Appendix: Feature Correlation Based RF Fingerprinting Identification for Walkie-Talkies Push-to-Talk Communications

Renhe Fan, Longcheng Li, Pengfei Sun, Yang Yu, Gang Wu

1 Feature Extraction Formula

- **Harmonic Features**: By performing a Fast Fourier Transform (FFT) on the signal, the amplitudes $A(f_0)$ and $A(3f_0)$ corresponding to the fundamental frequency f_0 and its third harmonic $3f_0$ can be obtained. The ratio of the amplitude of the third harmonic to the amplitude of the fundamental frequency can be used to indirectly reflect the nonlinear features associated with IIP3. Similarly, by calculating the amplitude $A(2f_0)$ corresponding to the frequency of the second harmonic $2f_0$, the ratio of the amplitude of the second harmonic to the amplitude of the fundamental wave can be obtained. This ratio can also reflect the degree of nonlinear distortion to a certain extent:

$$F_{\mathrm{H3}} = 10\log_{10}\left(\frac{A(3f_0)}{A(f_0)+\varepsilon}\right) \tag{1a}$$

$$F_{\mathrm{H2}} = 10\log_{10}\left(\frac{A(2f_0)}{A(f_0)+\varepsilon}\right) \tag{1b}$$

Here, a very small value ε is added to prevent the denominator from becoming zero.

- **I/Q Imbalance Features**: To quantify the amplitude imbalance between the I/Q channels, the standard deviations of the I and Q channel signals, σ_{I} and σ_{Q}, can be calculated. To reflect the phase difference between the I/Q channels, the average phase angle of the signal can be calculated:

$$F_{\mathrm{A}} = 20\log_{10}\left(\frac{\sigma_{\mathrm{I}}+\varepsilon}{\sigma_{\mathrm{Q}}+\varepsilon}\right) \tag{2a}$$

$$F_{\mathrm{P}} = \arg\left(\frac{1}{N}\sum_{n=1}^{N} s_n\right) \tag{2b}$$

Here, s_n is the complex signal at the n-th sample point.

- **DC Offset Features**: The average values of the I and Q channel signals, μ_{I} and μ_{Q}, are calculated separately to reflect the degree of DC offset:

$$F_{\mathrm{DCI}} = \mu_{\mathrm{I}} = \frac{1}{N}\sum_{n=1}^{N} \mathrm{I}_n \tag{3a}$$

$$F_{\mathrm{DCQ}} = \mu_{\mathrm{Q}} = \frac{1}{N}\sum_{n=1}^{N} \mathrm{Q}_n \tag{3b}$$

- **Power Spectral Density (PSD) Features**: The Welch method is used to calculate the PSD $\mathscr{P}(f)$, divide it into different frequency bands, and then sum the PSD of each frequency band as a feature:

$$F_{\mathrm{LP}} = \sum_{f=0}^{f_l} \mathscr{P}(f) \tag{4a}$$

$$F_{\mathrm{MP}} = \sum_{f=f_l}^{f_m} \mathscr{P}(f) \tag{4b}$$

$$F_{\mathrm{HP}} = \sum_{f=f_m}^{f_h} \mathscr{P}(f) \tag{4c}$$

Here, f_l, f_m, and f_h are the frequency band division points.

- **Spectral Features**: Calculate the spectral center to reflect the central position of the spectral energy distribution. Calculate spectral entropy to measure the complexity of the spectrum:

$$F_{\mathrm{SC}} = \frac{\sum_f f\cdot\mathscr{P}(f)}{\sum_f \mathscr{P}(f)} \tag{5a}$$

$$F_{\mathrm{SE}} = -\sum_f p(f)\log_2\left(p(f)+\varepsilon\right) \tag{5b}$$

Here, $p(f) = \frac{\mathscr{P}(f)}{\sum_f \mathscr{P}(f)}$.

- **Error Vector Magnitude (EVM) Features**: The mean square error between the calculated signal and the ideal signal is used to measure the modulation quality:

$$F_{\mathrm{EVM}} = 20\log_{10}\left(\frac{\sqrt{\frac{1}{N}\sum_{n=1}^{N}|s_n - s_{\mathrm{ref}}|^2}}{\sqrt{\frac{1}{N}\sum_{n=1}^{N}|s_{\mathrm{ref}}|^2}}\right) \tag{6}$$

Here, s_{ref} is the ideal signal.

- **Adjacent Channel Power Ratio (ACPR) Features**: The power ratio between the main and adjacent channels of the calculated signal is calculated to reflect the effect of nonlinear distortion on the spectrum:

$$F_{\text{ACPR}} = 10\log_{10}\left(\frac{\text{P}_m}{\text{P}_a+\varepsilon}\right) \tag{7}$$

Here, P_m and P_a are the power of the main channel and the adjacent channel, respectively.

- **Statistical Features**: For the kurtosis feature, the ratio of the fourth central moment to the fourth power of the standard deviation of the real part of the signal is calculated. This indicator measures the sharpness of the amplitude distribution of the signal and reflects nonlinear distortion in the time domain. For the skewness feature, the ratio of the third central moment to the cube of the standard deviation of the real part of the signal is calculated. This indicator measures the symmetry of the amplitude distribution of the signal, and changes in skewness can reflect the degree of DC offset:

$$F_{\text{K}} = \frac{\frac{1}{N}\sum_{n=1}^{N}(x_n-\mu)^4}{(\sigma^2+\varepsilon)^2} \tag{8a}$$

$$F_{\text{S}} = \frac{\frac{1}{N}\sum_{n=1}^{N}(x_n-\mu)^3}{(\sigma+\varepsilon)^3} \tag{8b}$$

Here, x_n is the real part of the signal, μ is the mean, and σ is the standard deviation.

- **Instantaneous Frequency Features**: Use the Hilbert transform to obtain the analyzed signal, calculate the instantaneous phase ϕ_n and instantaneous frequency f_n, and then calculate the mean and standard deviation of the instantaneous frequency:

$$F_{\text{MIF}} = \frac{1}{N-1}\sum_{n=1}^{N-1} f_n \tag{9a}$$

$$F_{\text{SIF}} = \sqrt{\frac{1}{N-1}\sum_{n=1}^{N-1}(f_n - F_{\text{MIF}})^2} \tag{9b}$$

Here, $\mathscr{H}$ denotes the Hilbert transform, $f_n = \frac{1}{2\pi}\cdot\frac{d\phi_n}{dt}$, and $\phi_n = \arg(s_n + j\mathscr{H}s_n)$.

- **Signal Envelope Features**: The average and standard deviation of the signal envelope are calculated. The signal envelope is affected by amplitude imbalance:

$$F_{\text{EM}} = \frac{1}{N}\sum_{n=1}^{N}|s_n| \tag{10a}$$

$$F_{\text{ES}} = \sqrt{\frac{1}{N}\sum_{n=1}^{N}(|s_n| - F_{\text{EM}})^2} \tag{10b}$$

- **Autocorrelation Features**: The autocorrelation function of the real part of the calculated signal is taken and its maximum value is used to reflect the periodicity and correlation of the signal:

$$F_{\text{ACmax}} = \max_k \sum_{n=1}^{N-k} x_n x_{n+k} \tag{11}$$