이중 행동 DQN 알고리즘 기반 저궤도 위성 통신 물리 계층 보안

이상철, 박지훈, 채승호*

한국공학대학교

shng9522@tukorea.ac.kr, wlgns4457@tukorea.ac.kr, *shchae@tukorea.ac.kr

Physical Layer Security of LEO Satellite Communication Based on Dual Action Deep Q-network

Sangcheol Lee, Jihoon Park, Seong Ho Chae*

Tech University of Korea

요 약

본 논문은 시간 상관도가 동적으로 변하는 시간 선택적 페이딩 채널 환경에서, 보안 전송률 최대화를 위해 최적의 Wyner 코드북 데이터율을 선택하는 방법을 제안한다. 이를 위해 이중 행동 DQN(Dual Action Deep Q-Network: DA-DQN) 알고리즘을 제안한다. 시뮬레이션을 통해 제안하는 알고리즘이 기존의 알고리즘보다 빠른 수렴성과 우수한 성능을 가짐을 확인하였다.

Ⅰ. 서 론

무선 통신은 신호의 브로드케스트로 인해 태생적으로 보안에 취약한 특성이 있으며, 이를 해결하기 위한 다양한 물리계층 보안 향상 기법들이 제안되었다[1]. 최근, 시간 선택적 페이딩 채널 환경에서 송신기가 시공간 블록 코드(space-time block code) 기반 보안 메시지를 전송할 때, 수신기/도청기가 LML(Linear Maximum Likelihood) 및 ZF(Zero-forcing) 신호 복호 방법 사용 시 획득 가능한 보안 전송율(secrecy transmission rate) 분석 연구가 이루어졌다[2]. 참고문헌[2]는 송신기의 완전 탐색(exhaustive searching) 기반 최적 보안 메시지 데이터율 및 전송 코드워드 데이터율의 선택이 보안 전송율을 최대화시킬 수 있음을 보였다. 하지만, 완전 탐색 방법은 많은 소요 시간 발생으로 인해 채널의 시간 상관도(time correlation)가 동적으로 변하는 환경에는 직접적인 적용이 어렵다. 이를 극복하기 위해, 참고문헌 [3]에서 DQN(Deep Q-Network) 기반 보안 메시지 데이터율 및 전송 코드워드 데이터율 선택 방법이 제안되었으나, 이는 매 학습마다 두 데이터율 중 하나만 선택적으로 변화시키는 방법으로 성능 및 학습 속도 저하 문제를 발생시킨다. 따라서, 본 논문에서는 동적으로 시간 상관도가 변화하는 환경에서, 보안 전송률 최대화를 위한 이중 행동 DQN(Dual Action Deep Q-Network: DA-DQN) 알고리즘을 제안하고, 기존 기법들과 성능을 비교 분석한다.

Ⅱ. 시스템 모델

그림 1은 본 논문에서 고려하는 통신 시스템을 보여준다. 여기서, 송신기(위성)는 2개의 안테나를 가지며, 수신기(지상국) 및 도청기(지상국)는 각각 1개의 안테나를 가진다. 송신기는 보안 메시지를 2개의 송신 안테나와 2개의 시간 슬롯을 활용하여 Alamouti 시공간 블록 부호를 통해 수신기에게 전송하고 도청기는 이를 도청한다. 수신기와 도청기는 LML과 ZF 신호 복호 방법 중 하나를 사용할 수 있다. 송신기는 수신기와 도청기의 채널 상태 정보(Channel State Information : CSI)는 가지고 있지 않으나, 수신기와 도청기의 이동속도 정보를 바탕으로 송신기-수신기 링크와 송신기-도청기 링크의 채널 시간 상관도(time correlation) 정보만 가지고 있음을 가정한다. 송신기-수신기 링크, 송신기-도청기 링크의 채널은 쉐도우드 라이시안 페이딩(Shadowed Rician Fading)을 따르며, 시간 t에서의 시간 상관도는 각각 $\rho_B^t \in [0,1]$와 $\rho_E^t \in [0,1]$를 가진다.

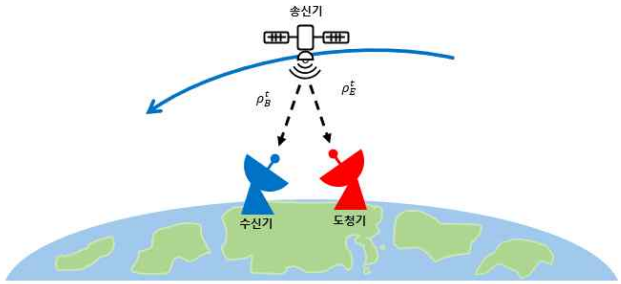

그림 1. 시스템 모델

Alice는 채널 상태 정보를 가지지 않음에 따라, 시간 t에서 Wyner 코드북 기반 보안 메시지 데이터율 $R_{\mathrm{S}}^t = R_T^t - R_E^t$를 가지는 코드워드를 전송한다. 여기서, $R_{\mathrm{T}}^t (\geq R_E^t)$와 R_{E}^t는 임의 선택이 가능하며, $0 \leq R_T^t, R_E^t \leq R_{\max}$의 제한 범위를 가진다. 수신기와 도청기가 각각 신호 복호 방법 $p,q \in \{LML, ZF\}$를 사용할 때, 수신 신호 대비 잡음 비율은 참고문헌[2]와 같이 주어지며, 다음 두 가지 사건들 $\mathrm{E}_{co}^{t,p}$와 $\mathrm{E}_{so}^{t,q}$를 정의할 수 있다. 여기서, $\mathrm{E}_{co}^{t,p}$는 송신기-수신기 링크 채널 용량이 전송율 R_{T}^t보다 작아 코드워드를 디코딩 할 수 없는 사건, $\mathrm{E}_{so}^{t,q}$는 송신기-도청기 링크 채널 용량이 전송율 R_{E}^t보다 커서 도청기가 보안 메시지 일부를 디코딩하는 사건이다. 보안 메시지 데이터율 R_{S}^t에 대한 보안 전송율(secrecy transmission rate)은 다음과 같이 정의된다[2].

$$T_{p,q}^t = R_{\mathrm{S}}^t (1 - \Pr[\mathrm{E}_{co}^{t,p}])(1 - \Pr[\mathrm{E}_{so}^{t,q}]). \quad (1)$$

따라서, 최적화 문제는 다음과 같이 주어진다.

$$\max_{\{R_T^t,\, R_E^t\}} \sum_{t=1}^{T} T_{p,q}^t, \text{ s. t. } R_{\mathrm{T}}^t \geq R_E^t. \quad (2)$$

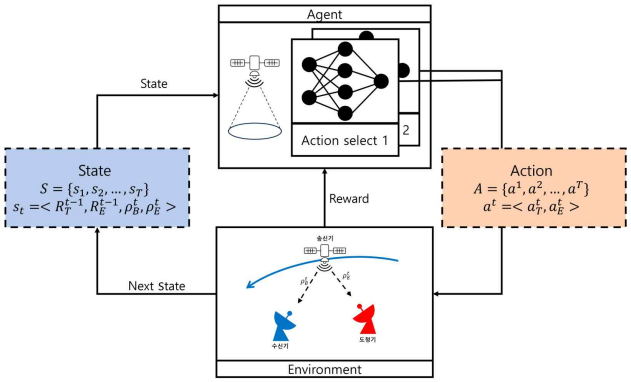

그림 2. DA-DQN 알고리즘 MDP

Ⅲ. 제안하는 데이터율 선택 알고리즘

본 장에서는 시간 상관도가 동적으로 변하는 환경에서, 보안 전송률을 최대화하기 위한 코드북 데이터율 선택 알고리즘을 제안한다. 그림 2은 본 논문에서 제안하는 DA-DQN 알고리즘의 MDP(Markov Decision Process)를 나타낸다.

- 에이전트*(Agent)*: 송신기가 에이전트이며, 송신기는 환경과 상호작용하며 채널 시간 상관도에 따라 Wyner 코드북 생성 데이터율을 제어한다.
- *상태(State)*: 시간 t의 상태는 $s_t =< R_T^{t-1}, R_E^{t-1}, \rho_B^t, \rho_E^t >$와 같이 정의된다.
- *행동(Action)*: 시간 t에서 송신기는 R_{T}^t, R_{E}^t의 값을 동시에 제어하며, $a_T^t \in \{\triangle R_T, -\triangle R_T\}$와 $a_E^t \in \{\triangle R_E, -\triangle R_E\}$로 정의된다.
- *보상(Reward)*: 시간 t에서 보상은 다음과 같이 정의된다.

$$r_t = \begin{cases} \varsigma, & \text{if } R_T^t > R_E^t \text{ and } (E_{co}^{t,p})^c \text{ and } (E_{so}^{t,p})^c, \\ \zeta, & \text{if } R_T^t < R_E^t \text{ or } E_{co}^{t,p} \text{ or } E_{so}^{t,p} \end{cases} \quad (6)$$

여기서, $\varsigma(>0)$는 Wyner 코드북이 성공적으로 생성되고 $\mathrm{E}_{co}^{t,p}$와 $\mathrm{E}_{so}^{t,p}$ 아웃티지가 발생하지 않고 보안 메시지가 성공적으로 전송되었을 때의 보상이고, $\zeta(<0)$는 코드북 생성 실패, $\mathrm{E}_{co}^{t,p}$ 또는 $\mathrm{E}_{so}^{t,p}$ 아웃티지가 발생했을 때 전송 실패에 따른 페널티를 나타낸다.

Ⅳ. 시뮬레이션 결과

본 장에서는 시뮬레이션을 통해 제안된 알고리즘의 성능을 검증한다. 수신기와 도청기는 각각 ZF, LML 신호 복호 방법을 사용함을 가정하고 $R_{\max} = 8$로 설정한다. 성능 검증을 위해 이상적 방법인 완전 탐색의 결과값을 기준으로 삼았다. 또한 다른 알고리즘과의 성능 비교를 위해 DQN, DDQN 알고리즘을 선정하여 학습 결과를 비교하였다. 시간 상관도는 송신기가 수신기, 도청기와 점차 가까워지다 다시 멀어지는 위성의 시나리오를 고려하여 에피소드 동안 ρ_B, ρ_E 값이 0.85에서 0.9로 증가하다가 다시 감소하도록 하여 학습을 진행하였다.

그림 3은 시간 상관도가 동적으로 변하는 시간 선택적 페이딩 채널에서, DA-DQN 알고리즘의 보안 전송률 성능을 비교한 그래프이다. 시뮬레이션 결과 DQN 알고리즘은 완전 탐색 대비 86%의 성능을, DDQN은 90%의 성능을 보여주는 반면, 제안하는 알고리즘은 96%의 성능을 보여주는 것을 확인하였다. 그림 4는 각 알고리즘의 학습 에피소드(Episodes)별 누적 합을 나타낸 그래프이다. 시뮬레이션 결과 제안하는 알고리즘은 75번의 에피소드 만에 수렴하였으며, DDQN과 DQN은 각각 100번과 175번의 에피소드가 필요한 것을 확인하였다.

Ⅴ. 결론

본 논문에서는 동적으로 시간 상관도가 변화하는 시간 선택적 페이딩 채널 환경에서, 보안 전송률을 최대화하는 최적의 Wyner 코드북 데이터율을 선택하기 위한 DA-DQN 알고리즘을 제안하였다. 시뮬레이션을 통해 제안한 알고리즘이 기존의 DQN, DDQN 알고리즘보다 빠른 속도로 수렴함을 알 수 있었고, 보안 전송률 또한 우수한 것을 확인할 수 있었다.

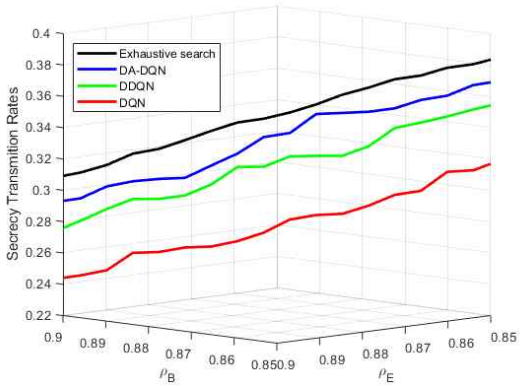

그림 3. 알고리즘별 Secrecy Transmission Rates 비교

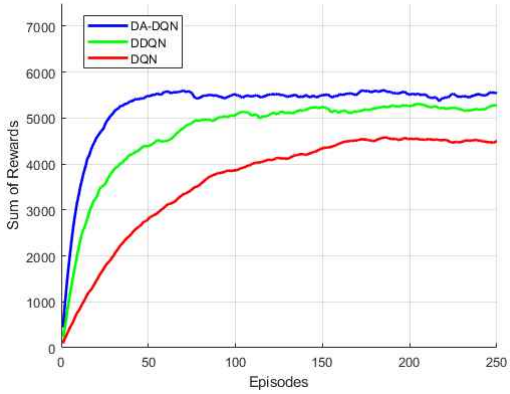

그림 4. 알고리즘별 누적 보상 값 비교

ACKNOWLEDGMENT

이 논문은 2022년도 정부(방위사업청)의 재원으로 국방기술진흥연구소의 지원을 받아 수행된 연구임(21-106-A00-007, 우주계층 지능통신망 특화연구실).

참 고 문 헌

[1] S. H. Chae and W. Choi, "Optimal power allocation for artificial noise in a Poisson interference field," *IEEE Commun. Letters*, vol. 20, no. 8, pp. 1671 - 1674, Aug. 2016.

[2] D. Kim, G. Kwon, H. Lee, and S. H. Chae, "On the achievable secrecy transmission rates by Alamouti space - time block coding in time-selective fading channels," *IET Electronics Letters*, vol. 58, no. 17, pp. 672-674, Aug. 2022.

[3] G. Kim, S. Lee, D. Kim, and S. H. Chae, "DQN based physical layer security enhancement in time-selective fading channels," in *Proc. of KICS Winter Conf*, Feb. 2023.

[4] S. H. Chae, W. Choi, J. H. Lee, and T. Q. S. Quek, "Enhanced secrecy in stochastic wireless networks: Artificial noise with secrecy protected zone," *IEEE Trans. Veh. Tech.*, vol. 69. no. 5, pp. 5647-5651, May 2020.