Poster Abstract: Cross-Domain Adaptation for RF Fingerprinting Using Prototypical Networks

Steven Mackey
California State University, Sacramento
Sacramento, California, USA
stevester94@gmail.com

Tianya Zhao
Florida International University
Miami, Florida, USA
tzhao010@fiu.edu

Xuyu Wang
Florida International University
Miami, Florida, USA
xuywang@fiu.edu

Shiwen Mao
Auburn University
Auburn, AL, USA
smao@ieee.org

ABSTRACT

Radio frequency (RF) fingerprinting is a hardware feature used in Internet of Things (IoT) applications to identify wireless devices. In this paper, we propose few-shot learning (FSL) and prototypical networks (PTNs) to create a new model that can adapt to a new domain with very few labeled examples. The proposed model can mitigate the domain shift caused by changing RF environments. Experimental results show the proposed method can improve the performance of RF fingerprinting over different domains.

CCS CONCEPTS

• **Security and privacy** → **Mobile and wireless security**; • **Computing methodologies** → **Machine learning**.

KEYWORDS

RF Fingerprinting, Prototypical Networks, Few-shot Learning

ACM Reference Format:
Steven Mackey, Tianya Zhao, Xuyu Wang, and Shiwen Mao. 2022. Poster Abstract: Cross-Domain Adaptation for RF Fingerprinting Using Prototypical Networks. In *The 20th ACM Conference on Embedded Networked Sensor Systems (SenSys '22), November 6–9, 2022, Boston, MA, USA.* ACM, New York, NY, USA, 2 pages. https://doi.org/10.1145/3560905.3568100

1 INTRODUCTION

The proliferation of low-power and cost-effective radio frequency hardware has enabled an unbridled explosion in the number of RF devices. However, this explosive growth also contributes to an ever-expanding attack surface. The most common mitigation of attacks comes from cryptographically secure authentication [1]. However, this solution is not valid for all devices, e.g., low-power devices. Wireless physical layer identification offers a new answer, which seeks to use the physical attributes of radio frequency (RF) devices to identify and authorize them for participation in an RF system[5].

SenSys '22, November 6–9, 2022, Boston, MA, USA

ACM ISBN 978-1-4503-9886-2/22/11…$15.00
https://doi.org/10.1145/3560905.3568100

An RF Fingerprint (RFF) is caused by slight imperfections in the analog circuitry of an RF emitter. Similar to the human fingerprint, the RFF can be employed to identify and classify wireless devices because these slight imperfections introduce non-linearities in the emitter's RF chain and cause a unique fingerprint on the transmitted waveform.

The current works do not consider the domain shift inherent in different RF environments (e.g., different distances, different scenarios). For instance, a substantial drop is in accuracy when training on one domain and testing on another with a traditional convolutional neural network (CNN) in [4]. The approaches such as traditional CNNs and autoencoders do not consider domain information for classification problems. In this paper, *we deploy few-shot learning (FSL) and Prototypical Networks (PTNs) to build a model that can adapt to a new domain with very few labeled examples.* Considering this feature, the model can mitigate the domain shift caused by changing RF environments without having to handcraft channel equalization algorithms.

2 SYSTEM DESIGN

The system assumes that all labels are known during training (supervised learning) and that there are no unseen labels during prediction. For this work, we define a domain to represent a distinct RF environment, delineated by either position in space or time, with time delineated domains differing in the order of hours to days and space delineated domains varying in the order of several feet.

We leverage three distinct datasets for our work. The first dataset is the ORACLE dataset [4], which consists of 16 USRP X310 devices transmitting 802.11a frames in an auditorium. A USRP B210 receiver was moved between 2 and 62 feet away from the transmitters, in an interval of 6 feet. This was repeated a second time, separating the dataset into ORACLE.run1 and ORACLE.run2. Moreover, the dataset was then pre-processed further in order to isolate the preambles of each frame in our method. These frames constitute separate datasets: oracle.run1.framed and oracle.run2.framed.

The second dataset is the CORES dataset [2] with 163 consumer WiFi cards arranged in a grid at the Orbit Testbed, where we use the 58 devices in all 4 days of the dataset. The third dataset is the WiSig dataset [3]. It was also collected in the Orbit Testbed with 41 unspecified USRP receivers capturing 174 COTS WiFi cards. In our work, we utilize data from one receiver and consider 130 emitters in all 4 days. Note that each day constitutes a separate domain.

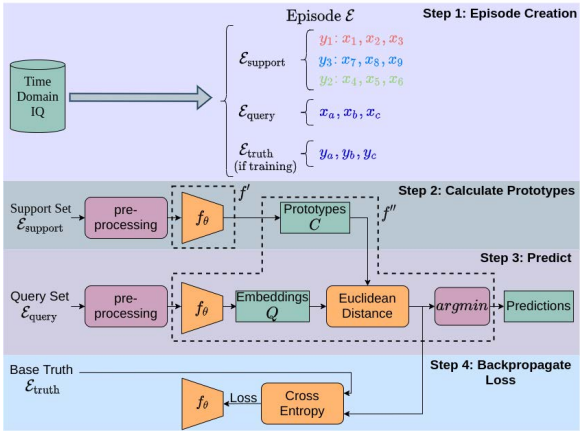

Figure 1: Proposed system steps

Fig. 1 shows the proposed system. The first step is episode creation. We divide the dataset into support set, query set, and base truth, respectively. Then, all examples from the support set are forward passed through the backbone network, yielding an embedding vector for each example in the support set. The embedding vectors for each label in the support set are averaged, resulting in a prototype for each label. Then we compare the Euclidean distance between each query set embedding and prototype for the prediction phase. We consider the distance between each query embedding and each prototype as a likelihood of classification. Then, the cross-entropy loss function is backpropagated to the network by stochastic gradient descent.

3 EXPERIMENTAL EVALUATION

We first apply an identical network except with the last layer configured for traditional classification. Specifically, the number of output nodes is equal to the number of classes for the respective dataset, and the negative log likelihood is applied to the outputs for the backpropagate loss. Overall, we can see from Fig. 2 that a traditional CNN does not generalize to the unseen target domains. We also notice the relatively high target accuracy of the CORES and WiSig datasets. This indicates their domains are not as diverse as the domains in ORACLE.

We now apply our prototypical network approach to the same datasets/domain splits. We also analyze the impact of the normalization. Fig. 3 shows the prototypical network results. A marked improvement in both source and target domain accuracy is visible across all datasets. The efficacy of isolating packets in the ORACLE dataset is also extremely clear. We also see a clear and nearly equal degradation in accuracy when performing either power or magnitude normalization across all datasets except for CORES and WiSig. We also see that run 1 and run 2 of the ORACLE dataset perform nearly identically under all circumstances.

4 CONCLUSION

In this paper, we presented prototypical networks as a viable approach for RF fingerprinting. Our results showed the proposed

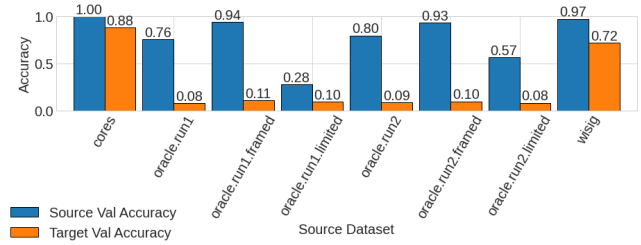

Figure 2: CNN baseline results

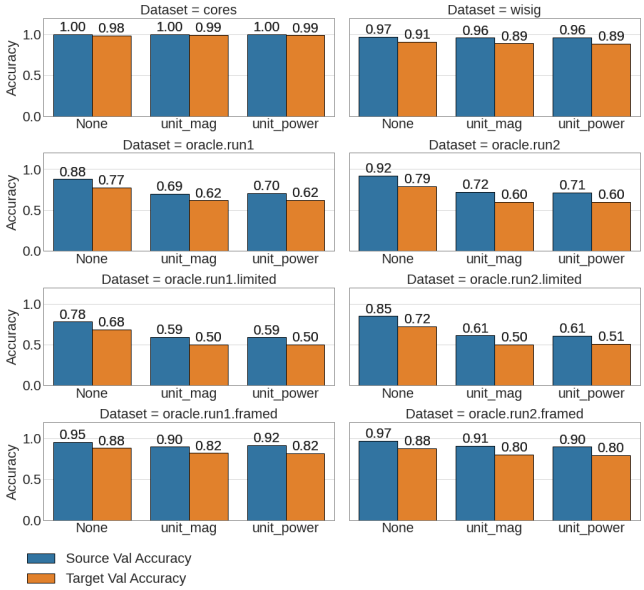

Figure 3: Prototypical network results

method is effective for few-shot learning over different RF environment domains, which can obtain a better performance than the traditional CNN-based method for RF fingerprinting.

ACKNOWLEDGMENTS

This work is supported in part by the NSF (CNS-2107190, CNS-2105416, and CNS-2107164).

REFERENCES

[1] Boris Danev, Davide Zanetti, and Srdjan Capkun. 2012. On physical-layer identification of wireless devices. *ACM Computing Surveys (CSUR)* 45, 1 (2012), 1–29.
[2] Samer Hanna, Samurdhi Karunaratne, and Danijela Cabric. 2021. Open Set Wireless Transmitter Authorization: Deep Learning Approaches and Dataset Considerations. *IEEE Transactions on Cognitive Communications and Networking* 7, 1 (2021), 59–72.
[3] Samer Hanna, Samurdhi Karunaratne, and Danijela Cabric. 2022. WiSig: A Large-Scale WiFi Signal Dataset for Receiver and Channel Agnostic RF Fingerprinting. *IEEE Access* 10 (2022), 22808–22818.
[4] Kunal Sankhe, Mauro Belgiovine, Fan Zhou, Shamnaz Riyaz, Stratis Ioannidis, and Kaushik Chowdhury. 2019. ORACLE: Optimized Radio clAssification through Convolutional neuraL nEtworks. In *IEEE INFOCOM 2019 - IEEE Conference on Computer Communications*. 370–378.
[5] Wenhao Wang, Zhi Sun, Sixu Piao, Bocheng Zhu, and Kui Ren. 2016. Wireless Physical-Layer Identification: Modeling and Validation. *IEEE Transactions on Information Forensics and Security* 11, 9 (2016), 2091–2106.