DEEP LEARNING FOR IDENTIFICATION OF RF EMITTERS BASED ON TRANSMITTER ARTIFACTS

Due Date: March 17th, 2023, at 08:00 AM EST

OVERVIEW

1. Team Name: Electromagnetics! Project Summary

RF Fingerprinting is an active area of research focused on developing solutions to support up to 1,000,000 Internet of Things (IoT) devices / km^2 - International Telecommunication Union (ITU) 5G wireless system Key Performance Indicator (KPI) (Watson, Woods, & Shyy, 2021)

MITRE Corporation's 2021 white paper "*6G and Artificial Intelligence & Machine Learning*" mentioned that one of ITU's "Key Performance Indicators (KPI)" for 5G wireless systems is the ability to "serve up to one million" Internet of Things (IoT) "devices per square kilometer" (Watson, Woods, & Shyy, 2021). Another key enabler of "massive IoT" the authors of this white paper highlight is "low-complexity" and "low-power consumption devices." However, a 2019 IEEE journal article "*IoT Device Security Using RF Fingerprinting*" cautions that "the computational complexity of cryptographic protocols and scalability problems make almost all cryptography-based authentication protocols impractical for IoT" (Nouichi, Abdelsalam, Nasir, & Abbas, 2019). For example, "*Ghost-in-ZigBee: Energy Depletion Attack on ZIgBee-Based Wireless Networks*" discusses the ramifications of an adversarial attack that aims to reduce a ZigBee device's battery life via "luring a node to do superfluous security-related computations" (Cao, et al., 2016).

Therefore, the need to enable the secure operation of large scale IoT device networks has motivated the development of RF fingerprinting (RFF) solutions. RFF is defined as the identification (classification) of a transmitter (TX) based off its emitted signal, and properties unique to that given transmitter and transmission (e.g., TX hardware artifacts). As an example of RFF in industry, an Association of Computing Machinery (ACM) WiseML 22' Conference Spotlight Session "*Systems View to Designing RF Fingerprinting for Real-World Operations*" mentioned that RF fingerprinting research has traditionally focused on "scaling up to large population sizes (e.g., 10,000 emitter populations)" (Kuzdeba, Robinson, Carmack, & Couto, 2022). Also, Captain Benjamin W. Ramsey, USAF, states in the introduction of his Air Force Institute of Technology (AFIT) PhD thesis "*Improved Wireless Security Through Physical Layer Protocol Manipulation and Radio Frequency Fingerprinting*" that "low-cost" devices can achieve ">90% authentication accuracy" via "RF fingerprinting" (Ramsey, 2014).

The goal of our project is for each team member identified in

Table 1 to implement a separate deep learning network based on prior motivating research. We selected the Northeastern University Institute for the Wireless Internet of Things "ORACLE: Optimized Radio Classification through Convolutional Neural Networks[1]" dataset for this project. Our rationale for this decision was to appropriately scope the difficulty of our semester project given the need to balance

[1] https://genesys-lab.org/oracle.

our time between preparing for upcoming quizzes, completing assignment #4, and our semester project (Sankhe, et al., 2019).

Table 1 Team Electromagnetics!

Team Member	Georgia Institute of Technology Email
Greg W Zdor	Gzdor3@gatech.edu
Jordan W Barker	Jbarker63@gatech.edu
Sean McCarty	smccarty30@gatech.edu

2. Related Work

- *Nokia Bell Labs – "While AI/ML will benefit 6G security immensely, we're also likely to see many AI-initiated attacks as well as attacks directed against the vulnerabilities of AI/ML-based mechanisms in our networks. What we need is a comprehensive set of security technology enablers for the 6G era, enhanced and supported by AI/ML, and new principles of cyber resilience" (Labs, 2023)*

"*A Comprehensive Survey on Radio Frequency (RF) Fingerprinting: Traditional Approaches, Deep Learning, and Open Challenges*" is a good introduction to RF Fingerprinting (Jagannath, Jagannath, & Kumar, 2022). The authors of this journal article identified three avenues that researchers have explored to solve this technical problem. First, they summarize "traditional approaches" as the combination of engineered features and machine learning algorithms (e.g., SVM & Random Forest). Second, researchers have evaluated the combination of manual feature engineering and neural networks. A limitation of this approach is that it does not take advantage of a neural network's ability to learn a data representation that yields optimal algorithm performance. Third, researchers have explored applying deep learning to RF signal data. This includes evaluating the performance of a 1-D Convolutional Neural Network (CNN)'s RF fingerprinting performance given unprocessed complex RF signal (i.e., I/Q) data. Current state of the art (SOTA) performance in RFF indicates at least for the dataset of consideration in this work, a 99% classification accuracy may be achieved when accounting for electromagnetic propagation specific augmentations and hardware augmentation in the training set. However, not accounting for these domain specific variations Sankhe et al. found performance to drop to 35% accuracy, demonstrating the importance of "careful impairment allocation" (Sankhe, et al., 2019).

3. Technical Approach

- *"Deep Learning for RF Fingerprinting: A Massive Experimental Study" describes the application of two 1-D Convolutional Neural Network (CNN) architectures to investigate "RF fingerprinting & scalability issues on very large device populations, in the range of 50-10,000 devices" (Jian, et al., 2020)*

We propose analyzing the effectiveness of applying 1-D Convolutional Neural Networks (CNNs) to learn and identify different radios. Our team will research the potential of modern neural network architectures and experiment with multiple hyperparameter configurations to find the most effective neural network for radio frequency fingerprinting. In addition to exploring different CNN-based architectures and hyperparameter configurations, if time allows, we will investigate the impact of RF-specific data

augmentation techniques like simulating a Rican fading channel[2]. Also, MathWorks Corporation's "Wireless Communications Onramp" self-paced training course is a good place to start to learn more about wireless system communications modeling & simulation[3].

Regarding domain specific data augmentations, inherent to over-the-air In-phase / Quadrature (IQ) data problems is modeling the electromagnetics propagation properties of the signal, and ensuring the classification model generalizes across wireless communication channels. Additionally, IQ data, a sequence of complex-valued samples, may be converted into a variety of higher-level feature representations, including the spectrogram, discrete Fourier transform plot, wavelets transform, constellation plot, and more. Each of these forms has its pros and cons, for example a spectrogram captures long time-varying features, for example burst hop pattern, while an IQ constellation representation captures better impairments such as DC offset and IQ imbalance, features directly relevant in RFF. Exploring which representation of IQ is optimal for RF fingerprinting will be assessed as time allows. Overall, this technical approach aims to demonstrate accurate, generalizable, and efficient radio frequency fingerprinting using deep learning techniques.

4. Dataset

Our selection of the Northeastern University Institute for the Wireless Internet of Things "ORACLE: Optimized Radio Classification through Convolutional Neural Networks" dataset was motivated by our need to balance our time between preparing for upcoming quizzes, completing assignment #4 and our semester project.

The dataset for this proposed effort comes from Northeastern University's GENESYS Laboratory and was released as part of their 2018 IEEE INFOCOM paper "*ORACLE: Optimized Radio clAssification through Convolutional NeuraL nEtworks*" (Sankhe, et al., 2019). This classification dataset that we summarize in Table 2 consists of in-phase quadrature (IQ) base banded, digitized, raw, over-the-air received samples of Wi-Fi bursts, namely IEEE802.11a protocol. This dataset used 16 different transmitters (hence 16 unique classes) and one B210 receiver. While assessing whether an RFF model has learned a particular receiver or has generalized across receivers is of interest in RFF, this work will focus on strictly on learning transmitter differences, since this dataset does not contain multiple receivers to evaluate generalization across.

Table 2 Dataset Details

Dataset producers	Northeastern University Institute for the Wireless Internet of Things https://genesys-lab.org/
Dataset Overview	https://genesys-lab.org/oracle (Download links on this page)
Associated Paper Link	https://ieeexplore.ieee.org/document/8737463
Data Size	• Sixteen classes (16 different USRP X310 Software Defined Radios) • Twenty million samples / class
Sampling Rate	Five Million Samples per Second (MSPS)
Center Frequency	2.45 GHz
Data Format	• Raw data in binary file format (read as np.complex128) • Metadata (labels) in standard Signal Metadata (SigMF) format[4]

[2] https://www.gaussianwaves.com/2020/08/rician-flat-fading-channel-simulation/
[3] https://matlabacademy.mathworks.com/details/wireless-communications-onramp/wireless
[4] https://github.com/sigmf/SigMF

BIBLIOGRAPHY

Cao, X., Shila, D. M., Cheng, Y., Yang, Z., Zhou, Y., & Chen, J. (2016, October). Ghost-in-ZigBee: Energy Depletion Attack on ZigBee-Based Wireless Networks. *IEEE Internet of Things Journal, 3 no. 5*, 816-829. doi:10.1109/JIOT.2016.2516102

Cekic, M., Gopalakrishnan, S., & Madhow, U. (2021). Wireless Fingerprinting via Deep Learning: The Impact of Confounding Factors. *IEEE 2021 55th Asilomar Conference on Signals, Systems, and Computers*, 677-684. doi:10.1109/IEEECONF53345.2021.9723393

Hanna, S., Karunaratne, S., & Cabric, D. (2021, March). Open Set Wireless Transmitter Authorization: Deep Learning Approaches and Dataset Considerations. *IEEE Transactions on Cognitive Communications and Networking, 7*(1), 59-72. doi:10.1109/TCCN.2020.3043332

Hermawan, A. P., Ginanjar, R. R., Kim, D.-S., & Lee, J.-M. (2020, May). CNN-Based Automatic Modulation Classification for Beyond 5G Communications. *IEEE Communications Letters, 24*(5), 1038-1041. doi:10.1109/LCOMM.2020.2970922

Jagannath, A., Jagannath, J., & Kumar, P. S. (2022). A comprehensive survey on radio frequency (RF) fingerprinting: Traditional approaches, deep learning, and open challenges,. *Computer Networks, Volume 219*, 1-27. doi:https://www.sciencedirect.com/science/article/pii/S1389128622004893

Jian, T., Rendon, B. C., Ojuba, E., Soltani, N., Wang, Z., Sankhe, K., . . . Ioannidis, S. (2020, March). Deep Learning for RF Fingerprinting: A Massive Experimental Study. *IEEE Internet of Things Magazine*, 50-57. doi:10.1109/IOTM.0001.1900065

Kuzdeba, S., Robinson, J., Carmack, J., & Couto, D. (2022). Systems View to Designing RF Fingerprinting for Real-World Operations. (A. f. Machinery, Ed.) *Proceedings of the 2022 ACM Workshop on Wireless Security and Machine Learning*, 33-38. doi:10.1145/3522783.3529520

Labs, N. B. (2023). *Envisioning a 6G future.* Retrieved from https://www.bell-labs.com/research-innovation/what-is-6g/6g-technologies/ai-native-air-interface#gref

Nouichi, D., Abdelsalam, M., Nasir, Q., & Abbas, S. (2019). IoT Devices Security Using RF Fingerprinting. *IEEE Advances in Science and Engineering Technology International Conferences (ASET)*, 1-7. doi:10.1109/ICASET.2019.8714205

Ramsey, B. W. (2014). *Improved Wireless Security through Physical Layer Protocol Manipulation and Radio Frequency Fingerprinting.* Wright Patterson AFB, OH: Air Force Institute of Technology (AFIT). Retrieved from https://scholar.afit.edu/etd/543

Sankhe, K., Belgiovine, M., Zhou, F., Riyaz, S., Ioannidis, S., & Chowdhury, K. (2019). ORACLE: Optimized Radio clAssification through Convolutional neuraL nEtworks. *IEEE INFOCOM 2019 - IEEE Conference on Computer Communications*, 370-378. doi:10.1109/INFOCOM.2019.8737463

Shen, G., Zhang, J., Marshall, A., & Cavallaro, J. R. (2022). Towards Scalable and Channel-Robust Radio Frequency Fingerprint Identification for LoRa. *IEEE Transactions on Information Forensics and Security*, 774-787. doi:10.1109/TIFS.2022.3152404

Watson, C., Woods, K., & Shyy, D. (2021). *"TW: 6G and Artificial Intelligence & Machine Learning".* Bedford, MA: The MITRE Corporation. Retrieved from https://www.mitre.org/sites/default/files/2021-11/pr-21-0214-6g-and-artificial-intelligence-and-machine-learning.pdf

Yin, G., Zhang, J., Shen, G., & Chen, Y. (2022). FewSense, Towards a Scalable and Cross-Domain Wi-Fi Sensing System Using Few-Shot Learning. *IEEE Transactions on Mobile Computing*, 1-16. doi:10.1109/TMC.2022.3221902