AWGN 채널에서 매개변수 설정에 강인한 GAN 기반 은밀 통신 연구

정현지, 한규민, 이진영
국립한국해양대학교

guswl1006@g.kmou.ac.kr, gksrbals0626@g.kmou.ac.kr, jylee120@kmou.ac.kr

A study on GAN-based Covert Communication Robust to Parameter Settings over AWGN Channels

Jung Hyun Ji, Han Gyu Min and Lee Jin Young
National Korea Maritime & Ocean University

요 약

본 논문은 통신 존재 자체를 숨기면서 신뢰성 있는 전송을 하기 위한 GAN 기반 물리 계층 은밀 통신 기법을 제안한다. 제안 방식에서는 GAN 구조에서 2 단계 학습을 통해 은닉 신호를 생성한다. 실험 결과, 제안 기법은 기존 기법과 동등한 수준의 은닉성을 유지하면서도, 매개변수 최적화 없이 안정적인 복원 성능을 보장하여 시스템의 강인함을 입증하였다.

Ⅰ. 서 론

은밀 통신은 합법 송신자(Alice)가 감시자(Willie)에게 통신 존재를 들키지 않으면서 합법 수신자(Bob)에게 메시지를 전달하는 통신 방식이다. Willie 의 신호 존재 탐지를 위한 수신 신호 모델은 아래와 같이 이진 가설 검정 모델로 정의할 수 있다.

$$y = \begin{cases} n, & H_0 \\ x + n, & H_1 \end{cases}, \tag{1.1}$$

여기서, n은 AWGN 이며, x는 Alice 의 송신 신호이다. 은밀 통신의 목표는 Willie 의 이진 가설 검정 모델의 오류율을 1 로 만들면서, Bob 에게 전달하는 데이터 전송을 최대화하는 것이다.

최근 물리 계층 보안 분야에서는 생성형 AI 를 활용한 데이터 기반 신호 생성 연구가 활발히 진행되고 있다 [1]. 특히, 생성적 적대 신경망 (Generative Adversarial Network, GAN) [2] 은 생성기와 판별기의 경쟁적 학습을 통해 채널 잡음과 구별 불가능한 신호를 생성할 수 있어 은밀 통신 시스템을 구현하는 데 효과적이다. 기존 GAN 기반 은밀 통신 연구 [3]는 Alice 와 Bob 의 손실을 하나의 손실로 정의하고 은닉성과 BER 간의 trade-off 를 조절하는 매개변수 λ와 μ를 통해 은닉성과 BER 사이의 균형점을 찾는 방식에 의존한다. 그러나 이러한 방식은 성능이 매개변수 설정에 영향을 받으며 최적의 성능을 얻기 위해 매번 λ와 μ를 조절해야 한다는 한계가 있다.

이에 본 논문에서는 이러한 의존성을 줄이고 학습 안정성을 높이기 위해 Willie 의 은닉성 최대화를 위한 학습과 Bob 의 데이터 복원 성능 최대화를 위한 학습으로 구성된 2 단계 학습 방식을 제안한다.

Ⅱ. 본론

은밀 통신의 목표인 은닉성의 정량적 평가를 위해 확률 분포 간의 차이를 측정하는 KL 발산($\mathcal{D}_{KL}$) 을 은닉성 지표로 사용한다 [4]. H_0 일 때의 수신 신호 분포를 P_0, H_1 일 때의 분포를 P_1 이라 할 때, $\mathcal{D}_{KL}$ 은 다음과 같이 정의된다.

$$\mathcal{D}_{KL}(P_1 \| P_0) = \int p_1(y) \log \frac{p_1(y)}{p_0(y)}\, dy. \tag{2.1}$$

본 논문은 GAN 기반 은밀 통신의 송수신 구조를 제안한다. 송신기 Alice(생성기, G)는 길이 M의 이진 메시지 $\mathbf{m} \in \{0,1\}^M$ 과 랜덤 벡터 $\mathbf{z} \in \mathbb{R}^M$ 을 결합하여 완전 연결 신경망(FCN)에 입력하며, 이를 통해 평균 전력이 P_s 로 제한된 송신 신호 $\mathbf{x} \in \mathbb{R}^N$ 를 생성한다. 감시자 Willie 는 수신 신호 $\mathbf{y}$를 통해 신호 존재 여부를 판별하는 판별기(D)로 동작한다. 수신기 Bob(B)은 $\mathbf{y}$로부터 원본 메시지 $\mathbf{m}$을 복원하는 역할을 수행하며, 결과적으로 Alice 와 Bob 은 하나의 오토인코더 구조를 형성하도록 설계된다.

GAN 의 학습을 위해, 1 단계에서 D, G 는 각각 아래 손실함수를 최대화, 최소화하도록 학습한다 [2], [6].

$$\mathcal{L}_D = B(D(n), H_0) + B(D(y), H_1), \tag{2.2}$$

$$\mathcal{L}_G = \lambda \cdot B(D(y), H_0) + \mu \cdot B(B(y), m), \tag{2.3}$$

여기서, $B(\cdot,\cdot)$ 는 Binary Cross-Entropy(BCE) 손실 함수이다. 이후, 2 단계에서는 G 의 파라미터를 고정한 상태에서 B 가 메시지 복원 오차를 줄이기 위해 아래의 손실 함수를 최소화하도록 학습한다.

$$\mathcal{L}_B = B(B(y), m). \tag{2.4}$$

Ⅲ. 실험

실험 환경은 메시지 길이 $M = 16$, 신호 길이 $N = 64$, 은닉층 2 개, 각 은닉층 뉴런 수 128, 평균 송신 전력 $P_s = 1$로 설정하였다. $\mathbf{n}$은 평균 0, 분산 σ^2을 따르는 AWGN 벡터 $\mathbf{n} \sim N(\mathbf{0}, \sigma^2\mathbf{I})$, SNR 은 P_s/σ^2로 정의하여 -10dB 부터 2dB 의 범위에서 실험하였다. [3]에서는 $\lambda = \mu = 1$ 로 설정하였지만, 본 논문(Proposed)은 $\lambda \in [0.2, 1.5]$, $\mu \in [0.5, 1.5]$ 범위에서 2 단계 학습을 적용하여 실험한다. 또한, 성능 지표 중 유효

전송률(Effective data rate, R_{eff})은 다음과 같이 정의한다.

$$R_{eff} = \frac{M}{N}(1 - BER). \quad (3.1)$$

그림 1 은 Proposed 와 [3]의 성능을 비교한 것으로 구체적인 수치 성능 비교는 표 1 과 같다. 실험 결과, 그림 (a)와 (c)에서 볼 수 있듯이, Proposed 는 전 SNR 구간에서 [3]과 일치하는 수준의 $\mathcal{D}_{KL}$과 R_{eff}를 보였다. 반면, 그림 (b)의 BER 의 경우 Proposed 가 [3] 대비 다소 증가하는 경향을 보였으나, 표 1 의 결과와 같이 전반적으로 [3]의 성능 수치에서 크게 벗어나지 않고 안정적인 값을 유지하였다. 이러한 실험 결과는 제안 기법이 매번 최적의 매개변수를 탐색하지 않더라도 일정 수준 이상의 은닉성과 복원 성능을 보장함을 보여준다.

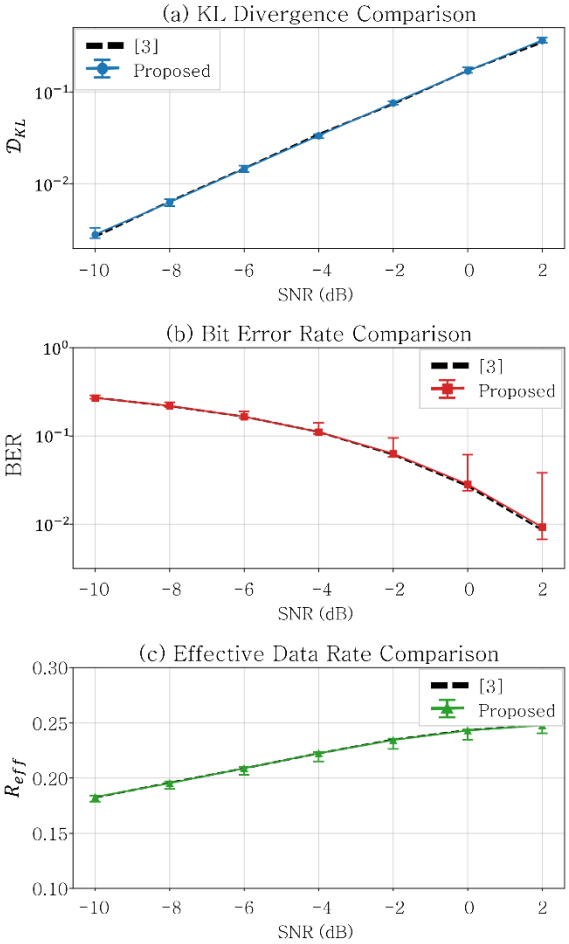

그림 1. Proposed 와 [3]의 성능 비교:
(a) SNR 에 따른 $\mathcal{D}_{KL}$, (b) SNR 에 따른 BER,
(c) SNR 에 따른 R_{eff}

Ⅳ. 결론

본 논문에서는 GAN 구조에서 2 단계 학습 기반 은밀 통신 신호 생성 기법을 제안한다. 해당 기법은 매개변수에 의존적이지 않은 강인한 은밀 통신 설계를 목표로 하며, 실험 결과 다양한 매개변수에서도 은닉성과 복원 성능이 안정적으로 유지됨을 입증하였다. 차후에는 BER 성능을 개선하고 매개변수에 의존적이지 않은 GAN 을 설계하고, 제안한 구조를 바탕으로 해양, 위성환경 등 현실적인 통신 환경에 적합한 은밀 통신을 제안할 예정이다.

표 1
λ와 μ에 따른 성능 결과

Method	λ	μ	$SNR(dB)$	$\mathcal{D}_{KL}$	BER	R_{eff}
[3]	1	1	-10	0.0028	0.2690	0.1828
			-4	0.0332	0.1109	0.2223
			2	0.3627	0.0085	0.2479
Proposed	0.6	0.5	-10	0.0029	0.2713	0.1822
			-4	0.0329	0.1113	0.2222
			2	0.3728	0.0087	0.2478
Proposed	1	1	-10	0.0032	0.2676	0.1831
			-4	0.0331	0.1065	0.2234
			2	0.3646	0.0071	0.2482
Proposed	1.5	1.5	-10	0.0028	0.2678	0.1831
			-4	0.0329	0.1089	0.2228
			2	0.3560	0.0078	0.2480

ACKNOWLEDGMENT

본 논문은 2022 년도 정부(방위사업청)의 재원으로 국방기술진흥연구소 (KRIT-CT-22-040, 이종 위성군 우주감시정찰 기술 특화연구센터)의 지원을 받아 수행된 연구입니다.

참 고 문 헌

[1] C. Zhao, X. Chen, J. An, Z. Xiong, N. Zhao, D. Niyato, and F. R. Yu, “Generative AI for secure physical layer communications: A survey,” IEEE Transactions on Cognitive Communications and Networking, vol. 11, no. 1, pp. 3–23, Feb. 2025.

[2] I. Goodfellow, J. Pouget-Abadie, M. Mirza, et al., “Generative adversarial networks,” in Advances in Neural Information Processing Systems 27, pp. 2672–2680, 2014.

[3] A. Ali, M. J. Piran, and H. Arslan, “Stealth signals: Multi-discriminator GANs for covert communications against diverse wardens,” arXiv preprint arXiv:2505.00399, May 2025.

[4] W. Xu, R. Zhu, and X. Ji, “Covert communication scheme for OOK in asymmetric noise systems,” Sensors, vol. 25, no. 8, article 2948, 2025.

[5] G. E. Hinton and R. R. Salakhutdinov, “Reducing the dimensionality of data with neural networks,” Science, vol. 313, no. 5786, pp. 504–507, Jul. 2006.

[6] R. Mehmood, R. Bashir, and K. J. Giri, “Mathematical analysis of loss function of GAN and its loss function variants,” International Journal of Advanced Technology and Engineering Exploration, vol. 9, no. 94, pp. 1327–1348, 2022.