

AQ-NAV: 수중통신에서 거리 추정을 이용한 강화 학습 기반 채널 접속 기법

박석현¹, 신경섭², 조오현^{3*}

¹충북대학교 소프트웨어학과 학생, ²상명대학교 컴퓨터과학과 교수, ³충북대학교 소프트웨어학과 교수

AQ-NAV: Reinforced Learning Based Channel Access Method Using Distance Estimation in Underwater Communication

Seok-Hyeon Park¹, Kyungseop Shin², Ohyun Jo^{3*}

¹Student, Department of Computer Science, Chungbuk National University

²Professor, Department of Computer Science, Sangmyung University

³Professor, Department of Computer Science, Chungbuk National University

요약 본 논문은 기존 강화학습 기반 수중통신 예약방식에서 성능 저하 요인 중 하나인 긴 학습 시간을 단축시킴으로써 에너지 소모를 감소시킬 수 있는 향상된 채널 접속 기법을 제안한다. 수중 무선 네트워크에서 노드 간 거리를 추정하여 이를 바탕으로 기존 강화 학습 기반 채널 접속 방법의 학습 범위의 최대, 최소치를 결정한다. 이는 기존 강화학습의 학습 범위를 줄일 수 있다. 수중 무선 네트워크 환경의 특성에 따른 거리 추정값의 오차를 고려하여 NAV 학습 범위를 고려하며, 이를 적용하기 위해 인위적으로 간섭의 크기를 변경시켜가며 학습 진행률에 대한 성능 테스트를 진행하였다. 실험 결과 기존 방법 대비 제안된 AQ-NAV 방안의 경우, 20-40회 학습에서도 360개의 학습 배열 중 평균 340-350개 이상의 학습 배열이 학습이 진행되었고 50회 이상 학습에서는 모든 학습 배열에 대하여 학습이 진행되었다. 반면, 기존 연구의 경우 학습이 120회 이상 진행되어도 360개의 배열 중 300-320개의 배열에 대한 학습이 진행되었다. 실험에서는 기존 대비 적은 횟수의 시도로 학습이 가능함을 보여준다. AQ-NAV가 수중 무선 네트워크에 적용될 경우 에너지 소비 절감을 통해 기존의 방안의 문제점을 완화하고 네트워크 성능 향상을 이룰 것으로 예상된다.

주제어 : 수중 무선 네트워크, 센서, 강화학습, IoT 네트워크, RTS, CTS, NAV

Abstract This work tackles the problem of conventional reinforcement learning scheme which has a relatively long training time to reduce energy consumption in underwater network. The enhanced scheme adjusts the learning range of reinforcement learning based on distance estimation. It can be reduce the scope of learning. To take account the fact that the distance estimation may not be accurate due to the underwater wireless network characteristics. this research added noise in consideration of the underwater environment. In simulation result, the proposed AQ-NAV scheme has completed learning much faster than existing method. AQ-NAV can finish the training process within less than 40 episodes. But the existing method requires more than 120 episodes. The result show that learning is possible with fewer attempts than the previous one. If AQ-NAV will be applied in Underwater Networks, It will affect energy efficiency. and It will be expected to relieved existing problem and increase network efficiency.

Key Words : Underwater Wireless Network, Sensor, Reinforcement Learning, IoT Network, RTS, CTS, NAV

This research was a part of the project titled 'Development of Distributed Underwater Monitoring and Control Networks', funded by the Ministry of Oceans and Fisheries, Korea.

*Corresponding Author : Ohyun Jo(ohyunjo@chungbuk.ac.kr)

Received June 26, 2020

Accepted July 20, 2020

Revised July 5, 2020

Published July 28, 2020

1. 서론

1.1 서론

자원의 고갈 문제를 해결하기 위한 해상 및 수중 자원의 발굴은 다양한 산업과 기술의 융합이 요구된다[1,2]. 수중 환경오염과 태풍, 지진, 해일 등 자연재해에 대한 분석 및 예측, 실시간 모니터링을 위한 기술 개발에 대한 수요가 증가하고 있다[3-5].

실제 국방 산업, 어업 등 다양한 수중 기술에 대한 수요가 증가하고 있으며 글로벌 IT 기업을 비롯하여 많은 기업 및 대학, 연구기관에서 실제 수중 통신망 구축 사업을 위한 연구 및 구현이 진행 중이다. 이러한 수중 산업 분야의 융합과 기술 혁신을 실현하기 위해서는 수중 통신 기술이 필수적이다. 수중 모니터링, 센서를 통한 감지, 환경 조사 등 다양한 응용 시나리오에서는 수중 관련 데이터를 생성하고 생성한 데이터를 수집해야 한다. 이러한 데이터를 수집하기 위해서는 통신 수단이 필요하다.

수중 환경에서는 무선 통신 시스템을 이용한다. 무선 시스템을 사용하는 이유는 해양 생물과 선박, 잠수부, 잠수함 등 수중에서 이동하는 매체가 있어 유선 환경을 사용하기 힘들기 때문이다. 또한 유선의 경우 바다 내의 염도에 의해 케이블 경화 및 부식이 발생할 수 있으며, 수중 내에 있다는 점에서 유지 보수가 힘들기 때문에 사용의 한계가 있다.

수중 통신은 특수한 환경에 의하여 지상에서의 통신 방식과는 차별화되는 다양한 특성을 갖는다. 우선 수중 통신에서는 주로 음파를 이용한다. 물론 수중 통신에서

도 RF(Radio Frequency) 통신, 가시광 통신 등 다양한 무선 통신 방식이 시도되었다. 하지만 이러한 매체를 사용하는 무선 통신의 경우 수중 네트워크에서 여러 취약점을 갖고 있다. 주로 짧은 전파 유효 사거리 문제나 소음 및 잡음, 굴절, 산란 등의 문제로 인해 사용 환경이 매우 제한적이다[5-7].

이러한 이유로 수중에서는 일반적으로 해당 환경에 적합한 음파를 사용한다. 음파는 느린 전파 속도와 소음, 잡음에 대한 통신 품질 저하, 높은 에너지 소비량, 배터리 사용 환경 등 다양한 문제점이 존재한다. 특히 배터리를 사용하는 환경은 수중 통신의 많은 제약을 주게 된다. 배터리를 사용하고 초기 설치 이후 유지 보수가 불가능한 수중 네트워크의 경우 노드 설치 이후 최대 한 오랫동안 유지되어야 한다. 실제로 수중 네트워크에서는 에너지 소모를 최소화하는 연구에 많은 관심이 쏠리고 있다.

하지만 그럼에도 불구하고 음파를 사용하는 이유는 비교적 수중 내에서 장거리 통신이 가능하다는 큰 이점을 갖는다.

1.2 연구 배경

본 논문에서는 경쟁 기반의 무선 채널 접속을 수행하는 수중 통신의 문제점 중 하나인 NAV(Network Allocation Vector) 최적화 문제를 해결하고자 한다. 수중 통신에서 사용하는 음파의 긴 전파 시간은 효율적이지 못한 통신 환경과 노드 간 통신 불공정성의 문제를 발생 시킨다[8-11].

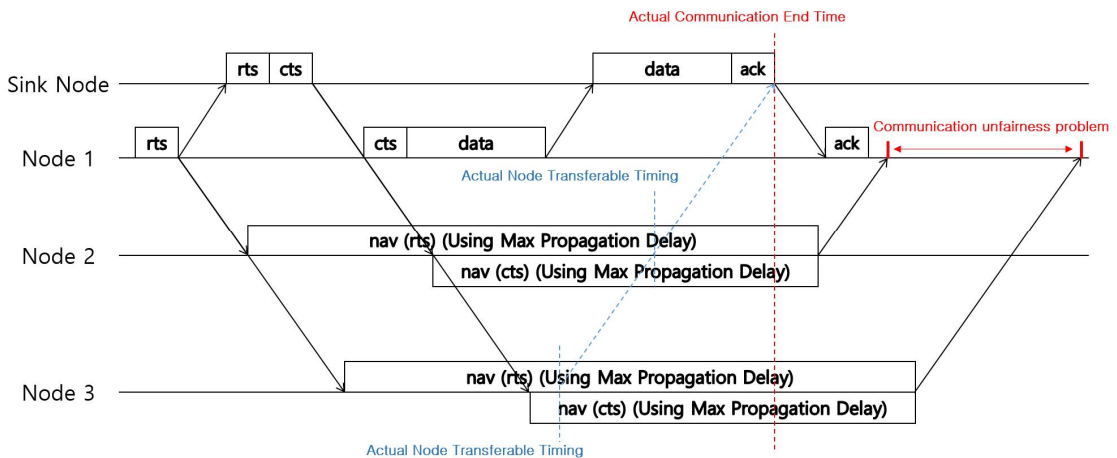


Fig. 1. Max Propagation Delay Problem of Underwater Wireless Networks.

Fig. 1은 최대 전파 지연시간을 사용하는 NAV의 비효율성 문제와 통신 불공정성 문제를 보여준다.

IEEE 802.11 프로토콜 기반의 수중 통신에서는 채널 예약을 위해 사용되는 RTS(Request to Send)/CTS(Clear to Send) 패킷 내부에 채널 점유를 위한 NAV 값을 계산하여 전송하게 되는데 [12], 수중 네트워크에서는 노드 사이의 거리를 정확하게 파악할 수 없기 때문에 최대 전파 지연시간을 NAV 값에 반영한다.

이는 전파 지연시간에 의해 시간 동기화 작업이나 노드 위치 계산 등이 어려워 최대 전파 지연시간을 사용하는 것이다. 하지만 최대 전파 지연시간을 이용한 통신방식은 수중 무선 네트워크에 통신 불공정성을 일으키고 비효율적인 통신 환경을 구성하게 된다[8].

위에서 언급한 문제들을 해결하기 위해 다양한 노력들이 이어져왔고, 다양한 방법의 NAV 최적화 방식이 제안되었다. 강화 학습을 활용한 NAV 값 계산에 대한 연구 [13]의 경우 시간이 지남에 따라 학습이 진행되면서 최적 NAV 값을 계산하였고 이로 인해 노드 간 불공정 문제와 불필요한 대기 시간문제를 해결할 수 있었다. 하지만 학습을 위해 초기에 임의로 NAV 값을 선택하게 되면서 학습이 진행되는 동안 패킷 충돌 문제가 발생할 수 있다. 충분히 큰 NAV 값을 선택하지 못하는 경우 통신 충돌이 일어나게 되고 이는 네트워크에 악영향을 끼칠 수 있다. 이러한 경우 재전송이 이루어져야 하며 이는 배터리를 최대한 아껴야 하는 수중 네트워크에서는 치명적인 단점이 될 수 있다.

기존 연구에서는 해당 환경에서의 모든 NAV 값에 대해 무작위로 선택하여 학습을 진행하였다. 본 논문에서는 강화 학습 시 선택하는 NAV 값을 RTT(Round Trip Time) 또는 RSS(Received Signal Strength)를 참조하여 거리를 계산하고 계산된 값 기반으로 학습 범위를 줄인다. 수중 환경에서 RTT 또는 RSS를 활용하여 거리를 계산하고 이를 통해 NAV 값을 설정할 경우 산란, 굴절, 간섭, 노이즈 등 수중 환경에 의해 정확한 NAV 값을 찾기 힘들 수 있다. 하지만 계산된 거리 값을 활용하여 무작위 선택의 범위를 줄여줄 수 있다면 NAV 값 학습에 필요한 시도 횟수를 감소시킬 수 있다. 이는 불필요한 시도로 인한 에너지 소비를 감소시킬 것으로 기대된다.

2. 기존 연구 및 연구 배경

2.1 Q-NAV

Q-NAV[13]는 수중 무선 네트워크에서 강화 학습 기반 NAV 설정 방법이다. 일반적인 통신 환경에서 RTS/CTS 패킷에 NAV 값을 설정하여 보내는 방식을 사용하지 않고 RTS/CTS를 수신한 노드가 무작위 NAV값을 선택하는 방식이다. 무작위 NAV 값의 범위는 해당 환경에서 설정한 통신 범위에 따라 달라진다. 수중 통신에서는 통신 범위가 제한적이고 주로 미리 범위를 설정해 두기 때문에 이를 참고한다. RTS/CTS를 수신한 노드가 무작위 선택을 통해 NAV 값을 선택하게 되고 NAV 값 만큼 대기 후 통신을 시도한다. 대기 이후 통신 시도에서 통신이 성공하는 경우 해당 NAV 값에 양의 보상 값을 부여한다. 통신에 실패하는 경우 음의 보상 값을 부여하는 방식을 통해 학습을 진행한다. NAV에 대한 보상을 참조하기 위해 노드는 배열을 생성하고 배열 안에 보상 값을 저장한다.

2.2 기존 연구에서의 문제

기존 연구에서 제안된 강화 학습 기반 NAV 결정 방식에서는 불필요한 NAV 대기 시간과 노드 간 통신 불공정성 문제를 해결하고자 하였다[13]. 기존에 제안된 방법은 일반적인 통신에서 사용되는 RTS/CTS에 NAV 값을 넣어 전송하는 방법을 사용하지 않는다. RTS/CTS 프레임에 NAV 값을 설정하지 않고 각 노드들이 모든 노드에 대한 NAV 값 학습 테이블을 가지고 이를 활용한다. 해당 환경의 노드 수와 통신 범위에 따라 테이블의 크기는 달라지며, 통신 초기에는 RTS/CTS를 수신한 상황에서 노드가 데이터를 송신하게 되는 경우, NAV값을 무작위로 선택하여 대기한 후 데이터를 송신한다. 이 송신한 데이터의 전송 결과에 따라 사용한 NAV 값에 해당하는 보상 테이블의 보상 값을 조정한다. 이후 충분한 학습이 진행되는 경우 NAV 값을 무작위 선택을 하지 않고 보상 테이블을 참조하여 NAV 값을 결정한다.

노드의 위치를 알기 어렵고 노드 간 시간 동기화가 어려운 수중 네트워크에서 정확한 NAV 값을 계산하기 어렵지만 제안된 연구에서는 정확한 NAV 값을 구할 수 있음을 보여주었다. 이는 불필요한 NAV 대기시간과 노드 간 통신 불공정성 문제를 해결하지만 통신 초기에

무작위로 선택하는 NAV 값의 범위가 넓기 때문에 학습이 충분히 완료되기 전까지 네트워크 자원의 낭비를 일으키게 된다. Fig. 2는 위에서 설명한 무작위 NAV 값 선택의 문제점을 보여준다. 너무 작은 NAV 값을 선택하여 통신의 충돌을 일으키거나 너무 큰 NAV 값을 선택하여 불필요한 대기시간을 낭비하게 된다. 학습이 빠르게 이루어지지 못하는 경우 네트워크에 전반적인 성능을 저하시킬 수 있다.

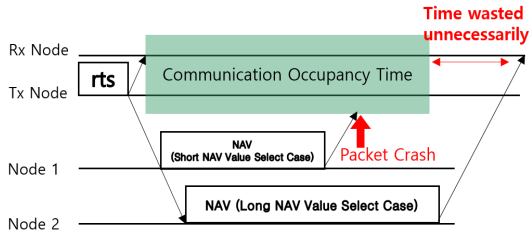


Fig. 2. Random NAV Selection Problem

3. 제안 알고리즘

본 논문에서는 기존에 제안된 강화 학습 기반 NAV 값 결정 방식인 Q-NAV 기반의 문제점을 개선한 방식을 제안한다. Q-NAV의 경우 최대 NAV 값의 모든 범위를 학습하기 때문에 학습에 많은 시간이 소요된다. AQ-NAV(Advanced Q-NAV)는 Q-NAV의 기본적인 프로토콜을 기반으로 학습이 될 때까지 소모되는 시간을 줄이기 위해 RTT 또는 RSS를 활용한다. 기본적인 원리는 RTT/RSS를 통해 거리를 유추하고 이를 이용하여 가능한 NAV의 선택 범위를 줄이는 방식이다.

3.1 수중 환경에서 거리 측정

수중환경에서의 거리 측정은 RTT(Round Trip Time), RSS(Received Signal Strength), TDoA(Include Time Difference of Arrival), AoA(Angle of Arrival) 등 다양한 기법이 제안되어 왔다[14, 15]. 하지만 수중 음파의 이동속도에 따른 시간 동기화의 어려움, 다중 경로 페이딩(Multi-path Fading) 문제, RF 사용의 어려움 등 측정이 어렵거나 부정확한 경우가 많다. 하지만 RSS 기반의 거리 측정 방법[15]은 수중 네트워크에서의 거리 측정이 높은 정확도 이루어질 수 있음을 보였다. 해당 연구에서는 수학적으로 RSS 기반 거리 계산을 증명하였다. 수중음향 전송 시 염도, 산도, 압력,

수온, 기포, 표면 반사 등에 의하여 발생하는 다양한 손실(Spreading loss, attenuation loss)을 활용하여 전송 손실(Transmission Loss: TL)에 대한 계수를 유도하고 Lambert W function을 활용하여 거리 측정이 가능함을 보여주었다. 해당 연구에서는 최종적으로 수식 (1)와 같이 거리($Distance$)를 측정하였다.

$$Distance = A_1 * W(A_2 * e^{A_3}),$$

where :

$$\begin{cases} A_1 = 10000/(\lambda\alpha) \\ A_2 = 1/A_1 \\ A_3 = \lambda TL \end{cases} \quad (1)$$

이러한 수중 네트워크에서의 거리 측정 연구가 정확하다면 NAV 값에 바로 적용할 수 있을 것이다. 하지만 앞서 기술한 바와 같이 수중 네트워크에서는 수많은 변수가 존재하며 통신 방해요소가 많다. 특히 음파를 사용하는 수중 네트워크에서 잡음, 간섭, 굴절, 산란 등 신호의 강도가 항상 일정하지 않고, 거리 측정이 정확하지 않을 수 있다. 따라서 본 연구에서는 신호 강도를 통해 측정된 대략적인 위치를 기반으로 AQ-NAV 무작위 선택 범위를 조절한다.

RSS 기반 거리 계산 [15] 외에도 RTT 방식을 활용하여 거리를 측정할 수 있다. RTT의 경우 통신 송, 수신 사이의 패킷 왕복 시간을 통해 구할 수 있다. 수중에서 사용하는 음파의 경우 약 1500m/s 속도를 갖는다. 노드 간 패킷 왕복 시간을 “Round Trip Time”이라 하고 이는 측정용 패킷을 상대측 노드에 전송하고 응답을 수신할 때까지의 시간을 측정하여 거리($Distance$)를 계산한다. 노드 간 거리($Distance$)는 식(2)과 같이 구할 수 있다.

$$Distance = \frac{(Round\ Trip\ Time/2)}{Acoustic\ Wave\ Speed} \quad (2)$$

3.2 AQ-NAV (Advanced Q-NAV)

AQ-NAV는 강화 학습의 학습량을 줄이기 위해 RSS를 통해 계산된 거리나 RTT를 통해 구해진 거리를 사용한다. RSS나 RTT와 같은 거리 측정 방식을 사용한 후 계산된 거리를 통해 NAV 값을 결정한다. 수중 네트워크 환경 특성상 정확하지 못할 수 있다는 점을 감안하여 AQ-NAV의 학습 테이블 범위의 최소, 최대한도를 설정한다. 수중에서 NAV 값은 식(3),(4)와 같이 구

할 수 있다.

$$NAV_{RTS} = (3 * Propagation Delay Time) + CTS + DataFrame + ACK \quad (3)$$

$$NAV_{CTS} = (2 * Propagation Delay Time) + DataFrame + ACK \quad (4)$$

결정된 NAV 값에서 AQ-NAV 최소값과 최대값을 결정한다. 결정 방식은 식(5),(6)와 같다.

$$AQNAV_{Min} = NAV - \alpha \quad (5)$$

$$AQNAV_{Max} = NAV + \alpha \quad (6)$$

거리 계산 및 실제 통신 오차를 감안하여 $AQNAV_{Min}$ 값과 $AQNAV_{Max}$ 값 사이의 범위 내에서 학습을 진행한다.

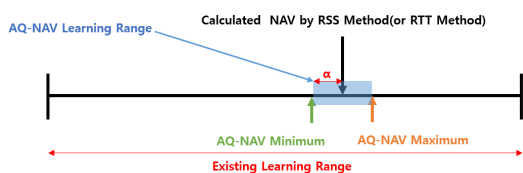


Fig. 3. Differences of NAV Searching Window between Existing Research and AQ-NAV

Fig. 3은 기존 강화 학습에 사용되던 학습 범위와 AQ-NAV의 학습 범위의 차이를 보여준다. 학습 범위는 환경 설정에 의해 달라질 수 있으나 기존 방식에 비해 더 작은 학습 선택폭을 갖는다. Fig.3. 에서와 같이 기존 방식은 해당 환경의 최대 전파 지연시간을 사용한 NAV 값을 학습 범위로 결정하나 제안하는 AQ-NAV는 측정된 거리를 기준으로 NAV 값을 계산하고 오차 범위를 고려하여 학습 범위를 결정한다. Fig 4.는 AQ-NAV의 진행 방식을 pseudocode로 표현한 것이다. E-Greedy 방식으로 정책을 결정하는데, 학습이 충분히 이루어질 때까지 NAV 값을 무작위로 선택하여 탐험적 학습을 수행하게 된다. 최대 NAV 값 내에서 무작위로 선택하던 기존 연구와는 다르게 AQ-NAV에서는 계산된 특정 범위 내에서 NAV 값을 무작위로 선택하게 된다.

AQ-NAV Algorithm :

```

if Receive_RTS/CTS():
    if random() < e:
        mean_nav = import_measured_nav()
        min_nav = mean_nav - alpha
        max_nav = mean_nav + alpha
        random_nav = random(min_nav, max_nav)
        if communication_failure():
            Q_table[random_nav] = Q_table[random_nav] - reward
        else:
            Q_table[random_nav] = Q_table[random_nav] + reward
    else:
        nav = select_max_reward_index(Q_table[node])
        if communication_failure():
            Q_table[nav] = Q_table[nav] - reward
        else:
            Q_table[nav] = Q_table[nav] + reward

```

Fig. 4. AQ-NAV Algorithm

RSS 또는 RTT로 계산된 NAV 값에서 오차 범위인 min_nav 와 max_nav 를 설정한다. min_nav 와 max_nav 사이의 값 중 임의의 값을 선택하여 실제 사용할 NAV 값으로 결정한다. 결정된 NAV 값을 사용하여 통신한 후 성공하는 경우 보상을 주는 방식을 채택한다. 학습이 진행됨에 따라 점차 무작위 선택 비중을 줄이고 보상 테이블의 최고 점수 NAV 값을 사용한다. 이와 같은 단순한 방식의 학습 방법은 낮은 컴퓨팅 파워와 에너지 제한적인 환경을 갖는 수중 무선 네트워크에서 강한 이점이 될 수 있다.

기존 연구에서는 학습을 해야 하는 범위가 AQ-NAV의 학습범위보다 넓어 학습에 많은 시간이 소요된다. 이는 결과적으로 학습 진행하는 동안 통신 성공률을 저하시키게 되고 실패한 통신을 재시도하면서 에너지를 추가적으로 소비하게 된다. 반면 AQ-NAV는 RSS/RTT를 통해 계산된 거리를 기반으로 학습 범위를 줄인다. 이는 학습 시 선택하는 NAV 값의 범위를 줄임으로써 학습에 필요한 시간을 줄일 수 있으며, 결과적으로 에너지 소비를 줄일 수 있다.

4. 검증 및 실험

성능을 평가하기 위해 시뮬레이션에서는 기존 연구인 Q-NAV[13]와 동등한 환경에서 비교하였다. Python을 이용하여 시뮬레이션 하였고, 총 10개의 노드가 각 노드끼리 통신을 하는 환경을 구성하였다. 실험 변수는 Fig 5.와 같으며 AQ-NAV 실험에서는 패킷 이동 시 실제 수중에서 발생 가능한 특성(간섭, 굴절 등)을 고려하

여 노이즈를 추가하였다.

Simulation Parameter
Nodes location : [0, 300, 800, 1200, 2100, 2700, 3300, 3500, 4300, 4900] (m),
Q table : 10 * 10 * MAX_NAV,
Control frame duration : 3.0 sec,
Data frame duration : 7.0 sec,
Acoustic wave speed : 1500 m/sec
Noise : 10%
Reward : +1, -3

Fig. 5. AQ-NAV Simulation Parameter

실험에 사용된 변수는 실제 수중 통신망에 활용이 가능하도록 우리 연구진이 진행 중인 수중 통신망 구축 사업에서 사용하는 모델을 기반으로 설정하였다. 노이즈는 각 패킷 이동마다 10% 범위 내의 무작위 오차를 갖도록 실험하였다. 실험에서는 NAV 선택 시 학습 범위 산출을 위해 거리 측정을 통해 구해진 NAV 값을 기준으로 음의 방향 20%, 양의 방향 20% 범위를 학습 범위로 선정하였다. NAV 선택 및 보상은 Fig. 4와 같은 방식으로 진행된다. Fig. 6은 채널 접속 및 보상 방식을 보여준다. NAV 시간 이후 데이터 전송을 시도할 때, 충돌하지 않는 경우 양의 보상 값을, 충돌하는 경우 음의 보상 값을 주는 방식을 보여준다.

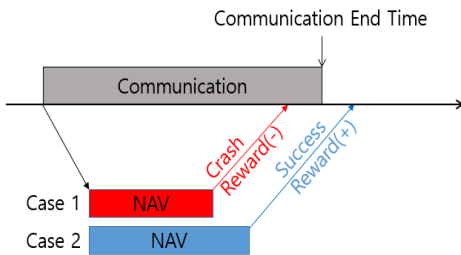


Fig. 6. Two Kind of Case in Communication about Reward

Fig. 7는 학습이 잘 진행되었는지를 알기 위해 무작위 선택 학습 횟수가 증가함에 따라 NAV Table에서의 학습이 이루어져야 하는 배열의 개수를 제외한 나머지 개수의 평균값을 측정한 것이다.

NAV Table에는 보상 값을 저장하기 위한 360개(3.0 Sec)의 배열이 존재한다. y 축은 360개의 보상 테이블 중 학습이 한 번이라도 진행된 배열의 개수를 의

미하며 모든 노드의 평균 개수 수치를 측정하였다. AQ-NAV의 경우 학습을 하지 않은 범위를 개수에 포함시키지 않는다.

x 축은 무작위 선택을 몇 번 진행했는지에 대한 수치를 나타낸다. Q-NAV의 경우 무작위 NAV 값 선택을 20회 진행하였을 때, 실험에 사용한 Q-NAV의 모든 노드의 시도 횟수 평균값이 20회 이하로 나타났다. 중복된 선택이 존재하기 때문에 20회보다 높게 측정될 수 없다. Q-NAV는 무작위 시도 횟수가 증가함에 따라 점차 학습 시도된 배열의 수가 증가하게 된다.

시뮬레이션의 Q-NAV는 무작위 선택을 1000회까지 하더라도 360개의 배열 중 평균적으로 20개 이상의 배열을 확인하지 못하는 것을 보여준다. 이러한 경우 정확하지 못한 학습이 진행될 수 있다. 또한 학습 시도를 많이 해야 학습에 대한 안정도가 보장된다. 이러한 경우는 수중 무선 네트워크에서 네트워크의 성능을 저하시키는 요인이 될 수 있다.

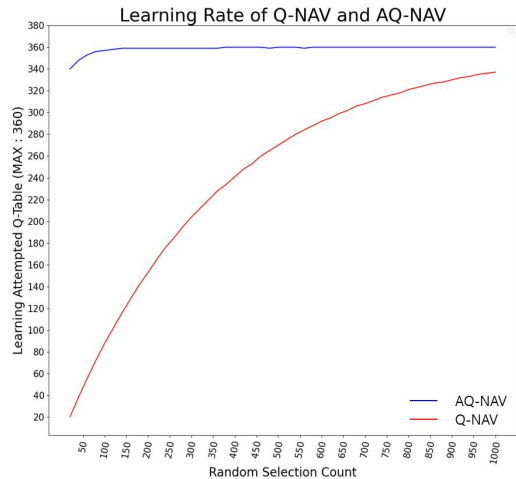


Fig. 7. Learning Rate Difference between Q-NAV and AQ-NAV

AQ-NAV의 경우 모든 배열에 대하여 학습을 진행하지 않는다. RSS/RTT로 측정된 거리를 통해 오차를 감안하여 Table의 특정 범위 내에서 학습을 진행하여 많은 양의 구간을 제외하고 학습을 진행한다. AQ-NAV의 경우 무작위 선택이 20회 된 경우 평균적으로 약 20개 정도의 NAV 값이 학습에 선택되지 못하였다. Q-NAV의 경우 약 340개 이상의 NAV 값이 학습되지

못하는 것 대비 적은 선택 횟수로도 충분한 학습 시도가 이루어진 것을 알 수 있다.

Fig. 8은 2번 노드와 10번 노드 쌍에 의한 통신의 RTS/CTS를 수신한 4번 노드가 10번 노드에게 데이터를 전송하기 위한 NAV 값을 결정하는 경우이다.

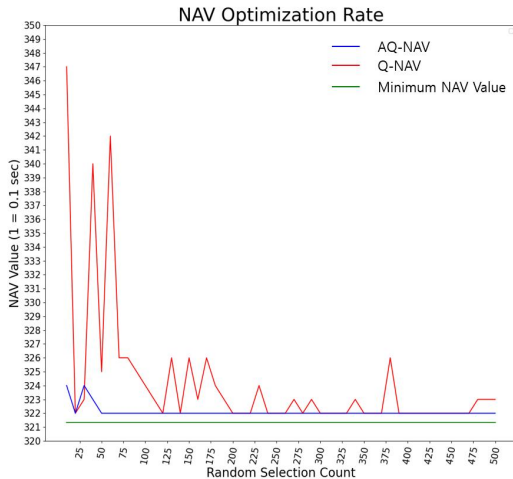


Fig. 8. an Example of Highest Reward NAV Selection

Fig. 8은 무작위 선택 횟수에 따른 최대 보상 NAV 값을 보여준다. 해당 통신 환경에서는 소수점 첫째 자리(0.1초 단위)까지 사용한다고 가정한다. 초록 선의 Minimum NAV Value는 32.133초로 RTS를 수신한 4번 노드가 실제 기다려야 하는 최소한의 NAV 값이다. 통신이 끝나는 시간이 RTS 수신 이후 32.133초인 것을 감안할 때, 32.2초가 선택될 수 있는 가장 짧은 NAV 값이 된다.

Q-NAV의 경우 무작위 선택을 120회 이상 하지 않는 경우, 학습이 완벽하게 이루어지지 않아 높은 NAV 값을 선택하는 경우가 생긴다. 120회를 넘어가면서 안정화되어 최적에 가까운 NAV를 선택하게 되지만 120회 이후에도 높은 수치를 선택하는 경우가 존재한다.

반면 AQ-NAV의 경우 20회에서 40회 사이에서도 충분히 최적에 가까운 NAV 값을 선택하며 이후에는 계속 최적의 NAV값을 선택하게 된다. 모든 노드에 대하여 동일한 실험을 진행하였으며 노드 위치에 따라 약간의 차이를 보이지만 결과는 비슷한 양상을 보였다. AQ-NAV의 경우 많은 학습 범위를 학습 이전에 제외하기 때문에 더 적은 횟수로 최적의 NAV 값을 찾는 모습을 Fig. 7에서 보여주었다. 또한 Fig. 8은 학습 횟수에

따른 NAV 값 최적화가 기존 Q-NAV 대비하여 전 구간에서 빠르고 안정적으로 이루어지는 것을 보여주었다.

5. 결론

본 논문은 기존에 제안된 강화 학습 연구에서 발생할 수 있는 네트워크 성능 저하를 줄이기 위하여 학습 범위를 줄이는 AQ-NAV를 제안하였다.

제안한 방법은 기존에 다양하게 제안된 수중 네트워크에서의 거리 측정 방식을 활용하여 노드 사이의 거리를 측정하고 측정된 거리를 통해 NAV 시간을 계산한다. 수중 통신 특성에 따른 오차율을 감안하여 계산된 NAV 값을 기준으로 학습 범위를 결정한다. 모든 NAV 범위에 대하여 무작위 선택을 하던 기존 연구 대비 범위를 획기적으로 줄여 더 빠른 시간 안에 최적의 NAV 값을 학습하는 AQ-NAV 방법을 제안하였다.

시뮬레이션 결과를 통해 제안하는 방식이 더 적은 시간에, 더 적은 횟수로 학습한다는 것을 보였다. 이는 강화 학습을 채택하면서도 적은 양의 연산으로 확실한 성능 향상을 보여주고 있다. 제안하는 방식을 통해 낮은 컴퓨팅 파워와 에너지 제한적인 환경을 갖는 수중 무선 네트워크에서 통신 기회 불공정성을 효과적으로 해결하고 비효율적인 네트워크 대기시간 문제를 포함해 학습에 걸리는 시간을 효과적으로 줄일 수 있음을 보여주었다.

REFERENCES

- [1] <https://water.usgs.gov/edu/earthhowmuch.html>
- [2] M. Mangel. (1983). Optimal search for and mining of underwater mineral resources. *SIAM Journal on Applied Mathematics*, 43(1), 99-106. DOI : 10.1137/0143008
- [3] N. Wakita, K. Hirokawa, T. Ichikawa & Y. Yamauchi. (2010). Development of autonomous underwater vehicle (AUV) for exploring deep sea marine mineral resources. *Mitsubishi Heavy Industries Technical Review*, 47(3), 73-80.
- [4] E. M. Sozer, M. Stojanovic & J. G. Proakis. (2000). Underwater acoustic networks. *IEEE journal of oceanic engineering*, 25(1), 72-83. DOI : 10.1109/48.820738
- [5] I. F. Akyildiz, D. Pompili & T. Melodia. (2004). Challenges for efficient communication in underwater acoustic sensor networks. *ACM Sigbed Review*, 1(2), 3-8. DOI : 10.1145/1121776.1121779

- [6] F. Yunus, S. H. Ariffin & Y. Zahedi. (2010). A survey of existing medium access control (MAC) for underwater wireless sensor network (UWSN). *In 2010 Fourth Asia International Conference on Mathematical/Analytical Modelling and Computer Simulation* (pp. 544-549). Bornea:IEEE.
- [7] X. Guo, M. R. Frater & M. J. Ryan. (2009). Design of a propagation-delay-tolerant MAC protocol for underwater acoustic sensor networks. *IEEE Journal of Oceanic Engineering*, 34(2), 170-180. DOI : 10.1109/JOE.2009.2015164
- [8] D. Shin & D. Kim. (2008). A dynamic NAV determination protocol in 802.11 based underwater networks. *In 2008 IEEE International Symposium on Wireless Communication Systems* (pp. 401-405). Reykjavik:IEEE.
- [9] S. Y. Shin & S. H. Park. (2007, December). UWA-NAV-Energy Efficient Error Control Scheme for Underwater Acoustic Sensor Network. *In International Conference on Embedded and Ubiquitous Computing* (pp. 505-514). Berlin, Heidelberg:Springer
- [10] Y. D. Chen, C. C. Li, R. T. Dai & K. P. Shih. (2011). On enhancing four-way handshake with stair-like NAV setting for underwater acoustic networks. *In OCEANS'11 MTS/IEEE KONA* (pp. 1-6). Waikoloa:IEEE.
- [11] J. Cho, E. Shitiri & H. S. Cho. (2017). Network allocation vector (NAV) optimization for underwater handshaking-based protocols. *Sensors*, 17(1), 32. DOI : 10.3390/s17010032
- [12] Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specification, IEEE Std. 802.11, 1999.
- [13] Seok-Hyeon Park, Ohyun Jo (2020). Q-NAV: NAV Setting Method based on Reinforcement Learning in Underwater Wireless Networks, *arXiv preprint arXiv:2005.13521*.
- [14] P. Karn & C. Partridge. (1987). Improving round-trip time estimates in reliable transport protocols. *ACM SIGCOMM Computer Communication Review*, 17(5), 2-7. DOI:10.1145/55483.55484
- [15] M. Hosseini, H. Chizari, C. K. Soon & R. Budiarto. (2010). RSS-based distance measurement in underwater acoustic sensor networks: An application of the Lambert W function. *In 2010 4th International Conference on Signal Processing and Communication Systems* (pp. 1-4). Gold Coast:IEEE.

박 석 현(Seok-Hyeon Park)

[학생회원]



- 2019년 2월 : 세명대학교 컴퓨터학과(학사)
- 2019년 2월 ~ 현재 : 충북대학교 컴퓨터학과 석사과정
- 관심분야 : 수중 통신, 인공지능
- E-Mail : seokhyeon@chungbuk.ac.kr

신 경 섭 (Kyungseop Shin)

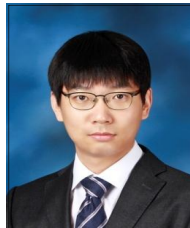
[정회원]



- 2011년 1월 : KAIST 전기 및 전자공학과 석사
- 2015년 1월 : KAIST 전기 및 전자공학과 박사
- 2015년 2월~2017년 8월 : KT 융합기술원 인프라연구소 5G TF 선임연구원
- 2017년 9월~2020년 3월 : 세명대학교 컴퓨터학부 조교수
- 2020년 3월~현재 : 상명대학교 컴퓨터학과 조교수
- 관심분야 : 차세대 이동통신, IoT 네트워크, 인공지능
- E-Mail : ksshin@smu.ac.kr

조 오 현(Ohyun Jo)

[정회원]



- 2005년 2월 : 한국과학기술원 전기 및전자공학(학사)
- 2007년 8월 : 한국과학기술원 전기 및전자공학(석사)
- 2011년 2월 : 한국과학기술원 전기 및전자공학(박사)
- 2011년 4월 ~ 2016년 2월 : 삼성전자 DMC 연구소
- 2016년 3월 ~ 2017년 7월 : 한국전자통신연구원
- 2017년 8월 ~ 2018년 2월 : 육군사관학교 전자공학과 조교수
- 2018년 3월 ~ 현재 : 충북대학교 소프트웨어학과 조교수
- 관심분야 : IoT 융합, 정보통신 및 네트워크, 기계학습
- E-Mail : ohyunjo@chungbuk.ac.kr