Compressed Sensing Based Fingerprint Identification for Wireless Transmitters

Caidan Zhao, 1 Xiongpeng Wu, 1 Lianfen Huang, 1 Yan Yao, 1 and Yao-Chung Chang 2

1 Department of Communication Engineering, Xiamen University, Xiamen, Fujian 361005, China
2 Department of Computer Science and Information Engineering, National Taitung University, Taitung 95092, Taiwan

Received 12 March 2014; Accepted 1 April 2014; Published 29 April 2014

Copyright © 2014 Caidan Zhao et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Most of the existing fingerprint identification techniques are unable to distinguish different wireless transmitters, whose emitted signals are highly attenuated, long-distance propagating, and of strong similarity to their transient waveforms. Therefore, this paper proposes a new method to identify different wireless transmitters based on compressed sensing. A data acquisition system is designed to capture the wireless transmitter signals. Complex analysis wavelet transform is used to obtain the envelope of the transient signal, and the corresponding features are extracted by using the compressed sensing theory. Feature selection utilizing minimum redundancy maximum relevance (mRMR) is employed to obtain the optimal feature subsets for identification. The results show that the proposed method is more efficient for the identification of wireless transmitters with similar transient waveforms.

1. Introduction

Fingerprint identification for wireless transmitters [1–3] is of great significance in safe radio communications [4–7], military communications confrontation, civilian radio monitoring, and so forth [8–11]. Owing to the discreteness of the wireless transmitter devices and inconsistent manufacturing levels [12, 13], the inherent fingerprint features can be extracted from the emitted signals. Hence, different transmitters can be identified from the captured signals in a fast way [14, 15].

The key process of individual identification is to extract the unique signal features that form a valid device fingerprint. The fingerprint of a transmitter should distinctly characterize it from the rest of the transmitters through its unique features presented in the signal waveform. The transient behavior varies from different transmitters due to imperfections in the analog components. Additionally, some wireless transmitter signals, such as the WiFi signal, are sent by frame under the conflict avoidance protocol, which facilitates transient signal acquisition. Therefore, the unique features extracted from transient signals can make a distinction among different transmitters by adopting the intercepted communication signals [16].

The transient feature extraction mostly utilizes the wavelet analysis or envelope analysis in the time-domain. Choe et al. have used the orthogonal Daubechies-4 wavelet to extract the transient signal features which are applied to identify two types of wireless transmitters at a recognition rate of 94.3% [17]. In the literature [18], wavelet coefficients are selected by genetic algorithms, and the rate of identification for communication stations is more than 90%. Statistical features are extracted from the energy envelopes of seven different manufacturers’ Bluetooth devices by Rehman et al., and the average recognition rate reaches up to 99.9% [19]. In the literature [20], the energy envelope, obtained from the instantaneous signal by spectrogram analysis, is polynomial fitted to get the fitting coefficient characteristics. The result illustrates that these coefficient characteristics can effectively distinguish four different network cards. Zhao et al. have utilized complex analytic wavelet transform to obtain the envelopes from the signals and extracted the fitting coefficient characteristics using the Gaussian fitting method. The identification rate is up to 93% [21]. However, the characteristics of wavelet coefficients extracted by the wavelet analysis methods are of little difference and with large number. It requires high recognition performance of the recognizer. In addition, the above mentioned methods result in large differences of the
signal waveforms, and most of their applications are for the short-distance WiFi signals and Bluetooth signals.

The compressed sensing theory can make the dimension of signals lower and all the information of original signal cannot be lost [22], and, therefore, this paper proposes a new method to identify different wireless transmitters based on compressed sensing to deal with the identification signals of high attenuation, long-distance transmission, and high transient waveform similarity. Experiments have been carried out in an outdoor environment. The distance between the transceivers is more than 3 km. Complex analytical wavelet transform is used to obtain the envelope of the transient signal, and features are extracted from the envelope based on the compressed sensing theory. A feature selection utilizing minimum redundancy maximum relevance (mRMR) is adopted to obtain optimal feature subsets for identification. The proposed method has also been compared to the techniques reported in the literatures.

The rest of this paper is organized as follows. In Section 2, signals acquisition is introduced. Furthermore, transient, envelope, and feature extraction of wireless transmitters is explained. Section 3 gives an overview of analysis of recognition results and draws a comparison between the proposed technique and the reported techniques. Finally, the conclusion follows.

2. Methodology

The key step of fingerprint identification for the wireless transmitter is to extract unique features from the signals that different transmitters emit. To use compressed sensing theory, the sparseness of the signal has to be satisfied. However, lots of high frequency coefficients under the sparse-domain exist in a high-frequency signal, such as the wireless signal. Besides, sparse transform from a transmitter signal envelope can greatly increase the sparseness and it is convenient to extract the features using the compressed sensing. The proposed technique of features extraction is shown in Figure 1. Firstly, the envelope is extracted from the signals of starting-point detected according to the complex analytic wavelet transform. Then, the features are obtained from the envelope by utilizing the compressed sensing theory.

2.1. Data Collection. Figure 2 shows the signal acquisition diagram. The sampling frequency is set as 5 MHz. The distance between the transmitting antenna and the receiving antenna is over 3 km. After being down-converted by the receiver, the intermediate-frequency signal is collected by an oscilloscope. The MATLAB instrument toolbox is used for transient and feature extraction. Due to the environmental impact, the receiving SNR of signals is around 10 dB to 20 dB. The transient signal waveforms of 8 different wireless transmitters, which are used for recognition analysis, are actually acquired as shown in Figure 3. The figure shows that there is no substantial difference among these waveforms. A total amount of 200 waveforms were captured and stored from each wireless transmitter during the device discovery mode.

2.2. Transient Extraction. Accurate transient detection is the key step to extract the features. Huang et al. have proposed a method of mean change point detection [23] to analyze the transient based on phase detection [24]. Its idea is to magnify the difference between the statistic of samples before and after the section, and the position of the maximum difference is determined to be the start of transient. The principle is as follows.

The sample sequence $x_1, x_2, \ldots, x_N$ is known ($x_i$ corresponds to the phase variance in the method of phase detection).
Firstly, let $i = 2, 3, \ldots, N$. The samples can be divided into two sections: $x_1, x_2, \ldots, x_{i-1}$ and $x_i, x_{i+1}, \ldots, x_N$ for each number of $i$; then we can calculate the mean and statistics of each section. Consider

$$S_i = \sum_{t=1}^{i-1} (x_t - \bar{X}_1)^2 + \sum_{t=i}^{N} (x_t - \bar{X}_2)^2.$$  \hspace{1cm} (1)

Secondly, $\bar{X}$ and $S$ are calculated, which represent the average and statistics of the original sample, respectively. Consider

$$S = \sum_{t=1}^{N} (x_t - \bar{X})^2.$$ \hspace{1cm} (2)

Thirdly, the curve of $S - S_i$ is plotted. The position of maximum of the curve would be the change point.

As shown in Figure 4, the detection for the start of wireless transmitter transient is accurate. Compared with the method of variance fractal dimension threshold detection [25], this method can increase the accuracy of detection without the limit of a threshold.

2.3. Envelope Extraction. In a wireless communication environment, because of the interference of the background noise, signal envelope extracted by conventional Hilbert transform, which is shown in Figure 5(b), cannot effectively eliminate the effect of the random noise. It causes an adverse influence on the subsequent feature extraction [26]. In the literature [19, 20], the authors have proposed a method of spectrogram analysis to obtain transient envelope. However, due to the limitation in short time window, the transient envelope exhibits an obvious distortion during the extraction process. This extraction result is shown in Figure 5(c).

Recently, many scholars have put forward the method to extract the envelope from the signal based on complex analytical wavelet transform [26–28]. The envelope obtained by this method has an inhibited effect on random noise and shows a smooth envelope curve whose distortion is relatively petty.

The simplified expression of Morlet wavelet is as follows:

$$\phi (t) = \frac{1}{\sqrt{2\pi} \sigma^2} e^{-\frac{\pi f_0^2}{2\sigma^2}} e^{-t^2/2\sigma^2}, \quad \sigma > 0,$$ \hspace{1cm} (3)

where $\sigma$ is the wavelet shape parameter (also known as bandwidth) and $f_0$ is the wavelet center frequency.
For time-discrete signal $x(t), t = n\Delta T, n = 1, 2, \ldots, N - 1$, where $\Delta T$ is the sampling period, Morlet wavelet transform of the signal can be expressed as follows:

$$W_x (b, a) (m) = \frac{\Delta T}{a \sqrt{2\pi}\sigma^2} \sum_{n=0}^{N-1} x(n) \exp \left[ -\frac{(n\Delta T - b)^2}{2a^2\sigma^2} \right] \times \exp \left( -i2\pi f_0 \frac{n\Delta T - b}{a} \right),$$
(4)

where $b = m\Delta T, m = 1, 2, \ldots, N - 1, a$ is a scale factor. When $a = f_0/f_1, W_x (b, a)(m)$ can get its maximum value. In the expression of $a = f_0/f_1, f_0$ is the wavelet center frequency and $f_1$ is the signal frequency. In order to ensure that the signal amplitude is equal before and after transformation, the envelope amplitude of the signal is $2|W_x (b, a)(m)|$.

To obtain a smooth envelope curve, the parameter $\sigma$ and $f_0$ of Morlet mother wavelet is adjusted. After lots of experiments, as shown in Figure 5(d), when $\sigma = 7$ and $f_0 = 1$, the envelope curve is similar to the original signal and smoother.

The envelope extraction of 8 different wireless transmitters using complex analytical wavelet transform is shown in Figure 6. It can be seen that the envelopes have some differences between each other.

2.4. Feature Extraction Based on Compressed Sensing. The goal of the feature extraction is to form a unique transmitter fingerprint, which distinctly characterizes a transmitter from all of the transmitters. The method of this paper is to extract a feature from the envelope based on compressed sensing.

An orthogonal (sparse) transformation such as DCT (discrete cosine transform) [29] and DWT (discrete wavelet transform) [30] is performed for the envelope signal.
The envelope of 8 different signals using wavelet method

Assuming the length of the envelope signal $y$ is $N$, $y$ is able to be represented by the linear combination of $\psi = \{\psi_1, \psi_2 \cdots \psi_N\}$. Consider

$$y = \sum_{n=1}^{N} \psi_n s_n = \psi s,$$

where $\psi$ is an $N \times N$ orthogonal matrix to form an orthogonal transformation and $s$ is an $N \times 1$ coefficient vector. When $s$ has $K \ (K \ll N)$ nonzero coefficients, the signal $y$ is sparse under the matrix $\psi$.

(1) Linear measurement is the key process of the compressed sensing theory. The signal $y$ is mapped to a measurement vector $z$ which is a lower dimensional data. Consider

$$z = \Phi y = \Phi \psi s,$$

where $z$ is an $M \times 1 \ (M < N)$ vector, namely, the extracted features and $\Phi$ is an $M \times N$ measurement matrix. There are four kinds of measurement matrices including Gauss matrix [22, 31], Bernoulli matrix [22, 31], Toeplitz matrix [32], and sparse matrix [33]. The measurement matrix is critical for feature extraction. It is required that the measurement matrix can make the dimension of feature vector $z$ as low as possible and all information of original signal cannot be lost.

(2) The reconstruction algorithm is used to confirm whether the features contain all information of the original signal, that is to say, whether the measurement $z$ can recover the original signal $y$. The most direct way to reconstruct the signal is optimization problems in $L_0$ norm:

$$\min \ |s|_0$$

s.t. \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ C = \Phi \psi s,$$

where $s$ is the sparse coefficients vector for reconstruction. Since the optimal computational complexity for reconstructing the signal is very high, Bayesian algorithm (BCS) [30] is usually used to find the approximate solution under acceptable complexity. Finally, the reconstruction signal $\hat{y}$ is solved by Formula [16].

It is depicted in Figure 7 that the reconstruction accuracy changes with the measurement matrices and the dimension.
Figure 8: Averages of 18 features extracted from 100 groups of signals for 8 wireless transmitters.

$M$ of measurement vector $z$. It can be clearly shown that the sparse matrix is best suited for feature extraction. Taking into account the reconstruction accuracy, the dimension of the feature vector is set to 230 for DWT, while the dimension of feature vector is set to 160 for DCT.

3. Recognition Results Analysis

In the stage of signal analysis, a recognizer combining multiclassification SVM recognizer [34] with BP neural network recognizer [35] is established. The “one-against-all” method needs more training time. And the appropriate parameters of multiclassification SVM are obtained by cross validation on the training data. To realize the multiclassification SVM, it is feasible to choose the “one-against-one” method, select the Gaussian radial basis function as Kernel, and set nuclear parameter and penalty factor as 1 and 1000, respectively. Furthermore, BP neural network recognizer is built by a three-layer neural network which contains only one hidden layer. The input layer neurons are set to be the number of dimensions of the transient fingerprint feature vector, and the output layer neurons are set to be the number of stations that need to be classified.

As presented in this paper, the number of features extracted by the compressed sensing is too much, which results in a long time identification. The recognition rate decreases for some irrelevant feature interferences. Thus, the use of mRMR [36] to optimize the features is considerable. To directly analyze the effectiveness of the features we select, the distribution of 18 features’ average values extracted from the 100 groups of signal of 8 wireless transmitters is shown in the Figure 8. It can be seen from the picture that the mRMR algorithm is an effective way to distinguish different wireless transmitters.

Figure 9 shows that the SVM identification results, which are through the envelope signal sparse by the DCT and DWT, with mRMR method to select features. The identification effectiveness is changing with the number of selecting features and different properties of sparse bases. In general, the convergence of DCT sparse base is better. If the number of the selecting features is under 50, the identification rate can be above 80%.

Through the envelope signal sparse by DCT sparse base, the SVM recognizer is utilized to identify 8 different wireless transmitters. For each transmitter, 100 groups are used for training and 100 groups for testing. Table 1 presents the specific identification results. The data illustrate that the algorithm referred to in this paper is effective to identify transmitter when the average identification rates of most transmitters are over 88% and the average identification rate is 87.75%.

### Table 1: Classification results for different wireless transmitters.

<table>
<thead>
<tr>
<th>Input</th>
<th>Classification</th>
<th>Average rate (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>TX1</td>
<td>90 2 0 0 3 5 0</td>
<td>90</td>
</tr>
<tr>
<td>TX2</td>
<td>1 80 6 1 5 7 0</td>
<td>80</td>
</tr>
<tr>
<td>TX3</td>
<td>1 9 88 1 0 1 0</td>
<td>88</td>
</tr>
<tr>
<td>TX4</td>
<td>3 0 0 91 0 2 4</td>
<td>91</td>
</tr>
<tr>
<td>TX5</td>
<td>0 5 1 0 93 0 1</td>
<td>93</td>
</tr>
<tr>
<td>TX6</td>
<td>3 2 2 5 3 79 3</td>
<td>79</td>
</tr>
<tr>
<td>TX7</td>
<td>5 0 2 3 0 2 88</td>
<td>88</td>
</tr>
<tr>
<td>TX8</td>
<td>1 0 1 2 1 0 2 93</td>
<td>93</td>
</tr>
</tbody>
</table>

Figure 9: The comparison of the recognition results based on different sparse bases.
Table 2: Comparison of identification of a transmitter using three different methods.

<table>
<thead>
<tr>
<th>Comparison metric</th>
<th>Mathematical statistics method</th>
<th>Fitting method</th>
<th>Compressed sensing method</th>
</tr>
</thead>
<tbody>
<tr>
<td>RF fingerprints</td>
<td>Area under the normalized curve, duration, maximum slope, kurtosis, skewness, and variance of curve</td>
<td>Gaussian fitting coefficients</td>
<td>Feature extracted as explained in subsection II</td>
</tr>
<tr>
<td>Envelope extraction</td>
<td>STFT slice method</td>
<td>Wavelet method</td>
<td>Wavelet method</td>
</tr>
<tr>
<td>Classifier</td>
<td>KNN</td>
<td>SVM and BP</td>
<td>SVM and BP</td>
</tr>
</tbody>
</table>

Table 3: Performance comparison for the envelope extraction and feature extraction.

<table>
<thead>
<tr>
<th>Envelope extraction</th>
<th>Feature extraction</th>
<th>Advantage</th>
<th>Disadvantage</th>
<th>Complexity</th>
<th>Time for calculation (s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>STFT slice</td>
<td>Mathematical statistics</td>
<td>Eliminate the effect of the noise</td>
<td>Envelope exhibits a distortion</td>
<td>$O(n^2)$</td>
<td>2.251245</td>
</tr>
<tr>
<td>Wavelet</td>
<td>Fitting</td>
<td>Envelope is smoother and precise</td>
<td>Parameters are determined by experience</td>
<td>$O(n^3)$</td>
<td>3.568356</td>
</tr>
<tr>
<td>STFT slice</td>
<td>Compressed sensing</td>
<td>Simple; the feature is easy to extract</td>
<td>Some features are not representative</td>
<td>$O(n)$</td>
<td>1.324417</td>
</tr>
<tr>
<td>Wavelet</td>
<td></td>
<td>Simple; the features are easy to extract</td>
<td>Fitting function is determined by experience</td>
<td>$O(n^2)$</td>
<td>2.419341</td>
</tr>
</tbody>
</table>

The method proposed in this paper is compared with the works of Rehman et al. [19] and Zhao et al. [20, 21]. Rehman’s work simply extracts the envelope statistic features from the energy spectrum envelope. And Zhao et al. put forward using the fitting coefficients of the envelope as the signal features for identification. Table 2 describes the comparison among Rehman’s work, Zhao’s work, and the proposed work in this paper. The term mathematical statistics method refers to Rehman et al.’s work, fitting method is used for Zhao et al.’s work, and compressed sensing method represents the work proposed in this paper. The 8 RF signal transmitting devices identified in this paper are similar in their transient waveforms. The recognition rates of either Rehman’s or Zhao’s method are relatively lower. Figure 10 shows the specific comparison result. As shown in the figure, the extracted envelope features method based on compressive sensing theory in this paper has many advantages over the existing statistic features and fitting coefficient characteristics methods. On the one hand, extracting envelope by complex analytic wavelet is superior to spectrum slicing. On the other hand, the recognition performance of SVM recognizer is better than both BP neural network recognizer and KNN recognizer.

According to the disadvantages of the previous methods and the merits of our proposed algorithm, performance comparison for the envelope extraction and feature extraction is provided as Table 3.

4. Conclusions

This paper has proposed a fingerprint identification method for wireless transmitter signal based on compressed sensing. Complex analytical wavelet transform is used to obtain the envelope of the transient signal, and features are extracted from the envelope using the compressed sensing theory. A feature selection utilizing minimum redundancy maximum relevance (mRMR) is employed to obtain optimal feature subsets for identification. Finally, the recognition of 8 wireless transmitters by the SVM recognizer and BP neural network recognizer is completely performed. From a series of experiments, it can be concluded that the method proposed in
the paper can effectively identify the transmitter signals, whose transient waveforms are difficult to distinguish with the naked eye. Especially, the method put forward in the paper has better performance in the recognition of wireless transmitter signals with serious interference and signal attenuation when compared to the previous methods, such as extracting statistic features directly from the envelope, fitting coefficient characteristic. What is more, this method may be applied to identify the signals which are transmitted by the transmitters with the same batch and type. According to the results, a future research on the optimization of the Morlet wavelet’s parameter is expected to improve the performance of the proposed method.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

This work is partly supported by NSFC (61271242, 61001072), NCETFJ, and the Fundamental Research Funds for the Central Universities (2012121028).

References


