T.C
BAHCESEHIR UNIVERSITY
GRADUATE SCHOOL OF
THE DEPARTMENT OF COMPUTER ENGINEERING

ENHANCING NETWOK LIFETIME OF UNDERWATER WIRELESS SENSOR NETWORKS WITH SMART ENERGY HARVESTING

PHD THESIS

ÇİĞDEM ERİŞ

ISTANBUL 2024

T.C

BAHCESEHIR UNIVERSITY

GRADUATE SCHOOL OF

THE DEPARTMENT OF COMPUTER ENGINEERING

ENHANCING NETWOK LIFETIME OF UNDERWATER WIRELESS SENSOR NETWORKS WITH SMART ENERGY HARVESTING

PHD THESIS

ÇİĞDEM ERİŞ

THESIS ADVISOR
ASSOC. PROF. DR. PINAR BOLUK

ISTANBUL 2024

T.C.

BAHÇEŞEHİR UNIVERSITY

GRADUATE SCHOOL

…/…/…

PhD THESIS APPROVAL FORM

Name Surname	ÇİĞDEM ERİŞ
Student Number	1266610
Program Name	COMPUTER ENGINEERING (ENGLISH - DOCTORATE)
Title of Thesis	ENHANCING NETWOK LIFETIME OF UNDERWATER WIRELESS SENSOR NETWORKS WITH SMART ENERGY-HARVESTING
Thesis Defense Date	

It has been approved by the Graduate School that this thesis has fulfilled the necessary conditions as a PhD thesis.

Assoc. Prof. Yücel Batu SALMAN

Director of the Graduate School

This Thesis has been read by us, it has been deemed sufficient and accepted as a PhD thesis in terms of quality and content.

PhD Thesis Defense Jury		
Thesis Defense Jury	**Title - Name / Surname**	**Signature**
Thesis Advisor	ASSOC. PROF. DR. PINAR SARISARAY BÖLÜK	
Member of Thesis Monitoring Committee	PROF. DR. ŞEBNEM BAYDERE	
Member of Thesis Monitoring Committee	ASST. PROF. DR. TARKAN AYDIN	
Member	ASSOC. PROF. DR. TEVFİK AYTEKİN	
Member	PROF. DR. TOLGA OVATMAN	

I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.

Name, Surname: ÇİĞDEM ERİŞ

Signature:

ABSTRACT

ENHANCING NETWOK LIFETIME OF UNDERWATER WIRELESS SENSOR NETWORKS WITH SMART ENERGY HARVESTING

ERİŞ, Çiğdem

Computer Engineering (English) PhD Program

Supervisor: Assoc. Prof. Dr. Pınar BÖLÜK

JUNE 2024, 75 pages

Underwater wireless sensor networks (UWSNs) play a crucial role in providing essential environmental data for various applications such as environmental research, disaster management, early warning systems, and industrial exploration and monitoring. However, the challenging underwater environment makes battery replacement unfeasible, requiring sensor nodes to rely on their limited batteries. Therefore, managing energy in these networks is a critical issue. Cluster-based network routing protocols have been extensively explored as a technology to address this energy management problem and optimize network energy consumption. Additionally, harvesting energy from the underwater environment has gained attention as a promising solution to extend the operational life of sensor nodes. In our thesis, we focus on stochastic energy harvesting processes at sensor nodes to address the energy-aware routing problem in underwater acoustic sensor networks. We introduce a novel cluster head (CH) selection approach and compare its performance with existing literature in the first part of the thesis. In the second part of the thesis, we extend our method by adapting the problem into a multi-agent reinforcement learning problem. The contribution of this thesis lies in proposing a novel reinforcement learning-based algorithm to optimize available energy in the underwater environment while determining cluster heads (CHs) considering not only the residual energy and locations of nodes but also the estimated amount of harvestable energy. In the final section of the thesis, a novel reinforcement learning-based algorithm is proposed for scheduling time slots in the medium access protocol of sensors, aiming to enhance energy efficiency in the underwater environment.

Keywords: Underwater Wireless Sensor Networks, Reinforcement Learning, Energy Harvesting, Lifetime

ÖZET

SUALTI DUYARGA AĞLARINDA AKILLI ENERJİ HASADI İLE YAŞAM ÖMRÜNÜN UZATILMASI

ERİŞ, Çiğdem

Bilgisayar Mühendisliği (İngilizce) Doktora Programı

Danışman: Doç. Dr. Pınar BÖLÜK

Haziran 2024, 75 sayfa

Su altı kablosuz sensör ağları (UWSN'ler), erken uyarı sistemleri, küresel ısınma çalışmaları ve endüstriyel keşif takip çalışmaları dahil olmak üzere çevresel verilere erişim olanağı sağlayarak felaketleri önleme potansiyeline sahiptir. Su altı akustik sensör ağları, gözlenmekte olan saha ile ilgili verilerin kesintisiz bir şekilde iletilmesinden sorumlu ana bileşenlerdir. Ancak, su altı ortamının zorlu doğası pil değişimini olanaksız hale getirir ve sensör düğümlerini pil kaynağına güvenmeye zorlar. Bu nedenle, enerji tüketimi su altı sensör ağlarında en önemli tasarım faktörüdür. Ağın enerji tüketimini minimize etmek için, kümeleme, su altı sensör uygulamalarında yoğun olarak çalışılmakta olan bir teknolojidir. Küme başları (CH'ler), veri birleştirme ve ağın enerji tüketimini azaltma amacıyla seçilir ve bu şekilde ağ ömrünü uzatır. Bu çalışmada, UWSN'de, kablosuz sensör ağlarında enerji bilincine sahip yönlendirme sorunu için her bir sensör düğümünde stokastik enerji toplama süreçleri göz önünde bulundurularak yaşam ömrünü iyileştirmesi amaçlanmıştır. Tezin birinci kısmında, yeni bir küme başı (CH) seçme tekniği önerilmiş ve CH'ler sadece düğümlerin konumunu ve kalan enerjisini değil, aynı zamanda hasatlanması tahmin edilen enerjiyi de göz önünde bulundurmuştur. Sayısal sonuçlar, önerilen tekniğin enerji tüketimini azalttığını ve ağ ömrünü önemli ölçüde uzattığını göstermektedir. Tezin ikinci bölümünde, yöntemimizi geliştirerek, problemi çok etmenli pekiştirmeli öğrenme problemine uyarladık. Bu çalışmanın katkısı, küme başlarını (CH'ler), yalnızca sensörlerin kalan enerjisini ve konumlarını değil aynı zamanda kazanılması öngörülen enerji miktarını göz önünde bulundurarak seçmesi, sensörlerin yaşam ömrünü uzatmaktır. Tezin son bölümünde sensörlerin ortam erişim protokolünde zaman aralıklarının planlanmasında, su altı ortamındaki enerjiden alınan verimi arttırmak için yeni bir takviyeli öğrenmeye dayalı algoritma önerilmektedir.

Anahtar kelimeler: Sualtı Kablosuz Sensör Ağları, Takviyeli Öğrenme, Enerji Hasatlama, Yaşam Ömrü

DEDICATION

This thesis is dedicated to my family.

ACKNOWLEDGMENTS

First and foremost, I would like to express my gratitude to my advisor, Assoc. Prof. Pınar Sarısaray Bölük, for her guidance, friendship, and unwavering support throughout my thesis. It was a great honor to work with her and shaped my academic and global perspective significantly.

I would also like to thank Prof. Şebnem Baydere and Assist. Prof. Tarkan Aydın for their guidance and insightful suggestions during the thesis monitoring committee meetings. In addition, I would like to thank Assoc. Professor Tevfik Aytekin and Prof. Tolga Ovatman for being on my thesis defense committee. I would also like to thank Prof. Vehbi Çağrı Güngör for his guidance and suggestions that greatly shaped the research of my thesis. My sincere gratitude also extends to Assoc. Prof. Ömer Melih Gül for his invaluable and seamless support during my studies. I owe him very much.

Finally, I would like to express my deepest gratitude to my family for their patience, continuous support and care throughout this thesis.

TABLE OF CONTENTS

LIST OF TABLES

TABLES

LIST OF FIGURES

FIGURES

LIST OF ABBREVIATIONS

EH-UWSN	Energy Harvesting Underwater Sensor Network
UWSN	Underwater Wireless Sensor Network
UASN	Underwater Acoustic Sensor Network
LEACH	Low-Energy Adaptive Hierarchical Clustering
GW	Gateway
CH	Cluster Head
CM	Cluster Member
BS	Base Station
ACO	Ant Colony Optimization
PSO	Particle Swarm Optimization
RL	Reinforcement Learning
MARL	Multi-Agent Reinforcement Learning
MAB	Multi-armed Bandit
TDMA	Time Division Multiple Access
CSMA	Carrier Sense Multiple Access
CDMA	Code Division Multiple Access
WSN	Wireless Sensor Network
EH-WSN	Energy Harvesting Wireless Sensor Network
UAV	Unmanned Autonomous Vehicle
AUV	Autonomous Underwater Vehicle
UW-A	Underwater Acoustic Channel
SNR	Signal to Noise Ratio

Chapter 1
Introduction

Underwater wireless sensor networks (UWSNs) consist of numerous sensor nodes and have diverse applications, including oceanographic data collection, pollution monitoring, offshore exploration, disaster prevention, assisted navigation, and tactical surveillance. Unmanned or autonomous underwater vehicles (UUVs, AUVs) equipped with sensors further enhance the exploration of natural underwater resources and facilitate the collection of scientific data through collaborative monitoring missions. (Akyildiz et al., 2005a). Underwater acoustic sensor networks (UASN) serve as the foundational technology for these applications. In this realm, it is imperative to establish reliable underwater communication capabilities among sensing devices. These devices require self-configuration capabilities, allowing them to coordinate their operations by exchanging configuration, location, monitoring information and relay observatory data to an onshore station.

In the literature, wireless sensor networks (WSNs) are extensively studied as the foundational technology for the Internet of Things (IoT) paradigm. The primary enabling technology for underwater sensor networks is long-range acoustic communication, whose distinctive characteristics have prompted researchers to reevaluate many networking paradigms specifically designed for wireless sensor networks in the IoT field. Acoustic communication technology presents significant challenges, including limited bandwidth, extremely high and variable propagation delay, severe impairment from multipath and fading, and high bit error rates with intermittent connectivity losses. (Akyıldız Fuat & Vuran Can, 2010). On the other hand, sensors within UASNs rely on the limited energy stored in their batteries. Consequently, the power demands associated with data transmission, reception, and computational tasks in challenging underwater conditions constrain the network's lifespan, making energy efficiency a critical design concern (Stefanov & Stojanovic, 2011a). The replacement of batteries is particularly challenging in underwater applications where subsea deployments are inaccessible to humans. It is imperative to develop communication protocols that prioritize energy efficiency. Recently, considerable attention has been given in the literature to address the limited energy supply of sensor nodes by emphasizing energy harvesting methods and alternative energy sources to supply device batteries (Pobering & Schwesinger, 2004). Recent advancements have shown promising potential in utilizing underwater ambient noise, pressure, and sea currents to significantly extend the operational lifespan of sensor devices. This groundbreaking approach harnesses the energy, and environmental factors present in underwater settings to enhance the durability and functionality of sensor technology, paving the way for more efficient and long-lasting underwater monitoring and data collection (Erdem et al., 2019; Erdem & Gungor, 2017). The transmission power needed in acoustic links varies based on the chosen frequency and distance. This significantly affects path loss and received power, often exceeding the amount of harvested energy. This highlights the increased importance of using ambient energy, presenting a significant challenge.

With the aim of energy efficiency, underwater sensor nodes dispersed across specific regions organized into clusters. These networks are configured to assign each node to a particular cluster. Rather than transmitting data directly to the surface sink, each node in the UWSN sends its data to a designated node known as the cluster head. The cluster head (CH) then relays the data collected from the sensor nodes within a cluster to the base station (BS) or the sink. Additionally, it may forward the data to another cluster head, which acts as a relay in multi-hop communication. The surface sink further disseminates the sensor data received from the cluster heads to the end user via the onshore station’s Internet infrastructure. The low-energy adaptive clustering hierarchy (LEACH) (W. B. Heinzelman et al., 2002; W. R. Heinzelman et al., 2000) stands as widely recognized clustering method introduced for WSNs and has garnered extensive attention in WSN research, several variants of LEACH proposed delving into its efficiency, coverage, connectivity, load balancing, scalability, energy and other aspects (Luo et al., 2021). Various studies investigated cluster-based data routing and the performance of cluster head selection and clustering algorithms on the lifetime of UASNs (Dhongdi et al., 2018; Domingo & Prior, 2007; Jha et al., 2024; M. F. Khan et al., 2021a; Mathankumar et al., 2021; Nguyen et al., 2021; Sandeep & Kumar, 2017; Sun et al., 2022a).

Recently, machine learning-based clustering techniques are becoming more prevalent in UASNs. K-means, a widely recognized clustering algorithm, exemplifies this trend. (Omeke et al., 2021), investigate K-means clustering to group the network, with cluster head selection determined by node density and proximity to centroids.

(Zhu et al., 2021) introduced a cooperative routing method for underwater mobile sensor networks, combining Q-learning and K-means clustering while considering node mobility, packet loss rate, and channel quality. Previous studies have shown a surge in different optimization algorithms for cluster head selection. For instance, (Fei et al., 2020) employs the fuzzy c-means clustering algorithm and selects cluster heads using the moth-flame optimization method. Many of the cluster-based routing algorithms prefer centralized and optimized clustering methods to form balanced clusters and minimize power consumption of the network (L. Li et al., 2022; Y. Li et al., 2014; Senturk, 2021). Yet, computational complexity and centralized approach of these methods can't cope with the spatial uncertainty of underwater medium (Eris et al., 2023; M. F. Khan et al., 2021b). For providing robust and adaptive data routing strategies, online learning algorithms proposed in cluster head selection and communication scheduling is investigated (Basagni et al., 2018; Eriş et al., 2024). Energy efficiency and accountability of energy harvesting methods are focus on the nowadays literature. While centralized and optimized methods are promising to improve the lifetime, considering the temporal and spatial dynamicity of UASNs it is important to develop robust, adaptive strategies in the realm of the UASNs applications.

1.1 Contribution in the Thesis

In a previous study, we conducted research on the lifetime analysis of Wireless Sensor Networks (WSN) in different smart grid radio propagation environments (Eris et al., 2014, 2016). These environments included 500 kV substations, main power control rooms, and underground network transformer vaults. Our main goal was to measure the impact of the smart grid propagation environment and sensor radio properties on the longevity of nodes in demanding smart grid settings. More precisely, the analysis examines sensor node lifespan concerning smart grid channel attributes (like path loss, shadowing variation, etc.), sensor operational states and modes, and network parameters (duty cycle, event occurrence rate, packet reception rate, frame length, distance, etc.). While there are existing analyses of sensor node longevity for various sensor hardware architectures, none of them investigate how different smart grid radio propagation environments impact the network lifespan of the associated smart grid application.

In our study (Eris et al., 2023), we propose a robust and adaptive cluster head selection policy, considering stochastic nature of piezoelectric energy harvesting process for 2D UASN, group of sensors anchored to the ocean bottom. Towards this aim, we examine stochastic energy harvesting processes at individual sensor nodes to address the energy-aware routing challenge in underwater acoustic sensor networks. Our study introduces a new method for selecting cluster heads (CHs), where CHs are determined based on factors including node positions, remaining energy, and expected harvested energy. Numerical findings demonstrate that this approach significantly reduces energy consumption and substantially extends the network's lifespan in case of accounting the expected level of harvested energy can be observed for cluster-head-candidate nodes.

In the study (Eriş et al., 2024), we present a novel reinforcement learning-based algorithm for identifying cluster heads (CHs) from an energy management perspective. We specifically examined piezoelectric harvester model considering the spatial uncertainty in underwater environments. This approach not only considers the locations and residual energy of the nodes but also accounts for the estimated harvested energy. Numerical results validate that our proposed approach remarkably reduces energy consumption by employing adaptive selection of cluster heads and significantly prolongs the network's operational lifespan. We introduce a reinforcement learning-based new cluster head selection algorithm tailored for UASNs. More specifically, our cooperative Q-learning algorithm considers factors such as expected harvested energy, node locations, and estimated energy consumption within each cluster. Cluster head sensors undertake energy-intensive operations within their respective regions when their remaining energy and proximity are adequate. Our proposed reward function accounts for each aspect, leading to notable enhancements in both the lifetime of sensor nodes and throughput, each by approximately 30%.

In our recent study, we addressed TDMA-based slot scheduling problem considering the spatial uncertainty of underwater ambient resources to enhance the effective utilization of available energy in time slot allocation perspective. We used stochastic model for piezoelectric energy harvesting and developed a novel distributed multi-agent reinforcement learning algorithm, considering realistic channel and environmental conditions. Nodes autonomously observe and learn energy peaks in their surroundings, enabling them to independently determine their communication slots without relying on a cluster head. Numerical results demonstrate that our proposed method enhances in-cluster communication time interval utilization and surpasses traditional TDMA-based time slot allocation methods in terms of throughput and energy harvesting efficiency.

1.2 Organization

The doctoral thesis continues with Chapter 2, related work for cluster-based data routing in wireless sensor networks and underwater wireless sensor networks considering energy aware and energy-harvesting aware methods proposed in the literature. Chapter 2 composed of sections which summarizes the background of the cluster-based data routing methods.

In Chapter 3 we present k-means clustering method and our cluster head selection policy is explained in detail. Then the chapter continues with comprehensive performance evaluation and comparison of our energy harvesting aware cluster head selection policy with related works in the literature. The last part concludes Chapter 2 with our findings.

In Chapter 4, we present our novel method, reinforcement learning based cluster head selection algorithm, namely implementation of distributive cooperative Q learning. After introducing the applied methods in literature, in the second part of the Chapter 4, we explain how we improved 3D UASN lifetime and throughput compared to our previous method. Before concluding the Chapter 4, we present performance comparison of two methods namely harvesting-aware moving average prediction and distributed Q-learning based cluster head selection. In Chapter 5 we investigated contention-free medium access control strategy, specifically time slot scheduling problem to maximize harvested energy in the network. We present multi-agent reinforcement learning based time slot allocation algorithm to improve the throughput and the lifespan of the UASNs.

Chapter 6 concludes the research presented in this thesis, providing remarks and exploring potential avenues for future investigation.

Chapter 2
Related Work

In this chapter we provide the related work about cluster-based data routing protocols that consider energy efficiency and energy harvesting in both wireless sensor networks and specifically underwater wireless sensor networks. We survey clustering and cluster head selection methods with the context of reinforcement learning methods aimed to improve performance of UWSNs in the literature.

2.1 Cluster Based Data Routing Protocols for EH-WSNs

Numerous research endeavors in wireless sensor networks have explored clustering, considering several factors such as node proximity, remaining energy levels, distance from the sink, and inter-cluster centrality, all with the aim of prolonging network functionality. The primary objective across these studies is to diminish energy consumption and distribute the workload among cluster heads effectively (Nakas et al., 2020). The Low-Energy Adaptive Clustering Hierarchy (LEACH) (W. R. Heinzelman et al., 2000) stands out as a prominent routing algorithm and extensively studied for facilitating energy-efficient data communication within Wireless Sensor Networks (WSN). Operating as a decentralized clustering approach, LEACH arranges clusters through the randomized selection of cluster heads, with sensor nodes independently opting to join the nearest cluster. However, this random selection process results in the cluster heads with close proximity, leading to imbalanced clusters. In (W. B. Heinzelman et al., 2002), a centralized version LEACH-C is introduced, where cluster formation and cluster head selection are orchestrated by the Base Station (BS). As a centralized version of hierarchical clustering method is proposed that the Base Station (BS) take responsibility for cluster formation and cluster head selection to avoid selecting low-energy nodes as cluster heads (W. B. Heinzelman et al., 2002). Younis et al., (Younis & Fahmy, 2004) introduced HEED to alleviate the load on the cluster heads to improve the scalability of LEACH. Providing a distributed clustering method and periodical selection for cluster heads, a hybrid method of remaining energy and node degree (the proximity of a sensor node to its neighbors) is examined and achieved uniform distribution of cluster heads in the network. Numerous studies delve into energy efficiency, scalability, load balancing, security and reliability to improve the performance of LEACH (Nakas et al., 2020). In (Gawade & Nalbalwar, 2016) a centralized energy efficient clustering protocol CEED is proposed to balance energy consumption of nodes in the network underlying the impact of number of cluster heads on the energy consumption in the network. A distributed cluster head selection strategy is proposed as the consumed energy of nodes in the current round. Energy levels is considered jointly with the randomized method of LEACH, CEED defines cluster head threshold accounting the remaining energy and proximity to the base station, hence, reduce the energy dissipation of nodes far from base station and balances the energy consumption of the network. R-LEACH, region-based LEACH protocol is defined by authors (Saha et al., 2019). R-LEACH propose dividing the network into rectangular sectors to prevent one major drawback of LEACH protocol results in proximity of cluster heads in the network and improve the clustering method in terms of lifetime and throughput. Machine learning based methods to extend operational time of WSNs is also investigated and Q learning method is extensively studied to improve clustering and cluster head selection performance in WSNs. (K. Li et al., 2019) introduced distributed clustering method, QLEC algorithm for energy efficient clustering problem for 3D WSNs. Every node in the network runs Q learning algorithm to select a cluster head as relay. The clustering of the network is compared with the performance of K-means and Fuzzy C-Means in terms of energy consumption and network lifetime. The reward function of QLEC designed to penalize unsuccessful transmissions to cluster head and information sharing about the energy cost of possessing cluster head role in terms of required energy. Q learning is also examined to enhance the performance of K-Means algorithm to provide optimal clustering and load balancing in WSNs (Sathyamoorthy et al., 2022). This study proposes two stages in the clustering of the network as clustering stage and node balancing stage. To alleviate the load on the cluster heads, author introduced Partition Heads (PH) as potential cluster heads and replace the exhausted cluster heads when their remaining energy is below a threshold determined priori.

Employing various optimization algorithms for cluster head selection is commonly used in the literature (Abraham & Vadivel, 2023a). In (Muthukkumar et al., 2022) a genetic algorithm is introduced for energy aware multi-hop clustering of heterogeneous WSNs. Heterogeneity in the energy levels is examined in this study. The optimal locations of cluster heads calculated based on node's residual energy and distance. The fitness function of this method considers the optimal distance of cluster heads, residual energy to avoid quick depletion of central nodes.

Although significant progress has been made in clustering techniques and cluster head selection methods as evidenced in existing literature, the energy of sensor nodes will inevitably deplete without an additional energy source. (ZHANG Gaoxi XIAO Hwee-Pink TAN et al., 2013) investigated various centralized and distributed clustering methods to optimize the lifetime of clustered WSNs via energy harvesting methods. They introduce a static clustering (clusters remain fixed once they are formed) and dynamic clustering algorithm and propose positioning of energy harvesting sensors. Their analytical results show the effect of the presence of nodes with energy harvesting capability on the lifetime of the network, which is defined as the certain amount of percentage of nodes in the network. Motivating with the energy harvesting contribution to sensor node lifetime, another research study, (Bozorgi et al., 2017) investigates the cluster head selection method based on the harvested energy by giving more chance to nodes observing larger energy harvesting rates that is based on random uniform distribution of lowest and highest possible energy levels. The authors introduced NEEC and employ a radius of competition, known as the neighborhood radius (cluster radius), for each node to establish clusters. Cluster formation occurs when nodes are selected as cluster heads based on their remaining energy and harvested energy. Additionally, the varied rate of harvested energy is considered in each node at the start of rounds. In (Sah & Amgoth, 2020) NEHCP a distributed hierarchical clustering protocol is introduced for EH-WSNs. The impact of two levels of solar power harvesting rate is examined in this study. The harvested energy rate is utilized in the cluster head selection probability calculation. Nodes use remaining energy and energy harvesting rates (based on low/high conditions) and decides acting as cluster head.

Prior research in the field encompasses centralized/distributed clustering methods employing optimization techniques and efficient search heuristics, alongside investigations into energy harvesting methods and the utilization of reinforcement learning algorithms to maximize sensor node operational time. In (Dwi Widodo HK et al., 2019), the authors proposed using intermediate nodes to forward data to cluster heads based on reinforcement learning, where rewards are determined by interference and transmission power. After clusters are formed, cluster heads calculate interference and broadcast this information to all nodes in the cluster. Nodes then decide whether to become an intermediate node based on the explore/exploitation threshold. In (M. Han et al., 2020), a decentralized reinforcement learning approach is introduced to enhance the network's throughput, with the goal of achieving sustainable data collection from EH-UWSNs. This work employs a multi-agent cooperative learning algorithm, specifically a multi-armed bandit approach, to optimize the contention window of the random-access medium while ensuring fairness within the cluster. However, it's worth noting that the reward function in this approach does not take energy harvesting into account; rather, it primarily focuses on enhancing throughput and fairness within the system. In (Ge et al., 2021), a reinforcement learning approach is applied to adapt to environmental changes in solar-powered WSNs. In this study, random cluster head selection is employed at the start of each round, with each node being trained based on its chosen data transmission rate. The objective is to maintain nodes in energy-neutral states while optimizing their data rates. This research highlights the significance of leveraging ambient harvesting resources to enhance the throughput and overall lifetime of a wireless sensor network using a reinforcement learning approach. However, it's important to note that in this study, cluster head selection is random, and learning is utilized solely to determine the optimal data rate selection for nodes. In (B. Han et al., 2022) an adaptive cluster-based routing protocol, namely HCEH-UC is introduced towards the aim of uninterrupted coverage of the area of interest utilizing energy harvesting cluster heads in EH-WSNs. Leveraging the hierarchical clustering, cluster formation and optimal number of clusters is defined as the confidence distance. The clustering of the network is terminated with calculation of confidence distance that is defined as the distance for maximum reliable data transmission, hence optimization of this distance as a threshold aimed to consider environment conditions.

Towards the aim of prolonging the lifespan of WSNs, energy management and resource allocation emerge as pivotal concerns. In the literature, another significant aspect of cluster-based data routing involves the use of autonomous robots for data collection, a concept extensively explored in wireless sensor networks. Various research endeavors have proposed the deployment of mobile sinks, including UAV robots, to efficiently collect data from clustered networks while optimizing the energy utilization of cluster heads (Gul & Erkmen, 2020, 2023). Towards the goal of extending the lifespan of WSNs, research in the literature primarily focuses on energy harvesting techniques along with energy management and resource allocation. Sustainable operation of WSNs by exploiting the ambient energy intelligently is attracting the researchers to explore adaptive and robust clustering methods.

2.2 Cluster Based Data Routing for UWSNs and EH-UWSNs

Given the constrained resources of sensors, the development of efficient data routing mechanisms for UASN presents a critical research challenge. (Sandeep & Kumar, 2017). Given that sensors operate on limited battery power, minimizing energy usage is crucial, particularly in challenging underwater communication conditions. Clustering has proven to be an effective strategy for enhancing network longevity by organizing nodes into groups (Y. Li et al., 2014; Sandeep & Kumar, 2017; Xing et al., 2021). This enables sensor nodes to communicate with a nearby sink and decrease their transmission power, thereby extending their operational lifespan. The efficiency of energy usage is regarded as a primary performance metric when evaluating the effectiveness of underwater wireless sensor networks. When tackling the challenging communication characteristics of underwater environments, minimizing energy consumption becomes a critical issue. This is because the energy storage capacity of sensors is constrained by the capacity of the battery. Utilizing clustering as a strategy has demonstrated to be an effective approach for prolonging underwater sensor network lifetime (Eris et al., 2023; M. F. Khan et al., 2021b; Sun et al., 2022b). Cluster head selection has been extensively investigated and has become the focal point of numerous studies in underwater wireless sensor networks. In these research endeavors, factors such as node localization, residual energy, distance from the surface station, and inter-cluster centrality have been considered to prolong the network's lifespan. The primary objective of these studies is to minimize energy consumption and alleviate the burden on the cluster heads to maintain an optimal operational level.

The research efforts for cluster-based data routing approaches in the literature can generally be classified into centralized and distributed categories. In centralized methods, a base station or gateway typically forms clusters and selects the optimal cluster heads. In distributed methods, nodes autonomously elect themselves as cluster heads, and cluster members join any cluster based on their proximity to the cluster heads within their communication range. In centralized methods, the BS requires the location and remaining energy of all nodes at the end of each round to efficiently generate clusters in terms of size and balance the energy consumption among sensors in the network. However, a centralized mechanism may not be able to obtain UASN node's energy states very timely, additionally computational cost of optimized data routing and relay selection methods may not adapt dynamic nature of underwater medium, hence frequent changes in the network cause usage of optimization methods challenging.

In distributed methods, nodes autonomously decide to become a cluster head considering its location, energy, proximity to the Base Station. Various methods in the literature have focused on enhancing the efficiency of LEACH for UWSNs. In (Domingo & Prior, 2007), a distributed clustering algorithm, DUCS, is proposed to identify optimal nodes for cluster head selection, alongside an efficient packet forwarding mechanism aimed at reducing delays, conserving energy, and enhancing data transmission reliability. In (Wu et al., 2021) a distributed hierarchical adaptive clustering method is introduced to reduce energy consumption of underwater acoustic sensor networks. Authors proposed to divide the 3D underwater sensor network into hierarchical sectors based on depth of sensors and defined competition radius to autonomously forms clusters varied in size. In (Nguyen et al., 2021) energy efficient cluster multi-hop protocol EECMR is introduced. Authors proposed a cluster head selection method based on the depth and remaining energy of sensors and defined a cluster relay to send aggregated signal from cluster heads reside in higher depths to the cluster heads located at lower depths.

Numerous studies have concentrated on various optimization methodologies for cluster head selection and cluster-based data routing (Fei et al., 2020; M. F. Khan et al., 2021a; Pengwei Li et al., 2017) recently. In (Fei et al., 2020), identifying the optimal cluster head is treated as an optimization problem. In (Fei et al., 2020) the moth-flame optimization method is employed to minimize the cost of an objective function, which accounts for factors such as the distance of the cluster head from the base station, energy depletion during data transmission/aggregation, and overall network energy consumption. Another approach utilized in cluster head selection algorithms is Particle Swarm Optimization (Pengwei Li et al., 2017), is integrated into cluster head selection algorithms as evolutionary-based computing and global search dynamics with the objective of identifying optimal nodes for clustering. (M. F. Khan et al., 2021a) introduced an adaptive clustering technique ANC-UWSNs with dragon fly optimization algorithm to produce optimal number of clusters and investigate the effect of cluster size, node density and sensor's communication radius to determine optimal cluster heads.

Table 1

Comparison of Related Research Studies

Related Studies	Distributed	Energy Harvesting (EH)	Clustering	Machine Learning Methods (ML)	Stochastic Energy Harvesting (EH)
(Dwi Widodo HK et al., 2019; Zhu et al., 2021)	✓	×	✓	✓	×
(Sun et al., 2022b)	✓	×	✓	×	×
(Erdem et al., 2019; Erdem & Gungor, 2017)	✓	×	✓	×	×
(Ge et al., 2021; M. Han et al., 2020)	✓	✓	×	✓	✓
(Fei et al., 2020; Narayan et al., 2022; Omeke et al., 2021; Pengwei Li et al., 2017)	✓	×	✓	×	×
(Ahmed et al., 2017)	×	×	✓	×	×

Table 2 (cont'd)

(Bozorgi et al., 2017; Eris et al., 2023; Ren & Yao, 2021)	✓	✓	✓	×	✓
(Eriş et al., 2024)	✓	✓	✓	✓	✓

Machine learning methods have seen increasing adoption in underwater wireless sensor networks (Uyan et al., 2023). The clustering technique known as K-Means has particularly gained significant attention recently (Eriş et al., 2024; Fujino et al., 2019; M. Han et al., 2020; M. F. Khan et al., 2021a; L. Li et al., 2022; Omeke et al., 2021, 2023; Sun et al., 2022b; Zhu et al., 2021). In (Omeke et al., 2021) introduced dynamic clustering protocol, DECKS, and proposed the selection of cluster heads based on the density of nodes within the network, as well as the distance between each node after dividing the network with the k-means algorithm. Another clustering strategy ECRKQ is based on k-means, proposed by the authors of the study (Zhu et al., 2021) and Q-learning is utilized for cooperative routing in underwater sensor networks. This method takes the remaining energy of nodes and their proximity to the centroids of their respective clusters into account. After the cluster head selection step, nodes compute a reward function that considers both the energy required for communicating with the cluster head and the energy cost of cluster heads communicating with the base station. A multi agent reinforcement learning based adaptive clustering method (Sun et al., 2022b) is investigated to find cluster heads in UWSNs. Sun et al., proposes autonomous selection of cluster heads based on the channel conditions. Nodes adaptively establish routes by learning the channel conditions and adjusting their routing strategies accordingly. (L. Li et al., 2022) introduced clustered routing algorithm LE-KCR considering the location and energy of nodes in K-means clustering of the network. A noteworthy novelty in their study is utilization of optical links and calculation of optimal number of clusters with gap statistical method instead of well-known elbow method in clustering.

Significant achievements in clustering and cluster head selection methods for

UASNs is published in the literature. On the other hand, energy harvesting methods, have undergone extensive research and their potential is investigated to improve sustainability of UASNs (Alamu et al., 2023; Erdem et al., 2019; Zhao et al., 2021). For instance, Toma et al. devised an underwater energy harvesting system utilizing plucked-driven piezoelectric, employing a cantilever-like structure with a piezoelectric material to generate electricity from water currents. Harnessing underwater energy have been extensively explored and studies demonstrated their potential to enhance the lifespan of underwater sensor nodes (Pobering & Schwesinger, 2004) in the literature. Furthermore, recent studies present harvesting ambient energy source is a promising method to improve the life cycle of underwater sensors (Erdem et al., 2019; Erdem & Gungor, 2017; Pobering & Schwesinger, 2004; Zhao et al., 2021; Zou et al., 2021). These studies have demonstrated that UASN applications must optimize energy management in routing by using ambient resources to capture energy to enhance sustainability of data collection in underwater sensor nodes. The main obstacle in tackling the distinct features of submerged communication channels is the constraints of sensing apparatus and harvesting machinery. Numerous studies have aimed to prolong the life cycle of network nodes by leveraging energy harvesting (Bereketli & Bilgen, 2012; McKenna et al., 2012; Vermaak et al., 2014a; Zou et al., 2021). In (Erdem et al., 2019), the joint optimization of compressive sensing (CS) and various harvester models' impacts on the lifetime of Underwater Acoustic Sensor Networks (UASNs) are analyzed. A noteworthy study is conducted by Basagni et al., (Basagni et al., 2018), introduced a data routing method to choose the optimal forwarding relay by considering both residual energy and anticipated tribune harvester energy. The relaying method, HyDRO (Harvesting-aware Data Routing), is designed to account for channel condition and overall route residual energy via local information sharing. The impact of energy harvesting contributes to the network lifetime remarkably, compared to other protocols lacking harvesting energy is examined in their study. Yet, HyDRO is analyzed in a network of 20-40 nodes in size, does not examine the clustering of the network, and the data relay relies on either multi-hop or single-hop routing. Similar relay selection method is conducted by (A. Khan et al., 2021) assessing the discovery of harvested energy from broadcast transmissions using piezoelectric harvester technology via hydrophones and performance results show that the impact of harvested energy is enhanced through intelligent relay calculation.

Table x provides the gap analysis in the related studies. Aiming to prolong the lifetime of UWSNs, energy management and resource allocation arises as the main issues. In the close literature, studies focus on centralized and distributed clustering strategies incorporating optimization and efficient search heuristics while others steer for energy harvesting methods and utilize learning algorithms to enhance the duration of operational time of sensor devices. Unlike other approaches in the literature, our study, (Eriş et al., 2024) proposes a reinforcement learning algorithm to determine cluster heads considering the stochastic nature of underwater environment and piezoelectric energy harvesting. Recent studies have applied reinforcement learning methods to utilization of energy in underwater environments. Diverging from existing literature, we implement a reinforcement learning algorithm for cluster head selection, integrating stochastic model of piezoelectric energy harvesting. Our approach differs from existing studies in the literature as it introduces cluster head selection to maximize accumulated harvested energy considering stochasticity of available energy in surrounding environment of nodes. .

2.3 MAC Protocols for Clustered UWSNs

There is a growing literature about medium access strategies to improve energy efficiency in underwater wireless sensor networks. In wireless sensor networks the aim of MAC protocols is to decrease latency in the network and focus on the management of high traffic and is extensively studied (Gabale et al., 2013). These efforts can be categorized as contention-based and contention-free protocols. In contention-based protocols, all nodes can transmit data packets simultaneously without any predefined scheduling algorithm. In (Y. Chen et al., 2023), a contention-based MAC protocol, TDTSPC-MAC is developed for three-dimensional Underwater Acoustic Sensor Networks (UASNs) facing challenges such as long propagation time, high bit error rate, and limited bandwidth. The protocol integrates time synchronization, power control, clustering, layering, and sleep mechanisms. TDTSPC-MAC utilizes hierarchical division of three-dimensional space and a distributed clustering algorithm, combining time synchronization and power control strategies to prevent collisions and reduce energy consumption through monitoring and sleep mode. Yet, in scenarios involving extensive data transmission or a significant number of nodes, data packets become susceptible to collisions. Large transmission delays and limited transmission rates inherent in underwater sensor networks increase collision probability, result in low energy utilization. In (Tomovic & Radusinovic, 2023), a delayed-reward ALOHA-Q protocol is designed for Underwater Acoustic Sensor Networks (UASNs). By leveraging Q-learning, DR-ALOHA-Q autonomously determines optimal time slots and offsets for data transmissions to improve channel utilization in the presence of large propagation delays. Contention-based methods require active listening and short control packet exchanges to mitigate long idle waiting times and improve overall energy consumption and transmission delay in UASNs. In (Alfouzan et al., 2019), challenges in underwater sensor networks (UASNs) related to high latency, low bandwidth, and high bit error rate in acoustic communication is investigated. It highlights the limitations of contention-based medium access control (MAC) protocols in UASNs and proposes a novel contention-free MAC protocol called DL-MAC. The DL-MAC protocol addresses underwater MAC challenges through depth-based layering, a distributed clustering algorithm, and TDMA-based scheduling, demonstrating superior performance in terms of throughput, packet delivery ratio, energy consumption, and packet loss.

On the other hand, contention-free MAC protocols can easily address the well-known hidden node problem, without extra control packet overhead and interference of neighboring nodes by utilizing scheduling nodes to distinct time slots to perform transmissions. A well-known contention-free MAC protocol, TDMA attracted researchers while other contention-free protocols such as FDMA suffers limited available bandwidth and CDMA faces challenges due to the near-far problem (Pompili et al., 2009). Hence TDMA-based MAC protocols have been the focus in the listerature for long-term surveillance application. Numerous TDMA-based medium access methods are proposed (Y.-D. Chen et al., 2011; G. Han et al., 2020; Ying. Wang et al., 2009) to leverage the benefits of its simplicity, fairness, and energy efficiency by mitigating issues such as collisions, idle listening, and transmission over-hearing. However, TDMA-based MAC protocols exhibit certain drawbacks, such as suboptimal channel utilization, challenges in dynamic scalability, and the need for time synchronization. In (Wang et al. 2017) Wang et al. proposed an online scheduling algorithm to calculate the optimal action, namely, decision of transmitting the packet or holding it back in the packet queue in each time slot with the aim of decreasing energy consumption in packet transmission. In this work, channel state prediction algorithm is employed to determine the optimal action. Aiming to minimize energy consumption and considering the channel conditions in transmissions, a transmission strategy is proposed in (Tomasi & Preisig, 2015). In (Lee & Jindal, 2009) suggested channel aware transmission scheduling considering duration of each slot and channel conditions and adjusting the amount of data to be transmitted. In (Ye et al., 2022), a deep reinforcement learning (DRL) algorithm is introduced, Delayed-Reward Deep Q-Network (DR-DQN), to improve throughput degrades due to long propagation delays in underwater acoustic networks (UANs). Delayed-Reward DRL Multiple Access (DR-DLMA) is proposed aiming towards optimal channel access strategy maximizing network throughput by efficiently utilizing available timeslots resulting from long propagation delays. In (Liu et al., 2021), authors suggest prediction of the channel state by capturing temporal and frequency correlations in one-dimensional convolutional neural network (CNN). In (Jagannath et al., 2013) a time slot allocation scheme for clustered underwater acoustic sensor networks, utilizing physical layer information to minimize energy consumption caused by unnecessary retransmissions. This enhancement aims to extend network lifetime and improve throughput while streamline the process and reduce overhead and computational complexity. In their two-phase approach, first, each member node autonomously determines the number of required time slots for the next intra-cluster cycle by solving a Markov decision process (MDP). Second, the cluster head optimizes scheduling decisions based on channel quality and an urgency factor.

Despite ongoing research efforts, existing studies mainly focus on the improvement of throughput and energy utilization considering only the channel state and fairness. On the other hand, the impact of energy harvesting in underwater medium is proved its importance in design of the communication protocols by extending the operational lifetime of underwater sensor devices considerably (Erdem et al., 2019; Erdem & Gungor, 2017). In the study (Gül, 2014), authors explore data collection in a single-hop wireless sensor network, employing a myopic policy based on a Round Robin structure. Findings reveal the sub-optimality of this approach for various energy harvesting processes, emphasizing limitations yet achieving optimality in specific instances. A noteworthy study conducted by Ge, Y. et.al.(Ge et al., 2021; M. Han et al., 2020), focused on the improvement of throughput and energy harvesting for solar-powered and clustered WSNs. This work proposed a multi-agent reinforcement learning algorithm to maintain nodes in an energy neutral state by balancing consumed and harvested energy of nodes. Yet, this study focuses on the WSNs, peculiar characteristics of underwater and the amount of available ambient energy needs to be considered for UASNs. A decentralized reinforcement learning approach for random access is proposed to improve throughput of the network considering energy harvesting for sustainable data collection from EH-UASNs (M. Han et al., 2020). In this work, multi-agent cooperative learning algorithm is used in adaptive choice of contention window size to optimize throughput while maintaining fairness in the network. However, authors focused on the throughput and the reward function considers only the fairness and throughput in the network. Maximization of the energy harvesting utilization is not investigated in this study.

When considering the clustered Underwater Acoustic Sensor Networks (UASNs) and TDMA-based protocol for intra cluster communication, it is imperative to allocate resources effectively to ensure reliable and efficient operation. Therefore, time slot allocation methods for underwater sensor networks need to be robust and accurate, as this greatly affects communication efficiency. In our study, we focus on optimal scheduling method for TDMA-based MAC protocol in clustered UASNs. While in the literature, research efforts focus on clustering and routing techniques to improve energy utilization in underwater sensors. Although several studies address lifetime maximization by balancing the data transmission rate adaptive to energy harvesting conditions and medium access opportunities, to the best of our knowledge there is no research study aiming to maximize harvested energy based on the channel access strategy. In our method we investigate dynamically adapting transmission slot choices based on learning from past slot selection rewards allowing for a more adaptive and learning-oriented approach to transmission slot selection.

Chapter 3
Cluster Based Data Routing for Energy-Harvesting Underwater Acoustic Sensor Networks

Technological advances in wireless sensor networks and seamless integration with diverse technologies have led to the proliferation of underwater wireless sensor networks and their application areas. Submarine missions are increasingly utilizing underwater sensor gear (Awan et al., 2019). Numerous vital applications, such as early warning systems, search and rescue missions, and ecosystem monitoring, can be carried out by these devices. With the collection and analysis of data in aquatic environments, sensors are crucial assets for intelligent systems that enable autonomous decision-making and data analysis. To fully realize the potential of the undersea realm, research focuses on addressing the limitations of the underwater medium to enable sustainable data collection in such applications. Underwater sensor nodes primarily employ acoustic communication technology since it can transmit data over long distances across water. However acoustic waves have high bit-error rates and route loss, therefore sensors need large transmission powers to avoid packet loss (Akyildiz et al., 2005b). Energy management is crucial for building UASNs since it raises concerns for maintenance and human intervention because sensor nodes are typically found in sub-sea zones and frequently run on batteries.

The use of autonomous robots for data collection is another interesting viewpoint, and it has been thoroughly investigated. By minimizing the energy consumption of cluster heads and UAV robots, several research projects suggested mobile sinks for data gathering from clustered networks (Abraham & Vadivel, 2023b; Gawade & Nalbalwar, 2016; Gul & Erkmen, 2020, 2023). Yet, in comparison to terrestrial wireless sensors, the underwater system has higher financial costs even though the assistance of autonomous robots lowers the energy cost of sensor equipment.

To maximize UASN lifetime, several studies in the literature focus on improving network clustering and cluster head selection methods (Domingo & Prior, 2007; Fei et al., 2020; M. F. Khan et al., 2021a; Mathankumar et al., 2021; Omeke et al., 2021; Sandeep & Kumar, 2017; Sun et al., 2022b; Xing et al., 2021; Zhu et al., 2021) There has been much effort invested in enhancing the energy efficiency and extend the lifetime of the underwater sensor networks. Energy harvesting is a promising solution to recharge the battery of sensor devices. Various technologies utilized to convert the different forms of ambient energy existing in the underwater environment into electricity. Piezoelectric harvesters are one of the most reliable options to increase the amount of energy supply to sensor nodes. However, the amount of harvested energy by piezoelectric bimorphs varies greatly depending on the underwater environment conditions. To address the highly varying source of energy we propose a realistic model for piezoelectric energy harvesting. In our work we use k-Means clustering of the network and propose energy harvesting aware cluster head selection method to prolong the operational lifetime of UASN nodes.

3.1 Background

Acoustic sensor nodes are deployed at the sea bottom and used in long-term underwater surveillance applications. These nodes collect sensor measurements, which are then transmitted to a fixed sink station located on the sea surface (Akyildiz et al., 2005a). Data transmission from the nodes to the sink can occur via single hop. All nodes are equipped with energy harvesting capabilities, but they can only harvest energy while in idle mode. UASN network comprises 100 sensor nodes (N_{node} = 100) distributed randomly within a 3D cube area measuring 250×250×250 meters (in a uniform distribution). At the sea surface, there is a single sink node positioned at the center top, and the network extends to a depth of 250 meters. The 2D static network architecture is illustrated in Figure 1. Given that the nodes are stationary at the ocean bottom, we assume that their positions are fixed over time.

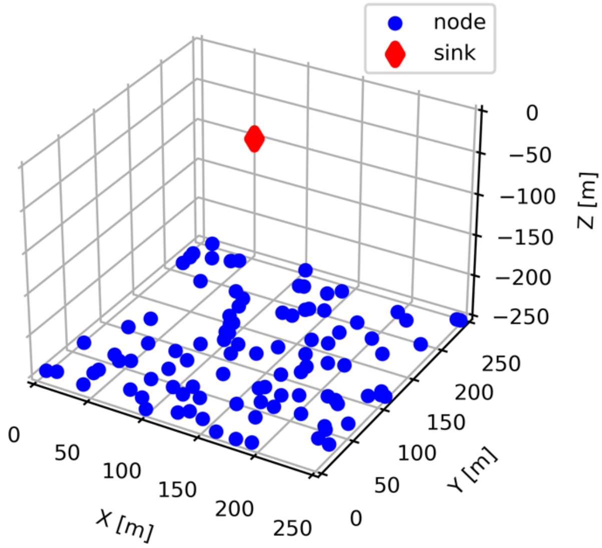

Figure 1. 2D Sample network architecture showing sink and node positions.

Cluster heads collect data within their respective clusters and transmit aggregated data to the sink. This process occurs over consecutive rounds, with each round comprising frames. During a frame, each node sends a single packet containing L bits. It's important to note that nodes consume energy during transmission, reception, and data aggregation stages. The time taken for packet transmission and reception is denoted by $\tau_{slot} = L/R_b$, where R_b represents the bit rate of the sensor nodes. Due to the significant propagation delay and constrained bandwidth of underwater channels, nodes utilize Time Division Multiple Access (TDMA) for communication within clusters (Molins & Stojanovic, 2006). Consequently, nodes engage in data transmission within their designated time slots, as scheduled by cluster head, and remain idle until the next frame in the round. During this idle period, all nodes harvest energy. In other words, nodes harvest energy all time except their designated time slot to send a data frame. Hence, knowing the idle duration of each node is essential for calculating the amount of energy generated in a single frame. The duration of a single frame varies according to the cluster size, where $\tau_{frame} = \tau_{slot} \times Q$, with Q representing the number of nodes in a cluster. Consequently, the idle time of nodes can be expressed as $\tau_{idle} = \tau_{frame} - \tau_{slot}$. The duration of a round is predetermined, on average, to allow each node to behave both as cluster head and a non-cluster head multiple times during its lifespan (W. R. Heinzelman et al., 2000). Consequently, each cluster in the network may have a different number of data frames in a single round.

3.1.1 Underwater acoustic propagation model. Underwater acoustic communication is significantly impacted by factors such as path loss, noise, multiple paths, Doppler spread, and variable propagation delay (Akyildiz et al., 2005a). These elements contribute to the temporal and spatial uncertainty of the acoustic channel, resulting in a limited bandwidth for the underwater acoustic channel (UW-A) that is heavily influenced by distance, time, and frequency. For instance, long-range systems spanning tens of kilometers may only possess a bandwidth of a few kilohertz, whereas short-range systems covering tens of meters may offer over 100 kHz of bandwidth. Despite the variation, these factors collectively result in a relatively low bit rate (J. A. Catipovic 1990), typically in the range of tens of kilobits per second.

The challenges posed by underwater channels for underwater sensor networking are a result of various factors that impact acoustic communications. These factors include path loss due to attenuation and geometric spreading of acoustic signals in water, ambient noise from water movement, such as tides, currents, storms, wind, and rain. Multipath propagation can significantly degrade acoustic communication signals by causing inter-symbol interference (ISI). The extent of multipath spreading varies based on link configuration, such as vertical or horizontal channels. Vertical channels exhibit minimal time dispersion, while horizontal channels can have extensive multipath spreads, influenced by depth and distance between transmitter and receiver (Freitag et al., 2005).

3.1.2 Physical layer model. Acoustic communication experiences significant degradation due to attenuation and absorption, especially over considerable distances. In contrast to radio models, where bandwidth is considered constant regardless of the distance between the transmitter and receiver, thus eliminating the need for additional information from other layers, underwater networks experience variable link bandwidth based on link length. To compensate, the source node adjusts its power output based on the distance to the destination node. Sonar equations are employed to define this relationship, and transmission power is determined in relation to sound pressure (E. Felemban, 2020) as

$$SL(d,f)_{ij} \ = \ A(d,f)_{ij} + N(f) + SNR_{dB} + DI \tag{1}$$

where $SL(d,f)_{ij}$ represents the ratio of sound intensity produced by a node i at a distance *d* (in meters) using an acoustic signal frequency *f* (in kHz), measured in terms of acoustic pressure referenced to dB ref μ Pa. $A(d,f)_{ij}$ denotes path loss of acoustic waves which escalates with the inter-distance d (in meters) between nodes i and *j* measured in dB (Stefanov & Stojanovic, 2011b) as

$$A(d,f)_{ij} \ = \ 10\kappa \, log_{10} d_{ij} + \alpha(f)\, d_{ij} 10^{-3} \tag{2}$$

where κ is spreading factor, $\alpha(f)$ is absorption coefficient in terms of dB/km, determined (Jornet et al., 2010) as

$$A(d,f)_{ij} \ = \ \frac{0.11f^2}{1 + \ f^2} + \frac{44f^2}{4100} + f^2 \ + 2.75 \, x \, 10^{-4} \, x \, f^2 \\ + \ 0.003 \tag{3}$$

$$N(f) \ = \ N(f)_t + \ N(f)_s \ + \ N(f)_{th} \ + \ N(f)_w \tag{4}$$

The ambient noise level, measured in dB ref μ Pa, is determined by aggregating the noise contributions from various sources, including water turbulence ($N(f)_t$), surface ships ($N(f)_s$), thermal activities ($N(f)_{th}$), and breaking waves ($N(f)_w$) (Stefanov & Stojanovic, 2011b). To maintain a predetermined Bit Error Rate (BER) on a link, the Signal-to-Noise Ratio (SNR) can be computed for 16-QAM modulation employing Orthogonal Frequency Division Multiplexing (OFDM) encoding as (M. Felemban & Felemban, 2013)

$$p^b{}_{ij} \ = \ \frac{3}{8} erfc\left(\sqrt{\left(\frac{4}{10}\frac{B_N}{R}\gamma(d)_{ij}\right)}\right) \tag{5}$$

where B_N represents the noise bandwidth (in Hz), $\gamma(d)_{ij}$ denotes the Signal-to-Noise Ratio (SNR) in linear form, and R represents the data rate in bits per second (bps). The electrical transmission power is then determined based on the sound intensity level at a given distance (I^t_{ij}) at 1 meter is calculated (E. Felemban, 2020) as

$$P^{tx}{}_{ij} = I^{tx}{}_{ij}\ x\ 2\pi\ x\ 1m\ x\ h_i \tag{6}$$

whereh_i is depth of node i in meters and the intensity is calculated based on the source level and reference intensity I_0 (E. Felemban, 2020) as

$$I^{tx}{}_{ij} = 10^{\frac{SL(d,f)_{ij}}{10}}\ x\ I_0 \tag{7}$$

The reference intensity I_0, is determined as 0.67 x 10^{-18}. Transmission power control methods are applicable in scientific modems. To estimate the underwater channel, typically probe packets are employed (Akyıldız Fuat & Vuran Can, 2010). The path loss of an acoustic signal $A(d, f)$ which depends on frequency f and distance, d is computed using equations 2 and 3. Ambient noise, $N(f)$ is determined in equation 2 and adopting attenuation and noise in equation 1, we can find the required transmit power output P^{tx} to satisfy particular SNR at the receiver with equations 5 and 7. Figure 2 illustrates the effect of carrier frequency on the attenuation and ambient noise of a transmitted signal. It is found that the center frequency is around 20-40 kHz varying with the distance.

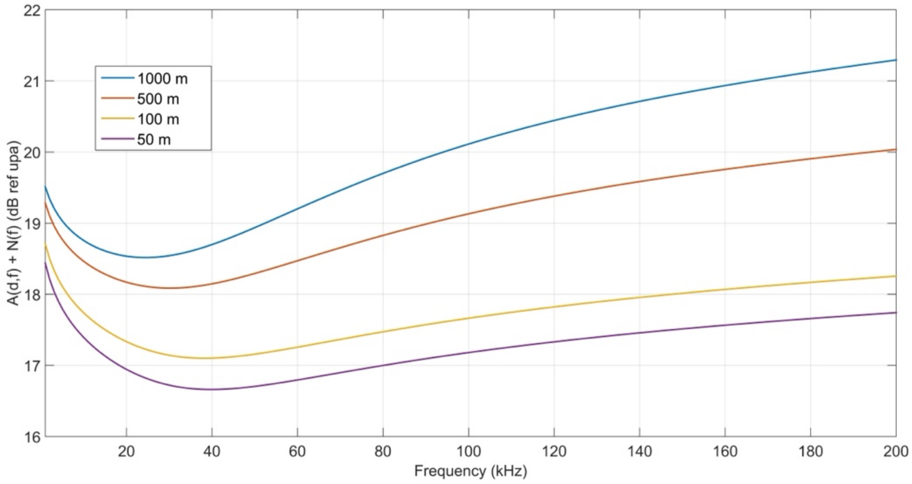

*Figure 2.*Attenuation A(d,f) and noise N(f) over frequency

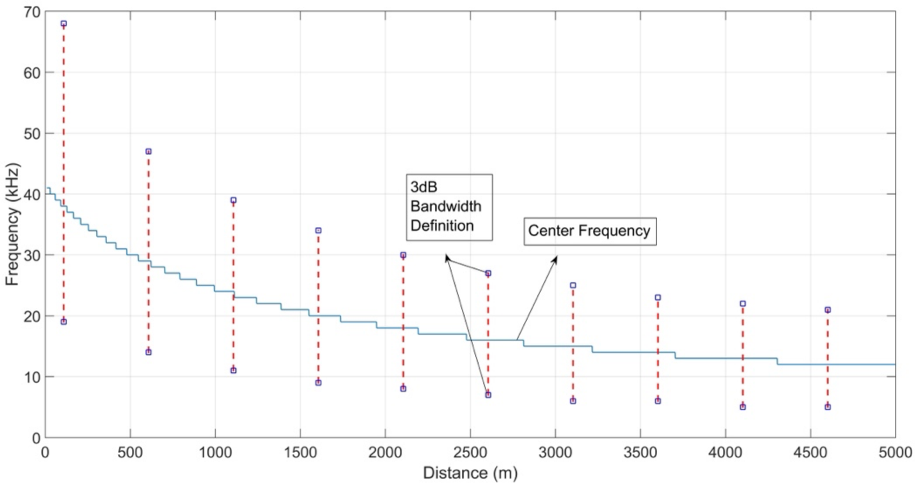

Figure 3. Center frequency and effect of distance on available bandwidth

The available bandwidth depends on the distance between the transmitter and receiver. Initially, the distance between the transmitter and receiver is used to identify the frequency with the minimum attenuation-noise (AN) factor. This frequency is defined as the center frequency for communication. Subsequently, the 3 dB definition of bandwidth is applied to determine the edges of the usable frequency band. Figure 2 shows available bandwidth and center frequency varying over distance with 3 dB bandwidth definition (Jornet et al., 2010). The line represents the center frequency and vertical lines represent available bandwidth. It is evident that available bandwidth decreases as the distance between the transmitter and receiver increases. In Figure 3 required transmission power over varying distances between two communication endpoints is denoted. In this figure, the required transmission power of central carrier frequency is calculated according to equations 2, 5, 6 to satisfy a certain SNR on link between two nodes varying with the distance.

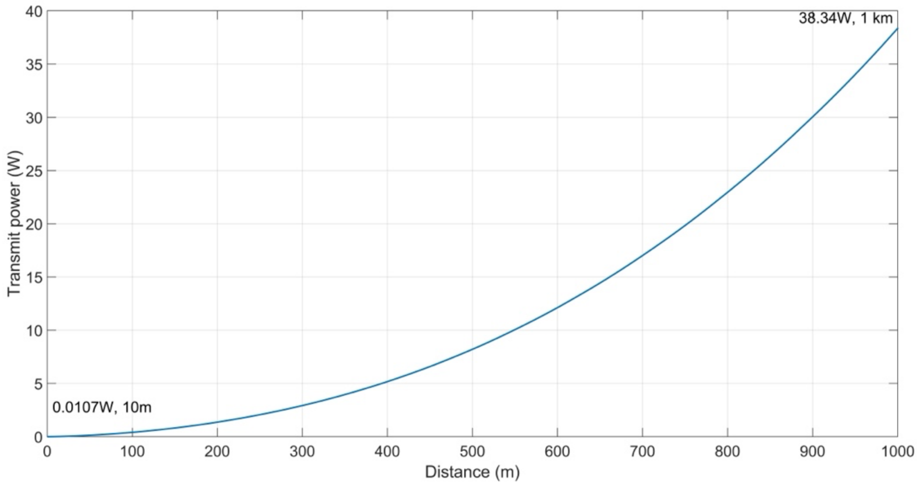

Figure 4. The effect of distance on transmission power

In commercial acoustic modems configuration provides adjustment of varied level of transmission power output. Table 1 gives commercial modem specifications of different manufacturers utilized by research community in the literature. Their power dissipation profile on idle, receiving (Rx), transmitting (Tx) modes is given. In scientific modems, it is possible to adjust transmission power based on the distance to the receiving end. This adjustment ensures a certain Signal-to-Noise Ratio (SNR) threshold is maintained hence, optimizing the power according to the distance. In Figure 4, the impact of distance on the necessary transmission power is depicted. The figure illustrates that the transmit power is below 1 W for distances of hundreds of meters, but it exceeds 30 W when the receiver is 1000 m away from the source node.

Table 3

Acoustic Modem Specifications and Power Profiles

Modem	Idle	Rx	Tx	Frequency Range	Modem Range	Data Rate
WHOI[1] Micromodem	0.63 – 2.93 W	0.63 – 2.93 W	Max 50W	9-15-25-4 kHz	1.5 km	80 bps – 5 kbps
Teledyne Benthos[2] (ATM-900)	N/A	N/A	50 W	9-27 kHz	2 - 6 km	80 bps (fh) 2.4 kbps(mfsk) 15 kbps(psk)
LinkQuest UWM1000[3]	8 mW	0.75 W	2 W	26.77 to 44.62 kHz	350 m	9.6 – 19.2 kbps
LinkQuest UWM3000H[4]	8 mW	0.8 W	3 – 12 W	7.5 - 12.5 kHz	3 km 6 km	80 – 320 bps
EvoLogics S2C R 18/34 Wise[5]	2.5 mW	1.3 W	2.8 – 8 – 35 – 65 W	18-34 kHz	3.5 km	13.9 kbps
EvoLogics S2C M HS[6]	0.5 mW	5 – 285 mW 0.8 W	3.5 W (0 - 200m) 8.5 W (max)	120 – 180 kHz	300 m	62.5 kbps

3.2 Energy Harvesting Methods

As low-power electronics continue to evolve with more efficient circuitry, the energy requirements have progressively diminished, making energy harvesting systems increasingly viable for powering self-sustaining systems. Of the various ocean regions with harvestable energy potential, areas with tidal and ocean currents, as well as shallow water regions experiencing significant wave-induced oscillatory currents at the sea floor, are particularly noteworthy. Wave-induced oscillatory currents at the sea bottom are intriguing for harvesting devices due to the dynamics of bodies placed in these regions. The movement of water displacement in gravity waves typically follows a circular motion.

3.2.1 Hydrokinetic tribune energy harvesting model. Hydrokinetic turbine harvesters are employed to transform the kinetic energy present in water flow into electrical energy. The potential energy harnessed from a water flow is determined by factors such as flow velocity (v_f), water density (ρ), and the surface area of the turbine (A) (Vermaak et al., 2014b). The overall power accessible from the turbine is computed as

$$P^{harv}{}_{tribune} = \frac{1}{2}\ x\ \rho \mathrm{\ x\ A}\ x\ v_f{}^3 x\ \eta_t \qquad (8)$$

where $A = \pi r_d^2$ is wing spanning area and r_d is wing radius η_t is power coefficient of tribune harvester, conversion fraction of total harvested power due to losses involved. v_f denotes the velocity of the fluid, experimental data can be retrieved in Global Oceans Currents Database (GOCD) and studies conducted in local areas (Jarosz et al., 2011).

3.2.2 Stochastic model of piezoelectric energy harvesting. The piezoelectric harvester converts the energy derived from water flows. It capitalizes on the turbulence within the flow, which generates vortices exhibiting opposite directions and varying rotational speeds. Variations in velocity on either side of the cantilever induce a pressure differential, prompting the cantilever to flutter and produce electrical energy. The underwater environment is characterized by high levels of variability and uncertainty. Sensor nodes are distributed across various depths, encountering a range of environmental factors, which causes variations in the produced electrical energy. The power output of a piezoelectric harvester is determined as

$$P_{harv} = 2 f_v (w_{el_{cw}} + w_{el_{ccw}}) \tag{9}$$

where f_v, $w_{el_{cw}}, w_{el_{ccw}}$ indicate vortex frequency, electrical energy in clockwise and counterclockwise directions respectively. The vortex frequency depends on the velocity of the flow v_f, bluff size (D), Strouhal Number (S_t) is determined as

$$f_v = \frac{S_t \ x \ v_f}{D} \tag{10}$$

Pressure difference in clockwise and counterclockwise directions are found using density and velocity of flow

$$\rho_{cw} = \frac{\rho}{2} {v_f}^2 \qquad \rho_{ccw} = -\frac{3}{2} \rho {v_f}^2 \tag{11}$$

The electrical energy produced by piezoelectric harvester varies with the cantilever specifications, piezo-material and pressure difference that is calculated as

$$w_{el} = \frac{1}{128}\rho^2 \, d_{31}^{p^2} \, \varepsilon_0 \varepsilon_r \frac{BL^5}{T_{pzt}^3} \tag{12}$$

where $d_{31}^{p^2}$, ε_0, ε_r represent piezoelectric constant, absolute and relative permittivity of piezoelectric material respectively. As flow speed varies with respect to time, available energy considering the past generated pattern can be predicted. B, L = 2.1D and T_{pzt} represent the cantilever width, length and thickness respectively. We utilize the Poisson distribution to model the fluctuations in flow velocity, accounting for the turbulence in the flow patterns. The average velocity of the flow is observed across various depths within the Bosphorus strait (Jarosz et al., 2011). Adapting the average velocity with Poisson mean inter-arrival rate, the energy harvested by a node i over time interval T is computed as an integral of power output over time for $t = 1 \ldots T$ as

$$E_i^h(T) = \frac{1}{128}\rho^2 \int_{t=1}^{T} 2f_v(t)\left(w_{el_{cw}}(t) + w_{el_{ccw}}(t)\right) dt \tag{13}$$

3.3 K-means Clustering and Our Cluster Head Selection Policy

Clustering of UASNs namely consists of two stages, cluster formation, cluster head selection. In cluster formation stage, our implementation first divides the network in k clusters based on Euclidean distance with K-Means algorithm. Following the cluster formation, cluster heads are selected based on residual energy and the predicted harvestable energy of nodes. Figure 5 illustrates the timeline of these stages. Cluster heads gather data from members within their clusters, perform data aggregation, and send the aggregated data to the surface station. In our approach, when a cluster head becomes exhausted, it computes a new potential candidate based on location, remaining energy, and estimated harvested energy to replace itself in the subsequent round. The selection of a new cluster head is triggered when the remaining energy of the current cluster head is inadequate for the next round of data collection. Figure 6 shows the flow diagram of our method, cluster formation (Set-up Phase) in the network and the rotation of cluster heads phase followed by data transmission phase (Steady Phase) in a round.

3.3.1 K-means clustering. The k-means algorithm partitions a sensor network into k clusters by iteratively assigning nodes to clusters and updating cluster centroids to minimize intra-cluster distances while implicitly aiming to maximize the distinctiveness of clusters. This results in a non-overlapping, optimized clustering of sensor nodes that enhances the overall efficiency of the network. For a sensor network with N nodes, the aim of the k-means algorithm is to divide the network into k clusters by minimizing the Euclidean distance between the nodes and their respective cluster centroids (intra-cluster distance) and maximizing the distance between different cluster centroids (inter-cluster distance).

$$J_{min} = \sum_{i=1}^{k} \sum_{c_j \in C} \left\| x_j - \mu_i \right\|^2 \tag{14}$$

where every cluster, c_i contains N_i nodes, x_j and μ_i represents j th node in the cluster i and the centroid of i th cluster respectively. The k-means centroids of each cluster in a 3D sensor network are determined by

$$\mu_i = \left(\frac{1}{N} \sum_{i=1}^{N} x_j , \frac{1}{N} \sum_{i=1}^{N} y_j , \frac{1}{N} \sum_{i=1}^{N} z_j \right) \tag{15}$$

In our approach, the clusters are organized in a way that enables each node to communicate directly with its cluster head.

3.3.2 Cluster head selection policy. Clustering is carried out prior to selecting cluster heads to minimize the energy expenditure involved in the cluster formation process. In the first round of the simulation, k-means clustering of the network is followed by the calculation of initial cluster heads. Cluster heads rotated in clusters and triggered by exhausting cluster head. The policy of cluster head selection involves four components, proximity to the cluster centroids, to the surface gateway, remaining energy of node and estimated harvested energy of node in the next round. Every node in the cluster calculates its cluster head score based on these factors and decide to become cluster head in the next round.

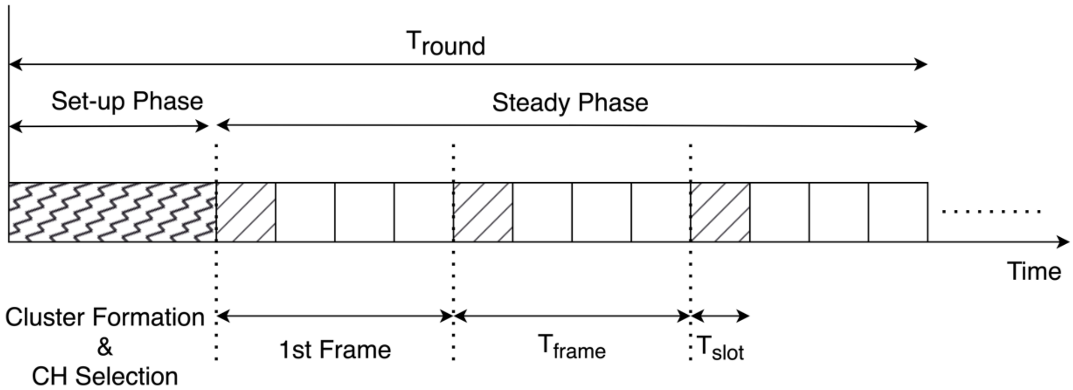

Figure 5. Timeline of round-based data collection in clustered networks.

To calculate the total anticipated energy cost in the cluster in case of selecting node *n* as a cluster head in the next round for a single data frame is determined (W. B. Heinzelman et al., 2002) as

$$E(CH) = \tau_{slot} x [P_{rx} Q + P_{tx}(d(n, GW), f)] + L \; x \; E_{DA} Q \qquad (16)$$

where τ_{slot} is the duration of transmission of a single data frame, Q is the number of nodes in the cluster, GW indicates the surface sink station, L is the packet size in bits, and $P_{tx}(d(n, GW), f)$ is the transmission power required to transmit a data frame successfully to the GW at a distance, $d(n, GW)$. f is the center frequency in kHz and E_{DA} is data aggregation energy in terms of Joules per bit. The total energy consumed by a cluster member node q to transmit a single packet to its cluster head node, n is determined as

$$E(n) = \tau_{slot} x \; P_{tx}(d(q, n), f) \; q \in C, q \neq n \qquad (17)$$

where $d(q, n)$ is the distance of cluster member node q to its cluster head n and C is the set of nodes in this cluster. The total number of frames, N_{frame}, transmitted to the cluster head varies according to the size of each cluster and as round duration is fixed and predetermined, calculated as $N_{frame=}\tau_{frame}/\tau_{slot}$. The estimated energy usage of both the cluster head and member nodes is evaluated internally. Nodes compute the potential energy expenditure within the cluster when operating as a cluster head, which is calculated as

$$E_{est}(n) = E_{CH}(n) + \sum_{q=1}^{Q} E(q) \quad q \in C, q \neq n \tag{18}$$

where $E_{est}(n)$ signifies the energy expenditure of the cluster head node n and the total energy usage within the cluster during a single frame, including data transmission, reception, and aggregation tasks. By predicting the remaining energy of the node with the total energy consumption of members within the cluster, the optimal selection of the cluster head is ensured among candidates possessing adequate energy levels. In the final stage, nodes forecast the available harvested energy, with cluster head node n being elected based on the highest calculated score as

$$CHScore(n) = \frac{E_{est}(n) + E_{est}^{harv}(n)}{E_{resd}(n) + E_{total}^{harv}(n)} \tag{19}$$

where $E_{resd}(n)$ is the residual energy of node *n*, $E_{est}^{harv}(n)$ is the predicted harvested energy in the following round and $E_{total}^{harv}(n)$ is total harvested energy of node in previous rounds. $E_{est}^{harv}(n)$ is determined with simple moving average algorithm (Eris et al., 2023). Figure 7 illustrates the power output of a piezoelectric harvester in relation to varied flow velocities, recorded power outputs of an underwater sensor.

Moving average is a commonly utilized algorithm in wireless sensor networks to forecast the solar power availability (Sah et al., 2023a). The moving average algorithm is computed incrementally, making it suitable for real-time applications where new data is continuously being observed. The sliding window method allows for efficient updates with minimal computational overhead making it a robust choice for short term predictions on real time data. Algorithm 1 represents sliding window implementation of moving average method to estimate power output of harvester and predict the total energy harvested by a cluster member node in each round. The window size (number of data points considered in each average) is a crucial parameter. It determines the balance between responsiveness to recent changes. The window size is defined as the number of time steps contained in a frame.

Algorithm 1 Moving Average Algorithm for Piezoelectric Energy Harvesting Prediction

Initialize:
  window_size ← T_{frame}
  E_i^{harv} ← []
function UPDATE_PREDICTION($E_i^h(t)$)
  E_i^{harv}.append($E_i^h(t)$)
  if length(E_i^{harv}) > window_size **then**
    E_i^{harv}.pop(0)
  end if
  if length(E_i^{harv}) == window_size **then**
    moving_average ← sum(E_i^{harv}) / window_size
    E_{est}^{harv}.append(moving_average)
    return moving_average
  else
    return None
  end if
end function
function PREDICT_ENERGY($E_i^h(T_{frame})$)
  prediction ← update_prediction($E_i^h(t)$)
  if $t < Tround$ **then**
    E_{est}^{harv} ← prediction
  else
    $t \leftarrow t + 1$
  end if
end function

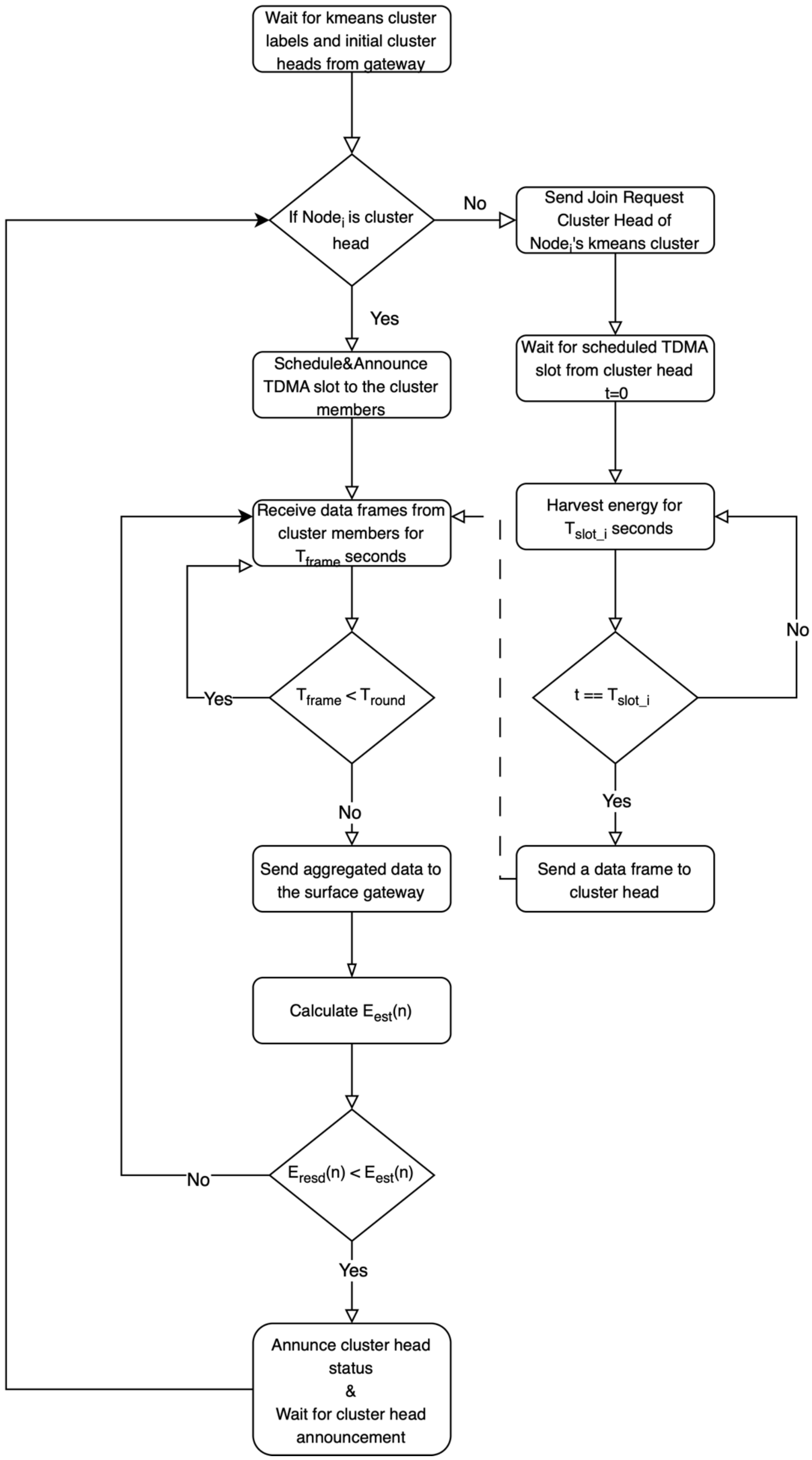

Figure 6. Flow diagram of a round and cluster head selection of our method

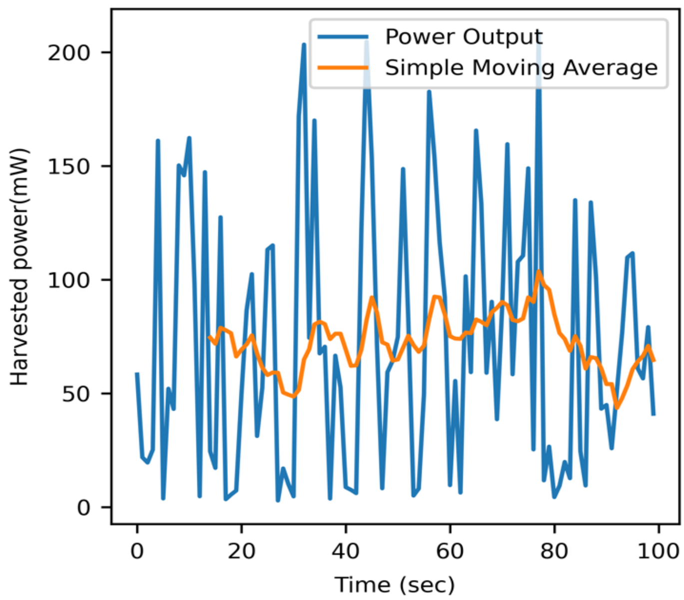

Figure 7. Piezoelectric harvested power output and moving average.

3.4 Performance Evaluation and Comparison

The performance and comparison of our method with the LEACH-based K-means algorithm is analyzed in both conditions, incorporating the piezoelectric harvester unit and without the harvester installation specifically for 2D UASNs through simulations using Python implementation of LEACH protocol (W. R. Heinzelman et al., 2000). A comprehensive description of the parameters utilized in our study is provided in Table 1. Performance metrics of cluster head selection methods for lifetime include round numbers at which the First Node Death (FND) and Last Node Death (LND) occurs in the network. In our investigation, we assess their performance network lifetime in relation to the number of rounds and with respect to the time steps until the last node dies. Figures 8, 9 depicts the efficiency of our k-means clustering method compared to LEACH, and two common cluster head selection methods in the UASN literature (Omeke et al., 2021; Pour & Javidan, 2021). In Figure 8, 9 Centroid Proximity curve denotes the performance of cluster head selection policy, takes only centroid proximity of nodes into account. This method forms better clusters in terms of size, hence outperforms LEACH in terms of FND and LND noticeably. It is observed that the FND and LND of the network is improved compared to the performance of Centroid Proximity when cluster head selection policy accounts for remaining energy of sensors and the distance of gateway station. It is also observed that centroid proximity or locality in the choice of cluster heads has very slight performance gap.

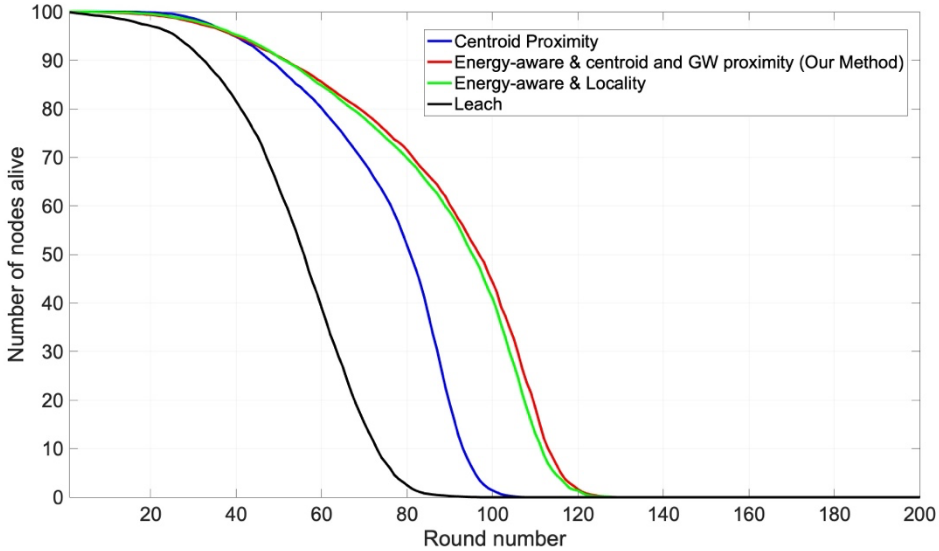

Figure 8. Performance comparison of ch selection methods w.r.t round

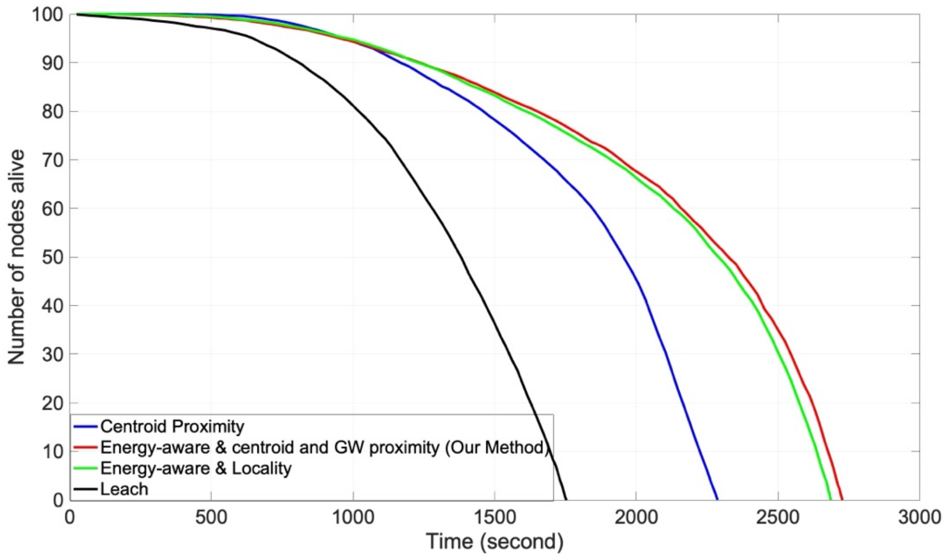

Figure 9. Performance comparison of ch selection methods w.r.t time

To investigate the effect of energy harvesting-aware cluster head selection policy, we proceed with k-means clustering and energy-aware & centroid and GW proximity cluster head selection method. Our findings illustrate the change in the number of alive nodes with respect to the round number, averaged over 100 simulation runs (Eris et al., 2023). We conducted a comparison between our method and the Leach- K-means clustering methods, as depicted in Figure 10. The marked curve as harvesting on denotes the change in alive nodes considering nodes with energy harvesting units installed and in idle times of nodes, they collect available energy in underwater medium. The curve marked as harvesting off represents the change in alive nodes when nodes don't have energy harvesting equipment. To assess network lifetime, we employed multiple frames of data collection in each round, with duration of 25 seconds. Because of non-uniform distribution of nodes, clusters vary in size, resulting distinct number of frames collected from every cluster in rounds. For this reason, in Figure 10, piezoelectric harvesters have a small impact to the number of alive nodes with respect to the number of rounds.

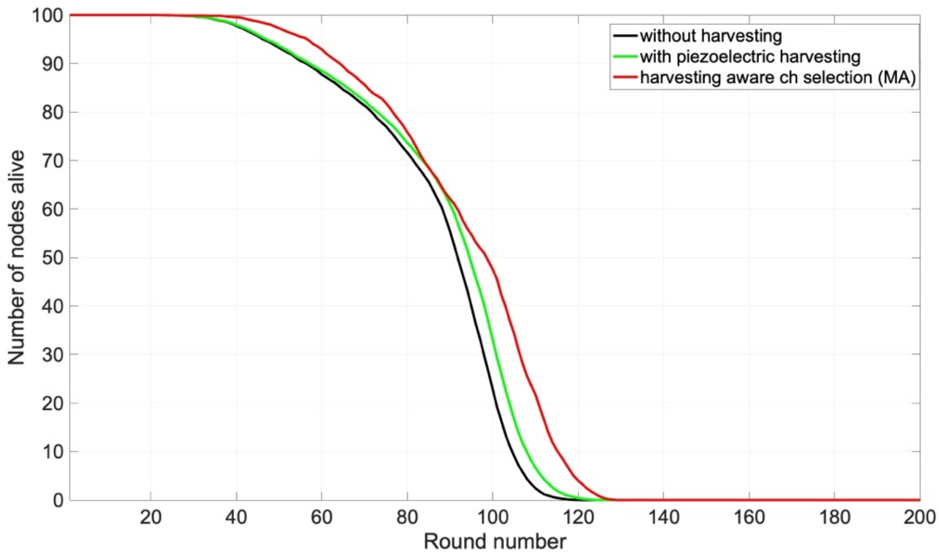

Figure 10. The impact of energy harvesting and moving average prediction based ch selection to the number of alive nodes over simulation rounds.

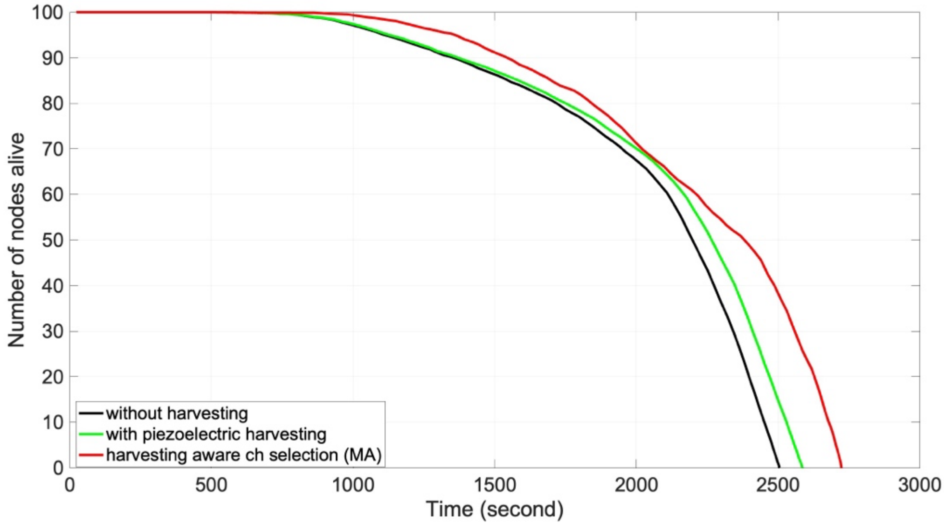

Figure 11. The impact of energy harvesting and moving average prediction based ch selection to the number of alive nodes over simulation time (seconds).

As illustrated in the Figure 11, piezoelectric harvesting contributes to extending the network lifetime, resulting in a higher First Node Death (FND) value, over simulation time. Although only a marginal discrepancy in lifetime is observed when harvesting unit is installed or not, the curve representing our policy utilizing simple moving average (SMA) prediction, indicates that if nodes can decide to become the cluster head with the predicted harvested energy, the network's lifetime improves. Many methods for determining harvested energy in piezoelectric models overlook temporal uncertainty of available energy and assume constant energy gain for the entire idle time of nodes, which may not hold true in real-life scenarios due to the highly variable underwater environment. Therefore, we investigated the effect of harvesting aware cluster head selection policy in lifetime with this approach and performance comparison of our simulation runs denotes considering the expected harvested energy results in a larger Last Node Death (LND) in the simulations.

Table 4

Simulation Settings of Moving Average Prediction-based Cluster Head Selection Policy

Network Parameters		
N_{node}	Number of nodes	100
M	Network Size (m^3)	$250 \times 250 \times 250$
h	Network Depth (mt)	250
$S_{(x,y,z)}$	Sink Location	(125,125,0)

Modem & Channel Parameters		
R	Data rate (kbps)	2
E_0	Initial Energy (kJ)	0.1
P_{idle}	Idle Power (mW)	2.5
E_{rx}	Energy dissipation in Receiver (nJ/bit)	50
f	Frequency (kHz)	25
p_{ij}^b	Bit error rate on link-(i,j)	10^{-10}
L	Data-Packet size in single frame (bits)	1024
κ	Spreading Factor	1.5
B_N	Noise Bandwidth (kHz)	1
E_{DA}	Data Aggregation Energy (nJ/bit/packet)	5

Piezoelectric Harvester Parameters		
L	Cantilever length (mm)	50
B	Cantilever width (mm)	3
ε_r	Relative permittivity	4000
ε_0	Absolute permittivity (F/m)	8.85×10^{-12}
T_{pzt}	Thickness (μm)	60
d_{31}^p	Piezoelectric constant (C/N)	300×10^{-12}
S_t	Strouhal number	0.2

3.5 Discussion and Future Work Directions

In this chapter we proposed moving average prediction to select cluster heads considering stochasticity in the underwater environment. Our findings reveal that if nodes predict their foreseeable energy and holds cluster head role when harvestable energy is low, it is shown that the overall lifetime of the network is improved. However, one must take into account that despite its simplicity, moving average can be sensitive to outliers or lead to delayed responses to energy fluctuations. When the energy harvesting trend is predictable and the change in the available energy is low, moving average can provide accurate predictions. To adapt to frequent changes in available energy, it is essential to engage additional strategies such as machine learning models to improve responsiveness of the algorithm. Regression models on historical observations of available energy or advanced learning methods are promising to provide nonlinearity and complex patterns to improve the prediction on the foreseeable energy that moving average can overlook. On the other hand, it is worth mentioning that our implementation investigates the contribution of energy harvesting to the lifetime within a set number of nodes and a fixed region, i.e. 250 meters of each dimension in the deployment area. The average distance between nodes in the deployment depends on the density of the topology hence, transmission power can vary and impact the effect of the harvested energy on lifetime. We plan to analyze the lifetime considering the density in deployment region. Furthermore, in addition to the piezoelectric harvester, different type of harvesters (i.e., hydrophone and tribune harvesters) will be focus on our future work.

3.6 Chapter Conclusion

We implemented an energy-harvesting aware cluster head selection policy for UASNs specifically long-term monitoring underwater applications. Underwater sensor applications primarily utilize sonar systems, which face challenges including high transmission loss, multi-path effects of acoustic signals, propagation delay, and limited bandwidth. Consequently, high transmission powers are often required to mitigate these challenges, resulting in energy-intensive data transmission tasks. Clustering has emerged as an effective solution to optimize network lifetime by partitioning nodes and reducing transmission powers. Efficient monitoring of vast sea areas by UASNs necessitates energy-efficient algorithms and techniques for long-lasting monitoring and surveillance systems. Significant efforts have been directed towards improving the energy efficiency and longevity of underwater sensor networks. Energy harvesting presents a promising solution for replenishing the batteries of sensor devices, leveraging various technologies to convert ambient energy forms in the underwater environment into electricity. Hydrokinetic turbines and piezoelectric cantilevers are among the technologies capable of harnessing energy from water currents, with piezoelectric harvesters offering reliability and cost-effectiveness. However, the harvested energy from piezoelectric bimorphs can vary significantly depending on underwater environmental conditions. To address this variability in energy sources, we propose a realistic model for piezoelectric energy harvesting. In this chapter, we address a cluster-based energy-aware routing issue in UASNs, assuming each sensor in the system harvests energy using piezoelectric harvesters. We propose a novel cluster head selection policy based on predicted harvested energy, location, and estimated energy consumption of sensor nodes. This is a crucial step towards ensuring sustainability of information retrieval for applications like early warning systems and industrial monitoring. Our proposed technique reduces energy consumption and extends the network's lifetime compared to existing approaches in the literature. Our main contributions in this work include being the first to address the stochastic consideration of energy harvesting processes in the cluster-based energy-aware routing problem within UASNs. We propose a novel cluster head selection method that incorporates a realistic model for the piezoelectric energy harvesting of underwater sensors. Additionally, we calculate the total energy needed for data collection tasks within clusters and select cluster heads based on factors such as estimated harvested energy, remaining energy, locality, and proximity to the sink. This approach aims to form energy-balanced clusters, thereby enhancing network lifetime.

Chapter 4
Reinforcement Learning Based Routing Algorithm for Energy Management in UASNs

Underwater wireless sensor networks (UASNs) have the potential to supply environmental data for various applications, including studies on environmental changes, early warning systems, and industrial monitoring. Maintaining continuous information delivery is crucial in these contexts, with UASNs serving as fundamental assets. However, the unique characteristics of underwater environments necessitate sensor nodes to rely on their limited battery reserves, making energy management a critical concern. Cluster-based network routing protocols have been extensively studied to minimize network energy consumption by employing cluster heads (CHs) to aggregate data and reduce overall energy usage, thereby extending the network's lifespan. Additionally, there has been a growing focus on harnessing energy from ambient resources underwater to prolong the operational life of sensor nodes. In this chapter we address the energy-aware routing problem in underwater acoustic sensor networks (UASNs) by considering the stochastic energy harvesting process at each sensor node. Our contribution lies in proposing a novel reinforcement learning-based algorithm for determining cluster heads (CHs), which considers not only the nodes' positions and residual energy but also the expected harvested energy. Our implementation demonstrate that our approach significantly reduces energy consumption and substantially extends the network's operational lifetime. To maximize the benefits of underwater environments, research endeavors concentrate on addressing the unique challenges posed by underwater mediums to facilitate sustainable data collection in various applications.

Underwater sensor nodes predominantly rely on acoustic communication technology, as it allows signals to propagate through water over extensive distances. However, acoustic waves are susceptible to high path loss and bit-error rates, necessitating sensors to use high transmission powers to prevent packet loss (Akyildiz et al., 2005a). Efficient energy management is primary concern in UWSN design since sensor nodes rely on batteries as their energy source and are often deployed in subsea areas, posing challenges for maintenance and human intervention. In underwater communication, data transmission stands out as the most energy-intensive operation for sensor nodes. Thus, to address energy consumption concerns, it's essential to minimize transmission power by employing multi-hop data forwarding or reducing the frequency of transmissions (Awan et al., 2019; Bholowalia & Kumar, 2014; Bozorgi et al., 2017; Dror et al., 2013; Erdem et al., 2019; Erdem & Gungor, 2017; Eris et al., 2023; Sah et al., 2023b). In the literature, significant efforts have been devoted to reducing the energy consumption of sensor nodes when operating in underwater channel conditions. To address this issue, clustering has emerged as a powerful technique for minimizing the transmission power of nodes by dividing the network into regions. Consequently, data is transmitted to a local cluster head (CH), and multi-sensor fusion techniques can be employed to reduce the volume of transmitted data to the surface gateway (Redhu & Hegde, 2019).

Energy harvesting serves as an effective approach to prolonging the operational lifespan of sensor nodes by harnessing energy from two primary sources of kinetic energy in the ocean: waves and currents. Nevertheless, the energy harvested from underwater environments is insufficient to meet the energy requirements of sensor transceivers (Zhao et al., 2021). Hence, in addition to optimizing the total energy consumption of underwater sensor networks, maximizing the utilization of ambient energy resources in these networks has emerged as a significant topic in this field. Numerous research endeavors suggest that leveraging ambient resources for energy harvesting enhances the lifespan of nodes in Underwater Acoustic Sensor Networks (UASNs), underscoring the importance of energy-efficient routing to facilitate sustainable data collection from underwater sensor nodes in UWSN applications. Several studies in the literature concentrate on enhancing network clustering and cluster head selection methods with the objective of maximizing the lifetime of Wireless Sensor Networks (WSNs).Various research initiatives highlight the potential of robust optimization methods in improving the lifetime of WSNs (Frikha et al., 2021; Gul & Demirekler, 2018; Gul & Uysal-Biyikoglu, 2014; Omeke et al., 2023; Pour & Javidan, 2021; Sah et al., 2023c; ZHANG Gaoxi XIAO Hwee-Pink TAN et al., 2013). The primary challenge in dealing with the unique aspects of underwater communication channels stems from the limitations of sensor devices and harvesting equipment.

In our previous cluster head selection method (Eris et al., 2023), we introduced a cluster head selection policy and examined the impact of predicted available energy on cluster head selection. This analysis incorporated stochastic energy harvesting processes on each sensor, addressing the energy-aware routing problem in Underwater Acoustic Sensor Networks (UASNs). This chapter builds upon our previous research and introduces a novel reinforcement learning-based approach for selecting cluster heads (CHs). This procedure takes various factors into account such as the nodes' positions, residual energy, predicted harvested energy, and learns from previous node decisions within the cluster. Our implementation of reinforcement learning aiming towards optimal cluster heads while incorporating piezoelectric energy harvesting for energy management in the cluster-based routing problem. We investigated the impact of energy harvesting on the cluster head selection mechanism on the lifetime of Underwater Acoustic Sensor Networks (UASNs). A multi-agent reinforcement learning method is employed in the cluster head selection mechanism and compared with our previous study, which utilized a cluster head selection policy based on moving average prediction of available energy in the underwater environment. Furthermore, to minimize energy dissipation in transmission we used the transmission power control (TPC) mechanism and conduct a more practical evaluation.

4.1 Background

In literature most studies that investigate the underwater long-term monitoring applications, consider the 3D architecture of UASNs. In 3D architecture, sensor nodes deployed in varied level of depths of underwater. Sensors are anchored to the ocean bottom and depth can be adjusted by buoyancy equipment of sensors. In this section we present the network model considered in our study and formulize the cluster head selection as an optimization problem.

4.1.1 Network model. In our network model, we consider one hundred of sensor nodes randomly deployed in varied level of depths in monitoring area of interest. Nodes are uniformly scattered in underwater and send periodical observatory data to the surface sink station located above the sea surface. Figure 12 shows a sample representation of 3D UASN network containing 100 nodes and one surface sink. Due to fixed replacement of sensor nodes as they are attached to the ocean bottom, we assume their location do not vary over the course of time. Cluster heads, also referred to as cluster administrators, are selected through a series of rounds to fulfill specific criteria. Each round comprises frames, with each node transmitting a single packet of L bits during a frame. Frames serve as the foundational units of rounds and nodes consume energy while transmitting, receiving, or aggregating data.

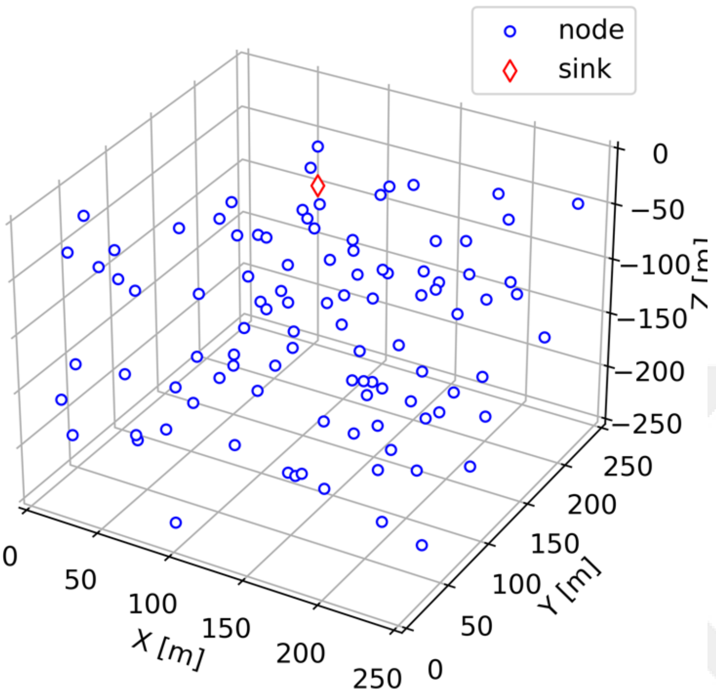

Figure 12. Example of 3D network architecture showing nodes and sink positions.

Cluster heads coordinate the transmissions of cluster member nodes using Time-Division Multiple Access (TDMA) for in-cluster communication (Molins & Stojanovic, 2006). TDMA is the medium access control protocol due to the high and variable propagation delays in acoustic signals. As a result, nodes remain idle until data transmission in the next frame at predefined intervals.

4.1.2 Optimization problem. Optimal cluster head selection is formulized as an optimization framework. We would like to find optimal cluster heads, cluster the network with these cluster heads. The objective function is defined as the cost of selecting k nodes as cluster head. Aiming to minimize the cost function is defined considering three aspects as

$$\min_{CH_i \in \{1,2,\ldots,N\}} \left(\begin{array}{c} \gamma_1 \sum_{i=1}^{k} dist(CH_i, GW) \\ + \gamma_2 \sum_{i=1}^{k} \sum_{j=1}^{Q_i} dist(n_j, CH_i) + \gamma_3 \sum_{i=1}^{N} \frac{1}{E_{res}(i)} \end{array} \right) \tag{20}$$

where Q_i represents the ith cluster and GW is the surface gateway. First part of the cost function aims to select nodes close to the GW since the cost of the communication with GW is less than compared to nodes that are located at larger distances to the GW. Second part of the cost function represents the cost of selection of CH_i to its respective cluster. The last part of the cost function indicates the cost of selected cluster heads in terms of every node's remaining energy. γ_1, γ_2, γ_3 depicts the weight factor of each cost function and $\gamma_1 + \gamma_2 + \gamma_3 = 1$. In the beginning of round, nodes selected aiming towards the minimized cost function.

4.2 Our Cooperative Q-learning Algorithm for Cluster Head Selection

Our proposed method initially, clusters the network with K-means algorithm presented in Chapter 2. The surface gateway performs K-means clustering of sensor nodes to partition the network. The gateway initializes clusters and selects initial cluster heads in the first round. In the following rounds, cluster head selection is triggered when cluster head determine that their remaining energy is inadequate for the next round. Nodes become cluster head based on a score calculation, which considers both remaining energy and predicted available energy. Nodes with higher scores, indicating lower delay time, assume the role of cluster heads by broadcasting their presence to their respective clusters. The prediction of available energy is achieved using a simple moving average method.

Table 5

Energy State, $State_{EH_n}$

Energy Harvesting State	
0	$Harv_{(t-1)} < 0.01$
1	$Harv_{(t-1)} \leq 0.03$
2	$Harv_{(t-1)} > 0.03$

Table 6

Harvested Energy State $State_{E_{ratio}}$

Energy State	
0	$E_{res}(t-1) > 0.9$
1	$E_{res}(t-1) > 0.5$
2	$E_{res}(t-1) > 0.2$
3	$E_{res}(t-1) > 0.1$
4	$E_{res}(t-1) < 0$

Our implementation of reinforcement learning framework consists of modelling the optimization problem from an agent perspective. We run Q-learning on every sensor in the network. Each underwater sensor node within the network operates as an agent in RL-perspective. The states of the nodes are defined as observations of nodes respective to their remaining energy and prediction of the amount of expected energy for the upcoming data collection round to optimize the sensor nodes' lifespan under varying environmental conditions. The state of each node is represented by the ratio of remaining energy to its initial energy and total harvested energy, denoted as $S_n = (Er_n, EH_n)$. Discrete states of remaining energy and available energy are represented in Table 3 and Table 4, respectively. Agents decide whether to serve as cluster heads for the current round. If an agent selects to be a cluster head all other nodes in the cluster act as cluster members. The action space is discrete and is either 0 or 1. The system aims to maximize network lifespan while maintaining a balance between energy consumption and energy harvesting, thereby enhancing throughput and preventing premature node failures. The reward for each node's decision to act as a cluster head or member is calculated as

$$Reward_n := \begin{cases} -100, & if \dfrac{E_n}{E_0} \leq 0.1 \\ E_n + E_{CH}(n_{CH}) + E_n^h, & otherwise \end{cases} \quad (21)$$

Algorithm 2 gives our implementation of cooperative Q learning method in cluster head selection phase.

Algorithm 2 Cooperative Q-Learning Algorithm for Cluster Head Selection

---

Require: $0 \leq \gamma \leq 1$, $\alpha \in (0, 1]$, small $\varepsilon > 0$, $n \geq 0$ where γ is the discount factor, α is the learning rate and ε is the exploration parameter set as 0.1.
  Initialize $State(n) \leftarrow (E_{ration_n}, EH_n) = (0, 0) \quad \forall n \in N$
  Initialize $Q(a, S) \leftarrow E_C(n_{CH}) \quad \forall a \in A$
  for each round or episode **do**
    for each node n in Clusters **do**
      if $random(n) \geq \varepsilon$ **then**
        Choose exploitation
        Select maximizing action $a_i = \arg\max Q(:, s)$
      else
        Choose exploration
        Select a random action a_i for node n
      end if
      Receive reward R from environment
      if $E_n/E_0 \leq 0.1$ **then**
        Receive reward, -100
      else
        Receive reward, R
        $E_n + E_C(n_{CH})^{-1} + E_{harvested}$
      end if
      Enter new state, S'
      Update $Q(a, S) \leftarrow \alpha * (R + \gamma * Q(a_{max}, S') - Q(a, S))$
    end for
  end for

4.3 Performance Evaluation and Comparison

We present comparison between our current findings and our prior research (Eris et al., 2023). In our earlier study, the selection of cluster heads was influenced by factors such as remaining energy and the moving average prediction of anticipated harvested energy for each node. We explored how stochastic modeling of piezoelectric energy harvesting impacts K-means clustered networks and devised a method for selecting cluster heads by predicting future energy availability. By conducting simulations in Python and averaging the results over 100 runs, we evaluated our approach against the conventional method of cluster head selection. As an advancement over our previous work, we enhance the performance of the moving average prediction technique in cluster head selection by incorporating reinforcement learning technique. The sustainable operation of sensor nodes significantly influences the efficacy of a sensor network. The failure or inactivity of a sensor node can greatly impact the overall performance of the network. To gauge the network's longevity, we measure the time until the occurrence of key events such as the first node death (FND), the time when half of the nodes have died (HND), and the time until the last node dies (LND). These metrics serve as benchmarks for comparing the lifespans of different networks. Figure 13 illustrates the performance comparison of cluster head selection methods in terms of network lifetime measured in rounds. The benchmark curve labeled as "none" represents the efficiency of k-means clustering and cluster head selection based solely on the locality of nodes and their remaining energy. The "EH" curve demonstrates the impact of energy harvesting on network lifetime without harvesting awareness. The "MA" curve depicts the performance of moving average prediction in cluster head selection, where harvesting awareness is incorporated by considering expected harvested energy in the selection of cluster head based on a score. However, in this method, nodes only use this parameter as a factor in the score calculation. With Q learning, significant improvements are observed in the three key metrics (FND, HND, LND) by an average of 23%, 40%, and 30%, respectively. The substantial improvement in FND can be attributed to the estimation of energy consumption within the cluster, which serves as an optimistic value to initialize each potential action for the Q-table of each node. Unlike initializing the Q-table with 0, this approach promotes better locality in selecting cluster heads, thereby avoiding the cost of exploration that may result from poor decisions made at the outset of the rounds. Figure 14 illustrates the impact of cluster head selection methods on network lifetime over the simulation period. We observe that the Last Node Death (LND) metric is increased by approximately 30% compared to the EH and MA curves. As the number of active nodes in the network decreases, more time is required for each round to be completed, given that the initial duration of a round remains fixed. Consequently, there is more time available for nodes to reach a steady-state phase within each round. This significant improvement in LND is notable compared to our previous study. Figure 15 depicts the number of active nodes versus the data signals received by the surface gateway (or Base Station, BS). It is evident that as the simulation progresses, the total amount of harvested energy accumulated by each node increases. Consequently, the rate of decrease in the number of active nodes diminishes, leading to a higher retention of active nodes and a sustained transmission of data signals. In the RL curve, there is an approximate 30% increase in transmitted data, indicating a more efficient utilization of available energy. On the other hand, in the MA curve, the transmitted data increases by approximately 7%. The K-means RL curve exhibits the most favorable outcome with a higher number of nodes remaining active in the later rounds. Consequently, the surface gateway (BS) receives approximately 22% more data signals compared to the MA method. Figure 16 indicates an increase in the amount of data received at the surface gateway per unit of time. When employing moving average prediction for cluster head selection, the cluster head score predominantly relies on factors like remaining energy and locality. In contrast, Q learning enables adaptation to change in ambient harvesting parameters by adjusting states based on harvested energy. Consequently, sensor nodes gain awareness of their environment and neighboring nodes. This heightened awareness of energy harvesting conditions leads to improved energy utilization, making central nodes less susceptible to rapid exhaustion.

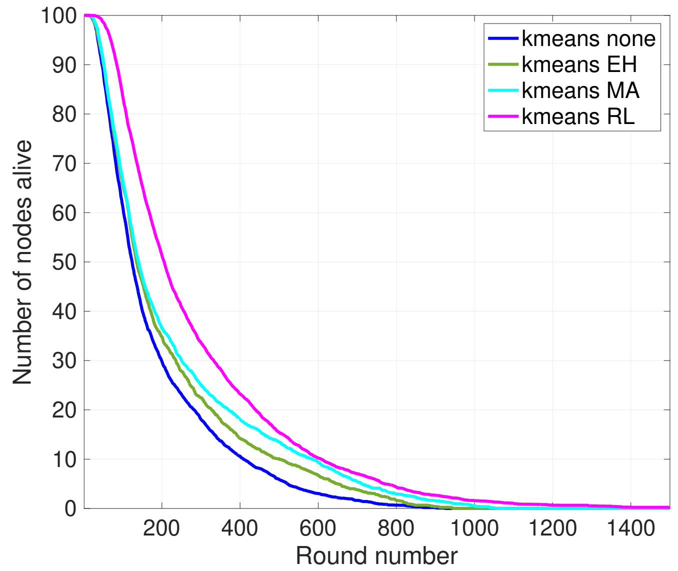

Figure 13. Number of alive nodes over simulation rounds.

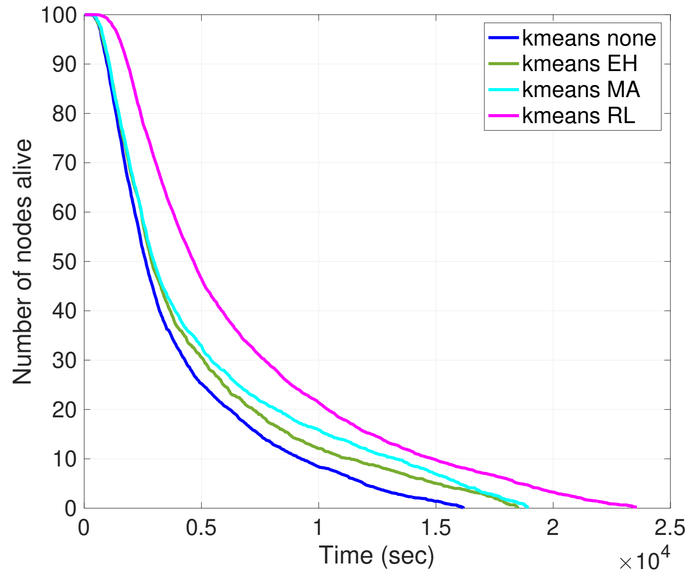

Figure 14. Number of alive nodes over time.

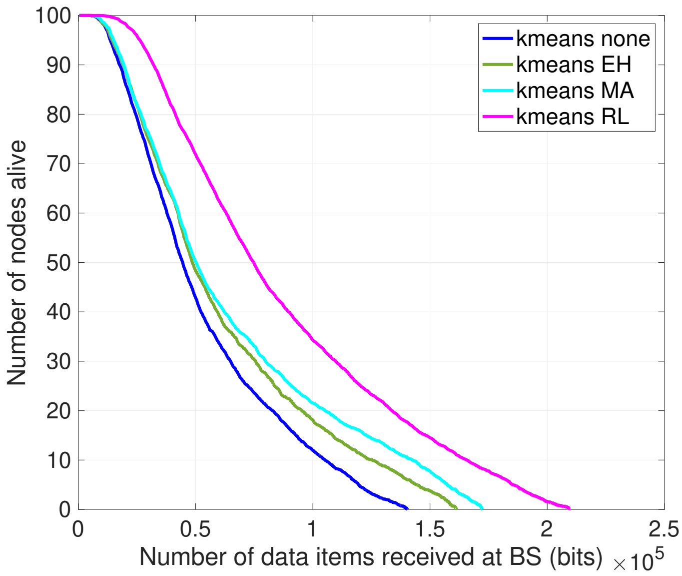

Figure 15. Data items received at the Base Station.

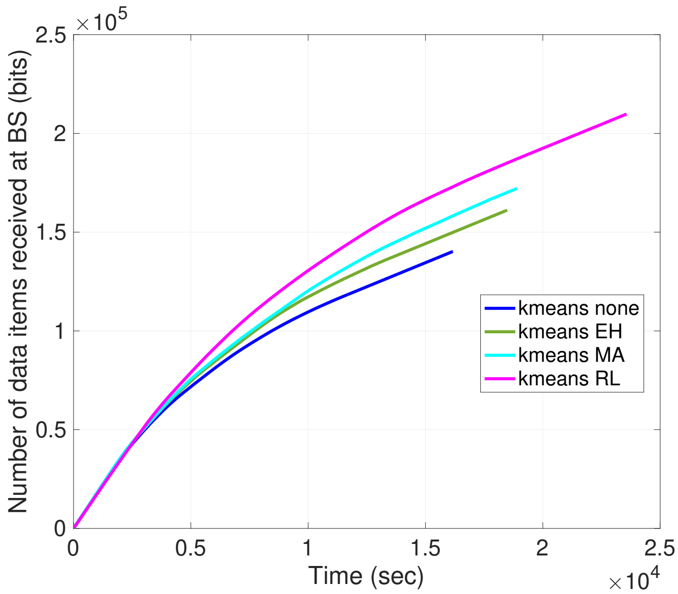

Figure 16. Number data signals received over time at the Base Station.

4.4 Discussion and Future Work Directions

Due to the hostile nature of underwater environment such as long propagation delays, high path loss of acoustic signals and temporal-spatial uncertainty, it is vital to develop adaptive strategies for energy efficient data routing for UASNs. For this reason, our method considers stochastic nature in the available energy and propose an adaptive, learning-based cluster head selection strategy for UASNs. As our cooperative Q-learning implementation defines the reward function that penalizes when nodes enter very low energy states after selecting a random action in the exploration. If more than a certain number of nodes explore the action to become cluster head simultaneously, the network may face inefficiencies despite the penalty defined in the reward function. Therefore, the reward definition can be tuned to appropriately select optimal number of cluster heads. As a future work, we plan to investigate the effect of optimal number of cluster heads incorporating energy harvesting on lifetime of UASNs. Another promising future work direction involves comparing various optimization methods such as Particle Swarm Optimization, Ant Colony Optimization proposed in the literature specifically addressing the cluster head selection problem without incorporation of energy harvesting. Comparing the stochastic optimization methods with our adaptive Q-learning approach and investigating the trade-offs in between these strategies is essential for achieving more reliable and sustainable operation in UASNs.

4.5 Chapter Conclusion

The design of underwater acoustic sensor networks has several challenges because of the limitations of their resources as well as harsh conditions of the underwater environment. Especially, monitoring large marine areas through UASNs requires effective energy management strategies to enable sustainable underwater surveillance applications. In this chapter, we consider a cluster-based routing problem in underwater wireless sensor networks, where the issue at hand is energy management. It is believed that piezoelectric harvesters are utilized by each sensor to recover its energy. To harness the available energy in ambient resources, a cluster head is designated adaptively, considering its energy harvesting state and selected as a relay for data transmission accordingly. For this problem, we propose a novel cluster head selection algorithm for UASN based on reinforcement learning, specifically, cooperative Q learning algorithm considers the expected harvested energy, location, estimated energy consumed by sensors in each cluster. Sensors take energy-intensive operation, acting as cluster head in their region while their remaining energy and locality is sufficient and available amount of energy to harvest is low. Our proposed reward function takes the amount of foreseeable energy of into consideration. It is shown that our approach improves lifetime of sensor nodes and throughput of the system both by around 30%. A crucial step that needs to be made is to ensure that the retrieval of information for applications such as long-term surveillance and industrial monitoring is carried out without interruption. Our proposed technique enhances the exploitation of the available energy in underwater medium while reducing energy consumption and enhancing lifetime of the network remarkably in comparison to the methods that are now being utilized and in the nearby literature.

Chapter 5
MARL Based TDMA Slot Allocation Method

Energy efficiency was the focus of underwater communication in the past decade, there exists considerable improvements in this research field to prolong network lifetime by suggesting mac and routing protocols to take limited energy resources of acoustic sensor nodes and high energy demand of underwater wireless medium into consideration. The primary reason for energy depletion in underwater acoustic sensor networks (UASNs) is environmental factors. The underwater environment is particularly prone to noise from various sources such as shipping activities, wind, and aquatic life. Consequently, this leads to a low signal-to-noise ratio (SNR), resulting in a high bit error rate (BER). Since acoustic communication is a dominant technology in underwater, high-energy demand of acoustic links is usually decreased by clustering methods. With the advancements in robotic technology, Autonomous Underwater Vehicles are common application assets to gather information collected by clusters in underwater applications. The network lifespan of underwater acoustic sensor networks (UASNs) holds significant importance for sustained long-term monitoring applications. Various learning algorithms to employ autonomous underwater vehicles and sensors to curate and process data over long-term surveillance applications, allows a comprehensive understanding of aquatic environments and exploit decision-making for ecological resource management. Main concerns are limitations on the battery of underwater sensors and reliability in low-rate acoustic links that shapes the design factors of underwater communication protocols. In this work, we considered the spatial uncertainty of underwater ambient resources to improve the utilization of available energy and used stochastic model for piezoelectric energy harvesting. Considering a realistic channel and environment conditions, a novel distributed multi-agent reinforcement learning algorithm is proposed. Nodes observe and learn the peaks in the available energy from their surroundings and autonomously decide their communication slots instead of relying on the cluster head. The numerical results show that our proposed method improves in-cluster communication time interval utilization and outperforms traditional TDMA-based time slot allocation methods in terms of throughput and energy harvesting efficiency.

5.1 Transmission Slot Scheduling Mechanism in TDMA-based MAC methods

In a clustered network, each node learns their clusters from the broadcast message of gateway station. Following the node calculation whether to become the cluster head for the current round or not, with prior knowledge of members in its cluster, cluster heads assign a unique time slot for each node in the cluster in contention-free TDMA-based medium access protocols. Cluster members send their data to cluster head in their programmed time slots to avoid collisions. After data transmission, each node become idle and harvests ambient energy until the next frame of data collection.

TDMA-based medium access protocols are preferred over collision avoidance-based protocols and broadly utilized by cluster-based data collection (Molins & Stojanovic, 2006) due to significant propagation delay and limited bandwidth of underwater channels. Furthermore, time slot allocation is utilized to reduce energy usage by preventing contention and collisions among packets without extra overhead for channel reservation that requires short control packets. For intra-cluster communication, nodes perform data transmission in their time slots that is designated by their cluster head and switch back to idle state until the next frame. However instead of centralized control of medium access, nodes can determine their slots by learning the environmental conditions of harvesting parameters.

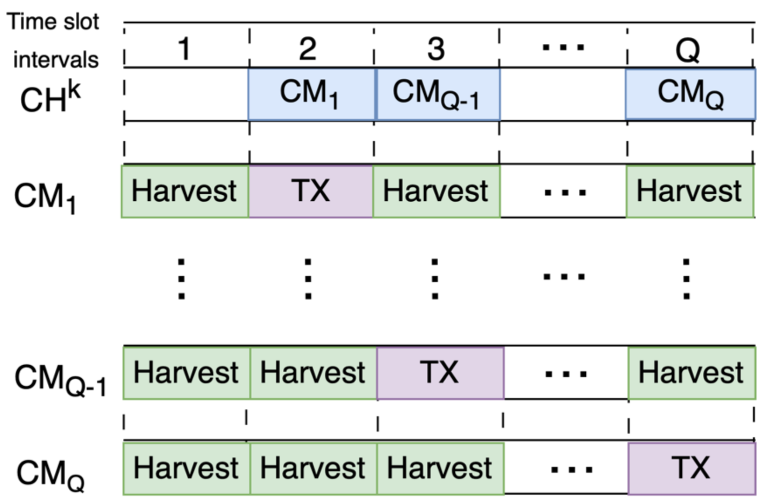

Figure 17. Example self-scheduling of time slot allocation representing low (TX)

and high-(Harvest) energy harvesting conditions.

In Figure 17 an example timeline showing our approach within a data frame in clustered networks. The duration of a round is predetermined and CH^k represents the cluster head of kth cluster with a time slot allocation table, contains record of each packet reception from its cluster members. Here, cluster members, denoted as CM decide their transmission slot for a packet by selecting one of Q available time slots in current frame. Nodes observe energy harvesting rates and schedule their transmission to a time slot with less available energy compared to other slots. For instance, CM_1 decides allocating time slot 2 according to its available energy. Depending on its value of Q table for time slot 2, transmission decision is calculated based on its harvested energy in previous frame, observed the lowest in time slot 2.

5.2 Transmission Slot Scheduling Problem Definition

As TDMA-based and round-robin transmission protocols in clustered sensor networks, each transmission round is divided into frames, *f*. In each frame cluster members send a fixed length (L) data packet to their cluster head. We consider data backlogging system where nodes capture measurements about their environment and always have data to transmit when it's their turn in their scheduled time slot. All nodes have the same data rate and allow to send same number of packets in each frame. We use transmission power control (TPC) so that each transmission is held with a transmission power by nodes to satisfy a predetermined SNR for successful packet reception (M. Felemban & Felemban, 2013).

We consider a network of non-uniformly dispersed 100 nodes in varied depths underwater and nodes maintain communication with their respective cluster heads over their scheduled time slots per frame. Cluster heads determine the time intervals and notify cluster members to their time slots for data transmission. We denote the set of N acoustic sensor nodes by $S = \{s_1, s_2, \dots, s_N\}$. We assume stochastic energy harvesting processes modeled by Poisson distribution at each node. The remaining energy and the amount of harvested energy of a sensor node s_i in a time slot t denoted as $E_i(t)$ and $E_i^h(t)$ respectively. $E_i^h(TS)$ represents total harvested energy per frame. Considering frames consist of TS many time slots, same as number of nodes in clusters, time slot set is represented as $TS = t_1, t_2,..., t_N$, the total amount of harvested energy in cluster within time steps from 1 to TS in frame, under a policy π where t_{tx_i} is the designated time slot index of sensor i, $I_{t_{tx_i}} \in \{0,1\}$ is the indicator function and harvested energy at time slot t_m can be written as $E_i^h(t_m) = E_i^h(t_m - 1)\ I_{t_{tx_i}}$. Overall harvested energy under a policy is determined as

$$V_i^{\pi}(TS) = \sum_{i=1}^{N} V_i^{\pi}(TS) = \sum_{i=1}^{N} \sum_{t=1}^{TS} E_i^h(t) \tag{22}$$

where $V_i^{\pi}(TS)$is the sum of harvested energy under policy π in cluster from starting time slot 1 through TS.

$$E_i^h(TS) = \sum_{t=1}^{TS} \left[E_i^h(t) - \sum_{t=t_{tx_i}}^{t_{tx_i}+1} E_i^h(t) I_{t_{tx_i}} \right] \tag{23}$$

Our objective is to maximize the total amount of harvested energy by each node over a finite time horizon,

$$\max_{\pi} \frac{1}{T} E[V_i^{\pi}(T)] \tag{24}$$

In clustered networks, optimal time slot scheduling algorithms employ cluster heads to schedule the transmission slots of each node. However, the cluster head may not select the optimal transmission slot without incurring communication overhead to gather knowledge on the energy harvesting states of nodes. Even when this information is provided to the cluster head, its timeliness might be compromised due to costs associated with complexity and computational delays, adopting a global optimal solution approach is not feasible. Therefore, instead of relying on centralized methods, delegating the transmission slot allocation task to cluster members not only delegates autonomy but also unveils for harnessing ambient energy through adaptive transitions between transmission and idle energy harvesting modes. Nodes can determine their slots to transmit packets, incorporating environmental conditions to determine the optimal timing for packet transmission or reverting to energy harvesting mode. We model this problem as multi agent cooperative learning problem. In our model, we used stochastic piezoelectric harvester, and the following assumptions are made. The statistical characteristic of all random quantities (flow velocity, voltage peaks, etc.) is known by observation and experiment (Jarosz et al., 2011; Pobering & Schwesinger, 2004; Zou et al., 2021). Harvested power in time slots are independent and identically distributed. As a well-known practice, we employ two different MAC methods on the low density (size of inter-cluster nodes) and high density (size of intra-cluster nodes) setting in clustered network topology. Namely, TDMA MAC protocol for intra-cluster communication and the CSMA MAC protocol for determining the average power consumption of cluster heads in inter-cluster communication.

5.3 Transmission Slot Selection Policy with Reinforcement Learning

The underwater medium introduces uncertainties regarding water depth and environmental conditions, impacting the magnitude of underwater vibrations and movements, thereby influencing the harvested energy over time. Scheduling methods based on time slot allocation often overlook environmental factors that influence the availability of harvestable energy in the underwater environment when assigning time slots to nodes within the cluster. For this reason, nodes transmit within their scheduled time periods, regardless of available energy, resulting deficiency of the expected amount of energy harvested. In our model we adapt a multi agent cooperative reinforcement learning algorithm running on each node to autonomously select their time slot per data frame and learn from its past decision-making choices influenced by ambient conditions. In single agent multi armed bandit problem, agent makes choice among *k* different actions modeled as arms of a slot machine (Sutton & Barto, 2021). In this slot machine, each time after action is performed, the agent receives a reward depending on a probability distribution in arms. The value of actions is recorded as mean of rewards over time steps. The agent maintains the estimated value of each action *a* in each time step, as $Q_t(a)$. The aim of the agent is to maximize the cumulative rewards over repeated actions in time steps. In a single data frame, $Rt[a_i]$ denotes the reward received after action a_i is selected in time step *t* and the estimated value of action a_i after *t-1* is Qt[a_i] defined as

$$Q_t[a_i] = Q_t[a_i] + \frac{1}{t-1}(R_{t-1}[a_i] - Q_{t-1}[a_i]) \tag{25}$$

On the other hand, Multi-Agent Systems (MASs), focusing on developing algorithms that enable agents to adaptive learning and optimize their behavior in changing environments. Multi-agent Reinforcement Learning (MARL) aims to educate multiple agents simultaneously within a common environment by employing Reinforcement Learning (RL) techniques. In a fully cooperative setting, agents work together toward a shared goal, like assembling machines (Yuan et al., 2023). Figure 18 represents the multi-agent system, *n* distributed agents as underwater sensor nodes, cooperatively play the same multi-armed bandit game. Each agent has its own record of Q table as

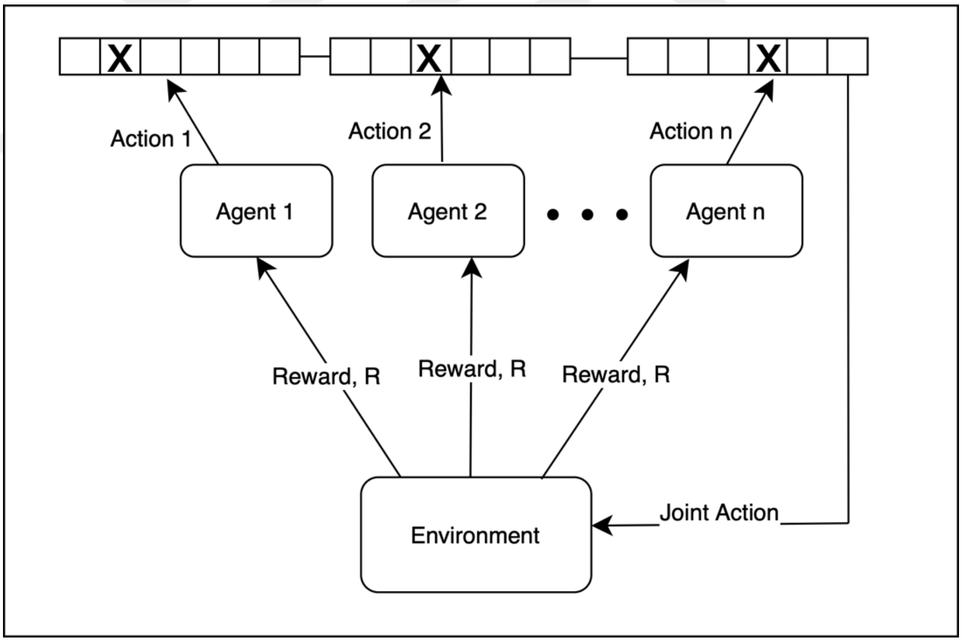

Figure 18. Training paradigm of multi agent reinforcement learning system.

time slot arrays. Agents choose actions independently, constructing a joint action array from received packets at the cluster head. We model k arms represent time slots, same as cluster size (total amount of cluster members). Note that each frame consists of k, i.e., TS many time slots. To find the best global combination of arms, agents wish to maximize a global return, in other words, minimize the overall regret. Associated with arms, mutually independent sequences of i.i.d. valued rewards are modeled. Specifically, the reward is the overall energy gained in the rest of arms except the arm selected for transmission. We run our model for a predetermined duration for a time slot. A detailed explanation of our method is given in Algorithm 3.

Algorithm 3 Distributed Multi-Agent Learning Algorithm for transmission slot selection

Require: $0 \leqslant \gamma \leqslant 1, \alpha \in (0,1]$, small $\varepsilon > 0, n \geq 0$
  Initialize $Q_i[a] \leftarrow 0 \forall a \in A_i$ and $\forall i \in N$
  Update Q table $Q_i[a] \leftarrow E_i^h(t) \forall t \in TS$
  for each time slot **t** in frame(TS) **do**
    for each node **n** in Cluster^k **do**
    if $random(n) \geq \varepsilon$ **then**
      Choose exploitation
      Select maximizing action $a_i = \arg\max Q_i[:]$
    else
      Choose exploration
      Select a random action a_i
    end if
    node **n** selects action a_i forms the joint action $\mathbf{a} \leftarrow \langle a^1, ..., a^N \rangle$
      Receive local reward from environment $R_t^n(a_i)$
    end for
    Return global reward from environment $R_t^N(\mathbf{a})$
    $R_t(a_i) \leftarrow \beta R_t^N(\mathbf{a}) + (1-\beta) R_t^n(a_i)$
    Update $Q_t[a_i] \leftarrow \alpha * (R_t(a_i) + \gamma * Q_t[a_{max}] - Q_{t-1}[a_i])$
  end for

We propose multi-agent reinforcement learning algorithm for the optimal policy in selection of time slots. Awareness of environmental conditions and observed energy fluctuations in ambient resources can be utilized in determination of timing for data transmission hence this approach potentially alleviates the load on the cluster head. Furthermore, the selection of the best transmission slots for each node in the cluster is a mutual interest of all nodes in the system, hence maximizing energy utilization in the cluster. With learning, each sensor can exploit the ambient resources to harvest energy to assure the maximal lifetime of the network. As stated in (Claus & Boutilier, 1998), convergence is guaranteed for a multi agent system if one state Q learning algorithm is used, and all agents have a common interest. In our method, we consider round based data collection. After one frame of data collection completes, all nodes in the cluster takes same reward for received number of data packets and local reward for sum of their harvested energy in the frame. We model RL as agents; every underwater sensor node in the network is an agent. Actions; is the selection of transmission slot to maximize the harvested energy during the current frame of data transmission. Transmission slot can be selected depending on the interaction of each sensor with its environment that effects the available energy. Transmission slots are designed as discrete time intervals. Action is the choice of time interval for data transmission. Reward calculation contains weighted sum of two parts namely the mutual reward and independent local reward. Mutual reward of a cluster member after selecting a transmission slot is total number of data packets arrived in the cluster head R_{CH} and it is broadcasted to cluster members at the end of each frame, the weight coefficient β is fixed and 0.4. Reward function penalizes conflicted slot selection with the decreased amount of received number of packets. Consequently, nodes adapt to their neighbor node's decision. The total harvested energy of the cluster member serves as the agent's local reward, calculated as the overall energy gain excluding the selected time interval for transmitting a single data frame. In simpler terms, it represents the accumulated harvested energy within the cluster during the idle time of nodes in the previous frame and defined as

$$Reward_n = \beta \sum_{t=1}^{TS} R_{CH} + (1 - \beta) \sum_{t=1}^{TS} E_{s_i}^{h}(t) \tag{26}$$

5.4 Performance Evaluation and Comparison

To compare efficiency of our transmission slot scheduling method with traditional TDMA approach. Additionally, the contribution of energy harvesting to the lifetime of a clustered underwater sensor network is also denoted in our simulations. We use Python for performance evaluation, stochastic model of piezoelectric energy harvesting and underwater communication channel in a round-based data collection system. We assume data is backlogged, i.e., nodes always have data to transmit to their cluster heads. Data packets are fixed in size and transmission power is adjusted according to the channel conditions and distance of cluster members to their respective cluster heads. We present performance comparison of 100 simulation runs in this section. In TDMA-based medium access methods, transmission slot of each node in the cluster is organized by cluster head and cluster members are notified about their scheduled slots for communication. We investigate the performance of TDMA and the effect of piezoelectric energy harvesting on clustered underwater sensor network.

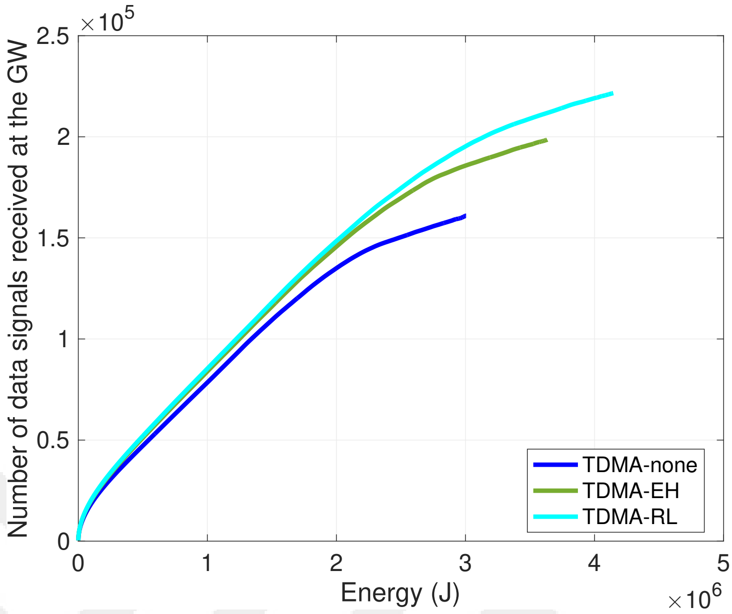

Figure 19. Number of data signals received at the surface station (GW) per energy.

TDMA-none curve indicates 100 nodes without energy harvesting equipment and using TDMA slot allocation for in-cluster communication. In TDMA-none curve cluster head schedules the transmission slots for cluster members. TDMA-EH curve presents 100 nodes with piezoelectric harvesting equipment and same transmission slot scheduling method as TDMA-none used. We present the performance of our method in TDMA-RL curve, transmission slot scheduling with cooperative learning and compare its efficiency to the traditional TDMA method. Figure 19 depicts that our proposed method improves energy utilization as more data signals received by the gateway station. In TDMA none curve, nodes perform transmissions in predetermined slots scheduled by cluster heads without energy harvesting equipment. TDMA-EH curve incorporates energy harvesting. It is observed that nodes collect more energy as in consecutive rounds. Therefore, compared to TDMA-none curve, TDMA-EH curve sends %20 more data signals to the gateway. TDMA-RL curve denotes the performance of our approach. As available energy changes during idle-tx states in frames, each node experience varying levels of energy within its surrounding environment. As the selection of transmission slots adapts to the harvesting conditions, the potential for collecting more energy during each frame increases. It is observed that %15 more data signals can be received by gateway in TDMA-RL curve compared to TDMA-EH curve. However, in traditional TDMA slot scheduling, nodes are unaware of their ambient conditions and energy is harvested in pre-established slots during frames and compared to the TDMA-none curve, it is observed that ~%35 more data signals can be transmitted by exploiting the available energy in underwater. Note that, y-axis in Figure 19 and 20 we present the total number of bits successfully received at the GW.

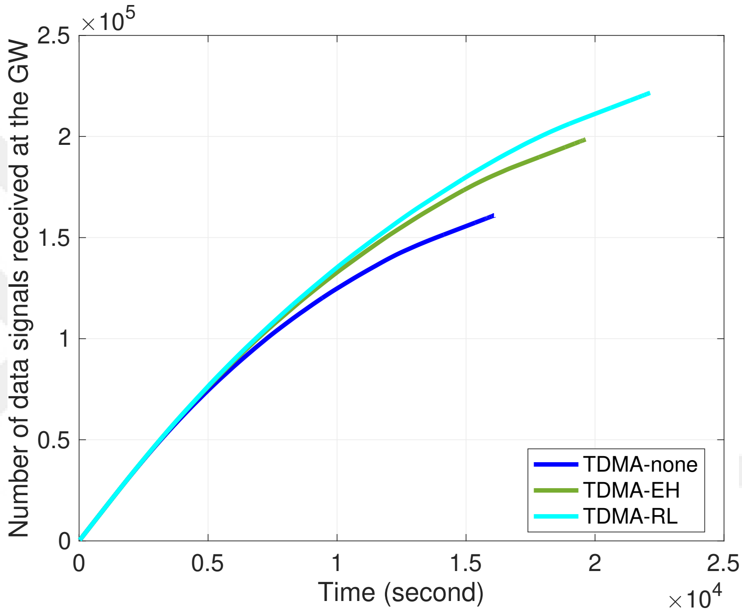

Figure 20. Number of data signals received at the surface station (GW) over time.

Figure 20 represents the performance of the previously mentioned methods in terms of throughput. It is observed that at the late stages of simulation in TDMA-RL method, more data signals received per unit of time compared to TDMA-none and TDMA-EH due to better utilization of available energy. It is also observed that all curves are identical in the early stages of the simulation. In the forthcoming rounds more energy is collected while nodes await idle and harvest energy until their communication slots in the cluster. Compared to TDMA-none method, in TDMA-EH and TDMA-RL, as more nodes remain active in rounds, %21 and %37 more data packets can be relayed to the gateway respectively. TDMA-RL and TDMA-EH curve have same characteristics while they collect more energy in successive rounds. On the other hand, TDMA-RL curve outperforms by selecting efficient slots to remain idle, observing the ambient available energy and selects the most inefficient time slot in terms of available energy and perform data transmission. Consequently, TDMA-RL captures more energy, whereas the TDMA-EH method neglects to exploit available energy by transmitting data during periods of high energy availability in time slots.

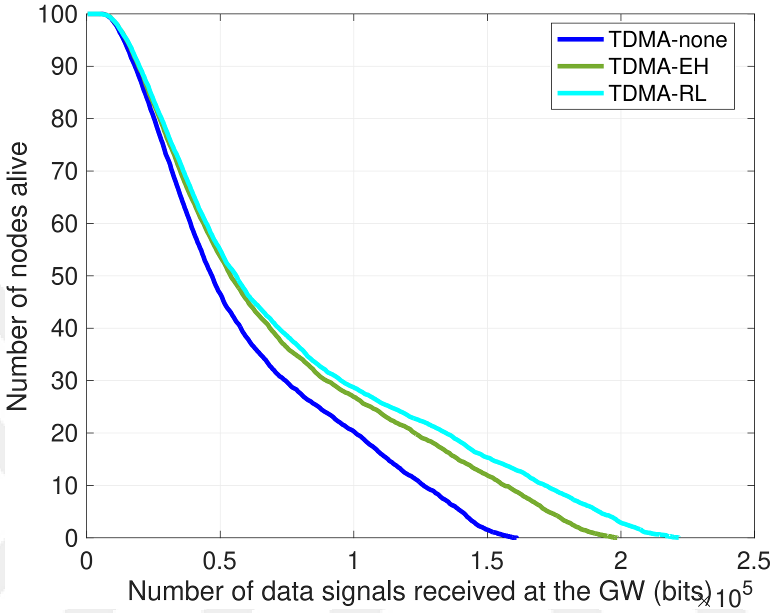

Figure 21. Number of data signals received at the surface station (GW) over time.

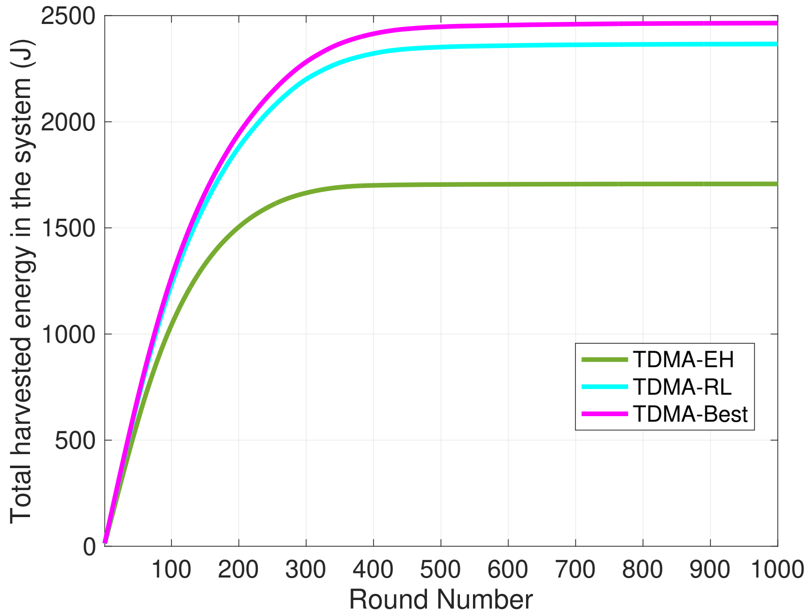

Figure 22. Total harvested energy in network in rounds.

Figure 21 shows the number of available nodes over data signals received at the base station. As more energy is captured from environment, lifetime duration of nodes is improved, and more data can be generated compared to TDMA-EH utilizing traditional TDMA approach. In TDMA slots, nodes communicate with their cluster heads without being aware of their ambient energy conditions. Therefore, less ambient energy can be collected. In the early phase of the simulation, it is observed that three methods have the same decrease rate in the number of alive nodes until %5 of nodes runs out of energy and few megabytes of data received at the base station. It is observed that if energy harvesting (EH) is involved in the system, number of data packet reception increases by %23 regardless of the medium access method. It is clear that harvested energy significantly contributes to the battery of nodes in successive rounds; hence nodes can perform more data transmission. TDMA-RL curve improves the amount of data packets by %11 compared to TDMA-EH method. We present the comparison of total harvested energy in the system in Figure 22. The TDMA-Best curve represents the maximum total energy captured by all nodes in the system, assuming nodes know least efficient transmission slots to harvest energy in advance. This curve serves as an upper bound for the total energy available in the environment. Since energy harvesting is not incorporated in the TDMA-none method, we compare TDMA-EH and TDMA-RL methods with this upper bound of available energy. As depicted in the plot, TDMA-RL captures %96 of the available energy in the system while TDMA-EH curve reveals total amount of energy in the system can only utilize the %56 of the available energy.

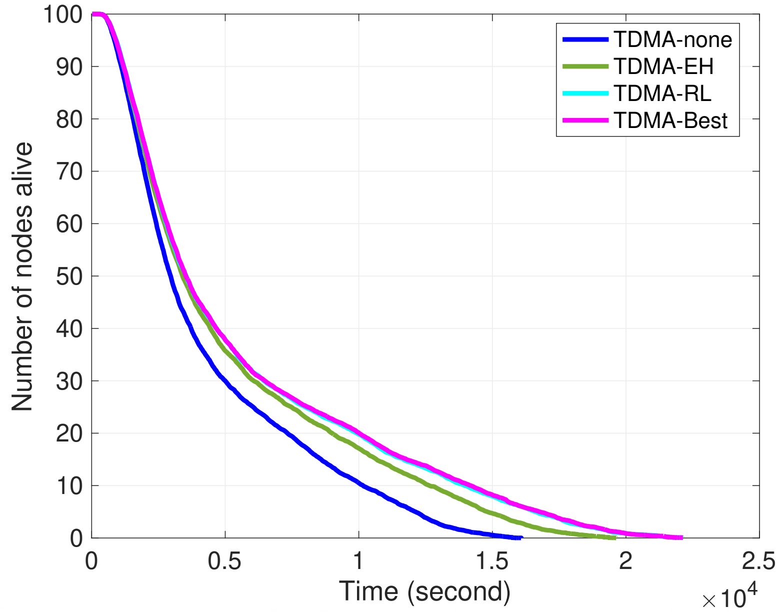

Figure 23. Number of alive nodes over simulation time steps.

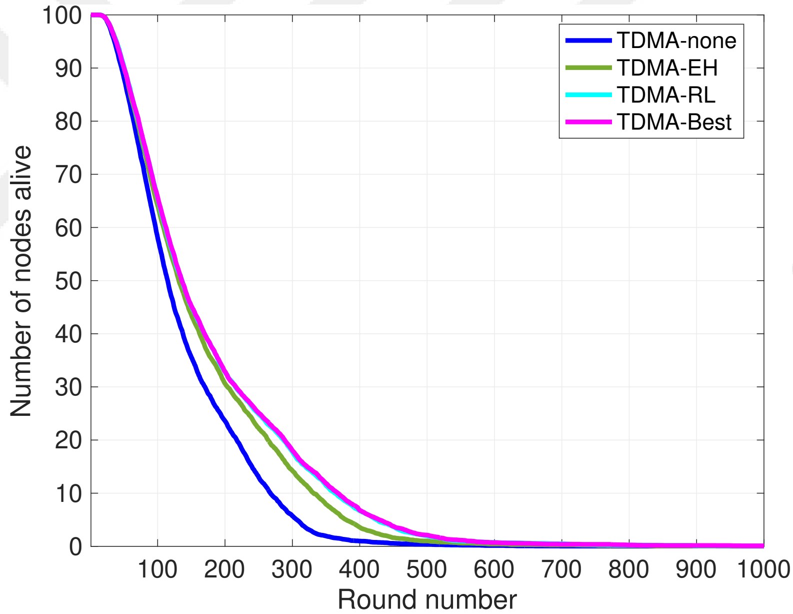

Figure 24. Number of alive nodes over simulation rounds.

Figure 23 and 24 depicts the performance of cooperative learning-based slot selection to the traditional TDMA protocol on lifetime basis. Considering the rounds consist of multiple frames, the duration of each round varies depending on the number of alive nodes. Both figures acknowledge that lifetime is significantly improved in case of cooperative selection of transmission slots in clusters. While energy harvesting improves the lifetime, adapting to the dynamic underwater environment and predicting the fluctuations in ambient energy it is observed that transmission slots can be selected with the aim of maximizing the utilization of the available energy in the environment. Furthermore, individual decision of transmission slot selection on each node minimizes energy wastage from overhearing, while also relaxing the need for strict synchronization. It's noted that in both figures, the curves for TDMA-Best and TDMA-RL exhibit marginal disparities, yet without significant alterations in the number of active nodes per unit time or across rounds. The network lifetime defined as the duration from the initialization of sensor network until the depletion of the last node's battery. To represent our contribution to lifetime, the performance of each method with statistics of 100 simulation run is depicted in Table 5. Results presented in Table 6, clearly illustrate the beneficial effect of TDMA-RL in increasing the number of active nodes over the course of rounds and enhance overall longevity of the network.

Table 7

Network Lifetime Over 100 Runs of Simulations.

FND	Min	Max	Median	Mean	Std
TDMA-none	12	46	25	25.53	6.5758
TDMA-EH	13	48	25.5	26.83	7.1873
TDMA-RL	13	48	25	25.87	7.0877
TDMA-Best	13	46	26	26.77	7.0437
HND					
TDMA-none	81	154	113	114.27	17.419
TDMA-EH	90	192	128.5	131.14	21.968
TDMA-RL	91	202	128.5	134.6	24.324
TDMA-Best	90	201	132.5	135.61	24.39
LND					
TDMA-none	222	1340	355.5	388.43	138.36
TDMA-EH	249	1500	442.5	475.38	168
TDMA-RL	286	1500	498.5	536	198.45
TDMA-Best	285	285	510	542.15	205.73

Note that due to the amount of harvested energy and early steps of learning in the beginning of rounds, TDMA-EH, TDMA-RL, TDMA-Best methods have similar central tendency in FND. HND is improved %14 due to the amount of harvested energy which incorporates the traditional TDMA method. HND in TDMA-EH is improved by %2 with TDMA-RL method. As more energy is collected towards the last stages of rounds, we observe the efficiency of our method on LND increased by %25 and %40 in TDMA-EH and TDMA-RL respectively.

5.5 Discussion and Future Work Directions

Underwater sensor networks encounter unique challenges, such as long propagation delays, limited bandwidth, and high attenuation, which can significantly decrease the performance of communication protocols. Therefore, traditional contention-free methods, specifically Time Division Multiple Access (TDMA) is commonly used in underwater sensor networks due to its simplicity and robustness in such demanding environments. Furthermore, clustering methods often rely on TDMA due to its efficiency of scheduling the intra-cluster communication, minimizing collisions and energy consumption, which is primary concern for the power-constrained nodes in underwater conditions. On the other hand, the dynamic nature of underwater environments causes significant variability in energy harvesting conditions. In our method, the primary objective is to augment the accumulated energy of underwater sensor nodes by dynamically selecting TDMA time slots. We specifically emphasized integration of harvesting awareness into the transmission slot selection strategy. Our investigation revolves around assessing the impact of harvesting-adaptive slot scheduling in response to varying level of available energy in underwater. Our results show that both TDMA-EH and TDMA-RL improve the network's initial robustness and longevity compared to the baseline TDMA method. The impact of energy harvesting in TDMA-EH extends the network's operational period, reflected in higher minimum and mean values for First Node Dead (FND). TDMA-RL further enhances performance by dynamically adapting time slots based on available energy, thus better sustaining network operations over time. However, TDMA-RL includes the added complexity of reinforcement learning, we would expect it to outperform the simpler TDMA-EH consistently. This similarity might indicate that the learning mechanism is not significantly improving the earliest phases of node operation over what is achieved by simple energy harvesting alone. One future investigation of a guided exploration strategy is to apply better actions in the early stages of the simulation.

In terms of Half Node Dead (HND), TDMA-EH and TDMA-RL significantly outperform the baseline, indicating that they can maintain a larger portion of the network for longer periods. TDMA-RL represents slightly higher mean HND, demonstrating the advantages of adaptive scheduling. On the other hand, TDMA-RL shows a higher standard deviation in both HND (24.324) and LND (198.45) compared to TDMA-EH (21.968 and 168, respectively). This suggests greater variability in network performance, which might be due to the adaptive nature of the method leading to inconsistent results in different simulation runs. While adaptive methods can improve performance, it shows that they can also define variability. Q-learning, specifically in the initial phase, is still learning and exploring the environment. The Last Node Dead (LND) results further confirm the superiority of TDMA-RL, achieving the highest median and mean LND values, indicating its effectiveness in prolonging the overall network lifetime. The TDMA-Best method is the ideal benchmark, setting the highest performance and illustrating the maximum potential of the network with optimal time-slot assignments. This benchmark emphasizes the effectiveness of the TDMA-RL method, showing that adaptive learning strategies can approach the optimal performance levels.

Overall, our study highlights the significant benefits of incorporating energy harvesting and adaptive scheduling into UASNs. TDMA-RL stands out for its ability to dynamically manage energy and time-slot allocation, leading to substantial improvements in network longevity and efficiency. The results suggest promising directions for future research, focusing on refining these methods and integrating advanced communication technologies to further enhance network performance and reliability.

As part of our future work directions, overhead of efficient information sharing and decentralized coordination mechanisms will be investigated while maintaining effective coordination among sensor nodes. We will then compare our harvesting-aware method in terms of throughput and energy utilization, thereby providing a comprehensive evaluation of our proposed approach. Moreover, the efficiency and performance of our transmission slot selection strategy can be improved with advanced multi-agent reinforcement learning algorithms, such as Proximal Policy Optimization (PPO) and Deep Q-Networks (DQN), along with their complexity, can be investigated. On the other hand, we intend to incorporate variations in channel state to offer a more realistic approach that accounts for the influence of energy harvesting. On the other hand, the rapid development of new communication techniques such as Rate-Splitting Multiple Access (RSMA) and Non-Orthogonal Multiple Access (NOMA) present opportunities. Furthermore, hybrid approaches, combining the strengths of multiple methodologies, such as integrating TDMA with non-orthogonal multiple access (NOMA) techniques is also a promising. To date, these techniques have not been extensively investigated in underwater communication systems. As a future research direction, their potential applicability in underwater environments may be explored. By pursuing these research directions, we aim to contribute to the advancement of UASNs, enhancing their performance and efficiency through innovative communication techniques and more accurate modeling of environmental factors.

5.6 Chapter Conclusion

In this study, we introduce a time slot scheduling policy and improve medium access control (MAC) protocol designed specifically for energy harvesting underwater acoustic sensor networks (EH-UASNs). The main goal of this policy is to optimize the network's lifespan while enhancing the throughput. We utilize a clustered topology leveraging its duty cycling and data aggregation nature to enhance the efficiency of network energy utilization. We investigate a Time Division Multiple Access (TDMA) scheme for intra-cluster communication. TDMA optimizes energy usage by enabling inactive nodes to transition into sleep or idle modes until their allocated time slot for communication arrives. Our main goal is to enhance the efficiency of energy harvesting within clusters of the network by leveraging the allocation of time slots. W introduce a novel fully cooperative multi agent reinforcement learning algorithm to schedule time slots of TDMA MAC protocol for EH-UASNs to maximize harvested energy and enhance throughput by adaptive time slot allocation in intra cluster nodes.

Chapter 6
Conclusions

In this thesis we addressed the stochastic energy harvesting process at underwater acoustic sensor nodes, specifically focusing on the energy-aware routing problem in underwater acoustic sensor networks (UASNs). Acoustic communication is essential in underwater environments due to its ability to provide long-range communication over distances typically spanning kilometers. Despite this advantage, UASNs encounter challenges due to the high energy demands of acoustic communication and the difficulty of replacing and recharging batteries underwater. Their potential to deliver critical environmental data for various applications, including environmental studies, early warning systems, and industrial monitoring renders continuous delivery of information is essential. Consequently, energy management in underwater sensor networks becomes a crucial resource allocation issue. To tackle this challenge, cluster-based network routing protocols have been extensively studied for their ability to minimize energy consumption. Cluster heads (CHs) are leveraged to aggregate data, thereby reducing overall energy use and extending the network's lifespan. In contrast to other cluster head selection techniques in the literature that primarily relying on centralized or optimization methods that prioritize remaining energy and node locality, we consider the stochastic nature of available energy and propose a robust cluster head selection policy. We introduced a novel policy for selecting cluster heads (CH). As a robust solution, we utilized moving average prediction on the foreseeable energy and constructed score formulization for candidate cluster heads incorporating the total energy dissipation for the data collection task, estimated harvested energy and remaining energy within a cluster. Besides, proximity to the surface sink, and local connectivity is also considered. Performance results showed that our proposed policy formed energy-balanced clusters and captured more ambient energy, thereby enhancing the network's lifespan. Numerical findings demonstrate that this approach significantly diminishes energy outage and prolongs the network's lifespan, shedding light on the importance of optimizing energy harvesting in cluster-based routing strategies.

Focusing on the exploitation of available ambient energy, we extend our cluster head selection policy and introduced a novel reinforcement learning-based method to form better clusters in terms of size and energy. We considered cluster head selection problem in multiple aspects, including node positions, remaining energy, anticipated harvested energy, and learning from past cluster head decisions within the cluster. We modeled cluster head selection problem as an optimization framework and implement a multi agent reinforcement learning method. In our performance evaluations we showed the effect of piezoelectric energy harvesting with temporal uncertainty in underwater environments, and we proved that the lifetime is more enhanced if estimated available energy of each individual node is considered in the cluster head selection algorithm. Our work is the first study in the literature that considers stochastic nature of energy harvesting in the process of cluster head selection to maximize lifetime. Our proposed technique leverages the available energy in the underwater environment, reducing energy consumption and extending both throughput and lifespan of the UASN by %30.

As a future work, we plan to consider different energy harvesting processes and the impact of underwater acoustic channel conditions to improve the lifetime of UASNs. Our method is proposed for stable link conditions, i.e., the required transmission power on a specific distance to obtain a certain SNR value is fixed over time, that is overlooked in practical perspective. Therefore, our method can be further enhanced in reliability perspective. In addition to that, with the advancements of reinforcement learning, cross-layer optimizations such as adaptive relay strategies consider physical layer parameters, medium access strategies and QoS can be considered while designing cluster head selection and cluster-based data routing algorithms. Mobility, on the other hand, is another design challenge in UASNs. Due to water currents and aquatic life, sensor nodes deployed in specific area of interest can shift and cause changes in topology of the network. Therefore, clustering methods that are adaptive to dynamic changes in the topology is a potential future research direction.

In our recent study, we proposed a novel multi-agent reinforcement learning based algorithm for intra-cluster TDMA time slot scheduling problem addressed in clustered underwater acoustic sensor networks to improve sustainability in long-term monitoring applications. In the literature this is the first study that exploits ambient energy in underwater in the context of medium access strategy. In literature most studies consider TDMA as medium access protocol for intra-cluster communications.

Due to energy efficiency, simplicity and fairness of contention-free TDMA protocol, it is widely utilized in clustered UASNs. However, time slot allocation strategy is investigated in most studies while cluster head is responsible for coordinating the communication in the cluster. Leveraging the distinct available energy states of dynamic underwater medium, we implemented a distributed, random-access strategy to reserve available time slots adaptively to node's surrounding environment that is initially calculated by cluster heads. We propose a learning algorithm for scheduling transmissions, where nodes autonomously determine their communication slots within their cluster by learning from their ambient energy resources and harvesting opportunities. Our method employs a distributed approach to alleviate the coordination responsibility and burden of scheduling on the cluster heads. Considering spatial-temporal uncertainty in available energy of piezoelectric harvester model, we demonstrate that the proposed adaptive scheme on the decision of time slots extends the network's lifespan by %15 while incorporation of energy harvesting only contributes %10 to the lifetime. To this end, we demonstrate the benefits of energy harvesting on traditional contention-free medium access methods. In future work, we plan to incorporate underwater channel state information and enhance our approach by considering hybrid methods that utilize contention-based channel access. We will also investigate the impact of this combined approach on throughput and energy efficiency.

The demand for underwater monitoring and exploration applications necessitates to design robust and energy efficient mechanisms, adapting to the significant constraints and challenges posed by harsh underwater environment. Advanced modulation techniques should be implemented in the microcontrollers used in UASN nodes to achieve reduced bit error rates in an energy-efficient manner to improve bandwidth utilization. New mechanisms and approaches are needed to fully leverage hardware advancements in the MAC layer. In the routing layer, to cope with frequent changes in the topology and reliability, state-of-the-art self-configuring techniques should be integrated by considering the constraints of UASNs into both existing routing protocols mainly developed for WSNs. Despite significant achievements in UASN research in close literature, several major challenges remain unresolved. Essentially, the development of simulation test beds incorporating the underwater propagation with ray tracing models is crucial for analyzing the performance of proposed methods in large-scale underwater communications and sensing. These developments are vital for accurate system characterization and performance analysis. Therefore, to enhance sustainability in UASNs, it is vital to improve realistic channel and mobility models, localization techniques, deployment methods to be further developed.

REFERENCES

Abraham, R., & Vadivel, M. (2023a). An Energy Efficient Wireless Sensor Network with Flamingo Search Algorithm Based Cluster Head Selection. *Wireless Personal Communications*, *130*(3), 1503–1525. https://doi.org/10.1007/s11277-023-10342-2

Abraham, R., & Vadivel, M. (2023b). An Energy Efficient Wireless Sensor Network with Flamingo Search Algorithm Based Cluster Head Selection. *Wireless Personal Communications*, *130*(3), 1503–1525. https://doi.org/10.1007/s11277-023-10342-2

Ahmed, G., Zou, J., Zhao, X., & Fareed, M. M. S. (2017). Markov chain model-based optimal cluster heads selection for wireless sensor networks. *Sensors (Switzerland)*, *17*(3). https://doi.org/10.3390/s17030440

Akyildiz, I. F., Pompili, D., & Melodia, T. (2005a). Underwater acoustic sensor networks: Research challenges. *Ad Hoc Networks*, *3*(3), 257–279. https://doi.org/10.1016/j.adhoc.2005.01.004

Akyildiz, I. F., Pompili, D., & Melodia, T. (2005b). Underwater acoustic sensor networks: Research challenges. *Ad Hoc Networks*, *3*(3), 257–279. https://doi.org/10.1016/j.adhoc.2005.01.004

Akyıldız Fuat, I., & Vuran Can, M. (2010). *Wireless Sensor Networks*.

Alamu, O., Olwal, T. O., & Djouani, K. (2023). Energy harvesting techniques for sustainable underwater wireless communication networks: A review. *E-Prime - Advances in Electrical Engineering, Electronics and Energy*, *5*. https://doi.org/10.1016/j.prime.2023.100265

Alfouzan, F. A., Shahrabi, A., Ghoreyshi, S. M., & Boutaleb, T. (2019). An energy-conserving collision-free mac protocol for underwater sensor networks. *IEEE Access*, *7*, 27155–27171. https://doi.org/10.1109/ACCESS.2019.2901646

Awan, K. M., Shah, P. A., Iqbal, K., Gillani, S., Ahmad, W., & Nam, Y. (2019). Underwater Wireless Sensor Networks: A Review of Recent Issues and Challenges. In *Wireless Communications and Mobile Computing* (Vol. 2019). Hindawi Limited. https://doi.org/10.1155/2019/6470359

Basagni, S., Di Valerio, V., Gjanci, P., & Petrioli, C. (2018). Harnessing HyDRO: Harvesting-aware data ROuting for underwater wireless sensor networks. *Proceedings of the International Symposium on Mobile Ad Hoc Networking and Computing (MobiHoc)*, 271–279. https://doi.org/10.1145/3209582.3209610

Bereketli, A., & Bilgen, S. (2012). Remotely powered underwater acoustic sensor networks. *IEEE Sensors Journal*, *12*(12), 3467–3472. https://doi.org/10.1109/JSEN.2012.2210401

Bholowalia, P., & Kumar, A. (2014). EBK-Means: A Clustering Technique based on Elbow Method and K-Means in WSN. In *International Journal of Computer Applications* (Vol. 105, Issue 9).

Bozorgi, S. M., Shokouhi Rostami, A., Hosseinabadi, A. A. R., & Balas, V. E. (2017). A new clustering protocol for energy harvesting-wireless sensor networks. *Computers and Electrical Engineering*, *64*, 233–247. https://doi.org/10.1016/j.compeleceng.2017.08.022

Chen, Y., Jin, Z., Xing, G., Zeng, Q., Chen, Y., Zhou, Z., & Yang, Q. (2023). An Energy-Efficient MAC Protocol for Three-Dimensional Underwater Acoustic Sensor Networks With Time Synchronization and Power Control. *IEEE Access*, *11*, 20842–20860. https://doi.org/10.1109/ACCESS.2023.3250452

Chen, Y.-D., Lien, C.-Y., Chuang, S.-W., & Shih, K.-P. (2011). *DSSS: A TDMA-based MAC Protocol with Dynamic Slot Scheduling Strategy for Underwater Acoustic Sensor Networks*.

Claus, C., & Boutilier, C. (1998). *The Dynamics of Reinforcement Learning in Cooperative Multiagent Systems*. www.aaai.org

Dhongdi, S., Bhandari, A., Singh, J., Kachhadia, S., & Joshi, V. (2018). Joint Clustering and Routing Protocol for 3-D Underwater Acoustic Sensor Network. *International Conference on Ubiquitous and Future Networks, ICUFN*, *2018-July*, 415–420. https://doi.org/10.1109/ICUFN.2018.8436816

Domingo, M. C., & Prior, R. (2007). A distributed clustering scheme for underwater wireless sensor networks. *IEEE International Symposium on Personal, Indoor and Mobile Radio Communications, PIMRC*. https://doi.org/10.1109/PIMRC.2007.4394038

Dror, E., Avin, C., & Lotker, Z. (2013). Fast randomized algorithm for 2-hops clustering in vehicular ad-hoc networks. *Ad Hoc Networks*, *11*(7), 2002–2015. https://doi.org/10.1016/j.adhoc.2012.02.006

Dwi Widodo HK, Adit Kurniawan, & M Sigit Arifianto. (2019). *Improving Topology of LEACH Cluster Using Reinforcement Learning Method*.

Erdem, H. E., & Gungor, V. C. (2017, June 27). Lifetime analysis of energy harvesting underwater wireless sensor nodes. *2017 25th Signal Processing and Communications Applications Conference, SIU 2017*. https://doi.org/10.1109/SIU.2017.7960419

Erdem, H. E., Yildiz, H. U., & Gungor, V. C. (2019). On the Lifetime of Compressive Sensing Based Energy Harvesting in Underwater Sensor Networks. *IEEE Sensors Journal*, *19*(12), 4680–4687. https://doi.org/10.1109/JSEN.2019.2900427

Eris, C., Gul, O. M., & Boluk, P. S. (2023). An Energy-Harvesting Aware Cluster Head Selection Policy in Underwater Acoustic Sensor Networks. *2023 International Balkan*

Conference on Communications and Networking, BalkanCom 2023. https://doi.org/10.1109/BalkanCom58402.2023.10168000

Eriş, Ç., Gül, Ö. M., & Bölük, P. S. (2024). A novel reinforcement learning based routing algorithm for energy management in networks. *Journal of Industrial and Management Optimization*, *0*(0), 0–0. https://doi.org/10.3934/jimo.2024049

Eris, C., Gungor, V. C., & Boluk, P. S. (2016). Analysis of battery-powered sensor node lifetime for smart grid applications. *2016 24th Signal Processing and Communication Application Conference, SIU 2016 - Proceedings*. https://doi.org/10.1109/SIU.2016.7496190

Eris, C., Saimler, M., Gungor, V. C., Fadel, E., & Akyildiz, I. F. (2014). Lifetime analysis of wireless sensor nodes in different smart grid environments. *Wireless Networks*, *20*(7), 2053–2062. https://doi.org/10.1007/s11276-014-0723-0

Fei, W., Hexiang, B., Deyu, L., & Jianjun, W. (2020). Energy-Efficient Clustering Algorithm in Underwater Sensor Networks Based on Fuzzy C Means and Moth-Flame Optimization Method. *IEEE Access*, *8*, 97474–97484. https://doi.org/10.1109/ACCESS.2020.2997066

Felemban, E. (2020). Acoustic Frequency Optimization for Underwater Wireless Sensor Network. In *IJACSA) International Journal of Advanced Computer Science and Applications* (Vol. 11, Issue 6). www.ijacsa.thesai.org

Felemban, M., & Felemban, E. (2013). *Energy-Delay Tradeoffs for Underwater Acoustic Sensor Network*. http://www.emadfelemban.net/uwsnmodel

Freitag, L., Stojanovic, M., Grund, M., & Singh, S. (2005). *Acoustic Communications for Regional Undersea Observatories*.

Frikha, M. S., Gammar, S. M., Lahmadi, A., & Andrey, L. (2021). Reinforcement and deep reinforcement learning for wireless Internet of Things: A survey. In *Computer Communications* (Vol. 178, pp. 98–113). Elsevier B.V. https://doi.org/10.1016/j.comcom.2021.07.014

Fujino, K., Sanada, K., & Mori, K. (2019). Enhanced Adaptive Cluster Control for Energy Harvesting Wireless Sensor Networks under Geographical Non-uniform Energy Harvesting Conditions; Enhanced Adaptive Cluster Control for Energy Harvesting Wireless Sensor Networks under Geographical Non-uniform Energy Harvesting Conditions. In *2019 International Conference on Electronics, Information, and Communication (ICEIC)*.

Gabale, V., Raman, B., Dutta, P., & Kalyanraman, S. (2013). A classification framework for scheduling algorithms in wireless mesh networks. *IEEE Communications Surveys and Tutorials*, *15*(1), 199–222. https://doi.org/10.1109/SURV.2012.022412.00068

Gawade, R. D., & Nalbalwar, S. L. (2016). A centralized energy efficient distance based routing protocol for wireless sensor networks. *Journal of Sensors*, *2016*. https://doi.org/10.1155/2016/8313986

Ge, Y., Nan, Y., & Guo, X. (2021). Maximizing network throughput by cooperative reinforcement learning in clustered solar-powered wireless sensor networks. *International Journal of Distributed Sensor Networks*, *17*(4). https://doi.org/10.1177/15501477211007411

Gül, Ö. M. (2014). *A LOW-COMPLEXITY, NEAR-OPTIMAL SCHEDULING POLICY FOR SOLVING A RESTLESS MULTI-ARMED BANDIT PROBLEM OCCURRING IN A SINGLE-HOP WIRELESS NETWORK A THESIS SUBMITTED TO THE GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES OF MIDDLE EAST TECHNICAL UNIVERSITY*.

Gul, O. M., & Demirekler, M. (2018). Asymptotically throughput optimal scheduling for energy harvesting wireless sensor networks. *IEEE Access*, *6*, 45004–45020. https://doi.org/10.1109/ACCESS.2018.2865451

Gul, O. M., & Erkmen, A. M. (2020). Energy-efficient cluster-based data collection by a uav with a limited-capacity battery in robotic wireless sensor networks. *Sensors (Switzerland)*, *20*(20), 1–35. https://doi.org/10.3390/s20205865

Gul, O. M., & Erkmen, A. M. (2023). Energy-Aware UAV-Driven Data Collection With Priority in Robotic Wireless Sensor Network. *IEEE Sensors Journal*, *23*(15), 17667–17675. https://doi.org/10.1109/JSEN.2023.3286877

Gul, O. M., & Uysal-Biyikoglu, E. (2014). *UROP: A Simple, Near-Optimal Scheduling Policy for Energy Harvesting Sensors*.

Han, B., Ran, F., Li, J., Yan, L., Shen, H., & Li, A. (2022). *A Novel Adaptive Cluster Based Routing Protocol for Energy-Harvesting Wireless Sensor Networks*.

Han, G., Wang, X., Sun, N., & Liu, L. (2020). A collision-free MAC protocol based on quorum system for underwater acoustic sensor networks. *Proceedings - 2020 16th International Conference on Mobility, Sensing and Networking, MSN 2020*, 671–676. https://doi.org/10.1109/MSN50589.2020.00112

Han, M., Duan, J., Khairy, S., & Cai, L. X. (2020). Enabling Sustainable Underwater IoT Networks with Energy Harvesting: A Decentralized Reinforcement Learning Approach. *IEEE Internet of Things Journal*, *7*(10), 9953–9964. https://doi.org/10.1109/JIOT.2020.2990733

Heinzelman, W. B., Chandrakasan, A. P., & Balakrishnan, H. (2002). An application-specific protocol architecture for wireless microsensor networks. *IEEE Transactions on Wireless Communications*, *1*(4), 660–670. https://doi.org/10.1109/TWC.2002.804190

Heinzelman, W. R., Chandrakasan, A., & Balakrishnan, H. (2000). Energy-efficient communication protocol for wireless microsensor networks. *Proceedings of the Hawaii International Conference on System Sciences*, 223. https://doi.org/10.1109/hicss.2000.926982

Jagannath, J., Saji, A., Kulhandjian, H., Sun, Y., Demirors, E., & Melodia, T. (2013). A hybrid MAC protocol with channel-dependent optimized scheduling for clustered underwater acoustic sensor networks. *Proceedings of the 8th ACM International Conference on Underwater Networks and Systems, WUWNet 2013*. https://doi.org/10.1145/2532378.2532382

Jarosz, E., Teague, W. J., Book, J. W., & Beşiktepe, Ş. (2011). On flow variability in the Bosphorus Strait. *Journal of Geophysical Research: Oceans*, *116*(8). https://doi.org/10.1029/2010JC006861

Jha, A. V., Appasani, B., Khan, M. S., & Song, H. H. (2024). A Novel Clustering Protocol for Network Lifetime Maximization in Underwater Wireless Sensor Networks. *IEEE Transactions on Green Communications and Networking*. https://doi.org/10.1109/TGCN.2024.3375011

Jornet, J. M., Stojanovic, M., & Zorzi, M. (2010). On joint frequency and power allocation in a cross-layer protocol for underwater acoustic networks. *IEEE Journal of Oceanic Engineering*, *35*(4), 936–947. https://doi.org/10.1109/JOE.2010.2080410

Josko A. Catipovic. (1990). *Performance Limitations in Underwater Acoustic Telemetry*.

Khan, A., Khan, M., Ahmed, S., Iqbal, N., Abd Rahman, M. A., Abdul Karim, M. K., Mustafa, M. S., & Yaakob, Y. (2021). EH-IRSP: Energy Harvesting Based Intelligent Relay Selection Protocol. *IEEE Access*, *9*, 64189–64199. https://doi.org/10.1109/ACCESS.2020.3044700

Khan, M. F., Bibi, M., Aadil, F., & Lee, J. W. (2021a). Adaptive node clustering for underwater sensor networks. *Sensors*, *21*(13). https://doi.org/10.3390/s21134514

Khan, M. F., Bibi, M., Aadil, F., & Lee, J. W. (2021b). Adaptive node clustering for underwater sensor networks. *Sensors*, *21*(13). https://doi.org/10.3390/s21134514

Lee, J., & Jindal, N. (2009). Energy-efficient scheduling of delay constrained traffic over fading channels. *IEEE Transactions on Wireless Communications*, *8*(4), 1866–1875. https://doi.org/10.1109/T-WC.2008.080037

Li, K., Huang, H., Gao, X., Wu, F., & Chen, G. (2019, August 5). QLEC: A machine-learning-based energy-efficient clustering algorithm to prolong network lifespan for IoT in high-dimensional space. *ACM International Conference Proceeding Series*. https://doi.org/10.1145/3337821.3337926

Li, L., Qiu, Y., & Xu, J. (2022). A K-Means Clustered Routing Algorithm with Location and Energy Awareness for Underwater Wireless Sensor Networks. *Photonics*, *9*(5).

https://doi.org/10.3390/photonics9050282

Li, Y., Wang, Y., Ju, Y., & He, R. (2014). Energy efficient cluster formulation protocols in clustered underwater acoustic sensor networks. *Proceedings - 2014 7th International Conference on BioMedical Engineering and Informatics, BMEI 2014*, 923–928. https://doi.org/10.1109/BMEI.2014.7002904

Liu, L., Cai, L., Ma, L., & Qiao, G. (2021). Channel State Information Prediction for Adaptive Underwater Acoustic Downlink OFDMA System: Deep Neural Networks Based Approach. *IEEE Transactions on Vehicular Technology*, *70*(9), 9063–9076. https://doi.org/10.1109/TVT.2021.3099797

Luo, J., Chen, Y., Wu, M., & Yang, Y. (2021). A Survey of Routing Protocols for Underwater Wireless Sensor Networks. In *IEEE Communications Surveys and Tutorials* (Vol. 23, Issue 1, pp. 137–160). Institute of Electrical and Electronics Engineers Inc. https://doi.org/10.1109/COMST.2020.3048190

Mathankumar, M., Thirumoorthi, P., & Viswanathan, T. (2021). An Improved Clustering Scheme for Underwater Sensor Network using Bayesian Linear Regression. *2021 International Conference on Advancements in Electrical, Electronics, Communication, Computing and Automation, ICAECA 2021*. https://doi.org/10.1109/ICAECA52838.2021.9675767

McKenna, M. F., Ross, D., Wiggins, S. M., & Hildebrand, J. A. (2012). Underwater radiated noise from modern commercial ships. *The Journal of the Acoustical Society of America*, *131*(1), 92–103. https://doi.org/10.1121/1.3664100

Molins, M., & Stojanovic, M. (2006). Slotted FAMA: A MAC protocol for underwater acoustic networks. *OCEANS 2006 - Asia Pacific*. https://doi.org/10.1109/OCEANSAP.2006.4393832

Muthukkumar, R., Garg, L., Maharajan, K., Jayalakshmi, M., Jhanjhi, N., Parthiban, S., & Saritha, G. (2022). A genetic algorithm-based energy-aware multi-hop clustering scheme for heterogeneous wireless sensor networks. *PeerJ Computer Science*, *8*. https://doi.org/10.7717/PEERJ-CS.1029

Nakas, C., Kandris, D., & Visvardis, G. (2020). Energy efficient routing in wireless sensor networks: A comprehensive survey. In *Algorithms* (Vol. 13, Issue 3). MDPI AG. https://doi.org/10.3390/a13030072

Narayan, V., Malaviya, M. M., Madan, D., Malaviya, M., Daniel, A. K., & Chaturvedi, P. (2022). *FGWOA: An EEcient Heuristic for Cluster Head Selection in WSN using Fuzzy based Grey Wolf Optimization Algorithm FGWOA: An Efficient Heuristic for Cluster Head Selection in WSN using Fuzzy based Grey Wolf Optimization Algorithm*. https://doi.org/10.21203/rs.3.rs-1725228/v1

Nguyen, N. T., Le, T. T. T., Nguyen, H. H., & Voznak, M. (2021). Energy-efficient clustering multi-hop routing protocol in a UWSN. *Sensors (Switzerland)*, *21*(2), 1–16. https://doi.org/10.3390/s21020627

Omeke, K. G., Abubakar, A. I., Zhang, L., Abbasi, Q. H., & Imran, M. A. (2023). How Reinforcement Learning is Helping to Solve Internet-of-Underwater-Things Problems. *IEEE Internet of Things Magazine*, *5*(4), 24–29. https://doi.org/10.1109/iotm.001.2200129

Omeke, K. G., Mollel, M. S., Ozturk, M., Ansari, S., Zhang, L., Abbasi, Q. H., & Imran, M. A. (2021). DEKCS: A Dynamic Clustering Protocol to Prolong Underwater Sensor Networks. *IEEE Sensors Journal*, *21*(7), 9457–9464. https://doi.org/10.1109/JSEN.2021.3054943

Pengwei Li, Shilian Wang, Hao Zhang, & Eryang Zhang. (2017). Improved Particle Swarm Optimization Algorithm of Clustering in Underwater Acoustic Sensor Networks. *Improved Particle Swarm Optimization Algorithm of Clustering in Underwater Acoustic Sensor Networks*.

Pobering, S., & Schwesinger, N. (2004). A novel hydropower harvesting device. *Proceedings - 2004 International Conference on MEMS, NANO and Smart Systems, ICMENS 2004*, 480–485. https://doi.org/10.1109/icmens.2004.1508997

Pompili, D., Melodia, T., & Akyildiz, I. F. (2009). A CDMA-based Medium Access Control for UnderWater Acoustic Sensor Networks. *IEEE Transactions on Wireless Communications*, *8*(4), 1899–1909. https://doi.org/10.1109/TWC.2009.080195

Pour, S. E., & Javidan, R. (2021). A new energy aware cluster head selection for leach in wireless sensor networks. *IET Wireless Sensor Systems*, *11*(1), 45–53. https://doi.org/10.1049/wss2.12007

Redhu, S., & Hegde, R. M. (2019). *Multi-Sensor Data Fusion for Cluster-based Data Aggregation in IoT Applications*.

Ren, Q., & Yao, G. (2021). Enhancing Harvested Energy Utilization for Energy Harvesting Wireless Sensor Networks by an Improved Uneven Clustering Protocol. *IEEE Access*, *9*, 119279–119288. https://doi.org/10.1109/ACCESS.2021.3108469

Sah, D. K., & Amgoth, T. (2020). A novel efficient clustering protocol for energy harvesting in wireless sensor networks. *Wireless Networks*, *26*(6), 4723–4737. https://doi.org/10.1007/s11276-020-02351-x

Sah, D. K., Hazra, A., Kumar, R., & Amgoth, T. (2023a). Harvested Energy Prediction Technique for Solar-Powered Wireless Sensor Networks. *IEEE Sensors Journal*, *23*(8), 8932–8940. https://doi.org/10.1109/JSEN.2022.3208730

Sah, D. K., Hazra, A., Kumar, R., & Amgoth, T. (2023b). Harvested Energy Prediction

Technique for Solar-Powered Wireless Sensor Networks. *IEEE Sensors Journal*, *23*(8), 8932–8940. https://doi.org/10.1109/JSEN.2022.3208730

Sah, D. K., Hazra, A., Kumar, R., & Amgoth, T. (2023c). Harvested Energy Prediction Technique for Solar-Powered Wireless Sensor Networks. *IEEE Sensors Journal*, *23*(8), 8932–8940. https://doi.org/10.1109/JSEN.2022.3208730

Saha, A. K., Uddin Khan, M. S., & Rahman, M. M. (2019). Region based low energy adaptive clustering hierarchy (R-LEACH) protocol: Prolonging the lifetime of WSN. *ACM International Conference Proceeding Series*, 127–132. https://doi.org/10.1145/3362966.3362978

Sandeep, D. N., & Kumar, V. (2017). Review on Clustering, Coverage and Connectivity in Underwater Wireless Sensor Networks: A Communication Techniques Perspective. In *IEEE Access* (Vol. 5, pp. 11176–11199). Institute of Electrical and Electronics Engineers Inc. https://doi.org/10.1109/ACCESS.2017.2713640

Sathyamoorthy, M., Kuppusamy, S., Dhanaraj, R. K., & Ravi, V. (2022). Improved K-Means Based Q Learning Algorithm for Optimal Clustering and Node Balancing in WSN. *Wireless Personal Communications*, *122*(3), 2745–2766. https://doi.org/10.1007/s11277-021-09028-4

Senturk, I. F. (2021). Clustered mobile data collection in WSNs: An energy-delay trade-off. *Turkish Journal of Electrical Engineering and Computer Sciences*, *29*(7), 3180–3196. https://doi.org/10.3906/ELK-2104-25

Stefanov, A., & Stojanovic, M. (2011a). Design and performance analysis of underwater acoustic networks. *IEEE Journal on Selected Areas in Communications*, *29*(10), 2012–2021. https://doi.org/10.1109/JSAC.2011.111211

Stefanov, A., & Stojanovic, M. (2011b). Design and performance analysis of underwater acoustic networks. *IEEE Journal on Selected Areas in Communications*, *29*(10), 2012–2021. https://doi.org/10.1109/JSAC.2011.111211

Sun, Y., Zheng, M., Han, X., Li, S., & Yin, J. (2022a). Adaptive clustering routing protocol for underwater sensor networks. *Ad Hoc Networks*, *136*. https://doi.org/10.1016/j.adhoc.2022.102953

Sun, Y., Zheng, M., Han, X., Li, S., & Yin, J. (2022b). Adaptive clustering routing protocol for underwater sensor networks. *Ad Hoc Networks*, *136*. https://doi.org/10.1016/j.adhoc.2022.102953

Sutton, R. S., & Barto, A. G. (2021). *Reinforcement learning : an introduction*.

Tomasi, B., & Preisig, J. C. (2015). Energy-Efficient Transmission Strategies for Delay Constrained Traffic with Limited Feedback. *IEEE Transactions on Wireless Communications*, *14*(3), 1369–1379. https://doi.org/10.1109/TWC.2014.2365819

Tomovic, S., & Radusinovic, I. (2023). DR-ALOHA-Q: A Q-Learning-Based Adaptive MAC Protocol for Underwater Acoustic Sensor Networks. *Sensors*, *23*(9). https://doi.org/10.3390/s23094474

Uyan, O. G., Akbas, A., & Gungor, V. C. (2023). Machine learning approaches for underwater sensor network parameter prediction. *Ad Hoc Networks*, *144*. https://doi.org/10.1016/j.adhoc.2023.103139

Vermaak, H. J., Kusakana, K., & Koko, S. P. (2014a). Status of micro-hydrokinetic river technology in rural applications: A review of literature. In *Renewable and Sustainable Energy Reviews* (Vol. 29, pp. 625–633). https://doi.org/10.1016/j.rser.2013.08.066

Vermaak, H. J., Kusakana, K., & Koko, S. P. (2014b). Status of micro-hydrokinetic river technology in rural applications: A review of literature. In *Renewable and Sustainable Energy Reviews* (Vol. 29, pp. 625–633). https://doi.org/10.1016/j.rser.2013.08.066

Wang, C., Wang, Z., Sun, W., & Fuhrmann, D. R. (2017). Reinforcement Learning-Based Adaptive Transmission in Time-Varying Underwater Acoustic Channels. *IEEE Access*, *6*, 2541–2558. https://doi.org/10.1109/ACCESS.2017.2784239

Wang, Ying., Zhou, Mengqi., Tan, Tieniu., & Institute of Electrical and Electronics Engineers. (2009). *Proceedings, 2009 IEEE Youth Conference on Information, Computing and Telecommunication : September 20-21, 2009 : Beijing, China*. IEEE Press.

Wu, J., Sun, X., Wu, J., & Han, G. (2021). Routing strategy of reducing energy consumption for underwater data collection. *Intelligent and Converged Networks*, *2*(3), 163–176. https://doi.org/10.23919/icn.2021.0012

Xing, G., Chen, Y., Hou, R., Dong, M., Zeng, D., Luo, J., & Ma, M. (2021). Game-Theory-Based Clustering Scheme for Energy Balancing in Underwater Acoustic Sensor Networks. *IEEE Internet of Things Journal*, *8*(11), 9005–9013. https://doi.org/10.1109/JIOT.2021.3055857

Ye, X., Yu, Y., & Fu, L. (2022). Deep Reinforcement Learning Based MAC Protocol for Underwater Acoustic Networks. *IEEE Transactions on Mobile Computing*, *21*(5), 1625–1638. https://doi.org/10.1109/TMC.2020.3029844

Younis, O., & Fahmy, S. (2004). HEED: A hybrid, energy-efficient, distributed clustering approach for ad hoc sensor networks. *IEEE Transactions on Mobile Computing*, *3*(4), 366–379. https://doi.org/10.1109/TMC.2004.41

Yuan, L., Zhang, Z., Li, L., Guan, C., & Yu, Y. (2023). *A Survey of Progress on Cooperative Multi-agent Reinforcement Learning in Open Environment*. http://arxiv.org/abs/2312.01058

ZHANG Gaoxi XIAO Hwee-Pink TAN, P., Zhang, P., Xiao, G., & Tan, H.-P. (2013). Clustering Algorithms for Maximizing the Lifetime of Wireless Sensor Networks with

Energy-Harvesting Sensors. In *Computer Networks* (Vol. 57, Issue 4). https://ink.library.smu.edu.sg/sis_research/2960

Zhao, T., Xu, M., Xiao, X., Ma, Y., Li, Z., & Wang, Z. L. (2021). Recent progress in blue energy harvesting for powering distributed sensors in ocean. In *Nano Energy* (Vol. 88). Elsevier Ltd. https://doi.org/10.1016/j.nanoen.2021.106199

Zhaohui Wang, Chaofeng Wang, & Wensheng Sun. (2015). *Adaptive Transmission Scheduling in Time-Varying Underwater Acoustic Channels*.

Zhu, J., Chen, Y., Sun, X., Wu, J., Liu, Z., & Xu, X. (2021). ECRKQ: Machine Learning-Based Energy-Efficient Clustering and Cooperative Routing for Mobile Underwater Acoustic Sensor Networks. *IEEE Access*, *9*, 70843–70855. https://doi.org/10.1109/ACCESS.2021.3078174

Zou, H. X., Li, M., Zhao, L. C., Gao, Q. H., Wei, K. X., Zuo, L., Qian, F., & Zhang, W. M. (2021). A magnetically coupled bistable piezoelectric harvester for underwater energy harvesting. *Energy*, *217*. https://doi.org/10.1016/j.energy.2020.119429