

Towards Distance-Invariant Radio Frequency Fingerprinting via Augmented Unsupervised Learning

Shiyue Huang, Yuchen Su, Hongbo Liu*, Zikang Ding, Xuewan He, Yanzhi Ren, Haitao Jia

University of Electronic Science and Technology of China
sy.huang@std.uestc.edu.cn, hongbo.liu@uestc.edu.cn

Abstract

Radio Frequency Fingerprinting (RFF) exploits inherent hardware-level imperfections of wireless transmitters as unclonable identifiers for device identification. These unique signatures, concealed in transmitted signals, inevitably experience complex distortions during wireless propagation (i.e., coupled with ambient noise and channel fading), making it extremely challenging for reliable extraction. Despite substantial research efforts dedicated to advancing effective fingerprint extraction techniques, current approaches still struggle in handling fingerprint robustness under distance variations, leading to severe SNR fluctuations and complex multipath effects. To address this gap, we propose the first unsupervised framework for distance-invariant radio frequency fingerprinting, eliminating dependence on labeled target domain data. Specifically, we first preprocess raw RF samples by confining them within a specified variation range and filtering noisy high-frequency components while avoiding aliasing. For source domain data, we then propose a set of physics-inspired data augmentation techniques designed to emulate realistic wireless signal propagation effects. Building on this, we introduce a dual alignment contrastive learning method to explicitly decouple identity-discriminative features, ensuring the model focuses on device-specific traits. Furthermore, we incorporate a pseudo-labeling-based domain adaptation module to refine the model for the unlabeled target domain, enhancing its generalization to unseen distances. Extensive experiments on public datasets show that our method achieves the identification accuracy outperforming state-of-the-art approaches by 40%, while maintaining computational efficiency suitable for edge deployment.

Introduction

The proliferation of Internet of Things (IoT) devices in critical infrastructure and daily life has intensified the demand for lightweight, secure identification mechanisms. Traditional software-based identifiers (e.g., MAC addresses) are vulnerable to spoofing, while public-key cryptography (PKC) imposes severe power and computational burdens, making it incompatible for resource-constrained IoT devices. As a promising non-cryptographic alternative, radio frequency fingerprinting (RFF) leverages unique, unclonable hardware impairments in RF transceivers as physical

*Corresponding author.

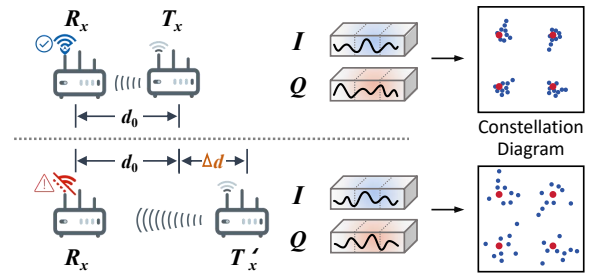


Figure 1: Illustration of distance-invariant RFF.

identifiers (Zhang et al. 2025), as shown in Figure 1. These hardware-imprinted signatures enable packet-level identification resistant to spoofing, making RFF ideal for high-security scenarios. Furthermore, RFF relies on hardware-intrinsic properties, eliminating the need for upper-layer protocols or additional computational overhead. Recognizing these advantages, the Defense Advanced Research Projects Agency (DARPA) has prioritized RFF research through its RF Machine Learning Systems (RFMLS) program (DARPA 2017), driving extensive studies on RFF for device identification.

Given above aforementioned merits, a critical limitation still persists: most RFF methods exhibit limited generalization under spatially variant channel conditions. Specifically, existing works (Wu et al. 2021; Yao et al. 2025; Zhang et al. 2025) have successfully mined discriminative features for stable-channel identification, but they often fail in dynamic channel. Domain adaptation techniques (Zhang et al. 2022; Li et al. 2022; Yin et al. 2023), although succeeded in dealing with temporal variations (e.g., cross-day identification), also struggle with spatial generalization. This challenge stems from two key factors: as the distance between devices increases, the signal-to-noise ratio (SNR) decreases under fixed transmission power, causing subtle fingerprint information to be easily masked by noise; meanwhile, the complex multipath effect caused by distance changes couples with the RF signal, further complicating fingerprint recognition. Alternative approaches have turned to few-shot learning (Zhao, Wang, and Mao 2024; Sun et al. 2025) to tackle domain shifts and scalability across different distances. However, these methods usually require labeled

training samples from the target domain. Collecting and annotating such data for each new separation distance is both labor-intensive and time-consuming, severely limiting their practical deployment. Thus, enabling distance-invariant RFF without relying on labeled target-domain data remains an open and critical challenge.

In this paper, we present the first distance-invariant RFF identification framework that eliminates reliance on labeled target-distance data. To obtain robust fingerprint representations, we develop several physics-inspired data augmentation strategies. Leveraging the augmented data, we propose a dual alignment framework that explicitly decouples identity-discriminative features, thereby extracting distance-invariant fingerprints. Additionally, we propose a progressive pseudo-labeling adaptation framework that iteratively refines model parameters using high-confidence predictions on unlabeled target samples. This strategy bridges domain gaps without requiring manual annotation, achieving state-of-the-art (SOTA) performance over existing methods. Our main contributions are summarized as follows:

- 1) To the best of our knowledge, this is the first work enabling distance-invariant RFF that generalizes to unseen distances without labeled target samples, addressing a critical gap in practical deployment.
- 2) We propose physics-inspired data augmentation strategies that explicitly model the wireless signal propagation effects, effectively expanding the source domain to cover diverse wireless channel conditions.
- 3) We design a dual alignment and progressive adaptation framework that jointly learns invariant representations through class-level and prototype-level alignments and progressively adapts to unlabeled target domains.
- 4) Extensive experiments on public datasets show that our framework achieves encouraging identification accuracy on 16 devices across different distances from the source domain, outperforming SOTA methods by 40%.

Related Work

Deep learning-based RFF. Deep learning has become a dominant paradigm in RF fingerprinting (RFF) due to its strong capability of automatic feature extraction from raw RF signals, with a variety of architectures validated for performance improvement. For instance, convolutional neural networks (CNNs) (Wu et al. 2021; Sankhe et al. 2019; Al-Shawabka et al. 2020) have been widely adopted to capture local discriminative patterns in signal representations. Recurrent models such as LSTM (Shen et al. 2021) and GRU (Shen et al. 2023) later emerged to model temporal dependencies in sequential IQ samples. More recently, Transformer-based models (Xu and Xu 2021; Han et al. 2025) and state-space alternatives like Mamba (Yao et al. 2025) have been proposed to handle long-range correlations in signals, while KAN networks (Chen et al. 2025) offer more interpretable fingerprint extraction. Despite these advancements in improving identification accuracy under fixed experimental settings (e.g., static locations and stable channels), most deep learning-based methods still face critical

generalization issues across spatially variant channel conditions. This limitation highlights the need for more robust frameworks that can decouple device-specific fingerprints from distance sensitive channel distortions.

Contrative Learning for RFF. Contrastive learning (CL), as a powerful unsupervised representation learning tool, has shown great potential in RFF identification. For instance, Zha et al. (Zha et al. 2023) treat receiver-induced distribution shifts as a data augmentation strategy, leveraging SimSiam-based contrastive pre-training to extract receiver-agnostic features and alleviate cross-receiver performance degradation. Wang et al. (Wang et al. 2024) further propose a CL-based heterogeneous feature fusion method, yielding more discriminative RFF representations and improving open-set classification performance. Despite these advances, critical challenges remain. Most CL approaches for RFF assume same place deployment, making them vulnerable to real-world spatial variations. While recent work explores residual contrastive learning between pre- and post-equalization signals to mitigate channel impacts (Pan et al. 2025), their validation is confined to simulations, lacking practical verification in real-world scenarios.

Domain Adaptation for RFF. Domain adaptation (DA) techniques (Li et al. 2024) have been explored to address distribution shifts in RFF authentication, primarily focusing on cross-modulation or cross-day scenarios. Yin et al. (Yin et al. 2023) proposed a few-shot DA framework for modulation-agnostic emitter identification, while Zhang et al. (Zhang et al. 2022) leveraged domain adversarial adaptation by contrasting real signals with demodulated ideal signals to improve robustness under variable modulations. In cross-day scenarios, RadioNet (Li et al. 2022) leveraged adversarial adaptation and a device rank metric to mitigate temporal drift. However, distance-invariant RFF adaptation remains under-explored. Most existing DA methods assume static device-receiver distances, yet spatial variations introduce SNR degradation and multipath distortions that obscure intrinsic RFF patterns (Wan et al. 2024). Wan et al. proposed the only cross-distance approach (VC-SEI) using semi-supervised DA, but it requires partial labeled target-domain data—a costly constraint for IoT deployments. Unlike these works, our method achieves fully unsupervised cross-distance adaptation without target-domain labels.

Methodology

Problem Definition

Assume we have a dataset from source domain $\mathcal{D}^s = \{(x_i^s, y_i^s)\}_{i=1}^{N_s}$, where $x_i^s \in \mathcal{X} \subseteq \mathbb{C}^T$ represents complex-valued RF samples acquired by a wireless receiver (Rx) with a fixed distance from the transmitter (Tx), and $y_i^s \in \{1, 2, \dots, K\}$ corresponds to the device identity labels spanning K distinct classes. For the target domain, we consider an unlabeled dataset $\mathcal{D}^t = \{x_j^t\}_{j=1}^{N_t}$ consists of N_t RF samples. These samples are collected at transceiver separation distances distinct from those in the source domain, which imply different multipath effects and SNR characteristics from the source domain. Our objective is to learn a feature extractor $f_e : \mathcal{X} \rightarrow \mathbb{R}^d$ that maps raw RF sam-

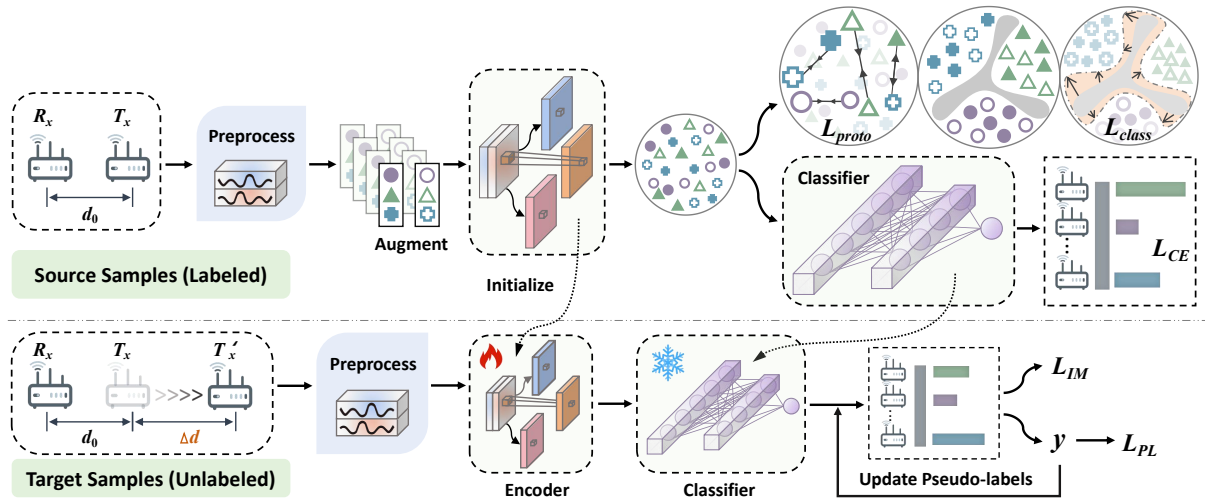


Figure 2: Framework of our method.

ples to d -dimensional feature embeddings, and a classifier $f_c : \mathbb{R}^d \rightarrow \mathbb{R}^K$ that predicts device labels from these embeddings. The feature extractor f_e captures spatially invariant, device-specific physical-layer features from \mathcal{D}^s , enabling the composite mapping $f_c \circ f_e$ to achieve high target-domain (\mathcal{D}^t) accuracy without requiring labeled \mathcal{D}^t samples.

Model Overview

As shown in Figure 2, our framework comprises three stages: *preprocessing*, *contrastive learning on the source domain*, and *unsupervised domain adaptation on the target domain*. First, *preprocessing* aims to mitigate domain shifts of raw input RF samples, establishing a consistent input space. Given the preprocessed RF samples, *contrastive learning* applies dual alignment to learn distance-invariant features, thereby constructing the source model. Finally, *target domain adaptation* facilitates adaptation to the unlabeled data in target domain with the feature encoder initialized from the source model.

Preprocessing

To reduce the impact of noise and multipath effects over wireless channel, we design the following preprocessing pipeline:

Normalization and Min-Max Scaling. Wireless signals experience different attenuation as the distance between transceivers varies, causing the amplitude of RF samples usually spanning over a large range. To eliminate such impact of distance on the RF samples for better domain adaptation, we confine the raw RF samples through normalization and min-max scaling. First, we normalize the sample using the root-mean-square (RMS) value:

$$\hat{x}_i = \frac{x_i}{\sqrt{\frac{1}{T} \sum_{t=1}^T |x_i[t]|^2}}, \quad \forall x_i \in \mathcal{X}, \quad (1)$$

where T denotes the length of each complex-valued sample $x_i \in \mathbb{C}^T$, and then we apply min-max scaling to bound the

sample within a standardized dynamic range:

$$\tilde{x}_i[t] = \frac{\hat{x}_i[t] - \min(\hat{x}_i)}{\max(\hat{x}_i) - \min(\hat{x}_i)}, \quad \forall t \in \{1, \dots, T\}. \quad (2)$$

Resampling. High-frequency components of RF fingerprints easily couple with interference and multipath effects on wireless channel, obscuring genuine feature extraction, while the low frequency components are more resilient to the external factors. Thus, we next resample the signal \tilde{x}_i at a lower sampling rate to restrict the RF fingerprints within low frequency band. In the meanwhile, it also dimensionally compresses the input data, thereby expediting subsequent feature extraction.

Source Model with Contrastive Alignment

To learn distance-invariant features from the source domain for robust fingerprinting, we design a dual contrastive learning framework to build the source model, combining both class-level and prototype-level alignment.

Backbone. Inspired by the MCNN (Cui, Chen, and Chen 2016), we propose a Multi-Scale Attention Network (MSAN) as the backbone network, specifically tailored for RF fingerprint extraction. As shown in Figure 3, the MSAN architecture consists of six cascaded encoder modules. Each module integrates three components: multi-scale convolutional layers to capture device-specific fingerprint across different time-frequency resolutions, a channel-wise attention module to suppress noise, and ReLU activation. To preserve gradient flow and mitigate vanishing gradients in deep stacks, we implement identity mappings that directly connect each block's output to subsequent layers. Detailed specifications and comparative experiments refer to the Appendix A and C.

Data Augmentation. Contrastive learning relies on maximizing agreement between augmented views of the same sample while minimizing similarity with negative samples, making the design of domain-specific augmentations

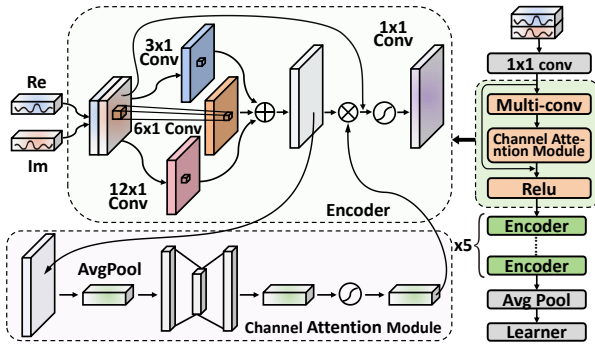


Figure 3: The structure of backbone.

critical for effective representation learning (Grill et al. 2020; Demirel and Holz 2023). To augment input RF samples, we propose a physics-inspired augmentation strategy through simulating realistic wireless channel effects with six transformations: *Signal scaling* applies random normal-distributed factors $\mathcal{N}(1, 0.1^2)$ to mimic wireless signal fluctuations caused by different path loss; *Permutation* (Liu and Chen 2024) partitions signals into multiple segments and reorders them randomly to emulate multipath effect; *Window slicing* extracts partial of the signals and expands it with linear interpolation, simulating the channel fading; *Time warping* (Zhang et al. 2024) introduces nonlinear distortions via random warping factors to replicate time-varying delays; *Magnitude warping* (Um et al. 2017) applies random scaling at knot points with cubic spline interpolation, modeling amplitude-varying channel fading; and *Window warping* (Rashid and Louis 2019) randomly select the segment of RF sample and stretch it through the linear interpolation, capturing local signal distortions. These augmentations systematically encode physical-layer invariances into the representation learning process, enabling the model to disentangle device-specific RF fingerprint from spatially variant propagation artifacts during contrastive optimization. During the training phase, we randomly sample and compose four distinct augmentations from the available set for each batch, denoted as $\mathcal{A}(\cdot)$.

Discrimination Loss. To achieve accurate device discrimination, we leverage cross entropy loss to directly optimize the model’s classification performance. For a batch of source samples $\{(x_b^s, y_b^s)\}_{b=1}^B$, after data augmentation $\mathcal{A}(x_b^s)$, the backbone $f_e(\cdot)$ generates feature embedding z_b^s . Classifier f_c then projects z_b^s to produce logits $\hat{y}_b^s = f_c(z_b^s)$ over K device classes. The cross entropy loss is computed as:

$$\mathcal{L}_{CE} = -\frac{1}{B} \sum_{b=1}^B \sum_{k=1}^K \mathbf{1}(y_b^s = k) \log \left(\frac{e^{(\hat{y}_b^s)_k}}{\sum_{j=1}^K e^{(\hat{y}_b^s)_j}} \right). \quad (3)$$

Here, $\mathbf{1}(y_b^s = k)$ is an indicator function for the true class label of sample b , and $(\hat{y}_b^s)_k$ denotes the k -th element of \hat{y}_b^s . \mathcal{L}_{CE} ensures the model learns maximally discriminative device-specific features.

Class-level Alignment Loss. Given \mathcal{L}_{CE} , device-specific features may still be entangled with the channel variations

in RF samples caused by the distance change. To resolve this problem, we introduce a supervised N-pairs contrastive loss (Khosla et al. 2020) that explicitly enforces all transformed views of the RF fingerprint from the same device to converge toward a unified representation. For a batch of source samples $\{(x_b, y_b)\}_{b=1}^B$ and their augmented embeddings $z_b = f_e(\mathcal{A}(x_b))$, the class-level alignment loss is defined as:

$$\mathcal{L}_{class} = -\frac{1}{B} \sum_{b=1}^B \frac{1}{|\mathcal{P}_b|} \sum_{p \in \mathcal{P}_b} \log \frac{\exp(\mathbf{z}_b^\top \cdot \mathbf{z}_p / \tau)}{\sum_{a \in \mathcal{A}_b} \exp(\mathbf{z}_b^\top \cdot \mathbf{z}_a / \tau)}, \quad (4)$$

where $\mathcal{A}_b \equiv I \setminus \{b\}$, $I \equiv \{1, \dots, B\}$ denotes the index set of all augmented samples; $\mathcal{P}_b \equiv \{p \in \mathcal{A}_b : \hat{y}_p = \hat{y}_b\}$ is the set of positive samples that belong to the same device as \mathbf{z}_b , and $|\mathcal{P}_b|$ is the cardinality of this set; τ is a temperature parameter controlling feature distribution concentration, and \cdot represents cosine similarity. \mathcal{L}_{class} promotes the clustering of RFF features from the same device, even when distorted by distance-related factors.

Prototype-level Alignment Loss. To further refine feature distributions for each device, we introduce prototype-level alignment, which anchors the features to class-specific references inspired by (Snell, Swersky, and Zemel 2017). This mechanism tightly clusters all features of a particular device around its prototype, enhancing both intra-class consistency and inter-class discriminability. Formally, we define a set of trainable prototypes $\mathcal{U} = \{\mathbf{u}_1, \dots, \mathbf{u}_K\} \in \mathbb{R}^{K \times d}$, where K is the number of devices and d matches the embedding dimension of the feature extractor f_e . For an augmented instance feature \mathbf{z}_b , we compute its prototype assignment distribution via:

$$\mathbf{p}_b^{(k)} = \frac{\exp(\mathbf{u}_k^\top \cdot \mathbf{z}_b / \tau)}{\sum_{j=1}^K \exp(\mathbf{u}_j^\top \cdot \mathbf{z}_b / \tau)}, \quad \forall k \in \{1, \dots, K\}, \quad (5)$$

where τ is the temperature parameter and \cdot denotes cosine similarity. To enforce consistency in the prototype space, we treat each \mathbf{z}_b as an anchor and construct positive pairs $(\mathbf{z}_b, \mathbf{z}_p)$, where \mathbf{z}_p is another augmented view of the same original sample as \mathbf{z}_b . In each training batch, the features from the same sample with different augmentations are split into labeled and unlabeled subsets. For a labeled anchor, its positive counterpart is randomly selected from the same device class; for an unlabeled anchor, its positive counterpart is the nearest neighbor in the embedding space. To ensure positive pairs exhibit consistent prototype association patterns, we define the prototype-level alignment loss as:

$$\mathcal{L}_{proto} = -\frac{1}{|\mathcal{P}_{pos}|} \sum_{i \in \mathcal{P}_{pos}} \log(\mathbf{p}_i \cdot \mathbf{p}'_i), \quad (6)$$

where \mathbf{p}_i and \mathbf{p}'_i are the prototype assignment distributions of the anchor \mathbf{z}_b and its positive instance \mathbf{z}_p , respectively, and \mathcal{P}_{pos} denotes all positive pairs in the batch. Optimizing \mathcal{L}_{proto} further refines feature representations and effectively suppresses residual distance-dependent biases.

Overall Objective in Source Domain. The complete objective function for source domain is formulated as:

$$\mathcal{L}_{source} = \mathcal{L}_{CE} + \lambda_1 \mathcal{L}_{class} + \lambda_2 \mathcal{L}_{proto}, \quad (7)$$

where λ_1 and λ_2 are hyperparameters used to balance the class and prototype loss, respectively.

Target Domain Adaptation

While the source-domain model achieves initial feature discrimination, its identification accuracy may still deteriorate when transceivers are separated at other distances from each other owing to inherent domain shifts. To address this issue, we aim to further adapt the model to the target domain where the separation distances between transceivers are different from those in source domain. Notably, in resource-constrained IoT settings, the devices are incapable in accessing the RF samples in source domain, making source-data-dependent methods (e.g., MMD (Yan et al. 2017) and adversarial alignment (Long et al. 2018)) infeasible. In the following, we elaborate on the target domain adaptation framework, which enables on-device adaptation without accessing source domain data.

Self-Supervised Pseudo-Labeling. Since no ground-truth labels are available for the target domain, we propose a dynamic cluster-based pseudo-labeling strategy to generate reliable supervision signals. Unlike conventional methods (Hu et al. 2021) that directly use source-domain classifiers for label assignment (suffering from noise induced by domain shift), our approach leverages target data’s intrinsic structure to refine pseudo-labels. Given unlabeled target data $\mathcal{D}^t = \{\mathbf{x}_j^t\}_{j=1}^{N_t}$, we initialize the adaptation process using the feature extractor f_e and the frozen classifier f_c from the source model. For each target sample \mathbf{x}_j^t , its initial classification output is $\hat{y}_j^t = f_c(f_e(\mathbf{x}_j^t))$. To reduce domain shift noise, we compute initial class prototypes $\mu_k^{(0)}$ weighted by the confidence scores from the source classifier:

$$\mu_k^{(0)} = \frac{\sum_{x_j^t \in \mathcal{D}^t} \delta_k(\text{Softmax}(\hat{y}_j^t)) \cdot f_e(x_j^t)}{\sum_{x_j^t \in \mathcal{D}^t} \delta_k(\text{Softmax}(\hat{y}_j^t))}, \quad (8)$$

where $\delta_k(\cdot)$ takes the k th logit, and $\mu_k^{(0)}$ denotes the initial prototype for class k . Then each target sample is assigned to the class corresponding to its closest prototype, generating initial pseudo-labels:

$$\hat{y}_t = \arg \min_k \cos(f_e(x_t), \mu_k^{(0)}). \quad (9)$$

To further strengthen the class structure in the target domain, we refine prototypes using these hard pseudo-labels and reassign labels with updated prototypes:

$$\mu_k^{(1)} = \frac{\sum_{x_j^t \in \mathcal{D}^t} \mathbf{1}(\hat{y}_t = k) f_e(x_j^t)}{\sum_{x_j^t \in \mathcal{D}^t} \mathbf{1}(\hat{y}_t = k)}, \quad (10)$$

$$\hat{y}_t = \arg \min_k \cos(f_e(x_t), \mu_k^{(1)}).$$

Here, $\mu_k^{(1)}$ is the refined prototype for class k . The final pseudo-labels \hat{y}_t are used to supervise the training of a target-domain classifier f_t (initialized from f_c) through a cross-entropy loss:

$$\mathcal{L}_{PL} = -\mathbb{E}_{x_t \in \mathcal{D}^t} \sum_{k=1}^K \mathbf{1}(\hat{y}_t = k) \log \delta_k(f_t(x_t)). \quad (11)$$

Information Maximization. Relying solely on pseudo-labeling may lead to error propagation, as the quality of pseudo-labels is inherently limited. Thus, we introduce an information maximization (IM) loss (Hu et al. 2017) that simultaneously maximizes per-sample confidence and class diversity. Let $y = \text{Softmax}(\hat{y}_j^t)$, the IM loss is defined as:

$$\mathcal{L}_{IM}(y) = \mathcal{L}_{\text{ent}}(y) - \mathcal{L}_{\text{div}}(y), \quad (12)$$

$$\text{where } \mathcal{L}_{\text{ent}}(y) = -\mathbb{E}_{x_i \in X_t} \left[\sum_{c=1}^C \delta_c(y) \log \delta_c(y) \right],$$

$$\mathcal{L}_{\text{div}}(y) = -\sum_{k=1}^K \bar{p}_k \log \bar{p}_k,$$

where $\bar{p}_k = \mathbb{E}_{x^t \in \mathcal{D}^t} [\delta_k(y)]$ is the average probability of class k across all target samples. Individual certainty \mathcal{L}_{ent} ensures that each sample is assigned to a single class with high confidence, and global diversity \mathcal{L}_{div} prevents mode collapse, where the classifier overemphasizes a subset of classes and ignores others.

Target Domain Overall Objective. To summarize, building on the source model $f_s = f_c \circ f_e$ and the pseudo-labels generated above, the target-domain adaptation optimizes the target classifier f_t using the following composite objective:

$$\mathcal{L}_{\text{target}} = \gamma \cdot \mathcal{L}_{PL} + \mathcal{L}_{IM}, \quad (13)$$

where $\gamma > 0$ is a hyperparameter that controls the balance between the loss components \mathcal{L}_{PL} and \mathcal{L}_{IM} .

Experiments

Experimental Setup

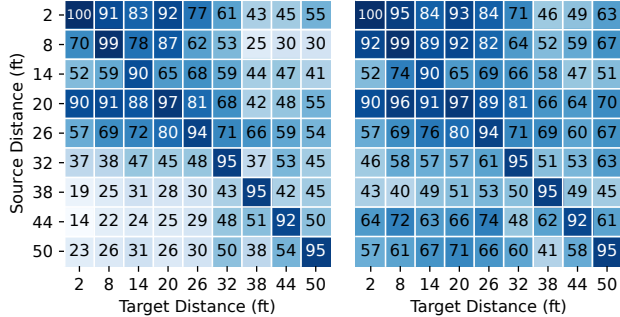
Datasets. We evaluate the distance robustness of our RFF model on two benchmark datasets: (1) **ORACLE** (Sankhe et al. 2019): This dataset contains raw IQ samples captured using 16 USRP X310 Tx and a USRP B210 Rx, with Tx–Rx distances of 2–62 feet (6ft intervals). Owing to significant frame decoding failure rates at the communication range 56/62ft (Zhao, Wang, and Mao 2024), we retain 9 valid distance (2–50 ft) for validation across its two subsets, ORACLE.1 and ORACLE.2, which were collected at different time periods. (2) **LoRa** (Elmaghoub and Hamdaoui 2021): A LoRaWAN-specific RFF dataset collected from 5 identical Pycom IoT devices (Tx) and a USRP B210 Rx. Samples were captured by the same receiver at four distances (5m, 10m, 15m, and 20m) across different transmitters.

To further evaluate the temporal robustness, we include two time-variant datasets: (3) **WiSig** (Hanna, Karunaratne, and Cabric 2022): Comprising wireless samples from 6 COTS Wi-Fi cards captured by the same USRP receiver over four days, with a fixed spatial setup. (4) **CORES** (Hanna, Karunaratne, and Cabric 2020): Samples from 6 Wi-Fi cards deployed in a static grid at the Orbit Testbed over five days.

Baselines. We compare against two baseline categories. For cross-domain RFF tasks, we choose three representative models: OpenRFI (Han et al. 2025), VC-SEI (Wan et al. 2024), and the modified prototypical network (MPTN) from

Dataset	Model <i>Avg.</i>	Ours			OpenRFI		ConvTran		VC-SEI		MPTN	
		S	T	T(Only S)	S	T	S	T	S	T	S	T
ORACLE.1	1S→8T	95.2	65.0	50.9	31.3	25.1	88.3	17.6	93.8	23.1	94.1	22.1
	2S→7T	95.9	81.0	70.6	46.8	36.7	91.7	56.7	93.8	41.7	95.7	55.9
ORACLE.2	1S→8T	95.3	66.0	53.9	29.7	24.6	83.0	20.7	94.6	25.0	93.8	24.8
	2S→7T	95.9	79.0	72.1	47.5	36.0	93.5	56.2	94.8	45.8	95.8	47.8
LoRa	1S→3T	99.3	93.4	74.0	32.0	26.7	59.3	41.1	94.3	43.0	93.8	76.0

Table 1: Distance robustness comparison of RFF models.



(a) Source-only training. (b) With target adaptation.

Figure 4: Distance robustness confusion matrices.

(Zhao, Wang, and Mao 2024). Additionally, we include ConvTran (Foumani et al. 2024), a well-known model in time series classification. We adopt the same preprocessing method as used in our model for the fairness of comparison.

Evaluation Metrics. In line with prior research in the same field, we adopt identification accuracy as the primary metric, defined as the percentage of device instances correctly classified.

Main Results

Overall Performance. To evaluate the distance robustness, we conducted experiments using the ORACLE.1 dataset, where models were trained on a single source domain and test across all other target domains where the samples are captured at different distances from that in source domain. Figure 4 presents the confusion matrices before and after applying domain adaptation, with the vertical axis representing source domains and the horizontal axis representing target domains. Diagonal entries thus reflect source-domain accuracy, while off-diagonal entries indicate cross-domain performance. As shown in Figure 4a, even when trained solely on source-domain data, the model achieves an average accuracy of 50.9% across all source-to-target domain transfers (mean of off-diagonal entries). After applying target-domain adaptation, as shown in Figure 4b, overall performance improves significantly: cross-domain accuracy increases by an average of 14.1%, reaching 65% overall. Additionally, the matrices reveal that models trained on the source domain captured at close distance exhibit stronger

Model	Dataset	ORACLE	LoRa	WiSig	CORES
	Domains	2	5	4	5
Ours	S/T	95/86	98/83	99/88	92/82
OpenRFI	S/T	28/19	80/71	99/77	91/68
ConvTran	S/T	95/36	72/58	99/74	90/66
VC-SEI	S/T	90/34	92/49	99/84	90/79
MPTN	S/T	94/24	91/70	98/75	99/72

Table 2: Temporal robustness for RFF models.

generalization to far-distance scenarios: the average accuracy of off-diagonal entries in the upper triangle (close-to-far distance transfers) reaches 67%. This phenomenon attributes to higher SNR of short-range wireless transmissions, where fingerprint information is less perturbed by noise and thus more reliably extracted. We also note that performance is relatively lower in the 32ft and 38ft rows, which likely stems from more complex multipath effects, increasing the difficulty of fingerprint extraction. More results can be found in the Appendix C.

Comparison Study. To benchmark the distance robustness, we compare our method against four SOTA RFF approaches across diverse experimental setups, as summarized in Table 1. The notation $nS \rightarrow mT$ denotes training on n source domains and testing on m target domains (e.g., 1S→8T: average accuracy training on any single source domain and testing across other eight targets; 2S→7T: training on the high-SNR 2ft source plus one additional source against seven targets). Across all setups, our methods consistently outperform competitors in both source-domain (S) and target-domain (T) accuracy. Specifically, our model trained solely on source-domain data (T(Only S)) already achieves SOTA cross-domain accuracy of 50.9%, validating effective decoupling of fingerprint features from channel variations. Using data from the 2ft and another source domain, target-domain accuracy further improves by 20%, with the 2S→7T setup reaching an average accuracy of 70.6%, demonstrating the benefit of leveraging multiple source domains. In contrast, VC-SEI and MPTN exhibit strong source performance but fail to transfer to targets, as they lack explicit mechanisms to decouple device-specific features from domain-specific noise. While OpenRFI shows moderate cross-distance capability, its overall ac-

Method		ORACLE.1 (Source Distance: 2ft)													
P	Augment					Target Distances					Avg				
	0	1	2	3	4	5	8ft	14ft	20ft	26ft		32ft	38ft	44ft	50ft
							8.5	3.4	0	8.1	0	15.9	6.3	1.7	5.5
	✓	✓	✓	✓	✓	✓	17.8	20.2	19.1	17.4	18.3	16	13.1	12.8	16.8
	✓	✓	✓	✓	✓	✓	18.3	23.1	18.3	17.5	17.9	14.8	15.4	12	17.2
✓							33.9	41.3	40	23.9	26.4	21.6	22.8	21.8	29.0
✓	✓						52.4	59.5	50	53.3	42.3	33.2	37.5	41	46.2
✓		✓					40.7	42.6	47.8	41.7	28.3	24.3	25.9	28	34.9
✓			✓				42.6	51.1	40.2	51.6	35.8	35	27.3	27.8	38.9
✓				✓			57.9	48.4	58.8	42.6	25.3	11.4	10.7	13.5	33.6
✓					✓		65.4	49.2	59.5	34.7	31.7	20.3	21.4	18.8	37.6
✓						✓	68.8	54.3	66.5	37	25.3	17.1	16.2	17.8	37.9
✓							49.3	46.1	54.8	36.2	19.3	15.3	19.4	21	32.7
✓	✓	✓	✓	✓	✓	✓	90.8	82.9	91.7	76.8	61.4	43.8	45.2	55.1	68.5
✓	✓	✓	✓	✓	✓	✓	94.7	83.6	91.9	83.1	64.4	44.3	35.8	58.1	69.5
✓	✓	✓	✓	✓	✓	✓	95	84.1	92.8	83.9	70.3	46.6	48.9	61.9	72.9
Full Model							95.3	84.2	93.1	84.4	71.1	46.1	49.4	62.8	73.3

Notation: P = Preprocessing; 0=Signal Scaling, 1=Permutation, 2=Window Slicing, 3=Time Warping, 4=Magnitude Warping, 5=Window Warping; $s = \mathcal{L}_{class} + \mathcal{L}_{proto}$; $c = \mathcal{L}_{PL}$; $i = \mathcal{L}_{IM}$.

Table 3: Ablation study on model components.

accuracy is much lower than ours. ConvTran, despite architectural advances, overfits to source-domain noise without dedicated enhancement strategies, leading to poor generalization. These results highlight our method’s advantage on distance robustness.

Temporal Robustness. We further validate our method’s robustness to temporal domain shifts—a widely studied RFF task where the surrounding environments drift over time while the spatial configuration of transceivers keep fixed. Beyond ORACLE (2 time domains, all distances) and LoRa (5 time domains, fixed distance), we include two dedicated temporal benchmarks: WiSig (4 time domains) and CORES (5 time domains). As shown in Table 2, our framework achieves 88% and 82% average cross-temporal accuracy on WiSig and CORES (i.e., any single domain trained to generalize across others), outperforming all SOTA methods. This observation aligns with our intuition that features robust to severe spatial perturbations naturally adapt to milder temporal shifts (e.g., gradual environmental drift), as distance-induced domain shifts are more disruptive.

Analytical Experiments

Ablation Study. To analyze the contribution of each component in our cross-distance RFF framework, we conducted ablation experiments using the ORACLE.1 dataset, with the source domain at 2ft and target domain at the distances ranging from 8ft to 50ft for generalization evaluation. As shown in Table 3, we incrementally added modules draw the following conclusions: (1) Preprocessing is an indispensable step. When disabled, the average accuracy drops to 17.2%. Normalization and filtering noisy high-frequency components stabilize subsequent feature learning, providing a consistent input space for downstream optimization. (2) Combinatorial data augmentations (0 to 5 types) play a criti-

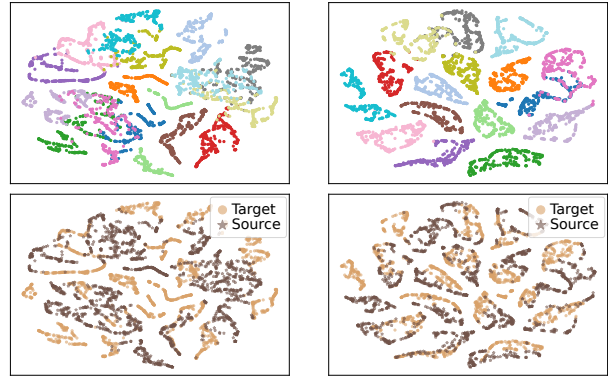


Figure 5: t-SNE visualization of features (ORACLE.1).

cal role in distance-invariant feature learning. By simulating channel effects induced by path loss and multipath effects, these augmentations enable the contrastive optimizer to extract robust features. (3) Our framework exhibits inherent cross-distance capability even without adaptation: using only the source model already achieves an average accuracy of 68.5%. When enhanced with domain adaptation, the overall performance improves to 73.3%.

Hyperparameter Analysis. We optimize hyperparameters via grid search, identifying $\lambda_1 = 1.5$, $\lambda_2 = 1$, and $\gamma = 1$ as optimal values. Detailed experimental setups and hyperparameter analysis are provided in Appendix C.

Visualization of Features Space. To demonstrate the effectiveness of fingerprint disentanglement and domain alignment, we visualize the feature distributions of source (2ft) and target (8ft) domains using t-SNE embeddings (Figure 5). For device separation (16 devices, each with a distinct color), the top-left shows source-only training without augmentation produces blurred device clusters, while the top-right reveals target-domain adaptation sharpens inter-class boundaries; for domain alignment, the bottom-left exhibits distinct source-target clusters under source-only training, whereas the bottom-right shows their blending after adaptation, indicating effective cross-domain feature alignment.

Conclusion

In this paper, we propose the first unsupervised framework for distance-invariant RFF, resolving the critical issue of cross-distance generalization in practical deployment. To this end, we design a preprocessing pipeline that operates directly on raw IQ signals, enhancing fingerprint saliency while suppressing noise. Building on this, our suite of data augmentation strategies simulates diverse channel effects; when integrated with contrastive learning, this enables the extraction of distance-invariant fingerprints. Furthermore, we introduce an unsupervised domain adaptation module leveraging pseudo-labeling and information maximization, achieving robust cross-distance generalization. Extensive experiments validate that our method outperforms SOTA approaches in both cross-distance and cross-temporal scenarios, addressing a critical generalization gap in RFF.

Acknowledgments

This work was supported by National Natural Science Foundation of China under Grant No.62172080, No.62271124, and National Key R&D Program of China No.2022YFB3103404.

References

- Al-Shawabka, A.; Restuccia, F.; D'Oro, S.; Jian, T.; Rendon, B. C.; Soltani, N.; Dy, J.; Ioannidis, S.; Chowdhury, K.; and Melodia, T. 2020. Exposing the fingerprint: Dissecting the impact of the wireless channel on radio fingerprinting. In *IEEE Conference on Computer Communications (INFOCOM)*, 646–655. IEEE.
- Chen, H.; Zhou, R.; Yuan, Q.; Guo, Z.; and Fu, W. 2025. KAN-ResNet-Enhanced Radio Frequency Fingerprint Identification with Zero-Forcing Equalization. *IEEE Sensors Journal*, 25(7): 2222–2252.
- Cui, Z.; Chen, W.; and Chen, Y. 2016. Multi-scale convolutional neural networks for time series classification. *arXiv preprint arXiv:1603.06995*.
- DARPA. 2017. The Radio Frequency Spectrum + Machine Learning = A New Wave in Radio Technology. <https://www.darpa.mil/news-events/2017-08-11a>.
- Demirel, B. U.; and Holz, C. 2023. Finding order in chaos: a novel data augmentation method for time series in contrastive learning. In *Proceedings of the 37th International Conference on Neural Information Processing Systems(NeurIPS)*. Red Hook, NY, USA: Curran Associates Inc.
- Elmaghbbub, A.; and Hamdaoui, B. 2021. LoRa Device Fingerprinting in the Wild: Disclosing RF Data-Driven Fingerprint Sensitivity to Deployment Variability. *IEEE Access*, 9: 142893–142909.
- Foumani, N. M.; Tan, C. W.; Webb, G. I.; and Salehi, M. 2024. Improving position encoding of transformers for multivariate time series classification. *Data Mining and Knowledge Discovery*, 38(1): 22–48.
- Grill, J.-B.; Strub, F.; Altché, F.; Tallec, C.; Richemond, P.; Buchatskaya, E.; Doersch, C.; Avila Pires, B.; Guo, Z.; Gheshlaghi Azar, M.; et al. 2020. Bootstrap your own latent: a new approach to self-supervised learning. In *Proceedings of the International Conference on Neural Information Processing Systems(NeurIPS)*, volume 33, 21271–21284.
- Han, Z.; Xiao, J.; Zhao, Q.; Cui, Z.; Wang, Y.; Zhang, D.; and Ding, W. 2025. Open-world Radio Frequency Fingerprint Identification via Augmented Semi-supervised Learning. In *Proceedings of the AAAI Conference on Artificial Intelligence(AAAI)*, 264–272. AAAI Press.
- Hanna, S.; Karunaratne, S.; and Cabric, D. 2020. Open set wireless transmitter authorization: Deep learning approaches and dataset considerations. *IEEE Transactions on Cognitive Communications and Networking*, 7(1): 59–72.
- Hanna, S.; Karunaratne, S.; and Cabric, D. 2022. WiSig: A Large-Scale WiFi Signal Dataset for Receiver and Channel Agnostic RF Fingerprinting. *IEEE Access*, 10: 22808–22818.
- Hu, W.; Miyato, T.; Tokui, S.; Matsumoto, E.; and Sugiyama, M. 2017. Learning discrete representations via information maximizing self-augmented training. In *Proceedings of the 34th International Conference on Machine Learning(ICML)*, volume 70, 1558–1567. PMLR.
- Hu, Z.; Yang, Z.; Hu, X.; and Nevatia, R. 2021. Simple: Similar pseudo label exploitation for semi-supervised classification. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 15099–15108.
- Khosla, P.; Teterwak, P.; Wang, C.; Sarna, A.; Tian, Y.; Isola, P.; Maschinot, A.; Liu, C.; and Krishnan, D. 2020. Supervised contrastive learning. In *Proceedings of the International Conference on Neural Information Processing Systems(NeurIPS)*, volume 33, 18661–18673.
- Li, H.; Gupta, K.; Wang, C.; Ghose, N.; and Wang, B. 2022. RadioNet: Robust deep-learning based radio fingerprinting. In *IEEE Conference on Communications and Network Security (CNS)*, 190–198. IEEE.
- Li, J.; Yu, Z.; Du, Z.; Zhu, L.; and Shen, H. T. 2024. A comprehensive survey on source-free domain adaptation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 46(8): 5743–5762.
- Liu, J.; and Chen, S. 2024. Timesurl: Self-supervised contrastive learning for universal time series representation learning. In *Proceedings of the AAAI conference on artificial intelligence*, 13918–13926. AAAI Press.
- Long, M.; Cao, Z.; Wang, J.; and Jordan, M. I. 2018. Conditional adversarial domain adaptation. In *Proceedings of the International Conference on Neural Information Processing Systems(NeurIPS)*, volume 31.
- Pan, R.; Chen, H.; Shen, G.; and Chen, H. 2025. Residual Channel Boosts Contrastive Learning for Radio Frequency Fingerprint Identification. *IEEE Wireless Communications Letters*, 14(6): 1728–1732.
- Rashid, K. M.; and Louis, J. 2019. Window-warping: A time series data augmentation of IMU data for construction equipment activity identification. In *Proceedings of the international symposium on automation and robotics in construction(ISARC)*, volume 36, 651–657. IAARC Publications.
- Sankhe, K.; Belgiovine, M.; Zhou, F.; Riyaz, S.; Ioannidis, S.; and Chowdhury, K. 2019. ORACLE: Optimized Radio cLAssification through Convolutional neural nEtworks. In *IEEE Conference on Computer Communications (INFOCOM)*, 370–378.
- Shen, G.; Zhang, J.; Marshall, A.; Peng, L.; and Wang, X. 2021. Radio frequency fingerprint identification for LoRa using deep learning. *IEEE Journal on Selected Areas in Communications*, 39(8): 2604–2616.
- Shen, G.; Zhang, J.; Marshall, A.; Valkama, M.; and Cavallo, J. R. 2023. Toward length-versatile and noise-robust radio frequency fingerprint identification. *IEEE Transactions on Information Forensics and Security*, 18: 2355–2367.
- Snell, J.; Swersky, K.; and Zemel, R. 2017. Prototypical networks for few-shot learning. In *Proceedings of the International Conference on Neural Information Processing Systems(NeurIPS)*, volume 30.

- Sun, M.; Teng, J.; Liu, X.; Wang, W.; and Huang, X. 2025. Few-Shot Specific Emitter Identification: A Knowledge, Data, and Model-Driven Fusion Framework. *IEEE Transactions on Information Forensics and Security*, 20: 3247–3259.
- Um, T. T.; Pfister, F. M.; Pichler, D.; Endo, S.; Lang, M.; Hirche, S.; Fietzek, U.; and Kulić, D. 2017. Data augmentation of wearable sensor data for parkinson’s disease monitoring using convolutional neural networks. In *Proceedings of the 19th ACM international conference on multimodal interaction(ICMI)*, 216–220. ACM.
- Wan, H.; Wang, Q.; Fu, X.; Wang, Y.; Zhao, H.; Lin, Y.; Sari, H.; and Gui, G. 2024. VC-SEI: Robust variable-channel specific emitter identification method using semi-supervised domain adaptation. *IEEE Transactions on Wireless Communications*, 23(12): 18228–18239.
- Wang, X.; Wang, Q.; Fang, L.; Hua, M.; Jiang, Y.; and Hu, Y. 2024. Radio frequency fingerprint authentication based on feature fusion and contrastive learning. *Expert Systems with Applications*, 255: 124537.
- Wu, W.; Hu, S.; Lin, D.; and Liu, Z. 2021. DSLN: Securing Internet of Things through RF fingerprint recognition in low-SNR settings. *IEEE Internet of Things Journal*, 9(5): 3838–3849.
- Xu, H.; and Xu, X. 2021. A transformer based approach for open set specific emitter identification. In *2021 7th International Conference on Computer and Communications (ICCC)*, 1420–1425. IEEE.
- Yan, H.; Ding, Y.; Li, P.; Wang, Q.; Xu, Y.; and Zuo, W. 2017. Mind the class weight bias: Weighted maximum mean discrepancy for unsupervised domain adaptation. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2272–2281.
- Yao, M.; Liu, C.; Xie, L.; and Chi, M. 2025. MambaRF: A Bi-directional Mamba Structure for Radio Frequency Signal Classification of Unmanned Aerial Vehicle. In *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 1–5. IEEE.
- Yin, L.; Fu, X.; Shi, S.; Liu, P.; Lin, Y.; Wang, Y.; and Sari, H. 2023. Few-shot domain adaption-based specific emitter identification under varying modulation. In *IEEE 23rd International Conference on Communication Technology (ICCT)*, 1439–1443. IEEE.
- Zha, X.; Li, T.; Qiu, Z.; and Li, F. 2023. Cross-receiver radio frequency fingerprint identification based on contrastive learning and subdomain adaptation. *IEEE Signal Processing Letters*, 30: 70–74.
- Zhang, J.; Ardizzon, F.; Piana, M.; Shen, G.; and Tomasin, S. 2025. Physical Layer-Based Device Fingerprinting For Wireless Security: From Theory To Practice. *IEEE Transactions on Information Forensics and Security*, 20: 5296–5325.
- Zhang, K.; Wen, Q.; Zhang, C.; Cai, R.; Jin, M.; Liu, Y.; Zhang, J. Y.; Liang, Y.; Pang, G.; Song, D.; et al. 2024. Self-supervised learning for time series analysis: Taxonomy, progress, and prospects. *IEEE transactions on pattern analysis and machine intelligence*, 46(10): 6775–6794.
- Zhang, X.; Li, T.; Gong, P.; Zha, X.; and Liu, R. 2022. Variable-modulation specific emitter identification with domain adaptation. *IEEE Transactions on Information Forensics and Security*, 18: 380–395.
- Zhao, T.; Wang, X.; and Mao, S. 2024. Cross-domain, scalable, and interpretable RF device fingerprinting. In *IEEE Conference on Computer Communications (INFOCOM)*, 2099–2108. IEEE.