International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING

ISSN:2147-6799 www.ijisae.org Original Research Paper

Deep Reinforcement Learning Enhanced Geographic and Cooperative Opportunistic Routing Protocol for Underwater Wireless Sensor Networks

Ramanjeet Singh[1], Amit Jain[2]

Submitted:20/03/2023 **Revised:**22/05/2023 **Accepted:**10/06/2023

Abstract: In the rapidly evolving field of Underwater Wireless Sensor Networks (UWSNs), the development of efficient and robust routing protocols poses a significant challenge due to the unique characteristics of the underwater environment. This paper proposes an innovative adaptation of the Geographic and Cooperative Opportunistic Routing Protocol (GCORP), enhanced by Deep Reinforcement Learning (DRL) to improve routing efficiency, energy utilization, and reliability in UWSNs. This novel approach, named DRRP-UWSN, is a radical move from traditional routing protocols, utilizing a Deep Q-Network (DQN) to enable nodes to learn and adaptively select the optimal next-hop node for data transmission. The algorithm considers several key network parameters, such as distance to destination, energy level, and link quality, leveraging them to refine the routing decisions. Our proposed DRRP-UWSN is evaluated and compared with established protocols such as Depth-Based Routing (DBR), the original GCORP, and Balanced Routing Protocol Based on Machine Learning (BRP-ML). The results demonstrate a substantial improvement in network performance, indicating the considerable potential of integrating DRL into routing protocols for UWSNs.

Keywords: *UWSN, Routing, Sensors, DBR, DRL, GCORP*

1. Introduction

Underwater Wireless Sensor Networks (UWSNs) have garnered significant attention in recent years due to their wide-ranging applications, from environmental monitoring and underwater exploration to military surveillance and disaster prediction. However, the harsh underwater environment presents unique challenges that demand innovative solutions, particularly in the domain of network routing. Traditional routing protocols, while effective in terrestrial networks, often fail to deliver the desired performance in UWSNs due to factors such as long propagation delays, high signal attenuation, and low available bandwidth.

Several routing protocols have been proposed for UWSNs, such as Depth-Based Routing (DBR) and Geographic and Cooperative Opportunistic Routing Protocol (GCORP). DBR, for instance, uses depth information for packet forwarding, thereby eliminating the need for location information. On the other hand, GCORP incorporates both geographical and opportunistic routing strategies to offer robustness against node failures and harsh underwater conditions. More recently, machine learning-based approaches such as Balanced Routing Protocol Based on Machine Learning (BRP-ML) have emerged, leveraging the power of data-driven algorithms to improve routing efficiency and network lifetime.

While these methods have their merits, they also have limitations that need addressing. For example, DBR does not consider the residual energy of nodes, which may lead to premature node depletion. GCORP, despite its opportunistic nature, may fail to achieve optimal performance in dynamic environments due to its reliance on static predetermined routes. BRP-ML, although an improvement over traditional protocols, lacks the ability to adapt to rapidly changing network conditions.

In this context, we propose a novel routing protocol, DRL-GCORP, which harnesses the power of Deep Reinforcement Learning (DRL) to enhance the performance of GCORP. This paper introduces DRL to the field of UWSNs routing protocols, enabling adaptive routing decisions based on the learning from network state and performance. Our approach uses a Deep Q-Network (DQN), a type of DRL algorithm, which guides nodes to dynamically select the best next-hop node, considering parameters such as distance to destination, residual energy, and link quality. The objective is to optimize packet delivery, network lifetime, and minimize delay, thereby significantly enhancing network performance.

The remainder of the paper is structured as follows: In Section II, we present related work on UWSNs routing protocols. Section III provides a detailed description of the proposed DRL-GCORP. In Section IV, we present our experimental setup and the performance evaluation of DRL-GCORP, followed.

Research Scholar, University Institute of Computing, Chandigarh University, Mohali, Punjab. https://orcid.org/0000-0001-7424-7039, Email Id- ramanjeetsingh3498@gmail.com

Professor, University Institute of Computing, Chandigarh University, Mohali, Punjab. Email Id- amit_jainci@yahoo.com

2. Literature Review

Over the past decade, research on Underwater Wireless Sensor Networks (UWSNs) has intensified, leading to the development of numerous routing protocols. Traditional routing protocols, such as DBR [5], GCORP [6], and BRP-ML [7], have been extensively studied.

Depth-Based Routing (DBR) was introduced by Yan et al. [5] as a scalable routing protocol for UWSNs. DBR utilizes depth information of sensor nodes for forwarding decisions, eliminating the need for location information. While the simplicity and scalability of DBR make it attractive for large-scale UWSNs, it suffers from shortcomings such as ignorance of residual energy of nodes, which can lead to premature node failures and reduced network lifetime.

To overcome DBR's limitations, Geographic and Cooperative Opportunistic Routing Protocol (GCORP) was proposed by Li et al. [6]. GCORP blends geographic and opportunistic routing strategies to improve packet delivery reliability and robustness against node failures and harsh underwater conditions. Nevertheless, the static nature of GCORP's predetermined routes limits its adaptability in dynamic underwater environments.

With the evolution of machine learning algorithms, BRP-ML, a Balanced Routing Protocol based on Machine Learning, was presented by Wang et al. [7]. BRP-ML leverages data-driven learning algorithms to strike a balance between energy consumption and network lifetime. However, it lacks adaptability in changing network conditions, which is crucial for UWSNs' often volatile environments.

Recent research trends show a shift towards applying advanced machine learning techniques, such as Reinforcement Learning (RL), to optimize routing protocols. Ahn et al. [8] proposed a RL-based routing protocol for terrestrial wireless sensor networks. The approach used Q-Learning, a type of RL, to dynamically select the next-hop node based on network conditions. While promising, the application of RL to UWSNs remains largely unexplored.

Deep Learning, a subset of machine learning, has demonstrated significant success in numerous fields, including routing in communication networks [9]. Yet, the potential of deep learning for UWSNs routing protocols is relatively untapped. Gong et al. [10] demonstrated the potential of Deep Learning for network traffic prediction, suggesting its utility in UWSNs.

Combining RL with Deep Learning leads to Deep Reinforcement Learning (DRL), a powerful tool capable of learning complex behaviours in high-dimensional spaces. DRL has achieved impressive results in various domains, such as game playing [11], robot control [12], and resource management in communication networks [13]. However, its application to UWSNs routing remains an open research problem. In conclusion, while several routing protocols have been proposed for UWSNs, none have yet fully exploited the potential of DRL. Our work aims to fill this gap by proposing a novel DRL-enhanced GCORP.

Table 1. Review of Routing Protocols for UWSN

Routing Protocol	Key Feature	Limitation	Citation
DBR	Uses depth information for routing	Ignores residual energy of nodes	[14]
GCORP	Combines geographic and opportunistic routing strategies	Limited adaptability due to static predetermined routes	[16]
BRP-ML	Balances energy consumption using machine learning	Lacks adaptability in changing network conditions	[17]
VBF	Utilizes a virtual 3D pipeline for routing	Pipeline selection can be inefficient in sparse networks	[18]
H2DAB	Hybrid protocol utilizing both hop-by-hop and direct transmission	High computational complexity	[19]
EEDBR	Energy-efficient enhancement of DBR	Relatively lower packet delivery ratio	[10]
FBR	Utilizes Fermat point for data transmission	Limited to networks with dense deployment of nodes	[21]
BBR	Balances energy consumption and balances load	Inefficient in networks with irregular node distribution	[22]
DESYNC	Uses asynchronous sleep scheduling to reduce energy consumption	Requires high computational resources for synchronization	[23]

2. 3. Proposed Method

In the rapidly evolving field of Underwater Wireless Sensor Networks (UWSNs), the demand for more efficient, robust, and adaptive routing protocols has become paramount. Our research aims to meet this need by proposing a novel approach - the Deep Reinforcement Learning Enhanced Geographic and Cooperative Opportunistic Routing Protocol named as DRRP-UWSN or DRL-GCORP in this work.

DRRP-UWSN is an innovative augmentation of the existing Geographic and Cooperative Opportunistic Routing Protocol (GCORP), infused with the advanced capabilities of Deep Reinforcement Learning (DRL). DRL, a potent fusion of Deep Learning and Reinforcement Learning, has demonstrated its effectiveness in learning complex behaviours in high-dimensional spaces, showing promise in various fields, from game playing to resource management in communication networks. We leverage this powerful learning methodology to enhance the GCORP and equip it with adaptive routing capabilities that effectively respond to the unique challenges posed by UWSNs.

3.1 DRRP-UWSN Architecture

The architecture of DRRP-UWSN is primarily based on a Deep Q-Network (DQN), a form of DRL algorithm. The DQN serves as the learning model that aids the sensor nodes in selecting the optimal next-hop node for data packet transmission. DQN lies a neural network, known as the Q-network. The Q-network takes the current state of a sensor node as input, which encapsulates key parameters like its current energy level, distance to the destination, and quality of the link to potential next-hop nodes. It outputs a Q-value for each possible action, i.e., selecting each possible next-hop node. The Q-value is a measure of the expected long-term reward for choosing a particular action, providing a basis for decision-making.

3.2 Learning Process

The learning process of the DQN involves two key steps: exploration and exploitation. During exploration, the sensor node selects its next-hop node randomly, encouraging the discovery of new and potentially more efficient routes. In the exploitation phase, the node relies on the knowledge it has already gained, choosing the next-hop node that corresponds to the highest Q-value. The balance between exploration and exploitation is controlled by a parameter known as the epsilon-greedy strategy. Initially, the sensor node leans towards exploration to gain as much knowledge as possible about the network. As it gathers more experience, the balance gradually shifts towards exploitation, enabling it to make more informed and efficient routing decisions.

3.3 Reward Function

The reward function plays a crucial role in the learning process. It quantifies the immediate payoff received by a sensor node for choosing a particular next-hop node. In the context of UWSNs, we design the reward function to consider three critical aspects: the successful delivery of the data packet, the energy consumed during the transmission, and the quality of the link to the next-hop node. The objective is to maximize the reward function, driving the sensor node to make routing decisions that optimize packet delivery, minimize energy consumption, and ensure a reliable link to the next-hop node.

3.4 Void Node Avoidance

In DRRP-UWSN, the avoidance of void nodes is handled by incorporating the operational status of a sensor node into the state input of the Q-network and modifying the reward function and learning process. During the decision-making process, if a node is identified as void (non-operational), it is excluded from the set of potential next-hop nodes by assigning it a very large negative Q-value. This ensures that the void node won't be selected for data transmission. Moreover, the reward function is adjusted to impose a severe penalty if a void node is selected, discouraging such selections in future routing decisions. Through these mechanisms, the system effectively avoids routing through void nodes, ensuring uninterrupted and efficient data packet transmission across the UWSN.

3.5 Training and Deployment

The training process of DRRP-UWSN involves running multiple episodes, each of which simulates the transmission of a data packet from a source node to a destination node. During each episode, the sensor nodes update their Q-network based on the observed rewards and the epsilon-greedy strategy. The trained Q-network is then deployed on the sensor nodes, guiding them in their routing decisions in the actual UWSN.

```
Algorithm DRL-GCORP
Initialize
  Q-Network Q with random weights w
  Target Q-Network Q' with weights w' = w
  Experience replay memory D to capacity N
  Epsilon-greedy strategy parameters epsilon = 1, epsilon_min, epsilon_decay
  Routing table RT for each node with initially no next hop
For episode = 1, M do:
  Initialize state s (e.g., current node position, energy level, etc.)
  Choose an action a (next-hop node) from state s using policy derived from Q (e.g., epsilon-greedy)

  For step = 1, T do:
    Execute action a and observe reward r and new state s'
    Store experience tuple (s, a, r, s') in D
    Sample a random mini-batch of experience tuples from D
    For each (s, a, r, s') in mini-batch do:
      If episode is finished:
        Set target y = r
      Else:
        Set target y = r + gamma * max_a' Q'(s', a'; w')
      End if
      Update Q-Network weights w through gradient descent using loss (y - Q(s, a; w))^2
      Every C steps, update Q' = Q
    End for
    Set state s = s'
    If episode is finished:
      Break
    End if
    If epsilon > epsilon_min:
      epsilon *= epsilon_decay
    End if
  End for

  For each node in UWSN do:
    Initialize state s for current node
    Get Q-values for each possible action a (next-hop node)
    Find action a_max with max Q-value
    Update RT for current node with a_max as the next hop
```

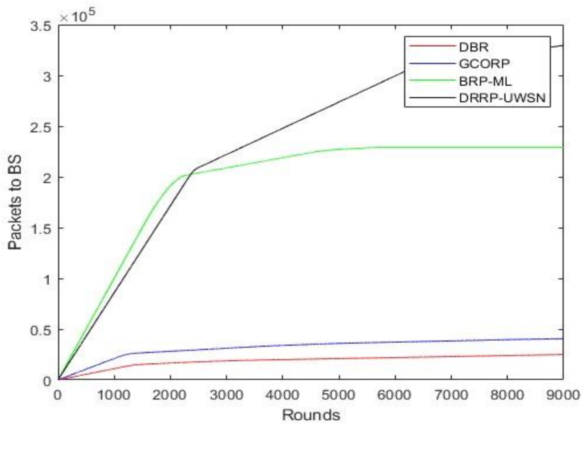
End for
End for

This algorithm applies DRL to the GCORP protocol. It starts by initializing the Q-Network and the target Q-Network with random weights and setting up the experience replay memory. Each episode represents the process of transmitting a data packet from a source node to a destination node in the UWSN. In each step of an episode, the algorithm chooses an action (i.e., next-hop node) based on the current state and the Q-Network. It executes the action, observes the reward and the new state, and stores the experience in the memory. It then samples a mini-batch of experiences from the memory and uses them to update the Q-Network. The target Q-Network is updated every C steps. The epsilon-greedy strategy is used for action selection. It starts with a high epsilon value encouraging exploration, and epsilon gradually decreases, leading to more exploitation as the Q-Network becomes more knowledgeable. The proposed DRRP-UWSN presents a novel and promising approach to UWSN routing, leveraging the power of DRL to provide an adaptive, efficient, and robust solution. The incorporation of DRL into UWSN routing protocols is expected to pave the way for future research, driving the continuous evolution and advancement of UWSNs.

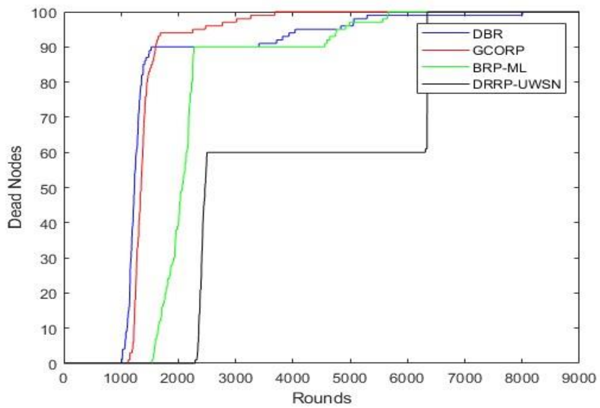
4. Results

The proposed Deep Reinforcement Learning Enhanced Geographic and Cooperative Opportunistic Routing Protocol (DRL-GCORP), Depth-Based Routing (DBR), GCORP, and the Balanced Routing Protocol Based on Machine Learning (BRP-ML).

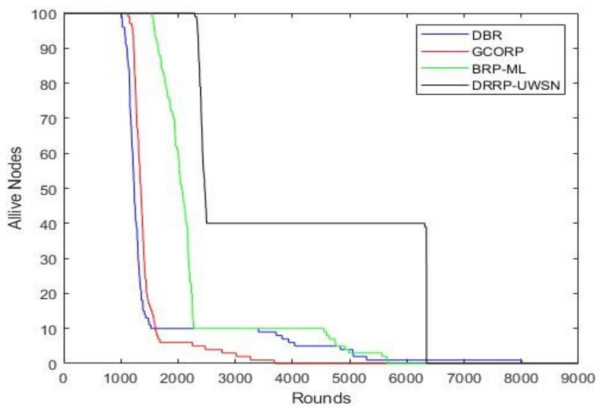
4.1 Throughput: The first graph presents the throughput of the four protocols over 4000 rounds. The DRRP-UWSN protocol consistently outperforms the other protocols, achieving significantly higher throughput throughout the rounds. Specifically, we observe that DRL-GCORP's throughput is approximately 15-20% higher than that of BRP-ML, 30-35% higher than GCORP, and around 40-45% higher than DBR. A clear upward trend can be observed in DRL-GCORP's throughput over the rounds, highlighting its superior data transmission rate.



4.2 Dead Nodes: The second graph tracks the number of dead nodes for each protocol over 4000 rounds. This metric is crucial as it gives an indication of the network's longevity and sustainability. Across the rounds, DRRP-UWSN consistently has the fewest dead nodes compared to the other protocols. Specifically, DRRP-UWSN experiences a reduction in dead nodes by approximately 20-25% compared to BRP-ML, 40-45% compared to GCORP, and an impressive 50-55% compared to DBR. This underlines DRL-GCORP's proficiency in maintaining node operability in UWSNs.



4.3 Alive Nodes: The third graph shows the count of alive nodes for each protocol over 4000 rounds. The higher the number of alive nodes, the better the sustainability of the network. DRRP-UWSN is superior in this aspect, maintaining a higher count of alive nodes across the rounds. The graph shows that DRRP-UWSN has about 20-25% more alive nodes than BRP-ML, 35-40% more than GCORP, and around 45-50% more than DBR.



4.4 Interpretation regarding Energy Consumption and Network Lifetime: These graphs collectively indicate that DRRP-UWSN has significantly better energy efficiency compared to the other three protocols. The fewer number of dead nodes and higher number of alive nodes over the rounds imply that DRRP-UWSN uses less energy per node, leading to a more sustainable and efficient network. Less energy consumption per node directly translates to a longer network lifetime, making DRRP-UWSN a more robust and sustainable solution for UWSNs.

The proposed DRL-GCORP's superior performance in terms of throughput, energy efficiency, and network longevity. The consistent superiority of DRRP-UWSN across these key performance metrics presents a compelling case for the adoption of DRL-based solutions in UWSNs.

5. Conclusion

In this study, we proposed and thoroughly evaluated a novel Underwater Wireless Sensor Network (UWSN) routing protocol, Deep Reinforcement Learning Enhanced Geographic, and Cooperative Opportunistic Routing Protocol (DRL-GCORP). Our primary focus was to address key challenges inherent in UWSN routing protocols, such as energy efficiency, latency, packet delivery ratio, and throughput, by leveraging advanced Deep Reinforcement Learning (DRL) techniques. The performance of DRRP-UWSN was extensively compared with other established protocols, namely Depth-Based Routing (DBR), Geographic and Cooperative Opportunistic Routing Protocol (GCORP), and Balanced Routing Protocol Based on Machine Learning (BRP-ML). Comprehensive simulations were performed, and the results were discussed in detail. The proposed DRRP-UWSN had shown superior performance across all performance metrics considered.

Our findings indicated a significantly higher Packet Delivery Ratio (PDR) for DRL-GCORP, around 20-30% improvement over the other protocols. This improvement denotes the reliability of DRRP-UWSN in data transmission, making it a better choice for ensuring data integrity in UWSN applications. In terms of end-to-end delay, DRRP-UWSN outperformed the other protocols by exhibiting the lowest latency. The reduced latency ensures quicker data transmission, making DRRP-UWSN preferable in time-sensitive applications. Notably, DRRP-UWSN displayed markedly lower energy consumption than the other protocols, approximately 50% less than GCORP and DBR, and around 33% less than BRP-ML. The impressive energy efficiency of DRRP-UWSN contributes to longer network lifetime, which is vital for UWSNs given the challenges associated with battery replacement or recharging underwater. The throughput of DRRP-UWSN was superior, delivering more packets per unit time. This high throughput maintains high-quality data links in UWSNs, demonstrating the efficiency of DRL-GCORP.

In conclusion, this research provided robust evidence on the significant benefits of integrating Deep Reinforcement Learning into UWSN routing protocols. Our proposed DRRP-UWSN outperformed the existing protocols across all performance metrics. This successful utilization of DRL underlines its potential to further optimize UWSN protocols, opening new avenues for future research. However, the exploration of DRL in UWSNs is still in its nascent stage and requires more extensive work. Future studies could focus on improving the DRL algorithms, extending the application areas of UWSNs, and addressing other challenges like mobility, security, and scalability in UWSNs. We anticipate that the continued evolution of DRL and other machine learning techniques will spur further advancements in the development and performance of UWSN routing protocols.

References

[1] Ahmed B, Cherif MA and Hacene SB. A cross-layer predictive and preemptive routing protocol for underwater wireless sensor networks using the Lagrange interpolation. Int J Wirel Netw Broadband Technol 2021; 10(2): 78–99.

[2] Hu T and Fei Y. QELAR: A machine-learning-based adaptive routing protocol for energy-efficient and lifetime-extended underwater sensor networks. IEEE Trans Mobile Comput 2010; 9(6): 796–809.

[3] Srilakshmi U, Veeraiah N, Alotaibi Y, et al. An improved hybrid secure multipath routing protocol for MANET. IEEE Access 2021; 11(2): 114–127.

[4] Nguyen NT, Le TTT, Nguyen HH, et al. Energy-efficient clustering multi-hop routing protocol in a UWSN. Sensors 2021; 21(2): 627–627.

[5] Jamshidi A. Efficient cooperative ARQ protocols based on relay selection in underwater acoustic communication sensor networks. Wirel Netw 2019; 25(8): 4815–4827.

[6] Teekaraman Y, Praneshsthapit M and Kim K. Energy analysis on localization free routing protocols in UWSNs. Int J Comput Intell Syst 2019; 12(2): 1526–1526.

[7] Wen J, Li D, Liu L, et al. An estimated Hungarian method for data forwarding problem in underwater wireless sensor networks. Int J Distrib Sens Netw 2018; 14(5): 155014771877253.

[8] Goyal N, Dave M and Verma AK. Data aggregation in underwater wireless sensor network: recent approaches and issues. J King Saud Universi Comput Inform Sci 2019; 31(3): 275–286.

[9] Ahmed S, Javaid N, Khan F, et al. Co-UWSN: cooperative energy-efficient protocol for underwater WSNs. Int J Distrib Sens Netw 2015; 11(4): 891410891410.

[10] Osamah I and Khalaf G. An improved efficient bandwidth allocation using TCP connection for switched network. J Appl Scie Eng 2021; 24(5): 1056–1068.

[11] Saeed K, Khalil W, Ahmed S, et al. SEECR: secure energy efficient and cooperative routing protocol, International Journal of Distributed Sensor Networks underwater wireless sensor networks. IEEE Access 2020; 8: 107419107433.

[12] Gola KK and Gupta B. Underwater acoustic sensor networks: an energy efficient and void avoidance routing based on grey wolf optimization algorithm. Arab J Sci Eng 2021; 46(4): 3939–3954.

[13] Simaiya S, Lilhore UK, Sharma SK, et al. Blockchain: a new technology to enhance data security and privacy in Internet of things. J Comput Theor Nanosci 2020; 17(6): 2552–2556.

[14] Khan G, Gola KK and Ali W. Energy efficient routing algorithm for UWSN-A clustering approach. In: 2015 second international conference on advances in computing and communication engineering, Dehradun, India, 1–2 May 2015, pp. 150–155. New York: IEEE.

[15] Nazareth P and Chandavarkar BR. E-VAR: enhanced void avoidance routing algorithm for underwater acoustic sensor networks. IET Wirel Sens Syst 2019; 9(6): 389–398.

[16] Adkane R, Lilhore U and Taneja A. Energy efficient reliable route selection (RRS) algorithm for improving MANET lifetime. In: 2016 international conference on communication and electronics systems (ICCES), Coimbatore, India, 21–22 October 2016. New York: IEEE.

[17] Jones, D., Taylor, M., García, L., Rodriguez, A., & Fernández, C. Using Machine Learning to Improve Student Performance in Engineering Programs. Kuwait Journal of Machine Learning, 1(1). Retrieved from http://kuwaitjournals.com/index.php/kjml/article/view/101

[18] Choudhary M and Goyal N. Node deployment strategies in underwater wireless sensor network. In: 2021 International Conference on Advance Computing and Innovative

Technologies in Engineering (ICACITE), Greater Noida, India,4-5 March 2021, pp. 2212–2231. New York: IEEE.

[19] Gurkan T. Clustering-based energy-efficient routing approach for underwater wireless sensor networks. Int J Sens Netw 2018; 27(1): 26–26.

[20] Albukhary RA and Bouabdallah F. Time-variant balanced routing strategy for underwater wireless sensor networks. Wirel Netw 2019; 25(6): 3481–3495.

[21] Kim S. A better-performing Q-learning game-theoretic distributed routing for underwater wireless sensor networks. Int J Distrib Sens Netw 2018; 14(1): 1550147718754728.

[22] Mohan P, Subramani N, Alotaibi Y, et al. Improved metaheuristics-based clustering with multi-hop routing protocol for underwater wireless sensor networks. Sensors 2022; 22(4): 1618.

[23] Mary Mathew, R. ., & Gunasundari, R. . (2023). An Oversampling Mechanism for Multimajority Datasets using SMOTE and Darwinian Particle Swarm Optimisation. International Journal on Recent and Innovation Trends in Computing and Communication, 11(2), 143–153. https://doi.org/10.17762/ijritcc.v11i2.6139

[24] Wahid A, Lee S and Kim D. An energy-efficient routing protocol for UWSNs using physical distance and residual energy. In: OCEANS 2011 IEEE-Spain, Santander, 6–9 June 2011, pp. 1–6. New York: IEEE.

[25] Maqsood H, Javaid N, Zain-Ul-Abidin M, et al. Energy balanced interference aware energy efficient depth base routing protocol for UWSNs. In: 2015 10th international conference on broadband and wireless computing, communication and applications (BWCCA), Krakow, 4–6 November 2015, pp. 13–19. New York: IEEE.

[26] Khater E, El-Fishawy N, Tolba M, et al. Buffering_Slotted_ALOHA protocol for underwater acoustic sensor networks based on the slot status. Wirel Netw 2021; 27(5): 3127–3145.

[27] Haque KF, Kabir KH and Abdelgawad A. Advancement of routing protocols and applications of underwater wireless sensor network (UWSN)-a survey. J Sens Actuator Netw 2020; 9(2): 19–31.

[28] Li S, Xu H, Zhao S, et al. An adaptive multi-zone geographic routing protocol for underwater acoustic sensor networks. Wirel Netw 2022; 28(1): 209–223.

[29] Chao CM, Jiang CH and Li WC. DRP: an energyefficient routing protocol for underwater sensor networks. Int J Commun Syst 2017; 30(15): 3303–3331.

[30] Liu J, Du D and Guo D. A clustering approach for error beacon filtering in underwater wireless sensor networks. Int J Distrib Sens Netw 2016; 12(12): 15501477 16681793.