Received: 7 May 2016 | Revised: 18 October 2016 | Accepted: 14 November 2016

DOI: 10.1002/dac.3257

WILEY

RESEARCH ARTICLE

LRP: Link quality-aware queue-based spectral clustering routing protocol for underwater acoustic sensor networks

Muhammad Faheem[1] | **Gurkan Tuna**[2] | **Vehbi Cagri Gungor**[1]

[1] Department of Computer Engineering, Abdullah Gül University, Kayseri, Turkey

[2] Department of Computer Programming, Trakya University, Edirne, Turkey

Correspondence
Gurkan Tuna, Department of Computer Programming, Trakya University, Edirne 22020, Turkey.
Email: gurkantuna@trakya.edu.tr

Funding information
Turkish Scientific and Technical Research Council, Grant/Award Number: 114E248

Summary

Recently, underwater acoustic sensor networks (UASNs) have been considered as a promising approach for monitoring and exploring the oceans in lieu of traditional underwater wireline instruments. As a result, a broad range of applications exists ranging from oil industry to aquaculture and includes oceanographic data collection, disaster prevention, offshore exploration, assisted navigation, tactical surveillance, and pollution monitoring. However, the unique characteristics of underwater acoustic communication channels, such as high bit error rate, limited bandwidth, and variable delay, lead to a large number of packet drops, low throughput, and significant waste of energy because of packets retransmission in these applications. Hence, designing an efficient and reliable data communication protocol between sensor nodes and the sink is crucial for successful data transmission in underwater applications. Accordingly, this paper is intended to introduce a novel nature-inspired evolutionary link quality-aware queue-based spectral clustering routing protocol for UASN-based underwater applications. Because of its distributed nature, link quality-aware queue-based spectral clustering routing protocol successfully distributes network data traffic load evenly in harsh underwater environments and avoids hotspot problems that occur near the sink. In addition, because of its double check mechanism for signal to noise ratio and Euclidean distance, it adopts opportunistically and provides reliable dynamic cluster-based routing architecture in the entire network. To sum up, the proposed approach successfully finds the best forwarding relay node for data transmission and avoids path loops and packet losses in both sparse and densely deployed UASNs. Our experimental results obtained in a set of extensive simulation studies verify that the proposed protocol performs better than the existing routing protocols in terms of data delivery ratio, overall network throughput, end-to-end delay, and energy efficiency.

KEYWORDS

energy efficiency, evolutionary approaches, routing protocol, spectral clustering, underwater acoustic sensor network

1 | INTRODUCTION

Routing, an integral part of underwater acoustic sensor networks (UASNs), faces several challenges because of the unique nature of underwater environments. Existing underwater routing solutions mainly focus on the traditional techniques that are not appropriate to obtain high-energy efficiency and low-memory requirement in case of large-scale network deployments. These routing solutions are generally not used in many underwater applications where link reliability is a critical issue due to inaccurate approximation for the delay and distance measurements. Moreover, most of the existing routing schemes designed for a specific underwater scenario do not consider loop-free topological issues and node over assignment costs while carrying information. All these above-mentioned factors intensely provide a platform in which a resource-aware routing approach plays a key role in accomplishing numerous application requirements with highly dynamic environmental challenges. Understanding the facts, significant importance is given to the development of reliable routing schemes to provide an effective route discovery between the source nodes and sink.

Energy-efficient routing protocols for ad hoc and terrestrial sensor networks (AHTSNs) have been heavily investigated in recent years.[1–8] However, existing terrestrial energy-efficient routing solutions are inappropriate for these networks because of the nature of the underwater environment and the requirements of time-critical applications. As a result, the idea of using tiny sensor nodes equipped with the capability of wireless communication in underwater environments has received increasing interest in the fields of commercial exploitation and scientific exploration, and hence UASNs have been used for many different applications including monitoring, control, surveillance, and measurement.[9] On the other hand, in some specific applications such as marine surveillance and oceanographic data collection, highly efficient time-critical and temporal-spatial continuous aquatic environment observing systems are essential. Even though the UASNs and AHTSNs have certain common properties, UASN applications significantly differ from the traditional ground-based sensor network applications. First, in the UASN applications, acoustic communication replaces radio communications since radio communications exhibit extremely diverse travel time and characteristics in underwater environments. Second, in the UASN applications, sensor nodes may move with water flows while most ground sensor nodes are immobile. Third, the UASN topology is highly dynamic due to the nature of underwater channels, and acoustic channel features high propagation delay, excessive bit error rate (BER), and low data rate.[10] Bit error rate is extremely high in the underwater environment compared to the AHTSNs because of extreme interference due to multipath data pass and unpredictable nature of underwater acoustic channels.

Although there are alternatives to overwhelm the difficulties encountered by acoustic wave based systems such as cabling, electromagnetic waves, and optical signals to witness under water natural resources, they have their own limitations.[11] For instance, cable-based systems call for high installation and maintenance cost along with frequency calibration. On the other hand, electromagnetic wave bands have extremely excessive attenuation in underwater wireless channels. Although just a minor portion of the long electromagnetic wave bands might pass through with relatively less attenuation, eg, 6-10 m distance with frequency 122KHz containing data rate of 1-8 Kbps. In addition, huge aerials and high-transmission power are also needed.[12] Optical signals seem to be one of the most prominent choices for underwater network scenarios and have been recognised as one of the most efficient and generic approaches to deal with high transfer data rates ranging from a few kilobits per second to megabits per second and for a distance of up to 100 m. However, the short-range optical signals get scattered poorly in the underwater environment. Therefore, they are not promising due to the highly dynamic network topology, the high absorption, and required high accuracy.[13] Moreover, although in recent years, many long-range optical modems for point-to-point communications have been developed for underwater wireless sensor networks (UWSNs); they are extremely costly; and thus for large-scale deployments where from a few hundreds to thousands sensor nodes are needed, this choice seems not feasible.

A review of the recent literature shows that, up till now, employing acoustic signals is the best choice for practical implementations of UWSNs because of their less attenuation, long distance communication with sophisticated reliability, and low cost for large deployments. Unfortunately, only a KHz bandwidth is available with a range of 100 m for sensor nodes for their effective communication, thus the bandwidth restriction is the major problem for UASNs. However, for the acoustic signals, available communication bandwidth is very restricted. Furthermore, like other types of underwater signals, acoustic signals also have to face high-attenuation problems and that upsurge with the rise in the frequency range.[9] In case of a long range, absorption also increases and this drops signal to noise ratio (SNR) severely. In addition, since the speed of acoustic signals is extremely slow in saline water (nearly 1500 m/s), which is $2 \times 100\,000$ as compared to electromagnetic waves, propagation delay is rather high.[10] Because of the above-mentioned challenging properties of the environment and the unique nature of the underwater applications, the existing routing schemes designed for the terrestrial sensor networks are not directly appropriate for UWSNs. In the past, many simple underwater-observing systems were introduced. However, they were only suitable for the small-scale networks and failed to achieve performance when the system generated huge data volumes in time-critical applications.

To sum up, the UWSN and more specifically UASN applications require more energy-efficient network solutions with extended buffering capacity to avoid congestion issues, availability in clustering and routing failures, repairing ability with low delay, and energy consumption for reliable data delivery. Therefore, in this study, we focus on the development of a link quality-aware queue-based spectral clustering routing protocol. Mainly, we have 4 contributions to the literature. First, we propose an energy-efficient evolutionary spectral clustering algorithm to solve clustering problems for reliable data collection over variable quality links. Second, to proportionally distribute energy load in the entire network by taking into account the nodes connectivity issues, we develop an energy-efficient evolutionary cluster-head selection algorithm. Third, we propose an energy-efficient evolutionary routing algorithm that follows a tree-like structure to provide reliable data gathering in UASNs. Fourth, an extensive set of simulations and experimental studies carried out in this study show that our proposed routing approach performs better than existing routing protocols in terms of major performance parameters such as data delivery ratio, end-to-end delay, network throughput, and energy efficiency.

The rest of the paper is organised as follows. An extensive literature survey is presented in Section 2. Section 3 introduces the statement of the problem addressed in this paper. Section 4 investigates methodologies and approaches to address the problem and explains the proposed approach. Section 5 reports simulation model and results of a set of simulation studies performed to prove the effectiveness of the proposed routing approach. Finally, the paper is concluded in Section 6.

2 | RELATED WORK

Because of the unique nature of the underwater environments and the requirements of time-critical applications, routing for reliable data transmission is a key issue in UWSN research. In the literature, a great deal of efforts have been devoted to routing protocols since their performances differ depending on the nature of underwater applications and the architecture of the underlying networks. As in 1 study,[14] a depth-based routing protocol called DBR is proposed to provide robust data packet delivery in UWSNs. In DBR, observed information is routed based on local depth information of the nodes. Depth-based routing protocol is suitable for small-sized networks where up to a few hundreds sensor nodes are involved to explore undersea miracles. However, in large network deployments, it experiences control message overheads due to excessive rerouting in addition to network delay because of information holding time at each hop-energy and high-energy consumption because of the lack of a mechanism, which considers appropriate link quality among nodes in the routing process.

In another study,[15] an energy-efficient depth-based routing protocol called EEDBR for UWSNs is discussed. In the knowledge acquisition phase of EEDBR, each sensor node shares its residual energy and depth with its neighbours. In the data-forwarding phase, each sender includes a list of its neighbouring nodes to the data packets. The set of the neighbouring nodes called forwarding set is selected based on the depth of the neighbouring nodes to carry information towards the sink. EEDBR performs well in terms of data delivery and energy consumption and does not require the localization of the sensor nodes in UWSNs. However, upon receiving data packets, forwarding nodes hold the packets for a certain time, and this increases delay in the network.

The authors in 1 study[16] proposed a distributed routing protocol for non-time-critical, long-term aquatic monitoring applications for underwater environments. The proposed routing protocol uses a data aggregation technique to eliminate redundant information in UWSNs, proactively decreases routing message exchanges, and avoids flooding. It achieves a very high packet delivery ratio (PDR) while significantly increasing the network throughput and reducing the network overhead. However, in dense deployments, it cannot distribute data traffic load evenly and cannot address memory overrun issues.

In another study,[17] authors try to minimise energy consumption of routing processes by using a novel trimming mechanism. The proposed trimming mechanism significantly improves PDR by avoiding additional spreading of the forwarded data packets. However, it consumes a considerable amount of energy because of forwarding a large number of redundant data packets in the network. Moreover, it fails to preserve highly reliable links among sensor nodes and hence excessive rerouting is experienced in the network.

The authors in 1 study[18] solve the problem of least network delay for UASN-based underwater applications. The proposed approach relies on multipath routing to improve data delivery robustness from a source node towards the sink. However, this multipath routing at lower layers increases redundant data packets, which lead to node buffer overflow problem in the network. Moreover, it fails to exploit new data paths in voids and shadow zones due to lack of considering dynamic adjustment power mechanism in UASNs.

A novel multipath energy-efficient routing approach is proposed in 1 study.[19] The approach efficiently integrates multipath communications to eliminate retransmission attempts and enhance reliability level in UWSNs. To reduce the transmission cost, it calculates overall packet error rate to decide the number of multiple paths that guarantee the desirable packet error rate. However, the proposed approach encounters high-memory overflow due to its flooding nature. In addition, because of lack of considering the dynamic data traffic load distribution, the nodes lie in cross layers die quickly. Thus, a huge amount of data packets can be lost because of voids.

A layer-based distributed routing protocol that addresses routing table complexity, node failures, and routing overheads is proposed.[20] The proposed protocol uses the idea of per-contact routing to carry data packets in UWSNs and significantly increases PDR in both dense and sparse deployments. However, especially in large-sized networks, during joining a new node in each new round and in a case when a source node cannot get a response from the other nodes, it has to wait for a notable amount of time. Therefore, network delay increases linearly from the source node towards the destination in the network.

A link-state based adaptive feedback routing protocol for UASNs is proposed.[21] The proposed protocol takes into account 3-dimensional (3D) direction of underwater sensor nodes, and the impact of beam width to get data transmission routes effectively. It is able to reduce the power consumption of the underwater sensor nodes and avoids energy consumption, which results from frequent routing table updates. However, it experiences control and data packet collisions due to inappropriate synchronization in the network.

In 1 study,[22] a reliable and energy-efficient routing protocol called R-ERP^{2}R is proposed for UASNs. The proposed protocol utilises distance as a routing metric to balance energy consumption among sensor nodes. It also considers link quality towards the forwarding nodes to provide required reliability and takes into account the residual energy of the forwarding nodes to extend network lifetime. However, its performance is degraded by the nonconnective interval of the layers and in sparse network deployments it occasionally fails to provide accurate measurement of the link quality and residual energy of the sensor nodes.

The authors in another study[23] propose a layer-by-layer angle-based flooding routing protocol to handle continuous sensor node movements and address end-to-end delay and energy consumption. For this goal, without using any explicit location information or configuration, every node can calculate its flooding angle to forward data packets toward the sinks. Although based on a flooding-based technique, it increases the reliability of the network, due to harsh underwater environments; it fails to calculate appropriate angles for the flooding zones to prevent unnecessary flooding of packets over the entire network. Thus, it degrades effectiveness of the minimum energy consumption.

The work in 1 study[24] proposes a delay-sensitive routing approach for UWSNs. The proposed approach employs an optimal weight function to compute transmission losses and speed of received signal. It achieves efficient data-forwarding performance and minimal relative transmissions in low-depth regions and decreases end-to-end delay with a small reduction in total network throughput. However, it fails to equally distribute energy load in the network.

In 1 study,[25] the authors propose an energy-efficient delay-aware routing protocol for UWSNs. The proposed protocol counts on a distance-bandwidth relationship and effectively handles intermittent connectivity problems. It involves an adaptable energy-efficient forwarding node selection mechanism and reduces collision rates in UWSNs. However, the increase in the number of relay nodes increases the channel interference related issues. Moreover, if multiple relay nodes are appointed at each layer, the network encounters high data redundancy.

In another study,[26] an adaptive hop-by-hop vector-based routing protocol is proposed. In the proposed protocol, during a transmission process, the radius of the virtual pipeline is changed hop-by-hop to restrict the forwarding range of data packets. Also, for energy efficiency, the transmission power is adjusted adaptively hop-by-hop in a cross-layer fashion. Moreover, the proposed protocol selects forwarding nodes based on the distance from current node to destination node. Therefore, it effectively reduces end-to-end delays. Finally, it guarantees transmission reliability even in sparse regions and reduces duplicate packets in dense regions.

Geographic forwarding based on geospatial division and greedy geographic forwarding based on geospatial division are proposed in 1 study.[27] In both of the proposed schemes, the entire underwater network is first divided into small cube spaces. Then, the data packets are transmitted as unit of cube spaces by selecting optimal sensor nodes as next hop nodes in 1 cube space. At the expense of longer paths and larger delay, the proposed schemes are able to achieve more route paths and expense less energy.

In another study,[28] authors proposed a pressure sensor-based reliable routing protocol for UWSNs. The proposed protocol benefits from a set of metrics provided by link quality estimator along with depth distance and residual energy level to select a single reliable node for data forwarding. Nevertheless, each sensor node holds information for a prominent time during finding an effective next hop relay, thus it degrades effectiveness of the minimum delay. Moreover, the proposed protocol does not consider interference issues during forwarding data packets in the network.

In another study,[29] a geographic and opportunistic routing protocol that routes data packets from sensor nodes to multiple sinks is proposed. To overcome the void regions, it offers a depth adjustment based topology control approach, which moves the void nodes to new depths. The proposed routing protocol significantly improves the network performance in both sparse and dense network deployments. However, it consumes a significant amount of energy and holding time at each hop to find routes in void regions.

In 1 study,[30] an intracluster and intercluster communication architecture for data aggregation in UWSNs is proposed. The proposed architecture uses fuzzy approach to effectively select cluster heads and estimate cluster size in UWSNs. Employing this approach, the network is able to provide energy-efficient data transmission between the sensor nodes and the clusters to extend the network lifetime. Although the proposed architecture performs well in terms of energy consumption, end-to-end packet delay and PDR, it is unable to address major issues such as appropriate cluster size, load balancing, and data redundancy in UWSNs when applied in a densely deployed scenario.

3 | PROBLEM STATEMENT

Routing in UASNs poses many challenges because of the peculiarities of the underwater channel. These challenges include excessive and mutable propagation delays, restricted bandwidth, high BERs, multipath effects, and temporary losses of connectivity. Though routing protocols for UWSNs exist, their scopes are limited, and the majority of those solutions (see Section 2) have been designed to meet application-specific design objectives and requirements in a particular scenario. Although these studies guide design decisions for exploring and monitoring UWSN-based applications, most of the existing routing schemes achieve 1 or 2 of these design objectives at the expense of others. For example, energy saving during data collection introduces excessive delay and overhead in the network. Moreover, they generally ignore the impact of external interference and noise on transmission reliability in harsh underwater environments. Furthermore, they ignore the large memory requirements for the time critical applications, which generate a huge amount of data in underwater environments. Therefore, the routing in UWSNs in the context of reliable data delivery with low-energy consumption and delay is still an important problem depending on the event characteristics. The existing network solutions due to their fixed or inefficient routing strategies fail to handle the underwater acoustic communication challenges; and thus, they are not resilient or efficient enough to provide desired reliable data delivery for UWSN-based applications. Therefore, in this study, particular importance is given to design a reliable link quality aware evolutionary queue-based spectral clustering routing scheme for UASNs. Extensive simulations and experimental studies performed confirm that our approach outperforms the existing well-known routing schemes in terms of high PDR, network throughput, and low network delay and energy consumption. The obtained results of the simulations also verify that the designed routing solution is efficient, robust, and practical for small-sized and large-sized network deployments.

4 | PROPOSED LRP SCHEME

4.1 | Role of evolutionary queue-based spectral clustering approaches in LRP

For the last couple of decades, Metaheuristics have been used to create a robust search tool to handle optimization and local optima problems in various domains. The evolutionary computation, a form of meta-heuristic, has become one of the most popular techniques in both theoretical analysis and industrial applications and is used as an optimization and search tool in several domains.[31] It is a fast growing interdisciplinary research field developed beyond its original meaning of "biological evolution" and has been successfully applied to a number of classes of complex combinatorial problems to solve large and dynamic issues in a unified framework. For solving real-world issues that are usually very challenging for the conventional computing techniques, the main purpose of the nature inspired-evolutionary algorithms (EAs) and techniques is to identify the mechanism of such complex computational systems and to develop an extremely malleable, proficient, and robust algorithms with a realistic execution time near optimal results.[32] In EAs, the fitness function is responsible for measuring the quality of an individual; the individual with superior quality has more chance to be elected as a solution of the initial population. Generally, a fitness function is a particular type of objective function used to measure the closeness of a design solution to achieving the set aims. The typical EA starts with the main loop of a selection procedure by granting that the better individuals have more opportunity to breed by imitating natural selection. In the mating pool, several available variation operators also can be used to alter recent information. The variation operator is named as mutation if the gene of an individual changes on itself. The variation operator is named recombination or crossover if the gene interchange is completed between 2 or more individuals. Generally, EA follows user-defined instructions under which recombination and mutation are made. These above-mentioned advantages of the evolutionary computation approaches make them extremely appropriate for the complex problems associated with the highly unpredictable environment. If the existing best solution is unpredictable in continuously changeable problematic environment, then there might be an alternative solution in the population, which fits the existing environment well.

Although clustering is indeed a challenging research problem, because of its distinct advantages, it has become an active topic of research in recent years. In the last couple of years, a new approach has started to get much attention namely spectral method that can produce high-quality clusterings. Spectral clustering is an attractive approach since it is easy to implement and reasonably fast and therefore has been applied in a wide-ranging set of applications to solve complex clustering problems.[33] Also, intrinsically, they do not suffer from the problem of local optima. Usually, the spectral methods for clustering involve taking the top eigenvectors of some matrix based on the distance or SNR between sensor nodes and then split them into multiple clusters where each sensor node belongs to only 1 cluster. Generally, spectral clustering algorithm forms an $n \times n$ affinity matrix and compute eigenvectors of this matrix with least computational complexity. Thus, for a given set of sensor nodes $\{Sn_i\}_{i=1}^{n}$, $Sni \in R^{1\times d}$, the main goal of the spectral clustering is to distribute the sensor nodes into disjoint groups based on their similarities between all pairs of sensor nodes Sni and Snj, where R is $l \times d$ matrix. In spectral clustering algorithm, the point of departure is a weighted similarity graph $g(v, e)$, where the weights corresponding to the pairwise similarities and vertices correspond to sensor nodes. Each edge (e_i, e_j) in the graph is assigned a weight $w(e_i, e_j)$ reflecting the original similarity matrix A between Sni and Snj as

$$w(e_i, e_j) = \begin{cases} \exp\left(-\dfrac{\left\|Sni - Snj^2\right\|}{2\sigma^2}\right) & \text{for } i \neq j \\ 0 & \text{for } i = j \end{cases}, \quad (1)$$

while the diagonal degree matrix D of sensor node Sni and Snj can be computed as

$$D = \sum_{j=1}^{n} w(Sni \cdot Snj), \quad (2)$$

in which $i = 1, 2, \ldots, n$ and σ is the scaling parameter indicates how fast $w(e_i, e_j)$ falls off with the Ecludian distance

(Ed) between Sni and Snj. The evolutionary spectral clustering techniques are based on this weight form Laplacian graph, which transforms the problem into a tractable eigenvector problem. To perform bipartitioning of the given data where large numbers of clusters are created, the developed algorithm uses the thresholding notion for the second eigenvector of the Laplacian. The eigenvector with the second least eigenvalue is nominated; however, to find clustering solution directly, herein developed algorithm uses multiple eigenvectors considering normalise cut (N_{cut}) described as

$$N_{cut}(Sni, \mathrm{Snj}) = \frac{cut(Sni, \mathrm{Snj}) + cut(Snj, Sni)}{w(Sni, Snj)}. \tag{3}$$

Particularly, the relaxation for N_{cut} is based on rewriting Equation 2 as a normalised quadratic form involving indicator vectors. These indicator vectors are then replaced with real-valued vectors, resulting in a generalised eigenvector problem that can be conveniently summarised in terms of the normalised graph Laplacian L of A defined as follows:

$$L = D^{-\frac{1}{2}}(D - A)D^{-\frac{1}{2}}, \tag{4}$$

in which $D = diag\{d_1, d_2, \ldots, d_n\}$ with $d_i = \sum_{j=1}^{n} w(e_i, e_j)$, $\mathrm{i} = 1, 2, \ldots, n$.

In addition, recently conducted experimental tests reveal that organizing sensor nodes in multiple centralised queues improve the clustering performance mainly thanks to its load-sharing characteristic. Also, it allows mutually exclusive access to all sensor nodes in queues to avoid bottleneck by reducing the number of processors in entire clustering network. Moreover, with the use of a round robin assignment strategy, the arriving sensor nodes are assigned to ready queues in a cyclic fashion, and the clustering performance is significantly improved compared to the traditional clustering schemes.

By inspiring the above-mentioned advantages of the evolutionary, spectral clustering, and queue-based approaches, we developed an energy efficient evolutionary queue-based clustering routing protocol, which is extremely appropriate to solve complex clustering and routing problems associated with the highly dynamic underwater environment.

4.2 | Network model

In this research, network model is represented as an undirected graph $g(v, \mathrm{e})$, where g is weight, v is the set of all underwater sensor nodes vertices denoted by $v = \{Sn_1, Sn_2, , \ldots, Sn_n\}$, and e represents the set of all connected edges $e = \{e_1, e_2, ..., e_n\}$ among the sensor nodes. For each pair of sensor node $Sni, \mathrm{Snj} \in \mathrm{v}$, a similarity $S(Sni, \mathrm{Snj}) = \mathrm{S}(Snj, \mathrm{Sni}) \geq 0$ is given. The similarities $\mathrm{S}(Sni, \mathrm{Snj})$ can be viewed as weights on the undirected edges $e(Sni, \mathrm{Snj})$ of a graph g over S. The matrix $A = [S(Sni, \mathrm{Snj})]$ plays the role of a "real-valued" adjacency matrix for g. The set of edges between 2 disjoint cluster sets $E, F \subseteq \mathrm{v}$ is called the edge cut or the normalise cut between E, F. A clustering $C = \{C1, C2, \ldots, Ck\}$ is a partitioning of v into the nonempty mutually disjoint subsets $\mathrm{C1}, C2, \ldots, Ck$. In the perspective of graph theory, a clustering represents a normalise multiway cut in the graph g. The distance among sensor node Sni and Snj can be denoted as $Sni \in Snj, \mathrm{d}(Sni, Snj) \leq d_{th}$, and the delay can be defined as $Sni \in Snj, D_e(Sni, Snj) \leq D_{e(th)}$. Herein, d_{th} and $D_{e\ (th)}$ show the distance and delay threshold, respectively. A 3D view of the designed network model is shown in Figure 1. For easiness, the designed 3D model is formally mapped to 2D model as shown in Figure 2. The designed scheme considers an m layer data collection underwater sensor network, which is composed of sensor nodes, relay nodes, and a sink. All sensor nodes and relay nodes are used randomly at different depths. All the sensor nodes are responsible for sensing and gathering local underwater activities and transmit the gathered data to the surface sink through multiple relays. The function of the relay nodes is to decide the path to the destination and forward the received packets to the next hop node. All the relay sensor nodes rely on a store-and-forward technique; hence, a packet received or generated in a slot can be forwarded from the next slot. In addition, the relay nodes collect data in some cases. The sink node is equipped with both acoustic modem and radio modem and is used in the middle of the water surface. Therefore, the sink node can both communicate with the underwater sensor nodes and the on-shore base stations or satellite. In this paper, we assume the following properties about the UASN. First, every underwater sensor node knows both its location and the location of the surface sink by means of location services as presented in 1 study.[34] Second, it is assumed that each node is static. Third, the acoustic channel is symmetric; therefore, the energy required to transmit a message from node Sn_i to Sn_j is the same as the energy required to transmit a message from node Sn_j to Sn_i for a given Ed or SNR. Fourth, every node can dynamically adjust

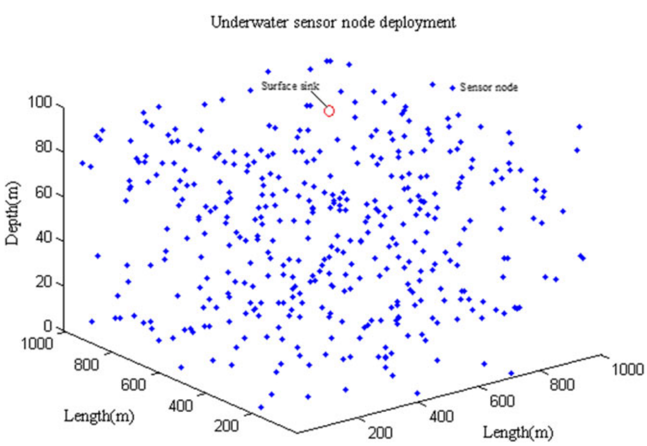

FIGURE 1 A 3D view of sensor node deployment in underwater

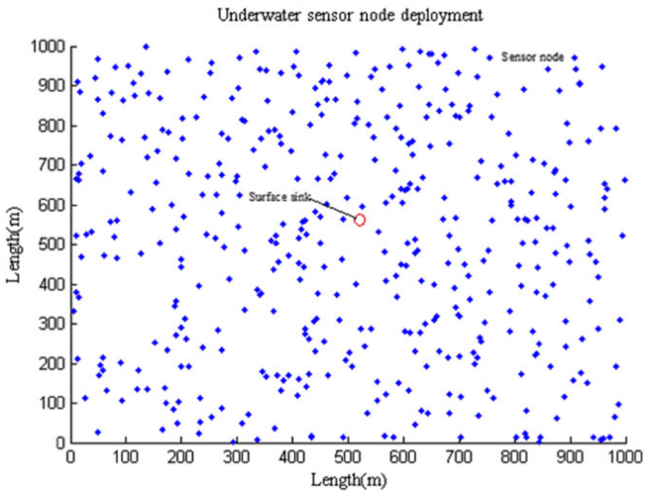

FIGURE 2 A 2D view of sensor node deployment in environment. Underwater environment

its transmission power. Fifth, the sink node is a superior node and can receive a large number of data packets at the same time without any collision. Sixth, the communication links between the sensor nodes are symmetrical. Moreover, to avoid data packet collision among nodes, a carrier sense multiple access (CSMA) mechanism is assumed. In addition, we also assumed that a sensor node can be a part of a pair of cluster heads; however, it will have different time slots assigned by each cluster head using time division multiplex access. Furthermore, it is also assumed that each node has at least 1 existing route to the sea surface sink going through reliable links. To analyse routing phenomenon more realistically, the arrival time follows a Poisson distribution while the service time is considered exponentially distributed.

4.3 | Network initialization

In the beginning, after sensor node deployment in an underwater environment, sea surface buoys is responsible to initiate the clustering process by disseminating multiple copies of a hello message to neighbouring nodes in range. After receiving information from the sink, each node is responsible to restricted broadcast this hello message to its neighbouring nodes by taking into account the carrier sense multiple access mechanism. This hello message consists of information about the sender node identity and residual energy. Then after receiving initiation message successfully, each receiver node replies to the sender node by sending an acknowledgment message, which guarantees that a message has been received successfully. This acknowledgment message contains the information of the receiver node identity, residual energy, received signal strength, and distance information measured during receiving information from the sender node. Herein, we noticed that because of harsh underwater environmental challenges, it is highly possible that a sender message can be lost to its neighbours. Therefore, we used tuple mechanism where each node is responsible to send maximum up to 3 hello messages to its neighbouring nodes if it does not receive acknowledgment message in specific time t_i. Consequently, in our scheme, it may also be possible that a sensor node may receive multiple copies of hello message copies from the same or different sender node neighbours. In this case, after receiving the first hello messages, the rest of the hello messages from the same sender will be dropped. Initially, throughout the sending and receiving process, each node is responsible to maintain the information about neighbouring node identity, residual energy, and received signal strength information. In each round, hello messages are generated to periodically share the updated information among sensor nodes deployed in underwater environment.

4.4 | Evolutionary queue based-spectral clustering algorithm (QSCA)

In proposed nature-inspired queue-based spectral clustering routing scheme, evolutionary spectral clustering algorithm plays an important role by providing a reliable clustering architecture in the entire UWSN. The proposed algorithm effectively reduces high-transmission energy-consumption cost and provides reliable data collection at extremely low-clustering complexity. In designed scheme, after nodes deployment in underwater environment, a surface buoys is responsible to run QSCA for partitioning the set of acoustic sensor nodes into multiple queues based on their similarity values. Then, an initial population of individuals (sensor nodes) is generated (of size $n = 400$), by using a random number generator $R_n \in [0, 1]$, and each individual is evaluated by a predefined fitness function. In current population, σ_0 denotes the dead or inactive individuals with insufficient energy and σ_1 for the active individuals with sufficient amount of residual energy. Each individual based on the fitness values is placed in predefined number of queues with diameter D and length L based on their similarity values in decreasing order. The total queuing system energy-consumption cost can be numerically written as

$$Q_{cost} = cost\sum_{i=1}^{k}\sum_{j=1}^{n}\left(\sum_{QDC}\mathbb{C}_{1i}(D_i)L_i + \sum_{QPC}\mathbb{C}_{2i}(P_i) + \sum_{QOC}\mathbb{C}_{3i}(O_i) + \sum_{QRC}\mathbb{C}_{4i}(R_i)\right)^{Q_j}, \tag{5}$$

in which QDC, QSC, QPC, QOC, and QRC show the queue designing, queue storage, queue processing, queue overhead, and queue repairing cost, respectively. On the other hand, $\mathbb{C}_{1i}$ is the cost for designing a queue having length L_i and diameter D_i, $\mathbb{C}_{2i}(P_i)$ is the processing cost of the ith queue have n number of sensor nodes, $\mathbb{C}_{3i}(O_i)$ is the cost of ith queue overhead, and $\mathbb{C}_{4i}(R_i)$ is the queue repairing cost in terms of a sensor node failure. Thus, the objective function of the developed clustering algorithm is to minimise entire queuing system and energy-consumption cost as

$$Fit(cost)_{Ci} = \min\left(\sum_{i=1}^{k} Q^{i}_{cost} + \mathbb{C}_i T_x + R_x\right) \quad , \tag{6}$$

where Q^{i}_{cost} is the ith queue cost and $\mathbb{C}_i$ is the transmission and reception energy cost during a set of clusters generation in the underwater environment.

Then, a looping phenomenon of selection, crossover, and mutation operators is applied on each individual (sensor node) to improve the quality of solution through the predefined probabilities until the termination criterion is satisfied for the complete clustering solution. By applying selection operator, developed algorithm produces new individuals in the new population. The selection operator is responsible for assigning each set of pair of individuals a positive weight to find a set of partition graph. The value of this assign weight is based on the SNR or Ed, which is proportional to the fitness function (*Fit*) of each individual in region R_i. Thus, the probability of a particular cluster *Ci* from the available set of clusters *Cj*, being selected with the highest weight based on the location information in region R_i, can be formally expressed as

$$\rho_{C(i)} = Fit_{Ci} / \sum_{j=1}^{n} Fit_{Cj} \quad . \tag{7}$$

The probability of each sensor node *Sni* being selected in cluster *Ci* with the highest weight of minimum delay and residual energy can be formally expressed as

$$\rho_{Sni} = Fit_{Sni} / \sum_{j=1}^{n} Fit_{Snj}. \tag{8}$$

To form a mating pool of k parents, selection operator randomly chooses best individuals from the current population set and repeat this process n times until the defined criteria satisfy. To produce a pair of offspring strings, crossover operator is responsible for partial swapping of the bits between 2 parent strings. In other words, in crossover, each gene in the offspring is copied from the same gene from one of the 2 parents with the same probability. For a given crossover probability ($\rho_c = 0.98$), each pair of clusters from the new generated population is randomly picked out by developed algorithm. Then, a random number is generated between 0 and 1. If the generated random number is less than the predefined value, then QSCA is responsible to apply the crossover operator for producing swapping generation, otherwise it does not apply crossover operator on these 2 particular clusters. Through applying the crossover operator, multiple crossover points in a predefined range $[C_{p1}, C_{p2,}\ldots, C_{pn}\ ]$ for $2 \leq n \leq 5$ are also randomly generated in the clustering network. A crossover operator plays a key important role for swapping individual values from 1 cluster to another pairwise cluster in the generated population for creating highly stable dynamic clusters in the entire network. Lastly, by applying the mutation operator, the value of each individual in a cluster changes in the new population. Herein, a mutation operator with the probability of ($\rho_m = 0.01$) is applied to perform the mutation operation by ensuring that no important genetic information is lost. Initially, in new clustering population, which is the result of reproduction and −crossover, the designed algorithm takes into account each individual bit-by-bit by generating a random number between 0 and 1. If the generated number is less than the predefined value, then it applies −mutation operator to the new individual; else, it does not consider mutation operator to that particular individual in a cluster. Herein, the core aim of the designed algorithm is to find the best clustering solution for the given population. If the better individual belongs to a cluster solution cannot be found in a given time over a predefined number of cycles (ie, limit), then that individual is supposed to be abandoned. Figure 3 shows the initial clusters generated by QSCA algorithm for optimization in UASNs. This entire process repeats until the termination criterion of the developed algorithm is satisfied for the best small-size cluster (*best_Ci*) solution in the entire UASNs, which can be numerically expressed as

$$\forall(\mathrm{C})^{1,2,\ldots,n} = \sum_{j=1}^{n} Fit(Cj) < \sum_{i=1}^{k} Fit(best\_{}_{Ci}). \tag{9}$$

The various queue-based network characteristics formally have been expressed from Equations 10 to 22, in Table 1.

4.5 | Evolutionary cluster-head selection algorithm (CHSA)

The key aim to develop evolutionary CHSA is to decrease energy-consumption load in the whole clustering network by appointing a set of cluster heads in the middle of each cluster in UASNs. Hence, less energy is spent during the

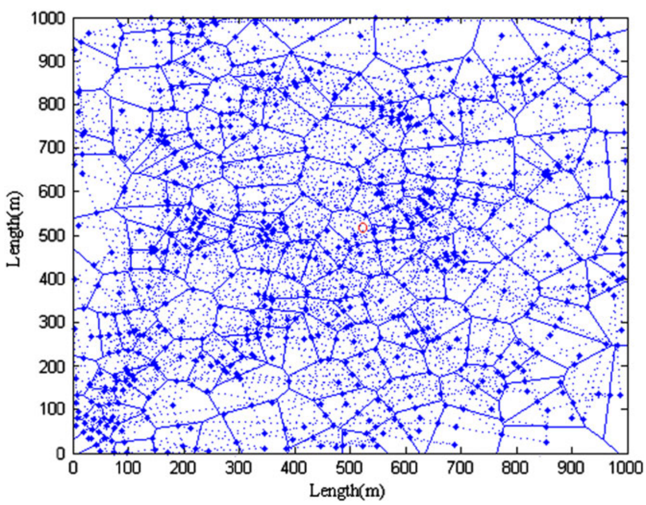

FIGURE 3 Small-size clusters are generated by queue based-spectral clustering algorithm for optimization in underwater acoustic sensor networks

TABLE 1 Formulation used in link quality aware queue-based spectral clustering routing protocol

Definition	Formulation	Explanation	Equation no.
Sum of all queue diameter (D_{sum})	$D_{sum} = \left(D_1, D_2, \ldots. D_n = \sum_{i=1}^{n} C(D_i, L_i)\right)$	*Fit* is the fitness function, D is the diameter of the queue, L is the length of the queue, D_i and L_i are the diameter and queue length of the ith queue, respectively.	(10)
Mean queue length (*MQL*) including nodes being served	$MQL = \sum_{l=1}^{L} \sum_{d=1}^{D} \left[\frac{[A(Sni)]^2}{(P(Sni)-A(Sni))}\right]^{Q_i}$	A is the arrival rate of a sensor node *Sni*, and P is the processing rate of a sensor node *Sni* in a queue Q_i	(11)
Expected number of sensor nodes $E(Sni)$ in each queue (including being served)	$E(Sni) = \sum_{i=1}^{n} \left[\frac{A(Sni)}{P(Sni)-A(Sni)}\right]^{Q_i}$	–	(12)
Queue robustness constraint	$\sum_{i=1}^{n} Sni(Q_{in}) - \sum_{j=1}^{n} Snj(Q_{out}) = \sum Q_p$	Q_{in} represents the number of sensor nodes entered in the queuing system, and Q_{out} is the number of leaving sensor nodes from the queuing system during processing under the constraint that the value of Q_p is always positive.	(13)
Queue looping constraint	$\sum_{i=1}^{n} Sni(Q_j) - \sum_{j=1}^{n} Snj(Q_j) = 0$	Total number of sensor nodes enter in a queue Q_j must be equal to total number of sensor nodes leaving the queue Q_j at the end of the process.	(14)
Average waiting time (w_t) for each sensor node in a queue	$w_t = \sum_{i=1}^{n} \left[\frac{A(Sni)}{P(Sni)(P(Sni)-A(Sni))}\right]^{Q_i}$	$A(Sni)$ and $P(Sni)$ are the arrival and departure rate of the sensor nodes in a queue Q_i, respectively.	(15)
Probability of spending time for each sensor node in a queue	$\rho(w_t \leq T) = \sum_{i=1}^{n} \left[1 - \rho e^{-(1-\rho)P(Sni)t}\right]^{Q_i}$	T is the maximum spending time of a sensor node in a queue.	(16)
Probability of more than defined sensor nodes in each queue	$\rho(Sni > N) = [\rho^{N+k}]_{Qi}$	N is the maximum number of sensor nodes in a queue, and k is the additional sensor nodes added to a queue.	(17)
$Ed_{min}(w)$	$Ed_{min}(w) = \sum_{i=1}^{n} \sum_{j=1}^{k} w(Sni, Snj)^{x \times y} \left\Vert Sni - Snj^2 \right\Vert \leq Ed_{maxi}$	$Ed_{min}(w)$ is the minimum Euclidean distance in terms of weight w between sensor node *Sni* and *Snj*, and x and y are the coordinates of a sensor node *Sni* and *Snj*. While Ed_{maxi} is the maximum distance threshold value.	(18)
Euclidean distance error (*Ede*)	$Ede = \sum_{k=1}^{n} \frac{\left\Vert Sni - Snj^2 \right\Vert}{Snk} \leq Ede_{maxi}$	*Snk* is the number of neighbouring sensor nodes associated to a sensor node *Sni*, and Ede_{maxi} is maximum acceptable value.	(19)
Signal to noise ratio(*SNR*) between a pair of nodes	$SNR(Sni, \mathrm{Snj}) = \frac{P_{Signal}}{P_{noise}} \geq SNR(Sni, \mathrm{Snj})_{min}$	P_{Signal} and P_{noise} are the power of the received signal and power of the noise, respectively. $SNR(Sni, \mathrm{Snj})_{min}$ is the minimum acceptable value of SNR.	(20)
Delay (*De*)	$De_{min}(w) = \sum_{i=1}^{n} \sum_{j=1}^{k} w(Sni, Snj) \Vert d_t(ni - Snj) \Vert^2$	The minimum delay $De_{min}(w)$ is the time difference d_t in terms of weight w between a pair of sensor nodes *Sni* and *Snj*. This difference is based on the packet time of arrival (TOA) and time of departure (TOD).	(21)
Residual energy (*Re*)	$Re_{max}(w) = \sum_{i=1}^{n} \sum_{j=1}^{k} w(Sni, Snj) \Vert E_i \Vert^2$	$Re_{max}(w)$ is the maximum residual energy of a sensor node in terms of weight. E_i is the energy consumed of a sensor node *Sni* after sending a request or data packets to its neighbouring node *Snj*.	(22)
Intradistance ($I_{dis.}$)	$I_{dis.}(w) = \sum_{i=1}^{n} \sum_{j=1}^{k} (Sni, Cj)^{x \times y} \, \vert\vert Sni - Cj \vert\vert^2$	$Ed_{min}(w)$ is the minimum Euclidean distance in terms of weight w between sensor node *Sni* and cluster center *Cj*. x and y are the coordinates of a sensor node *Sni* to *Cj*.	(23)
Interdistance between clusters (*IDC*)	$IDC = k\left(\sum_{i=1}^{n} \sum_{j=1}^{n} \frac{\vert Snij \vert^2}{\vert Ci_t \vert \cdot \vert Cj_{t-1} \vert} - 1\right)$	$\vert Snij \vert$ is the number of sensor nodes that are both in *Ci* and *Cj* at time t and $t - 1$, respectively. k is a scaling factor set to 1.	(24)

intracluster, and intracluster data processing computed using Equations 23 and 24. This way, more energy can be preserved for the inter-cluster relay traffic. The core aim of designed evolutionary CHSA is to provide an effective way for reliable data collection and extending the lifetime of a UASN. Therefore, the designed algorithm considers a double check mechanism of Ed and SNR for providing reliable intracluster data collection in the entire underwater clustering network. Moreover, it selects cluster heads with more residual energy and rotates each cluster head periodically to distribute the energy consumption among nodes in each cluster and prolong the network lifetime. In addition, in the existing schemes (see Section 2), the cluster heads nearer to the sink are burdened with heavy relay traffic and deplete their energy much faster than the other cluster heads. Therefore, those areas of the network become uncovered, and the network becomes partitioned. To cope this hot spot problem, we appointed a set of 2-hop cluster heads with lower distance to the sea surface sink iteratively by the designed algorithm. After dividing the entire network into multiple small-size clusters, in the second phase of our proposed scheme, an initial population of individuals is generated using a random number generator $R_n \in [0, 1]$. Each individual in the current population is evaluated by a predefined fitness function, which consists of higher residual energy; SNR lower intradistance and distance to the sink (d_{sink}) can be numerically indicated as

$$Fit(Sni)_{CHi} = Fit_{Sni} / \sum_{j=1}^{n} Fit_{Snj} \left(Re^{Snj} + I_{dis.}^{Snj} + d_{sink}^{Snj} \right). \quad (25)$$

The selection operator selects a pair of best individuals in each cluster by measuring the fitness value by taking into account the abundant individuals indicated as σ_0. Then a looping phenomenon of multipoints crossover and mutation operators are applied to each individual through the predefined probabilities of $P_c = 0.95$ and $P_m = 0.03$. In each round of the proposed routing protocol, the cluster-head formation phase generates an initial population of solutions, the fitness of which is then evaluated, and the parents are selected to generate a new population via recombination and mutation operators based on the fitness values. During the association phase, a best individual (CH) is selected in each cluster based on the fitness value. Figure 4 shows the initial cluster heads generated by CHSA algorithm for optimization in UASNs. Moreover, herein, by sharing mutual information between each pair of clusters, the quality of each cluster is measured using following Equation.

$$MI = \frac{\sum_i^k \sum_j^k Sn(i,j) \log\left(\frac{Sn(i,j)n}{Sn_{Ca}^i Sn_{Cb}^j}\right)}{\sqrt{\left(\sum_{i=1}^k Sn_{Ca}^i \log \frac{Sn_{Ca}^i}{n}\right)\left(\sum_{j=1}^k Sn_{Cb}^j \log \frac{Sn_{Cb}^j}{n}\right)}}, \quad (26)$$

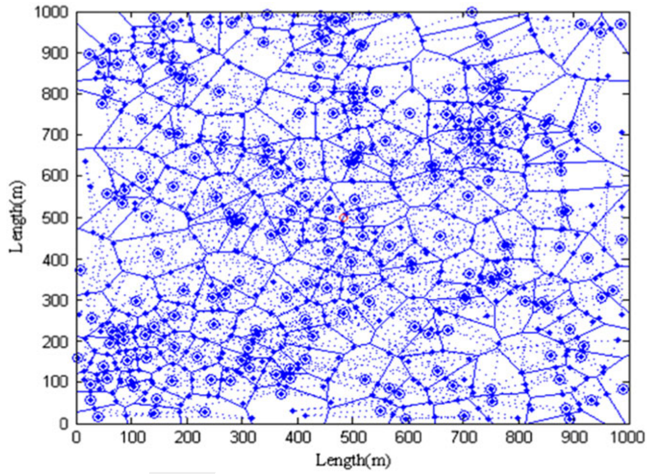

FIGURE 4 Clustering heads generated by cluster head selection algorithm for optimization in underwater acoustic sensor networks

in which MI is the mutual information used for comparing a pair of clusters. Sn_a^i is the number of sensor nodes in cluster Ca, Sn_b^j is the number of sensor nodes in cluster Cb, and n is a constant set to 1. This process repeats until the termination condition of the EA is satisfied for the complete cluster-head solution in each cluster of the clustering network and can be formally shown as

$$Fit(CHi)_{Ck} = \sum_{k=1}^{n} \left(best_{FitSni} / \sum_{j=1}^{m} Fit_{Snj} \right)^{Ck,}. \quad (27)$$

4.6 | Evolutionary multihop minimum spanning tree routing algorithm (MTRA)

One of the main objectives of proposed routing algorithm is to route data packets over loop free-spanning tree from the source towards the sink in the network. In the proposed routing algorithm, observed information is routed over a set of cluster heads by considering a double check mechanism of high SNR or minimum distance and over-assignment cost (OC) to reduce high-transmission energy-consumption and packet loss. The minimum distance for a next hop-relay node is computed as

$$RCN(CHi) = \sum_{j=1}^{k} d(CH_i, CH_j) + d(CH_j, S_{ink}), \quad (28)$$

in which $RCN(CHi)$ represents the relay cluster-head node CHi and d is the distance between a pair of cluster heads CH_i to CH_k and then to sea surface sink, such that $k \le n$, where n is the maximum number of relay nodes in a route. Generally, during route discovery process, when a source cluster-head node needs to initiate a communication, it sends route discovery packets (RDP) to its neighbouring cluster-head nodes in range. When the destination node receives the first RDP packet it calculates its residual energy, distance to sender cluster head, distance to sink cost, and its over-assignment cost of participating in the route and sends it to the sender using reply route request (RRDP) packet. The receiver node calculates the cost of each reply RRDP massage from CHs and selects the best next hop-relay node with low cost to route data packets. Thus, in the routing, each cluster head (gateway) is supposed to send its aggregated data to the sink directly if the sink within its range, otherwise it sends the data through appropriate next hop cluster-head relay nodes. Moreover, routing information of a cluster is shared with only other cluster heads. Therefore, the number of transmissions carried out for distributing routing information reduces. In evolutionary implementation, after dividing the entire sensor nodes into multiple clusters, an initial population of individuals (cluster heads) is generated, using a random number generator, and each individual is evaluated by a predefined fitness function. A set of individuals in the initial population with greater fitness value is selected as the current best solution, ie, the best next hop CHs leading to the lowest energy and over-assignment cost for robust data delivery over reliable links towards sink. These objectives enable routes with intermediate cluster-head nodes with higher battery reserve to be selected. This helps to balance the packet-forwarding task in the network so that a few sensor nodes do not run out of their battery power earlier compared to other CHs. Moreover, buffer length as a metric helps to select routes that are not congested, thereby decrease congestion loss and end-to-end delay resulting from queuing delay. Finally, link stability metric enables the route selection process to select relatively long-lived and more stable routes and guarantees a lower packet loss rate as a result of the movement of intermediate CHs and fewer route failures. Thus, the fitness function for an appropriate next hop cluster head i as a relay node formally can be described as

$$Fit(CHi) = Fit_{CHi} / \sum_{j=1}^{n} Fit_{CHj}\left(Re^{CHj} + OC^{CHj} + d_{sink}^{CHj}\right), \quad (29)$$

in which Re^{CHj}, OC^{CHj}, and d_{sink}^{CHj} are the residual energy, the over-assignment cost, and the distance to sink of a next hop cluster-head node CHj, respectively. In each round of the proposed routing protocol, to improve the quality of individuals (an appropriate next hop-relay CH node with low-energy consumption and over-assignment cost), we applied a loop of multipoint crossover and mutation operators to each parent individual through the predefined probabilities as $\rho_c = 0.99$ and $\rho_m = 0.002$. This entire process repeats iteratively until the termination criterion is satisfied for the complete cluster-based routing solution over a set of shortest path cluster heads in a greedy manner from the source towards sink. Let us consider a fully connected undirected network graph; selection operator assigns each cluster head in the population a set of weight $g(v, e)$, where a finite set of vertices $v = \{1, 2, \ldots, n\}$ and edge $e = \{(i, j) | i, j \in \mathrm{v}\}$ are connected between each cluster-head node in the system. The each edge in fully connected graph assigned a positive real number value denoting distance among cluster heads and can be represented as $w = \{w_1, w_2, \ldots, w_{(N-1)N}\}$, where decision variable can be expressed as

$$CH_{(ij)} = \begin{cases} 1 & \text{if edge}(i,j) \text{ is slelected} \\ 0 & \text{otherwise} \end{cases} \quad \text{for} \quad \forall \text{edges } (i,j) \in e. \quad (30)$$

Let vector v of graph g represents the spanning tree and ω is the set of all such vectors in graph g corresponds to spanning tree, the minimum spanning tree (MST) problem presented in 1 study,[35] in improved form and can be expressed as

$$\mathrm{MST} = \min\left\{\sum_{i=1}^{n-1}\sum_{j=i+1}^{n} w_{CH(ij)}\left(v_{ij}\right) | v \in \omega\right\}. \quad (31)$$

To solve the problem of the edge constraints on each vertex v, it is assumed that the edge of each vertex is not always arbitrary in the spanning tree such that at each vertex v_i the corresponding edge value e_k is at most a given value ∂_k which is based on distance and residual energy of a pair of cluster head nodes i and j in terms of weight $w_{CH(ij)}$. Therefore, the problem can be expressed as

$$\mathrm{MST} = \min\left\{\sum_{i=1}^{n-1}\sum_{j=i+1}^{n} w_{CH(ij)}\left(v_{ij}\right) | e_k \leq \partial_k; \; \mathrm{i}, j, k \in v; \; \mathrm{v} \in \omega\right\} \quad (32)$$

consequently, if there are N total numbers of clusters in the minimum spanning tree network and there are n numbers of sensor nodes in each cluster where a cluster head can be represented as CH_i in each individual cluster. Then the average shortest path (ASP) for N number of cluster heads is the length of a shortest path (ASP) from a cluster-head node i to sensor node k towards the sink is calculated as

$$ASP = \frac{1}{N(N-1)} \sum_{k=1}^{n} [CH_1 +, \ldots, CH_k + S_{ink}], \quad (33)$$

in which N is the number of cluster heads over a shortest path from a cluster head node CH_1 to cluster head node CH_n. In MTRA, for measuring CH_i over-assignment cost, the entire network is divided into internal and external traffic by considering M/M/1 and M/M/N queuing models as discussed in 1 study,[36] respectively. The sensor nodes in the same cluster generate internal traffic while the external data traffic (EDT) cost is from the other cluster heads is calculated using Equation 34. While the Jobs blocking probability (JBP) of a cluster head CH_i when queue is full is computed using Equation 35.

$$EDT\ (CHi) = \sum_{k=1}^{n}\sum_{j=1}^{n} Ck\cdot\lambda_j(1-P_b), \tag{34}$$

$$JBP(CHi) = \sum_{k=1}^{n}\sum_{j=1}^{n} L_k(Ck)\cdot \mathrm{C}_j\lambda_j(1-P_b), \tag{35}$$

in which λ_j is the packets arrival rate from a set of neighbouring clusters Ck with probability P_b, and L_k is the queue length of a cluster-head node for storing data packets. To represent the connection between a cluster head to its neighbouring cluster head, we simply draw a $n\times$n topological matrix as l_1 and $n\times$m topological matrix as l_2. To show the connectivity among a set of neighbouring clusters and also to cluster member sensor nodes (Sn_i), we use $l_1 i\rightarrow j$ and $l_2 i\rightarrow j$ in the entire network where the spanning tree matrix $[l_1 l_2]$ define as $n\times(n+l)$. Then, we can formulate a topological problem as

$$l_{1ij} = \begin{cases} 1 & \text{if } Sn_i \in CH_j \\ 0 & \text{otherwise} \end{cases} \text{ for } \quad i,j = 1,2,\ldots,n, \tag{36}$$

$$l_{2ij} = \begin{cases} 1 & \text{if } CH_i \in CH_J \\ 0 & \text{otherwise} \end{cases} \text{ for } \quad i,j = 1,2,\ldots,n, \tag{37}$$

$$\min \frac{1}{\delta_t}\left[\sum_{i=1}^{n}\frac{T_{CHi}}{DTC_{CHi}-T_{CHi}} + \sum_{i=1}^{n}\sum_{j=1}^{n} d_{CH(ij)}gT_{CH(i,j)}\right], \tag{38}$$

$$\min\left[\sum_{i=1}^{n-1}\sum_{j=i+1}^{n} w_{lij}\cdot g\cdot l_{1ij} + \sum_{i=1}^{n}\sum_{j=1}^{m} w_{2ij}\cdot g\cdot l_{2ij}\right], \tag{39}$$

$$\sum_{j=1}^{k} l_{1ij} < Max_i, \quad \mathrm{i} = 1,2,\ldots,n, \tag{40}$$

$$\sum_{i=1}^{l} l_{2ij} < Max_i, \quad \mathrm{j} = 1,2,\ldots,n, \tag{41}$$

where δ_i denotes the entire offered data traffic, T_{CHi} shows the maximum data traffic at cluster head i, $T_{c(ij)}$ denotes total data traffic through the link (k, l), DTC_{CHi} represents the data traffic capacity of cluster head i, $d_{CH(ij)}$ indicates the between cluster head node i and cluster head node j, g is a constant set to 0.1, and w_{1ij} and w_{2ij} represent the weight of the link between cluster head j and sensor node i and cluster head node i and cluster head j, respectively. Overall genetic representation of the proposed scheme is shown in Figure 5.

5 | COMMUNICATION CHANNEL MODEL

The communication channel model derived in 1 study[37] for underwater acoustic channels is used in this study. The path loss, $PL(d,f)$, of acoustic signal with transmission range d in meters and frequency f in KHz is given in dB as

$$PL(d,f) = \mathrm{k}\log(d) + \alpha(f)\ \mathrm{d}\times 10^{-3}, \tag{42}$$

where $\alpha(f)$ is the absorption coefficient and k is the spreading factor caused by energy spreading. Herein, we assume orthogonal frequency division multiplexing encoding technique and use the quadrature amplitude modulation scheme to calculate the amplitude and phase of the sub-carrier. In our model 16–quadrature amplitude modulation with orthogonal frequency division multiplexing transmission is considered. Bit error rate is computed as

$$P_b^{16QAM} = \frac{3}{2k}\ \mathrm{erfc}\left(\sqrt{\frac{k}{10}\frac{E_b}{N_0}}\right), \tag{43}$$

SNR is calculated as:

$$SNR = 10^{SNR(d,f)/10} \tag{44}$$

5.1 | Simulation model and metrics

In this study, experiments were conducted using a network simulator based on MATLAB[38] to evaluate the performance of LRP against the existing underwater protocols such as R-ERP^{2}R,[22] EEDBR,[15] and DBR[14] using 4 widely used metrics: (1) PDR, (2) delay, (3) throughput, and (4) energy consumption. Packet delivery ratio is regarded as the ratio between the number of packets successfully received by the sink and those sent by the source nodes.[39] Delay is the time taken by a packet to reach any destination node from a source node. Throughput is the number of data packets processed in a given time (bits per second) defined in another study.[40] Residual energy is the amount of remaining energy after energy consumed during successful transmission of data packets by the nodes. A 3D sensor network of length × width × height with values 1000 × 1000 × 100 is created, and the nodes are deployed randomly at different depths. The total number of deployed underwater sensor nodes ranges from 50 to 400 nodes. One sink is floating on the surface of the sea. The carrier frequency is 40 kHz for bandwidth efficiency of 20 bits per second. All sensor nodes are identical with the same capabilities, and each of them is assumed to be equipped with a microphone with a sensing range of 120 m, and acoustic communication capabilities with communication range far-field distance is set to 80 m with an initial energy value of 5 J. A total of 5 sensor nodes positioned at the bottom of the deployment region were selected as the source nodes. In the network, each source node generated a packet every 10 seconds. The size of the data packets was fixed at 32 bytes. As soon as a packet is generated, it is associated with a source node randomly selected among all the nodes except for the sink. The destination of all packets is the sink in each simulation round. The reception energy waste is 0.24 W, and the transmission energy consists of 2 levels for the selected value of the carrier frequency and the bandwidth: 1.5 W to keep the transmission electronics turned on at their highest level and 1 W for low-level transmission with respect to hop distance.

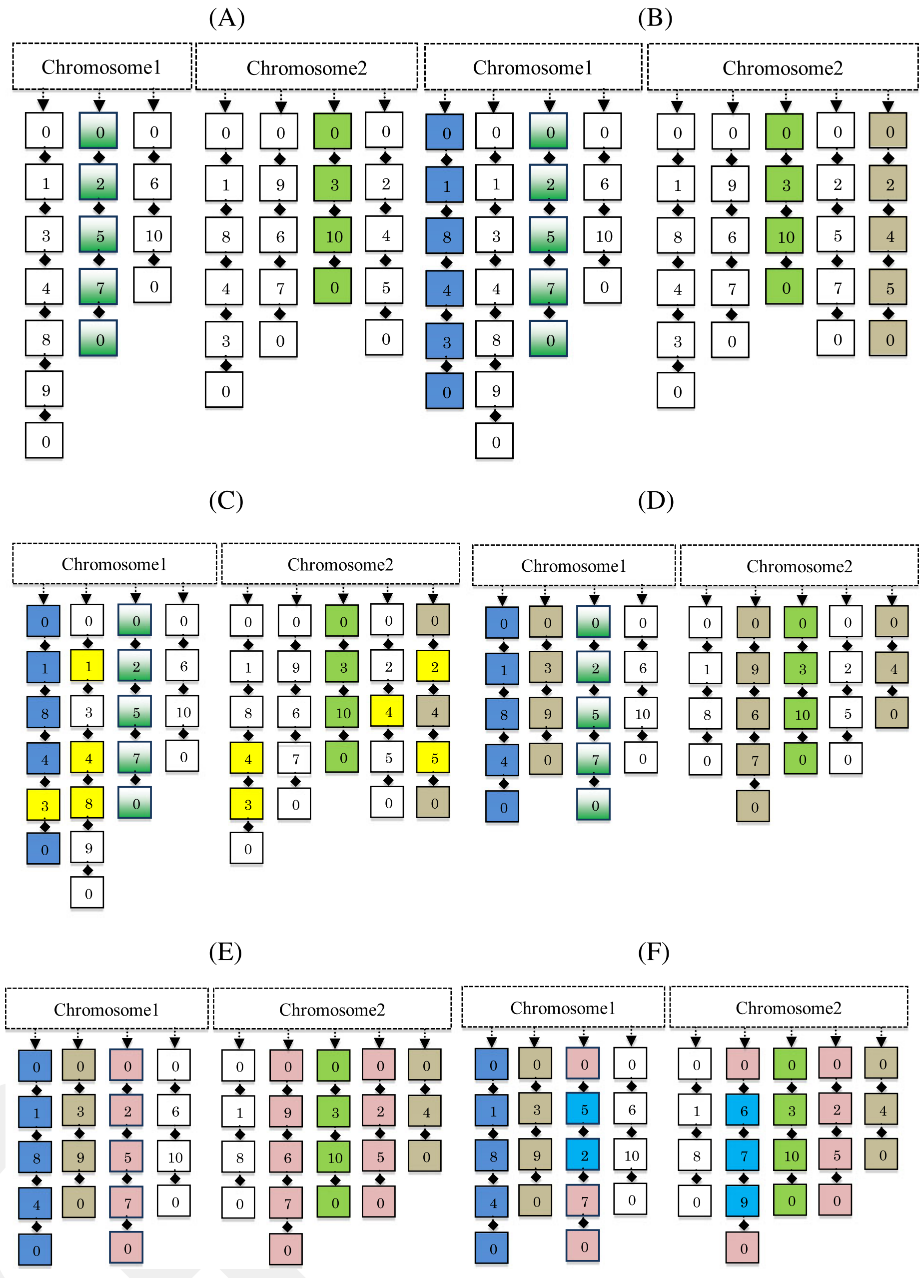

FIGURE 5 Shows the genetic representation of our proposed LRP scheme. In the scheme, 0 to 10 indicates the different number of acoustic sensor nodes involved in genomic process. A, Shows the selection of chromosomes in the current population. B, Indicates the selection of chromosomes when new cromosomes information is added in the current population. C, Represents the removal of duplications cromosomes after multipoint crossover. D, Demonstrates the random shuffling of cromosomes using a random number generator. E, Illustrates the selection of the chromosome for mutation process. F, Shows the best individuals (sensor nodes/cluster heads) after multipoint mutation to find complete clustering and routing solution in the current population

Moreover, the idle listening power and sleep power were set to 0.064 and 3×10^{-16} W, respectively. The values for the acoustic propagation model factors, such as temperature in °C, salinity, acidity, and cylindrical spreading coefficient factor were set to 35, 8, 22, and 10, respectively. Fifty simulation rounds are conducted to obtain statistic average performance of the simulated protocols for each topology. Each of the simulation run lasted until one of the nodes ran out of energy. The hello packet interval was set to 13 seconds. After every 13 seconds, the SNR, distance, and link values are computed, and the residual energy information is shared among the neighbour nodes for updating routing tables.

5.2 | Performance results and discussions

Figure 6 shows the measurement of PDR for 4 protocols with varying number of rounds between 1 and 2500. Here, it can be seen that the PDR in all routing schemes increases with the increase in number of rounds between 1 and 2500. In the beginning, the performance of EEDBR is better than DBR and R-ERP^{2}R routing schemes in terms of high PDR. This is because of exploiting its aptitude to carry information by following alternative data paths in case of a route node failure. However, when the network size becomes larger between round numbers 1000 and 2500, the PDR of

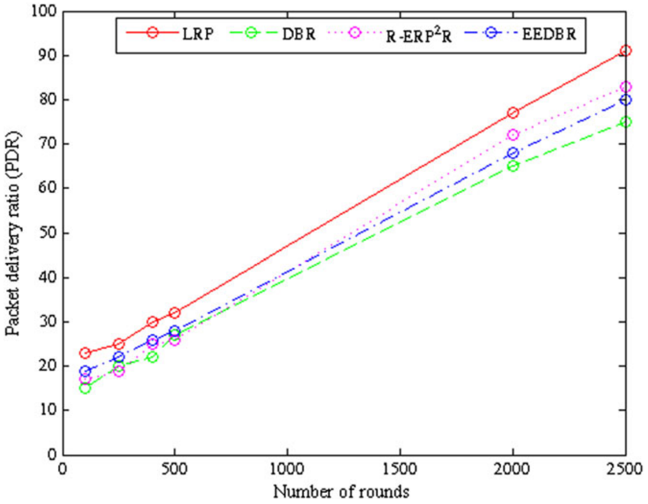

FIGURE 6 Packet delivery ratio vs number of rounds between 1 and 2500. DBR indicates depth-based routing protocol called; EEDBR, energy-efficient depth-based routing protocol; LPR, link quality aware queue-based spectral clustering routing protocol; R-ERP^{2}R, reliable and energy-efficient routing protocol

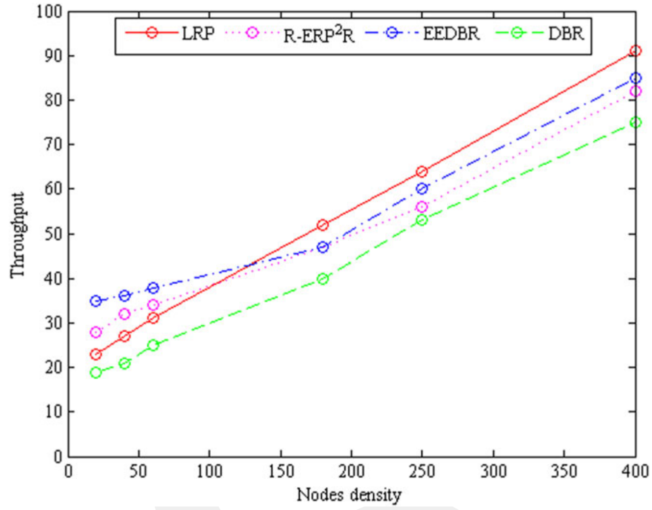

FIGURE 8 Throughput vs nodes density. DBR indicates depth-based routing protocol called; EEDBR, energy-efficient depth-based routing protocol; LPR, link quality aware queue-based spectral clustering routing protocol; R-ERP^{2}R, reliable and energy-efficient routing protocol

EEDBR becomes poor than LRP and R-ERP^{2}R routing schemes. This is because of its forwarding data packets over a set of nodes having low-residual energy and node memory overrun issues, which is found more in DBR routing scheme. In the meanwhile, the performance of R-ERP^{2}R to achieve network delay is observed slightly better than both EEDBR and DBR routing schemes as shown in Figure 7. Herein, we observed that this better performance of R-ERP^{2}R is with the cost of low network throughput compared to EEDBR and LRP routing schemes as shown in Figure 8. The residual energy profile of each routing scheme is shown in Figure 9. Herein, it is clearly shown that the upkeeping residual energy profile of R-ERP^{2}R is found better than both EEDBR and DBR routing schemes. However, observed poor than LRP routing scheme. This low upkeeping residual energy performance in R-ERP^{2}R is due to its high-energy consumption during sending a notable amount of request messages to find a reliable data path towards the sink in the network. Herein, it is witnessed that in terms of high PDR, throughput and upkeeping residual energy, the graphs of our proposed scheme LRP is usually higher than all other routing schemes. During performing experimental studies, we witnessed a number of possible reasons that may possible deteriorates the network performance of R-ERP^{2}R, EEDBR, and DBR routing scheme in the network: First, DBR and EEDBR routing schemes do not have effective and efficient routing table management mechanism in terms of updating and route finding. Second,

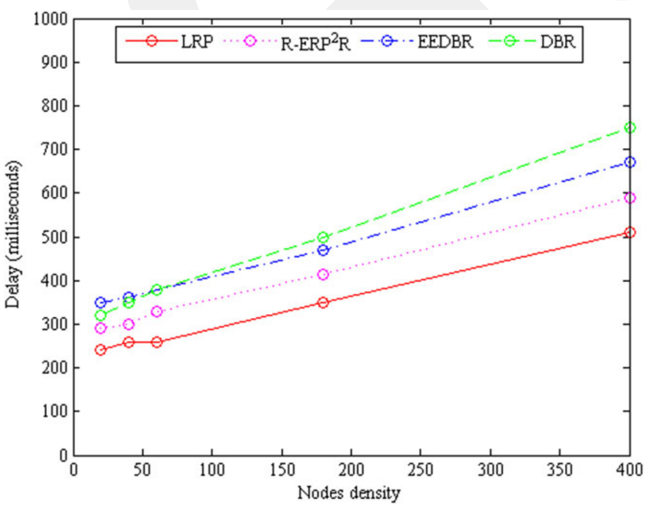

FIGURE 7 Delay vs nodes density between 1 and 400. DBR indicates depth-based routing protocol called; EEDBR, energy-efficient depth-based routing protocol; LPR, link quality aware queue-based spectral clustering routing protocol; R-ERP^{2}R, reliable and energy-efficient routing protocol

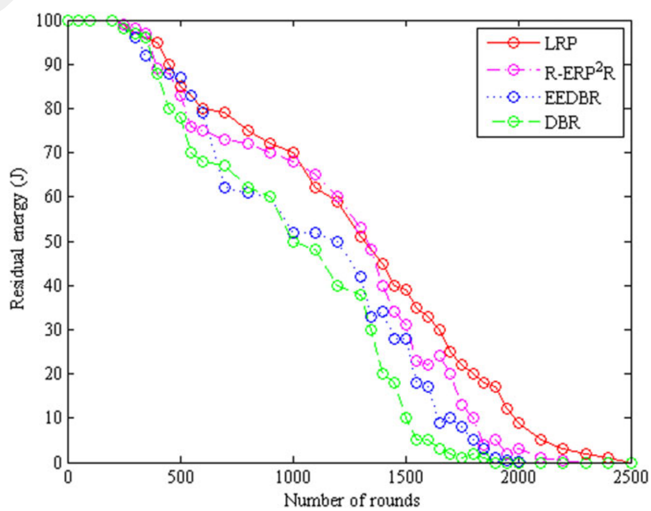

FIGURE 9 Residual energy vs number of rounds between 1 and 2500. DBR indicates depth-based routing protocol called; EEDBR, energy-efficient depth-based routing protocol; LPR, link quality aware queue-based spectral clustering routing protocol; R-ERP^{2}R, reliable and energy-efficient routing protocol

one of the main reasons, which causes highly inconsistent performance in both DBR and EEDBR routing schemes, is their huge amount of generating redundant data packets in the network. Thus, each hop near to the base station has to face congestion management issues, which lead to high data packet collision in the network compared to R-ERP^{2}R and LRP routing schemes. Third, the other main reason, which rapidly increases the packet loss ratio and network delay in DBR, EEDBR, and R-ERP^{2}R routing schemes, is because of their not considering the node over-assignment costs during forwarding information. This may cause node memory runoff problems result in huge amounts of packets lost. Fourth, during experimental studies, it is detected that the stable link reliability between each set of pair of nodes is an important, challenging issue in R-ERP^{2}R routing scheme leading to a certain amount of data packets drop in the network. Fifth, excessive multihop data packet transmission is another main challenging issues in DBR and EEDBR compared to R-ERP^{2}R routing scheme, which causes excessive data packet delay and individual node energy in the network. Thus, seeing the local data traffic at each node, sensed information has to wait for some time at each hop to get attention, as well as the time required to forward the data to next hop or sink, which makes routing schemes prone to longer delay. Moreover, both DBR and EEDBR routing schemes are suffering from an intelligent suitable mechanism to guarantee link reliability between each set of node pairs in harsh nature of the underwater environment. As the distance among sensor nodes increase, the link quality becomes poorer, which result in excessive message exchanges to cope link stability problem compared to R-ERP^{2}R routing scheme. This excessive message sharing consumes a significant amount of energy in DBR, EEDBR, and R-ERP^{2}R routing schemes leading to poor upkeeping nodes residual energy in the network. In addition, high interference among sensor nodes is another challenging issue in R-ERP^{2}R routing scheme, which generates a significant amount of corrupted data packets in the network. These corrupted packets are routed at each hope consume a notable amount of individual node energy. To achieve better delay level performance in LRP, we find that the behavior of QSCA, CHSA, and MTRA mechanisms play important roles to increase the nodes upkeeping residual energy by minimizing energy consumption and network delay to provide a robust and link reliability aware multihop greedy routing architecture for UASNs. To provide quality aware links among nodes, the designed scheme considers a double check mechanism of SNR and Ed information and adopts the best among the available one, opportunistically. In the first phase of LRP, QSCA starts to generate small-size highly reliable dynamic clusters in the entire network. Herein, the core aim of organizing sensor nodes in multiple short hierarchical centralised queues is to provide better clustering performance mainly because of its load-sharing characteristic. Moreover, the use of a round robin assignment strategy, in which the arriving sensor nodes are assigned to ready queues in a cyclic fashion, improves the clustering performance significantly compared to the traditional clustering schemes. In addition, it allows mutually exclusive access to all sensor nodes in queues to avoid bottleneck by reducing the number of processors in the clustering network. Typically, the numbers of generated queues are less than the number of sensor nodes in the network. In the second phase, CHSA appoints a set of cluster-head nodes from the normal nodes in the center of each cluster based on high residual energy and minimum intradistance to distribute energy load evenly in the clusters. This substantially reduces the battery drainage of each sensor node, which only needs to communicate with their respective CHs over relatively short distances. Moreover, a set of appointed cluster heads are periodically rotated in each round of the designed scheme. The main aim of rotation is to balance the intracluster and intercluster energy consumption among cluster heads and member nodes. This mechanism avoids hot spots and network partitioning problems caused due to unbalanced energy consumption at different CHs. Herein, the designed clustering algorithm aims to generate small-size clusters in such a way that any node in any cluster is single hop away from the cluster head to achieve better load balancing. Consequently, the proposed algorithm has a time complexity of $O(\mu \cdot n)$, where μ is the number of initializations and n is the total number of nodes. The computational time complexity of the clustering algorithm is linear to μ. In the third phase, MTRA divides the entire network into multiple loop-free virtual LANs like architecture to route information in the form of minimum spanning tree from the source towards the destination. The designed MTRA routing architecture is responsible to evenly distribute energy load in the network by appointing a set of CHs as backbone nodes by considering their residual energy, minimum distance, and node over-assignment cost. Thus, it takes the advantage of shortest path routing near optimal number of hops by considering minimum distance values in the network, which prevents unnecessary multihop and data path looping in the network as well as maintain node residual energy when deployed for UASNs applications. For shortest path routing, the designed routing algorithm has a time complexity of $O(\mu \cdot CH_n)$ compared to existing routing algorithms, which are suitable only for networks with a small number of sensor nodes, where n is the number of cluster heads. This makes the new routing algorithm suitable for networks of small and large number of nodes. In case of a backbone route node failure, routing table helps to find an appropriate next hop relay node with low network delay, energy consumption, and node over assignment cost to convey information towards the sink with low network delay. Herein, it is important to note that to find an appropriate next hop node, first, sender will look into its low-transmission single-hop routing table (it stored maximum up to 8 hops neighbouring nodes in range with transmission power $(10 \le d_{th} \le 30)$ meters); if fail to find, then it route information using high-transmission single-hop routing table information (it stored maximum up to 5 hops neighbouring nodes in range with transmission power ($10 \geq d_{th} \geq 30$) meters) in the network. Thus, in sum, each CH is responsible to maintain dual information in its routing table called low-level and high-level single-hop routing information differentiated by a unique pointer. This mechanism reduces the impact of management complexity at extremely low cost; as a result, high throughput is reported for UASNs applications in harsh nature underwater. This routing information of a new relay node is then shared with only other cluster heads or cluster gateways, which reduces the number of transmissions realized for distributing routing information. Also, it minimises the information processed by sensor nodes and data stored in sensor nodes. Furthermore, as stated before, since this process extremely reduces the impact of data path looping and unnecessary multihop data transmission for UASNs applications. Therefore, the impact of control message overheads is reduced indirectly, and a significant amount of network energy is saved in LRP compared to all the other routing schemes in UWSNs.

6 | CONCLUSION AND FUTURE WORK

Underwater acoustic sensor networks are getting growing interest because of their wide range of applications. These applications include pollution monitoring, offshore exploration, tsunami warnings, and tactical surveillance. On the other hand, the unique characteristics of the underwater acoustic channel, such as limited bandwidth capacity, high error rate, and variable delays, impose many challenges that limit the utilization of underwater sensor networks for these applications. Therefore, designing of an efficient routing protocol that provides reliable data delivery over longer periods in a timely and efficient manner for UASNs is very challenging. To address these problems and challenges, in this paper we proposed a novel nature-inspired evolutionary link quality-aware queue-based spectral clustering routing protocol for UASN-based underwater applications.

In the proposed approach, to decrease the total transmission power aggregated over the nodes in the selected path and balance the load between the nodes for extending the network lifetime, we used a highly stable clustering architecture. The clustering architecture designed in this study reduces unnecessary retransmissions, delay, and the residual energy to provide energy balancing in UASNs. On the basis of a distributed architecture, the proposed approach successfully distributes network data traffic load evenly in harsh underwater environment. The results of our extensive simulations confirm that compared to the existing approaches, the proposed approach expends significantly less energy and decreases communication delay for successful packet delivery and improves throughput. However, its shortest queue strategy used in the clustering architecture creates complexity overhead, which can be addressed by mechanisms that dynamically select appropriate queue size. Future work of this study consists of designing mechanisms to address the impact of node density on reducing the probability of collisions in queues in the MAC layer so that the performance of the proposed approach can be improved in a cross-layer fashion.

ACKNOWLEDGEMENT

This work was supported by the Turkish Scientific and Technical Research Council under grant 114E248.

REFERENCES

1. Faheem M, Abbas MZ, Tuna G, Gungor VC. EDHRP: Energy efficient event driven hybrid routing protocol for densely deployed wireless sensor networks. *J Netw Comput Appl*. 2015;58(12):309–326.
2. Gungor VC, Sahin D, Kocak T, et al. A survey on smart grid potential applications and communication requirements. *IEEE Trans Ind Inf*. 2013;9:28–42.
3. Gungor VC, Korkmaz MK. Wireless link-quality estimation in smart grid environments. *Int J Distrib Sens Netw*. 2012;2012:1–10.
4. Liu B, Ren F, Shen J, Chen H. Advanced self-correcting time synchronization in wireless sensor networks. *IEEE Commun Lett*. 2010;14:309–311.
5. Zhang W, Yin Q, Chen H, Gao F, Ansari N. Distributed angle estimation for localization in wireless sensor networks. *IEEE Trans Wirel Commun*. 2013;12:527–537.
6. Cheng L, Cao J, Chen C, Chen H, Ma J. Towards intelligent contention-based geographic forwarding in wireless sensor networks. *IET Commun*. 2011;5:1711–1719.
7. Wang C, Chen H, Yin Q, Feng A, Molisch AF. Multi-user two-way relay networks with distributed beamforming. *IEEE Trans Wirel Commun*. 2011;10:3460–3471.
8. Huang P, Chen H, Xing G, Tan Y. SGF: A state-free gradient-based forwarding protocol for wireless sensor networks. *ACM Trans Sensor Netw (TOSN)*. 2009;5:14
9. Ahmed M, Salleh M, Channa MI. Routing protocols for underwater wireless sensor networks based on data forwarding: A review. *Telecommun Syst*. 2016;1–15.
10. Ayaz M, Baig I, Abdullah A, Faye I. A survey on routing techniques in underwater wireless sensor networks. *J Netw Comput Appl*. 2011;34:1908–1927.
11. Stojanovic M, Freitag L. Recent trends in underwater acoustic communications. *Mar Technol Soc J*. 2013;47:45–50.
12. Zahedi Y, Ngah R, Abdulrahman A, et al. Experimental measurement and analysis of electromagnetic communication in underwater wireless sensor networks. *J Comput Theor Nanosci*. 2015;12:6069–6076.
13. Rao J, Yao W, Wen L. Using the combination refraction-reflection solid to design omni-directional light source used in underwater wireless optical communication. *Appl Optics Photonics China (AOPC2015)*. 2015;96790F–96790F-9.
14. Yan H, Shi ZJ, Cui J-H. DBR: depth-based routing for underwater sensor networks. In: Das A, Pung HK, Bu Sung Lee F, Wai Choong Wong L, eds. *NETWORKING 2008 Ad Hoc and Sensor Networks, Wireless Networks, Next Generation Internet*. Berlin, Germany: Springer; 2008, 72–86.
15. Wahid A, Kim D. An energy efficient localization-free routing protocol for underwater wireless sensor networks. *Int J Distrib Sens Netw*. 2012;2012:1–11.
16. Domingo MC. A distributed energy-aware routing protocol for underwater wireless sensor networks. *Wirel Pers Commun*. 2011;57:607–627.
17. Huang C-J, Wang Y-W, Liao H-H, Lin C-F, Hu K-W, Chang T-Y. A power-efficient routing protocol for underwater wireless sensor networks. *Appl Soft Comput*. 2011;11:2348–2355.

18. Zhou Z, Peng Z, Cui J-H, Shi Z. Efficient multipath communication for time-critical applications in underwater acoustic sensor networks. *IEEE/ACM Trans Netw*. 2011;19:28–41.

19. Xu J, Li K, Min G. Reliable and energy-efficient multipath communications in underwater sensor networks. *IEEE Trans Parallel Distrib Syst*. 2012;23:1326–1335.

20. Ayaz M, Abdullah A, Faye I, Batira Y. An efficient dynamic addressing based routing protocol for underwater wireless sensor networks. *Comput Commun*. 2012;35:475–486.

21. Song Z, Deshi L, Jian C. A link-state based adaptive feedback routing for underwater acoustic sensor networks. *IEEE Sensors J*. 2013;13:4402–4412.

22. Wahid A, Lee S, Kim D. A reliable and energy-efficient routing protocol for underwater wireless sensor networks. *Int J Commun Syst*. 2014;27:2048–2062.

23. Ali T, Jung LT, Faye I. End-to-end delay and energy efficient routing protocol for underwater wireless sensor networks. *Wirel Pers Commun*. 2014;79:339–361.

24. Javaid N, Jafri M, Ahmed S, et al. Delay-sensitive routing schemes for underwater acoustic sensor networks. *Int J Distrib Sens Netw*. 2015;2015:1

25. Yang C-H, Ssu K-F, Yang C-L. A collision-analysis-based energy-efficient routing protocol in 3D underwater acoustic sensor networks. *Comput Commun*. 2015;66:25–35.

26. Yu H, Yao N, Liu J. An adaptive routing protocol in underwater sparse acoustic sensor networks. *Ad Hoc Networks*. 2015;34:121–143.

27. Jiang J, Han G, Guo H, Shu L, Rodrigues JJ. Geographic multipath routing based on geospatial division in duty-cycled underwater wireless sensor networks. *J Netw Comput Appl*. 2016;59:4–13.

28. Tariq M, Latiff A, Shafie M, Ayaz M, Coulibaly Y, Wahid A. Pressure sensor based reliable (PSBR) routing protocol for underwater acoustic sensor networks. *Adhoc Sens Wirel Netw*. 2016;32:175–196.

29. Coutinho RW, Boukerche A, Vieira LF, Loureiro AA. Geographic and opportunistic routing for underwater sensor networks. *IEEE Trans Comput*. 2016;65:548–561.

30. Goyal N, Dave M, Verma AK. Energy efficient architecture for intra and inter cluster communication for underwater wireless sensor networks. *Wirel Pers Commun*. 2016;1–21.

31. Dasgupta D, Michalewicz Z. *Evolutionary Algorithms in Engineering Applications*. Berlin, Germany: Springer Science & Business Media; 2013.

32. Shokouhifar M, Jalali A. A new evolutionary based application specific routing protocol for clustered wireless sensor networks. *AEU-Int J Electron C*. 2015;69:432–441.

33. Cuimei G, Sheng Z, Yaocheng X, Wei H. A survey on spectral clustering. *World Autom Congress (WAC)*. 2012;2012:53–56.

34. Liu B, Chen H, Zhong Z, Poor HV. Asymmetrical round trip based synchronization-free localization in large-scale underwater sensor networks. *IEEE Trans Wirel Commun*. 2010;9:3532–3542.

35. Zhou J, Chen L, Wang K. Path optimality conditions for minimum spanning tree problem with uncertain edge weights. *Int J Uncertainty, Fuzziness Knowledge-Based Syst*. 2013.

36. Allen AO. *Probability, Statistics, and Queueing Theory*. San Diego, CA, USA: Academic Press; 2014.

37. Felemban M, Felemban E. Energy-delay tradeoffs for underwater acoustic sensor networks. In *2013 First International Black Sea Conference on Communications and Networking (BlackSeaCom)*. Batumi, Georgia: IEEE; 2013: 45–49.

38. https://www.mathworks.com/products/matlab/

39. Azad S, Casari P, Zorzi M. Multipath routing with limited cross-path interference in underwater networks. *IEEE Wirel Commun Lett*. 2014;3:465–468.

40. Azad S, Hasan KT, Nandi D, Pathan A-SK. A high-throughput routing metric for multi-hop underwater acoustic networks. *Comput Electr Eng*. Elsevier, 2015;44: 24–33.

How to cite this article: Faheem M, Tuna G, Gungor VC. LRP: Link quality-aware queue-based spectral clustering routing protocol for underwater acoustic sensor networks. *Int J Commun Syst*. 2017;30:e3257. https://doi.org/10.1002/dac.3257