the adaptive jammer, the throughput of MAB-ACA algorithm is 7.75 times higher than myopic channel access.

Among three kinds of jammers, the throughput under myopic jammer is the highest because the myopic jammer always tries to maximize current reward. Such attacking strategy is easy to learn through our MAB-ACA algorithms. The throughput under random jammer and adaptive jammer is similar. It shows the anti-jamming game between SUs and adaptive jammers ends in a draw. The effect of adaptive learning strategy is just the same as the simple random access.

In UCANs, it takes a long time for an SU receiver to receive a complete message. The message delay performance is evaluated by analyzing the CDF of expected time to achieve message delivery. Assume a message is divided into 10 packets, Figure 4 shows the CDF of expected time to achieve message delivery. For our MAB-ACA algorithm, the whole message can be
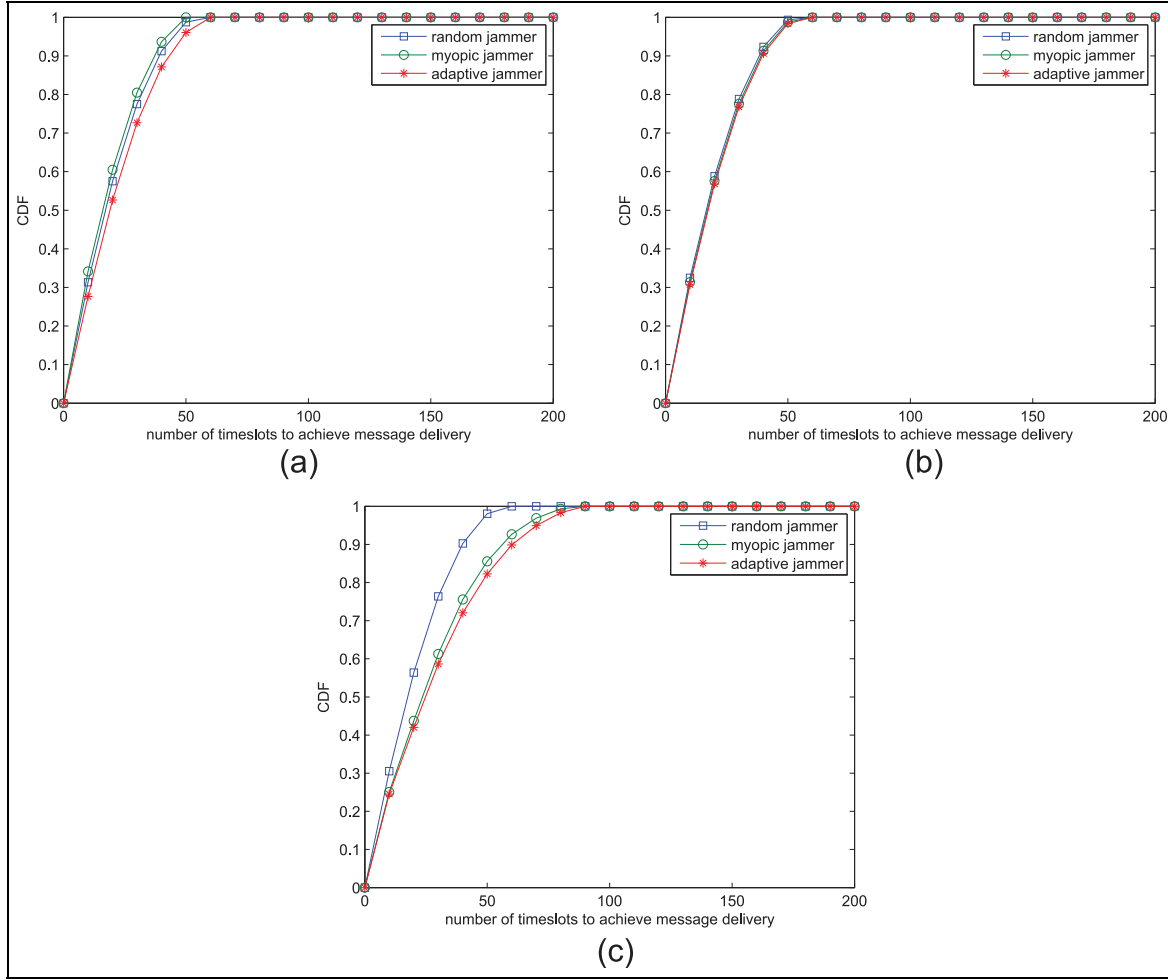
Figure 4. CDF of expected time to achieve message delivery under different approaches: (a) CDF of expected time to achieve message delivery under MAB-ACA algorithm, (b) CDF of expected time to achieve message delivery under random access, and (c) CDF of expected time to achieve message delivery under myopic algorithm.

delivered with high probabilities within 40 time slots. When random or myopic jamming adopted, the CDF performance of random access is similar to MAB-ACA algorithm. Our MAB-ACA algorithm gives better CDF performance under adaptive jamming. When random jamming adopted, the CDF performance of myopic algorithm is similar to MAB-ACA algorithm. Our MAB-ACA algorithm outperforms myopic algorithm under myopic and adaptive jamming.

In the anti-jamming game, it is possible that more than one jammer will participate to jamming the idle acoustic channels. Compared to one jammer scenario, the throughput of our MAB-ACA algorithm is lower. Such performance degradation will be more serious as the number of jammers increases. Figure 5 shows the performance degradation in the multi-jammer scenario. The throughput drops a lot when multiple random jammers exist. For the myopic jammers and adaptive
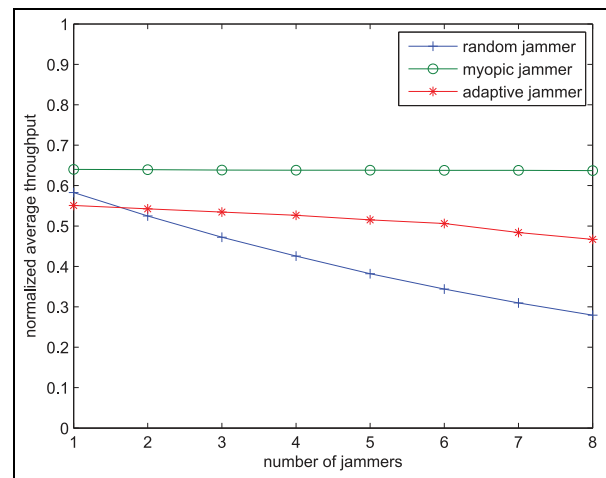

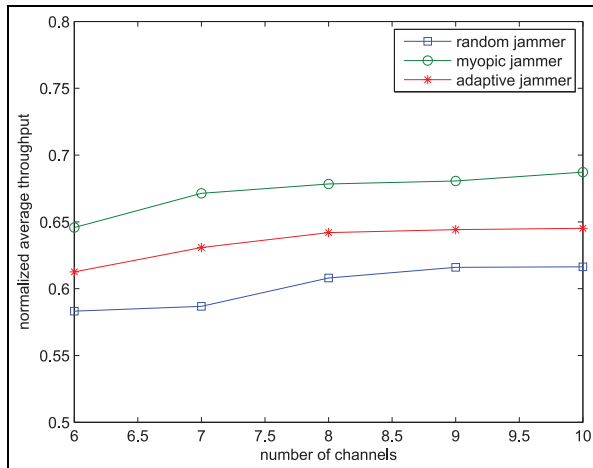**Figure 5.** Normalized average throughput of multiple jammers.

Figure 6. Normalized average throughput with respect to number of channels.

jammers, such performance degradation is not so obvious. The reason lies in the similar actions in the myopic jamming and adaptive jamming strategies. Since this kind of jammers hope to maximize their jamming effect, the actions of different jammers may be identical. As a result, our MAB-ACA can learn the actions of different jammers in a same way, which impairs the effect of multiple jammers.

In UCANs, the number of channels that the SUs can access also affects the throughput performance. Figure 6 shows that the normalized average throughput will increase when there are more channels. In the anti-jamming game, the SUs hop between different channels to avoid jamming attack. More channels can provide more hopping choices, which increases the attack difficulty.

Conclusion

Cognitive acoustic communication is a promising technique for the underwater acoustic networks. In such technique, the SUs utilize the idle channels of PUs to implement their data transmission. The acoustic communication capacity can be significantly improved. However, the jamming attack can block the communication of legitimate users. The existing anti-jamming approaches for CRNs are not suitable for UCANs due to the constraints in the underwater acoustic networks. In this article, we introduced the security issues in the cognitive acoustic networks and proposed jamming-resilient channel access algorithm called MAB-ACA. At first, the channel state is estimated through Baum–Welch genetic algorithm. Then the anti-jamming channel access strategy is made based on MAB algorithm. Under different types of jamming attack, the jamming-resilient approach can achieve higher throughput.

References

1. Stojanovic M. On the relationship between capacity and distance in an underwater acoustic communication channel. In: *Proceedings of WUWNet'06*, Los Angeles, CA, 25 September 2006, pp.41–47. New York: ACM.
2. Domingo M. Overview of channel models for underwater wireless communication networks. *Phys Commun* 2008; 1(3): 163–182.
3. Stojanovic M and Preisig J. Underwater acoustic communication channels: propagation models and statistical characterization. *IEEE Commun Mag* 2009; 47(1): 84–89.
4. Zhang S, Wang Z, Liu M, et al. Energy-aware routing for delay-sensitive underwater wireless sensor networks. *Sci Chin Inf Sci* 2014; 57(10): 1–14.
5. Luo Y, Pu L, Zuba M, et al. Challenges and opportunities of underwater cognitive acoustic networks. *IEEE T Emerg Topic Comput* 2014; 2(2): 198–211.
6. Haykin S. Cognitive radio: brain-empowered wireless communications. *IEEE J Sel Area Comm* 2005; 23(2): 201–220.
7. Wang B and Liu K. Advances in cognitive radio networks: a survey. *IEEE J Sel Top Signa* 2011; 5(1): 5–23.
8. Baldo N, Casari P and Zorzi M. Cognitive spectrum access for underwater acoustic communications. In: *Proceedings of the IEEE international conference on communications workshops*, Beijing, China, 19–23 May 2008, pp.518–523. New York: IEEE.
9. Bicen A, Sahin A and Akan O. Spectrum-aware underwater networks: cognitive acoustic communications. *IEEE Veh Technol Mag* 2012; 7(2): 34–40.
10. Biagi M, Petroni A, Colonnese S, et al. On rethinking cognitive access for underwater acoustic communications. *J Ocean Eng* 2016; 41(4): 1045–1060.
11. Lal C, Petroccia R, Conti M, et al. Secure underwater acoustic networks: current and future research directions. In: *Proceedings of the 2016 IEEE third underwater communications and networking conference*, Lerici, 30 August–1 September 2016, pp.1–5. New York: IEEE.
12. Domingo M. Securing underwater wireless communication networks. *IEEE Wirel Commun* 2011; 18(1): 22–28.
13. Wang Q, Ren K and Ning P. Anti-jamming communication in cognitive radio networks with unknown channel statistics. In: *Proceedings of the 19th IEEE international conference on network protocols*,

Vancouver, BC, Canada, 17–20 October 2011, pp.393–402. New York: IEEE.

14. Wang B, Wu Y, Liu KJ, et al. An anti-jamming stochastic game for cognitive radio networks. *IEEE J Sel Area Comm* 2011; 29(4): 877–889.

15. Wu Y, Wang B and Liu K. Optimal defense against jamming attacks in cognitive radio networks using the Markov decision process approach. In: *Proceedings of the 2010 IEEE global telecommunications conference*, Miami, FL, 6–10 December 2010, pp.1–5. New York: IEEE.

16. Su H, Wang Q, Ren K, et al. Jamming-resilient dynamic spectrum access for cognitive radio networks. In: *Proceedings of the IEEE international conference on communications*, Kyoto, Japan, 5–9 June 2011, pp.1–5. New York: IEEE.

17. Su H and Qiu M. Secure wireless communication system for smart grid with rechargeable electric vehicles. *IEEE Commun Mag* 2012; 50: 62–68.

18. Singh S and Trivedi A. Anti-jamming in cognitive radio networks using reinforcement learning algorithms. In: *Proceedings of the 2012 ninth international conference on wireless and optical communications networks*, Indore, India, 20–22 September 2012, pp.1–5. New York: IEEE.

19. Yang D, Xue G, Zhang J, et al. Coping with a smart jammer in wireless networks: a Stackelberg game approach. *IEEE T Wirel Commun* 2013; 12(8): 4038–4047.

20. Wang Q, Ren K, Ning P, et al. Jamming-resistant multi-radio multichannel opportunistic spectrum access in cognitive radio networks. *IEEE T Veh Technol* 2016; 65(10): 8331–8344.

21. Zhao Q, Krishnamachari B and Liu K. On myopic sensing for multichannel opportunistic access: structure, optimality, and performance. *IEEE T Wirel Commun* 2008; 7(12): 5431–5440.

22. Ahmad S, Liu M, Javidi T, et al. Optimality of myopic sensing in multichannel opportunistic access. *IEEE T Inform Theory* 2009; 55(9): 4040–4050.

23. Wood A and Stankovic J. A taxonomy for denial-of-service attacks in wireless sensor networks. In: Ahson SI (ed.) *Handbook of sensor networks: compact wireless and wired sensing systems*. Boca Raton, FL: CRC Press, 2004, pp.739–763.

24. Auer P, Cesa-Bianchi N, Freund Y, et al. The multi-armed bandit problem: decomposition and computation. *SIAM J Comput* 2002; 32(1): 48–77.

25. Cabric D, Mishra SM and Brodersen RW. Implementation issues in spectrum sensing for cognitive radios. In: *Proceedings of the conference record of the thirty-eighth Asilomar conference on signals, systems and computers*, Pacific Grove, CA, 7–10 November 2004, p.776. New York: IEEE.

26. Ghasemi A and Sousa ES. Spectrum sensing in cognitive radio networks: requirements, challenges and design trade-offs. *IEEE Commun Mag* 2008; 46(4): 32–39.

27. Yucek T and Arslan H. A survey of spectrum sensing algorithms for cognitive radio applications. *Commun Surv Tutor* 2009; 11(1): 116–130.

28. Atapattu S, Tellambura C and Jiang H. Energy detection based cooperative spectrum sensing in cognitive radio networks. *IEEE T Wirel Commun* 2011; 10(4): 1232–1241.

29. Thorp W. Analytic description of the low-frequency attenuation coefficient. *J Acoustl Soc Am* 1967; 42(1): 270.

30. Filar J and Vrieze K. *Competitive Markov decision processes*. Berlin: Springer, 1997.

31. Baum LE, Petrie T, Soules G, et al. A maximization technique occurring in the statistical analysis of probabilistic functions of Markov chains. *Ann I Stat Math* 1970; 41(1): 164–171.

32. Baum LE and Eagon JA. An inequality with applications to statistical estimation for probabilistic functions of a Markov process and to a model for ecology. *B Am Math Soc* 1967; 73(3): 360–363.

33. Baum LE and Sell GR. A maximization technique occurring in the statistical analysis of probabilistic functions of Markov chains. *Pac J Math* 1968; 27(2): 211–227.

34. Li R, Zheng J and Pei C. Text information extraction based on genetic algorithm and hidden Markov model. In: *Proceedings of the international workshop on education technology and computer science*, Wuhan, China, 7–8 March 2009, pp.334–338. New York: IEEE.

35. Oudelha M and Ainon RN. Hmm parameters estimation using hybrid Baum-Welch genetic algorithm. In: *Proceedings of the 2010 international symposium in information technology*, Kuala Lumpur, Malaysia, 15–17 June 2010, pp.542–545. New York: IEEE.