BULLETIN OF THE POLISH ACADEMY OF SCIENCES
TECHNICAL SCIENCES, Vol. 69(2), 2021, Article number: e136741
DOI: 10.24425/bpasts.2021.136741

Accurate identification on individual similar communication emitters by using HVG-NTE feature

Ke LI [1,2,3], Wei GE [1,2], Xiaoya YANG [1,2], and Zhengrong XU [1*]

[1] School of Information and Computer, Anhui Agricultural University, Hefei, Anhui, 230036, China
[2] Anhui Provincial Engineering Laboratory for Beidou Precision Agriculture Information, Anhui Agricultural University, Hefei, Anhui, 230036, China
[3] Key Laboratory of Specialty Fiber Optics and Optical Access Networks, Shanghai University, Shanghai, 200072, China

Abstract. Individual identification of similar communication emitters in the complex electromagnetic environment has great research value and significance in both military and civilian fields. In this paper, a feature extraction method called HVG-NTE is proposed based on the idea of system nonlinearity. The shape of the degree distribution, based on the extraction of HVG degree distribution, is quantified with NTE to improve the anti-noise performance. Then XGBoost is used to build a classifier for communication emitter identification. Our method achieves better recognition performance than the state-of-the-art technology of the transient signal data set of radio stations with the same plant, batch, and model, and is suitable for a small sample size.

Key words: communication emitter; identification; feature extraction; HVG; NTE.

1. Introduction

Individual identification of similar communication emitters refers to recognizing particular individuals of communication emitters of the same type [1]. In the military field, using prior information to identify from which communication emitter of the enemy the signal comes can provide an important basis for accurately predicting the enemy's strategic and tactical intentions [2, 3]. Meanwhile, in the civil field, this technology also has important application prospects in wireless communication network security [4], cognitive radio [5], and mechanical fault diagnosis [6].

When the transmitter is turned on or off or the communication mode is switched on, the transmitted signal will undergo a transient process, called the transient signal [7]. The transient signal is the impulse response generated by each component, which contains rich fingerprint information [8, 9]. Previous research has mainly extracted transient signal features from instantaneous amplitude, phase, and frequency [10, 11], such as wavelet analysis [12] and fractal analysis [13].

Dubendorfer et al. were the first to point out that individual identification of wireless transmitter emitters can be performed using time-frequency analysis methods [14]. Reising used discrete Gabor transform analysis to extract the statistical features of the normalized amplitude coefficient sequences in the WiMax transient signal to identify individual emitters. In the case of SNR > -3 db, the correct recognition rate exceeded 90%, but the experiment was carried out in an office environment, which could not meet the actual requirements [15]. Yuan used the Hilbert-Huang Transform [16] to obtain the time-frequency energy distribution of transient communication signals, but the process of obtaining the eigenmode function from HHT was inefficient, and there was a serious boundary effect [17].

Carroll was the first to point out that since the communication transmitter contains a semiconductor-based power amplifier, the transmitter system can be considered as a nonlinear system, and the RF fingerprint of the communication emitter can be extracted from the perspective of nonlinear dynamics [18]. Sun proposed a method based on singular spectrum analysis to analyze the transient signals of mobile phones, which is effective even in the case of low SNR [19]. But the experiment used different types of emitters and did not use the same type of communication emitters. Jia takes different transmitters as chaotic systems and extracts one-dimensional component natural measures as subtle features, which can solve the problem of small data volume and low sampling rate [20]. But the SNR of all data samples is about 18db, which is not applicable under the condition of low SNR.

In this paper, we propose an individual identification method of communication emitters based on system nonlinearity. The method includes two stages: feature extraction and identification recognition. In the feature extraction stage, the improved phase method is used to detect the transient starting point, then the adaptive threshold method based on HHT energy trajectory is used to detect the transient endpoint, and finally, the fingerprint features of communication emitters are obtained by the HVG-NTES feature extraction method to improve the anti-noise performance. In the identification and recognition stage, a classifier for individual identification of communication emitters is constructed based on XGBoost. Our method achieves

*e-mail: xzr_525@ahau.edu.cn

Manuscript submitted 2020-08-19, revised 2020-12-08, initially accepted for publication 2021-01-16, published in April 2021

better recognition performance than the state-of-the-art technology of the transient signal data set of radio stations with the same plant, batch, and model.

2. The proposed method

Visibility graph (VG) [21] is a tool for analyzing nonlinear time series and is a technique for mapping time series into graphs [22]. Horizontal visibility graph (HVG) [23] is a simplification of VG. Its method is simple, and the calculation speed is faster than VG, so the horizontal visibility algorithm is more practical in the actual communication confrontation environment. The definition is as follows [24]:

Let $\{y_t\}$ $(t = 1,2,\ldots,T)$ be the time series of T data and assign each data of the sequence to a node in the HVG. In the time series connecting y_i and y_j, if a horizontal line can be drawn in the time series that does not intersect with any intermediate data height, the two nodes (t_i, y_i) and (t_j, y_j) in the graph are connected. Therefore, if the following geometric criteria are met within the time series, then (t_i, y_i) and (t_j, y_j) are two connected nodes:

$$y_i,\ y_j > y_l\,. \tag{1}$$

Figure 1 shows an intuitive example of HVG, where the upper half draws a time series, and the lower half represents the graph generated by the horizontal visibility algorithm. Each data in the sequence corresponds to a node in the graph. If the height of the data corresponding to the two nodes is greater than the height of all the data between them, the two nodes are connected.

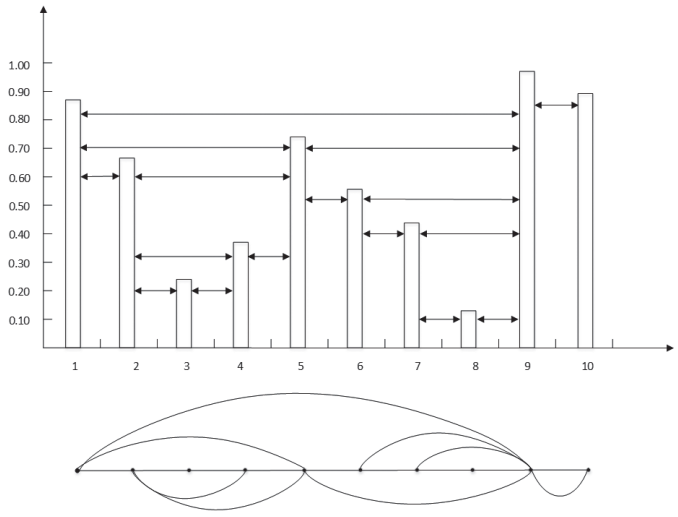

Fig. 1. Illustration of the HVG

We can extract the information on the graph structure after finding the graphs of the time series. Here we distinguish them by the degree distribution of the chaotic process [25]. For a given graph G with T nodes, the degree of node k is the number of edges associated with the node. The degree distribution $P(k)$ is the proportion of nodes with degree k.

Because of the short duration, short data length, and non-stationary characteristics of the transient signal of communication emitter, it is generally considered as a non-extensive system. In this paper, the non-extensive entropy of the degree distribution, Tsallis entropy (TE) [26], is used to quantify the shape of the degree distribution as a subtle feature of the communication emitter. For the discrete distribution $P = \{p_i, i = 1,2,\ldots,T\}$, TE is defined as follows:

$$S_q(P) = -\sum_{j=1}^{T} p_j^q \ln_q p_j \quad (q \neq 1), \tag{2}$$

where: $\ln_q p_j = \dfrac{p_j^{1-q} - 1}{1-q}$; the parameter q is a non-negative real number. The normalized TE is obtained as follows:

$$H_q(P) = \frac{S_q(P) - S_{q\_\min}(P)}{S_{q\_\max}(P) - S_{q\_\min}(P)}\,. \tag{3}$$

The NTE calculated from the HVG degree distribution is expressed by horizontal visibility graph-normalized Tsallis entropy (HVG-NTE). In this paper, the detected transient signal is regarded as a nonlinear time series. The first three are selected from the intrinsic mode functions (IMFs) obtained by empirical mode decomposition of the transient signals [27], which contain more nonlinear complexity components, and then the HVG-NTE is extracted. For each transient signal data, the following feature vectors can be obtained:

$$\mathbf{v} = [v_1, v_2, v_3]\,, \tag{4}$$

where: v_1, v_2, v_3 are the HVG-NTES obtained from the first three IMFs, respectively.

3. Experimental results and discussion

3.1. Signal acquisition. Figure 2 shows the schematic diagram of communication emitter signal receiving and collecting the subsystem. In this experiment, 4 short-wave handheld radio stations with the model of QYT CB were used as the identifi-

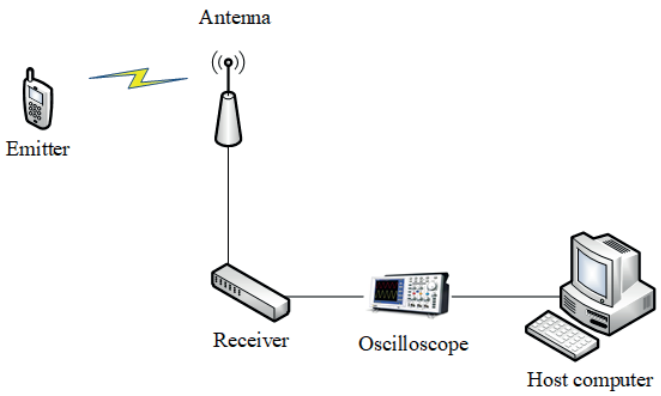

Fig. 2. Schematic diagram of communication emitter signal receiving and collecting subsystem

cation objects, and the communication frequency was 50 MHz. The signal acquisition equipment is the Egret RM200 receiver, the sampling frequency band is set at 49.75–50.25 MHz, the sampling rate is 7 KHz, and the radio frequency signal is converted to an intermediate frequency of 70 KHz. The sampled data is stored in the host computer in hexadecimal data format. The host computer model is Lenovo ThinkCentre M910T, equipped with an i7-7700 processor and 16G memory. It should be noted that the experimental data was obtained under the condition of the outfield, considering the signal power and noise level interference, both the radio and the receiver use antennas to send and receive signals, rather than an ideal laboratory environment.

3.2. Transient detection. The typical transmission signal waveform is shown in Fig. 3, including noise, transient signal, and steady-state signal.

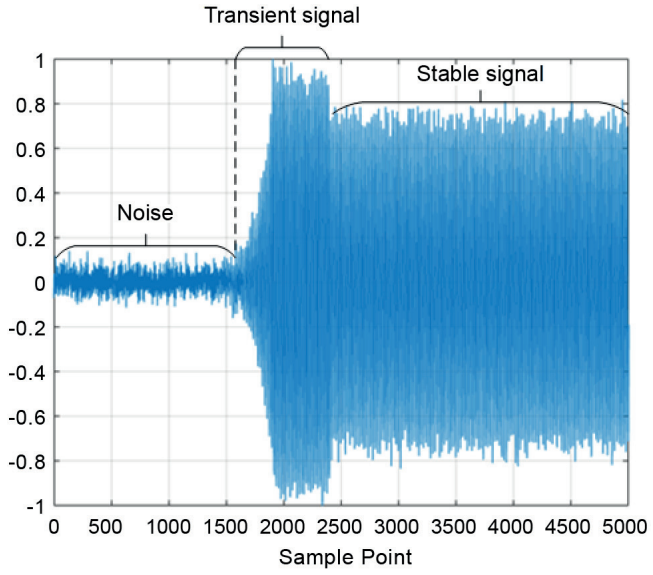

Fig. 3. The waveform of the transmission signal

The duration of the transient signal is very short. If there is an error in the detection of the starting and endpoints of the transient signal, it will affect the accuracy of identification recognition. We used an improved phase-based method to detect the transient starting point [28]. The difference between each successive phase variance is calculated twice to obtain the fractal trajectory (FT). At the beginning of the transient, the fractal trajectory should be zero, that is, the transient starting point is detected as the 90th frame, corresponding to the 1800th point in the original signal, as shown in Fig. 4.

The adaptive threshold method based on energy trajectory detects the transient endpoint [29]. It can be seen from Fig. 5 that after the endpoint P_E, a positive slope value must exist reasonably as the energy is stable and almost equal. After the first point with a positive slope, the point where the energy trajectory is greater than the maximum value plus the standard deviation is the endpoint.

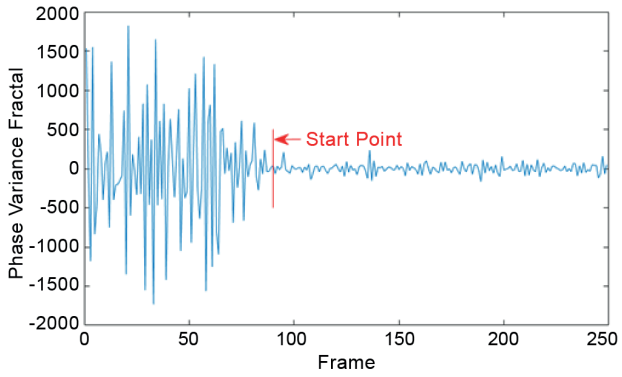

Fig. 4. Fractal trajectory

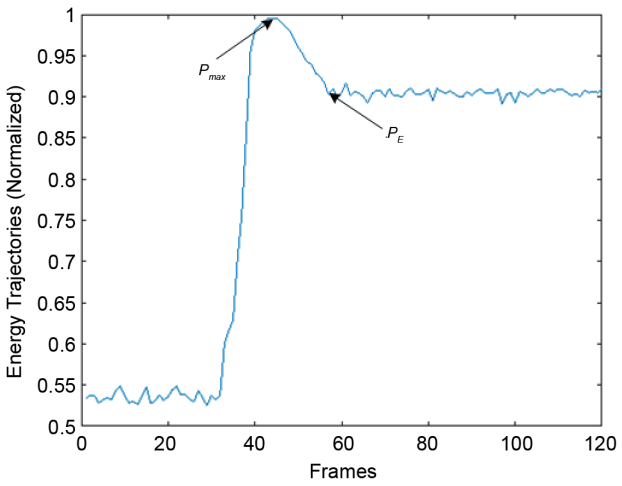

Fig. 5. Energy trajectories based on HHT

3.3. Robustness and anti-noise performance of NTE. Here, the performance of the proposed NTE is mainly proved by the Rossler map and compared with Shannon entropy (SE). The Rossler map [30] is given by:

$$dx/dt = -z - y, \tag{5}$$

$$dy/dt = x + 0.15y, \tag{6}$$

$$dz/dt = 0.2 + K(zx - 5). \tag{7}$$

Four different dynamic systems were selected respectively, with parameters $K = 0.6$, 0.7, 0.8, and 0.9, and the nonlinear complexity of the system increased continuously. The time series is obtained by integrating with increments of 0.01, and the range of the parameter q is set to 0.1 to 1.5, and T is 1000. The results are shown in Fig. 6.

As can be seen from Fig. 6, when there is no noise, SE can improve the recognition performance when increasing the q value appropriately, but as q continues to increase, the discrimination of the dynamic system will decrease. NTE can clearly distinguish these four dynamic systems under different non-extensive coefficients q.

In order to evaluate the anti-noise performance of the two algorithms, the noise is superimposed on each time series by adding gaussian random variables with different noise levels.

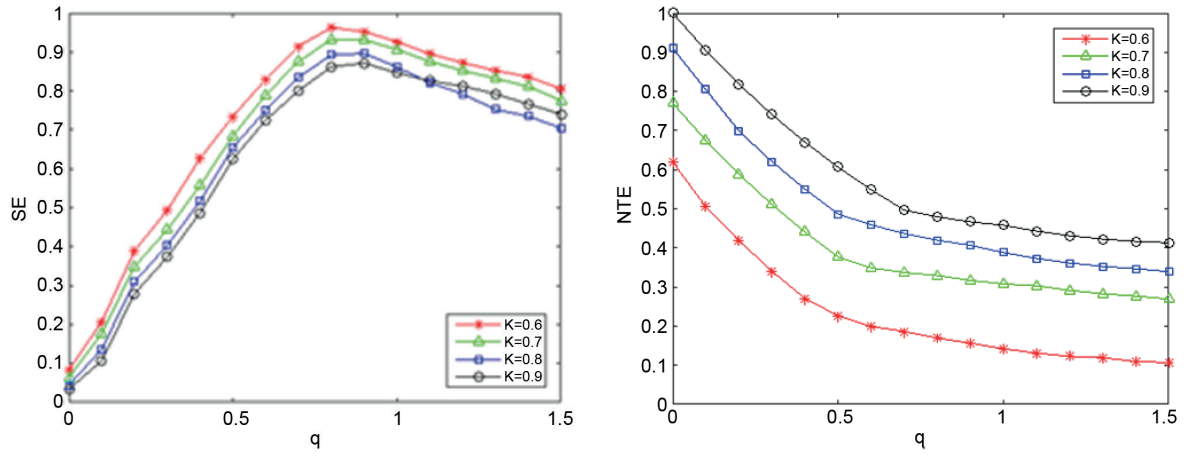

Fig. 6. Results of NTE and SE of four dynamic systems changes with q

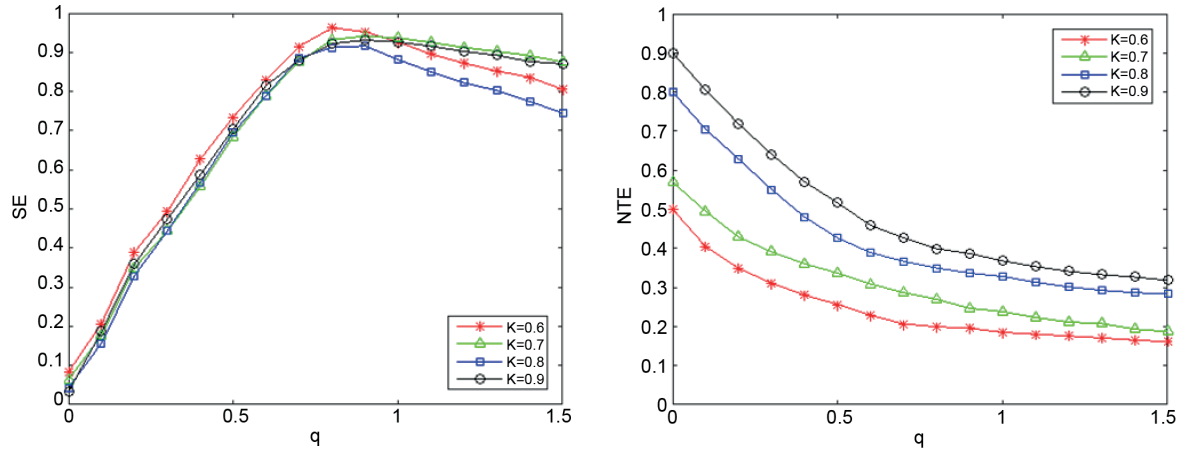

Fig. 7. Results of NTE and SE of four dynamic systems changes with q when the noise level is 0.2

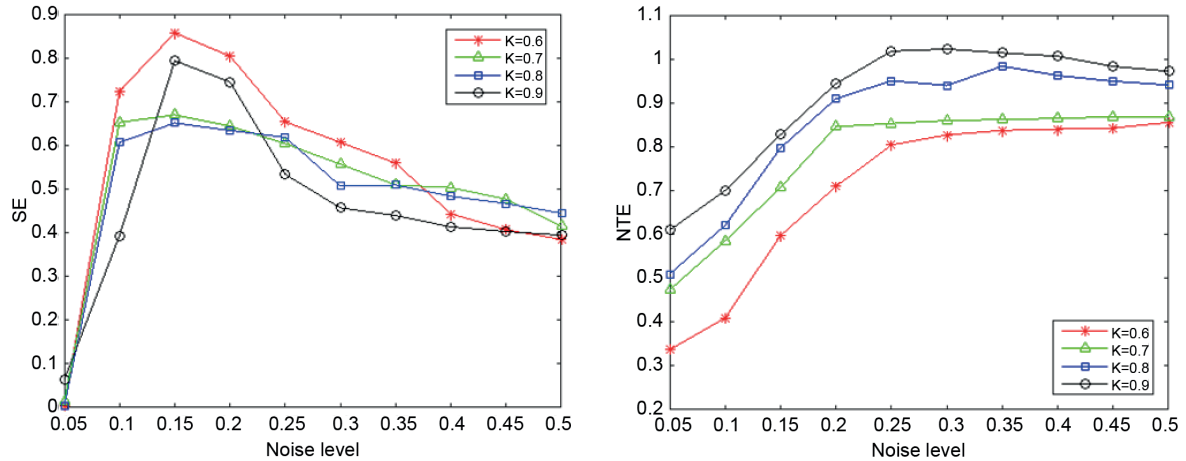

Fig. 8. Results of NTE and SE of four dynamic systems changes with noise level when q is 0.5

Figure 7 shows the entropy of the Rossler map of noise pollution (noise level is 0.2), and it can be seen that NTE has a good resolution ability in the range of 0.1∼1.5. However, no matter how much q is taken, SE can hardly distinguish these four different systems.

Figure 8 shows the performance of the two entropy statistics to distinguish the Rossler system at different noise levels ($q = 0.5$). It can be seen that NTE can correctly identify the sequence generated by the Rossler system with different parameters K until the noise level increases to 0.5. However, when the

noise level is greater than 0.1, SE is unable to distinguish these systems. Therefore, NTE is an effective nonlinear complexity extraction algorithm with robustness and noise immunity.

4. Fingerprint feature extraction

We extracted the HVG-NTES features from the transient signal of four handheld radio stations. In order to indicate the difference in fingerprint features between communication emitters, we visualized the feature distribution.

As can be seen from Fig. 9, the feature sets of the four stations are significantly different, which indicates that the four stations can be distinguished by the method based on HVG-NTE.

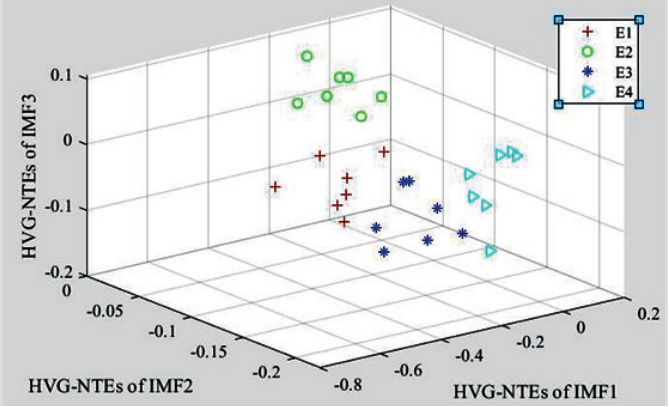

Fig. 9. HVG-NTE values for 4 handheld stations

5. Emitter identification

XGBoost [31] generates a second-order Taylor expansion for the loss function, and obtains the optimal solution for the regular term outside the loss function, making full use of the parallel computing advantages of the multi-core CPU to improve the accuracy and speed. Therefore, we use XGBoost to establish an individual recognition classifier for communication emitters. We use the grid search method with the optimal parameters n_estimators = 300, learning_rate = 0.01, max_depth = 20. 100 samples were collected from each handheld station, and the samples from each station were divided in a 1:1 scale, that is, 50 samples were taken as the training set, and the remaining 50 samples were taken as the test set. The experiment was performed 20 times independently and the average result was calculated.

To illustrate the effectiveness of the algorithm, we compared our method with NPE [32], SSA [33], and Natural Measure [20].

Figure 10 shows the average recognition rates of the four methods for handheld radio stations under the SNR of $5 \sim 25$ dB. The recognition rate increases as the SNR increases. However, when the SNR is lower than 10 dB, the average recognition rates of NPE, SSA, and Natural Measure are all lower than 85.2%. Even if the SNR reaches 25 dB, the average recognition rate does not exceed 94%.

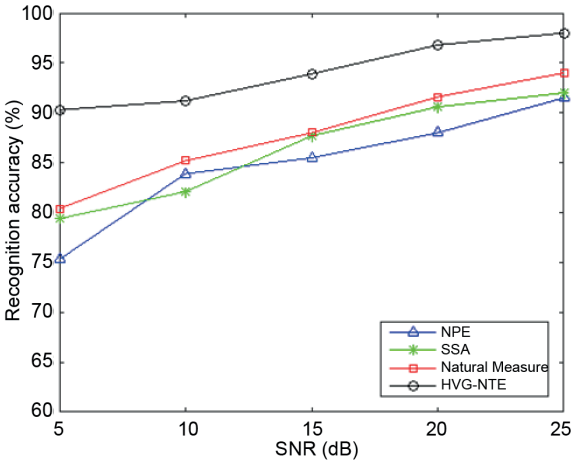

Fig. 10. Comparison of average recognition rate between HVG-NTE-based method and NPE-based method

The results of normalized permutation entropy (NPE) will be affected by the embedding dimension. A reasonable dimension must be set, while a small dimension will lower the recognition accuracy [32]. Singular spectrum analysis (SSA) can filter out noise well under high SNR, but as the SNR decreases, its denoising ability also decreases [33]. We need to use emitters of different brands when using SSA. The SNR of natural measure is fixed at 18dB, and its identification objects are emitters of four different manufacturers [20].

However, our method is more efficient and robust than the other three methods. The correct recognition rate can still reach 90% when the SNR is 5 dB, and the recognition rate reaches 98% when the SNR is 25 dB, which achieves fairly good performance and reduces the negative impact of noise. The experimental results show that our method performs better than the other three methods in the case of noise interference and a small number of samples and can effectively identify the communication emitters with the same plant, batch, and model.

6. Conclusions

This paper proposes a method for individual identification of communication emitters based on system nonlinearity. This method obtains IMFs of transient signals through empirical mode decomposition, extracts the degree distribution of the first three IMFs from the received signals based on the HVG algorithm, and uses NTE to quantify the shape of the degree distribution as the fingerprint features of the communication emitters. The anti-noise performance and robustness of the fingerprint features are proved by the Rossler map of different non-extensive coefficient q and noise level. Then a classifier based on XGBoost for individual identification of communication emitters is constructed [31]. In this paper, the method of identifying radios with the same plant, batch, and model is implemented under the condition of small sample size and low SNR, whose performance is better than the state-of-the-art method.

Acknowledgments. This work was supported by the Provincial-level Natural Science Research Key Project of Colleges and Universities in Anhui Province. (No. KJ2018A0145)

References

[1] J. Dudczyk, "Radar emission sources identification based on hierarchical agglomerative clustering for large data sets", *J. Sens.* 2016, 1879327 (2016).

[2] G. Manish, G. Hareesh, and M. Arvind, "Electronic Warfare: Issues and Challenges for Emitter Classification", *Def. Sci. J.* 201161(3), 228–234 (2011).

[3] J. Dudczyk and A. Kawalec, "Specific emitter identification based on graphical representation of the distribution of radar signal parameters", *Bull. Pol. Acad. Sci. Tech. Sci.* 63(2), 391–396 (2015).

[4] Q. Xu, R. Zheng, W. Saad, and Z. Han, "Device Fingerprinting in Wireless Networks: Challenges and Opportunities", *IEEE Commun. Surv. Tutor.* 18(1), 94–104 (2016).

[5] P.C. Adam and G.L. Dennis, "Identification of Wireless Devices of Users Who Actively Fake Their RF Fingerprints With Artificial Data Distortion", *IEEE Trans. Wirel. Commun.* 14(11), 5889–5899 (2015).

[6] N. Zhou, L. Luo, G. Sheng, and X. Jiang, "High Accuracy Insulation Fault Diagnosis Method of Power Equipment Based on Power Maximum Likelihood Estimation", *IEEE Trans. Power Deliv.* 34(4), 1291–1299 (2019).

[7] S. Guo, R.E. White, and M. Low, "A comparison study of radar emitter identification based on signal transients", *IEEE Radar Conference,* Oklahoma City, 2018, pp. 286–291.

[8] Q. Wu, C. Feres, D. Kuzmenko, D. Zhi, Z. Yu, and X. Liu, "Deep learning based RF fingerprinting for device identification and wireless security", *Electron. Lett.* 54(24), 1405–1407 (2018).

[9] A. Kawalec, R. Owczarek, and J. Dudczyk, "Karhunen-Loeve transformation in radar signal features processing", *International Conference on Microwaves,* Krakow, 2006.

[10] B. Danev and S. Capkun, "Transient-based identification of wireless sensor nodes", *Information Processing in Sensor Networks,* San Francisco, 2009, pp. 25–36.

[11] R.W. Klein, M.A. Temple, M.J. Mendenhall, and D.R. Reising, "Sensitivity Analysis of Burst Detection and RF Fingerprinting Classification Performance", *International Conference on Communications, Dresden,* 2009, pp. 641–645.

[12] C. Bertoncini, K. Rudd, B. Nousain, and M. Hinders, "Wavelet Fingerprinting of Radio-Frequency Identification (RFID) Tags", *I IEEE Trans. Ind. Electron.* 59(12), 4843–4850 (2012).

[13] Z. Shi, X. Lin, C. Zhao, and M. Shi, "Multifractal slope feature based wireless devices identification", *International Conference on Computer Science and Education,* Cambridge, 2015, pp. 590–595.

[14] C.K. Dubendorfer, B.W. Ramsey, and M.A. Temple, "ZigBee Device Verification for Securing Industrial Control and Building Automation Systems", *International Conference on Critical Infrastructure Protection* ,Washington DC, 2013, pp. 47–62.

[15] D.R. Reising and M.A. Temple, "WiMAX mobile subscriber verification using Gabor-based RF-DNA fingerprints", *International Conference on Communications,* Ottawa, 2012, pp. 1005–1010.

[16] Y. Li, Y. Zhao, L. Wu, and J. Zhang, "Specific emitter identification using geometric features of frequency drift curve", *Bull. Pol. Acad. Sci. Tech. Sci.* 66, 99–108 (2018).

[17] Y. Yuan, Z. Huang, H. Wu, and X. Wang, "Specific emitter identification based on Hilbert-Huang transform-based time-frequency-energy distribution features", *IET Commun.* 8(13), 2404–2412 (2014).

[18] T.L. Carroll, "A nonlinear dynamics method for signal identification", *Chaos Interdiscip. J. Nonlinear Sci.* 17(2), 023109 (2007).

[19] D. Sun, Y. Li, Y. Xu, and J. Hu, "A Novel Method for Specific Emitter Identification Based on Singular Spectrum Analysis", *Wireless Communications & Networking Conference,* San Francisco, 2017, pp. 1–6.

[20] Y. Jia, S. Zhu, and G. Lu, "Specific Emitter Identification Based on the Natural Measure", *Entropy* 19(3), 117 (2017).

[21] L. Lacasa, B. Luque, J. Luque, and J.C. Nuno, "The visibility graph: A new method for estimating the Hurst exponent of fractional Brownian motion", *Europhys. Lett.* 86(3), 30001–30005 (2009).

[22] M. Ahmadlou and H. Adeli, "Visibility graph similarity: A new measure of generalized synchronization in coupled dynamic systems", *Physica D* 241(4), 326–332 (2012).

[23] S. Zhu and L. Gan, "Specific emitter identification based on horizontal visibility graph", *IEEE International Conference Computer and Communications,* Chengdu, 2017, pp. 1328–1332.

[24] B. Luque, L. Lacasa, F. Ballesteros, and J. Luque, "Horizontal visibility graphs: Exact results for random time series", *Phys. Rev. E.* 80(4), 046103 (2009).

[25] W. Jiang, B. Wei, J. Zhan, C. Xie, and D. Zhou, "A visibility graph power averaging aggregation operator: A methodology based on network analysis", *Comput. Ind. Eng.* 101, 260–268 (2016).

[26] M. Wajs, P. Kurzynski, and D. Kaszlikowski, "Information-theoretic Bell inequalities based on Tsallis entropy", *Phys. Rev. A.* 91(1), 012114 (2015).

[27] J. Liang, Z. Huang, and Z. Li, "Method of Empirical Mode Decomposition in Specific Emitter Identification", *Wirel. Pers. Commun.* 96(2), 2447–2461, (2017).

[28] A.M. Ali, E. Uzundurukan, and A. Kara, "Improvements on transient signal detection for RF fingerprinting", *Signal Processing and Communications Applications Conference (SIU),* Antalya, 2017, pp. 1–4.

[29] Y. Yuan, Z. Huang, H. Wu, and X. Wang, "Specific emitter identification based on Hilbert-Huang transform-based time-frequency-energy distribution features", *IET Commun.* 8(13), 2404–2412 (2014).

[30] D.R. Kong and H.B. Xie, "Assessment of Time Series Complexity Using Improved Approximate Entropy", *Chin. Phys. Lett.* 28(9), 90502–90505 (2011).

[31] T. Chen and C. Guestrin, "XGBoost: A Scalable Tree Boosting System", *Knowledge Discovery and Data Mining,* San Francisco, 2016, pp. 785–794.

[32] G. Huang, Y. Yuan, X. Wang, and Z. Huang, "Specific Emitter Identification Based on Nonlinear Dynamical Characteristics", *Can. J. Electr. Comp. Eng.-Rev. Can. Genie Electr. Inform.* 39(1), 34–41 (2016).

[33] D. Sun, Y. Li, Y. Xu, and J. Hu, "A Novel Method for Specific Emitter Identification Based on Singular Spectrum Analysis", *Wireless Communications and Networking Conference,* San Francisco, 2017, pp. 1–6.