Sensors and Materials, Vol. 35, No. 5 (2023) 1619–1630
MYU Tokyo

S & M 3275

Multi-objective Game Learning Algorithm Based on Multi-armed Bandit in Underwater Acoustic Communication Networks

Hui Wang[1*] and Liejun Yang[2,3]

[1]Minnan Normal University, School of Physics and Information Engineering,
No. 36, Xianqianzhi Street, Zhangzhou 363000, China
[2]Ningde Normal University, School of Information and Mechanical & Electrical Engineering,
No. 1, College Road, Ningde 352000, China
[3]Key Laboratory of Intelligent Ecotourism and Leisure Agriculture, Ningde Normal University,
Fujian Province University, Ningde 352100, China

(Received January 7, 2023; accepted April 25, 2023)

Keywords: underwater acoustic communication, reinforcement learning, power allocation, multi-armed bandit

To address the challenges of interference in underwater multi-node communication and enhance the efficiency of underwater acoustic communication, we propose a multi-objective game learning algorithm based on the multi-armed bandit framework. Firstly, the multi-objective optimization problem is constructed as a multi-node multi-armed bandit (MAB) game model. Secondly, we incorporate the overall network interference level and nodes' power cost in the utility function to achieve the desired optimization objectives. Thirdly, we establish the existence and uniqueness of the Nash equilibrium point of the game model and introduce an improved greedy strategy MAB learning algorithm to determine the equilibrium solution. Finally, our simulation results demonstrate that the proposed algorithm effectively optimizes interference management while enhancing the nodes' adaptive capabilities.

1. Introduction

With the continuous utilization and development of marine resources, the phenomenon of coexistence of multiple operating networks has begun to appear in the ocean.(1,2) In this way, underwater acoustic communication networks (UACNs) will include not only users inside the network, but also those outside the network who use acoustic waves to operate. The diverse types of users constitute generalized UACNs with multiple users. Without uniform resource allocation, the performance of UACNs can be adversely affected.(3,4) One of the pressing challenges in the field of UACNs is to minimize interference and maintain the information rate of users in the communication network. Therefore, finding ways to effectively suppress interference in UACNs has become an urgent area of research.

[*] Corresponding author: e-mail: wh1953@mnnu.edu.cn

https://doi.org/10.18494/SAM4305

ISSN 0914-4935 © MYU K.K.
https://myukk.org/

To improve the quality of underwater acoustic communications, experts have conducted extensive research on the suppression of interference among users in multi-user communication networks, resulting in significant achievements. As an illustration, a distributed medium access control (MAC) protocol regulates transmit power to optimize the throughput of UACNs.(5) Moreover, a distributed power allocation algorithm is proposed to reduce energy consumption for different network densities.(6) In terms of multi-objective optimization, implementing spectrum management schemes that concurrently maximize throughput and minimize delay can significantly increase the spectrum utilization and data transmission rates.(7) Another example involves a power-rate joint allocation algorithm for orthogonal frequency division multiplexing (OFDM)-based UACNs, which optimizes node transmit power and improves network transmission rates.(8) However, most of the existing technical solutions rely on traditional optimization ideas, and some operations require manual participation, such as the selection of model parameters. Consequently, the problem of network interference remains unsolved in many applications. Furthermore, current UACNs have generated substantial amounts of available data from environmental interaction or multi-user communication. It is expected that high-density communication networks will generate even more real-time data in the future. However, existing algorithms have yet to fully utilize these data to improve the network's performance.

Network intelligence represents a future development trend and offers an effective means to meet the performance indicators of future networks.(9–12) Reinforcement-learning-based problem models are highly similar to the human learning process, as they both rely on the interactive operation between the decision-making body and the environment, and use the limited online feedback information actively obtained from the environment to gradually improve their performance. Specifically, they emphasize the trade-off between exploring the unknown and leveraging existing experience.(9,10) Furthermore, the interactive module of reinforcement learning has increasingly become a crucial part of performance improvement in neural networks. Theoretical research on this topic has also become the research focus of machine learning for improving theoretical performance.(11,12)

The model characteristics of reinforcement learning are well-suited to the needs of intelligent algorithms in UACNs, particularly to compensate for the lack of information obtained by nodes owing to the unknown configuration environment. As such, online reinforcement learning algorithms have emerged as effective tools for solving the self-organization problem of underwater networks and have been the subject of significant research.(13,14) In reinforcement learning, the multi-armed bandit (MAB) model has become the focus of research in the theoretical analysis part of reinforcement learning in recent years owing to its simplicity of setting and rich extensibility.(15,16) In this study, we address the issue of transmit power allocation among communication nodes in the context of underwater acoustic communication. We present a novel distributed power allocation game algorithm that takes into account the quality of the communication service. Firstly, we establish a multi-node game model and demonstrate the existence of the Nash equilibrium solution. Secondly, we introduce the MAB model and its algorithm as a practical means of implementing our approach in the multi-node scenario. Finally, we conduct a system simulation to evaluate the performance of the proposed algorithm.

The remainder of this paper is organized as follows. In Sect. 2, the system model is introduced. In Sect. 3, we outline the problems that need to be addressed and present the corresponding solutions. In Sect. 4, we evaluate the research scheme proposed in this study. Finally, in Sect. 5, we provide a summary of the entire paper.

2. Related Work

One of the most challenging issues in node deployment in UACNs is to set the power of communication nodes appropriately to ensure network coverage while minimizing interference between communication nodes within the network.(17,18) The underwater environment is characterized by high dynamism, and the coherence time of the acoustic channel is typically shorter than the processing time. Hence, conventional power allocation schemes tend to be outdated and ineffective, and lack scalability.

Machine learning has shown excellent performance in computer vision, natural language processing, and other fields, making it a popular choice for solving resource allocation problems in complex wireless networks. Dynamic power control is a typical application of machine learning to maximize the overall rate in wireless networks. In Ref. 19, a deep-learning-based optimization scheme is proposed to accelerate power allocation in the presence of interference. Specifically, the scheme employs the weighted minimum mean squared error (WMMSE) algorithm to generate power allocation sets, which are subsequently used as labels to train a deep neural network. This approach leverages a considerable amount of global channel state information and allows the network to allocate power effectively. To address the issue of demanding instantaneous global channel state information, a deep-reinforcement-learning-based algorithm is formulated in Ref. 20. This algorithm assumes that adjacent nodes in the network can share local information through cooperation. The algorithm optimizes power allocation in a trial-and-error fashion and converges to the performance of the WMMSE algorithm after sufficient trials. The proposed approach shows promise in enhancing power allocation efficiency and effectiveness in challenging underwater environments. However, further research is needed to assess the scalability and generalizability of these methods in a diverse range of scenarios.

While previous works have shown good performance in power management, they may not be suitable for highly dynamic underwater environments. One of the main advantages of the MAB learning algorithm over supervised learning methods is that it does not require correct input/output data during the training phase. Currently, some works have made significant progress in this area.(21,22) Dynamic power management is the process of quickly adjusting the power of nodes to adapt to environmental changes within a short time frame. In this process, different power values can be modeled as different "rocker arms," and the benefit function is a performance function related to the signal-to-interference and noise ratio (SINR). In Ref. 21, the channel and power selection problem in a heterogeneous communication network is modeled as a multi-user adversarial MAB model, and the algorithm is used to achieve an equilibrium state to maximize the overall performance of the network. In another notable study, the authors investigate the problem of distributed relay station selection and power control using a state-varying MAB model.[22] The work takes into account the state of relevant links and relay nodes, which is modeled as a Markov process. By optimizing the long-term cumulative performance of the network and balancing the overall network rate and running time, the proposed approach offers a promising solution to address the issue of distributed relay station selection and power control in dynamic and challenging underwater environments. However, future research is necessary to evaluate the performance of this approach under different scenarios and to explore its scalability and generalizability.

On the basis of previous work, we propose to combine the dynamic power management problem with the MAB learning algorithm. By doing so, we aim to realize interactive operation between the decision-making body and the environment, and actively utilize the limited online feedback information obtained from the environment to improve the algorithm's performance gradually.

3. Multi-node MAB Game Model

3.1 UACN model

The model for UACNs is depicted in Fig. 1, which comprises M receiving nodes $R_j, j \in M$, and N transmitting nodes $S_i, i \in N$. During communication, a signal transmitted by node Si is received by node R_i, which then forwards it to the surface base station. In scenarios where multiple nodes in the network employ the same frequency band, the interference between nodes, both within and across layers, is inevitable. Therefore, in this study, we focus on addressing the issue of power allocation in UACNs under interference conditions.

SINR received by node R_j in UACNs can be expressed as

$$\gamma_j = \frac{p_j h_{jj}}{\sum_{k \neq j, k=1}^{N} I(d_{kj} < d) p_k h_{kj} + \sigma^2}, \tag{1}$$

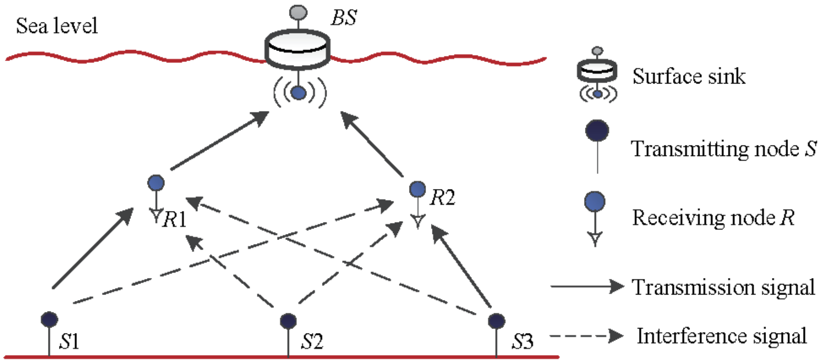


Fig. 1. (Color online) UACN model.

where d_{kj} denotes the distance between the sender k and the receiver j. The communication radius of the sender is denoted by ϱ. If $d_{kj} < \varrho$, then $I(d_{kj} < \varrho)$ is equal to 1; otherwise, it is 0. The variable p_i represents the transmit power of node S_i, whereas p_{-i} indicates the power strategy of other transmitting nodes, excluding node S_i. The channel gain of the interference caused by transmitting node S_k to receiving node R_i is denoted as h_{kj}. Additionally, $\sum_{k\neq j,k=1}^{N} p_k h_{kj}$ represents the interference resulting from other transmitting nodes that use the same frequency channel as receiving node R_i. In UACNs, the channel gain h can be denoted as(15)

$$h = A_0^{-1} d^{-sp} (\alpha(f))^{-d}, \tag{2}$$

where the normalization coefficient, denoted as A_0, accounts for the reference amplitude of the wave. The transmission distance d represents the distance traveled by the wave through the medium. The communication frequency f refers to the frequency of the acoustic wave used for communication. The spread loss d^{-sp} characterizes the attenuation of the wave due to spreading. The spread coefficient sp, which takes a value of 1.5 in this study, describes the rate at which the wave spreads through the medium. Additionally, the absorption coefficient, denoted as $\alpha(f)$, is determined using Thorp's empirical formula(23)

$$10\alpha(f) = \frac{0.11f^2}{1+f^2} + \frac{44f^2}{4100+f^2} + 2.75\times10^{-4} f^2 + 0.003. \tag{3}$$

Moreover, we assume that the noise power of all underwater acoustic channels in our study is uniform and denoted as σ^2.(24) Given the knowledge of channel parameters for all nodes, the channel capacity of the jth channel link can be obtained using the Shannon theorem formula as

$$C_j = \frac{B}{2}\log_2(1+\gamma_j),\ j = 1,2,\cdots,N, \tag{4}$$

where B denotes the channel bandwidth.

3.2 Multi-node MAB game model

During underwater transmission, each transmitter aims to achieve a higher quality of service by transmitting data at a higher power, given the noise conditions of the underwater acoustic channel. However, if nodes increase their transmit power excessively, it can lead to interference between users and result in a higher level of network interference, which ultimately reduces the quality of service for users. Therefore, a balance must be struck between user service quality and network interference level to optimize system performance. Game theory can provide a solution to this problem. A game-theoretic model of power allocation in UACNs can be expressed as

$$G=\left\{N,\left\{B_i\right\}_{i\in N},\left\{U_i(\bullet)\right\}_{i\in N}\right\}. \tag{5}$$

Note that the set of players in our power allocation game is the N transmitting nodes, and each transmitting node $i \in N$ has a strategy represented by its transmit power denoted by $p_i \in B_i$. The set of all strategies for all players is also denoted by B_i, and the reward obtained by transmitting node $i \in N$ using its own strategy in the game is represented by U_i.

For each player, the power allocation problem can be treated as a MAB problem, where the strategy of node (player) $i \in N$ is its power allocation strategy, denoted by p_i. If node $i \in N$ has s feasible strategies, then its feasible strategy set is $P_i = \left\{p_{i,1}, p_{i,2}, \ldots, p_{i,s_i}\right\}$. Furthermore, the reward in the MAB problem corresponds to the utility in the game problem. Throughout the solution process, players can find the optimal strategy without requiring any direct information exchange or prior knowledge of the channel state information.

4. Problem Description and Proposed Scheme

4.1 Problem formulation

In a MAB game with multiple players, each player aims to maximize their own utility. We construct the following utility function for player $i \in N$:

$$U_i(p_i) = B\log_2\left(1+\frac{p_i h_{ii}}{\sum_{i\neq j,j=1}^{N} I(d_{ji}<d)p_j h_{ji}+\sigma^2}\right)-\alpha_i p_i, \tag{6}$$

where B denotes the channel bandwidth, p_i represents the transmit power of user i, h_{ii} denotes the channel gain from transmitting node i to receiving node j, α_i represents the price factor, and p_{max} denotes the maximum transmit power of a node.

On the basis of the aforementioned analysis, we can express the optimization problem for each player i in the multi-player MAB game problem as

$$\begin{aligned}&\max_{p_i} \log_2\left(1+\frac{p_i h_{ii}}{\sum_{i\neq j,i=1}^{N} I(d_{ji}<d)p_j h_{ji}+\sigma^2}\right)-\alpha_i p_i \\ &s.t. \quad p_i \in B_i \\ &\qquad\ 0 \le p_i \le p_{max}\end{aligned}. \tag{7}$$

As indicated in Eq. (7), in this study, we investigated the power allocation problem among multiple nodes while taking interference constraints into consideration. Owing to the competitive nature of the interaction between nodes, a node's satisfaction (utility) depends not only on its own strategy, but also on the strategies of other nodes. Consequently, the objective of all players is to adjust their strategies to maximize their own utility. Thus, we seek a Nash equilibrium (NE) as the solution to this game.

4.2 Power assignment algorithm description

From the system model, it is known that the problem of solving the Nash equilibrium can be abstracted as a single-step gain maximization problem in the context of reinforcement learning. The reinforcement learning algorithm is to select a set of parameters with the largest benefit value from the action space through interactive learning. Among the current mainstream reinforcement algorithms, MAB is a learning algorithm that finds a single-step optimal action by interacting with the environment. MAB generally contains two important concepts: exploration and exploitation.(25,26) Exploration refers to taking actions for which the current reward is unknown or underestimated when interacting with the environment to estimate the reward of the unknown action; utilization refers to taking the action with the largest estimated reward among the current actions. The purpose of exploration is to find potential optimal actions, and the purpose of exploitation is to guarantee the reward obtained during the learning process. MAB generally finds the optimal action by compromising exploration and utilization, and the commonly used compromise methods generally include greedy and softmax algorithms.

In this study, we propose an approach to address the multi-node MAB game problem by employing ε-greedy, which offers an improved solution for selecting the power allocation strategy. Specifically, each decision is made using the probability ε, which allows for the exploration of non-optimal solutions and selects the action with the highest current reward with probability $1-\varepsilon$. Typically, ε is set to 0.1. However, this approach has some limitations that should be taken into account. For a fixed reward distribution, an excessively high ε will lead to too many random exploration processes, resulting in lower average returns. However, in the case of random rewards, if ε is very small, the strategy will not be able to fully explore the action to obtain a local optimal action, and the final reward cannot be maximized.

Algorithm 1: $\varepsilon(t)$-greedy-based game learning algorithm

```
1: For all actions $a \in P_i$, initialize the counter Count(a) = 0 and the expected reward estimate
   Q(a) = 0.
2: Initialize the maximum sampling probability $\varepsilon_{start} = 0.9$, the minimum sampling probability
   $\varepsilon_{end} = 0.02$, and the learning time T.
3: for $t = 1 \rightarrow T$ do
   a)  Calculate the sampling probability $\varepsilon$ at time t via Eq. (8).

       $\varepsilon \leftarrow \varepsilon_{start} - \frac{t}{T} \times (\varepsilon_{start} - \varepsilon_{end})$        (8)

   b)  Select the action according to Eq. (9).
```

$$a_t = \begin{cases} \text{random selection from } P_i & \text{Sampling probability:} 1-\varepsilon \\ \arg\max_{a \in P_i} Q(a_t) & \text{Sampling probability:} \varepsilon \end{cases} \tag{9}$$

c) Calculate the reward $U_i(a_t)$ via Eq. (6).

d) Update counter $Count(a_t)$.

$$Count(a_t) \leftarrow Count(a_t) + 1 \tag{10}$$

e) Update expected reward $Q(a_t)$.

$$Q(a_t) \leftarrow \frac{Q(a_t) \times Count(a_t) + U_i(a_t)}{Count(a_t) + 1} \tag{11}$$

4: **end for**

During the action selection process, it is common to increase the number of explorations during the early stage and decrease it in the later stage while relying more on exploration conclusions. This means that ε is typically inversely proportional to the number of explorations, T. An improved algorithm, $\varepsilon(t)$-greedy, is based on ε-greedy and can be obtained. The power allocation process is described in detail in Algorithm 1.

5. Simulation and Performance Evaluation

To evaluate the effectiveness of the proposed algorithm, we assume UACNs consisting of three transmitting nodes and three receiving nodes with coordinates as shown in Table 1. In this study, we analyze the algorithm's performance in terms of the (a) effects of different values of the price factor α on the utility function, (b) the impact of varying the number of nodes in UACNs on the algorithm, (c) convergence analysis when three transmitting nodes coexist, and (d) a comparison between the proposed algorithm and the random strategy. The node's maximum transmit power is set to p_{max} = 10 W, the system bandwidth to B = 10 kHz, the propagation coefficient to sp = 1.5, the carrier frequency to f = 20 kHz, the noise power to $\delta^2 = 1.5 \times 10^{-7}$, and the number of iterations to 50000.

Table 1
Coordinate information for multiple transmitter–receiver pairs.

	S_1	S_2	S_3	R_1	R_2	R_3
x (km)	0.1	0.1	0.8	0.5	0.5	0.3
y (km)	0.1	0.5	0.2	0.5	0.1	0.25
z (km)	0.1	0.12	0.1	0.1	0.12	0.12

In this work, we first compare the effects of ε-greedy and $\varepsilon(t)$-greedy on the algorithm. Figure 2 shows that the stability of $\varepsilon(t)$-greedy becomes stronger as the number of learning increases. For exploration using ε-greedy, an action is always chosen randomly with the same probability, which makes the utility curve fluctuate greatly, so the utility estimate obtained is not always optimal. Therefore, the cumulative effect curve of ε-greedy when the time gap reaches 150 is significantly lower than that of $\varepsilon(t)$-greedy. Moreover, if a fixed value of ε is used, the utility estimate becomes more accurate as more rounds are selected, and at this time, we should reduce the number of explorations, i.e., gradually reduce ε. In this way, much exploration will be carried out at the very beginning, and eventually, the focus will be on utilization, so in this work, we use $\varepsilon(t)$-greedy.

Next, we analyze the convergence of the utility and power curves of node S_1 when the number of nodes in UACNs is varied. As shown in Fig. 3, the utility is maximal when the number of nodes is Num = 1, and it decreases as the number of nodes increases, with the minimal value occurring when Num = 3. Similarly, the power is also highest when Num = 1 and lowest when Num = 3. This can be attributed to the fact that as the number of nodes increases, the interference at node S_1 increases, leading to a continuous decrease in its utility.

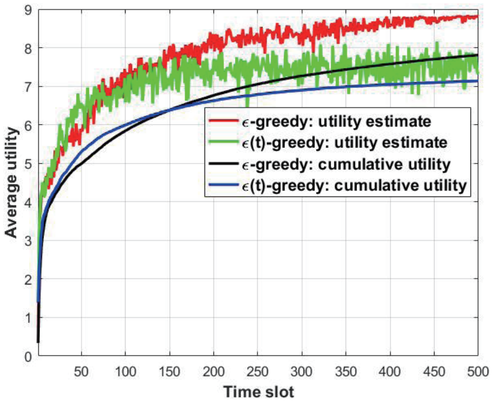
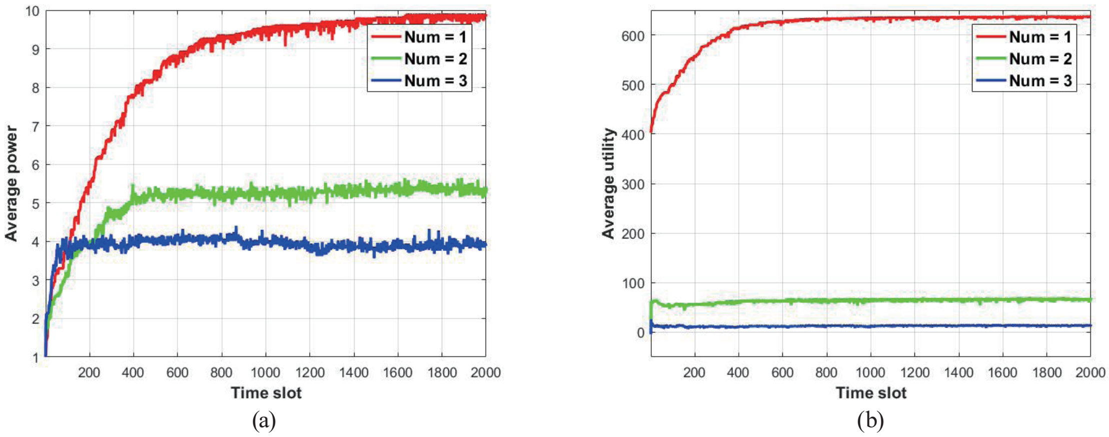


Fig. 2. (Color online) Performance comparison of ε-greedy and $\varepsilon(t)$-greedy.

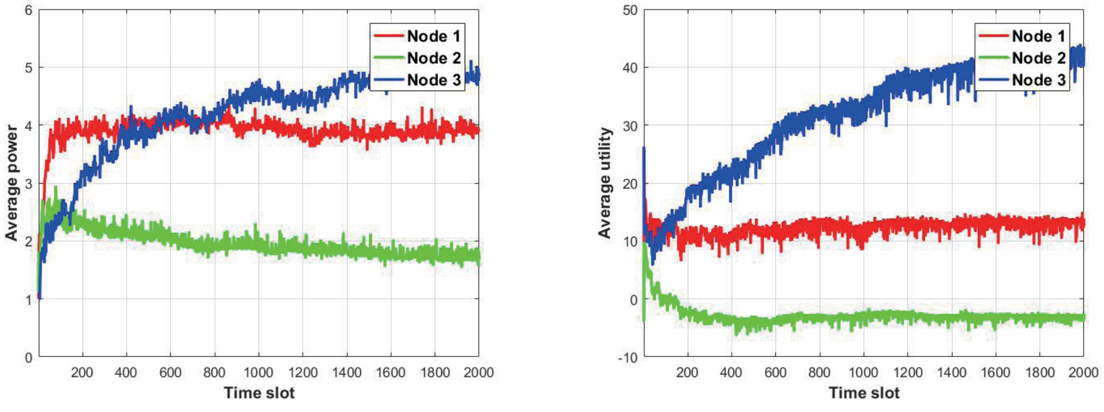


(a) (b)

Fig. 3. (Color online) Performance comparison of (a) power and (b) utility curves of node S_1 as the number of nodes, Num, increases.

In this section, we analyze the convergence of the utility and power curves of nodes S_1, S_2, and S_3 when $Num = 3$. As depicted in Fig. 4(a), the algorithm converges to equilibrium after 200 iterations and obtains the optimal transmit power of each node. It is worth noting that the fluctuation of the transmit power curve of each node decreases with the algorithm iteration, indicating the effectiveness of the proposed algorithm. Moreover, Table 1 shows that the interference of node S_3 increases with the proximity of its receiving node to nodes S_1 and S_2, resulting in the smallest transmit power during equalization, whereas node S_1 has the largest transmit power. Correspondingly, Fig. 4(b) reveals that, upon reaching equilibrium, node S_3 attains the smallest utility, while node S_1 achieves the largest utility.

Finally, we compare the proposed $\varepsilon(t)$-greedy strategy with the random strategy. To mitigate the effect of randomness on algorithm performance, we repeat each algorithm 100 times and observe their average performance. Assuming that there are two transmitter–receiver pairs in the network, Fig. 5 shows the comparison of the utility curve variation under the guidance of $\varepsilon(t)$-greedy and the random strategy as the distance between nodes S_2 and R_1 increases. It is



(a) (b)

Fig. 4. (Color online) Performance comparison of (a) power and (b) utility curves in the case of three groups of nodes.

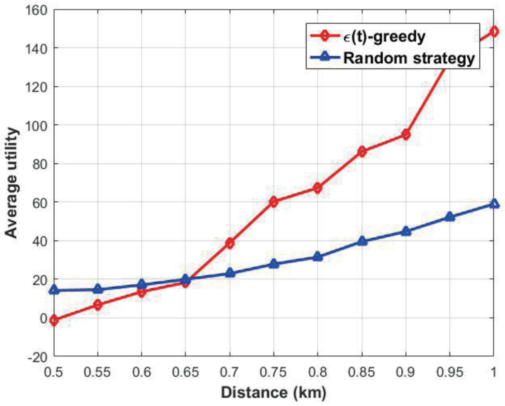


Fig. 5. (Color online) Performance comparison of $\varepsilon(t)$-greedy and random strategy.

evident that as the interference decreases (distance increases), $\varepsilon(t)$-greedy promotes node S_1 to achieve higher utility, reflecting its strong environmental adaptability. This result is primarily attributed to the balanced mechanism of exploration and exploitation in $\varepsilon(t)$-greedy, enabling the node to continuously adjust its transmit power based on the environment.

6. Conclusions

In this study, we proposed a distributed power allocation game algorithm to balance node interference in UACNs, modeled as a multi-node MAB game. First, the transmitting node is considered as an agent and its transmit power is divided into action sets with a utility function fitted on the basis of SINR and power cost. The existence and uniqueness of the constructed multi-node MAB game model Nash equilibrium are then verified. An improved search strategy, $\varepsilon(t)$-greedy, is proposed to achieve the search of multi-node equilibrium points. Simulation and comparative analysis demonstrate that the proposed algorithm can adaptively adjust the transmit power according to the environment and optimize the SINR level of the entire system, leading to an improved network quality of service and preventing the nodes from falling into vicious competition. However, owing to the limited energy supply of the nodes, the transmit power decreases as the battery life decreases. Future work will focus on energy efficiency in UACNs to enhance the nodes’ survival time.

Acknowledgments

This research was supported by several funding sources, including the National Natural Science Foundation of China (Grant no. 62001199), the Fujian Provincial Natural Science Fund (2019J01842 and 2022J011226), the project of the Education Department of Fujian Province (JT180596), the principal fund project of Minnan Normal University (KJ2020003), and the Open funding project of Fujian University Key Laboratory of Intelligent Ecotourism and Leisure Agriculture (LN202201).

References

1 P. Casari and M. Zorzi: Comput. Commun. **34** (2011) 17. https://doi.org/10.1016/j.comcom.2011.06.008

2 R. Diamant, P. Casari, and S. Tomasin: IEEE Trans. Wireless Commun. **18** (2019) 954. https://doi.org/10.1109/TWC.2018.2886896

3 J. Partan, J. Kurose, and B. Levine: ACM SIGMOBILE Mob. Comput. Commun. Rev. **11** (2007) 23. https://doi.org/10.1145/1347364.1347372

4 K. Chen, M. Ma, E. Cheng, F. Yuan, and W. Su: IEEE Commun. Surv. Tutorials **16** (2014) 1433. https://doi.org/10.1109/SURV.2014.013014.00032

5 Y. Su, Y. Zhu, H. Mo, J. H. Cui, and Z. Jin: Ad Hoc Netw. **26** (2015) 36. https://doi.org/10.1016/j.adhoc.2014.10.014

6 J. M. Jornet, M. Stojanovic, and M. Zorzi: IEEE J. Ocean. Eng. **35** (2010) 936. https://doi.org/10.1109/JOE.2010.2080410

7 Y. Luo, L. Pu, H. Mo, Y. Zhu, and Z. Peng: IEEE Trans. Mobile Comput. **16** (2017) 198. https://doi.org/10.1109/tmc.2016.2544757

8 Y. M. Aval, S. K. Wilson, and M. Stojanovic: IEEE J. Ocean. Eng. **40** (2015) 785. https://doi.org/10.1109/JOE.2015.2451251

9 C. Watkins and P. Dayan: Mach. Learn. **8** (1992) 279. https://doi.org/10.1007/BF00992698

10 E. Vargo and R. Cogill: IEEE Trans. Autom. Control. **59** (2014) 2796. https://doi.org/10.1109/TAC.2014.2314527.

11 S. Yin and F. R. Yu: IEEE Internet Things J. **9** (2022) 2933. https://doi.org/10.1109/JIOT.2021.3094651

12 T.-Y. Tung, S. Kobus, J. P. Roig, and D. Gündüz: IEEE J. Sel. Areas Commun. **39** (2021) 2590. https://doi.org/10.1109/JSAC.2021.3087248

13 H. Wang, Y. Li, and J. Qian: IEEE Internet Things J. **7** (2020) 2816. https://doi.org/10.1109/JIOT.2019.2962915

14 Y. Zhang, Z. Zhang, L. Chen, and X. Wang: IEEE Trans. Veh. Technol. **70** (2021) 2756. https://doi.org/10.1109/TVT.2021.3058282

15 M. M. Drugan: IEEE Trans. Neural Netw. **30** (2019) 2493. https://doi.org/10.1109/TNNLS.2018.2885123

16 H. Zhao, X. Li, L. Yan, S. Han, and J. Yu: IEEE Sens. J. **22** (2022) 7961. https://doi.org/10.1109/JSEN.2022.3154974

17 P. Pandey, M. Hajimirsadeghi, and D. Pompili: IEEE J. Ocean. Eng. **39** (2014) 189. https://doi.org/10.1109/JOE.2013.2293932

18 X. Zhong, F. J, F. Chen, Q. Guan, and H. Yu: IEEE Internet Things J. **7** (2020) 9930. https://doi.org/10.1109/JIOT.2020.2990414

19 H. Sun, X. Chen, Q. Shi, M. Hong, X. Fu, and N. D. Sidiropoulos: IEEE Trans. Signal Process. **66** (2018) 5438. https://doi.org/10.1109/TSP.2018.2866382

20 Y. S. Nasir and D. Guo: IEEE J. Sel. Areas Commun. **37** (2019) 2239. https://doi.org/10.1109/JSAC.2019.2933973

21 C. Luo, F. R. Yu, H. Ji, and V. C. M. Leung: Proc. 2010 IEEE Int. Conf. Communications (IEEE, 2010) 1–5. https://doi.org/10.1109/ICC.2010.5502231

22 P. Blasco and D. Gunduz: IEEE J. Sel. Areas Commun. **33** (2015) 585. https://doi.org/10.1109/JSAC.2015.2391852

23 J. M. Jornet, M. Stojanovic, and M. Zorzi: IEEE J. Ocean. Eng. **35** (2010) 936. https://doi.org/10.1109/JOE.2010.2080410

24 Z.-Q. Luo and W. Yu: IEEE J. Sel. Areas Commun. **24** (2006) 1426. https://doi.org/10.1109/JSAC.2006.879347

25 W. Wang, A. Kwasinski, D. Niyato, and Z. Han: IEEE Trans. Commun. **66** (2018) 2588. https://doi.org/10.1109/TCOMM.2018.2799616

26 Y. Gai and B. Krishnamachari: IEEE Trans. Signal Process. **62** (2014) 6184. https://doi.org/10.1109/TSP.2014.2360821

About the Authors

Hui Wang earned his B.S. degree from Nanyang Institute of Technology, Henan, China in 2009, his M.S. degree in computer application technology from Minnan Normal University in 2013, and his Ph.D. degree in communication and information systems from Ningbo University, Ningbo, China in 2019. From 2013 to 2016, he worked as an assistant professor at Ningde Normal University, Ningde, China. Since 2020, he has been an associate professor at Minnan Normal University. His research interests lie in the field of underwater acoustic communication networks. (wh1953@mnnu.edu.cn)

Liejun Yang obtained his B.S. degree in 2002 from Fujian Normal University, China, and his M.S. degree in 2011 from Xiamen University, China. From 2002 to 2008, he was a teaching assistant at Ningde Normal University, China. Since 2009, he has been working as a lecturer at Ningde Normal University. His research interests are focused on embedded systems, IoT engineering, and sensors. (ylj@ndnu.edu.cn)