

수중 센서 네트워크에서 강화학습 기반의 라우팅 프로토콜

김희원*, 조호신°

Reinforcement Learning-Based Routing Protocol in Underwater Sensor Networks

Hee-won Kim*, Ho-Shin Cho°

요약

본 논문은 수중 환경에서 네트워크 토폴로지를 고려한 강화학습 기반 라우팅 프로토콜을 제안한다. 센서노드들은 차기 전달노드를 선택할 때, 네트워크 토폴로지를 고려한 가치값을 이용하여 목적지 반대방향으로의 패킷 전달을 회피하고 자원낭비를 줄일 수 있다. 시뮬레이션 결과, 지연 및 배터리 소모 측면에서 기존의 QELAR에 비해 우수한 성능을 보인다.

Key Words : Underwater Sensor Networks, Routing, Q-learning, Latency, Energy consumption

ABSTRACT

This paper proposes a reinforcement learning based routing protocol to consider network topology in underwater sensor networks. Utilizing the reward value reflecting network topology, sensor node selects the next-forwarding node and thereby escaping the opposite direction routing to the sink and accordingly reducing the waste of resource. Computer simulation result shows that the proposed scheme outperforms QELAR in terms of latency and battery consumption.

I. 서론

데이터 수집을 목적으로 하는 수중센서 네트워크에서 보다 효율적인 다중 홉 전달을 위해서 강화학습 기반의 라우팅 프로토콜이 제안되었다^[1]. 센서 노드들은 라우팅 경로 상의 차기 노드를 선택하고 그 선택에 대한 보상(reward)을 받게 되며 최종 목적지인 싱크까지의 보상값의 총합인 가치값(value)이 최대가 되도록 경로를 설정하게 된다. 기존기법에서는 노드의 잔여 배터리 량, 지연시간 등을 고려하여 보상이 결정된다. 이 경우, 그림 1 (LV는 계층레벨로서 싱크노드까지의 홉수를 나타냄)에서 보듯이 노드 4번이 싱크노드와 반대 방향에 위치한 노드 5번을 차기노드로 선택하는 경우가 발생할 수 있다. 이는 최종 목적지인 싱크노드에 근접할수록 배터리가 소모가 심한 현상, 즉 깔때기 현상(funnel effect)^[2]으로 인해 싱크노드로부터 먼 하위 레벨 이웃노드들이 차기노드로 선호될 수 있기 때문이다.

본 논문에서는 네트워크 토폴로지를 고려하여 싱크노드에 가까운 상위레벨의 노드들이 높은 가치값을 가지도록 하여 반대방향으로의 패킷 전달을 회피하고 지연 및 배터리 소모를 줄일 수 있는 강화학습(Q-learning)기반 수중 라우팅 프로토콜을 제안한다.

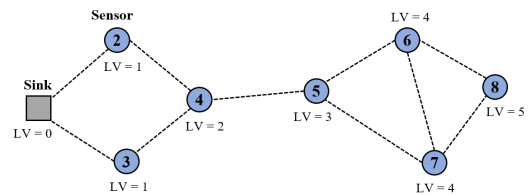


그림 1. 센서 네트워크 토폴로지 예시
 Fig. 1. Example of sensor network topology

II. 제안 기법

각 센서노드는 네트워크 초기화 시 자신의 식별자(ID)와 계층 레벨(LV)을 안다고 가정하며, 이후 네트워크 토폴로지 변화에 따라 자신의 LV뿐만 아니라 이웃노드들의 ID, LV, 차기노드 후보군, 배터리 잔여량을 수시로 갱신한다. 이러한 이웃정보는 엇들게 되는

※ 본 연구는 2020년 해양수산부 재원으로 한국해양과학기술진흥원의 지원을 받아 수행되었습니다(수중 관측제어망용 네트워크 개발).
 • First Author : Samsung Electronics, hwkim028@gmail.com, 정희원
 ° Corresponding Author : (ORCID:0000-0002-6949-0904)School of Electronics Engineering, Kyungpook National University, hscho@ee.knu.ac.kr, 정교수, 종신회원
 논문번호 : 202007-147-B-LU, Received July 6, 2020; Revised July 20, 2020; Accepted July 27, 2020

패킷 헤더들을 참조함으로써 획득된다. 차기노드 선택 시 이웃노드 정보가 바탕이 된다. 패킷 전달이 성공할 때마다 특정량의 보상이 주어지며, 이는 차기노드의 배터리 잔량과 패킷 전달이 성공할 때까지의 전송 횟수에 기반한다.

2.1 차기노드 후보군 탐색

$P(X)$ 와 $P^*(X)$ 를 각각 노드 X 의 이웃노드군과 차기노드 후보군이라 하자. 노드 X 는 데이터 발생 시 먼저 $P^*(X)$ 를 생성한다. $P^*(X)$ 에 복수의 노드들이 존재하면 가장 높은 가치값을 지닌 노드가 차기노드로 선택된다. 노드 Y 가 $P^*(X)$ 에 포함되기 위해서는 다음 세 조건들을 모두 만족해야 한다.

- (a) $Y \in P(X)$
- (b) $X \notin P^*(Y)$
- (c) $P(X) \cap P^*(Y) = \emptyset$

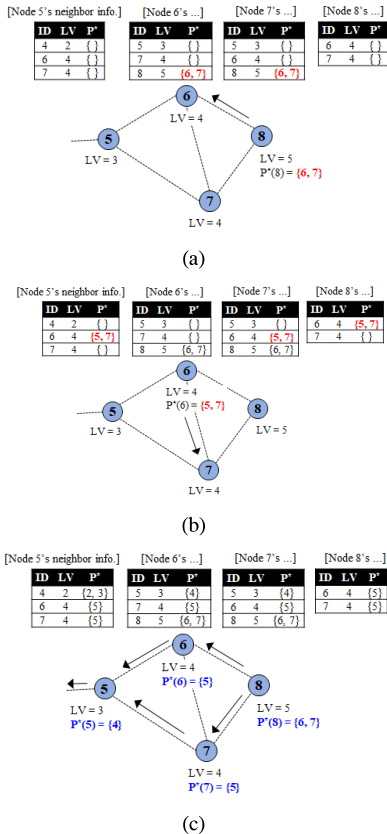


그림 2. 차기노드 탐색 과정
Fig. 2. Searching procedure for next-forwarder

세 조건들을 종합하면, 노드 Y 가 노드 X 에게 싱크 노드 방향으로의 경로를 제공함을 뜻한다. 그림 2는 차기노드 후보군 탐색 과정에 대한 예시이다. 처음에는 모든 노드들이 이웃노드의 P 정보를 가지고 있지 않다고 가정한다. 그림 2a에서, 노드 8번이 데이터를 획득하면 노드 6번/7번이 위의 세 조건들을 만족함을 안다. 즉, $P^*(8) = \{6,7\}$ 이다. 아직 노드 6번/7번의 가치값을 모르므로, 노드 8번은 임의로 노드 6번을 차기노드로 선택한다. 그림 2b에서, 노드 6번은 8번으로부터 데이터를 받은 후 (그 과정에서 데이터 헤더에 기록된 $P^*(8) = \{6,7\}$ 정보 획득) 노드 5번/7번/8번 중 노드 5번/7번이 위 세 조건들을 만족함을 안다. 즉, $P^*(6) = \{5,7\}$ 이다. 노드 8번이 이미 차기노드 후보군에 노드 6번을 포함하였으므로 ($P^*(8) = \{6,7\}$), 조건 b에 의해 노드 8번은 $P^*(6)$ 에서 제외된다. 노드 7번은 6번에 의해 차기노드로 임의 선택된 후, 같은 방법으로 $P^*(7) = \{5\}$ 를 생성한다. 그림 2c와 같이, 차기노드 후보군 탐색은 모든 노드들이 최신 이웃노드 정보를 가질 때까지 갱신된다. 그림 2c에서 노드 6번/7번은 서로에 대해 조건 a/b를 만족하나, 공통된 차기노드를 가지므로 조건 c에 위배됨에 유의한다.

2.2 보상 계산

수신노드의 응답 메시지를 받으면 패킷 전송 성공으로 간주되고, 다음과 같은 보상 r_s 가 주어진다.

$$r_s = \alpha_1 \times R_f + \alpha_2 \times R_b + \alpha_3 \times R_t$$

R_f, R_b, R_t 는 각각 고정 보상, 에너지 보상, 지연 보상을 나타내고, $\alpha_1, \alpha_2, \alpha_3$ 은 각 보상의 가중치이다. 각 보상은 다음과 같이 정의된다.

$$R_f = -1$$

$$R_b = \frac{E_r}{E_i}, \quad 0 \leq R_b \leq 1$$

$$R_t = 1 - \frac{N_t - 1}{N_m - 1}, \quad N_m > 1, \quad 0 \leq R_t \leq 1$$

E_r, E_i, N_t, N_m 은 각각 차기노드의 에너지 잔여량, 초기 에너지, 수신노드의 응답 메시지를 받기까지 패킷 전송 횟수, 최대 패킷 전송 가능 횟수이다. R_f 는 매 패킷 전송 행위에 대한 페널티를 부여하기 위한 것이며, 결국 싱크노드까지의 총 전달 홉 수에 상응한다.

R_b 는 선택된 차기노드의 잔여 에너지에 의존하는데, 잔여 에너지가 많은 이웃노드가 더 많은 보상을 받도록 한다. 채널 상태가 나쁘면 패킷 재전송 횟수가 더 많아지고 지연이 길어지므로, 더 낮은 R_i 가 주어진다. 패킷 전송에 최종 실패하는 경우 다음과 같은 보상 $r_f (= \beta \times R_f)$ 가 주어진다. β 는 가중치 인자이다.

III. 모의실험 결과

표 1은 시뮬레이션 파라미터를 나타낸다. 수중 채널 특성 반응을 위해, 실증적 수중 음향 채널 모델^[3]이 적용되었다. 패킷 전송속도와 송/수신 전력의 경우, 상용 수중모뎀인 Benthos ATM-903의 것이 사용되었다. 간섭, 비트 오류 정정, 유희상태 전력 소모 등은 고려되지 않았다. 각 센서노드는 푸어송 분포를 따르는 패킷 생성률을 가진다. 빠른 강화학습을 위해 학습률 0.9, 미래지향적 그리고 근시안적 학습의 중간 단계로 감소율 0.5를 사용하였다.

총 16개의 고정 노드들이 각 격자점을 중심으로 반경 20% 이내의 임의 위치에 배치된다. 그 중 가장 자리에 있는 특정 노드가 싱크노드가 되며, α_1 과 β 는 1로 고정된다. 제안 기법에서 α_2 의 증가는 하위 레벨 노드들이 차기노드로서 선택될 가능성을 높이는 것을 의미하며, α_3 을 크게 설정하는 것은 센서노드들이 패킷 오류율이 높은 지역 쪽을 회피하도록 패킷을 전달하는 것을 의미한다.

비교 기법인 QELAR는 보상 계산 시 지연 보상을 고려하지 않고, 고정 보상과 에너지 보상만을 고려한다. 특히 에너지 보상의 경우, 노드의 절대적인 에너지 잔여량과 이웃노드들 간 상대적인 에너지 잔여량 분포를 반영한다. 절대적인 에너지 잔여량의 가중치 (γ_1, δ_1)와 상대적인 에너지 잔여량 분포의 가중치 (γ_2, δ_2)^[1]는 본 시뮬레이션 환경에서의 최적값인 0.5와 0.1

표 1. 모의실험 파라미터
Table 1. Simulation parameter

파라미터	값
전송속도	2,400 bps
전파속도	1,500 m/s
전송반경	1,200 m
송신전력/수신전력	20 W / 756 mW
초기에너지	104 J
최대패킷전송회수	5
노드당 트래픽 부하	0.001 packet/s

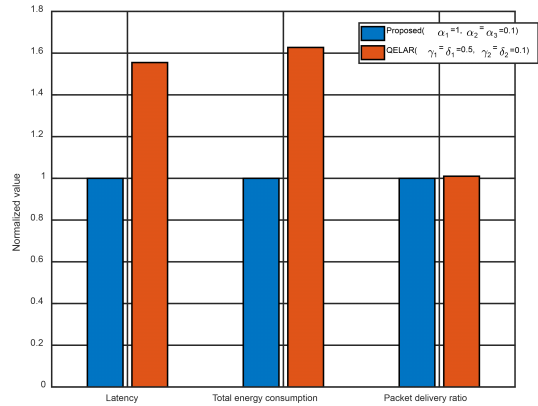


그림 3. 성능비교 결과
Fig. 3. Performance comparison results

로 설정되었다. 그림 3에서 α_2 와 α_3 이 0.1로 설정되었을 때, 제안 기법이 더 낮은 지연과 더 적은 에너지 소비에 기반하여 QELAR보다 더 나은 성능을 보인다. 제안 기법의 성능 이득은 네트워크 토폴로지를 고려하여 차기노드를 선정함으로써 발생한다. 그리고, QELAR와 달리 보상 함수에 포함된 지연 보상이 패킷 오류율이 높은 네트워크 영역을 회피하는 데 기여한다. 패킷 전달률의 경우 두 기법 모두 98% 이상의 높은 수치를 보였다.

IV. 결론

본 논문에서는 수중 센서 네트워크에서 토폴로지를 고려한 강화학습 기반 라우팅 프로토콜이 제안되었다. 센서노드들은 이웃노드들의 LV 및 차기노드 후보군 정보를 기반으로 자신의 차기노드 후보군을 먼저 추려낸 후 그 중 가치값이 가장 높은 이웃노드를 차기노드로 선택한다. 그 결과 패킷이 의미 없거나 잘못된 방향으로 전달되는 것이 방지된다. 또한, 보상 함수는 고정 보상과 잔여 에너지와 채널 상태로 귀결되는 패킷 전송 횟수를 고려한다. 시뮬레이션 결과, 제안 기법이 지연과 에너지 소모 측면에서 QELAR보다 나은 성능을 보였다.

References

[1] Z. Jin, Y. Ma, Y. Su, S. Li, and X. Fu, "A q-learning-based delay-aware routing algorithm to extend the lifetime of underwater sensor networks," *Sensors*, vol. 17, p. 1660, Jul.

2017.

- [2] G.-S. Ahn, S. G. Hong, E. Miluzzo, A. T. Campbell, and F. Cuomo, "Funneling-MAC: A localized, sink-oriented mac for boosting fidelity in sensor networks," in *Proc. 4th Int. Conf. Embedded Netw. Sensor Syst.*, pp. 293-306, Oct. 2006.
- [3] M. Stojanovic and J. Preisig, "Underwater acoustic communication channels: Propagation models and statistical characterization," *IEEE Commun. Mag.*, vol. 47, pp. 84-89, Feb. 2009.