QERP: Quality-of-Service (QoS) Aware Evolutionary Routing Protocol for Underwater Wireless Sensor Networks

Muhammad Faheem, Gurkan Tuna, and Vehbi Cagri Gungor

***Abstract*—Quality-of-service (QoS) aware reliable data delivery is a challenging issue in underwater wireless sensor networks (UWSNs). This is due to impairments of the acoustic transmission caused by excessive noise, extremely long propagation delays, high bit error rate, low bandwidth capacity, multipath effects, and interference. To address these challenges, meet the commonly used UWSN performance indicators, and overcome the inefficiencies of the existing clustering-based routing schemes, a novel QoS aware evolutionary cluster based routing protocol (QERP) has been proposed for UWSN-based applications. The proposed protocol improves packet delivery ratio, and reduces average end-to-end delay and overall network energy consumption. Our comparative performance evaluations demonstrate that QERP is successful in achieving low network delay, high packet delivery ratio, and low energy consumption.**

***Index Terms*—Bioinspired, clustering, evolutionary, routing, underwater wireless sensor networks (UWSNs).**

I. Introduction

OVER the last few years, numerous quality-of-service (QoS) aware routing protocols have been proposed for underwater wireless sensor networks (UWSNs), such as vector-based forwarding (VBF) [1], depth-based routing protocol (DBR) [2], a link state based adaptive feedback routing protocol (LAFR) [3], distance-based reliable and energy efficient (DREE) routing protocol [4], reliable energy efficient routing protocol [5], aided efficient data gathering [6], and layer-by-layer angle-based flooding routing protocol [7]. In recent years, tendency toward using the clustering mechanisms has led to design of various novel routing solutions which balance the data traffic load and improve the overall scalability for remote monitoring applications [8], and accordingly some novel cluster solutions have been proposed, such as energy efficient signal-to-noise ratio (SNR) based clustering with reliable data encryption routing protocol [9], energy efficient and balanced energy consumption cluster based routing protocol [10], and sparsity-aware energy efficient clustering protocol for reliable data delivery [11]. Although the main aim of these routing schemes is to provide reliable and efficient data delivery in harsh underwater environment and these routing schemes present design objectives for UWSN-based applications, most of the existing routing schemes generally ignore the impact of external interference on transmission reliability in harsh underwater environments and achieve some design objectives at the expense of others. In addition, most of the existing routing schemes do not consider the issue of unnecessary multihop data packet transmission and data path looping for reliable data delivery in underwater networks and are unable to uniformly distribute energy consumption load around the whole network. Moreover, in the existing clustering-based routing schemes, due to the unbalanced data aggregation and forwarding traffic load, cluster heads (CHs) consume available resources more rapidly than others nodes. This unbalanced traffic distribution among the CHs results in spending more energy and may lead to early death of some nodes, which partitions the network and thereby degrades the overall network performance. Finally, the majority of the existing routing schemes have been designed to meet application-specific design objectives and requirements in a particular underwater scenario. The contribution of this paper is threefold. First, we propose an evolutionary clustering based routing protocol (QERP) for UWSNs to overcome the inefficiencies of the existing routing schemes. Second, the key performance requirements of UWSN applications, such as low delay, high packet delivery ratio (PDR) and low overall energy consumption, can be met by the proposed protocol. Third, the challenges caused by excessive noise, high interference, multipath effects, high bit error rate (BER), long propagation delay, and low bandwidth can be addressed if the proposed protocol is integrated into UWSN applications. Comparative performance evaluations demonstrate that QERP is successful in attaining the commonly used UWSN performance indicators. This paper is organized as follows. QERP is explained in Section II. Section III presents the simulation model and gives the results of the performance evaluation study. Finally, Section IV concludes the paper.

Manuscript received February 8, 2016; revised October 21, 2016, December 2, 2016, January 24, 2017, and February 21, 2017; accepted February 21, 2017. Date of publication March 8, 2017; date of current version August 23, 2018. This work was supported by the Turkish Scientific and Technical Research Council under Grant 114E248.

M. Faheem and V. C. Gungor are with the Department of Computer Engineering, Abdullah Gul University, Kayseri 38039, Turkey (e-mail: smfaheem@ymail.com; cagri.gungor@agu.edu.tr).

G. Tuna is with the Department of Computer Programming, Trakya University, Edirne 22020, Turkey (e-mail: gurkantuna@trakya.edu.tr).

Digital Object Identifier 10.1109/JSYST.2017.2673759

II. Proposed Routing Scheme

In our model, a network of underwater sensors at different depths is deployed to obtain cooperative sampling of three-dimensional (3-D) ocean scenario. In this architecture, entire 3-D modeling architecture is mapped to 2-D and the

following properties are assumed. First, all randomly deployed sensor nodes have the same capabilities in terms of transmission range and initial energy, and sensor nodes near to the sea surface buoys (sink) can directly communicate with each other. Second, each sensor node knows its own location by means of a location service [12]. Third, every node only moves in the horizontal direction at random, but also moves slightly in the vertical direction. Fourth, the acoustic channel is symmetric so the energy required to transmit a message from sensor 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 SNR. Fifth, every sensor node can dynamically adjust its low and high transmission power. Sixth, we assume carrier sense multiple access (CSMA) mechanism. Moreover, it is also assumed that the sink node is rich in resources and is capable of receiving multiple packets at the same time.

A. Generic Representation of Scheme

In our proposed scheme, we consider a chromosome (routing path), which consists of sequences of nonnegativity integers that denote the IDs of genes (CH nodes) through which a routing path passes. In a routing path, the order of each CH is represented by the locus of chromosome in which the gene of the first locus is always reserved for the source CH nodes. The length of each constructed routing path is variable; however, it must not be greater than the total number of defined CH nodes to avoid data path loops in the network. Thus, a chromosome encodes the problem from its origin CH node toward destination by listing up the node IDs depending on topological information obtained from the constructed routing table in underwater acoustic network. Physically, several alternative partial routes from the source CH node to the destination are generated by the proposed scheme in a greedy manner. The one with better fitness value based on high residual energy, SNR, least number of member nodes, and minimum distance to the sink is accepted as a final route. After the route selection, each CH node is responsible for maintaining history of the relay nodes in its routing table. This phenomenon avoids the energy depletion of the same CH nodes being selected twice in different routing paths.

B. Population Initialization

In population initialization phase of any clustering-based routing solution, initial population size and procedures to initialize the population are two main issues to be addressed. To generate the initial population, heuristic or random initializations are preferred. Due to its robustness in complex problematic search space [13], we consider random initialization to generate an initial population. After the deployment, surface buoys are responsible for running an evolutionary algorithm to generate an initial population of individuals by using a random number generator ϕ_r. Initially, it disseminates multiple copies of an initiation message ($\mathrm{init\_msg}$) to neighboring nodes in underwater environment. After receiving information from the sink, each node is responsible for disseminating initiation message to its one hop or maximum two hops neighboring nodes by taking into account the CSMA mechanism. This initiation message contains information about the sender node identity, and residual energy. After receiving the initiation message successfully, each receiver node replies to the sender node by sending an acknowledgment message ($\mathrm{ack\_msg}$), which guarantees that a message has been received successfully. This acknowledgment message contains the information of the receiver node identity, residual energy, and received signal strength indicator (RSSI) measured while receiving information from the sender node. Throughout the sending and receiving process, each node is responsible for maintaining the information on neighboring node identity, residual energy, and RSSI. Then, a looping feature of selection (ϕ_s), crossover (ϕ_c), and mutation (ϕ_m) operators is applied on each individual to improve the quality of the solution through the predefined probabilities (ϕ_p) until the termination criterion is satisfied for the set of complete clustering solution indicated as $\mathcal{C}_n = \{\mathcal{C}_1, \mathcal{C}_1, \ldots, \mathcal{C}_k\}$.

C. Selection of Parents

The selection process determines which of the chromosomes from the current population undergoes recombination to create offspring to populate the next generation in order to survive. The selection weight defines the selection probability ratio of the elite chromosomes to that of average chromosomes in the population. Among various selection schemes, tournament selection is found robust to preserve the selection noise at extremely low as possible. In the tournament selection, nonoverlapping random sets of chromosomes are chosen from the population sets in such a way that the fittest chromosome from each set is preserved as parents for the next generation [14]. We consider a pairwise tournament selection without replacement that randomly chooses two individuals from the population. It selects the best individuals with higher fitness values to be parents by avoiding the good-quality individuals that can be numerically expressed as

$$\phi_s = \frac{\psi_{f(j)}}{\sum_{i=1}^{n} \psi_{f(i)}}, \quad i = 1, 2, \ldots\ldots, n \tag{1}$$

where $\psi_{f(j)}$ is the fitness of the individual j and n is the population size.

D. Crossover

The crossover process examines the current solutions for generating offspring from the selected parents in a way that combines and maintains the desirable features from both parents. In this study, we prefer basic crossover operators, including multiparent and universal parent operators to exchange parent chromosomes information. In a multiparent crossover, two or three parental strings are randomly picked out from the created mating pool (ϕ_{mp}). One of them is called mother string (ϕ_{ms}), which generates offspring after crossover with the remaining parental strings (ϕ_{ps}). To generate a new offspring, mother string is chosen from the created mating pool in universal–parent crossover and then combined with the universal parent (a complete set of clusters) at very low probability. In all, nine categories of different crossover operators such as simple uniform (ϕ_{su}), universal parent uniform (ϕ_{upu}), multiparent uniform (ϕ_{mpu}), simple

single point ($\phi_{\rm sp}$), universal parent single point ($\phi_{\rm ups}$), multiparent single point ($\phi_{\rm mps}$), simple two point ($\phi_{\rm s2p}$), universal parent two point (ϕ_{u2p}), and multiparent two-point (ϕ_{m2p}) are considered. In $\phi_{\rm upu}$ crossover, first $\phi_{\rm ms}$ and two others $\phi_{\rm ps}$ are selected and then one by one, members are randomly selected with equal probability from the $\phi_{\rm ps}$ in the $\phi_{\rm mp}$. In ϕ_{su} crossover, while creation members in the $\phi_{\rm mp}$ are randomly chosen one by one either from the $\phi_{\rm ms}$ or other $\phi_{\rm ps}$ with equal probability. In $\phi_{\rm sp}$ crossover, first a single crossover point c_1 is randomly chosen then substrings on the left and right side of the crossover point are combined in order to produce an offspring by employing $\phi_{\rm ms}$ and another $\phi_{\rm ps}$, respectively. In $\phi_{\rm s2p}$ crossover, first two points (c_1, c_2) are randomly chosen then left, right, and in between sides substrings are combined in order to produce an offspring by employing $\phi_{\rm ps}$, $\phi_{\rm ms}$, and another $\phi_{\rm ps}$, respectively. In $\phi_{\rm ups}$ crossover, first a signal crossover point is randomly chosen then on the left side of the crossover point one by one members in the $\phi_{\rm mp}$ are randomly selected with equal probability from the $\phi_{\rm ms}$. Similarly, on the right side of the crossover point one by one members in the $\phi_{\rm mp}$ are randomly elected with equal probability from another $\phi_{\rm ps}$. In ϕ_{u2p} crossover, first two crossover points are randomly chosen then in between these two crossover points one by one members in the $\phi_{\rm mp}$ are randomly selected with equal probability from the $\phi_{\rm ps}$. Whereas on the left and right side of the crossover points one by one members in $\phi_{\rm mp}$ are randomly selected with equal probability from the $\phi_{\rm ms}$. The implementation of multiparent operators ($\phi_{\rm mpo}$) is same as the universal parent operators ($\phi_{\rm upo}$) explained above. The proposed scheme assumes a member between two crossover points c_1 and c_2 from another $\phi_{\rm ps}$ and the remaining $\phi_{\rm ms}$. The implementation employs $\phi_{\rm ms}$, another $\phi_{\rm ps}$, and a universal parent string containing the set of feasible cluster sizes. Since the universal parent containing variable size clusters, it offers an opportunity to select appropriate size clusters from the candidate solution at low selection probability. To generate a new population of individuals, we define a decreasing individual crossover rate probability as $\phi_c^1 = 0.95$ and $\phi_c^2 = 0.90$ is the parameters that affect the rate at which the crossover operator is applied. In this way, new clusters are introduced more quickly into the population. To generate a new clustering solution, an individual is randomly selected from the mating pool. For a given crossover probability, each pair of strings from the new individual population is randomly picked out by the proposed algorithm. Then, a uniformly distributed random number is generated between 0 and 1. If the generated random number is less than the predefined value, then the algorithm is responsible for applying the crossover operator to produce swapping generation. Otherwise, it does not apply crossover operator on these two particular strings. Through applying the crossover operator, a $\phi_{\rm ms}$ is chosen with crossover probability $\phi_c^1 = 0.95$ or $\phi_c^2 = 0.90$. Then, to select a crossover operator a random number between 0 and 9 is generated. For a given crossover probability, on average clusters crossover in each generation is given as: $\phi_c^1 = 100 \times 0.95 = 95$ and $\phi_c^2 = 100 \times 0.90 = 90$, depending upon the operator selected. By assuming that if $y^1 = (y_1^1, y_2^1, \ldots, y_{\rm nd}^1)$ and $y^2 = (y_1^2, y_2^2, \ldots, y_{\rm nd}^2)$ are the two parents selected for crossover, then two offspring which are generated can be indicated as follows:

$$x^k = \left(x_1^k, x_2^k, \ldots, x_{\rm nd}^k\right),\ k = 1, 2 \tag{2}$$

subjected to

$$x_i^1 = \phi_c^i\, y_i^1 + \left(1 - \phi_c^i\right) y_i^2 \ \&\&\ x_i^2 = \phi_c^i\, y_i^2 + \left(1 - \phi_c^i\right) y_i^1 \tag{3}$$

in which x_i^1 and x_i^2 are the two offsprings generated by applying crossover probability ϕ_c^i ($i = 1$ or 2) on two selected parents y_i^1 and y_i^2, respectively.

E. Mutation

The mutation process plays a significant role to enhance the fitness of the solution [15]. We use a nonuniform mutation operator instead of uniform mutation operator. The probability of the mutation rate increases from extremely low ($\phi_m^1 = 0.01$) to its maximum value ($\phi_m^n = 0.05$) in order to perform the mutation operation with the assertion that no important genetic material is lost. On average, an individual gene (one clustering solution) is improved in each offspring with very low ($\phi_m^1 = 100 \times 0.01 = 1\%$) to high ($\phi_m^n = 100 \times 0.05 = 5\%$) probability of mutation, respectively. After applying crossover, each string in the new generation is evaluated bit by bit by generating a random number between 0 and 1. If the random number value is less than the predefined value, then the developed algorithm is responsible for applying the mutation operator to the new individual else, which does not ponder mutation operator to that particular node. This genomic type representation eases the creation of dynamic clusters trickily in the whole network. Thus, sensitivity to mutation probability can be expressed as

$$\phi_m = \frac{\psi_k}{\psi_l} \tag{4}$$

where ψ_k is the number of mutation and ψ_l is the chromosome length (in genes).

By assuming that if $x(k)$ is an offspring and x_i^k is the randomly selected gene for mutation, then gene combines (z_i^k) after mutation can be indicated as follows:

$$z_i^k = x_i^k + N\left(\phi_m^i, 0\right) \tag{5}$$

where N is the random number selected from zero to defined maximum value of nonuniform mutation operator ϕ_m^i.

F. Fitness Function

In each generation, a set of fittest individuals receives higher fitness values and thus has a higher chance of surviving in the resulting generation. Better fitness function leads to lower clustering cost and more stable link quality for reliable data transmission with minimum energy consumption. Here, the selection operator is responsible for assigning each set of pair of nodes a positive weight $\psi_{w(ij)}$ in the population in order to find a partition graph (G) for generating dynamic clusters, such as $G \in \psi_{w(ij)}$. The value of this assigns weight based on the RSSI, which is proportional to the fitness function of each individual

in the cluster $\mathcal{C}_i$. Hence, the probability of fitness function that involves computational efficiency and accuracy at low cost for each cluster $\mathcal{C}_j$ is defined as follows:

$$\phi_{\rho(\mathcal{C}_j)} = \psi_{\mathcal{C}_j} / \sum_{i=1}^{n} \psi_{\mathcal{C}_i} \tag{6}$$

in which $\phi_{\rho(\mathcal{C}_j)}$ is the probability for the cluster $\mathcal{C}_j$, $\psi_{\mathcal{C}_j}$ is the fitness value for a cluster $\mathcal{C}_j$, and $\sum_{i=1}^{n} \psi_{\mathcal{C}_i}$ is the set of available clusters. The probability of an individual node i being selected as a member in cluster $\mathcal{C}_j$ with the highest fitness value is given as

$$\phi_{(Sn_i)} = \psi_{Sn_j} / \sum_{i=1}^{n} \psi_{Sn_i} \tag{7}$$

where $\phi_{(\mathrm{Sn}_i)}$ is the probability for the individual (sensor node) Sn_i, ψ_{Sn_j} is the fitness value for an individual Sn_j from the set of individuals $\sum_{i=1}^{n} \psi_{\mathrm{Sn}_i}$. We take into account RSSI to guarantee link reliability among sensor nodes in the network and can be numerically indicated as

$$\phi_{LQ(\mathrm{Sn}_{ij})} = \frac{\psi^{j}_{\mathrm{RSSI}}}{\psi_{s\max}} \in [1, 0] \tag{8}$$

where $\psi_{s\max}$ is the maximum signal strength value and ψ^{j}_{RSSI} is the RSSI of the neighboring node j measured at node i while receiving control or data packet. The value 1 and 0 shows probability of the best and poor link quality, respectively. Thus, the minimum Euclidean distance $E_d(\min)$ between each set of pair of individuals i and j in terms of weight w_{ij} can be calculated as follows:

$$\phi_{E_d(\min)} = \sum_{i=1}^{n} \sum_{j=1}^{k} \psi_{w(i,j)} ||Sn_i - Sn_j||^2. \tag{9}$$

The delay (D_e) between each set of pair of individuals i and j in terms of weight w_{ij} can be defined as

$$\phi_{D_e} = \sum_{i=1}^{n} \sum_{j=1}^{k} \psi_{w(i,j)} ||\mathrm{diff}_{ij}||^2 \tag{10}$$

where $\psi_{w(i,j)} \in \mathrm{diff}_{ij}$ is the difference between individuals Sn_i and Sn_j is based on the packet arrival and departure time. The residual energy (R_e) of each individual in terms of weight $\psi_{w(i,j)}$ can be numerically written as

$$\phi_{R_e(\mathrm{maxi})} = \sum_{i=1}^{n} \sum_{j=1}^{k} \psi_{w(i,j)} ||R_e(Sn_i, Sn_j)||^2. \tag{11}$$

To measure the similarity of the individuals Sn_i associated to the jth best individual in a cluster based on RSSI can be represented as

$$\phi_{y(j)} = \frac{1}{n} \sum_{i=1}^{n} \psi_{w(Sn_i, Sn_j)} \left(Sn_i,\ \psi_{x(i)}(Sn_j)\right) \tag{12}$$

where $\phi_{y(j)}$ is the centre of the $j\mathrm{th}$ cluster, $\psi_{x(i)}$ is the position of the $j\mathrm{th}$ individual in a cluster, and $\psi_{w(\mathrm{Sn}_i, \mathrm{Sn}_j)}$ is weight of individuals associated to the cluster centre (c_j) can be either 1 or 0, such that

$$\psi_{w(\mathrm{Sn}_i, \mathrm{Sn}_j)} = \begin{cases} 1, & \text{if } \mathrm{Sn}_i \in c_j \\ 0, & \text{otherwise.} \end{cases} \tag{13}$$

The maximum $(\max)$ cohesion associated between a set of individuals (sensor nodes) and a cluster centre can be mathematically defined as

$$\phi_{\mathrm{RSSI(Cohesion)}}(\mathrm{Sn}_i,\ M) = \max \sum_{i}^{n} \left[\sum_{\forall Sn_i \in C_j} \mathrm{RSSI}(\mathrm{Sn}_i, c_j) / |\mathrm{Sn}_i(c_j)| \right] \tag{14}$$

in which M illustrates the individual sensor node clustering domain characteristic value of the matrix and $|\mathrm{Sn}_i(c_j)|$ is the number of sensor nodes in the cluster including CH. The average RSSI of all the individuals to their associated clusters can be numerically defined as

$$\phi_{\mathrm{RSSI}_{\mathrm{Sn}_i\ (\mathcal{C}_n)}} = \frac{\sum_{i,j=1}^{n} \left[\sum_{\forall \mathrm{Sn}_i \in \mathcal{C}_i} \mathrm{RSSI}(\mathrm{Sn}_i,\ c_j)\right] / |\mathrm{Sn}_i(c_j)|}{N_c} \tag{15}$$

in which $\mathcal{C}_i$ shows the total number of clusters in the entire search space, Sn_i is the number of individuals belong to a cluster $\mathcal{C}_i$, and $\mathrm{RSSI}(\mathrm{Sn}_i,\ c_{ij})$ shows the RSSI between a set of sensor nodes to their associated cluster centre. Thus, for the m CHs and n number of sensor nodes in each cluster $\mathcal{C}_i$, the standard deviation of CH load is given by

$$\phi_{\mathrm{CH(Load)}} = \sqrt{\frac{\sum_{i=1}^{m} (\psi_{\propto} - \psi_j)^2}{m}} \tag{16}$$

in which $\psi_{\propto}$ is the average load $= \sum_{j=1}^{n} \mathrm{Sn}_{j(\mathrm{Load})}/m$, $\mathrm{Sn}_{j(\mathrm{Load})}$ is the load a of sensor node j, and ψ_j is the overall load of the CH $\mathrm{CH}_{i(\mathrm{Load})}$. We choose the fitness function values as the reciprocal of the standard deviation of the CH load $\phi_{(\mathrm{CH}_i)=1/\phi_{\mathrm{CH(Load)}}}$. The maximum separation between any pair of clusters can be numerically written as

$$\phi_{\mathrm{RSSI(Separation)}} = \max i._{\forall \mathcal{C}_{j1}, \mathcal{C}_{j2},\ \mathcal{C}_{j1} \neq \mathcal{C}_{j2}} \times \sum_{i}^{n} \{\mathrm{RSSI}(c_{j1}(\mathrm{Sn}_i), c_{j2}(\mathrm{Sn}_i))\} \tag{17}$$

in which c_{j1} and c_{j2} are two different clusters that have different numbers of sensor nodes $\mathrm{Sn}_i, i \in \{1, 2, \ldots, n\}$. The total energy consumed while appointing CHs in the network is measured as the sum of the energy dissipated during communication. Thus, for $\mathrm{CH}^i, i \in \{1, 2, \ldots, n\}$ the total number of CH generated in the network can be written as

$$\phi_{E(\mathrm{CH})_{i,\ldots,n}} = \sum_{j=1}^{nc} \sum\nolimits_{C \in K_i} \sum\nolimits_{(T_x + R_x)_{\mathcal{C}_j}} \mathrm{CH}_i \tag{18}$$

where nc in $\sum_{j=1}^{nc} \sum_{C \in K_i}^{i} E_{(T_x+R_x)_{\mathcal{C}_j}} \mathrm{CH}_i$ represents the total number of CHs, $C \in K_i$ is a non-CH sensor node associated

with the ith cluster head node. T_x and R_x are the transmission and receiving power consumption, respectively. Thus, the fitness function in terms of each individual cluster while appointing can be written as

$$\phi_{\mathrm{Fit}_i} = 1/E_i = \sum_{j=1}^{n} \sum_{(T_x+R_x)_{\mathcal{C}i}} \mathrm{CH}_i. \tag{19}$$

CHs selection cost in terms of fitness function becomes

$$\phi_{\mathrm{Fit}} = 1/\sum_{j=1}^{n} \sum_{c\in k_i} \sum_{(T_x+R_x)_{\mathcal{C}i}} \mathrm{CH}_i. \tag{20}$$

The probability of a particular CH node with minimum data traffic load and energy consumption is selected as a relay node to convey information is given as

$$\phi_{\rho b\ (\mathrm{CH}_j(\mathrm{Load}))} = \mathrm{CH}_j / \sum_{i=1}^{\mathcal{C}_n} \mathrm{CH}_i. \tag{21}$$

The algorithm used in our scheme works iteratively and in each iteration cluster centroids $\mathcal{C}_{\mathrm{ci}}$ and cluster associated function $F_{\mathcal{C}_i}$ are updated for each cluster $\mathcal{C}_i$ as follows:

$$\phi_{F(\mathcal{C}_i)} = \left(\sum_{j=1}^{\mathcal{C}} \left(\frac{\mathrm{RSSI}_{\mathcal{C}_i}}{\mathrm{RSSI}_{(\mathrm{Sn}_i,\ \mathrm{CH}_j)}} \right)^{2/(m-1)} \right)^{-1} \quad \forall m$$
$$= 2, .., n \tag{22}$$

$$\phi_{\mathrm{RSSI}(\mathcal{C}_i)} = \left(\sum_{t=1}^{n} \left(t_{\mathrm{CH}_j} - \mathcal{C}_{\mathrm{CH}_j} \right)^2 \right)^{-\frac{1}{2}} \tag{23}$$

$$\phi_{\mathcal{C}(c_i)} = \left(\sum_{i=1}^{p} \left(\phi_{F(\mathcal{C}_i)} \right)^m \cdot t_{\mathrm{CH}_j} \right) \left(\sum_{i=1}^{p} \left(\phi_{F(\mathcal{C}_i)} \right)^m \right)^{-1} \quad \forall p$$
$$= 2, \ldots, n \tag{24}$$

where $\mathrm{RSSI}_{\mathcal{C}_i}$ is the RSSI value between a cluster centroid c_i and sensor node i and t_{CH_j} is the maximum active time of a CH. The iteration process terminates when the magnitude of the change in the cluster associated values decreases below a defined threshold value ($0 \leq \phi_{F(\mathcal{C}_i)} \leq 1$).

Relay nodes can change their transmission power level to save energy but a relay node can switch to its higher transmission power level only if it fails to find an appropriate neighboring relay node. The transmission range of a relay node can be numerically written as

$$\phi_{\mathcal{C}_{R_i(\max)}} \cong \max. \sum_{(i,j)\ \in\ \mathcal{C}} \mathcal{C}_{ij} \cdot \phi_{\mathcal{TP}_{ij}} \tag{25}$$

in which $\phi_{\mathcal{C}_{R_i(\max.)}}$ is transmission range of the relay node by considering transmission power $\phi_{\mathcal{TP}_{ij}}$ between a set of pair of CHs node i and j in the routing network. The cost of a relay node in terms of energy consumption can be expressed as

$$\phi_{C_{\mathrm{RN}_i}} = \frac{E_{\max}}{E(i)} \tag{26}$$

$$\phi_{C_{\mathrm{RN}_i}} = \frac{E_i(0)}{\left(E_{\mathrm{RN}_i}(0) - \sum_{k=0}^{n} E_{\mathrm{RN}_i \cdot k}\right.} \tag{27}$$

where $\phi_{C_{\mathrm{RN}_i}}$ is the node cost, $E_{\max}$ is the maximum energy utilization, $E(i)$ is the available energy at a relay node i, and E_{RN_i} is the energy consumed by a relay node RN_i for the kth update. Here, it is important to note that with high R_e of a relay node, the value of C_{RN_i} decreases rapidly, which means that it has more probability to be allocated to a higher transmitting power. The sum of the transmitting power $(\phi_{\mathcal{TP}_{\mathrm{Sum}}})$ for an individual relay node RN_i in a chain (C_i) like routing network can be formally shown as

$$\phi_{\mathcal{TP}_{\mathrm{Sum}}} = \sum_{\mathrm{RN}_i \in N} \mathcal{TP}_{\mathrm{RN}_i} \left(\psi_{\mathcal{TP}_{\mathrm{low}}} \psi_{\mathcal{TP}_{\mathrm{high}}} \right) \tag{28}$$

subjected to the number of links in a chain (C_i) like multihop routing network having length (L) between relay nodes RN_i and RN_j, i.e.,

$$C_i = \mathrm{RN}\left(L_{ij} + {L_{ij}}^2 + , \ldots, + {L_{ij}}^{|N|-1} \right). \tag{29}$$

The number of relay nodes appointed as active nodes in the clustering network in round R_i by considering $\psi_{\mathcal{TP}_{\mathrm{low}}}$ and $\psi_{\mathcal{TP}_{\mathrm{high}}}$ can be shown as

$$\mathrm{CN}_{\mathrm{RN}} = \sum_{i=1}^{N} \mathrm{RN}^i_{\mathrm{NS-SLN=ARN}} \tag{30}$$

where NS shows the total number of active relay nodes, SLN are the sleeping or inactive relay nodes, and ARN are the active relay nodes available for data transmission in the network.

G. Termination Criterion

The fitness of each new generated individual is computed by using fitness function and each individual with lower fitness is replaced with the one with higher fitness value. The decision variable to select the best individual i for the current round j can be expressed as

$$\phi_i^j\ (R_i) = \begin{cases} \psi^j_{i(\mathrm{new})}, & \psi^j_{i(\mathrm{previous})} < \psi^j_{i(\mathrm{new})} \\ \psi^j_{i(\mathrm{previous})}, & \text{otherwise} \end{cases} \tag{31}$$

where $\phi_i^j(R_i)$ is the fittest individual in region R_i in the deployed network, $\psi^j_{i(\mathrm{new})}$ indicates the new individual, and $\psi^j_{i(\mathrm{previous})}$ indicates the parent old individual.

III. Simulation Model and Performance Evaluation

The communication channel model derived in [16] is used in this study. The path loss, $\mathrm{PL}(d, f)$, of acoustic signal with transmission range d in meters and frequency f in kHz is given in dB as

$$\mathrm{PL}\ (d, f) = k \log(d) + \alpha(f)\ d\ \times\ 10^{-3} \tag{32}$$

where $\alpha(f)$ is the absorption coefficient and k is the spreading factor caused by energy spreading. We assume orthogonal frequency division multiplexing (OFDM) encoding technique and use the quadrature amplitude modulation (QAM) scheme to calculate the amplitude and phase of the subcarrier. In our model, 16-QAM modulation with OFDM transmission is considered.

BER is computed and SNR is calculated as

$$P_b^{16\mathrm{QAM}} = \frac{3}{2k}\ erfc\left(\sqrt{\frac{k}{10}\frac{E_b}{N_0}}\right) \tag{33}$$

$$\mathrm{SNR} = 10^{\mathrm{SNR}(d,f)/10}. \tag{34}$$

In QERP, we define PDR numerically as

$$\mathrm{PDR} = \frac{\mathrm{DP}_{\mathrm{received}}}{\mathrm{DP}_{\mathrm{send}}} \tag{35}$$

where $\mathrm{DP}_{\mathrm{send}}$ and $\mathrm{DP}_{\mathrm{received}}$ are the total number of data packets send by the source nodes and data packets received successfully at the sink, respectively,

$$T_{\mathrm{avg}}^{P_i} = \mathrm{DP}_{\mathrm{Sn}_i}^{\mathrm{Source}} + T\sum_{i=1}^{n}\mathrm{RDP} + \mathrm{DP}_{\mathrm{Sn}_i}^{\mathrm{Sink}} \tag{36}$$

in which $T_{\mathrm{avg}}^{P_i}$ is the average time to send a data or request packet, $\mathrm{DP}_{\mathrm{Sn}_i}^{\mathrm{Source}}$ is the time when a source node Sn_i generates data or control packet, $T\sum_{i=1}^{n}\mathrm{RDP}$ is the time taken by relaying data packets, and $\mathrm{DP}_{\mathrm{Sn}_i}^{\mathrm{Sink}}$ is the time when data packets reached successfully at the sink. The delay probability of the CHs-based multihop chain like routing network, i.e., $\rho_b(C_n)$ can be computed as

$$P_b\left(C_n\right) = \prod_{\rho_b \in C_i} P_b\left(D_e C_n < T_d\right) \tag{37}$$

in which $\rho_b(D_e C_n < T_d)\ \rho_b((D_e C_n < T_d))$ is the delay probability and C_n is the sum of individuals chain paths C_i each of length less than defined threshold T_d. Then, using the Erlang-k distribution method [17], it can be calculated as

$$P_b\ \left(C_n\right) = \frac{\mu^k {T_d}^{k-1} e^{-\mu T_d}}{(k-1)!}. \tag{38}$$

The fittest number of hops involved to convey information from the source toward sink can be estimated. Let the average path length of the message and communication range among two following relay nodes $\mathrm{RN}_{\mathrm{CH}_i,\mathrm{CH}_j}$ in the network is P_l and R_c, respectively. Then, projected sum of hops from the source toward destination ($H_{s\to d}$) can be numerically defined as

$$H_{s\to d} = \frac{P_l}{R_c}\ . \tag{39}$$

Here, the path length P_l function depends on the relay nodes traffic pattern, relay nodes density, and network coverage area. If the next hop relay nodes have the equal probability, then a relay node selects next hop node opportunistically. Otherwise, a next hop relay node with minimum transmission distance near to the destination is preferred to convey information. R_e can be numerically defined as

$$R_e = E_{\mathrm{initial}}\ - \mathrm{TEC} \tag{40}$$

where E_{initial} is the initial energy and total energy consumption (TEC) is the sum of various factors which can be numerically defined as

$$\mathrm{TEC} = \sum_{i=1}^{n} C_{\mathrm{lc}} + C_{\mathrm{Hc}} + C_R + \sum_{j=1}^{n} C_{\mathrm{comm}} + C_{\mathrm{DI}} + C_{\mathrm{DA}} \tag{41}$$

where C_{lc} is the sum of the clustering cost, C_{Hc} indicates the CHs appointing cost, C_R shows the routing cost, C_{comm} is the communication cost, C_{DI} is the ideal energy consumption, and C_{DA} indicates the data aggregation energy consumption cost in cluster-based routing network.

In our performance evaluation study, conducted using MATLAB 7.0 [18], we analyze and compare the performance of QERP with DBR and VBF schemes. The standard deviation of the 53 sets of simulations with 95% confidence intervals is considered using batch mean method [19] in order to provide the consistent results. We assume that a total number of 350 sensor nodes have the ability of communicating each other though acoustic channel. These sensor nodes are deployed in 3-D underwater area with dimensions in meters $(1000\ (\mathrm{l}) \times 1000(\mathrm{w}) \times 100(\mathrm{h}))$. To satisfy the entire network coverage problem, the maximum communication range of each sensor node is set to 50 m, whereas the average communication range is set to 40 m. Initial energy, data aggregation, ideal listening, sleeping, and receiving power are set to 3.5 J, 0.043 W, 0.063 W, 3 $\times$ 10^{-6} W, and 0.065 W, respectively. We divided the communication energy consumption into two categories 0.93 W for short and 1.31 W for long range communication and set the maximum data rate, packet length, and control packet size to 15 kb/s, 1024 b, and 28 b, respectively. The values for temperature in °C, salinity, acidity and cylindrical spreading coefficient factor were set to 35, 8, 22, and 1, respectively.

It is shown in Fig. 1(a) that PDR in all routing schemes increases linearly with the increase in round numbers between 1 and 3500. Initially, when the network size is small between round numbers 1 and 1000 the performance of VBF in achieving higher PDR is relatively better than DBR routing scheme. However, when the network size grows up between round numbers 1100 and 2300, the performance of DBR in achieving higher PDR is a little better than VBF. The similar behavior is observed in the ending region when network size becomes extremely large between round numbers 2400 and 3500. In all the regions, the PDR of QERP is found to be significantly better than DBR and VBF routing schemes. On the other hand, we notice that the PDR of DBR is relatively better than VBF due to forwarding data packets over relatively better links among nodes in UWSNs. Fig. 1(b) shows that the delay performance in all routing schemes rapidly decreases with the increase in nodes between 1 and 350. When the network consists of up to 350 sensor nodes, the delay level performance of DBR is better than VBF. In all three regions, the delay performance of QERP is better compared to both DBR and VBF routing schemes. We also observe that the delay performance of DBR is relatively better than VBF due to considering shortest path while forwarding data packets. As shown in Fig. 1(c), in all of the routing schemes, residual energy of sensor nodes decreases with the increase in round numbers between 1 and 3500. While in the beginning, the

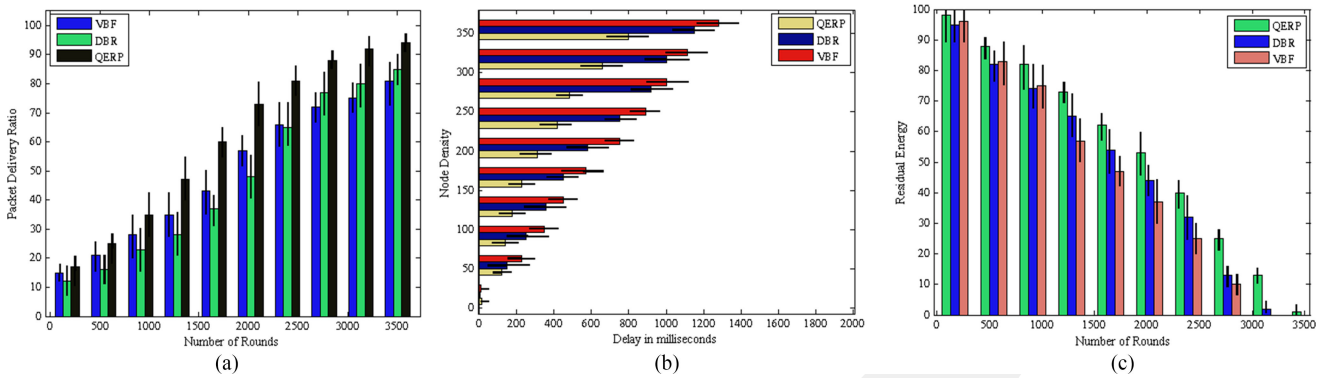

Fig. 1. (a) PDR versus round numbers between 1 and 3500, (b) the delay versus node density between 1 and 350, and (c) the residual energy versus number of round numbers between 1 and 3500.

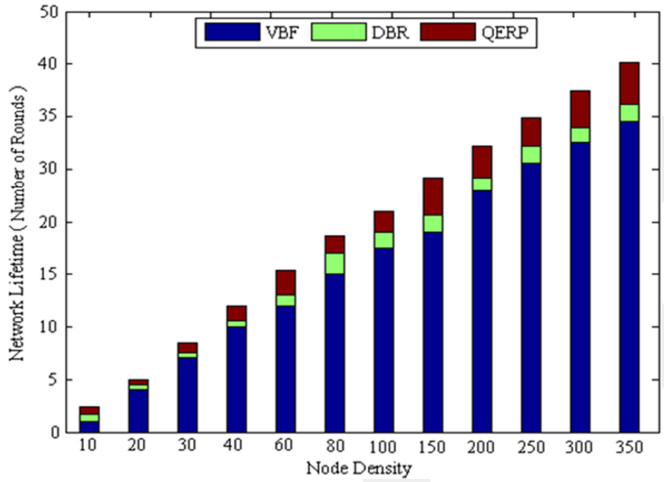

Fig. 2. Network lifetime versus node density between 1 and 350. The bars indicate the average of the 53 sets of simulations.

energy consumption profile of VBF is better than DBR, in round numbers between 1100 and 2200 and between 2300 and 3150 the residual energy performance of DBR is better than VBF. On the other hand, in terms of residual energy performance, QERP is better than both DBR and VBF since it keeps sensor nodes alive in various rounds. Finally, in terms of network lifetime, due to its novel clustering mechanism, small-size clusters are generated to organize sensor nodes into a connected hierarchy in order to balance load and prolong network lifetime. As shown in Fig. 2, in terms of network lifetime, QERP is better than both DBR and VBF.

In sum, due to its highly stable and reliable clustering and routing mechanisms, in all the regions the performance of QERP is better than all the other routing schemes in terms of achieving a high PDR, high residual energy, and low delay. In the clustering mechanism of QERP, small-size clusters are generated to create a connected hierarchy for load balancing and prolonging lifetime. Moreover, QERP utilizes residual energy, locations, and end-to-end delay to search for the next-hop relay CH node in a greedy manner and considers multihop data transmission among the CHs, and this brings significant improvement in terms of energy consumption and data traffic load balancing. With its distinct features, QERP maintains the reliability of data transmission and provides loop free robust data delivery by considering energy efficiency and delay requirements to maximize throughput of the network compared to existing routing schemes. To sum up, a number of reasons degrade network performance in both DBR and VBF routing schemes: First, in case of large network deployments, excessive message retransmission rate causes more energy consumption in VBF and DBR routing schemes due to data packet losses caused by node buffer overflows. Second, in VBF, unstable link quality between sensor nodes results in rapid energy depletion. Third, in large network deployments, DBR fails to maintain its stability period and selects next hop nodes with less residual energy and a notable amount of data packets become lost or invalid due to not reaching or late reaching at the sink. Fourth, both in DBR and VBF, a large number of hops are typically setup between a source node and the sink and this may lead to greater delay. Hence, energy in sensor nodes lying at low depths far from the sink exhaust more rapidly and this may cause voids problem.

IV. Conclusion and Future Work

In this paper, we proposed a QERP to improve reliability in data transfer for energy and delay for real-time UWSN-based underwater applications. In QERP, highly stable small clustering mechanism is used to organize sensor nodes into a connected hierarchy for distributing energy and data traffic load evenly in the network. QERP successfully exploits highly reliable link quality information among the CHs in a greedy manner leading to successful transmissions toward the sink. Moreover, due to its QoS aware shortest path selection mechanism, it effectively reduces data path loops, network delay, and energy consumption. Moreover, in case of a route node failure, its routing table and simple dynamic power adjustment mechanisms lead to find an appropriate next hop relay node by avoiding connectivity voids. Thus, QERP significantly decreases the probability of packet loss and preserves high link quality among nodes in highly dynamic underwater environments. The performance evaluation results show that QERP achieves excellent performance in

terms of the metrics, such as the PDR, average end-to-end delay, and energy consumption. Future work of this study aims to improve the proposed algorithm so that it takes mobility scenarios into account, and focus the impacts of relationship between the network traffic, node density, and data transmission rate to reduce the probability of collisions on MAC layer for time-critical applications of UWSNs.

REFERENCES

[1] P. Xie, J.-H. Cui, and L. Lao, "VBF: Vector-based forwarding protocol for underwater sensor networks," in *Proc. 5th Int. IFIP-TC6 Conf. Netw. Technol., Serv. Protocols/Perform. Comput. Commun. Netw./Mobile Wireless Commun. Syst.*, 2006, pp. 1216–1221.

[2] H. Yan, Z. J. Shi, and J.-H. Cui, "DBR: Depth-based routing for underwater sensor networks," in *Proc. 7th Int. IFIP-TC6 Netw. Conf. AdHoc Sensor Netw., Wireless Netw., Next Gener. Internet*, 2008, pp. 72–86.

[3] S. Zhang, D. Li, and J. Chen, "A link-state based adaptive feedback routing for underwater acoustic sensor networks," *IEEE Sensors J.*, vol. 13, no. 11, pp. 4402–4412, Nov. 2013.

[4] M. Tariq, M. S. Latiff, M. Ayaz, Y. Coulibaly, and N. Al-Areqi, "Distance based reliable and energy efficient (DREE) routing protocol for underwater acoustic sensor networks," *J. Netw.*, vol. 10, pp. 311–321, 2015.

[5] A. Wahid, S. Lee, and D. Kim, "A reliable and energy-efficient routing protocol for underwater wireless sensor networks," *Int. J. Commun. Syst.*, vol. 27, pp. 2048–2062, 2014.

[6] N. Javaid *et al.*, "An efficient data-gathering routing protocol for underwater wireless sensor networks," *Sensors*, vol. 15, pp. 29149–29181, 2015.

[7] T. Ali, L. T. Jung, and I. Faye, "End-to-end delay and energy efficient routing protocol for underwater wireless sensor networks," *Wireless Pers. Commun.*, vol. 79, pp. 339–361, 2014.

[8] M. Faheem, M. Z. Abbas, G. Tuna, and V. C. Gungor, "EDHRP: Energy efficient event driven hybrid routing protocol for densely deployed wireless sensor networks," *J. Netw. Comput. Appl.*, vol. 58, pp. 309–326, 2015.

[9] S. Bandita and K. P. Mohan, "Energy efficient SNR based clustering in underwater sensor network with data encryption," in *Proc. Int. Conf. Distrib. Comput. Internet Technol.*, 2016, pp. 137–141.

[10] A. Majid *et al.*, "An energy efficient and balanced energy consumption cluster based routing protocol for underwater wireless sensor networks," in *Proc. IEEE 30th Int. Conf. Adv. Inf. Netw. Appl.*, 2016, pp. 324–333.

[11] I. Azam *et al.*, "SEEC: Sparsity-aware energy efficient clustering protocol for underwater wireless sensor networks," in *Proc. IEEE 30th Int. Conf. Adv. Inf. Netw. Appl.*, 2016, pp. 352–361.

[12] P. Carroll, K. Mahmood, Z. Shengli, Z. Hao, X. Xiaoka, and C. Jun-Hong, "On-demand asynchronous localization for underwater sensor networks," *IEEE Trans. Signal Process.*, vol. 62, no. 13, pp. 3337–3348, Jul. 2014.

[13] A. Sadollah, H. Eskandar, D. G. Yoo, and J. H. Kim, "Approximate solving of nonlinear ordinary differential equations using least square weight function and metaheuristic algorithms," *Eng. Appl. Artif. Intell.*, vol. 40, pp. 117–132, 2015.

[14] H. M. Pandey, A. Shukla, A. Chaudhary, and D. Mehrotra, "Evaluation of genetic algorithm's selection methods," in *Information Systems Design and Intelligent Applications*. New York, NY, USA: Springer, 2016, pp. 731–738.

[15] Z. Michalewicz, "A survey of constraint handling techniques in evolutionary computation methods," *Evol. Program.*, vol. 4, pp. 135–155, 1995.

[16] M. Felemban and E. Felemban, "Energy-delay tradeoffs for underwater acoustic sensor networks," in *Proc. 1st Int. Black Sea Conf. Commun. Netw.*, 2013, pp. 45–49.

[17] R. Nelson, *Probability, Stochastic Processes, and Queueing Theory: The Mathematics of Computer Performance Modeling*. New York, NY, USA: Springer-Verlag, 1995.

[18] Jan. 3, 2016. [Online]. Available: https://www.mathworks.com/products/matlab

[19] G. S. Fishman and L. S. Yarberry, "An implementation of the batch means method," *INFORMS J. Comput.*, vol. 9, no. 3, pp. 296–310, 1997.

Muhammad Faheem, photograph and biography not available at the time of publication.

Gurkan Tuna, photograph and biography not available at the time of publication.

Vehbi Cagri Gungor, photograph and biography not available at the time of publication.