

A Novel Statistical Analysis Method for Radiation Source Classification

Haobo Geng*, Yaoyao Li, Weiping Tong, Youwei Meng, Houpu Xiao, Yicong Liu
School of Electronics and Information Engineering, Beihang University, Beijing, China
haobogeng@buaa.edu.cn

Abstract—With the rapid advancement of electronic information technology, the number and variety of unknown radiation sources have increased significantly. Some of these sources share common characteristics, which offers the potential to effectively address the challenge of identifying unknown radiation sources. However, research on the classification of radiation sources remains relatively limited. This paper proposes a big data analysis method that combines linear discriminant analysis (LDA) with a rough neighborhood set (NRS) for radiation source classification, and its effectiveness is validated on the RadioML 2018 dataset. The results indicate that, under certain constraints, all modulation types can be categorized into four distinct classes, laying a foundation for further research on cognitive interference signal cancellation.

Index Terms—radiation source, classification, LDA, NRS, electromagnetic environment

I. INTRODUCTION

With the rapid development of electronic information technology, the electromagnetic environment (EME) has become increasingly complex. The number and variety of unknown radiation sources continue to grow, posing challenges to traditional signal analysis. However, classifying these signals is essential for applications such as electronic warfare and communication security.

Although many studies focus on identifying the types of radiation sources using methods such as hypothesis testing [1] and feature engineering [2], modulation schemes often exhibit shared characteristics, suggesting the possibility of grouping them.

This paper combines linear discriminant analysis (LDA) with neighborhood rough set theory (NRS) to analyze radiation source data. Experiments on the RadioML2018.01A dataset [3] demonstrate effective clustering and classification of radiation source signals, providing a foundation for future research on radiation source recognition and interference suppression.

II. RADIATION SOURCE CLASSIFICATION METHODS

A. Selected Feature

In the field of electromagnetic compatibility (EMC), signal characteristics are typically described in the time, spatial, frequency, energy, and code domains. This paper focuses on target signal features, with spatial propagation effects already reflected in signal parameters such as multipath and fading, which are therefore not considered further. To characterize the temporal dynamics, frequency density, energy fluctuations, and modulation complexity of the electromagnetic environment, the following features are selected [4].

1) Temporal Dynamics

(a) Temporal Kurtosis: Measures peak sharpness in the time domain.

(b) Singular Spectrum Entropy: Reflects time-series complexity.

2) Frequency Density

(a) Bispectral Integration: Captures amplitude and phase in the frequency domain.

(b) Bandwidth Factor: Ratio of effective bandwidth to central frequency.

3) Energy Fluctuation

(a) Energy Concentration: Indicates energy distribution in the frequency domain.

(b) Fluctuation Index: Quantifies the variation in signal strength over time.

4) Modulation Complexity

(a) Fractal Dimension: Describes detail and complexity across scales.

(b) Lempel-Ziv Complexity: Quantifies the structural complexity of the signal sequence.

Clearly, the features exhibit redundancy, and the complexity of the high-dimensional feature space necessitates dimensionality reduction. To retain the physical significance of the features, a quantitative reduction approach is applied. NRS proves to be an effective tool for reducing continuous data features. This paper utilizes eight features from the RadioML 2018 dataset for subsequent analysis.

B. NRS Reduction Based on LDA

NRS attribute reduction relies on prior classification results for radiation source signals. This paper first uses LDA for initial dimensionality reduction and classification.

Also known as Fisher linear discriminant, LDA is a supervised technique that maximizes between-class variance while minimizing within-class variance to enhance class separability. Unlike unsupervised methods like principal component analysis, LDA utilizes class labels, making the reduced features better suited for classification tasks [5].

The dataset is $S = \{(\mathbf{x}_1, y_1), (\mathbf{x}_2, y_2), \dots, (\mathbf{x}_m, y_m)\}$, where $\mathbf{x}_i \in \mathbb{R}^n$ represents an arbitrary sample as an n -dimensional column vector, and $y_j \in \{C_1, C_2, \dots, C_K\}$ denotes the class label of the j -th sample, with a total of K classes.

Define μ_j as the mean vector of the j -th class samples.

$$\mu_j = \frac{1}{N_j} \sum_{\mathbf{x}_i \in S_j} \mathbf{x}_i, j = 1, 2, \dots, K \quad (1)$$

where N_j denotes the number of samples in the j -th class, and S_j represents the sample set of the j -th class.

The within-class scatter matrix measures the dispersion of samples within the same class, reflecting how samples are distributed around the mean.

$$\mathbf{S}_W = \sum_{j=1}^K \sum_{\mathbf{x}_i \in S_j} (\mathbf{x}_i - \boldsymbol{\mu}_j)^T (\mathbf{x}_i - \boldsymbol{\mu}_j) \quad (2)$$

The between-class scatter matrix measures the dispersion between samples of different classes, representing the distances between the centroids of different classes.

$$\mathbf{S}_B = \sum_{j=1}^K N_j (\boldsymbol{\mu}_j - \boldsymbol{\mu})^T (\boldsymbol{\mu}_j - \boldsymbol{\mu}) \quad (3)$$

where $\boldsymbol{\mu}$ is the overall mean vector of all samples.

LDA aims to find the optimal projection direction by maximizing the ratio of the between-class to within-class scatter matrices. The multi-class objective function is

$$\underbrace{\arg \max}_{\mathbf{W}} J(\mathbf{W}) = \frac{\prod_{diag} \mathbf{W}^T \mathbf{S}_B \mathbf{W}}{\prod_{diag} \mathbf{W}^T \mathbf{S}_W \mathbf{W}} \quad (4)$$

where $\arg \max$ represents finding a $\mathbf{W}$ that maximizes $J(\mathbf{W})$, $\prod_{diag} \mathbf{A}$ denotes the product of the diagonal elements of $\mathbf{A}$, and $\mathbf{W}$ is the projection matrix.

The maximum value of (4) is the largest eigenvalue of $\mathbf{S}_W^{-1} \mathbf{S}_B$. The matrix $\mathbf{W}$ consists of the eigenvectors corresponding to the largest d eigenvalues of $\mathbf{S}_W^{-1} \mathbf{S}_B$, given by

$$\mathbf{W} = \text{eig}(\mathbf{S}_W^{-1} \mathbf{S}_B) \quad (5)$$

where $\text{eig}()$ calculates eigenvalues and eigenvectors. For d -dimensional reduction, the top d eigenvectors are selected as the projection directions.

After dimensionality reduction, each sample is transformed as $\mathbf{z}_i = \mathbf{W}^T \mathbf{x}_i$. And the output dataset is $\mathbf{S}' = \{(\mathbf{z}_1, y_1), (\mathbf{z}_2, y_2), \dots, (\mathbf{z}_m, y_m)\}$. The reduced result retains the most relevant information for classification, providing reliable prior information for attribute reduction in neighborhood rough sets.

Next, we introduce NRS, an extension of traditional rough sets. Unlike precise partitioning, NRS uses "neighborhoods" to address data ambiguity, grouping similar samples [6]. The feature vectors and class labels from the previous section correspond to the conditional and decision attributes, respectively.

Let $NDS = (U, R, V, f)$ be an information system, where $U = \{z_1, z_2, \dots, z_m\}$ is the universe of objects, $R = C \cup D$ is the attribute set with condition attributes C and decision attributes D , and $V = \bigcup V_a$ represents the attribute value set, where V_a is the range of attribute a . The information function $f : U \times R \rightarrow V$ assigns values such that $\forall a \in R, z \in U, f(z, a) \in V_a$.

Specifically, for a neighborhood rough set decision system, given $z_i \in U$ and a subset of condition attributes $B \subseteq C$, the neighborhood of z_i under B is defined as

$$\delta_B(z_i) = \{z_j | z_j \in U, \Delta^B(z_i, z_j) \leq \delta\} \quad (6)$$

where Δ is defined as $\Delta(z_i, z_j) = \left[\sum_{k=1}^n (z_{ik} - z_{jk})^2 \right]^{\frac{1}{2}}$, where z_{ik} and z_{jk} represent the k -th attribute values of

objects z_i and z_j , respectively, and n is the total number of condition attributes. $\Delta(z_i, z_j)$ represents the Euclidean distance between z_i and z_j , while δ denotes the neighborhood radius.

Let Z be a subset of U . For $\forall B \subseteq C$, the lower approximation of Z in neighborhood space is defined as

$$\underline{\text{apr}}_B(Z) = \{z_i | \delta_B(z_i) \subseteq Z, z_i \in U\} \quad (7)$$

For the NRS decision system $NDS = (U, R, K)$, if D partitions U into K equivalence classes $(Z_1, Z_2, \dots, Z_K)$, the lower approximation of D with respect to B is defined as

$$\underline{\text{apr}}_B(D) = \bigcup_{i=1}^K \underline{\text{apr}}_B(Z_i) \quad (8)$$

The dependency degree of D on the conditional attribute subset B is defined as

$$\gamma_B(D) = \frac{|\underline{\text{apr}}_B(D)|}{|U|} \quad (9)$$

If $a \in (C - B)$, the significance of a relative to B is

$$\text{SIG}(a, B, D) = \gamma_{B \cup a}(D) - \gamma_B(D) \quad (10)$$

This achieves the most critical application of rough set theory: knowledge extraction and attribute reduction, thereby avoiding dimensional explosion. When B satisfies $\forall a \in B, \gamma_{B-a}(D) < \gamma_B(D)$ and $\gamma_B(D) = \gamma_C(D)$, B is called the minimal subset of attributes.

III. EXPERIMENTAL VERIFICATION

A. Public Dataset

This paper uses the RadioML 2018 dataset, a benchmark in radio signal processing, covering 24 modulation types with simulated channel effects. From this dataset, 1,000 samples with varying signal-to-noise ratios were randomly selected for each modulation type, totaling 24,000 samples for analysis.

B. Validation Process

To reduce external interference and improve data reliability, power normalization is applied to the complex signal $z(n)$ with N sampling points, minimizing the effect of signal strength on classification accuracy.

$$\bar{z}(n) = \frac{z(n)}{\sqrt{\frac{1}{N} \sum_{n=1}^N |z(n)|^2}} \quad (11)$$

Using Section II.A's definitions, eight features were computed to form the dataset $\mathbf{S}$. Modulation schemes of the same type but different orders share similar principles, leading to common post-processing methods across various domains. Thus, in LDA, $\mathbf{S}$ groups modulation schemes of the same type but different orders into a single initial category, as shown in Table I.

The $\mathbf{W}$ was computed using (5). To evaluate the impact of dimensionality d on accuracy, ten-fold cross-validation was performed, and the results are shown in Fig. 1.

From Fig. 1, it can be seen that the classification accuracy based on the original dataset is 79.3%. When $d = 3$, the accuracy approaches the ideal level. Therefore, we choose the 3-dimensional projection matrix, and the dimensionality-reduced results are visualized in Fig. 2. Based on this

TABLE I
INITIAL CLASSIFICATION

Modulation mode	Class
FM	1
GMSK	2
2ASK, 4ASK, 8ASK	3
16APSK, 32 APSK, 64 APSK, 128 APSK	4
BPSK, QPSK, OQPSK, 8PSK, 16PSK, 32PSK	5
16QAM, 32 QAM, 64 QAM, 128 QAM, 256 QAM	6
AM-DSB-SC, AM-DSB-WC, AM-SSB-SC, AM-SSB-WC	7

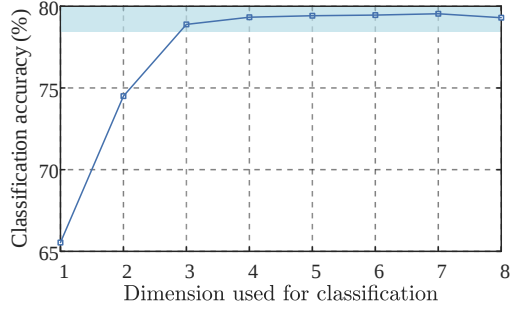


Fig. 1. Classification accuracy varies with the dimensionality.

clustering result, the class attributes are updated accordingly.

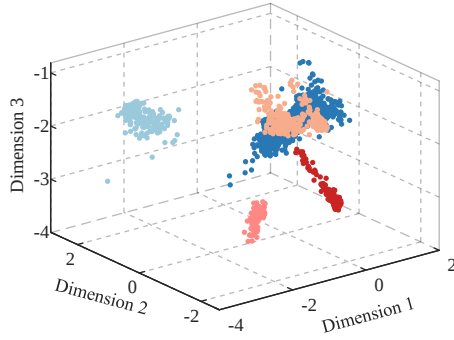


Fig. 2. Visualization of samples reduced to three dimensions using LDA.

Since LDA relies on 8 features, which is computationally intensive and redundant, we reduce the features using NRS based on LDA's classification results. The reduction depends on the neighborhood radius δ , so we conduct experiments under different δ values. The importance distribution of the 8 features is shown in Fig. 3.

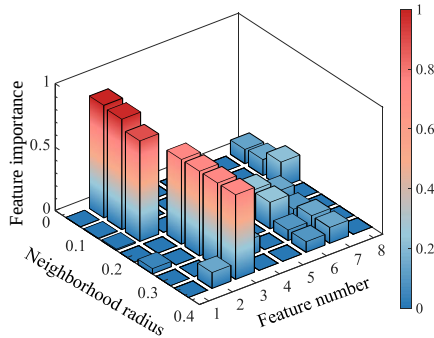


Fig. 3. The importance calculation results of the 8 features under various neighborhood radius conditions.

Based on the reduction results under various δ conditions, the second and third attributes, namely singular spectral

entropy and bispectral integral, are selected as the final classification features. The updated classification results are shown in Fig. 4, with detailed classifications in Table II.

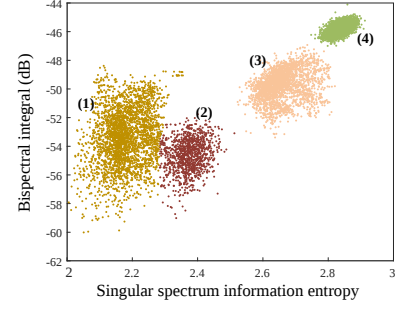


Fig. 4. The classification results of the samples based on the reduced features.

TABLE II
THE RESULT OF CLASSIFYING THE SAMPLE ACCORDING TO THE REDUCED FEATURES

Modulation mode	Class
AM-SSB-SC, AM-SSB-WC, FM	1
AM-DSB-SC, AM-DSB-WC	2
2ASK, 4ASK, 8ASK	3
16APSK, 32 APSK, 64 APSK, 128 APSK, BPSK, QPSK, OQPSK, 8PSK, 16PSK, 32PSK, 16QAM, 32 QAM, 64 QAM, 128 QAM, 256 QAM, GMSK	4

Thus, we have developed an understanding of signal types based on feature engineering and statistical analysis. Future research will build upon this work to further explore cognitive interference signal cancellation.

IV. CONCLUSION

This paper conducts an in-depth analysis of signal features using the RadioML 2018 dataset and proposes a simplified radiation source classification method that integrates linear discriminant analysis with neighborhood rough set theory. This study offers novel technical support for interference cancellation in electronic warfare and communication security, demonstrating both substantial theoretical significance and practical application potential.

ACKNOWLEDGMENT

This work was supported by the National Key Research and Development Program under Grant 2023YFB3306900.

REFERENCES

- [1] F. Wang and X. Wang, "Fast and Robust Modulation Classification via Kolmogorov-Smirnov Test," in *IEEE Transactions on Communications*, vol. 58, no. 8, pp. 2324-2332, Aug. 2010.
- [2] G. Z. Zhu, N. Zhou, Z. Wang, J. Liang and Z. Wang, "MSGNet: A Multi-Feature Lightweight Learning Network for Automatic Modulation Recognition," in *IEEE Communications Letters*, vol. 28, no. 11, pp. 2553-2557, Nov. 2024.
- [3] T. J. O'Shea, T. Roy and T. C. Clancy, "Over-the-Air Deep Learning Based Radio Signal Classification," in *IEEE Journal of Selected Topics in Signal Processing*, vol. 12, no. 1, pp. 168-179, Feb. 2018.
- [4] H. Geng, Y. Li, Y. Meng, H. Xiao, and S. Cai, "A novel method for predicting antenna port electromagnetic susceptibility under broadband interference via function-performance association," *IEEE Trans. Electromagn. Compat.*, Early Access, doi: 10.1109/TEM.2025.3583370.
- [5] M. A. Akbar et al., "An Empirical Study for PCA- and LDA-Based Feature Reduction for Gas Identification," in *IEEE Sensors Journal*, vol. 16, no. 14, pp. 5734-5746, Jul. 2016.
- [6] L. Sun, L. Wang, W. Ding, Y. Qian and J. Xu, "Feature Selection Using Fuzzy Neighborhood Entropy-Based Uncertainty Measures for Fuzzy Neighborhood Multigranulation Rough Sets," in *IEEE Transactions on Fuzzy Systems*, vol. 29, no. 1, pp. 19-33, Jan. 2021.