Adversarial AI-Driven Covert Communications in Bistatic Backscatter Systems: An Overview

Hao Hoang Tran, Quang Tuan Do, Tung Son Do, Junsuk Oh, and Sungrae Cho
School of Computer Science and Engineering
Chung-Ang University
Seoul 06974, South Korea
Email: {hhtran, dqtuan, tsdo, jsoh}@uclab.re.kr, srcho@cau.ac.kr

***Abstract*—Bistatic backscatter communication is a promising technology for low-power Internet of Things (IoT) applications. However, its broadcast nature poses significant security risks. Covert communication aims to mitigate these risks by concealing the very existence of a transmission. This paper provides an overview of the evolution of covert communication schemes in bistatic backscatter systems, from classical analytical models to modern adversarial AI frameworks. We first review the foundational approach, where system parameters are optimized against a simple, model-based warden. We then detail the paradigm shift required to counter an intelligent warden that employs machine learning (ML) for detection. We propose a framework based on Generative Adversarial Networks (GANs) and Reinforcement Learning (RL), where a legitimate RL agent learns to generate signal characteristics that are indistinguishable to an advanced, data-driven warden. By surveying key literature, we compare these approaches, highlighting the necessity of AI-driven strategies for ensuring robust covertness in future wireless networks.**

***Index Terms*—Bistatic Backscatter, Covert Communication, Reinforcement Learning, Generative Adversarial Network (GAN), Physical Layer Security.**

I. Introduction

Bistatic backscatter communication is a cornerstone technology for energy-efficient Internet of Things (IoT) networks by enabling passive tags to transmit data by reflecting signals from a dedicated carrier emitter [1]. While effective, this paradigm's broadcast nature creates a significant vulnerability: an adversarial warden can perform traffic analysis to detect ongoing transmissions, compromising user privacy and security. To address this, covert communication seeks to hide the very existence of a transmission, making the signal statistically indistinguishable from ambient noise to a warden [3].

The design of such covert schemes has undergone a significant evolution, moving from predictable analytical models to adaptive AI frameworks. Foundational works approached the problem by assuming a classical warden that uses a simple power detector (radiometer). Within this model, system parameters like transmit power and a tag's reflection coefficient could be analytically optimized to maximize the warden's detection error [1]. However, this approach is not robust against a modern, intelligent adversary that employs machine learning (ML) for detection [2]. Countering this advanced threat requires a paradigm shift to an adversarial framework, naturally modeled by Generative Adversarial Networks (GANs) [9]. In this new paradigm of this paper, the legitimate system is framed as a Reinforcement Learning (RL) agent that learns a policy to create covert signal waveforms that can fool the warden's ML-based classifier. This paper provides a structured overview of this critical evolution, comparing the classical and adversarial AI paradigms and outlining future research directions.

II. Classical Model-Based Approach to Covertness

The foundational work in covert bistatic backscatter systems relies on precise mathematical modeling and analytical optimization. This approach provides fundamental insights and performance bounds under specific assumptions.

A. System Model and Warden's Detection

The foundational model for covert bistatic backscatter communication, as established in [1], comprises four key entities: a dedicated Carrier Emitter (CE), a passive information-bearing Tag, a legitimate Reader, and an adversarial Warden. The system model is illustrated in Fig. 1. The communication process is initiated by the CE, which transmits a carrier signal, potentially embedded with artificial noise (AN), to create channel uncertainty. The passive Tag leverages this incident energy to transmit its own information by modulating its reflection coefficient. The Reader is tasked with decoding this modulated, backscattered signal while canceling the interference from the CE.

Concurrent to this legitimate communication, the Warden's objective is to determine whether the Tag is active. This task is framed as a binary hypothesis testing problem, where the Warden must decide between two possible states:

$$\mathcal{H}_0: \text{ The Tag is silent (signal is absorbed).} \quad (1)$$

$$\mathcal{H}_1: \text{ The Tag is active (signal is reflected).} \quad (2)$$

The fundamental challenge for the legitimate system is to ensure the Warden cannot reliably distinguish $\mathcal{H}_1$ from $\mathcal{H}_0$.

AS Warden's perspective, To minimize detection error, it utilizes a Likelihood Ratio Test (LRT). Under the standard assumption that the channel noise and artificial noise follow Gaussian distributions, this LRT simplifies mathematically to an energy detection strategy. Consequently, the classical warden is modeled as a radiometer that compares the average received signal power against a fixed threshold τ.

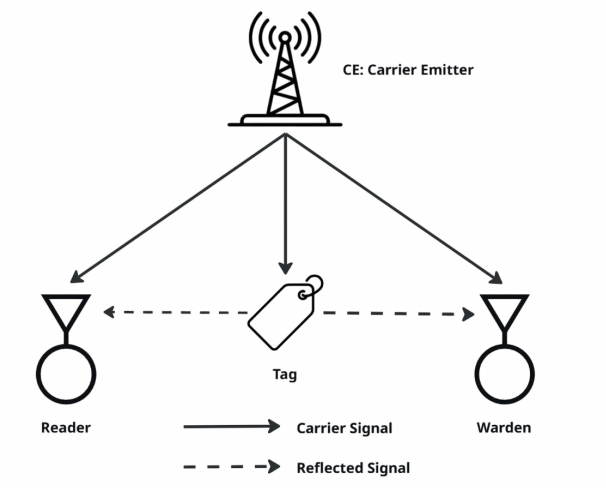

Fig. 1. System model for classical covert bistatic backscatter communication, adapted from [1]. The Warden attempts to detect the transmission from the Tag to the Reader.

In this classical setting, we assume:

- The Warden has perfect knowledge of the channel statistics (e.g., noise variance, average channel gain).
- The Warden does *not* know the instantaneous realization of the artificial noise generated by the CE.
- The environment is static, meaning channel statistics do not change during the detection window.

This reliance on fixed statistical models makes the radiometer analytically tractable but vulnerable to manipulation if the signal statistics deviate from the Warden's expectations.

B. Analytical Performance Optimization

The core of the classical approach is to leverage the well-defined system model to analytically optimize for energy efficiency while satisfying strict security and reliability requirements. This is achieved by first deriving closed-form expressions for the system's key performance indicators (KPIs) and then formulating a constrained optimization problem.

The two primary KPIs are defined from the perspectives of the adversary and the legitimate users, respectively:

- *Covertness Metric:* The effectiveness of the covert scheme is quantified by the Warden's minimum sum of error probabilities, ξ^*, which represents the highest level of uncertainty the legitimate system can induce at the Warden. As derived in [1], by exploiting the channel uncertainty introduced by the CE's artificial noise, a closed-form expression for ξ^* can be obtained. The goal is to forcing it as close as possible to the point of random guessing ($\xi^* \to 0.5$ for equal priors).
- *Reliability Metric:* The quality of the legitimate communication link is measured by the outage probability at the Reader, denoted as P_{out}. This metric captures the likelihood that the received signal-to-interference-plus-noise ratio (SINR) falls below the threshold required for successful decoding.

With these analytical metrics established, an optimization problem is formulated to achieve an energy-efficient design. The primary objective is to minimize the CE's transmit power, P, which is the main source of energy consumption in the system. The optimization variables are the CE's transmit power P and the Tag's reflection coefficient β. The problem is formally stated as:

$$\min_{P,\beta} P \tag{3}$$

$$\text{subject to } \xi^*(P,\beta) \geq 1-\epsilon, \tag{4}$$

$$P_{out}(P,\beta) \leq \delta, \tag{5}$$

$$0 \leq P \leq P_{max}, \quad 0 \leq \beta \leq 1. \tag{6}$$

Here, (4) represents the covertness constraint, where ϵ is the maximum tolerable detection probability by the Warden. Constraint (5) ensures the reliability of the legitimate link, where δ is the maximum allowable outage probability. Finally, (6) defines the physical constraints on the system variables. By solving this problem, one can find the optimal static operating point (P^*, β^*) that guarantees covert and reliable communication with minimal energy expenditure under the assumed classical warden model.

C. Limitations

Despite its analytical elegance, the classical approach is constrained by several fundamental limitations that curtail its applicability in realistic scenarios. Its primary drawback lies in the assumption of a simplistic and predictable adversary. By modeling the warden as a static, model-based power detector, the framework fails to account for intelligent adversaries that can leverage machine learning to learn and adapt their detection strategies from observed data. Furthermore, the model is inherently static, neglecting the impact of dynamic channel conditions, mobility, and other real-world environmental factors. This, coupled with its analytical complexity, renders the approach intractable for large-scale IoT scenarios involving multiple heterogeneous wardens or numerous communicating tags. These constraints collectively underscore the need for a new paradigm—one that is adaptive, scalable, and resilient enough to counter intelligent threats in dynamic wireless networks.

III. Adversarial AI Framework for Enhanced Covertness

The classical model's reliance on a predictable, non-learning warden creates a significant vulnerability. A modern, sophisticated adversary will not use a simple, static power detector but will instead employ data-driven Machine Learning (ML) techniques to identify subtle statistical patterns in the electromagnetic spectrum, rendering analytical covertness schemes ineffective [2]. To counter this intelligent threat, the legitimate system must also become intelligent, shifting the paradigm from a static optimization problem to a dynamic, adversarial game. This section details a framework for achieving robust covertness based on the principles of Generative Adversarial Networks (GANs) and Reinforcement Learning (RL).

A. The Intelligent Warden as a Deep Learning Discriminator

In the adversarial AI framework, the threat model evolves significantly. The warden is no longer a simple radiometer constrained by fixed statistical assumptions. Instead, it acts as a Deep Learning (DL) classifier (Discriminator) capable of extracting complex features from the raw I/Q signal samples. Its objective is to analyze the raw received signal and distinguish a covert transmission ($\mathcal{H}_1$) from ambient noise ($\mathcal{H}_0$), as identical to the discriminator in a GAN [9]. The warden's intelligence is embodied in a Deep Neural Network (DNN), typically a 1D Convolutional Neural Network (CNN) optimized for time-series analysis, as conceptually illustrated in Fig. 2.

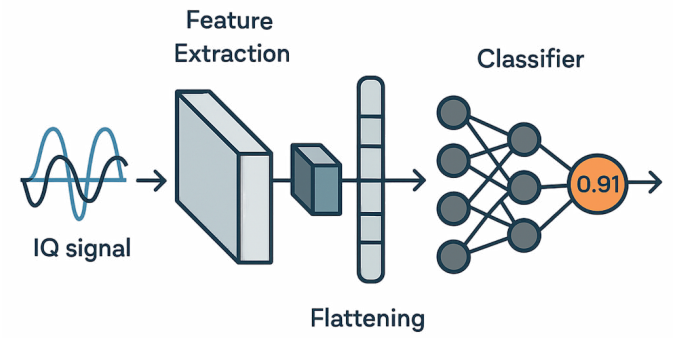

Fig. 2. Conceptual architecture of the intelligent Warden's DNN. The network processes the raw I/Q signal through feature extraction (CNN layers) and classification (dense layers) to produce a probabilistic detection score.

The warden's decision-making process can be deconstructed into several key stages, as shown in the figure:

1) **Input Signal:** The process begins with the raw, high-dimensional I/Q signal received at the warden's antenna. This complex waveform contains both the potential covert signal and environmental noise, making simple energy-based distinctions unreliable.
2) **Feature Extraction:** The signal is first passed through a series of convolutional layers. This is the core of the warden's discriminative power. Unlike a classical radiometer which only measures aggregate power, the CNN automatically learns to extract subtle, high-level features and patterns—such as faint modulation traces, specific spectral signatures, or temporal correlations—that are characteristic of the legitimate user's transmission.
3) **Classification:** After the relevant features have been extracted and condensed, they are "flattened" into a vector and fed into a set of dense, fully-connected layers. This classifier stage weighs the extracted features to make a final judgment.
4) **Probabilistic Output:** The network's output is not a hard binary decision, but a probabilistic score (e.g., 0.91) representing its confidence that the input contained a covert signal. This continuous output is crucial, as it provides a granular feedback signal that the legitimate RL agent can use to calculate its reward and incrementally improve its own policy.

This sophisticated, multi-stage architecture empowers the warden to transcend simple energy detection and learn the specific statistical fingerprint of the transmitter, making it a far more formidable adversary. Specifically, we assume an advanced threat model where the warden possesses data-driven learning capabilities, utilizing access to labeled historical or synthetic signal datasets to train its neural network. Unlike a classical radiometer, this feature extraction process employs CNNs to identify subtle temporal or spectral correlations that distinguish the backscatter signal from noise. Furthermore, the warden operates in a model-free manner, learning decision boundaries directly from data without requiring prior knowledge of specific channel parameters. The challenge is further amplified in realistic scenarios with multiple, heterogeneous wardens [2], where each adversary may possess distinct channel conditions, hardware capabilities, and classifier architectures. This collective intelligence represents a "worst-case" scenario for the legitimate user, as the adversary is capable of adapting to simple obfuscation techniques, thereby necessitating the employment of the adversarial training dynamic described below.

B. The Legitimate System as an RL Agent

To navigate this adversarial environment, the legitimate system is modeled as a Deep Reinforcement Learning (DRL) agent. Unlike static optimization, this agent interacts with the environment to learn a dynamic policy π that maps states to actions. The formulation is defined as follows:

- **State Space ($\mathcal{S}$):** The state vector $\mathbf{s}_t$ encapsulates the agent's environmental awareness at time t. It typically comprises the estimated Channel State Information (CSI) of the forward and backscatter links, the residual energy level of the tag (if battery-powered), and the historical decoding feedback (ACK/NACK) from the Reader.
- **Action Space ($\mathcal{A}$):** To allow for fine-grained control, the agent operates in a continuous action space. The action vector $\mathbf{a}_t = [P_t, \beta_t]$ consists of the Carrier Emitter's transmit power P_t and the Tag's reflection coefficient β_t. Continuous control algorithms, such as Deep Deterministic Policy Gradient (DDPG), are essential here.
- **Reward Design (R):** The reward function is the critical driver of behavior, balancing three competing objectives: reliability, covertness, and energy efficiency. The immediate reward r_t is formulated as:

$$r_t = \underbrace{w_1 \log_2(1+\gamma_R)}_{\text{Reliability (Rate)}} + \underbrace{w_2 \log(1-D(\mathbf{y}_w))}_{\text{Covertness (GAN Loss)}} - \underbrace{w_3 P_t}_{\text{Energy}} \tag{7}$$

where γ_R is the SINR at the Reader, and $D(\mathbf{y}_w) \in [0,1]$ is the Warden's estimated probability that the signal $\mathbf{y}_w$ contains information. The first term encourages high data rates. The second term rewards the agent when the Warden misclassifies the transmission as noise (i.e.,

$D(\mathbf{y}_w) \to 0$). The third term penalizes energy consumption.

C. Adversarial Training Loop and Stability Considerations

The core of the proposed framework is the minimax adversarial training loop, which embodies the dynamics of a Generative Adversarial Network (GAN) and establishes a competitive interplay between the legitimate reinforcement learning (RL) agent (Generator) and the machine learning–based Warden (Discriminator), as depicted in Fig. 3. The training progresses in an alternating manner. First, the Warden's deep neural network is trained on labeled data comprising genuine noise samples ($\mathcal{H}_0$) and signal samples produced by the current policy ($\mathcal{H}_1$), optimizing binary cross-entropy loss to maximize classification accuracy. Subsequently, the Warden's parameters are frozen, and the RL agent updates its policy network to maximize the expected cumulative reward r_t, which integrates both communication performance and covertness objectives. The covertness component is informed by the gradient feedback from the Warden's discriminator, effectively guiding the agent toward generating signal constellations that induce maximum uncertainty in the Warden's detection process.

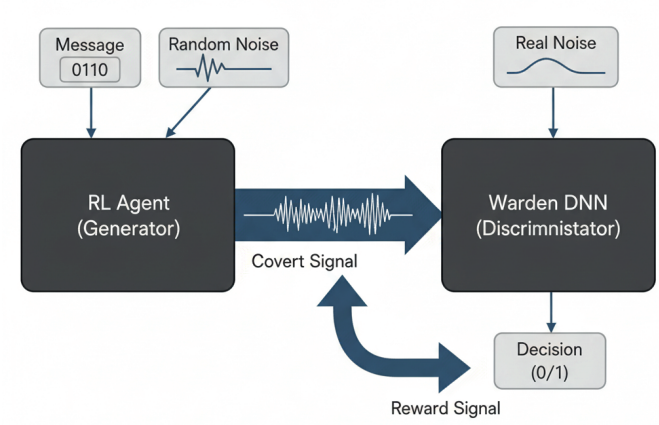

Fig. 3. Conceptual illustration of the adversarial training loop. The RL Agent (Generator) learns to produce signals that deceive the Warden's DNN (Discriminator), which, in turn, continuously improves its detection capability.

This alternating optimization constitutes a two-player, zero-sum game, wherein the Generator endeavors to minimize the Discriminator's classification accuracy, while the Discriminator simultaneously seeks to maximize it. Theoretical convergence is achieved when the system reaches a Nash equilibrium, corresponding to an operational state in which the Generator's outputs become statistically indistinguishable from background noise. In this equilibrium, the Warden's performance approaches random guessing, and the detection error probability converges to its upper bound, representing perfect covertness. Through this adversarial dynamic, the RL agent evolves beyond naive noise injection strategies, discovering non-trivial, statistically grounded signal generation policies that optimize covert communication robustness.

Despite its capability, adversarial training is inherently unstable and susceptible to phenomena such as mode collapse, wherein the agent's policy converges to a restricted subset of signal patterns. To enhance training stability and ensure reliable convergence, the framework incorporates two critical mechanisms. First, **experience replay** is employed, wherein transitions $(\mathbf{s}_t, \mathbf{a}_t, r_t, \mathbf{s}_{t+1})$ are stored in a replay buffer and randomly sampled to decorrelate training data, thus stabilizing gradient updates. Second, **target networks** are utilized as slowly updated copies of the actor and critic models to compute target values, mitigating oscillations induced by rapidly changing network parameters and promoting convergence in continuous state–action spaces.

These stability components render the proposed framework both scalable and extensible to more sophisticated covert communication paradigms. In particular, multi-discriminator architectures have been demonstrated to further strengthen covert resilience by enabling a single Generator to simultaneously deceive multiple heterogeneous Warden models [2]. Such extensions highlight the adaptability of the GAN-inspired training paradigm, which has been effectively applied to diverse physical-layer security challenges, including intelligent jamming for UAV-assisted satellite communications [4] and secure transmission design in cooperative cognitive radio networks [6].

IV. Comparative Analysis and Discussion

The transition from a classical, analytical framework to an adversarial AI-driven approach marks a fundamental evolution in designing secure communication systems. This section provides a comparative analysis of the two paradigms, highlighting the distinct advantages and trade-offs inherent in each methodology. A direct comparison is summarized in Table I.

The classical approach offers mathematical elegance and provides theoretical performance bounds, but its real-world applicability is constrained by its rigid assumptions. Its primary strength lies in establishing a fundamental understanding of the problem under idealized conditions, defining the baseline against which more advanced systems can be measured.

In stark contrast, the adversarial AI framework sacrifices analytical tractability for a profound increase in robustness, adaptability, and scalability. By learning directly from data and interaction, it makes no strong assumptions about the warden's internal strategy, other than its ability to learn. This model-free nature allows it to discover novel and complex covert signaling strategies that would be impossible to derive analytically. For instance, instead of merely optimizing power levels, the RL agent can learn to craft specific waveform shapes or artificial noise structures that exploit the learned "blind spots" of the warden's DNN classifier.

Furthermore, the AI approach is inherently better suited for the complexities of modern wireless environments. The framework presented in [2] demonstrates that a single generator can learn to be covert against a swarm of diverse, intelligent wardens—a scenario that is analytically intractable but is a natural extension of the multi-discriminator GAN architecture.

TABLE I
COMPARATIVE ANALYSIS OF COVERT COMMUNICATION FRAMEWORKS

Feature	Classical Analytical Approach [1]	Adversarial AI Approach [2]
Warden Model	Assumes a simple, static power detector (LRT-based radiometer). The warden's strategy is fixed, predictable, and reliant on perfect channel knowledge.	Models an intelligent, adaptive Deep Learning classifier (Discriminator). The warden's strategy is data-driven, evolving, and capable of learning from observed signals.
Methodology	Relies on deriving closed-form mathematical expressions for error probabilities and solving a constrained optimization problem.	Employs an adversarial training loop (GAN/RL) where the agent learns an optimal policy through continuous interaction and feedback (trial-and-error).
Solution Form	A set of optimal static parameters, typically the carrier emitter's transmit power (P^*) and the tag's reflection coefficient (β^*).	A trained neural network policy that dynamically generates optimal signal waveforms and power levels based on the current system state.
Adaptability	Low. The solution is fixed to the initial system assumptions. It is brittle and fails if the environment changes or if the warden adopts a new strategy.	High. The agent can continuously adapt its policy online to counter new warden strategies, mobility, and changing channel conditions via Meta-Learning.
Scalability	Poor. The analytical complexity becomes mathematically intractable when considering multiple tags or a swarm of heterogeneous wardens.	High. The framework naturally scales to multiple adversaries by incorporating them as additional discriminators in the multi-discriminator GAN architecture.

This scalability is crucial for securing large-scale IoT deployments where multiple potential adversaries may be present. While the AI framework introduces its own challenges, such as training complexity and the need for sufficient data (either real or simulated), its ability to address the core limitations of the classical model makes it the definitive path forward for achieving practical and resilient covert communications in contested environments.

V. FUTURE RESEARCH DIRECTIONS

The adversarial AI framework represents a significant leap forward in achieving robust covert communications. However, as adversaries evolve and network demands grow, several cutting-edge research directions emerge. This section outlines key frontiers that will shape the future of intelligent and secure backscatter systems.

A. Meta-Learning for Rapid Adversarial Adaptation

The current adversarial training paradigm produces an RL agent optimized against a specific set of warden classifiers. A critical limitation is the agent's inability to adapt quickly if the warden drastically changes its detection strategy or if the channel environment shifts unexpectedly. Extensive retraining in real-time is often infeasible. A highly promising solution lies in *Meta-Reinforcement Learning*, as explored in [8]. The goal is to train an agent not merely to master one specific covert policy, but to "learn how to learn." By training across a wide distribution of simulated warden types and channel models, a meta-trained agent can rapidly adapt its transmission policy to a novel, unseen adversary with only a handful of interactions. This capability transforms the legitimate system from a reactive entity into a proactive one, capable of maintaining covertness in a highly dynamic arms race.

B. Resource-Efficient AI for Edge Deployment

A primary challenge in transitioning these advanced AI frameworks from theory to practice is their implementation on resource-constrained hardware. The computationally intensive nature of DNNs and RL algorithms is fundamentally at odds with the low-power design philosophy of backscatter tags. The initial goal of energy efficiency [1] must be revisited in the AI context. Future work must focus on developing resource-efficient AI solutions, such as distributed and adaptive communication frameworks designed specifically for heterogeneous IoT environments [17]. This includes research into lightweight neural architectures and model compression techniques to reduce the computational footprint. Furthermore, drawing inspiration from federated learning (FL) security [5], a federated adversarial learning approach could be explored. In such a system, multiple legitimate tags could collaboratively train a powerful global covert policy using energy-efficient, DDPG-based algorithms, similar to those developed for modern federated IoT networks [16].

C. Multi-Agent Systems for Cooperative Covertness

Real-world networks are rarely single-user systems. Future research must extend the framework to complex multi-agent environments to leverage cooperative dynamics. As demonstrated in cognitive radio [6] and D2D systems [11], friendly jammers or relays can significantly enhance covertness. To manage this complexity, Multi-Agent Reinforcement Learning (MARL) is essential. A prime example of this is seen in autonomous systems where multiple agents learn to cooperate to achieve a common goal, such as reliable surveillance [13]. In the context of covertness, a MARL framework would enable agents (tags, CEs, and jammers) to learn optimal cooperative strategies, deciding not only on their individual transmission policies but also on how to best assist their peers through intelligent jamming or relaying, potentially managed by robust auction mechanisms [12].

D. Joint Trajectory and Signal Optimization

Finally, the agent's action space can be dramatically expanded beyond signal modulation. In mobile scenarios involving UAVs acting as CEs or tags [4], the RL agent

could learn to jointly optimize its physical trajectory and its signal characteristics . This allows the agent to physically maneuver to locations that are simultaneously advantageous for communication with the reader and disadvantageous for the warden's detection. This concept, which has been successfully demonstrated in learning-based cooperative mobility control for autonomous drones [14], adds a physical dimension to the adversarial game, forcing the warden to contend with both a changing signal and a changing physical environment. Finally, as surveyed in [3], a GAI-powered agent could learn to transmit a signal that is semantically valid and decodable by the warden but contains innocuous information, while the true, secret message is embedded in subtle, goal-oriented features only decodable by the intended reader. This would render traditional signal detection-based wardens entirely obsolete.

VI. Conclusion

This paper has surveyed the progression of security techniques for covert communications in bistatic backscatter systems. We began with the classical, model-based optimization approach, which provides valuable theoretical insights but is brittle against intelligent adversaries. We then detailed the necessary evolution to an adversarial AI framework, where a legitimate RL agent learns to generate covert signals against one or more ML-powered wardens. This GAN-inspired approach transforms the problem from a static optimization to a dynamic, adaptive arms race. By leveraging the power of generative models, this framework provides a scalable and robust path toward securing the next generation of low-power IoT networks.

VII. Acknowledgement

This work was supported by the IITP (Institute of Information & Communications Technology Planning & Evaluation) - ITRC (Information Technology Research Center) (IITP-2026-RS-2022-00156353, 50% / IITP-2026-RS-2023-00258639, 50%) grants funded by the Korea government (Ministry of Science and ICT).

References

[1] Y. Wang, S. Yan, W. Yang, Y. Huang, and C. Liu, "Energy-Efficient Covert Communications for Bistatic Backscatter Systems," *IEEE Transactions on Vehicular Technology*, vol. 70, no. 3, pp. 2906-2911, Mar. 2021.

[2] A. Ali, Md. J. Piran, and H. Arslan, "Stealth Signals: Multi-Discriminator GANs for Covert Communications Against Diverse Wardens," *arXiv preprint arXiv:2505.00399*, 2025.

[3] C. Zhao, H. Du, D. Niyato, J. Kang, Z. Xiong, D. I. Kim, X. Shen, and K. B. Letaief, "Generative AI for Secure Physical Layer Communications: A Survey," *IEEE Transactions on Cognitive Communications and Networking*, vol. 11, no. 1, pp. 3-26, Feb. 2025.

[4] S. Jia, L. Xiaomeng, L. Xiaomin, T. Zhuangzhuang, and H. Junfan, "Covert LEO Satellite Communication Aided by Generative Adversarial Network Based Cooperative UAV Jamming," *China Communications*, vol. 21, no. 9, pp. 27-39, Sep. 2024.

[5] Y. Feng, Y. Jiang, and Y. Wang, "GAN-Based Covert Communications Against an Adversary with Uncertain Detection Threshold in Federated Learning Networks," in *2023 International Conference on Networking and Network Applications (NaNA)*, 2023, pp. 613-618.

[6] Y. Wen, Y. Huo, J. Li, J. Qian, and K. Wang, "Generative Adversarial Network-Aided Covert Communication for Cooperative Jammers in CCRNs," *IEEE Transactions on Information Forensics and Security*, vol. 20, pp. 1278-1291, 2025.

[7] Q. Yang, H.-M. Wang, Y. Zhang, and Z. Han, "Physical Layer Security in MIMO Backscatter Wireless Systems," *IEEE Transactions on Wireless Communications*, vol. 15, no. 11, pp. 7547-7560, Nov. 2016.

[8] N. H. Chu, N. V. Huynh, D. N. Nguyen, D. T. Hoang, S. Gong, T. Shu, E. Dutkiewicz, and K. T. Phan, "Countering Eavesdroppers With Meta-Learning-Based Cooperative Ambient Backscatter Communications," *IEEE Transactions on Wireless Communications*, vol. 23, no. 10, pp. 13678-13693, Oct. 2024.

[9] K. S and M. Durgadevi, "Generative Adversarial Network (GAN): a general review on different variants of GAN and applications," in *2021 6th International Conference on Communication and Electronics Systems (ICCES)*, 2021, pp. 1-8.

[10] L. Gonog and Y. Zhou, "A Review: Generative Adversarial Networks," in *2019 14th IEEE Conference on Industrial Electronics and Applications (ICIEA)*, 2019, pp. 505-510.

[11] S. Feng, X. Lu, S. Sun, D. Niyato, and E. Hossain, "Securing Large-Scale D2D Networks Using Covert Communication and Friendly Jamming," *IEEE Transactions on Wireless Communications*, vol. 23, no. 1, pp. 592-606, Jan. 2024.

[12] I. Lotfi, H. Du, D. Niyato, S. Sun, and D. I. Kim, "On the Robustness of Channel Allocation in Joint Radar and Communication Systems: An Auction Approach," *IEEE Transactions on Mobile Computing*, vol. 23, no. 4, pp. 3466-3483, Apr. 2024.

[13] W. J. Yun, D. Kwon, M. Choi, J. Kim, G. Caire, and A. F. Molisch, "Cooperative Multiagent Deep Reinforcement Learning for Reliable Surveillance via Autonomous Multi-UAV Control," *IEEE Transactions on Industrial Informatics*, vol. 18, no. 10, pp. 7086-7096, Oct. 2022.

[14] S. Park, C. Park and J. Kim, "Learning-Based Cooperative Mobility Control for Autonomous Drone-Delivery," *IEEE Transactions on Vehicular Technology*, vol. 73, no. 4, pp. 4870-4885, April 2024.

[15] D. Kwon, J. Jeon, S. Park, J. Kim and S. Cho, "Multiagent DDPG-Based Deep Learning for Smart Ocean Federated Learning IoT Networks," *IEEE Internet of Things Journal*, vol. 7, no. 10, pp. 9895-9903, Oct. 2020.

[16] M. C. Ho, A.-T. Tran, D. Lee, J. Paek, W. Noh, and S. Cho, "A DDPG-based Energy Efficient Federated Learning Algorithm with SWIPT and MC-NOMA," *ICT Express*, vol. 10, no. 3, pp. 600-607, Jun. 2024.

[17] J. Oh, D. Lee, D. S. Lakew, and S. Cho, "DACODE: Distributed Adaptive Communication Framework for Energy Efficient Industrial IoT-based Heterogeneous WSN," *ICT Express*, vol. 9, no. 6, pp. 1085-1094, Dec. 2023.