# A Beamforming Attack on Deep Learning-Based Spectrum Violator Localization

Chibuikem Ezemaduka
*Electrical, Computer, and Systems Engineering*
*Rensselaer Polytechnic Institute*
Troy, NY, USA
ezemac@rpi.edu

Alhussein A. Abouzeid
*Electrical, Computer, and Systems Engineering*
*Rensselaer Polytechnic Institute*
Troy, NY, USA
abouzeid@ecse.rpi.edu

***Abstract*—We investigate the robustness of deep learning-based transmitter localization techniques to an adversarial machine learning (AML) attack in a realistic scenario, where spectrum violators aim to evade localization without falsifying sensor data. In our proposed attack model, violators strategically apply beamforming to produce authentic but deliberately crafted received signal strength (RSS) measurements at sensors, which when processed by the localization model, misleads it. Our evaluation results show that even with modest antenna array sizes, the attack can increase localization error by up to 5 times relative to non-adversarial conditions, and up to 10 times with larger arrays. We also show that the attack remains effective under stringent conditions such as full sensor coverage, achieving up to 3 times higher localization error in such settings. Our results underscore the need for robust defense mechanisms to secure deep learning-based localization systems in wireless spectrum management.**

***Index Terms*—Adversarial attack, deep learning, localization, interference management, spectrum sharing, beamforming**

## I. Introduction

Spectrum violators transmitting without authorization, pose a serious threat to spectrum sharing systems by causing harmful interference to primary users [1]. To address this challenge, researchers have developed transmitter localization techniques designed to accurately estimate the positions of such unauthorized transmitters. Recently, low-cost methods solely based on received signal strength (RSS) measurements gathered from crowdsourced sensors, and processed by convolutional neural network (CNN) models, have gained prominence and demonstrated promising performance [2].

As these techniques continue to advance, adversaries are likely to develop strategies that would allow them to evade localization, while continuing to transmit illegally. Deep learning models are well known to be susceptible to adversarial machine learning (AML) attacks [3] [4]. These attacks generally involve an adversary carefully perturbing the input data to mislead the model, resulting in significant errors in its predictions. While AML attacks have been widely explored across various domains, their applicability to learning-based transmitter localization remains largely unstudied. Existing research in this area, e.g., [5], primarily adopts the conventional data falsification threat model, where adversaries are assumed to directly modify RSS measurements before they are processed by the localization model. However, such direct data manipulation may not be practically feasible due to stringent security measures aimed at preventing data breaches and alterations [6].

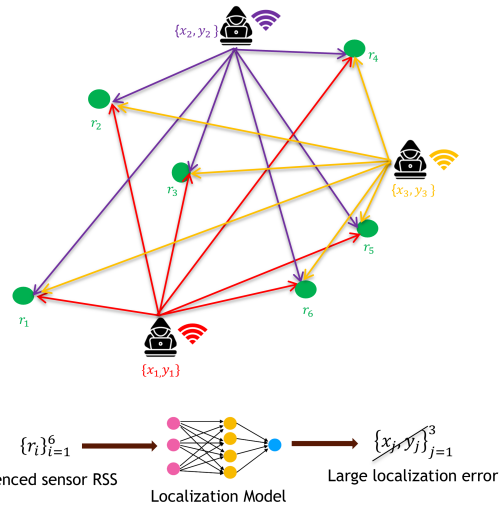


Fig. 1: Attack scenario involving 3 violators and 6 sensors. The violators use coordinated beamforming to influence sensors' RSS measurements, misleading the localization model and concealing their true positions.

The objective of this work is thus to investigate a different attack approach, where spectrum violators attempt to evade localization without having to falsify sensor data. Specifically, we explore a setting where violators attempt to successfully disrupt the localization process purely through clever transmission patterns designed to confuse the localization algorithm. As illustrated in Fig. 1, violators, without having to modify sensor data directly, can, through intelligent beamforming, produce real but targeted RSS measurements, which when processed by the CNN, degrades its localization accuracy. To the best of our knowledge, this work is the first to investigate AML attacks targeting deep learning-based transmitter localization through beamforming rather than data falsification.

In this paper, we assess the robustness of a CNN-based

localization method under this practical adversarial scenario. Our evaluation considers various factors, including the amount of sensor data available to the localization model and the beamforming capabilities of the violators. Our primary goal is to quantify how effectively an adversary can degrade localization accuracy through strategic beamforming.

The remainder of this paper is organized as follows. Section II discusses related works. Sections III and IV outline the attack formulation and implementation. Sections V and VI present our evaluation setup and results. Finally, Section VII concludes the paper and outlines avenues for future research.

## II. Related Work

AML attacks have gained considerable attention in wireless communications, with applications in areas such as modulation and signal classification, spectrum sensing, channel state information (CSI) feedback, and network slicing [7] [8] [9]. Such attacks generally aim to deceive a model into producing incorrect outputs for the wireless task it was trained to perform. Despite extensive research in these domains, AML attacks targeting deep learning-based transmitter localization remain largely unexplored. Most existing localization methods primarily utilize RSS measurements obtained from sensors, which are then processed by CNN models to estimate transmitter locations [10] [11] [12] [13] [14], while a few employ graph neural networks to locate spectrum violators [15] [16]. AML research in this context has primarily focused on the spectrum sensing data falsification (SSDF) threat model, in which adversaries directly manipulate sensor data inputs during model training or inference [5]. However, SSDF scenarios may be impractical in environments with stringent security controls preventing unauthorized access to sensor data. Moreover, the SSDF threat model has been extensively studied, leading to various robust countermeasures [17].

This paper introduces an AML attack mechanism that targets transmitter localization without data falsification. This is, arguably, more practical as it does not require access to the stored RSS data measurements. Instead, spectrum violators evade localization by carefully shaping their transmission patterns, exploiting vulnerabilities inherent in the localization algorithm itself. While prior works [18] [19] have investigated transmission pattern adjustments to spoof single transmitter locations (though not specifically in the context of violator localization), these works have not addressed deep learning-based localization techniques. Deep learning methods present unique challenges that require attackers to precisely tailor their strategies to exploit weaknesses in trained models. In this work, we consider an attack model in which violators can craft beam patterns which are specifically tuned for exploiting the vulnerabilities of a trained CNN model. Additionally, unlike prior works, our attack framework accommodates the possibility of multiple violators collaboratively coordinating to evade localization.

## III. Attack Formulation

Consider $K$ violators, each with a transmitter (TX) located at coordinate $\mathbf{A}_v \in \mathbb{R}^2$ , $\forall v = 1, \ldots, K$. Let $\mathbf{A} \triangleq [\mathbf{A}_1, \ldots, \mathbf{A}_K]^T$. Each violator's TX is equipped with a circular phased array of $N$ elements. The antenna weight vector $\mathbf{w}_v \in \mathbb{C}^N$, allows violator $v$ to control its radiation pattern.

Within the spectrum area, $S$ monitoring sensors ($S \in \mathbb{Z}^+$) are evenly distributed and each sensor $i \in \{1, \ldots, S\}$ measures the RSS at its location. Let $\mathbf{h}_{v,i} \in \mathbb{C}^N$ denote the steering vector of violator $v$'s TX towards sensor $i$.

Let $g_{v,i}$ denote the beamforming gain, and $c_{v,i}$, the channel gain from violator $v$'s TX to sensor $i$. The beamforming gain is defined as

$$g_{v,i} = \left|\mathbf{w_v}^H \mathbf{h}_{v,i}\right|^2 \tag{1}$$

The channel gain $c_{v,i}$ can be computed using any suitable propagation model. For simplicity, we adopt a distance-based path loss model given by $c_{v,i} = {d_{v,i}}^{-\alpha}$, where $d_{v,i}$ is the distance between the violator's TX and sensor $i$, and $\alpha$ is the path loss exponent.

Similarly, the beamforming and channel gains between violator $v$'s TX and its own receiver (RX) are denoted by ${g^R}_{v,v}$ and ${c^R}_{v,v}$. Gains between violator $v$'s TX and the RX of another violator $u$, are represented as ${g^R}_{v,u}$ and ${c^R}_{v,u}$ respectively.

$${g^R}_{v,v} = \left|\mathbf{w_v}^H \mathbf{h}_{v,v}\right|^2 \tag{2}$$

$${g^R}_{v,u} = \left|\mathbf{w_v}^H \mathbf{h}_{v,u}\right|^2 \tag{3}$$

and the channel gains by ${c^R}_{v,v} = {d^R}_{v,v}{}^{-\alpha}$ and ${c^R}_{v,u} = {d^R}_{v,u}{}^{-\alpha}$.

The RSS measured at sensor $i$, denoted $r_i$, is computed as:

$$r_i = \sum_{v=1}^{K} g_{v,i} c_{v,i} + n_i \tag{4}$$

where $n_i$ represents sensor noise.

The signal-to-interference-plus-noise ratio (SINR) at violator $v$'s RX, is denoted by $\gamma_v$:

$$\gamma_v = \frac{{g^R}_{v,v} {c^R}_{v,v}}{n_v + \sum_{u \neq v}^{K} {g^R}_{u,v} {c^R}_{u,v}} \tag{5}$$

where $n_v$ is the receiver noise. Let $\gamma_v^{\min}$ denote the minimum SINR required for acceptable quality-of-service (QoS) at violator $v$'s RX.

To localize violators, the RSS measurements from all sensors, $\mathbf{R} = [r_1, r_2, \ldots, r_S]^T$, are input to a CNN trained to produce location estimates $\hat{\mathbf{A}} = [\hat{\mathbf{A}}_1, \ldots, \hat{\mathbf{A}}_K]^T$. The localization function is given by $\hat{\mathbf{A}} = f(\mathbf{R}; \mathbf{\Theta})$ where $\mathbf{\Theta}$ represents the CNN weights.

Let the CNN's loss function, $L(\mathbf{R}, \mathbf{A}, \mathbf{\Theta})$ be:

$$L(\mathbf{R}, \mathbf{A}, \mathbf{\Theta}) = \sum_{v=1}^{K} \|\mathbf{A}_v - \hat{\mathbf{A}}_v\|^2. \quad (6)$$

The objective for the violators is to design beamforming patterns that maximize the loss function $L$, thereby misleading the CNN, while maintaining acceptable SINR and adhering to power constraints $p_v^{\max}$. Formally, the beamforming optimization problem is:

$$\begin{aligned} \underset{\mathbf{W}}{\text{maximize}} \;\; & L(\mathbf{R}, \mathbf{A}, \mathbf{\Theta}) \\ \text{s.t.} \quad & \gamma_v \geq \gamma_v^{\min} \qquad \forall v = 1, \ldots, K \\ & \|\mathbf{w}_v\|^2 \leq p_v^{\max} \qquad \forall v = 1, \ldots, K \end{aligned} \quad (7)$$

where $\mathbf{W} = [\mathbf{w}_1, \ldots, \mathbf{w}_K]^T$

Our attack scenario assumes a white-box setting, where violators possess complete knowledge of sensor positions and the CNN parameters. In scenarios where CNN parameters are unavailable (black-box scenario), attackers can instead train a surrogate model. Previous studies have demonstrated that AML attacks using surrogate models remain effective even if trained on different datasets with similar distributions [20]. We intend to investigate this direction in future work.

## IV. Attack Implementation

As is typical with most AML problems, the optimization in (7) is inherently non-convex due to the high non-linear nature of CNNs. To address this, we propose a heuristic gradient-based approach to solve the problem. We begin by reformulating the constrained optimization into an unconstrained form by incorporating the SINR constraints directly into the objective function through a penalty method.

The augmented objective function is defined as:

$$Z = L(\mathbf{R}, \mathbf{A}, \mathbf{\Theta}) - \sum_{v=1}^{K} \beta_v \max(0,\ \gamma_v^{\min} - \gamma_v)^2 \quad (8)$$

Here, $\beta_v > 0$ are penalty coefficients that softly enforce the SINR constraints for each violator $v \in \{1, \ldots, K\}$. The penalty term discourages solutions that violate the SINR requirement. Meanwhile, instead of adding the power constraints as penalty terms which would introduce additional complexity, we handle them via projection. Specifically, at each iteration, we apply an $\mathcal{L}_2$-norm projection to ensure that the transmit power remains within the allowable limit.

The resulting optimization problem is:

$$\underset{\mathbf{W}}{\text{maximize}} \;\; Z \quad (9)$$

The procedure for solving (9) is outlined in Algorithm 1, which uses a projected gradient ascent approach, as discussed, to maximize the objective function.

**Algorithm 1** AML attack against violator localization

**Input:** Initialized weight vectors $\mathbf{W}^{(0)}$; $\mathbf{h}_{v,\cdot}$, $\forall v = 1, \ldots, K$; $c_{v,\cdot}$, $\forall v = 1, \ldots, K$; initial penalty parameters $\beta_v^{(0)} > 0$, $\forall v = 1, \ldots, K$; step size $\rho > 0$; penalty increase factor $c > 1$; $f(\cdot; \mathbf{\Theta})$; $\mathbf{A}$; $p_v^{\max}$, $\forall v = 1, \ldots, K$; $\gamma_v^{\min}$, $\forall v = 1, \ldots, K$; max no. of iterations $T$; stopping criterion $\epsilon$; violator RX positions, $\forall v = 1, \ldots, K$; sensor positions, $\forall i = 1, \ldots, S$.

**Output:** Optimized antenna weight vectors $\mathbf{W}^*$.

$\mathbf{W} \leftarrow \mathbf{W}^{(0)}$, $t \leftarrow 0$
$\hat{\mathbf{A}}^{(0)} \leftarrow f\big(\mathbf{R}^{(0)}; \mathbf{\Theta}\big)$
$L^{(0)} \leftarrow \sum_{v=1}^{K} \big\|\mathbf{A}_v - \hat{\mathbf{A}}_v^{(0)}\big\|^2$
$\beta_v \leftarrow \beta_v^{(0)}$, $\forall v = 1, \ldots, K$
**while** $t \leq T$ **do**
  Compute $\gamma_v^{(t)}$, $\forall v = 1, \ldots, K$
  Compute $Z^{(t)}$
  Compute $\nabla_{\mathbf{W}} Z^{(t)}$
  $\mathbf{W} \leftarrow \mathbf{W} + \rho \nabla_{\mathbf{W}} Z^{(t)}$
  **for all** $v = 1, \ldots, K$ where $\|\mathbf{w}_v\|^2 > p_v^{\max}$ **do**
    $\mathbf{w}_v \leftarrow \mathbf{w}_v \times \dfrac{\sqrt{p_v^{\max}}}{\|\mathbf{w}_v\|}$
  **end for**
  Compute $\mathbf{R}^{(t+1)}$
  $\hat{\mathbf{A}}^{(t+1)} \leftarrow f\big(\mathbf{R}^{(t+1)}; \mathbf{\Theta}\big)$
  $L^{(t+1)} \leftarrow \sum_{v=1}^{K} \|\mathbf{A}_v - \hat{\mathbf{A}}_v^{(t+1)}\|^2$
  **if** $|L^{(t+1)} - L^{(t)}| < \epsilon$ **then**
    **break**
  **end if**
  $\beta_v \leftarrow c\beta_v$, $\forall v = 1, \ldots, K$
  $t \leftarrow t + 1$
**end while**
$\mathbf{W}^* \leftarrow \mathbf{W}$

TABLE I: Hyperparameter values used in the simulation.

| Parameter | Value |
|---|---|
| $T$ | 200 |
| $\rho$ | $10^{-1}$ |
| $\beta^{(0)}$ | 1.01 |
| $c$ | 1.5 |
| $\epsilon$ | $10^{-6}$ |

## V. Evaluation Setting

We use Python to simulate the attack within a 100 m x 100 m spectrum area, discretized into 1 m x 1 m grid cells. In this paper, we focus on the simplest scenario involving a single violator ($K = 1$), with plans to extend the evaluation to $K > 1$ in future work. The positions of the sensors, the violator's TX, and its corresponding RX are randomly assigned to the centers of the grid cells.

For the localization model, we adopt a CNN architecture inspired by the Deeptxfinder algorithm [14], which employs multiple CNNs, each with up to four convolutional layers and two dense layers, to estimate the locations corresponding to different predicted numbers of transmitters. Since our

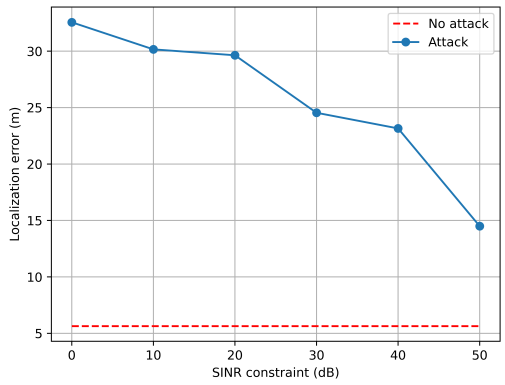


Fig. 2: Effect of SINR constraint on localization error.

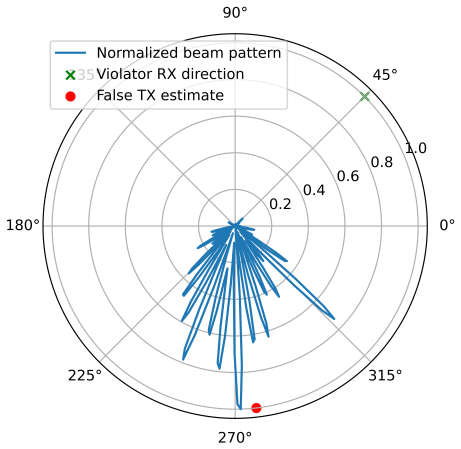


Fig. 3: Adversarial beam pattern (low SINR constraint).

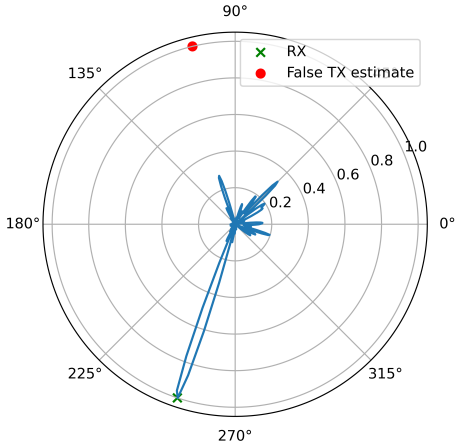


Fig. 4: Adversarial beam pattern (high SINR constraint).

evaluation focuses on a single violator, we train only one CNN to output the estimated position vector of a transmitter within the spectrum area. The training dataset consists of 10,000 samples generated through randomized simulations, where the transmit power is uniformly sampled between 1 and 5 dBm, and the sensor density varies across samples. The wireless channel is modeled using a path loss exponent of $\alpha = 2$, with additive white Gaussian noise (AWGN).

Our primary metric for evaluating attack performance is the localization error, defined as the Euclidean distance between the estimated and actual transmitter position. The hyperparameters that achieved the best empirical performance in our experiments are summarized in Table I.

We evaluate the attack under various conditions, including different SINR constraints, number of antenna elements in the violator's array, and sensor densities. For each setting, the localization error observed when no attack is applied, serves as the baseline.

Note that in some scenarios, the optimization algorithm may fail to converge due to infeasible SINR constraints. Since our primary objective is to assess attack effectiveness under successful attack conditions, we report results only for cases where convergence is achieved. In practical deployments, a violator may choose a more relaxed SINR constraint based on operational trade-offs. In our simulations, the transmit power is selected to ensure that the SINR requirement is at least theoretically satisfiable.

The required gradients for optimization are obtained using the automatic differentiation capabilities of PyTorch. Unless otherwise specified, all simulations assume a default configuration where the violator uses a 16-element antenna array and targets an SINR of 20 dB. The default sensor density is 10%, and the results for each experimental setting are averaged over independent trials.

## VI. Evaluation

In this section, we present and analyze the results of our evaluation under various experimental settings.

### A. Localization error - SINR tradeoff

Fig. 2 illustrates the attack's impact on localization accuracy across varying SINR constraints imposed by the violator. Since our evaluation considers only a single violator, we assume an interference-free environment and compute SINR solely based on AWGN.

As expected, the effectiveness of the attack generally diminishes with stricter SINR requirements. When the SINR constraint increases, a larger portion of the violator's antenna gain must be directed toward its own receiver to meet the quality-of-service condition. This is evident in the beam pattern of a 128-element violator array shown in Fig. 4, where a higher SINR constraint (50 dB) results in less flexibility for misleading the CNN. In contrast, Fig. 3 shows that with a much more relaxed SINR requirement (5 dB), the violator can focus more of its gain towards misleading the model by simulating a plausible but fake TX location.

In general, the attack is quite effective, increasing the localization error by up to 5 times at a realistic constraint of 20 dB.

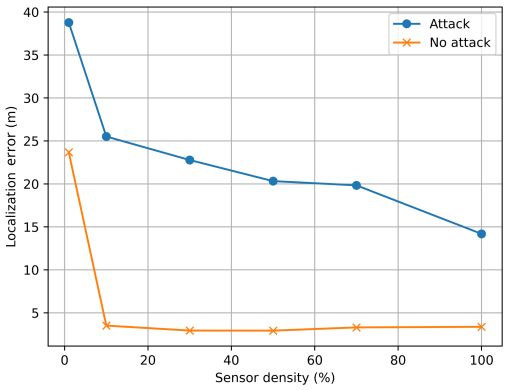


Fig. 5: Impact of sensor density on localization error.

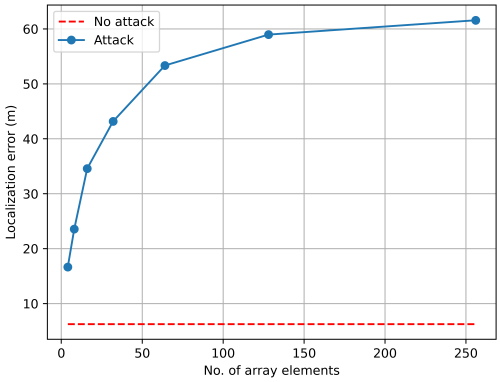


Fig. 6: Impact of array size on localization error.

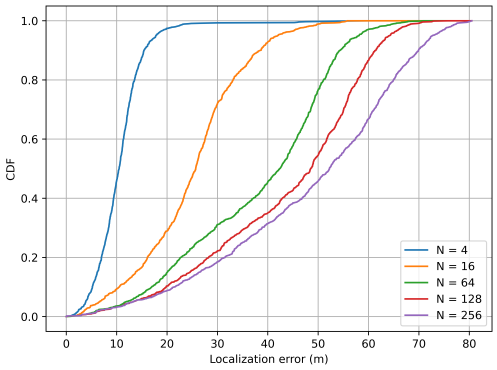


Fig. 7: CDF of localization error for different array sizes.

Interestingly, in some simulation runs, we observed that under stricter SINR constraints, Alg. 1 tended to steer the fake location in the same direction as the violator's RX. This behavior suggests that the algorithm may attempt to exploit beamforming efficiency by using the high gain required for maintaining SINR to simultaneously mislead the CNN model. However, this introduces the risk of revealing the receiver's location, even though the localization model is only trained to estimate TX positions.

The violator could circumvent such scenarios by introducing an additional term into the loss function: $\sum_{v=1}^{K} \|\mathbf{B}_v - \hat{\mathbf{A}}_v\|^2$, where $\mathbf{B}_v$ denotes the location of the violator's RX. This modification would discourage the algorithm from generating fake locations that are close to either the TX or the RX, thus better preserving the anonymity of both entities.

### B. Impact of Sensor Coverage

We evaluate the effectiveness of the attack across 1%, 10%, 30%, 50%, 70%, and 100% sensor densities, using a smaller array size of 8 elements. As observed in Fig. 5, the attack's effectiveness generally diminishes with increasing sensor density. This reduction in effectiveness can be attributed to richer and more diverse sensor data, which enhances the robustness of the localization algorithm and makes it less susceptible to deception.

Although the baseline error without an attack remains relatively stable across sensor densities from 10%, there is a steady decline in the localization error that the attack can achieve. Nonetheless, even at full sensor coverage (100%), the attack still remains effective, achieving localization error of up to 3 times greater than the baseline.

These observations suggest that enhancing sensor coverage can improve the robustness of CNN-based localization systems against adversarial attacks. However, deploying extensive sensor networks is typically expensive and may not always be practical. Therefore, it is crucial to explore alternative methods to enhance model robustness without incurring the substantial costs associated with increased sensor deployment.

### C. Impact of Array Size

We evaluate the impact of antenna array size on attack effectiveness, considering $N \in \{4, 8, 16, 32, 64, 128\}$ elements. As shown in Fig. 6, an increase in the number of array elements significantly enhances the attack's effectiveness. Specifically, with a 128-element array, the violator can achieve a more than 10 times increase in localization error. This improvement can be intuitively explained; a larger array offers increased beamforming flexibility, enabling the formation of more complex and diverse beam patterns which can better mislead the CNN.

Fig. 7, showing the cumulative distribution function (CDF) of the localization error for different array sizes, further illustrates the impact that the array size has on improving the effectiveness of the attack.

## VII. Conclusion and Future Work

We demonstrated the effectiveness of an AML attack against deep learning-based transmitter localization without falsifying sensor data. By intelligently leveraging beamforming, spectrum violators can evade localization by influencing real sensor measurements, which when processed by a localization model, leads to substantial localization

error. Our simulations indicate that with a moderate array size of 16 elements, an attacker can cause up to 5 times increase in localization error, and up to 10 times increase with a much larger array of 128 elements. In future work, we intend to explore more complex scenarios involving multiple cooperating violators with diverse capabilities and operational constraints, as well as black-box attack models where violators lack knowledge of the CNN parameters and sensor positions, and scenarios with more realistic wireless channel models. Additionally, we plan to investigate robust defensive strategies to enhance the resilience of localization algorithms against such adversarial attacks.

## ACKNOWLEDGMENT

This material is based upon work supported in part by the National Science Foundation under grant number 2007454.

## REFERENCES

[1] A. Dutta and M. Chiang, ""see something, say something" crowdsourced enforcement of spectrum policies," *IEEE Transactions on Wireless Communications*, vol. 15, no. 1, pp. 67–80, 2016.

[2] F. Mitchell, J. P. Smith, S. Sarkar, N. Patwari, A. Bhaskara, and S. K. Kasera, "Localizing spectrum offenders using crowdsourcing," in *Network Security Empowered by Artificial Intelligence*. Springer, 2024, pp. 237–264.

[3] C. Szegedy, W. Zaremba, I. Sutskever, J. Bruna, D. Erhan, I. Goodfellow, and R. Fergus, "Intriguing properties of neural networks," 2014. [Online]. Available: https://arxiv.org/abs/1312.6199

[4] Y. Vorobeychik and M. Kantarcioglu, *Adversarial machine learning*. Morgan & Claypool Publishers, 2018.

[5] F. Mitchell, P. Smith, A. Bhaskara, and S. K. Kasera, "Exploring adversarial attacks on learning-based localization," in *Proceedings of the 2023 ACM Workshop on Wireless Security and Machine Learning*, ser. WiseML'23. New York, NY, USA: Association for Computing Machinery, 2023, p. 15–20. [Online]. Available: https://doi.org/10.1145/3586209.3591398

[6] S. Shi, Y. Xiao, W. Lou, C. Wang, X. Li, Y. T. Hou, and J. H. Reed, "Challenges and new directions in securing spectrum access systems," *IEEE Internet of Things Journal*, vol. 8, no. 8, pp. 6498–6518, 2021.

[7] D. Adesina, C.-C. Hsieh, Y. E. Sagduyu, and L. Qian, "Adversarial machine learning in wireless communications using rf data: A review," *IEEE Communications Surveys & Tutorials*, vol. 25, no. 1, pp. 77–100, 2023.

[8] J. Liu, M. Nogueira, J. Fernandes, and B. Kantarci, "Adversarial machine learning: A multilayer review of the state-of-the-art and challenges for wireless and mobile systems," *IEEE Communications Surveys Tutorials*, vol. 24, no. 1, pp. 123–159, 2022.

[9] W. Zhang, M. Krunz, and G. Ditzler, "Stealthy adversarial attacks on machine learning-based classifiers of wireless signals," *IEEE Transactions on Machine Learning in Communications and Networking*, vol. 2, pp. 261–279, 2024.

[10] C. Zhan, M. Ghaderibaneh, P. Sahu, and H. Gupta, "Deepmtl pro: Deep learning based multiple transmitter localization and power estimation," *Pervasive and Mobile Computing*, vol. 82, p. 101582, 2022. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1574119222000311

[11] F. Mitchell, A. Baset, N. Patwari, S. K. Kasera, and A. Bhaskara, "Deep learning-based localization in limited data regimes," in *Proceedings of the 2022 ACM Workshop on Wireless Security and Machine Learning*, ser. WiseML '22. New York, NY, USA: Association for Computing Machinery, 2022, p. 15–20. [Online]. Available: https://doi.org/10.1145/3522783.3529529

[12] C. Ezemaduka and A. A. Abouzeid, "Privacy-aware deep learning based localization of spectrum violators," in *2023 IEEE 20th International Conference on Mobile Ad Hoc and Smart Systems (MASS)*, 2023, pp. 98–106.

[13] Yapar, R. Levie, G. Kutyniok, and G. Caire, "Real-time outdoor localization using radio maps: A deep learning approach," *IEEE Transactions on Wireless Communications*, vol. 22, no. 12, pp. 9703–9717, 2023.

[14] A. Zubow, S. Bayhan, P. Gawłowicz, and F. Dressler, "Deeptxfinder: Multiple transmitter localization by deep learning in crowdsourced spectrum sensing," in *2020 29th International Conference on Computer Communications and Networks (ICCCN)*, 2020, pp. 1–8.

[15] Y. Zhang, T. Li, and Y. Zhang, "Gnn-sml: Graphic neural network-based spectrum misuser localization," in *IEEE INFOCOM 2025 - IEEE Conference on Computer Communications*, 2025, pp. 1–10.

[16] C. Ezemaduka and A. A. Abouzeid, "Spec-gnn: Spectrum enforcement through graph neural networks in dynamic spectrum access systems," *IEEE Transactions on Cognitive Communications and Networking*, pp. 1–1, 2025.

[17] Q. Wang, H. Sun, R. Q. Hu, and A. Bhuyan, "When machine learning meets spectrum sharing security: Methodologies and challenges," *IEEE Open Journal of the Communications Society*, vol. 3, pp. 176–208, 2022.

[18] T. Wang and Y. Yang, "Analysis on perfect location spoofing attacks using beamforming," in *2013 Proceedings IEEE INFOCOM*, 2013, pp. 2778–2786.

[19] Y. Lin, Y. Ye, and Y. Yang, "Preserving incumbent user's location privacy against environmental sensing capability," in *2019 IEEE International Symposium on Dynamic Spectrum Access Networks (DySPAN)*, 2019, pp. 1–10.

[20] N. Papernot, P. McDaniel, I. Goodfellow, S. Jha, Z. B. Celik, and A. Swami, "Practical black-box attacks against machine learning," in *Proceedings of the 2017 ACM on Asia Conference on Computer and Communications Security*, ser. ASIA CCS '17. New York, NY, USA: Association for Computing Machinery, 2017, p. 506–519. [Online]. Available: https://doi.org/10.1145/3052973.3053009