The Physical Layer Identification of Communication Devices Based on RF-DNA

Ying Li[1], Xiang Chen[2], Jie Chang[1], and Yun Lin[1](✉)

[1] College of Information and Communication Engineering,
Harbin Engineering University, Harbin 150001, China
linyun_phd@hrbeu.edu.cn

[2] State Key Laboratory of Complex Electromagnetic Environment Effects
on Electronics and Information System (CEMEE),
Luoyang 471003, Henan, China

Abstract. Traditional methods of improving wireless network security are through software-level device identification, such as IP or MAC addresses. However, these identifiers can be easily changed by software, making wireless network communication a high risk. In response to these risks, radio frequency fingerprinting technology has been proposed. Since the radio frequency fingerprint is an essential feature of the physical layer of the wireless communication device and is difficult to be tampered with, it is widely used to improve the security of the wireless network. Based on the physical layer characteristics of the communication system, this paper has established a relatively complete RF fingerprint identification system to realize the identification and classification of the devices. Two signal starting point detection methods and two RF fingerprint feature extraction methods are studied in this paper. The detailed results are obtained by combining the dimensionality reduction and classification methods. Finally, an optimal identification scheme was found to achieve a classification accuracy of more than 90% when the signal-to-noise ratio is greater than 15 dB.

Keywords: RF fingerprinting · Physical layer identification · Feature extraction · Device classification

1 Instruction

Wireless network security protocols based on cryptographic mechanisms are vulnerable to malicious attacks and face the risk of password leakage. For these risks, people proposed radio frequency fingerprinting technology in 1994 to improve network security. Radio frequency fingerprinting refers to extracting features from radio frequency signals to construct the radio frequency fingerprint of the transmitter, thereby realizing the ID card authentication of the device. RF-DNA refers to the extraction of statistical features from the characteristics extracted from RF signals, which can comprehensively characterize the signal details. Physical layer device identification is a commonly used identification method to identify and classify devices with hardware defects in circuit. Hardware defects include many types, such as time interval errors

© ICST Institute for Computer Sciences, Social Informatics and Telecommunications Engineering 2019
Published by Springer Nature Switzerland AG 2019. All Rights Reserved
S. Liu and G. Yang (Eds.): ADHIP 2018, LNICST 279, pp. 452–460, 2019.
https://doi.org/10.1007/978-3-030-19086-6_50

caused by imperfect clock hardware [1] and sampling errors caused by DAC module hardware defects [2]. Inadequacy of the construction of the local frequency synthesizer will also cause phase shift in the mixing process and cause errors [3]. The nonlinear distortion of the power amplifier can lead to in-band distortion and spectrum regeneration of the digital modulated signal, which is the most considered in RF fingerprint extraction [4–7]. And power amplifiers also have some applications in the spectrum sharing field [8]. In addition, the polarization of the transmit and receive antennas can also be used to study RF fingerprints [9, 10]. The modulator sub-circuit in the device [11] and the multipath effect of the wireless channel [12–14] can also be studied as radio frequency fingerprints. Hardware defects cause the unique physical layer characteristics of the device. These unique features are carried by the transmitted signal to the receiving end. We identify and classify different types of devices by studying the characteristics of the signals, and identify malicious users to achieve the purpose of improving network security. It is also a good choice to add deep learning to the research of equipment classification [15]. But researching classification problems requires a large number of data sets [16]. And it is necessary to consider the movement of the device during communication, which is also an important challenge [17]. RF fingerprints are the physical layer unique characteristics of wireless communication devices and difficult to be tampered with, so they have broad prospects for development.

2 Research Methods

2.1 Overall Framework

This paper has established a relatively complete RF fingerprint identification system to study the identification of wireless devices. As shown in Fig. 1, the system includes signal acquisition, signal detection, feature extraction, feature dimension reduction, classification and recognition. The data set used is the measured data in the laboratory. First we preprocess the signal, that is, detect the change position of the transient signal and intercept valid signals for feature extraction. This step is necessary because the effectiveness of the intercepted signal will affect the final classification accuracy. Variance trajectory detection and Bayesian detection are used to complete this work in this paper. The signal after detecting can be used to extract features. Time domain and wavelet domain methods are applied to it. Since the extracted features have higher dimensions and make the computational complexity larger, the features are reduced in dimension by PCA and LDA. The features after dimension reduction are used for classification with KNN and SVM. Finally, the extracted features are compared with the features in the fingerprint database. Accidental error can be avoided by using the cross-validation in classification process.

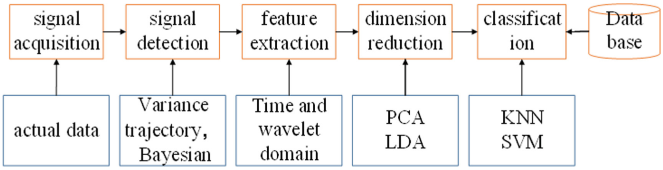


Fig. 1. Fingerprint feature recognition system block diagram

2.2 Signal Acquisition

The signal was captured from 10 transmitters of the same model but different serial numbers at a sampling rate of 40 MHz. The acquisition equipment was a high-performance Agilent oscilloscope. In order to reduce the influence of environmental noise on the signal, the receiver and transmitter are connected by cables. Since the oscilloscope acquisition signal is divided into I/Q paths, one of the two is selected for experiment. For 10 devices, we collect 50 signals from each one and totally 500 signals. Finally, Gaussian white noise was added to the signal.

2.3 Signal Detection

Variance Trajectory Detection. By setting a window function of a certain size, the mean variance of the data in each window is calculated separately, and the difference between adjacent windows constitutes a variance change trajectory. When the change of successive windows is greater than a certain threshold set by experience, the position is marked as a change point position. The variance trajectory sequence is shown in the following formulas.

$$VTx\,(i) = |Wx\,(i) - Wx\,(i+1)|, i = 1, 2, \ldots, L - 1 \tag{1}$$

$$W_x(m) = \frac{1}{N_w} \sum_{k=1+(m-1)N_s}^{1+(m+1)N_s+N_w} [x(k) - \mu_w]^2,\ m = 1,\ 2,\ \ldots,\ L \tag{2}$$

N_w is the length of the signal, N_s is the length of the window calculated at each step, μ_w is the mean of $\{x_{\mathrm{w}}(k)\}$.

Bayesian Detection. Bayesian detection is mainly to equivalent the received signal to a simple piecewise function model, and obtain the maximum value of the probability density function based on the basis function matrix, and the maximum value is the change point position of the signal. A prior knowledge is not necessary for the model to set the threshold and only implements the maximum a posteriori estimate of the change point based on the observed data. The relevant formula is as follows.

$$p(\{w\}|\mathbf{d},\mathbf{I}) \propto \frac{\left[\mathbf{d}^{\mathbf{T}}\mathbf{d} - \mathbf{d}^{\mathbf{T}}\mathbf{G}\left(\mathbf{G}^{\mathbf{T}}\mathbf{G}\right)^{-1}\mathbf{G}^{\mathbf{T}}\mathbf{d}\right]^{-\frac{N-M}{2}}}{\sqrt{\det\left(\mathbf{G}^{\mathbf{T}}\mathbf{G}\right)}} \tag{3}$$

$$G^T = \begin{bmatrix} 1,1,1,1,1,\ldots,1,0,0,0,\ldots,0 \\ 0,0,0,0,0,\ldots,0,1,1,1,\ldots,1 \end{bmatrix} \tag{4}$$

Where d is the signal, G is a diagonal array, N and M are the breakpoint positions of the piecewise function.

2.4 Feature Extraction

Time Domain Features. For the time domain feature extraction, the signal was performed Hilbert transform and then extracted the standard deviation, variance, skewness and kurtosis of the instantaneous amplitude as statistical features. Then standardize the features and remove redundancy. Skewness and kurtosis are represented by the following two formulas, respectively.

Skewness:

$$s = \frac{E(x-\mu)^3}{\sigma^3} \tag{5}$$

Kurtosis:

$$s = \frac{E(x-\mu)^4}{\sigma^4} \tag{6}$$

Where μ is the mean and σ is the standard deviation of the signal.

Wavelet Domain Feature. This part applies the method of multi-scale discrete wavelet transform. By extracting the multi-degree coefficient and taking it as a whole feature set. Then extracting the energy value of each coefficient as the final statistical feature. This iteration can be expressed as the inner product of the sampling signal and the wavelet function. Where j is the scale parameter and k is the translation parameter. K is the number of wavelet coefficients and n is the maximum scale of the wavelet transform. By n-scale decomposition of the signal, n detail coefficients and one approximation coefficient are obtained. The n + 1-dimensional feature vector is obtained by the following formula.

$$C(j,k) = \sum_{n \in Z} z(n)\psi_{j.k}(n) \tag{7}$$

$$F_i = \sqrt{\frac{1}{K}\sum_{k=1}^{K} W_{ik}^2} \tag{8}$$

2.5 Dimension Reduction and Classification

Dimension Reduction. PCA and LDA are two kinds of effective methods of dimensionality reduction. The dimensionality reduction criterion of PCA is to reduce the dimension while retaining the original data information as much as possible, so-called the principal component contribution rate. Following are the principal component contribution rate of PCA for time domain and wavelet domain feature extraction. Through the information of the table and the accuracy requirements of the classification, we choose to reduce to 7 dimensions in this paper (Table 1).

Table 1. The principal component contribution rate of PCA.

Dimension	Time domain	Wavelet domain
3	0.85	0.78
5	0.91	0.86
7	0.97	0.91

Classification. In this paper, KNN and SVM are used to classify features after dimensionality reduction. For the KNN classifier, based on experience and the size of the data set, we set K = 5. SVM has many kernel function types to choose from. In this paper, the SVM is set as Gaussian kernel function.

3 Result Analysis

3.1 Result Analysis of Signal Detection

In order to verify the effect of the two signal detection methods, we generate two kinds of analog signals with a length of 1000 for simulation. One of them is a gradual signal, that is, the value of the first 400 points of the signal is 0, and the value of the last 600 gradually changes from 0 to 1. The second type of signal is a step signal, that is, the value of the first 400 points of the signal is 0, and the value of the last 600 is all 1. The mutation points of both types of signals are set at the 400th point. It is detected by variance trajectory detection and Bayesian detection method respectively. The detected change point position and the running time required for detection are shown in Tables 2 and 3.

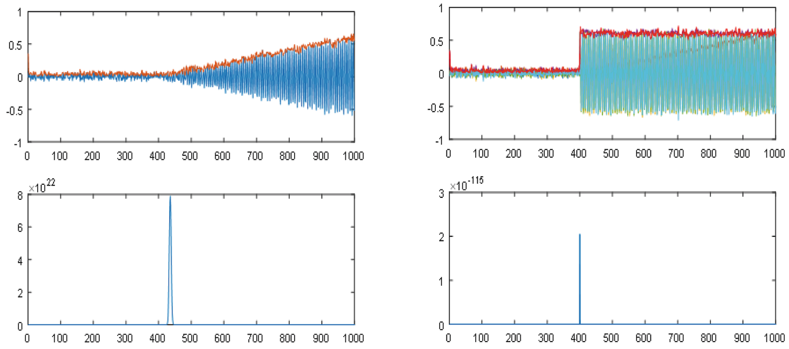
Table 2. Running time and change position for gradual signal.

Methods	Mutation position	Running time
Variance trajectory	427	0.305972
Bayesian detection	415	14.725243

Table 3. Running time and change position for step signal.

Methods	Mutation position	Running time
Variance trajectory	408	0.323541
Bayesian detection	401	0.051159

From the two tables we can see that Bayesian detection method has higher accuracy but longer running time than variance trajectory detection for gradual signals. Both methods have higher detection accuracy for step signal. And the Bayesian detection method has shorter running time for step signal than gradual signal. The following is the simulation result of Bayesian detection method for gradual signal (a) and step signal (b) (Fig. 2).

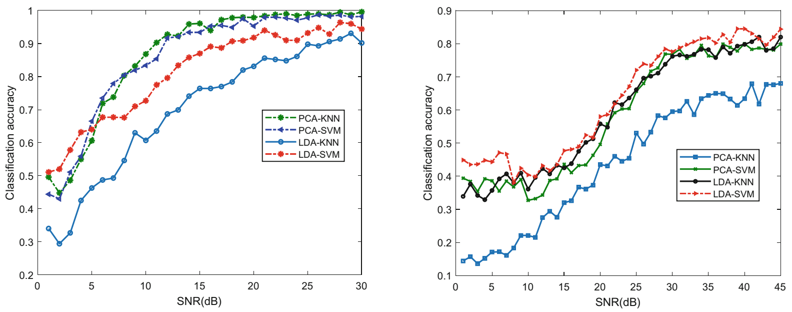


(a) (b)

Fig. 2. Bayesian gradient point detection and step detection results

3.2 Classification Results of Time Domain and Wavelet Domain

In Fig. 3, (a) is the classification accuracy curve of time domain feature under the SNR of 1 to 30(dB) and (b) is the classification accuracy curve of wavelet domain feature under the SNR of 1 to 45(dB). In (a), the top two curves represent the classification accuracy of using the PCA to reduce the dimensionality of the time domain features, which is obviously better than the following two, indicating that the PCA dimensionality reduction method is more effective for the time domain features. Similarly, in (b), the curve with the highest classification accuracy represents the processing of wavelet domain features by LDA dimension reduction and SVM classifiers. From the two figures, we can see that for the time domain feature, PCA dimension reduction and KNN classifier processing is the best choice, which can achieve more than 90% classification accuracy. For the wavelet domain feature, the best feature recognition combination is LDA dimension reduction and SVM classifier, which can achieve more than 80% classification accuracy. So in the end we found an optimal classification scheme, which extracts time domain features from the received signals and uses LDA dimension reduction and KNN classifiers to process the device classification.

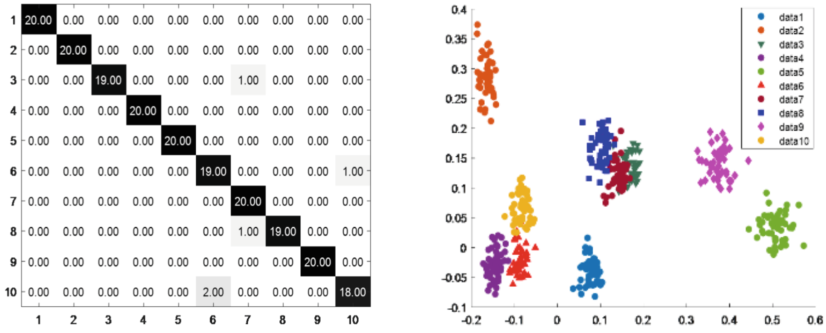


(a) (b)

Fig. 3. Classification accuracy of time domain feature and wavelet domain feature

3.3 Analysis of the Results of the Best Classification Scheme

The classification confusion matrix and scatter plot of the best classification scheme at 15 dB are shown in Fig. 4.



	1	2	3	4	5	6	7	8	9	10
1	20.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
2	0.00	20.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
3	0.00	0.00	19.00	0.00	0.00	0.00	1.00	0.00	0.00	0.00
4	0.00	0.00	0.00	20.00	0.00	0.00	0.00	0.00	0.00	0.00
5	0.00	0.00	0.00	0.00	20.00	0.00	0.00	0.00	0.00	0.00
6	0.00	0.00	0.00	0.00	0.00	19.00	0.00	0.00	0.00	1.00
7	0.00	0.00	0.00	0.00	0.00	0.00	20.00	0.00	0.00	0.00
8	0.00	0.00	0.00	0.00	0.00	0.00	1.00	19.00	0.00	0.00
9	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	20.00	0.00
10	0.00	0.00	0.00	0.00	0.00	2.00	0.00	0.00	0.00	18.00

(a) (b)

Fig. 4. Results of an optimal classification scheme

In Fig. 4, the two graphs (a) and (b) are the classification confusion matrix and the classification scatter plot obtained by the best classification scheme of time domain features respectively. It can be seen from Figure a that most of the sample devices are correctly classified, and only a few samples are incorrectly classified into other categories. From the scatter plot, the distribution of the characteristics of various devices on the coordinate axes can be visually seen. The classification of the 10 types of devices proves the effectiveness of the method used in this paper.

4 Conclusion

This paper completes the process of device identification through a complete physical layer device identification model and focuses on the time domain and wavelet domain feature extraction methods. The effects of different dimensionality reduction and classification methods on classification results are summarized. Finally, a set of optimal classification identification schemes is obtained, that is, the time domain feature extraction is combined with PCA dimensionality reduction and KNN classification. The classification accuracy is more than 90% when SNR is 15 dB. The wavelet domain feature combined with LDA and SVM is also effective, which achieves an accuracy of more than 80% at high SNRs. One possible reason why the time domain feature extraction method is superior to the wavelet domain feature extraction is that the wavelet transform has translation sensitivity, a small disturbance in the signal will have a great influence on the transformation. Finally, although many RF fingerprint identification methods can accurately classify transmitter devices, the same type of devices produced by the same manufacturer is difficult to specialize in the prior art because the similarity of their fingerprint characteristics of the devices are extremely high, so finding more effective fingerprint features may be a promising research direction for RF fingerprinting in the future.

Acknowledgment. This work is supported by the National Natural Science Foundation of China (61771154) and the Fundamental Research Funds for the Central Universities (HEUCFG201830).

This paper is also funded by the International Exchange Program of Harbin Engineering University for Innovation-oriented Talents Cultivation.

Meantime, all the authors declare that there is no conflict of interests regarding the publication of this article.

We gratefully thank of very useful discussions of reviewers.

References

1. Szu, H.H.: Novel identification of intercepted signals from unknown radio transmitters. In: Proceedings of SPIE - The International Society for Optical Engineering, vol. 2491, no. 1, pp. 504–517 (1995)
2. Toonstra, J., Kinsner, W.: Transient analysis and genetic algorithms for classification. In: Conference Proceedings of the IEEE Communications, Power, and Computing IEEE 1995, WESCANEX 95, vol. 2, pp. 432–437. IEEE (1995)
3. Desmond, L.C.C., Yuan, C.C., Tan, C.P., et al.: Identifying unique devices through wireless fingerprinting. In: ACM Conference on Wireless Network Security, WISEC 2008, Alexandria, VA, USA, 31 March–April, pp. 46–55. DBLP (2008)
4. Gao, K., Corbett, C., Beyah, R.: A passive approach to wireless device fingerprinting. In: IEEE/IFIP International Conference on Dependable Systems and Networks, IEEE 2010, pp. 383–392. IEEE (2010)
5. Gard, K.G., Larson, L.E., Steer, M.B.: The impact of RF front-end characteristics on the spectral regrowth of communications signals. IEEE Trans. Microw. Theory Tech. **53**(6), 2179–2186 (2005)

6. Polak, A.C., Dolatshahi, S., Goeckel, D.L.: Identifying wireless users via transmitter imperfections. IEEE J. Sel. Areas Commun. **29**(7), 1469–1479 (2011)
7. Polak, A.C., Goeckel, D.L.: RF fingerprinting of users who actively mask their identities with artificial distortion. In: Signals, Systems and Computers, IEEE 2013, pp. 270–274. IEEE (2013)
8. Zhao, N., Yu, F.R., Sun, H., et al.: Adaptive power allocation schemes for spectrum sharing in interference-alignment-based cognitive radio networks. IEEE Trans. Veh. Technol. **65**(5), 3700–3714 (2016)
9. Nguyen, N.T., Zheng, G., Han, Z., et al.: Device fingerprinting to enhance wireless security using nonparametric Bayesian method. In: Proceedings IEEE, INFOCOM 2011, IEEE 2011, vol. 34, pp. 1404–1412. IEEE (2011)
10. Danev, B., Capkun, S.: Transient-based identification of wireless sensor nodes. In: International Conference on Information Processing in Sensor Networks IEEE 2009, pp. 25–36. IEEE (2009)
11. Polak, A.C., Goeckel, D.L.: Wireless device identification based on RF oscillator imperfections. In: IEEE International Conference on Acoustics, Speech and Signal Processing IEEE, vol. 10, pp. 2492–2501. IEEE (2014)
12. Merchant, K., Revay, S., Stantchev, G., et al.: Deep learning for RF device fingerprinting in cognitive communication networks. IEEE J. Sel. Top. Signal Process. **12**(1), 160–167 (2018)
13. Li, Z., Xu, W., Miller, R., et al.: Securing wireless systems via lower layer enforcements. In: ACM Workshop on Wireless Security ACM, pp. 33–42. ACM (2006)
14. Liang, X., Greenstein, L., Mandayam, N., et al.: Fingerprints in the ether: using the physical layer for wireless authentication. In: IEEE International Conference on Communications IEEE, pp. 4646–4651. IEEE (2009)
15. Tu, Y., Lin, Y., Wang, J., et al.: Semi-supervised learning with generative adversarial networks on digital signal modulation classification. CMC-Comput. Mater. Continua **55**(2), 243–254 (2018)
16. Zhou, J.T., Zhao, H., Peng, X., et al.: Transfer hashing: from shallow to deep. IEEE Trans. Neural Netw. Learn. Syst. **29**(12), 6191–6201 (2018)
17. Zheng, Z., Sangaiah, A.K., Wang, T.: Adaptive communication protocols in flying Ad Hoc network. IEEE Commun. Mag. **56**(1), 136–142 (2018)