A Novel Individual Radio Identification Algorithm Based on Multi-dimensional Features and Gray Relation Theory

Hui Han[1], Jingchao Li[2(✉)], and Xiang Chen[1]

[1] State Key Laboratory of Complex Electromagnetic Environment Effects on Electronics and Information System (CEMEE), Luoyang 471003, Henan, China
[2] Electronic Information College, Shanghai Dianji University, Shanghai 201306, China
lijc@sdju.edu.cn

Abstract. With the advent of the Internet of Things, the number of mobile, embedded, and wearable devices are on the rising nowadays, which make us increasingly faced with the limitations of traditional network security control. Hence, accurately identifying different wireless devices through Hybrid information processing method for the Internet of things becomes very important today. To this problem, we design, implement, and evaluate a robust algorithm to identify the wireless device with fingerprints features through integral envelope and Hilbert transform theory based PCA analysis algorithm. Integral envelope theory was used respectively to process the signals first, then the principal component features can be extracted by PCA analysis algorithm. At last, gray relation classifier was used to identify the signals. We experimentally demonstrate effectiveness of the proposed algorithm ixin differentiating between 500 numbers of wireless device with the accuracy in excess of 99%. The approach itself is general and will work with any wireless devices' recognition.

Keywords: Individual radio recognition · Hilbert transform
Integral envelope theory · PCA analysis · Gray relation theory

1 Introduction

Internet of things as an important branch for the information technology, refers to the combination of all kinds of information sensing devices (such as radio frequency identification (RFID) devices, global positioning systems, infrared sensors, laser scanners) with the Internet to form a huge network, whose purpose is to make all the items connected with the network to facilitate the identification and management. Accurately identifying different wireless devices for the Internet of things becomes very important today. It is difficult to draw a conclusion that which factors affect the development of complex system. In 1982, combining the idea of system theory, information theory and cybernetics, Deng proposed the gray system theory, the study object is to extract valuable information by exploiting the limited information from uncertain system with few samples. It can give an effective description of the complex system behavior and development [1].

G. Sun and S. Liu (Eds.): ADHIP 2017, LNICST 219, pp. 130–136, 2018.
https://doi.org/10.1007/978-3-319-73317-3_16

The gray relational analysis theory (GRA) uses the gray relational grade (GRG) model to calculate and analyze, is an essential concept in the grey system theory. It is the fundamental theory for gray system analysis, prediction and decision. Analyzing a system means that it should distinguish the primary and secondary factors and finding the factors which enhance or restrict the development of system. GRA analyzes the system with few samples by quantifying and ordering the factors. In order to quantify the relationship between factors, various forms of correlation coefficient were proposed, such as canonical correlation coefficient and resemblance correlation coefficient. These coefficients are based on the mathematical statistic theory, so they require a great deal of data because it seems difficult to find out the statistical rules or some kinds of typical probability distribution from few data. But in the practical work, it is hard to meet the requirement, the statistical data are limited and the degree of grey information is large, it is not easy to use the method based on mathematical statistic theory to analyze systems. GRA covers shortcoming to some extent, it has no requirement on the sample size and distribution [2]. The following are some common models: Deng's gray relational grade model, proposed by Deng based on the GRA axioms, known as the most fundamental model [3]; B-type gray relational analysis (B-GRA), given by Wang (1989) according to the proximity and similarity between two objects [4]. The gray absolute relational grade (GARG), proposed by Mei in 1992 based on the adjacency degree of absolute trends and relative trends between the factors' time series curves, the T-type gray relational analysis (T-GRA), produced by Tang in 1995 according to the approaching degree of the relative changing trend between the time series curves of factors [5, 6]. The grey slope relational grade (GSRG), created by Dang in 1994 and later developed by Dang and Sun in 2007 [7, 8]. The generalized relational grade was discussed on generalized interval by Wang and Guo (2005) [9]. Although each kind of relation calculation method is improved, they also have defects. The references [4–6, 10] proposed the defects of grey relational grade and corresponding improvement but the improved relational grade has new defects. This paper proposed a new Hybrid information processing method based on Integral envelope and Hilbert transform theory with PCA analysis theory together to extract the features of individual devices which will establish a good basis for the development of Internet of things.

2 Basic Theory

2.1 Integral Envelope Theory

The signal is integrated at any starting time with a certain length of time, the waveform of the integral value is called the signal integral envelope [11].

If the signal in time domain is:

$$s(t) = A(t)\exp(j(2\pi f_0 t + \varphi(t))) \tag{1}$$

where $A(t)$ is the instantaneous amplitude of the signal, f_0 is the signal carrier frequency, $\varphi(t)$ is the phase modulation function. If $\varphi(t)$ only take 0 and π, the signal is

BPSK signal. If $\varphi(t)$ does not change, the single is carrier frequency signal. Therefore, the If the signal as shown in formula (1) is single carrier frequency signal, the integral envelope can be written as

$$\begin{aligned} G^2 &= \left| \frac{1}{\Delta T} \int_{t'}^{t'+\Delta T} s(t)dt \right|^2 = \frac{1}{\Delta T^2} \left| \int_{t'}^{t'+\Delta T} \exp(j(2\pi f_0 t + \varphi(t)))dt \right|^2 \\ &= \frac{1}{(2\pi f \Delta T)^2}(2 - 2\cos(2\pi f_0 \Delta T)) = \frac{(\sin(\pi f_0 \Delta T))^2}{(\pi f \Delta T)^2} = (\sin c(\pi f_0 \Delta T))^2 \end{aligned} \tag{2}$$

where $\sin c(x) = \frac{\sin(\pi x)}{\pi x}$.

The signal integral envelope G is related to the integral time interval ΔT and the signal carrier frequency f_0, as shown in formula (2). For a fixed frequency signal, when the interval of integration time is fixed, the integral envelope of the signal is a constant value. The G conforms to the rule of $|\sin c(x)|$ function when the signal carrier frequency changes.

If the LFM signal is

$$s(t) = A\exp(j(2\pi f_0 t + kt^2/2)) \tag{3}$$

where k is frequency modulation rate. When the interval of integration time is small, the integral envelope process of the LFM signal is equivalent to processing the single carrier frequency signal, and the integral carrier frequency is changed in each integral transformation. So, the waveform of the integral conforms to the rule of $|\sin c(x)|$ function in time domain. When $f_0 \Delta T$ is an integer, the integral envelope of the signal is zero. It is important.

2.2 PCA Analysis Theory

Principal component analysis (PCA) can be described as a regression-type optimization problem. Not only in the data processing, but also in dimensionality reduction, PCA has a wide range of applications [12].

PCA makes the derived variables capture maximal variance over seeking the linear combinations of the original variables. For computing the PCA, we can by obtaining the data matrix's singular value decomposition (SVD). In detail, we have an n-by-p matrix **X**, here, n is the number of samples and p is the number of variables. In order to generalization, we assume that the means of every column are all 0. The singular value decomposition of **X** can be written [13].

$$\mathbf{X} = \mathbf{U}\mathbf{D}\mathbf{V}^T \tag{4}$$

where **D** is the singular values of **X**, an rectangular diagonal matrix which size is n-by-p; **U** is the left singular matrix of **X**, an n-by-n matrix, whose columns are orthogonal unit vectors with the length of n; and **V** is the right singular matrix of **X,** a p-by-p matrix, the columns of which are also orthogonal unit vectors.

We use the singular value decomposition the score matrix $\mathbf{Y}^T$ as

$$\mathbf{Y}^T = \mathbf{X}^T\mathbf{U} = \mathbf{V}\mathbf{D}^T\mathbf{U}^T\mathbf{U} = \mathbf{V}\mathbf{D}^T \tag{5}$$

It is easily seen that the left singular vectors of $\mathbf{X}^T$ is $\mathbf{V}$, so $\mathbf{Y}^T$ is given by the left singular matrix $\mathbf{V}$ multiplied by the transpose of singular values matrix. The polar decomposition of $\mathbf{Y}^T$ is also expressed as this.

It is no need to form the matrix $\mathbf{X}^T\mathbf{X}$ to calculate the SVD of $\mathbf{X}^T$ in efficient algorithms exist, so calculating a principal component analysis from a data matrix is the standard way for computing the SVD, unless we require a handful of components.

At last, PCA makes the greatest variance according to transforming the data to a new coordinate system, some projection of data which are called the first principle component comes to lie on the first coordinate, and on the second coordinate, the second greatest variance appears, and so on.

Finally, multiply the first l largest singular values and corresponding singular vectors, we can get a truncated $n \times l$ score matrix $\mathbf{Y}_l$ of $\mathbf{Y}$:

$$\mathbf{Y}_l = \mathbf{U}_l^T\mathbf{X} = \mathbf{D}_l\mathbf{V}^T \tag{6}$$

After construction, the size of the transformed data matrices is n-by-l.

3 Simulation and Analysis

The feature extraction algorithm proposed in this paper is used to extract the characteristics of wireless devices, and then gray relation theory was used to classify the signals. The steps of the algorithm are as follows:

Firstly, 50 sets of signals of 10 groups of devices are calculated by integral envelope algorithm respectively, and then instantaneous amplitude features can be obtained as data sets.

Second, the dimensions are decreased based on PCA analysis method. According to the contribution rate of the main components, the corresponding number of principal components are selected as the features to be identified.

At last, gray relation algorithm was used to extract the principal component characteristics of 50 sets of 10 devices in order to classify and identify them. While, 200 sets of data were randomly selected to train and 300 sets of data were selected to test, and the recognition rate was calculated.

According to the above procedures, take two kinds of fingerprint features as example, the algorithm simulation results are shown in Fig. 1:

From the simulation results we can see that, the contribution rate of the first principal component is 89.85%, the contribution rate of the second principal component is 5.69%, so the total contribution rate is 95.54%. According to the definition of PCA analysis, when the contribution rate of principal component is between 85% and 95%, the obtained principal features can approximately represent all the characteristics of the original signal, in order to realize the accurate classification of the signal. Therefore, Fig. 1 shows the two-dimensional principal components of different

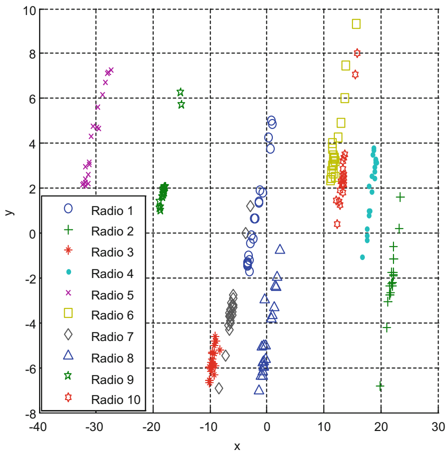

Fig. 1. Two-dimensional principal components of 10 wireless devices' subtle features based on integral envelope algorithm under no-noisy environment

wireless devices' features, where the abscissa x represents the first principal component value and the ordinate y represents the second principal component value. It can be seen that there is a good aggregation degree and interclass separation degree.

And then gray relation algorithm was used to identify the obtained features, and the recognition result is 96.78%, which basically realize the accurately recognize of signals. If the recognition effect needs to be improved, take one more principal components to form three-dimensional features, whose contribution rate is 3.94%. Plus the first two principal components together, the total contribution rate is 99.48%. They can more accurately describe all the characteristics of the original signals. The same method was used to identify the signals, and the recognition rate is 98.33%, comparing with using two principal components, the recognition result is improved.

Similarly, change the simulation condition to 20 dB SNR environment, then using the algorithm proposed in this paper to extract and classify the characteristics of the signals. Simulation results are shown in Fig. 2. Where the abscissa x represents the first principal component value and the ordinate y represents the second principal component value. It can be seen from the simulation results that, some characteristics of wireless devices have a certain degree of overlap, it is difficult to achieve accurately classification.

According to the results of PCA analysis, the contribution rate of the first principal component reached 88.26%, the contribution rate of the second principal component reached 5.39%, and the contribution rate of the third principal component reached 3.95%. Due to the noise added to the signals, the contribution rate of each reduced dimensional principal component are relatively dispersion. So the total contribution rate of the first two principal component features used in Fig. 2 is only 93.65%, and the recognition result obtained by gray relation classifier is 82.67%. In order to improve the recognition result, another principal component is added together, and the total contribution rate is 97.6%, then the recognition result can be calculated as 88.45%.

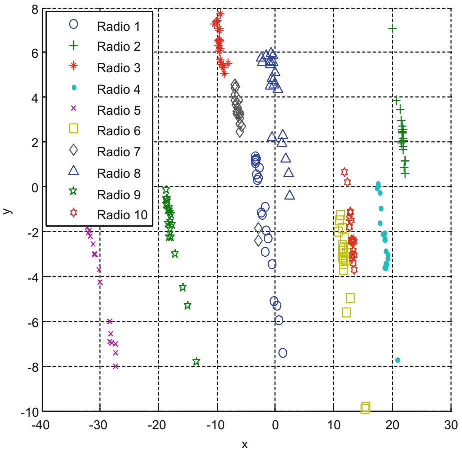

Fig. 2. Two-dimensional principal components of 10 wireless devices' subtle features based on integral envelope algorithm under the SNR of 20 dB environment

4 Conclusion

In this paper we have addressed the fundamental issue of wireless device identification for the Internet of things. We designed, implemented and evaluated the integral envelope theory based PCA analysis algorithm, a technique that can identify the subtle features of wireless devices. Unlike the previous techniques, this algorithm can accurately identify the fingerprints features of different wireless devices. Simulation results show that, the proposed algorithm can recognize the subtle features of signals between 500 numbers of wireless devices with the accuracy in excess of 99%. The approach itself is general and will work with any wireless devices' recognition.

Our evaluation of the strengths and weaknesses of the proposed Hybrid information processing algorithm suggests that it could be especially useful for wireless device's subtle features' identification. And it also can be used in some other related fields such as image processing, fault diagnosis and so on.

Acknowledgements. The research of the paper is supported by the National Natural Science Foundation of China (No. 61603239) and (No. 61601281).

References

1. Liu, S., Forrest, J., Yang, Y.: A brief introduction to grey systems theory. Grey Syst.: Theory Appl. **2**(2), 89–104 (2012)
2. Zhou, X.: The study on grey relational degree and its application. Jilin University Press, Changchun (2007)
3. Deng, J.L.: Grey System. China Ocean Press, Beijing (1988)
4. Wang, Q.: The grey relational analysis of B-mode. J. Huazhong Univ. Sci. Tenchnol. **6**, 77–82 (1989)

5. Mei, Z.: The concept and computation method of grey absolute correlation degree. Syst. Eng. **5**, 43–44+72 (1992)
6. Tang, W.: The concept and the computation method of T's correlation degree. Appl. Stat. Manag. **1**, 34–37+33 (1995)
7. Dang, Y.: The research of grey slope relational grade. Syst. Sci. Compr. Stud. Agric. **10** (Supplement), 331–337 (1994)
8. Sun, Y., Dang, Y.: The improved model of grey slope relational grade. Stat. Decis. **15**, 12–13 (2007)
9. Wang, Q., Guo, L.: Generalized relational analysis method. J. Huazhong Univ. Sci. Technol. (Nat. Sci. Edn.) **8**, 97–99 (2005)
10. Zhang, S.: Comparison between computation models of grey interconnect degree and analysis on their shourages. Syst. Eng. **3**, 45–49 (1996)
11. Liu, X., Si, X., Lu, M., Cai, Z.: Quick estimation to parameters of LPI radar-signals based on integral-envelope. Syst. Eng. Electron. **10**, 2031–2035 (2010)
12. Peng, K., Zhang, M., Li, Q., et al.: Fiber optic perimeter detection based on principal component analysis. In: 2016 15th International Conference on Optical Communications and Networks (ICOCN), pp. 1–3. IEEE (2016)
13. Ying, Y., Cao, Y., Li, S., Li, J., Guo, J.: Study on gas turbine engine fault diagnostic approach with a hybrid of gray relation theory and gas-path analysis. Adv. Mech. Eng. 8(1) (2016). https://doi.org/10.1177/1687814015627769