Context-Aware Evasion of PMU-Based Intrusion Detection Systems in Wide-Area Control

Muneeba Asif*, Prabin Mali†, Sumit Paudyal† and Mohammad Ashiqur Rahman*
*Analytics for Cyber Defense (ACyD) Lab, †Power System Computational Laboratory
Florida International University, USA, {masif004, pmali004, spaudyal, marahman}@fiu.edu

***Abstract*—Phasor measurement units (PMUs) improve wide-area control by providing synchronized, high-rate measurements for grid awareness. Recent learning-based intrusion detection on PMU streams includes recurrent and convolutional models, and more recent adversarially trained GAN-LSTM detectors that report stronger robustness under gradient tests and PMU-specific patterns. However, existing work focuses largely on gradient-based and white-box scenarios, and rarely evaluates an adaptive black-box adversary that updates inputs using only detector scores while keeping changes small, physically plausible, and within reporting latency. To address this gap, we model evasion as a query-based black-box problem with limited feedback and adopt reinforcement learning (RL) to learn a context-aware attack policy when gradients are not available and the objective must balance higher detector output with small, temporally smooth changes. In particular, we use proximal policy optimization (PPO) with an eight-parameter grouped-action policy that adjusts four PMU feature groups via one rescaling level and one sign-control level per group. Validation on the IEEE 118-bus system, having a GAN-LSTM detector with simulation on OPAL-RT, shows that the learned policy increases the fraction of attacks classified as benign from about 5% to about 95% at a median-benign threshold while keeping perturbations modest and smooth.**

***Index Terms*—intrusion detection system evasion, adversarial attacks, black-box attacks, proximal policy optimization, reinforcement learning, WAMPAC, cybersecurity, smart grid**

I. Introduction

A. Motivation

Wide-area monitoring, protection, and control (WAMPAC) systems rely on phasor measurement units (PMUs) for time-synchronized measurements that support situational awareness, estimation, and fast control. Hence, the integrity and timing of PMU streams directly affect system stability. Almas et al. experimentally assess time-synchronization spoofing and show that GPS-based phase-angle errors can impair wide-area applications [1], while Chen et al. review WAMPAC cybersecurity and show that false-data injection and timing attacks can compromise estimation and control, motivating defenses beyond classical bad-data tests [2]. Consequently, learning-based detection for synchrophasor data has been explored. Aligholian et al. develop a generative model for event and anomaly detection in streaming PMU data [3]. More recent evaluations benchmark detectors for PMU-oriented false-data injection and find that traditional residual checks are often insufficient for stealthy variants: Sahu et al. study LSTM and GNN-LSTM state-prediction detectors, Bitirgen et al. report gains with a hybrid CNN–LSTM for FDIA discrimination, and Asif et al. introduce a GAN–LSTM IDS tailored to PMU-enhanced WAMPAC operations [4]–[6]. These works illustrate a shift to temporal and spatio-temporal modeling for defense.

Nevertheless, detectors trained on gradients or static residuals remain vulnerable to adaptive adversaries, where limited-query attacks have been shown in both general ML and time-series studies to reliably fool temporal classifiers without access to gradients [7], [8]. Furthermore, recent PMU-specific work designs structured perturbations that exploit these properties and evaluate the robustness of learning-based detectors, including GAN-LSTM IDS, under both gradient-driven and design-time settings [6]. PMU data have temporal correlations, phase coupling, topology structure, and strict reporting latencies. These properties allow small, smooth, and physically plausible changes to stay stealthy while still influencing closed-loop behavior. However, most WAMPAC studies evaluate white-box or gradient-driven attacks and do not consider adaptive black-box adversaries that update inputs using only detector outputs while respecting reporting rates, residual budgets, and placement constraints. The resulting gap is the lack of systematic evaluation and defense against query-based, control-aware evasion under closed-loop timing and limited feedback. This is crucial as WAMPAC operates at sub-second timescales, so temporally coherent manipulations can shift damping or remedial-action timing without alarms.

To address this gap, we propose a reinforcement learning (RL) driven context-aware evasion framework that operates under limited feedback, where the attacker observes the detector output at each step and must respect PMU reporting rates and residual budgets. This setting reflects practical WAMPAC conditions, where adversaries have little internal knowledge, face tight latency, and must keep perturbations small and physically plausible. The proposed approach trains a proximal policy optimization (PPO) agent with a grouped-action design over four PMU feature groups. Each group features rescaling and sign control, and the objective balances higher benign classifications with penalties on magnitude and temporal roughness. During training, latency limits and plausibility bounds are enforced so that learned behaviors align with closed-loop operation. This setup provides a systematic assessment of query-based, control-aware evasion that existing evaluations lack. Evaluation of the IEEE 118-bus system in OPAL-RT with a GAN-LSTM IDS reveals significant gains in benign classifications of attack data at median thresholds, while maintaining modest and smooth perturbations. We also

979-8-3315-8138-1

study query budgets, generalizability of proposed framework to alternative IDSs, and discuss practical operator defenses.

B. Related Works

Evasion of PMU-driven analytics in wide-area settings has been demonstrated in both power-system and time-series security studies. Biswal et al. showed that coordinated, small PMU perturbations can mislead a wide-area ML supervisory layer in a black-box regime, establishing the feasibility of query-constrained evasion in WAMPAC workflows [9]. In parallel, black-box attacks with limited queries have matured in the general machine learning (ML) literature, where bandit and zeroth-order methods achieve strong success with partial feedback and no gradients; these results motivate feedback-limited strategies relevant to operational grids [7]. Furthermore, the time-series community has confirmed that query-based black-box attacks can reliably fool temporal classifiers, hence, showcasing the vulnerability of streaming detectors that resemble PMU analytics [8].

Beyond generic evasion, several works have developed system-aware FDIAs that respect the grid structure or limit the attacker's knowledge. Higgins et al. leveraged topology learning to mount stealthy attacks under partial information, illustrating how physical constraints can guide undetected manipulations [10]. Other studies examine optimized FDIA constructions that trade off impact and stealth, further highlighting how constrained perturbations can bypass conventional checks and stress operations [11]. Moreover, evaluations focused on PMU measurements confirm that multiple FDIA variants can evade residual-based detection and degrade situational awareness, reinforcing the need to study adaptive evasion in wide-area contexts [12]. Furthermore, the power-systems community has begun to explore black-box attacks tailored to event classifiers, including hybrid, query-efficient methods that transfer across models [13]. Time-series security studies likewise report score-based black-box techniques that succeed with lightweight perturbations, further supporting the practicality of feedback-limited evasion against streaming models [14]. Finally, PMU-specific adversarial evaluations on event identification demonstrate that modest, structured noise can flip decisions in synchrophasor pipelines, providing concrete attack patterns for wide-area applications [15].

The existing works establish that (i) query-limited, gradient-free attacks are effective on temporal models and (ii) physics- and topology-aware FDIAs can remain stealthy under operational constraints; however, the literature lacks a unified, adaptive black-box evasion framework that (a) learns an online policy under limited feedback, (b) enforces PMU reporting rates and residual budgets during learning, and (c) is validated in closed-loop, real-time WAMPAC. Our study addresses this gap by training an RL policy that produces small, temporally smooth, and physically plausible perturbations under these constraints and by evaluating its query efficiency, transfer, and operator-oriented mitigations. To the best of our knowledge, this is the first systematic study of control-aware, query-limited RL evasion targeting PMU-driven IDSs in WAMPAC.

C. Contributions

This work makes the following key contributions:

- We formalize PMU-specific, query-based black-box evasion under reporting-rate, residual-budget, and plausibility constraints, and define an eight-parameter grouped-action space over four PMU feature groups (rescaling and sign per group) to enforce small, smooth perturbations.
- We develop a PPO-augmented RL-driven evasion agent for WAMPAC IDS that optimizes the reward combining detector's benign score, perturbation magnitude, and temporal variation, enabling adaptive attacks with only score feedback and real-time constraints.
- We validate the proposed framework on the IEEE 118-bus system with an OPAL-RT real-time simulator and a GAN-LSTM IDS, showing the learned policy raises attacks misclassified as benign from 5% to 95% at the median threshold while keeping perturbations modest and smooth; we also analyze query budgets, cross-model transfer, and simple operator mitigations, and discuss incorporating adversarial training with query-generated samples to improve IDS robustness.

II. System Model and Problem Formulation

This section formalizes the system and threat models, the grouped action space, and the MDP for our RL agent, aligned with the OPAL-RT evaluation on the IEEE 118-bus system.

A. System Model - WAMPAC Pipeline and Detector

We consider a WAMPAC pipeline in which N_p PMUs publish synchrophasors at a rate of f_r frames per second. At discrete time t, the (clean) feature vector is $x_t \in \mathbb{R}^d$, stacking per-PMU quantities (e.g., voltage/current magnitudes and angles, and ROCOF as used in our implementation). The detector $\mathcal{D}_\theta(\cdot)$ is the GAN–LSTM discriminator operating on a sliding window shown as follows:

$$X_t \triangleq [x_{t-k+1}, \ldots, x_t] \in \mathbb{R}^{k\times d} \tag{1}$$

and outputs a scalar score $s_t \in [0, 1]$ interpreted as a benign score. We adopt the standard convention from our prior training, where higher s_t indicates a more benign state. Thresholds (e.g., median) used in evaluation are defined in Section IV. Let $T_{\max}$ denote the end-to-end sensing-to-decision latency budget; our real-time loop respects this budget during evaluation. The attacker then adds δ_t to the clean frame x_t to form the attacked frame $\tilde{x}_t \triangleq x_t + \delta_t$, and the attacked window $\tilde{X}_t$. The detector is queried on $\tilde{X}_t$ to produce $s_t = \mathcal{D}_\theta(\tilde{X}_t)$.

B. Threat Model and Assumptions

In our OPAL-RT setup, the attacker observes only the GAN-LSTM window score per query and perturbs PMU stream under same reporting, and latency constraints as the system.

1) ***Goal:*** Increase the fraction of attack sequences classified as benign (push $s_t \geq \tau$ for an operating threshold τ) while keeping perturbations (i) small, (ii) temporally smooth, and (iii) physically plausible, under real-time constraints.

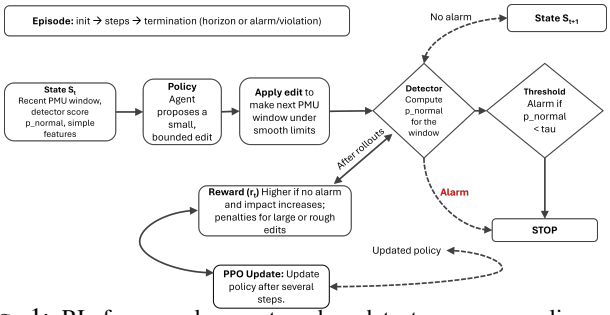

Fig. 1: RL framework: agent probes detector score, applies small grouped adjustments to data, and repeats until score crosses threshold.

*2) **Knowledge and Access**:* (i) *Black-box*: the adversary has no access to parameters or gradients of $\mathcal{D}_\theta$ and only observes scalar score s_t after each query; (ii) *Query-limited*: at most one score query per frame, consistent with the PMU reporting rate f_r; (iii) *Actionable controls*: perturbations are applied to the PMU feature stream immediately upstream of the detector.

*3) **Operational Constraints**:* The perturbation must satisfy:

$$\|\delta_t\|_2 \le \varepsilon \quad \text{(magnitude budget)} \tag{2}$$

$$\|\delta_t - \delta_{t-1}\|_2 \le \rho \quad \text{(temporal smoothness)} \tag{3}$$

$$\tilde{x}_t \in \mathcal{C} \quad \text{(per-feature plausibility bounds)} \tag{4}$$

$$\text{time}(\delta_t) \le T_{\max} - T_{\text{pipe}} \quad \text{(real-time budget)} \tag{5}$$

where $\mathcal{C}$ encodes admissible ranges and simple couplings (e.g., angle wrap and magnitude sanity), and T_{pipe} is the measured pipeline overhead excluding the action computation.

C. Grouped Action Space

To promote physical coherence and limit degrees of freedom, we partition features into four PMU groups,

$$\mathcal{G} = \{g_1, g_2, g_3, g_4\} \tag{6}$$

(e.g., voltage magnitude, angle, current magnitude, and angle/ROCOF etc.). At each t, the action selects, per group g, a rescaling level α_g and a sign control σ_g:

$$a_t = \big[(\alpha_{g_1}, \sigma_{g_1}),\ (\alpha_{g_2}, \sigma_{g_2}),\ (\alpha_{g_3}, \sigma_{g_3}),\ (\alpha_{g_4}, \sigma_{g_4})\big] \tag{7}$$

where $\alpha_g \in [\underline{\alpha}, \overline{\alpha}]$ and $\sigma_g \in \{-1, +1\}$. Update for group g is

$$\widehat{\delta}_t^{(g)} \ = \ \lambda\, \sigma_g\, \alpha_g \odot \phi^{(g)}(x_t) \ + \ \beta\, \delta_{t-1}^{(g)} \tag{8}$$

with step size $\lambda > 0$, smoothing factor $\beta \in [0, 1)$, and $\phi^{(g)}(\cdot)$ a group-specific normalized direction. λ controls the step size; β carries forward a fraction of the previous perturbation to encourage smoothness; $\phi^{(g)}(\cdot)$ gives a per-group direction in feature space. The final perturbation applies a projection to enforce the constraints (2)–(4):

$$\delta_t^{(g)} \ = \ \text{Proj}_{\varepsilon,\rho,\mathcal{C}}\big(\widehat{\delta}_t^{(g)}\big), \qquad \delta_t \ = \ \sum_{g \in \mathcal{G}} \delta_t^{(g)} \tag{9}$$

This yields eight action parameters per step (4 α_g and 4 σ_g).

D. RL Formulation - MDP Under Limited Feedback

We model the evasion as a PPO problem.

1) State: $o_t = \big[\tilde{X}_t,\ s_{t-1},\ b_t\big]$, where $\tilde{X}_t$ is the current attacked window, s_{t-1} is the last observed detector score, and b_t are optional budget indicators (normalized counters summarizing remaining magnitude/smoothness allowance). Consistent with our practice, $\tilde{X}_t$ is preprocessed using the same normalization as the detector.

2) Action: a_t is the 8-parameter grouped action in (II-C). In PPO, we implement this with mixed heads: four bounded continuous heads for α_g and four categorical heads for σ_g.

3) Transition: Given a_t, the environment computes δ_t via (8)–(9), forms $\tilde{x}_t = x_t + \delta_t$, advances the window to $\tilde{X}_{t+1}$, queries the detector to obtain $s_t = \mathcal{D}_\theta(\tilde{X}_t)$, and updates b_t, all within the per-frame time budget (5).

4) Reward: We shape rewards to increase benign classifications while penalizing size and variation:

$$\begin{aligned} R_t \ &= w_1\big(\mathbb{1}[s_t \ge \tau] + \eta\, s_t\big) \\ &\quad - w_2\, \|\delta_t\|_2 - w_3\, \|\delta_t - \delta_{t-1}\|_2 - w_4\, \mathbb{1}[\tilde{x}_t \notin \mathcal{C}] \end{aligned} \tag{10}$$

with tuning weights $w_i \ge 0$ and a small $\eta > 0$ for shaping. The objective is to maximize the discounted return $\mathbb{E}\big[\sum_t \gamma^t R_t\big]$ (Fig. 1). Rollout collection and action computation ensure time(δ_t) satisfies (5) at PMU rate f_r, while projection and grouped actions keep the OPAL-RT loop lightweight.

III. Proposed PPO-Based Evasion of WAMPAC IDS

The design of the proposed PPO-based RL agent uses mixed action heads, a shared encoder for the PMU window, and a fast projection operator to enforce budgets and plausibility.

A. Observation Encoding

The input $\tilde{X}_t \in \mathbb{R}^{k \times d}$ is normalized with the same scaler used for the GAN–LSTM IDS. We use a causal temporal encoder (stacked gated/conv or LSTM layer) to produce $h_t \in \mathbb{R}^m$. The final observation is $o_t = [h_t;\ s_{t-1};\ b_t]$.

B. Action Heads and Constraint Projection

The policy outputs an 8-parameter grouped action $a_t = [(\alpha_g, \sigma_g)]_{g \in \mathcal{G}}$, with

$$\begin{aligned} &\alpha_g = \alpha_{\max}\, \tanh\big(\mu_g + \sigma_g^{(\text{std})}\, \epsilon\big), \\ &\sigma_g \in \{-1, +1\}, \qquad \mathbb{P}(\sigma_g = +1) = p_g \end{aligned} \tag{11}$$

where $(\mu_g, \sigma_g^{(\text{std})}, p_g)$ are produced from o_t. A safety operator projects the candidate update to constraints (2)–(5):

$$\delta_t \ \leftarrow \ \text{Proj}_{\varepsilon,\rho,\mathcal{C}}\left(\sum_{g \in \mathcal{G}} \lambda\, \sigma_g\, \alpha_g \odot \phi^{(g)}(x_t) \ + \ \beta\, \delta_{t-1}\right) \tag{12}$$

C. PPO Objective and Losses

With advantage $\hat{A}_t$ (GAE), the clipped PPO objective is

$$\mathcal{L}_{\text{ppo}} \ = \ \mathbb{E}\Big[\min\Big(r_t(\theta)\, \hat{A}_t,\ \text{clip}\big(r_t(\theta),\, 1 - \epsilon,\, 1 + \epsilon\big)\, \hat{A}_t\Big)\Big] \tag{13}$$

where $r_t(\theta) = \pi_\theta(a_t|o_t)/\pi_{\theta_{\text{old}}}(a_t|o_t)$. The total loss is

$$L_{\text{total}} = L_{\text{ppo}} + c_v L_V - c_H \mathcal{H}[\pi_\theta] \tag{14}$$

with value loss L_V and entropy regularization $\mathcal{H}$. The reward R_t is formulated as previously shown in Eq. 10.

D. Training Procedure and Latency Handling

We train the PPO agent in fixed-length rollouts under the per-frame timing budget and projection constraints.

1) Rollouts and updates: Each iteration collects a horizon of H interaction steps, then performs K epochs of minibatch SGD over buffered data. We use Adam with learning rate η, PPO clip ϵ, value loss c_v, and entropy c_e. Advantages $\hat{A}_t$ are computed with generalized advantage estimation (GAE) $(\gamma, \lambda_{\text{GAE}})$ and are standardized per batch. We apply gradient clipping (global norm $g_{\max}$) and reward normalization.

2) Action parametrization: The policy has two outputs. For each group g, a continuous head gives a raw value that we pass through $\tanh$ and scale into $[-\alpha_{\max}, \alpha_{\max}]$ to get α_g. A discrete head gives the probability $p_g = \Pr(\sigma_g = +1)$ (via a sigmoid) and we sample $\sigma_g \in \{-1, +1\}$. We add entropy term so policy does not collapse to a single choice too early.

3) Safety projection and budgets: At each step, candidate update is projected onto intersection of magnitude, smoothness, and plausibility sets (Eqs. 2- 5). Budget indicators b_t track residual magnitude and smoothness to inform the policy.

4) Latency handling: We measure per-step action time and enforce $\text{time}(\delta_t) \leq T_{\max} - T_{\text{pipe}}$. If a step risks violation, we (i) reduce the step size λ, (ii) increase smoothing β, or (iii) skip the update and apply $\delta_t = \delta_{t-1}$.

5) Stability tricks: We use observation normalization (same scaler as the IDS), advantage centering, early stopping on the PPO KL distance, and periodic evaluation on held-out attack sequences to monitor benign-rate and constraint violations.

IV. Evaluation

Here, we describe the experimental setup and show how the RL agent degrades performance and generalizes to other IDS.

A. Experimental Setup

We use the IEEE 118-bus system implemented in OPAL-RT as our WAMPAC testbed. PMUs are placed at selected buses and stream synchrophasor measurements, including voltage magnitudes, angles, frequency, and ROCOF, at a fixed reporting rate. The GAN–LSTM IDS operates on sliding PMU windows (X_t) as defined in Sec. II-A, producing a benign score $(s_t \in [0, 1])$ per window. It is first trained offline on labeled benign and attack sequences with the same preprocessing and normalization as the RL agent. The RL agent then interacts with the detector under the threat model in Sec. II-B: at each step it observes the current attacked window, previous score, and budget indicators, chooses an eight-parameter grouped action over PMU feature groups, and the environment builds the perturbation, enforces magnitude/smoothness/plausibility constraints, and returns the new score and reward. We report score distributions, miss rate vs. decision threshold, ROC, and perturbation statistics (variance and minimal line-search scale).

B. Evasion of GAN–LSTM IDS

We first quantify how the RL agent degrades the state-of-the-art GAN–LSTM [6]. Fig. 2(a) shows the score distributions: original attack windows are concentrated near low

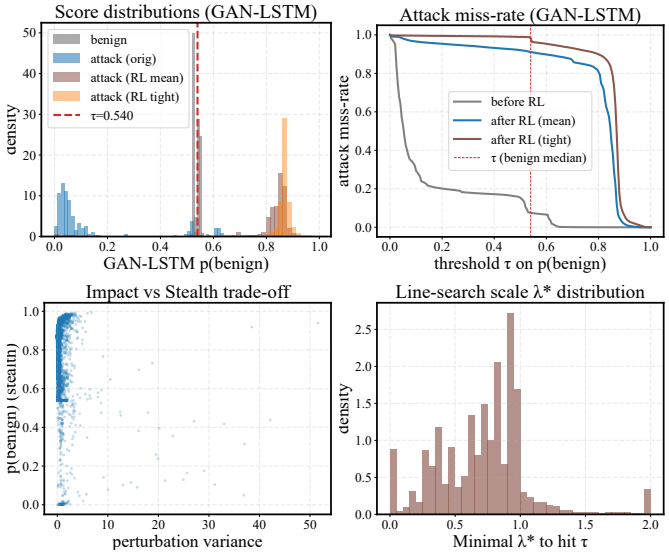

Fig. 2: RL-driven black-box evasion of the GAN–LSTM IDS: score distributions, attack miss-rate versus threshold, impact–stealth trade-off, and line-search scale (λ^*) distribution.

benign scores and are clearly separated from the benign cluster, whereas RL-perturbed attacks (mean and tight variants) are shifted into the benign region. In Fig. 2(b), we plot the attack miss-rate versus the decision threshold τ. Around the benign median threshold $\tau = 0.54$, the miss-rate increases from roughly 5% before RL to about 95% after RL, meaning that most attack sequences are now classified as benig. Large increases in miss-rate directly mean undetected disturbances or delayed corrective actions in the closed-loop WAMPAC pipeline. Figs. 2(c) and 2(d) highlight the impact–stealth behavior of the learned policy. Fig. 2(c) plots, for each attacked window, the perturbation variance against the resulting benign score $p(\text{benign})$; most points lie in the upper-left region, indicating that high benign scores are achieved with relatively small variance. Fig. 2(d) shows the distribution of the minimal line-search scale $\lambda^\star$ required to cross the benign threshold when we rescale the RL perturbation along its direction; the mass near $\lambda^\star \approx 1$ suggests that the raw policy outputs are already close to the decision boundary. Finally, Fig. 3 reports the ROC of the GAN–LSTM before and after RL, together with the distribution of PPO perturbation energy across PMU feature groups. The ROC curve collapses toward the diagonal under RL (AUC drops from 0.919 to 0.098), and the energy plot shows that most perturbation budget is spent on voltage magnitude features, consistent with their impact on IDS.

C. Transfer to CNN–LSTM IDS

To test whether the proposed framework generalizes beyond the GAN–LSTM baseline, we evaluate the same RL perturbations on a CNN–LSTM IDS trained on the same PMU dataset. This model represents a widely used state-of-the-art architecture, so it provides a natural baseline for comparison. We keep the RL policy fixed and simply pass the RL-perturbed windows through the CNN–LSTM. Fig. 4(a) shows that original attacks remain clustered near low benign

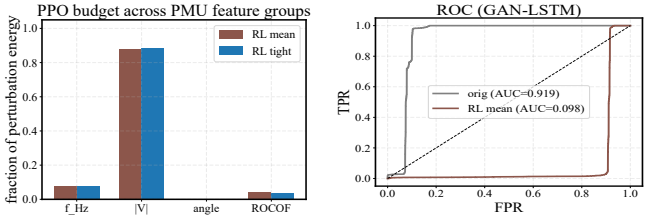

Fig. 3: (a) ROC of the GAN–LSTM IDS under RL evasion and (b) distribution of PPO perturbation energy across PMU feature groups.

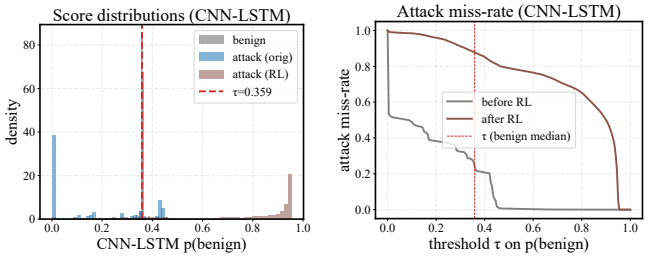

Fig. 4: RL perturbations transferred to the CNN–LSTM IDS: (a) score distributions and (b) attack miss-rate versus threshold

scores, while RL-perturbed attacks move closer to the benign score region. The corresponding attack miss-rate curves in Fig. 4(b) confirm this effect: near the benign median threshold $\tau = 0.359$, the miss-rate for the CNN–LSTM increases significantly once RL perturbations are applied, so many sequences that were previously detected now cross the benign threshold. These results indicate that the proposed RL-based grouped-action policy framework does not overfit a single detector, but generalizes well to another strong IDS architecture while still keeping perturbations modest and physically plausible.

D. Discussion and Defenses

Our results show that a query-based RL adversary can reliably evade a PMU-driven IDS in a realistic WAMPAC setting, even under magnitude, smoothness, and plausibility constraints. In terms of defenses, the IDSs can be coupled with moving-target defense (MTD) mechanisms that introduce controlled diversity in the cyber and control layers. For example, operators could periodically vary routing paths between PMU and controller hosts, rotate PMU-to-controller assignments, or randomize which subsets of PMU features feed a given analytic. By changing these configurations over time, MTD can limit how quickly a black-box agent learns a stable policy and can force the attacker to continually readapt. In addition, our findings motivate several complementary countermeasures. IDS ensembles that combine heterogeneous models (e.g., GAN–LSTM, CNN–LSTM, and residual-based checks) can reduce the risk of overfitting to a single discriminator. Adversarial training with RL-generated samples under different operating thresholds may further harden the detector against query-based fine-tuning. A systematic design and evaluation of such MTD- and ensemble-based defenses, using the proposed RL framework, is a direction for future work.

V. Conclusion

In summary, we presented a control-aware RL framework for query-based black-box evasion against PMU-based IDSs in WAMPAC, modeling the attacker as a PPO agent with an 8-parameter grouped action over PMU feature groups. On the IEEE 118-bus system in OPAL-RT, the learned policy raises the attack miss-rate at the benign median threshold from about 5% to about 95% with perturbations meeting timing constraints. The same RL perturbations degrade ROC score and generalize to other IDSs, indicating that current learning-based IDSs remain vulnerable to adaptive adversaries.

Acknowledgement

This work is supported by the Department of Energy (DOE) under Award# DE-CR0000024. Any opinions, findings, conclusions, or recommendations expressed in this material are those of the authors and do not reflect the views of the DOE.

References

[1] M. S. Almas, L. Vanfretti, R. S. Singh, and G. M. Jonsdottir, "Vulnerability of synchrophasor-based wampac applications' to time synchronization spoofing," *IEEE Transactions on Smart Grid*, vol. 9, no. 5, pp. 4601–4612, 2017.

[2] B. Chen, S.-i. Yim, H. Kim, A. Kondabathini, and R. Nuqui, "Cybersecurity of wide area monitoring, protection, and control systems for hvdc applications," *IEEE Transactions on Power Systems*, vol. 36, no. 1, pp. 592–602, 2020.

[3] A. Aligholian, A. Shahsavari, E. Cortez, E. Stewart, and H. Mohsenian-Rad, "Event detection in micro-pmu data: A generative adversarial network scoring method," in *2020 IEEE Power & Energy Society General Meeting (PESGM)*, pp. 1–5, IEEE, 2020.

[4] A. Sahu, T. Nguyen, K. Chen, X. Zhang, and M. Hassanaly, "Detection of false data injection attacks (fdia) on power dynamical systems with a state prediction method," *IEEE Access*, 2024.

[5] K. Bitirgen and Ü. B. Filik, "A hybrid deep learning model for discrimination of physical disturbance and cyber-attack detection in smart grid," *International Journal of Critical Infrastructure Protection*, vol. 40, p. 100582, 2023.

[6] M. Asif, P. Mali, M. Z. Haider, L. Ren, S. Paudyal, and M. A. Rahman, "Adversarially robust intrusion detection for pmu-enhanced power grids: Detecting stealthy attacks in a wide area control framework,"

[7] A. Ilyas, L. Engstrom, A. Athalye, and J. Lin, "Black-box adversarial attacks with limited queries and information," in *International conference on machine learning*, pp. 2137–2146, PMLR, 2018.

[8] D. Ding, M. Zhang, F. Feng, Y. Huang, E. Jiang, and M. Yang, "Black-box adversarial attack on time series classification," in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 37, pp. 7358–7368.

[9] M. Biswal, S. Misra, and A. S. Tayeen, "Black box attack on machine learning assisted wide area monitoring and protection systems," in *2020 IEEE Power & Energy Society Innovative Smart Grid Technologies Conference (ISGT)*, pp. 1–5, IEEE, 2020.

[10] M. Higgins, J. Zhang, N. Zhang, and F. Teng, "Topology learning aided false data injection attack without prior topology information," in *2021 IEEE Power & Energy Society General Meeting (PESGM)*, pp. 1–5, IEEE, 2021.

[11] P. L. Bhattar and N. M. Pindoriya, "False data injection attack with max-min optimization in smart grid," *Computers & Security*, vol. 140, p. 103761, 2024.

[12] S. Paudel, "An evaluation of methods for detecting false data injection attacks in the smart grid," *Frontiers in Computer Science*, vol. 6, p. 1504548, 2024.

[13] Y. Cheng, K. Yamashita, N. Yu, and Y. Liu, "A hybrid query-efficient black-box adversarial attack on power system event classifiers," in *2024 IEEE International Conference on Communications, Control, and Computing Technologies for Smart Grids (SmartGridComm)*, pp. 359–365, IEEE, 2024.

[14] S. Liu and Y. Luo, "Score-based black-box adversarial attack on time series using simulated annealing classification and post-processing based defense," *Electronics*, vol. 13, no. 3, p. 650, 2024.

[15] Y. Cheng, K. Yamashita, and N. Yu, "Adversarial attacks on deep neural network-based power system event classification models," in *2022 IEEE PES Innovative Smart Grid Technologies-Asia (ISGT Asia)*, pp. 66–70, IEEE, 2022.