Device identification based on distortion of power amplifiers excited by swept sine

Tomoaki Hiraoka, Yuichi Miyaji, and Hideyuki Uehara
Dept. of Electrical and Electronic Information Engineering, Toyohashi University of Technology
Hibariga-oka, Tempaku, Toyohashi, Aichi 441-8580, Japan
a) miyaji@ee.tut.ac.jp

Abstract: This letter presents a device identification method based on the distortion of a power amplifier (PA). The distortion of PA is measured by a logarithmic swept sine, which can distinguish harmonic distortion for each order. In this letter, we propose a new feature extraction utilizing each harmonic distortion to accurate device identification in a multipath environment. The proposed method can reduce computational complexity because it requires no calculation of the pseudoinverse matrix and iterative calculation. Experimental results show that the proposed method can achieve a higher identification ratio in an indoor environment.

Keywords: device identification, power amplifier, harmonic distortion, feature extraction

Classification: Sensing

References

[1] M. W. Liu and J. F. Doherty, “Specific emitter identification using non-linear device estimation,” 2008 IEEE Sarnoff Symposium, pp. 1–5, April 2008. DOI:10.1109/SARNOF.2008.4520119
[2] K. Kim, C. M. Spooner, I. Akbar, and J. H. Reed, “Specific emitter identification for cognitive radio with application to IEEE 802.11,” IEEE GLOBECOM 2008 - 2008 IEEE Global Telecommunications Conference, pp. 1–5, Nov 2008. DOI:10.1109/GLOCOM.2008.ECP.404
[3] M. Vanhoef, C. Matte, M. Cunche, L. S. Cardoso, and F. Piessens, “Why MAC address randomization is not enough: An analysis of Wi-Fi network discovery mechanisms,” Proceedings of the 11th ACM on Asia Conference on Computer and Communications Security, ASIA CCS ’16, pp. 413–424, 2016. DOI:10.1145/2897845.2897883
[4] B. R. Koti and G. R. Kumar, “Information leakage detection and protection of leaked information by using the MAC-IP binding technique,” International Journal of Engineering & Technology, vol. 7, no. 1.7, pp. 230–235, 2018. DOI:10.14419/ijet.v7i1.7.10700
[5] O. Ureten and N. Serinken, “Wireless security through RF fingerprinting,” Canadian Journal of Electrical and Computer Engineering, vol. 32, no. 1, pp. 27–33, Winter 2007. DOI:10.1109/CJECE.2007.364330
[6] K. B. Rasmussen and S. Capkun, “Implications of radio fingerprinting on the security of sensor networks,” 2007 Third International Conference on Security and Privacy in Communications Networks and the Workshops
1 Introduction

With the development of information technology, various wireless devices have been used everywhere. Identifying these devices is useful for application such as intrusion detection of unauthorized devices, obtaining statistical information by collecting user information, and cognitive radios [1, 2]. Identity (ID) information such as media access control (MAC) address is used for devices identification. However, the devices identification using the ID information may be considered as a risk of ID spoofing [3, 4]. Therefore, identification methods using the transmitter’s external components characteristics seem acceptable approach. For example, a method using electromagnetic characteristics of the radio frequency waveform [5, 6], I/Q offset and error vector magnitude [7], and the power spectrum density of the received signal [1, 8] have been proposed.

In [9], they focus on the nonlinearity characteristics of a power amplifier (PA) included in the transmitter. Because of alternative designs and fabrication randomness generating different nonlinearity characteristics, the nonlinearity characteristics are unique to every transmitter. However, this method is not effective in a multipath environment because this method is based on
Table I. Comparison of each methods.

| Method       | Transmission signal | Multipath | Iterative | Experimental environment |
|--------------|---------------------|-----------|-----------|--------------------------|
| Z. Yu [9]    | Training sequence   | Unsupported | No        | Simulation               |
| M. W. Liu [10] | Log-SS            | Supported  | Yes       | Simulation               |
| proposed     |                     | Supported  | No        | Real                     |

a line-of-sight (LOS). In [10], they propose a method corresponding to the multipath environment. This method requires calculation of pseudoinverse matrix and iterative calculation to obtain the nonlinearity characteristics. Calculation of the pseudoinverse matrix has high time complexity and there is a problem of trade-off between accuracy and processing time.

In this letter, we measure the harmonic distortion of PA using a Logarithmic-Swept Sine (Log-SS) signal [11, 12], which is one of the sweep signal. Table I shows the characteristics of each device identification methods using the nonlinearity characteristics of a power amplifier. By using the Log-SS signal, the magnitude of the distortion can be measured for each order and the distortion can be calculated without using the iterative calculation. Moreover, proposed method supports multipath environment by defining the feature not depending on the propagation path. Evaluation experiments are done on real devices, not simulations. By this experiment, it can be evaluated whether the devices can be identified by fabrication randomness.

2 Nonlinear distortion analysis

2.1 Logarithmic swept sine (Log-SS)

A Log-SS signal with period $N$ of the frequency domain representation is defined as

$$S[k] = \begin{cases} 
1 & (k = 0), \\
\frac{1}{\sqrt{k}} \exp(-j\alpha k \log k) & (1 \leq k \leq \frac{N}{2}), \\
S^*[N - k] & (\frac{N}{2} < k < N),
\end{cases}$$

$$(1)$$

$$\alpha = \frac{2m\pi}{N \log \frac{N}{2}},$$

where $k$ is a discrete frequency, and $m \in \mathbb{N}$ is a stretch factor. The length of the time interval is called the effective length $J$, and $J = 2m$.

The phase characteristic of the Log-SS signal is expressed as $\phi(k) = -\alpha k \log k$. The group delay characteristics of the Log-SS signal $\tau(k)$ is expressed as

$$\tau(k) = \frac{N}{2\pi} \frac{d\phi(k)}{dk} = \frac{J}{\log(N/2)}(\log k + 1).$$

$$(2)$$

Therefore, the time-frequency characteristic of the Log-SS signal $k(n)$ is expressed as

$$k(n) = e^{-1} \cdot \exp \left[ \frac{\log(N/2)}{J} n \right]$$

$$(3)$$
where $n$ is a discrete time. Eq. (3) represents that the frequency of the Log-SS signal transitions exponentially. By multiplying both sides of the Eq. (3) by $p$, we obtain

$$pk(n) = e^{-1} \cdot \exp\left[\frac{\log(N/2)}{J} \left(n + \frac{J\log p}{\log(N/2)}\right)\right].$$  (4)

Eq. (4) represents that the $p$th order distortion having the frequency $p$ times the fundamental frequency is observed $J\log p/\log(N/2)$ faster than the fundamental response. By using the Log-SS signal, it is possible to measure separately the magnitude of distortion for each order.

2.2 Feature extraction

We define a feature value independent of the propagation characteristics for terminal identification. Let $S_p[k]$ be the log-SS signal with the sweep rate $p$ times Eq. (1). $S_p[k]$ can be expressed as

$$S_p[k] = \begin{cases} 1 & (k = 0) \\ \frac{1}{\sqrt{p}} \exp(-jak \log k/p) & (1 \leq k \leq \frac{N}{2}) \\ S_p[N - k] & (\frac{N}{2} < k < N). \end{cases}$$  (5)

When the Log-SS signal is input to a PA with maximum $P$th order distortion, its output $x[n]$ can be written:

$$x[n] = \sum_{p=1}^{P} a_p s_p[n]$$  (6)

where $a_p$ is the gain of $p$th order distortion, and $s_p[n]$ is the time domain representation of $S_p[k]$. The received signal $R[k]$ can be expressed as

$$R[k] = \sum_{p=1}^{P} a_p H[k] S_p[k] = \sum_{p=1}^{P} R_p[k]$$  (7)

where $H[k]$ is the propagation characteristics, and $R_p[k] = a_p H[k] S_p[k]$ is a $p$th order distortion term included in $R[k]$. $R_p[k]$ is a value observable by using the Log-SS signal. In addition, $S_p[k]$ is a known value. By defining the feature value for identification as shown in Eq. (8), we can obtain the feature value independent of the propagation characteristics as follows:

$$a_p' = 1_{L_h} \sum_{k=0}^{L_h-1} |a_p| = \frac{1}{L_h} \sum_{k=0}^{L_h-1} |R_p[k]/(H[k]S_p[k])| = \frac{1}{L_h} \sum_{k=0}^{L_h-1} |R_p[k]/S_p[k]|$$  (8)

where $L_h$ is the number of samples for the observed impulse response of the $p$th order distortion.

3 Performance evaluation

In order to evaluate the classification accuracy of the proposed method, we set the devices in an indoor environment and calculated the feature value of the proposed method included in PA of the transmitter. Fig. 1 shows the measurement environment. We set a metal plate between transmitter and receiver to block direct waves and create a non-LOS environment. Universal Software Radio Peripheral (USRP)
2 which is a kind of software defined radio was used for transmitter and receiver. The USRP2 used in this experiment contains IRM046U7 which is a PA designed for 5 GHz W-LAN [13].

We transmitted the Log-SS signal from the transmitter and calculated the feature value of PA of the transmitter from the received signal. The period of the Log-SS signal was set to \( N = 65536 \). The sampling frequency was set to 10MHz and the center frequency was set to 5.015 GHz. The feature value was measured 25 times per transmitter and the position of the device was moved by more than half wavelength so that the propagation characteristics changed every measurement. The transmitter USRP was changed and the feature values for a total of five USRPs were measured. Even if the USRP of the transmitter was changed, the USRP of the receiver did not change and consistently used the same one.

Fig. 2 shows the distributions of the feature values. In order to compare with the case without equalization, that is, the feature value is not independent of the propagation characteristics, we calculate the feature value when \( a_1 = 1 \) is set in Eq. (8). Figs. 2(b),(d) show Figs. 2(a),(c) as viewed from the positive \( a_5 \) axis direction, respectively. Before the equalization, the distribution of the feature values varies even with the same USRP. We consider that it was caused by fluctuation of channel gain due to device position fluctuation because the feature values are linearly distributed. In contrast, the variation in the feature values is small after equalization. This is because the proposed feature value is not affected by channel variation.

In order to evaluate the classification accuracy, we classified the feature values by the nearest neighbor method. For the classification, we used 6 dimensional features up to \( P = 7 \) and calculated the similarity between the data by evaluating the Euclidean distance. Classification accuracy was evaluated by \( K \)-fold cross validation with \( K = 5 \). The feature values which are not equalized achieved 60% classification accuracy, while the feature values which are equalized achieved 100% classification accuracy. Under the situation with channel fluctuation, the proposed method can obtain higher classification accuracy.

Moreover, we compared the processing time required to obtain the feature values. Ref [10] contains iterative calculation and pseudoinverse calculation. The time complexity per iteration is \( O((N + L_h)^3) \). The proposed method contains fast Fourier transform of size \( N \) and size \( L_h \) and there is no iterative calculation. The time complexity is \( O(N \log N + L_h \log L_h) \). Therefore, the proposed method is superior in time complexity. We evaluated processing time by computer simulation using a Mac mini with 1.4 GHz dual-core CPU. We also set training symbol period and maximum delay time of propagation path as \( N = 256 \) samples and \( L_h = 3 \) samples, respectively. The processing time is 165 ms for [10] and 0.241 ms for the proposed method. Since the proposed method does not require a pseudo inverse matrix and
iterative calculation, the processing time is reduced.

4 Conclusion

In this letter, we focused on a distortion of the PA and measured the distortion by the Log-SS signal. By using the fact that the distortion is temporally separated by using the Log-SS signal, we reduce the time complexity by obtaining the magnitude of distortion for each order. We proposed a feature value independent of propagation path and evaluate device identification accuracy using it. Performance evaluation shows that the proposed method can achieve high discrimination accuracy in multipath environments. Moreover, processing time can be reduced compared with the conventional method.