Reinforcement Learning-Based Connectivity Restoration in an Ocean Network of Fishing Vessels

Simi Surendran*, Alberto Montresor‡, Maneesha Vinodini Ramesh†, Paolo Casari‡
*Department of Computer Science and Engineering, Amrita Vishwa Vidyapeetham, Amritapuri, India
†Amrita Center for Wireless Networks & Applications, Amrita Vishwa Vidyapeetham, Amritapuri, India
‡Department of Information Engineering and Computer Science, University of Trento, via Sommarive 9, 35131 Trento, Italy

***Abstract*—An Offshore Communication Network (OCN) is a network of fishing vessels at sea aimed at providing wireless Internet access over the ocean. The connectivity of fishing vessels is essential to disseminate messages, monitor emergency management, and provide information services. The impact of extreme weather conditions on wireless signals, the inability to deploy additional infrastructure, the movements induced by sea waves, the expanded mobility freedom at sea, and the non-uniform density of nodes create connectivity holes in an OCN.**

This paper proposes a reinforcement learning-based connectivity restoration scheme for OCNs. During the planning phase, a node examines the history of its contacts with other nodes and estimates their mobility vector to determine the expected location. Mobile nodes discover the best spots to re-establish connectivity and the most appropriate path to reach these spots via a trial-and-error strategy. In a reinforcement learning framework, we simulate actions to move toward the expected contact locations and learn the optimal movement directions without guessing the contact's actual position. During the control phase, these learned policies are utilized to relocate isolated OCN nodes and restore high-quality connectivity. Simulation results show that our scheme improves the nodes' connectivity probability.

I. Introduction

One of the critical challenges in deep-sea fishing is the lack of low-cost offshore communication solutions capable of working over long distances from the shore. Existing conventional communication technologies like cellular networks and marine radio only provide offshore connectivity up to 20 km. However, fishing operations typically spread far beyond 100 km from the shore. Although satellite telephone systems provide communication facilities everywhere, these are very costly and thus impractical. Due to the lack of low-cost and real-time communication facilities, fishermen cannot connect with the external world even in emergencies. Rao *et al.* proposed the concept of an ad hoc *offshore communication network* (OCN) for fishing vessels to provide Internet access over the ocean [1]. Through an OCN, smartphones on the vessels can achieve relayed access to the Internet via the onboard access network.

Compared to vehicular ad hoc networks, whose mobility patterns are constrained by the road infrastructure, OCN nodes have broader freedom of movement at sea. On the other hand, building an OCN presents unique challenges, such as the sea-wave-induced mobility of fishing vessels, the impact of extreme weather conditions on wireless signals, the inability to deploy additional infrastructure, and the misalignment of directional antennas. Even when an ad hoc network has been established, nodes may experience unpredictable movements due to sea waves. The topology may change rapidly due to the antenna orientation, the rocking movement of vessels, and the propagation effects leading to abrupt changes in the link quality. Therefore, real-time connectivity maintenance is crucial to ensure adequate network connectivity.

One possible solution to restore connectivity is movement-based location optimization without affecting the actual task of the nodes. The nodes need previous contact details to identify connectivity locations when recovering after a disconnection. There is no guarantee that the node will be connected when it reaches the point where connectivity was expected. In this paper, we use reinforcement learning to learn the path to be followed by the vessel to connect to the base station after a disconnection. The algorithm includes two phases: planning and reorientation. A Q-learning approach is applied to learn paths without actual movement in the planning phase. The nodes are positioned virtually to learn the path followed for expected connectivity and compute the post-movement reconnection probability. Based on this probability, the node performs the reorientation phase.

The rest of the paper is organized as follows: Section II reviews previous works related to movement based connectivity restoration. In Section III, we present the architecture of an OCN. Section IV describes the reinforcement learning model for position reorientation. Section V presents simulation results, followed by the concluding remarks of Section VI.

II. Related Work

Wireless networks are prone to frequent connectivity failures due to the nature of the environment in which they operate. To improve the quality of communications among nodes, methods to restore and improve the connectivity have been studied [2]–[4]. For example, nodes can be deployed to establish k-connected networks [2], [5], [6]; such methods, however, are more suitable for static wireless networks. Another approach is to use additional infrastructure such as unmanned areal vehicles and relay nodes to restore or maintain connectivity [3], [7], [8].

To maintain connectivity, it is also possible to reposition special nodes, called actor nodes [9], [10]. In this case, the objective is to reconnect the network with a minimum number of relay nodes. These methods assume that the nodes are stationary and additional nodes are always available to deploy. Transmit power adjustments and topology control can also ensure strong connectivity [4], [11], both locally and globally [12], [13].

A different approach performs movement-based connectivity restoration. For example, Basu and Redi propose to rearrange the network topology by moving existing nodes in a one-dimensional robotic network unfortunately requiring centralized control to avoid connectivity holes [10]. Abbasi *et al.* discussed a method to restore connectivity failure due to a single node through relocation towards a neighbor with the lowest node degree [14]. However, this approach can not tolerate multiple node failures simultaneously.

Heuristic algorithms were used to reconnect the networks based on the total distance traveled [15]. This technique requires identifying the relay points in advance and then filling connectivity holes based on a greedy heuristic. Kim *et al.* proposed a neighborhood-aware restoration algorithm in a drone network, where each node moves to its last connected neighbor location [16]. Nevertheless, this method does not guarantee connectivity because the node may not be available at the last known location. Instead of waiting until disconnection, proactive approaches were also applied to prepare backup plans in advance [17], but typically require a much more significant coordination overhead.

Most of the node (re)deployment strategies proposed for terrestrial wireless network connectivity restoration cannot be employed for OCNs. Since additional infrastructure or relay nodes are unavailable to fill connectivity holes in an OCN, role-based connectivity restoration is also not applicable. Moreover, movement control algorithms may not always be appropriate because they depend on the node's fishing tasks and communication context. Therefore, OCNs require an adaptive node position relocation scheme to restore connectivity. Our work in the following proceeds along this line.

III. Architecture of OCN

The *offshore communication network* (OCN) is a vehicular network that provides Internet access to fishing vessels over the ocean [1], [18]–[20]. The network is designed to extend the wireless connectivity range over 100 km offshore and uses a distributed architecture integrated with edge computing. Fishing vessels are considered edge nodes in this architecture that processes the data collected locally, avoiding dependency on the base station for analysis.

Currently, fishermen resort to limited hand-held radios for communications, barely reliable in rough sea conditions. Cellular network coverage is restricted to 15 km from the shore. Conversely, OCN enables fishermen to access applications such as WhatsApp and communicate directly through their smartphones. The onboard Internet access facilitates the use of information services provided by the Government to the fishermen at sea and emergency communication to the shore. Fig. 1 shows the architecture.

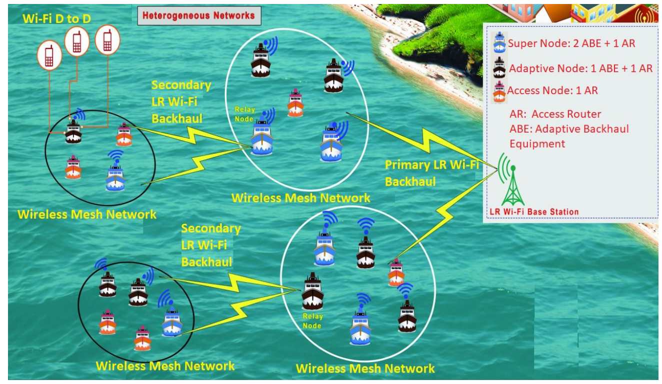

Fig. 1. Architecture of an Offshore Communication Network.

OCN edge nodes are grouped into three categories: access nodes, adaptive nodes, and super nodes, depending on the communication resources available in the fishing vessels. *Access nodes* are vessels that contains only a wireless *access router* (AR); *adaptive nodes* contain *adaptive back-haul equipment* (ABE), while *super nodes* equipped with ABEs and one AR. Adaptive and super nodes are also called *long-range* (LR). Each AR has an omnidirectional antenna to provide WiFi access up to 500 m, connecting devices such as smartphones and other nearby ARs. The ABEs are provided with 120° sector antennas that provide connectivity of nearly 20 km, using long-range WiFi links.

The architecture is partitioned into three layers:

- *Layer 0* provides Wi-Fi with a mesh network of access nodes;
- *Layer 1* provides LR Wi-Fi with an ad hoc backbone network of LR nodes;
- *Layer 2* is the network of base stations on the shore.

A test implementation of the OCN architecture has been evaluated over the Arabian Sea from a coastal village in Kerala, India. In the field tests, the network provided a 50+ km range in the first hop and 20+ km in every succeeding hop.

IV. Reinforcement Learning-based Connectivity Restoration

Maintaining connectivity between nodes is a fundamental problem since OCNs operates over a large geographical area, and the network is sparse. In case of disconnection, recommendations on the direction and distance to relocate for better connectivity will immensely benefit fishermen at sea.

A. Communication Requirements

The communication requirements in each fishing state are identified based on the interviews conducted with over a hundred fishers. Although the ability to make audio calls is essential, fishermen also use video calls and entertainment facilities depending on the fishing phase. According to the fishing tasks, we can divide the states of fishing into four groups: sailing, searching, fishing, and resting. Fig. 2 summarizes the mobility and required communication level in each state.

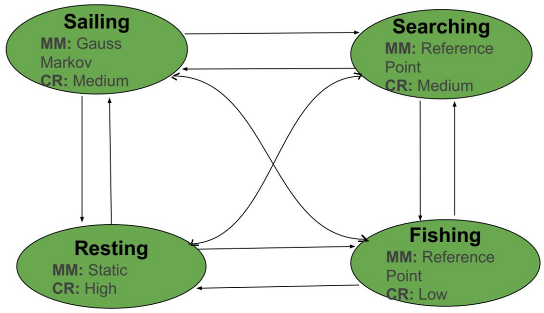

Fig. 2. OCN fishing states and communication requirements. MM: Mobility Model, CR: Communication Requirements. (From [21].)

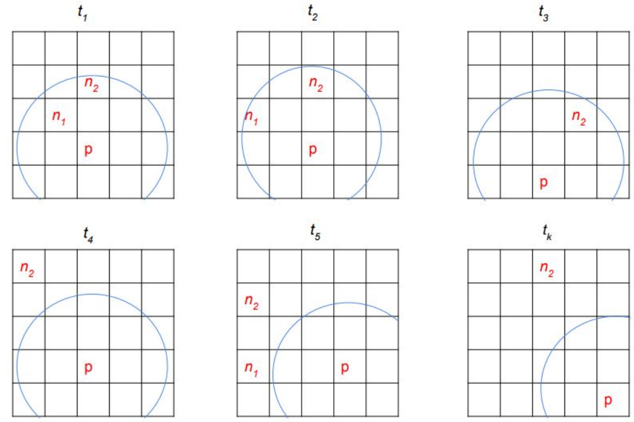

Fig. 3. Sample scenario for connectivity restoration after a disconnection.

The *sailing state* includes the journey towards the high sea and the return to the shore. In this case, the communication is predominantly with peer groups, and can be classified it as a medium requirement. However, the base station connectivity is occasionally required. When the required sea depth is reached, the vessels start searching the fishing zones, and the state is called *searching state*. This stage focuses more on discovering potential fishing zones. The connectivity requirement is medium, as more communication occurs merely to the peer nodes. Since the fishermen work as a group to discover the fishing zones, the vessels are seen in clusters, and the cluster members follow the path of the cluster head. Nodes can adjust their positions within the cluster to obtain more reliable connectivity. Once a fishing zone has been found, the state changes to *fishing state*. The communication requests are very low in this state since the fishermen engaged totally in the fishing task as a group. The mobility is also limited. During the night, the vessels stop fishing and move to *resting state*. Since mobility is significantly less, this stage is considered a static state, and all others are dynamic. Base station connectivity is more demanding here as they mostly communicate with the shore or use mobile entertainment applications. In a few cases, peer connectivity is also required. Hence, the requirement is categorized as high.

B. Formulation of Position Reorientation Model

Connectivity restoration after disconnection in an OCN is formulated as a Markov Decision Process (MDP). The MDP is represented using 5-tuples $\langle S, A, T, R, \gamma \rangle$, where S is the set of environmental states, A is the distinct set of agent actions, T is the state transition function $T : S \times A \times S \rightarrow [0, 1]$, R is the reward function of the agent $R : S \times A \rightarrow \mathbb{R}$ and γ is the discount factor. To solve the MDP, we apply reinforcement learning.

Reinforcement Learning (RL) is a machine learning technique that learns a system by interacting with the environment via trial-and-error [22]. The learner in an RL system is termed as an agent. The agent constantly interacts with the environment by selecting arbitrary actions. The environment returns feedback on how good the selection is for each action. At each stage t, the agent selects an action that transforms the state of the environment from s_t to s_{t+1} and gains a reward. It is incapable of delivering the best performance at the initial learning phase, since the environment is unknown. It explores possible actions to achieve the goal and examines the impact of those actions. After the initial phase, the agent operates on the learned parameters and increases the expected reward.

Fig. 3 illustrates a sample connectivity restoration scenario of an OCN. Assume that the ocean environment can be viewed as a grid where each cell has size 1 km$\times$1 km. The position information of all nodes in an OCN is available via GPS. Nodes communicate with each other either through long-range (LR) WiFi or standard WiFi links. Link quality is classified into five levels, from 1 to 5, depending on the received signal strength. The thresholds for the link classification are derived from data collected from field trials performed over the Arabian Sea. Fig. 4 shows the signal strength collected from an LR link between the base station and a fishing vessel. From the analysis of our signal strength vs. distance measurement dataset, we divided the link quality into the above five grades, and determine the corresponding distances as shown in Fig. 4.

When two nodes come in contact, they exchange their current position and movement direction. Fig. 3 shows the locations of three nodes: p, n_1 and n_2 at time epochs t_1 to t_k. During the time interval from t_1 to t_3, node p is connected to both n_1 and n_2 at time t_1 and only n_2 at time t_2. After t_3, p lost connectivity and wants to reconnect. Therefore node p will look up the history of previously connected nodes. The history includes the locations of last contact nodes, last contact time, and contact mobility vector with speed and direction of motion. Here, node p has two previous contacts n_1 and n_2. Node p computes the expected location of the contact n_i as a function of time and average velocity as

$$\hat{\mathbf{x}}_{n_i}(t_k) = \alpha\, \bar{\mathbf{v}}_{n_i}(t_k - t_c) + \mathbf{x}_{n_i}(t_c), \tag{1}$$

where t_k is the time at which node p wants to reconnect, t_c is the last contact time with neighbor and $\bar{\mathbf{v}}_{n_i}$ is the average velocity of neighbor i. A decision making scenario of node p is shown in Fig. 5. It can select the direction of previous contacts n_1 and n_2. The moving node must follow a sequence of grid positions to reach a well-connected area. This sequential

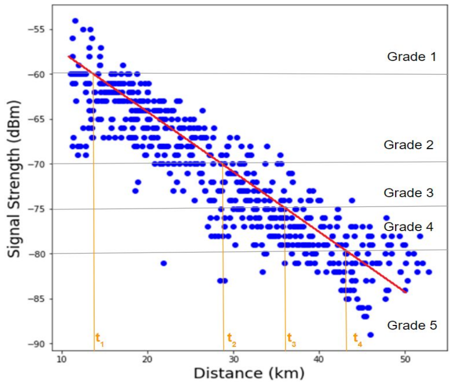

Fig. 4. Signal strength vs. distance data collected from sea trial experiments and the threshold selection for grading link quality.

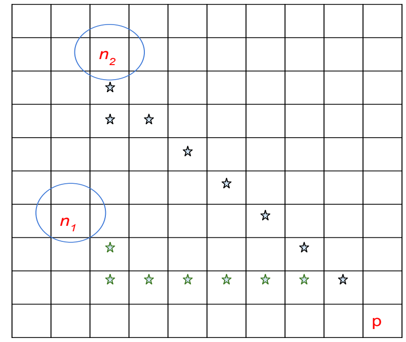

Fig. 5. Decision making context.

decision-making problem is represented using an MDP with the following state, action, and reward function.

States—Nodes store the history of previous contacts, including node connectivity quality, location, and contact time. The state of the agent is defined as $(\hat{\mathbf{x}}_{n_i}, \Delta t_i\})$ where $\hat{\mathbf{x}}_{n_i}$ denote the location of the connected contact i in a threshold time interval. Δt_i is the difference between the time connectivity restoration is initiated and time of the last contact with node i.

Actions—An action is defined as choosing one of the grid cells (i, j) to move from the current position. Possible actions from the current grid include moving to north, south, east, west, north-east, north-west, south-east, and south-west. Also, the node can move in the direction of the base station or to one of the available potential fishing zones (PFZ), where more vessels are expected. Hence the action space $A = \{N, S, E, W, NE, NW, SE, SW, BS, PFZ_k\}$.

Reward function—The environment returns a reward to the agent conveying the impact of its action. The agent's main objective is to achieve a connection to the base station (BS). The reward is assigned based on the quality of the radio link between the moving node and the BS, $\mathcal{Q}_{\text{com}}(p, BS)$, which maps to the probability of actually establishing a connection successfully (P_{conn}) as follows:

$$R_1 = \begin{cases} +50, & \text{if } \mathcal{Q}_{\text{com}}(p, BS) = \text{grade 1 } (P_{\text{conn}} = 1), \\ +30, & \text{if } \mathcal{Q}_{\text{com}}(p, BS) = \text{grade 2 } (P_{\text{conn}} = 0.8), \\ +20, & \text{if } \mathcal{Q}_{\text{com}}(p, BS) = \text{grade 3 } (P_{\text{conn}} = 0.7), \\ +10, & \text{if } \mathcal{Q}_{\text{com}}(p, BS) = \text{grade 4 } (P_{\text{conn}} = 0.5), \\ +5, & \text{if } \mathcal{Q}_{\text{com}}(p, BS) = \text{grade 5 } (P_{\text{conn}} = 0.4), \\ -1, & \text{otherwise } (P_{\text{conn}} = 0). \end{cases} \tag{2}$$

Based on data collected from sea experiments, we analyzed the connectivity quality for each signal strength range and designed the reward function. Good-quality communications are rewarded with high positive values. For grade 1 communication, we give a high positive reward of +50 in the simulations. A grade 2 quality radio link is still good but its signal strength is less stable, whence a reward of +30. Similarly, we proceed down to a +5 reward for communication grade 5. For all other communication grades where the quality of the signals was poor, a negative reward was used as a penalty. This reward range is chosen considering the total cumulative reward obtained in each simulation episode. Note that the BS connectivity can be restored via either a single-hop or a multi-hop link. In the multi-hop case, one or more contact nodes act as relays to and from the BS when the moving node reaches their communication range, and the connection quality is the lowest quality of all links along the multi-hop path. For instance, with two relays n_1 and n_2, we have $\mathcal{Q}_{\text{com}}(p, BS) = \max\{\mathcal{Q}_{\text{com}}(p, n_1), \mathcal{Q}_{\text{com}}(n_1, n_2), \mathcal{Q}_{\text{com}}(n_2, BS)\}$, and the resulting grade maps to its corresponding connection probability as in Eq. (2) above.

Let the initial location of node p be $\mathbf{x}_{\text{init}}$, the desired location to get connectivity be $\mathbf{x}_{\text{tgt}}$ and location in next time step be $\mathbf{x}_{\text{next}}$. Based on the signal availability of the next location and travel distance minimization constraints, we can define reward as

$$\mathcal{R}_t = \mathcal{R}_1 + \frac{d(\mathbf{x}_{\text{init}}, \mathbf{x}_{\text{tgt}}) - d(\mathbf{x}_{\text{next}}, \mathbf{x}_{\text{tgt}})}{d(\mathbf{x}_{\text{init}}, \mathbf{x}_{\text{tgt}})}, \tag{3}$$

where $d(\mathbf{x}_1, \mathbf{x}_2)$ represents distance between $\mathbf{x}_1$ and $\mathbf{x}_2$.

Learning—We apply an off-policy Q-learning algorithm to learn the actions that lead to better connectivity. When a node wants to reconnect to the network, it computes the expected locations of the last contacts using the history. Then the node performs a planning phase by choosing the grid positions to reach the target. In every step, it selects the highest Q-valued action with $1 - \varepsilon$ probability. The action value is updated as

$$\mathcal{Q}(s_t, a_t) \leftarrow (1-\alpha)\mathcal{Q}(s_t, a_t) + \alpha[\mathcal{R}_t + \gamma \max_a \mathcal{Q}(s_{t+1}, a)] \tag{4}$$

Algorithm 1 details the steps of our relocation scheme.

V. Results

We have simulated the ocean environment by modeling it as a grid of size 200 km×200 km, where each cell in the grid has size 1 km×1 km. When a node needs to relocate, it performs a planning phase to learn the Q-table. In the actual reorientation, the node uses this learned Q-table to reach locations yielding better connectivity. We consider a scenario with 12 LR nodes and a single base station at the shore for the learning phase. The convergence of the learning phase is plotted in Fig. 6, where the algorithm converges after 2000 iterations. After this

Algorithm 1: Adaptive Connectivity Restoration

Initialization
State $S = \{(\hat{\mathbf{x}}_{n_i}, \Delta t)\}$
Action $A = \{N, S, E, W, NE, NW, SE, SW, BS, PFZ_k\}$
Exploration coefficient ε = 0.01
if *Node x has to restore connectivity* **then**
  a. $\mathcal{C}_x \leftarrow$ Contacts of x over interval $[t_c, t_k]$
  b. Compute expected node locations using (1)
  c. Compute signal strength from the BS to the expected location of all contacts using the channel model
  d. Define terminal states $\mathcal{T}$ of the MDP based on the communication range of selected contacts
  e. **foreach** $i \in s$, $a \in A$ **do**
    $Q(i, a)$ = 0 ;
  f. **while** *$iterations \leq max\_episodes$* ***or*** *$current\_state \in terminal\_states$* **do**
    Generate a random number $p \in [0, 1]$
    if $\varepsilon \leq p$ **then**
      $\mathbf{x}_{\text{next}}$ = random grid position
    else
      $\mathbf{x}_{\text{next}}$ = grid with highest Q values
    Update Q value using (4)

planning phase, the agent takes steps towards the best contact locations using the learned Q-table. Fig. 7 shows a sample of the path that the agent follows to reach a contact's expected location over subsequent learning steps.

We test two main metrics: the probability that the reorienting node can connect to a contact along its path towards the reorienting destination, and the probability that this actually translates into a (possibly multi-hop) connection to the base station. We consider five test cases for our Q-learning-based framework, as listed in Table I. Furthermore, we compare against two baseline schemes: (i) the node reorients to a random location in the map, or (ii) the node reorients towards one of its previous contacts, chosen at random. These cases are labeled 6 and 7. The probability that the reorienting node successfully connects to a contact as a result of its own movement is shown in Fig. 8. When the contact nodes follow the same trajectory as in the planning phase (Case 1), the relocating node's connectivity probability to the contact and the base station is in the range $[0.975, 0.99]$. This result confirms that the node learns where to relocate. Moreover, it provides a realistic baseline for the case where several vessels tend to move around the same fishing area, as their trajectories would typically approach known potential fishing zones.

After the learning phase, we evaluated the connectivity probability with six unseen contacts. In this case, the connectivity probability to the contact and the base station is in the range $[0.75, 0.85]$. This confirms that the learning framework allows a node to restore its connection with high probability by moving to a well-chosen location. Fig. 9 shows the connectivity probability of the reorienting node to the base station. We observe that the results have a slightly higher statistical dispersion than those related to the contact nodes in Fig. 8. This is because the quality of the (typically multi-hop) connection from the reorienting node to the BS is bounded by

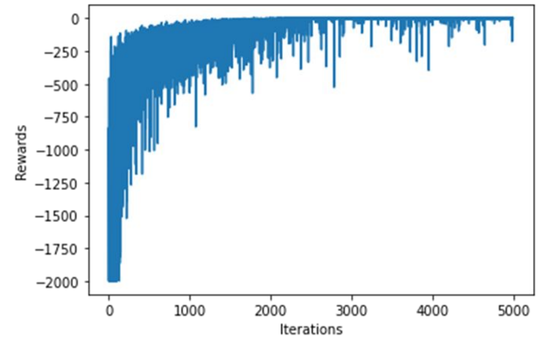

Fig. 6. Reward acquired vs number of iterations.

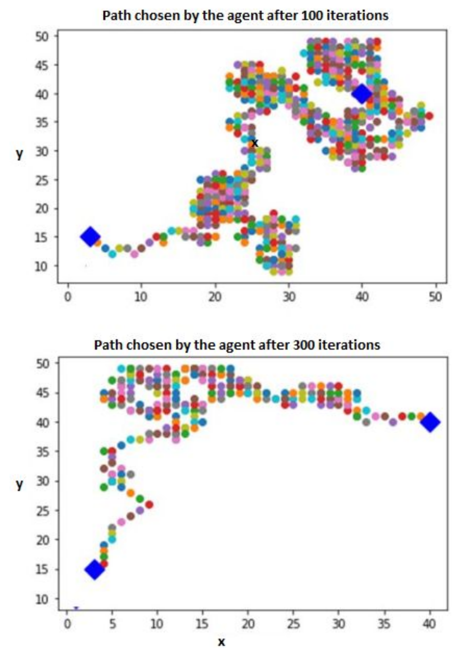

Fig. 7. Path followed by the agent at different learning steps from a start location to the expected location of a contact.

TABLE I
SIMULATION SCENARIOS

Scenario	Learning phase	Testing phase
Case 1	12 nodes	6 nodes from training phase
Case 2	12 nodes	6 new nodes with different moving patterns
Case 3	12 nodes	12 new nodes with different moving patterns
Case 4	18 nodes	6 new nodes with different moving patterns
Case 5	18 nodes	12 new nodes with different moving patterns
Case 6	—	12 nodes in the topology and the node selects a random target location
Case 7	—	12 nodes in the topology and the node selects a random neighbor from its previous contacts

the lowest-quality link, and the corresponding probability of successfully establishing a connection behaves as in Eq. 2.

To check the impact of increasing the number of nodes in the test scenario during reorientation, we placed 12 unseen nodes in the environment. Due to the increase in density of the contact nodes, the relocating node gets better connectivity to the contact and base station, as seen in case 3 of Fig. 8 and 9. In the following test scenario, case 4,

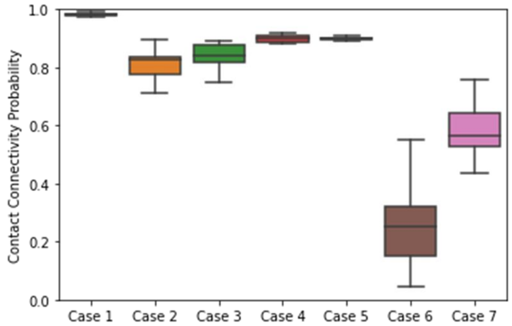

Fig. 8. Connectivity probability between the reorienting node and the contact in different test scenarios.

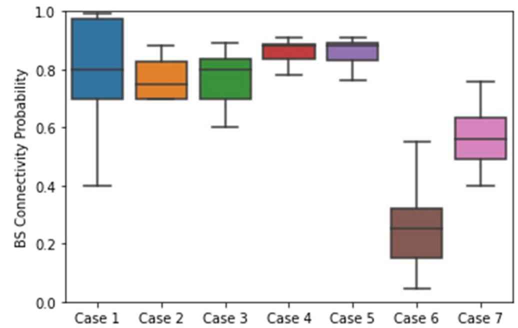

Fig. 9. Connectivity probability between the reorienting node and the base station in different test scenarios.

we increased the number of nodes in the training phase to 18 and observed better connectivity than in case 3. When a node chooses a random reorienting direction, the connectivity probability is low compared to the other cases discussed in Table I. After disconnection, a node is unaware of the location where reconnection is possible. Moving to a random location without knowing the current topology reduced the connection probability. However, moving towards any one of the previously connected contacts increased reconnection probability compared to a random reorientation. Instead, RL-based relocation tries possible paths in the learning phase, and those learned paths are utilized in actual movement. This path planning improved the connectivity probability in RL-based relocation. Moreover, the results in Fig. 9 indicate that the connectivity probability is improved with the increase in node density in the reorientation scenario.

VI. Conclusions

The connectivity of fishing vessels is critical for information services and message dissemination among fishing vessels at sea. While an offshore communication network enables long-range connectivity for fishing vessels, OCN nodes tend to undergo abrupt connectivity failures due to the extreme environment in which the network operates.

In this paper, we proposed a reinforcement learning algorithm to restore connectivity in an OCN after a disconnection. In the first planning phase of the algorithm, a node utilizes its history of contacts to estimate the locations expected to offer the best connectivity, and computes the most suitable path to reach these zones. The paths learned from the planning phase are used in the second phase of the algorithm to reach potential connectivity spots. Our simulation results show that the connectivity probability improves by using our RL-based approach, compared a random search of the best location.

References

[1] S. N. Rao *et al.*, "Mobile infrastructure for coastal region offshore communications and networks," in *Proc. IEEE GHTC*, 2016.

[2] B. Gupta and A. Gupta, "On the k-connectivity of ad-hoc wireless networks," in *Proc. IEEE SOSE*, 2013, pp. 546–550.

[3] S.-Y. Park *et al.*, "Dronenetx: Network reconstruction through connectivity probing and relay deployment by multiple uavs in ad hoc networks," *IEEE Trans. Veh. Technol.*, vol. 67, no. 11, pp. 11 192–11 207, 2018.

[4] Z. Nutov, "Approximating minimum-power k-connectivity." *Ad hoc & Sensor Wireless Networks*, vol. 9, 2010.

[5] M. Segal and H. Shpungin, "On construction of minimum energy k-fault resistant topologies," *Ad Hoc Networks*, vol. 7, no. 2, pp. 363–373, 2009.

[6] F. Deniz *et al.*, "An adaptive, energy-aware and distributed fault-tolerant topology-control algorithm for heterogeneous wireless sensor networks," *Ad Hoc Networks*, vol. 44, pp. 104–117, 2016.

[7] R. Magán-Carrión *et al.*, "A dynamical relay node placement solution for MANETs," *Computer Communications*, vol. 114, pp. 36–50, 2017.

[8] L. Liu *et al.*, "Optimal relay node placement and flow allocation in underwater acoustic sensor networks," *IEEE Trans. Commun.*, vol. 65, no. 5, pp. 2141–2152, 2017.

[9] R. W. Pazzi and A. Boukerche, "Mobile data collector strategy for delay-sensitive applications over wireless sensor networks," *Computer Communications*, vol. 31, no. 5, pp. 1028–1039, 2008.

[10] P. Basu and J. Redi, "Movement control algorithms for realization of fault-tolerant ad hoc robot networks," *IEEE Netw.*, vol. 18, no. 4, pp. 36–44, 2004.

[11] X. Han *et al.*, "Fault-tolerant relay node placement in heterogeneous wireless sensor networks," *IEEE Trans. Mobile Comput.*, vol. 9, no. 5, pp. 643–656, 2009.

[12] V. K. Akram and O. Dagdeviren, "On hardness of connectivity maintenance problem in drone networks," in *Proc. IEEE BlackSeaCom*, 2018.

[13] D. Cai *et al.*, "Distributed global connectivity maintenance and control of multi-robot networks," *IEEE Access*, vol. 5, pp. 9398–9414, 2017.

[14] A. A. Abbasi *et al.*, "Movement-assisted connectivity restoration in wireless sensor and actor networks," *IEEE Trans. Parallel Distrib. Syst.*, vol. 20, no. 9, pp. 1366–1379, 2008.

[15] I. F. Senturk *et al.*, "An effective and scalable connectivity restoration heuristic for mobile sensor/actor networks," in *Proc. IEEE GLOBECOM*, 2012, pp. 518–523.

[16] G.-H. Kim *et al.*, "Self-recovery scheme using neighbor information for multi-drone ad hoc networks," in *Proc. APCC*, 2017, pp. 1–5.

[17] M. Imran *et al.*, "Resource efficient connectivity restoration algorithm for mobile sensor/actor networks," *EURASIP Journal on Wireless Communications and Networking*, vol. 2012, no. 1, pp. 1–16, 2012.

[18] S. N. Rao *et al.*, "Realizing cost-effective marine internet for fishermen," in *14th Int. Symposium on Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks*. IEEE, 2016, pp. 1–5.

[19] J. Dhivvya *et al.*, "Towards maximizing throughput and coverage of a novel heterogeneous maritime communication network," in *Proc. of the 18th Int. Symposium on Mobile Ad Hoc Networking and Computing*. ACM, 2017, p. 39.

[20] S. Simi and M. V. Ramesh, "Intelligence in wireless network routing through reinforcement learning," *International Journal of Communication Networks and Distributed Systems*, vol. 23, no. 2, pp. 231–251, 2019.

[21] S. Surendran *et al.*, "Modelling communication capability and node reorientation in offshore communication network," *Computers & Electrical Engineering*, vol. 87, p. 106781, 2020.

[22] R. S. Sutton and A. G. Barto, *Reinforcement learning: An introduction*. MIT press, 2018.