Machine Learning in Adversarial RF Environments

Debashri Roy, Tathagata Mukherjee, and Mainak Chatterjee

For RFML tasks, deep feature learners with an inherent recurrent structure have been shown to perform well. Even so, the field of RFML is still very young, and a lot needs to be done to bridge the gap between the ML community and the wireless community for RFML to be successfully applied for solving large-scale real-life problems.

Abstract

With more and more autonomous deployments of wireless networks, accurate knowledge of the RF environment is becoming indispensable. Various techniques have been developed over the years that can not only assess the RF environment but can also characterize the various radio transmitters (sources) that define the ambient RF environment. Machine learning techniques have shown promise for such characterizations through the development of RF machine learning (RFML) systems delivering autonomous control. Although classical machine learning techniques work well for a large variety of tasks, they have not done as well for RFML. For RFML tasks, deep feature learners with an inherent recurrent structure have been shown to perform well. Even so, the field of RFML is still very young, and a lot needs to be done to bridge the gap between the ML community and the wireless community for RFML to be successfully applied for solving large-scale real-life problems. This article is a step in that direction.

Introduction

The ubiquitous use of wireless services and applications has become ingrained in every aspect of our lives. We depend on wireless technologies not only for our smartphones but also for other applications like telemetry, surveillance, emitter location, radio navigation, jamming, anti-jamming, radar detection, unmanned aerial vehicle (UAV) surveillance, navigation, and location tracking. Applications like teleconferencing and telemedicine, which traditionally depended on wired networks, have migrated to the wireless domain. With such large-scale dependence on use of the RF spectrum, it becomes imperative that we manage and use the limited available spectrum in the most efficient manner possible. In order to do that, one needs to better understand the ambient signal characteristics for optimal deployment of wireless infrastructure and efficient resource provisioning.

Design of new wireless technologies and deployment of wireless networks using those technologies must take into consideration several factors including detection and monitoring of encroachment, ability to predict RF propagation and coverage, techniques to mitigate noise, policies enabling spectrum sharing, characterization of frequencies and waveforms, coverage analysis for optimal deployment, detection and de-confliction, and more importantly, identification of adversarial RF signals. Furthermore, the advent of dynamic spectrum access enabled by the use of software defined radios is pushing the frontiers of wireless communications. These radios are expected to constantly monitor the radio environment and the resulting data can be used to *learn* about their surroundings so that they can intelligently adapt their RF parameters (e.g., operating frequency, bandwidth, waveform, modulation, noise mitigation) to meet their desired objectives.

In order to best use the radio resources in both the spatial and time domains, and to maximize the spectral efficiency, past and current knowledge of the RF signals are important. Although sensing mechanisms can be used to assess the current environment, learning techniques are typically used to analyze the past observations and predict future occurrences of events related to a signal. With the proven success of machine learning (ML) techniques in various domains, such techniques are also being sought for characterizing and understanding the RF environment. Some of the goals of the learning techniques in the RF domain are emitter fingerprinting, emitter localization, modulation recognition, feature learning, attention and saliency, autonomous RF sensor configuration, and waveform synthesis.

ML techniques allow radios to learn and adapt their RF parameters so as to optimize their respective objectives. Such adaptation by the radios is achieved by exposing their configuration options to make the operational parameters flexible and tunable. As a consequence, the configurability and adaptability features open up avenues for manipulation as well where a radio can be induced to learn false information by adversaries [1]. This creates a unique set of challenges in the domain of RFML systems, which makes implementing ML algorithms for RF systems far more challenging.

In this article, we discuss the recent trend of facilitating learning in the RF domain using different ML approaches. These ML techniques have their own strengths and weaknesses depending on the context and the type of dataset being used. We start by broadly classifying the ML techniques into supervised and unsupervised learning, and highlight the various schemes that have been used in the RF domain. Then we discuss five techniques that are currently being widely explored as they exhibit promising results for future implementations. These are:

Digital Object Identifier:
10.1109/MCOM.2019.1900031

Debashri Roy and Mainak Chatterjee are with the University of Central Florida; Tathagata Mukherjee is with the University of Alabama.

0163-6804/19/$25.00 © 2019 IEEE

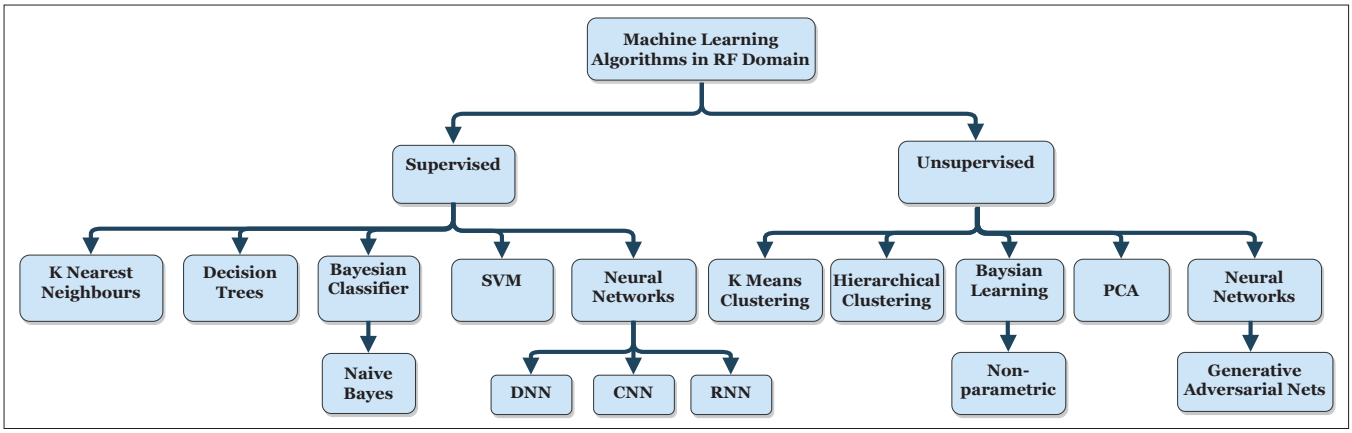

Figure 1. Classification of machine learning algorithms in the RF domain.

- Support vector machine (SVM)
- Deep neural network (DNN)
- Convolutional neural network (CNN)
- Recurrent neural network (RNN)
- Non-parametric Bayesian classifier

SVM and the three neural networks (NNs) perform better with continuous and multi-dimensional datasets, which could be leveraged when the RF signals contain multiple attributes. However, applicability of SVM and NN comes at a cost: they exhibit high variance sensitivity. On the other hand, Bayesian classifiers are advanced statistical techniques for classifying and identifying features. We discuss how a non-parametric version of the Bayesian classifier with an infinite Gaussian mixture model can be used for emitter identification without prior knowledge of the number of transmitters (or classes) or their characteristics. In spite of these developments, we argue that the ML techniques have their limitations in an adversarial setting (i.e., if adversarial RF signals are present during the learning and/or classification process). We show that approaches based on the relatively new generative adversarial nets (GANs) are well suited for learning in adversarial RF environments, and are able to distinguish between adversarial and trusted signals and sources. We implement the discriminative and generative models of a GAN on software defined radios (SDRs) and show that the model could classify four transmitters as well as detect rogue transmitters with very high accuracy.

Machine Learning for RF Signals

The basic work flow of any ML algorithm starts with digesting the feature space. Most often, the quality of the model learned by the algorithm depends heavily on the quality of the features used as input to the algorithm, which in turn depends on the problem. Thus, for example, even though the RF data may consist of the received signal strength indicator (RSSI) values for a given problem, using the raw input in this form may not result in the optimal model. In such cases it might be helpful to transform the raw input features into a higher/lower dimensional feature space that succinctly captures the essence of the problem, thereby making it easier to learn better ML models for the problem.

Figure 1 shows the classification of different ML algorithms that have been used for learning in the RF domain. All learning techniques broadly fall under either supervised or unsupervised training mechanisms.

Supervised Learning in the RF Domain

Supervised learning algorithms are a set of learning algorithms where a set of mutually exclusive labeled data is used for building the learning model (also called training). One of the simplest supervised learning algorithms is based on the *K*-nearest neighbor (KNN) search. The KNN algorithm classifies previously unseen data based on the labels of the nearest data samples that are included in the training set. Algorithms based on KNN searches are usually very computation-intensive, especially in the higher dimensions where the dimension of the feature space factors into the running time of the known algorithms for computing KNNs. This also makes algorithms based on KNN search unsuitable for the RF domain. Furthermore, KNN may not provide optimal performance for a higher-dimensional dataset, which is often the case for RF signal data. Another set of supervised learning algorithms are based on the idea of decision trees. Decision trees are used to assign specific class labels to the items using predictive modeling based on the values of the features associated with the item of interest. The decision trees perform well with high-dimensional data, which could be advantageous for handling multi-dimensional RF data; however, longer training time and lack of accuracy could hinder usefulness in mission-critical applications, as is most often the case with RF processing. Another important class of supervised algorithms are based on the idea of Bayesian classifiers. Bayesian classifiers predict the probability of a given sample belonging to a particular class using a priori knowledge. A specific category of Bayesian classifier is the Naive Bayes, where one assumes that each feature contributes toward the classification and are mutually correlated. Since different parameters from the same kind of radios can be considered as mutually correlated, Naive Bayes can potentially be used successfully in the RF domain. Naive Bayes also works reasonably well with limited training data and is less likely to suffer from overfitting. However, the predictions are less accurate compared to the other methods like neural networks [3, 4]. We defer our discussions on SVMs, DNN, CNN, and RNN to a later section as they have shown promising results in

Unsupervised learning algorithms are a set of algorithms where there is no explicit training phase for building a model from labeled data as is done with the supervised algorithms. These algorithms make inference from unlabeled data, most often exploiting the observed variance of the data and their association relative to each other.

the RF domain and thus deserve to be discussed in detail.

Unsupervised Learning in the RF Domain

Unsupervised learning algorithms are a set of algorithms where there is no explicit training phase for building a model from labeled data as is done with the supervised algorithms. These algorithms make inference from unlabeled data, most often exploiting the observed variance of the data and their association relative to each other. *K*-means clustering is one such learning algorithm where the observed data is partitioned into clusters using information about the distance between the data points (or more generally using the similarity between the observed data points). New data is assigned to a cluster such that the data point is closest to a particular cluster mean. Another set of clustering algorithms are based on the idea of *hierarchies*. Here the items of interest are progressively partitioned into a hierarchy of clusters; the clusters that are higher up in the hierarchy are coarse, whereas the ones that are lower are more fine-grained. Hierarchical clustering can be used for transmitter identification and classification where the number of entities are unknown. Another set of unsupervised algorithms concerns the problem of *dimensionality reduction*. Dimensionality reduction aims to identify a subspace of the feature space such that the projection of the data into the subspace (which has a lower dimension than the feature space) would explain most of the variation observed in the data.

Principal component analysis (PCA) is one such method for dimensionality reduction and data compression. PCA can be very useful for the multi-dimensional RF data as it can be leveraged to speed up the underlying classification or regression algorithms when faster online learning and processing is required.

The final and most important category of unsupervised algorithms for the RF domain is the generative adversarial networks (GANs) [2]. It is a relatively new class of algorithms that have been shown to perform well in adversarial settings. GANs use a generative model that enables the realistic generation of samples from a given distribution, which can then be used to train a discriminator for identifying real samples drawn from the distribution as opposed to fake ones obtained from the generator. We dedicate a section to describing GANs as they have been shown to perform very well in adversarial settings. Likewise, non-parametric Bayesian learning is discussed later due to its relevance in the RF domain.

Important ML Algorithms for RF Domain

In this section, we describe the ML techniques that have been most successful in characterizing the RF environment by identifying and differentiating signals from different kinds of transmitters including broadcast radio, local and wide area data and voice radios, radars, and so on.

Support Vector Machine

SVM has been one of the most successful classical ML algorithms and has been applied to a vast array of problems. SVM is a discriminative classifier using supervised learning and generates an optimal hyperplane for data classification and identification. At its heart, SVM is a binary classifier which assumes that the data is linearly separable and computes the optimal separating hyperplane by solving a quadratic program on the space defined by the training data.

SVM has been applied to the RF domain for transmitter identification using fingerprinting. In [5], Kroon *et al.* presented an RF fingerprinting technique for banning prohibited transmitters from accessing the cellular base stations. They proposed a customized ensemble classifier based on the probability density of an SVM classifier, which achieved 97 percent true positive and 80 percent true negative rates.

However, in recent times, the research trends on RF fingerprinting are shifting toward using raw signal data as compared to using hand engineered features. This has also been facilitated by the availability of automatic feature learning systems like the multi layer perceptron. In [4], Youssef *et al.* investigated different ML strategies including SVM and neural networks using raw I/Q data rather than using any hand-engineered features. I/Q data consists of a complex-valued in-phase (I) and quadrature (Q) component in a signal data constellation. Their implementation produces good training accuracy of 87.6 percent but poor test accuracy of 67.6 percent.

SVM classifiers are relatively easy to implement and can be extended for higher-dimensional data. However, achieving high accuracy remains a challenge. Furthermore, SVM implementations do not typically consider any adversarial situation; thus, when a malicious entity tries to pass as a trusted device, the SVM classifier has no way of determining its true identity and thus would incorrectly classify it as one of the trusted devices. This makes the possibility of using SVMs in adversarial settings quite low. Moreover, SVMs require hand-engineered features even when raw I/Q data is used. As a result of these drawbacks and the recent resurgence of NNs, attention has turned to using the same for RFML applications.

Deep Neural Networks

Deep neural networks have revolutionized the field of artificial intelligence in the last few years. The problems tackled by DNNs range over computer vision, natural language processing, speech processing, and so on. They have been shown to perform better than humans for some of these problems. More recently, [3, 6] have shown the efficacy of using DNNs in RF communication systems. In [3], Shea *et al.* presented a modulation recognition approach using a DNN, achieving nearly 82–87 percent accuracy. They used a synthetic dataset consisting of 11 widely used modulations: 8 digital and 3 analog modulations. Their RF fingerprinting was based on the modulation type, which was used as the primary input feature for the network, which then computes more complex features for model learning.

The automatic features learned by DNNs are often better and more informative than the hand-engineered features used for SVMs, and hence the DNNs yield better accuracy. However, DNNs do not perform as well for datasets with spatial and temporal correlations, which is the

case for the RF domain. RF signals exhibit high spatial correlation (e.g., modulation schemes), high temporal correlation (e.g., I/Q signal data), or both, and in order for a system to work well with RF data, these correlations need to be exploited by the learning system. Furthermore, DNNs have proven to be susceptible to malicious attacks and fail to distinguish rogue transmitters from trusted ones [7] in the presence of active adversaries.

Convolutional Neural Networks

Fully connected DNNs are like standard NNs but with a lot more hidden layers. However, these networks can be augmented with *convolutional layers* for faster training and for enabling the network to learn more compact and meaningful representations. Deep convolutional neural networks (DCNNs) have been shown to be effective for several different tasks in communication. There have been quite a few attempts at using DCNNs for learning different RF parameters. One such effort is presented in [7], where Shea *et al.* presented an optimized DCNN model (with 18 layers) for recognition of modulation schemes for a large synthetic dataset (consisting of 24 modulation schemes), as well as over-the-air data. 94 percent accuracy on synthetic data and 87 percent accuracy on over-the-air data were achieved. Inspired by this and buoyed by the success of the application of CNNs in communication, Youssef *et al.* presented a CNN architecture [4] for deep feature embedding, transmitter identification, and classification.

A CNN works with several filters that capture the *spatial* correlation within the input features for the purpose of computing lower-level features that effectively capture the spatial correlation between the input features. A cascade of such filters are used for propagation of these features through multiple layers, in effect computing more low-level features, thereby giving the best solution for datasets having spatially correlated features. However, it does not perform well in general for datasets where the features are uncorrelated, for example, the *rise time*, which is a feature that can be computed for many RF datasets. CNNs may not perform well if the nature of the correlations change from the training data to the test data. This is because in such cases the features computed through the different filters on the training data will not be applicable to the test data. It is also not recommended for time series data where temporal correlations might exist. Moreover, CNNs have the same limitations as DNNs in regard to immunity from malicious attacks.

Recurrent Neural Networks

Recurrent neural networks are capable of predictions with *time series data*. They have been shown to work well with speaker recognition tasks [8]; inspired by this and based on the fact that the raw RF signal data from the transmitter represent a time series, the RNN can be considered as a potential model to build a system for learning transmitter embedding and classification. One variant of the RNN is long short-term memory (LSTM) [9], which has been successfully used for modeling *temporal data* such as speech. The problem of estimating the noise from the signal requires analysis of the temporal data because the noise characteristics can only be estimated by looking at the received signal over time.

In order to estimate the noise, any system needs to "listen" to the underlying signal for some time and "remember" the same for noise estimation. Previously, NNs lacked this capability. Another issue was the problem of *vanishing gradients* when trying to use *back propagation* with temporal data. However, both these problems were solved by the invention of gated units, such as the LSTM, gated recurrent unit (GRU), and their variants.

In [10], Sreeraj *et al.* presented a 2-layer LSTM model to perform modulation recognition over the synthetic dataset of 11 modulation schemes. Accuracies close to 90 percent were achieved for data with high signal-to-noise ratio coupled with time domain data. This implementation establishes the feasibility of using RNN models for learning RF features related to time domain analysis. However, the effectiveness of LSTM models for learning other RF parameters is still an open research challenge.

The RNN performs well with temporally correlated datasets through the implementation of the concepts of "memory" and "gates." However, it responds poorly to spatially correlated data. RNN implementations also incorrectly classify rogue data as trusted when active adversaries spoof the signals as coming from trusted sources in cases where they are not.

Non-Parametric Bayesian Classifier

A non-parametric Bayes classifier is an unsupervised learning strategy that uses a probability density estimator to determine the probability of an observation belonging to a particular class. Nguyen *et al.* presented a non-parametric Bayesian learning [11] approach to identify wireless devices by characterizing their fingerprints. They considered a device-dependent feature space modeled as multivariate Gaussian distribution, which includes frequency difference and phase shift difference as dominant features. The non-parametric Bayesian learning approach is then employed over the generated distributions and used to identify the unknown number of clusters. Experimental results reveal that the proposed method was capable of clustering 1 to 4 Zigbee transmitting devices with a 100 percent hit rate. Next, they showed that the proposed RF fingerprinting approach can be applied for intrusion detection for Sybil and masquerade attacks by spoofing the medium access control (MAC) address. There is still a dearth of methods that are resilient to any general type of attacks for RFML systems.

Comparison of Different ML Techniques

Analyzing the potential of different kinds of ML approaches in RF parameter learning, we summarize that:

- SVM struggles to achieve high accuracies for higher dimensional datasets
- The DNN is best suited for fixed valued RF parameters (e.g., *rise time*).
- The CNN works well with RF parameters that exhibit high spatial correlations (e.g., *modulation techniques*).

Recurrent neural networks are capable of predictions with time series data. They have been shown to work well with speaker recognition tasks; inspired by this and based on the fact that the raw RF signal data from the transmitter represent a time series, the RNN can be considered as a potential model to build a system for learning transmitter embedding and classification.

A premise for ML techniques to work is that the data used to train the systems is accurate and faithfully represents the distribution that describes the underlying data generation process. Hence, the model learned from this data is able to generalize. However, in an adversarial setting, we cannot anticipate the adversary's moves; thus, there is no "clean" training data to begin with.

- The RNN is the best option for RF parameters with high temporal correlations (e.g., *I/Q signal data*).
- Non-parametric Bayesian classifiers are limited to specific types of applications and datasets.

It must be noted that all the aforementioned ML techniques are susceptible to attackers. Once the attacker gets to know the features used by the learning engine, it becomes easy for the attacker to mislead the learning process. The same applies in the RF domain where an RF transmitter can spoof the signals of others and remain undetected.

Next, we discuss the importance of adversarial learning and how it overcomes the shortcommings of the other ML approaches. We show how GANs are not only robust to the ever-changing wireless environment but are also able to operate in the presence of active adversaries.

Adversarial Learning via GAN

Most of the ML techniques mentioned above can handle static (data) attacks after the model has been trained. Thus, once the trained model is available, the algorithm will not be affected by malicious changes to the test data. This is true for RF data as well. Note, however, that active attacks, where the adversary can modify the training data, are more problematic. In fact, detection of adversarial effects on RF data that exhibit dynamic statistical properties is extremely challenging and nontrivial. A premise for ML techniques to work is that the data used to train the systems is accurate and faithfully represents the distribution that describes the underlying data generation process. Hence, the model learned from this data is able to generalize. However, in an adversarial setting, we cannot anticipate the adversary's moves; thus, there is no "clean" training data to begin with. Moreover, there are ways to pollute the training examples intentionally. Thus, most ML techniques would fail in such settings, as was made evident in [12]. To make matters worse, sophisticated adversaries would exploit the online learning mechanisms used by the ML systems and therefore be able to adopt strategies to mislead the learning process. This motivates the design of *adversarial machine learning* algorithms that can combat the presence of adversaries during and/or after the training phase. This makes it a difficult problem from both the perspective of ML as well as that of communications. Thus, it merits consideration and recently has become an open research area.

Generative Adversarial Network

As mentioned before, one of the main problems of using DNNs for learning in adversarial settings is the possibility that the learned model is inaccurate because of interference of the adversary. In order to circumvent this problem, one possibility is to use DNNs in conjunction with GANs.

Commonly used ML techniques use a *discriminative* model, which learns a function that maps the input data to some desired output class label (i.e., they learn the conditional distribution). On the contrary, a *generative* model tries to learn the joint probability of the input data and labels *simultaneously*, which can be converted to conditional probability for classification via Bayes rule. Generative adversarial networks, introduced by Goodfellow *et al.* in 2014 [2], are a class of artificial intelligence algorithms used in unsupervised ML. In a GAN, a system of two NNs contend with each other. One is called the generator $\mathcal{G}$, and the other is called the discriminator $\mathcal{D}$. The generator's job is to create fake data that would have similar attributes to those in the actual training dataset. The discriminator's job is to identify the generated data as real or fake. As this process continues, both $\mathcal{G}$ and $\mathcal{D}$ get better at generating fake data and identifying them, respectively. Both models learn from each other and are driven toward an equilibrium under certain assumptions.

In the context of RF signals, the adversary is like the generator $\mathcal{G}$, which tries to generate fake RF data and evade detection at the same time. The defender is the discriminator $\mathcal{D}$, which tries to distinguish the fake signal from the real ones. Both the adversary and the defender act in a round-robin manner, where the adversary learns to come up with new strategies to overpower the defense mechanisms employed by the defender. Both learn and adopt better strategies over time. The objective is to train $\mathcal{D}$ to maximize the probability of assigning the correct label (i.e., the detection problem) and train $\mathcal{G}$ simultaneously. Since $\mathcal{G}$'s costs depend on $\mathcal{D}$'s parameters and vice versa, none can control the other's parameters. This situation is best represented as a *game*, more specifically a *min-max game*. The objective of this min-max game is to find the Nash equilibrium that minimizes the maximum loss. GANs are different from most of the current mechanisms that emphasize strengthening the defense mechanism and ignore how the adversary might also strengthen its strategies over time.

GAN Implementation

In order to prove the efficacy of GANs in an adversarial setting, we implement the discriminator $\mathcal{D}$ and the generator $\mathcal{G}$ for the problem of transmitter identification through the use of RF fingerprints.

Data Collection and Design: We design the discriminator and generator separately, as shown in Fig. 2 As for the training phase, we collect over-the-air radio signal data from four software defined radios (`T1–T4`). We used Universal Software Radio Peripheral (USRP) B210s [13]. The adversarial network helps to strengthen both the generative and discriminative models through multiple iterations.

Generator: The adversary helps the discriminator to be robust enough to fight against future real adversaries. We model the generator to learn the distribution of different features from the sample space. The generator is also fed with additive white Gaussian noise along with the features to generate fake signal data. Over time, the generator learns the probability distribution over the sample space of the input.

Discriminator: The discriminator is trained over both real and fake data created by the generator. The model learns by minimizing a cost function during training, which depends on both the generator and the discriminator. In each iteration the discriminator is trained to become smarter to identify the signals generated by the generator. Although in theory the game between the generator and the discriminator can go on indefinitely, depending on the ratio of data and

model density, the discriminator overpowers the generator most of the time. This happens due to the vanishing gradient for the generator. In practice, this phenomenon results in generating more accurate ML classifiers. The discriminator is typically implemented as a deeper network than the generator to get a purposeful implementation of the GAN framework.

Configuration: We train both the generator and discriminator through iterative sequential learning to strengthen the generative model over time. We model the generator as a three-layer NN with *sigmoid* activation, generating `Fake` radio signal data. The discriminator is modeled as a five-layer deep network with a *softmax* activation at the end. We use categorical cross-entropy training on an *Adam* optimizer for gradient-based optimization. We implement the GAN in Python using ML libraries *Keras* as the frontend and *Tensorflow* as the backend.

Results: We notice that the discriminator is able to detect the fake transmitters (generated by the generator) with 50 percent accuracy *before* the GAN training. After several epochs (< 50) of training, the optimal discriminator ($\mathcal{D}^*$) is able to detect the `Fake` transmitters with about 90 percent accuracy, as shown in the confusion matrix in Fig. 3. The proposed model is able to recognize the rogue transmitters as well as categorize the trusted ones (`T1–T4`) with similar accuracy. It is evident that the number of false positives and false negatives in the confusion matrix are low and looks promising for future advancement of GAN oriented research in the RF domain.

Conclusions

In order to optimally use the radio resources, it is imperative to develop techniques that can learn, characterize, and predict the wireless RF environment efficiently. In this article, we discuss some of the supervised and unsupervised learning strategies that have been used to investigate different types of problems in the RF domain. However, most of these methods suffer from an adversarial disadvantage. We demonstrate the efficacy of a GAN-based approach for identification of transmitters in adversarial RF environments.

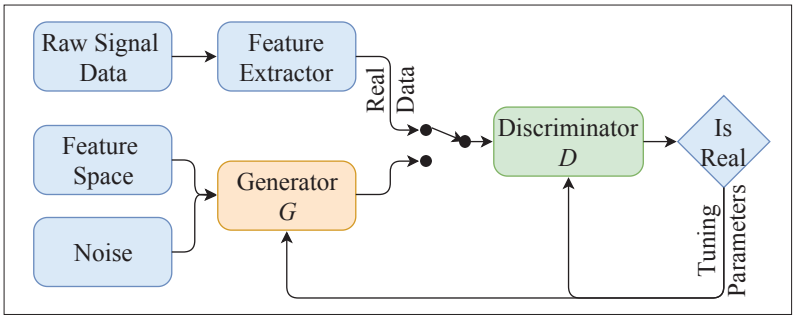

Figure 2. A simplified view of GAN implementation.

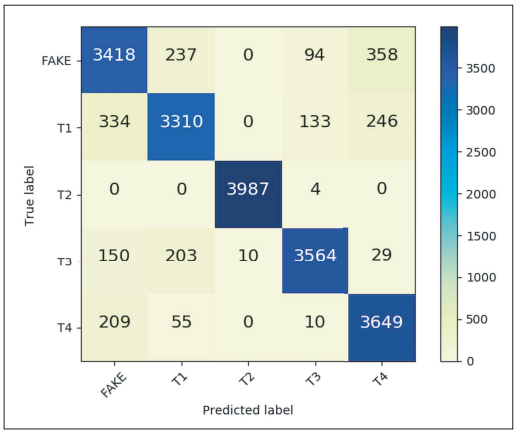

Figure 3. Confusion matrix for determining four trusted and one fake transmitters.

References

[1] S. Bhattacharjee, S. Sengupta, and M. Chatterjee, "Review: Vulnerabilities in Cognitive Radio Networks: A Survey," *ELSEVIER COMCOM*, vol. 36, no. 13, 2013, pp. 1387–98.

[2] I. Goodfellow *et al.*, "Generative Adversarial Networks." *NIPS*, 2014, pp. 2672–80.

[3] T. O'Shea, J. Corgan, and T. Clancy, "Convolutional Radio Modulation Recognition Networks,," *EANN*, 2016, pp. 213–26.

[4] K. Youssef *et al.*, "Machine Learning Approach to RF Transmitter Identification," *CoRR*, vol. arXiv:1711.01559, 2017.

[5] B. Kroon *et al.*, "Steady State RF Fingerprinting for Identity Verification: One Class Classifier Versus Customized Ensemble," *AICS*, 2010, pp. 198–268.

[6] T. O'Shea and J. Hoydis, "An Introduction to Deep Learning for the Physical Layer," *IEEE TCCN*, vol. 3, no. 4, 2017, pp. 563–75.

[7] T. O'Shea, T. Roy, and T. Clancy, "Over-the-Air Deep Learning Based Radio Signal Classification," *IEEE JSTSP*, vol. 12, no. 1, 2018, pp. 168–79.

[8] J. Ren *et al.*, "Look, Listen and Learn-A Multimodal LSTM for Speaker Identification," *AAAI*, 2016, pp. 3581–87.

[9] F. Gers, J. Schmidhuber, and F. Cummins, "Learning to Forget: Continual Prediction with LSTM," IDSIA, tech. rep., 1999.

[10] S. Rajendran *et al.*, "Deep Learning Models for Wireless Signal Classification With Distributed Low-Cost Spectrum Sensors," *IEEE TCCN*, vol. 4, no. 3, 2018, pp. 433–45.

[11] N. T. Nguyen *et al.*, "Device Fingerprinting to Enhance Wireless Security Using Nonparametric Bayesian Method," *IEEE INFOCOM*, 2011, pp. 1404–12.

[12] N. Papernot *et al.*, "Practical Black-Box Attacks Against Machine Learning," *ACM ASIACCS*, 2017, pp. 506–19.

[13] D. Roy *et al.*, "Detection of Rogue RF Transmitters Using Generative Adversarial Nets," *IEEE WCNC*, 2019.

Biographies

Debashri Roy (debashri@cs.ucf.edu) received her M.S. degree in computer science from the University of Central Florida, where she is currently a Ph.D. candidate. Her research interests are in the areas of machine learning, wireless communication, and video QoE.

Tathagata Mukherjee (tm0130@uah.edu) is an assistant professor of computer science at the University of Alabama in Huntsville. He obtained his M.S. and Ph.D. in computer science from Florida State University. His interests are in cyber security, adversarial machine learning, cognitive radio networks, optimization, and graph theory. Prior to his current position ,he was the chief scientist at Intelligent Robotics Inc., a non-profit DoD research lab.

Mainak Chatterjee (mainak@cs.ucf.edu) is an associate professor in the Department of Computer Science at the University of Central Florida, Orlando. He received his B.Sc. degree in physics (Hons.) from the University of Calcutta, his M.E. degree in electrical communication engineering from the Indian Institute of Science, Bangalore, and his Ph.D. degree in computer science from the University of Texas at Arlington. His research interests include economic issues in wireless networks, applied game theory, cognitive radio networks, dynamic spectrum access, and mobile video delivery.