


Research Article

Robust Graph Factorization for Multivariate Electricity Consumption Series Clustering

Kaihong Zheng ¹, Honghao Liang,² Lukun Zeng,¹ Xiaowei Chen,² Sheng Li,¹ Hefang Jiang,² Qihang Gong,¹ Sijian Li,² Jingfeng Yang,³ and Shangli Zhou¹

¹Digital Grid Research Institute, China Southern Power Grid, Guangzhou 510663, China

²Shenzhen Power Supply Bureau Co., Ltd., Shenzhen 518001, China

³China Southern Power Grid Co., Ltd., Guangzhou 510663, China

Correspondence should be addressed to Kaihong Zheng; zheng_kh@163.com

Received 24 May 2021; Accepted 13 August 2021; Published 26 August 2021

Academic Editor: Andriette Bekker

Copyright © 2021 Kaihong Zheng 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.

Multivariate electricity consumption series clustering can reflect trends of power consumption changes in the past time period, which can provide reliable guidance for electricity production. However, there are some abnormal series in the past multivariate electricity consumption series data, while outliers will affect the discovery of electricity consumption trends in different time periods. To address this problem, we propose a robust graph factorization model for multivariate electricity consumption clustering (RGF-MEC), which performs graph factorization and outlier discovery simultaneously. RGF-MEC first obtains a similarity graph by calculating distance among multivariate electricity consumption series data and then performs robust matrix factorization on the similarity graph. Meanwhile, the similarity graph is decomposed into a class-related embedding and a spectral embedding, where the class-related embedding directly reveals the final clustering results. Experimental results on realistic multivariate time-series datasets and multivariate electricity consumption series datasets demonstrate effectiveness of the proposed RGF-MEC model.

1. Introduction

Recently, multivariate electricity consumption series clustering has been an important issue in machine learning fields [1–3]. In multivariate electricity consumption series, each instance consists of multiple time series from different sources which often contain information related to each other [4]. For example, multivariate electricity consumption series are composed of global-active-power series, global-reactive-power series, voltage series, and global-intensity series [5, 6]. Therefore, multivariate electricity consumption series clustering needs to analyze the relationship among these series [7, 8]. In other words, the objective of multivariate electricity consumption series clustering is to discover the relationship among multiple series and divide instances into groups.

There have been many works on time-series clustering, including univariate time-series clustering [9, 10] and multivariate time-series clustering [7, 11]. It is well known

that k-shape [12], k-DTW Barycenter Averaging (KDBA) [13], kAVG + ED [14], and k-Spectral Centroid (KSC) [15] are effective distance-based univariate time-series clustering algorithms. k-shape [12] utilized a normalized version of cross-correlation measure in order to consider shapes of time series. KDBA [13] was a global averaging method for dynamic time warping, where a new strategy was used to reduce the length of the resulting average sequence. kAVG + ED [14] made use of an efficient indexing method to locate 1-dimensional subsequences within a collection of sequences. KSC [15] can effectively find cluster centroids with a similarity measure, which applied an adaptive wavelet-based incremental approach to clustering. Because k-shape, KDBA, kAVG + ED, and KSC have achieved great success in univariate time-series clustering, researchers considered the relationship among multiple series and extended these methods for multivariate time-series clustering (i.e., m-kShape, m-kDBA, m-kAVG + ED, and m-KSC) [16].

In addition to extending existing univariate time series methods, researchers have proposed some unsupervised feature learning works which can learn informative features of multivariate time series. For example, two-dimensional singular value decomposition [15], variable-based principal component analysis (VPCA) [17], common principal component analysis (CPCA) [18–20], and deep encoder networks [21, 22] are used to reduce dimensionality of multivariate time series and learn informative features for clustering. He et al. [17] proposed a spatial weighted matrix distance-based fuzzy clustering (SWMDFC) model for multivariate time-series clustering. SWMDFC made use of VPCA to achieve dimensionality reduction and reduce computational consumption and then utilized spatial weighted matrix distance to compute distance among multivariate time-series data. Li [20] proposed a multivariate time-series clustering method based on CPCA, which used CPCA to construct projection coordinate features from multivariate time series and reconstructed multivariate time series from coordinate features. Franceschi et al. [21] proposed an unsupervised scalable representation learning model (USRL) for multivariate time series, which utilized a deep encoder network formed by dilated convolutions to generate informative features. However, noisy data and outliers are ubiquitous in realistic multivariate electricity consumption series data. Moreover, most existing works ignore the existence of outliers and noisy data, which can significantly affect clustering performance on multivariate electricity consumption series data.

In this paper, we develop a novel robust graph factorization framework (RGF-MEC) based on multivariate electricity consumption series, which is a joint learning framework for multivariate electricity consumption series clustering and discovery of outliers in multivariate electricity consumption series. After calculating the similarity matrix among multivariate electricity consumption series data, RGF-MEC performs robust orthogonal and nonnegative matrix factorization on this similarity graph structure. In addition, it can directly reveal clustering results without using k-means clustering that is sensitive to initialization. Therefore, the main contributions of this paper can be summarized as follows:

- (1) This paper proposes a novel multivariate electricity consumption series framework (RGF-MEC) to simultaneously factorize the graph structure and discover outliers in multivariate electricity consumption series data
- (2) RGF-MEC utilizes robust matrix factorization to simultaneously learn nonnegative class-related representations and orthogonal spectral representations
- (3) We perform experiments on realistic datasets to verify the proposed RGF-MEC model, and the results demonstrate that RGF-MEC is an effective multivariate electricity consumption series clustering framework

The remainder of this paper is organized as follows. Section 2 introduces related works, including spectral clustering, symmetric nonnegative matrix factorization (SymNMF), and orthogonal and nonnegative matrix

factorization (ONMF). Section 3 details the proposed RGF-MEC framework, including specific derivation process and complexity analysis. In Section 4, experimental results show the feasibility of the proposed RGF-MEC framework. Finally, some conclusions are given in Section 5.

2. Related Works

This section first explains the relationship between spectral clustering and symmetric nonnegative matrix factorization and then introduces the ONMF model.

Suppose \mathbf{X} denotes original instances, \mathbf{X}_i denotes the i th instance, \mathbf{L} denotes normalized Laplacian matrix of original instances, and $\mathbf{F} \in \mathbb{R}^{N \times K}$ denotes spectral embedding. Spectral clustering decomposes eigenvalues of normalized Laplacian matrix \mathbf{L} and constructs clusters by processing eigenvectors [23], which has the advantage of clustering on a sample space of arbitrary shape and converging to an optimal solution. Therefore, the objective function of spectral clustering can be defined as

$$\min_{\mathbf{F}^T \mathbf{D}^{-1} \mathbf{F} = \mathbf{I}} \text{Tr}(\mathbf{F}^T \mathbf{L} \mathbf{F}), \quad (1)$$

where $\mathbf{L} = \mathbf{I} - \mathbf{D}^{-1/2} \mathbf{S} \mathbf{D}^{-1/2}$, \mathbf{S} denotes similarity graph of original instances where an element S_{ij} is similarity between \mathbf{X}_i and \mathbf{X}_j , \mathbf{D} is the degree matrix with a diagonal element $D_i = \sum_j S_{ij}$, and clustering results are obtained by performing k-means clustering on the embedding $\mathbf{D}^{-1/2} \mathbf{F}$.

Kuang et al. [24] proposed SymNMF on the basis of nonnegative matrix factorization, which can simplify the high-dimensional graph structure to low-dimensional embedding while keeping the information as unchanged as possible. The objective function of SymNMF is defined as

$$\min_{\mathbf{F} \geq 0} \|\mathbf{G} - \mathbf{F} \mathbf{F}^T\|_F^2, \quad (2)$$

where \mathbf{G} denotes the symmetric high-dimensional data. Relaxing \mathbf{F} to be orthogonal, the above problem becomes the problem of spectral clustering.

Since the solution of (2) is difficult to compute, Han et al. [25] first transformed the problem of spectral clustering into the problem of symmetric nonnegative matrix factorization,

$$\begin{aligned} (1) &\Leftrightarrow \max_{\mathbf{F}^T \mathbf{F} = \mathbf{I}} \text{Tr}(\mathbf{F}^T \mathbf{G} \mathbf{F}) \\ &\Leftrightarrow \min_{\mathbf{F}^T \mathbf{F} = \mathbf{I}} -\text{Tr}(\mathbf{G}^T \mathbf{F} \mathbf{F}^T + \mathbf{F} \mathbf{F}^T \mathbf{G}) \\ &\Leftrightarrow \min_{\mathbf{F}^T \mathbf{F} = \mathbf{I}} \text{Tr}(\mathbf{G}^T \mathbf{G} - \mathbf{G}^T (\mathbf{F} \mathbf{F}^T) - (\mathbf{F} \mathbf{F}^T) \mathbf{G} + (\mathbf{F} \mathbf{F}^T) (\mathbf{F} \mathbf{F}^T)) \\ &\Leftrightarrow \min_{\mathbf{F}^T \mathbf{F} = \mathbf{I}} \|\mathbf{G} - \mathbf{F} \mathbf{F}^T\|_F^2, \end{aligned} \quad (3)$$

and then introduced an auxiliary nonnegative variate \mathbf{H} to approximate the orthogonal variate \mathbf{F} :

$$\min_{\mathbf{H} \geq 0, \mathbf{F}^T \mathbf{F} = \mathbf{I}} \|\mathbf{G} - \mathbf{H} \mathbf{F}^T\|_F^2, \quad (4)$$

where \mathbf{H} is a nonnegative embedding. With orthogonal and nonnegative constraints, the reconstructed graph of ONMF

naturally has clear structure about clusters. It can be seen that ONMF can simplify the calculation by orthogonalizing \mathbf{F} while ensuring clustering accuracy.

3. Proposed Frameworks

In this paper, a robust graph factorization framework (RGF-MEC) is proposed for multivariate electricity consumption clustering, in which graph factorization and outlier discovery are performed at the same time.

A multivariate electricity consumption series dataset $\{\mathbf{ME}^{(i)}\}_{i=1}^N$ is given, where $\mathbf{ME}^{(i)} \in \mathbb{R}^{M \times D}$, M denotes the number of sequences in $\mathbf{ME}^{(i)}$, and D denotes the length of each sequence in $\mathbf{ME}^{(i)}$. RGF-MEC must first calculate the similarity graph structure \mathbf{G}^{ME} of $\{\mathbf{ME}^{(i)}\}_{i=1}^N$. If $\mathbf{ME}^{(j)}$ belongs to k -nearest neighbors of $\mathbf{ME}^{(i)}$, an element G_{ij}^{ME} of \mathbf{G}^{ME} is equal to

$$G_{ij}^{\text{ME}} = \frac{d(\mathbf{ME}^{(i)}, \mathbf{ME}^{(k+1)}) - d(\mathbf{ME}^{(i)}, \mathbf{ME}^{(j)})}{\sum_{j'=1}^k d(\mathbf{ME}^{(i)}, \mathbf{ME}^{(k+1)}) - d(\mathbf{ME}^{(i)}, \mathbf{ME}^{(j')})} \quad (5)$$

$$\min_{\mathbf{H}^{\text{ME}} \geq 0, \mathbf{F}^T \mathbf{F} = \mathbf{I}} \sum_{n=1}^N 1_{RG_n \leq \varepsilon} \cdot \|\mathbf{G}_n^{\text{ME}} - \mathbf{H}_n^{\text{ME}} \mathbf{F}^T\|_F^2 + \frac{\lambda}{2} \sum_{i,j} \|\mathbf{H}_i^{\text{ME}} - \mathbf{H}_j^{\text{ME}}\|_F^2 G_{ij}, \quad (6)$$

where \mathbf{H}^{ME} is the nonnegative class-related embedding, \mathbf{F} is the orthogonal spectral embedding, λ is a regularization parameter, $RG_n = \|\mathbf{G}_n^{\text{ME}} - \mathbf{H}_n^{\text{ME}} \mathbf{F}^T\|_F^2$ denotes the reconstruction error of n th multivariate electricity consumption series, and ε is a parameter to filter outliers (where ε is determined according to the reconstruction error $\{\|\mathbf{G}_n^{\text{ME}} - \mathbf{H}_n^{\text{ME}} \mathbf{F}^T\|_F^2\}_{n=1}^N$ and a hypothetical ratio r of outliers, and we need to sort $\{\|\mathbf{G}_n^{\text{ME}} - \mathbf{H}_n^{\text{ME}} \mathbf{F}^T\|_F^2\}_{n=1}^N$ from large to small and assign reconstruction error ranked in r to the parameter ε). If $RG_n \leq \varepsilon$, $1_{RG_n \leq \varepsilon} = 1$, otherwise $1_{RG_n \leq \varepsilon} = 0$. It can be seen that the objective function of RGF-MEC consists

$$\min_{\mathbf{H}^{\text{ME}} \geq 0, \mathbf{F}^T \mathbf{F} = \mathbf{I}} \sum_{n=1}^N 1_{RG_n \leq \varepsilon} \cdot \|\mathbf{G}_n^{\text{ME}} - \mathbf{H}_n^{\text{ME}} \mathbf{F}^T\|_{2,1} + \frac{\lambda}{2} \sum_{i,j} \|\mathbf{H}_i^{\text{ME}} - \mathbf{H}_j^{\text{ME}}\|_F^2 G_{ij}. \quad (7)$$

It can be seen that the main difference between RGF-MEC $_{F_2}$ and RGF-MEC $_{\ell_{2,1}}$ is a norm of robust graph factorization term. We provide the learning procedure in the following two sections and then give a convergence analysis of RGF-MEC.

where $d(\mathbf{ME}^{(i)}, \mathbf{ME}^{(j)}) = \sum_{m=1}^M \|\mathbf{ME}_m^{(i)} - \mathbf{ME}_m^{(j)}\|_F^2$ and $\mathbf{ME}^{(k+1)}$ is the $k+1$ th nearest neighbor of $\mathbf{ME}^{(i)}$, otherwise an element G_{ij}^{ME} is equal to zero. It is known that there are many works that can reduce dimensionality of multivariate time series, such as CPCA [20], VPCA [17], and deep encoder networks [21]. RGF-MEC can reduce dimensionality of the data to obtain more effective features before calculating the graph structure. However, in order to verify robustness of the proposed model, this paper does not perform dimensionality reduction operation.

RGF-MEC takes outliers into consideration while performing matrix factorization on the similarity graph \mathbf{G}^{ME} . The objective function of RGF-MEC can be defined as

of two parts, the robust graph factorization term and the graph regularization term. In other words, \mathbf{H}^{ME} also satisfies the graph constraint of $\{\mathbf{ME}^{(i)}\}_{i=1}^N$.

In (6), RGF-MEC utilizes a squared F-norm to constrain robust graph factorization term, named RGF-MEC with squared F-norm (RGF-MEC $_{F_2}$). At present, researchers have proved that $\ell_{2,1}$ -norm nonnegative matrix factorization is effective and has a mathematical meaning [26–28]. Therefore, we also present RGF-MEC with $\ell_{2,1}$ -norm (RGF-MEC $_{\ell_{2,1}}$). The objective function of RGF-MEC $_{\ell_{2,1}}$ can be defined as

3.1. Learning Procedure of RGF-MEC $_{F_2}$. RGF-MEC $_{F_2}$ can make use of the coordinate descent method to solve problem (6), where two embeddings \mathbf{H}^{ME} and \mathbf{F} are updated in turn. In addition, a detailed learning procedure of RGF-MEC $_{F_2}$ is provided in Algorithm 1.

Input: a set of a multivariate electricity consumption series dataset $\{\mathbf{ME}^{(i)}\}_{i=1}^N$, the parameters λ, r .

Output: the class-related embedding \mathbf{H}^{ME} .

- 1: Calculate the similarity graph structure \mathbf{G}^{ME} of $\{\mathbf{ME}^{(i)}\}_{i=1}^N$ based on (5).
- 2: Initialize \mathbf{F} , \mathbf{H}^{ME} , and ε in turn.
- 3: **while** Not convergent **do** % \mathbf{H}^{ME} is updated based on (18) in RGF-MEC $_{\ell_{2,1}}$.
- 4: Fix \mathbf{F} and calculate \mathbf{H}^{ME} based on (10). % \mathbf{F} is updated based on (21) in RGF-MEC $_{\ell_{2,1}}$.
- 5: Fix \mathbf{H}^{ME} and calculate \mathbf{F} based on (13).
- 6: Calculate ε based on the parameter r and reconstruction errors $\{\|\mathbf{G}_n^{\text{ME}} - \mathbf{H}_n^{\text{ME}}\mathbf{F}^T\|_F^2\}_{n=1}^N$.
- 7: **end while**
- 8: **return** \mathbf{H}^{ME} .

ALGORITHM 1: Learning procedure of RGF-MEC $_{F_2}$

3.1.1. *Fix \mathbf{F} and Update \mathbf{H}^{ME} .* If \mathbf{F} is fixed, problem (6) will become

$$\min_{\mathbf{H}^{\text{ME}} \geq 0} \sum_{n=1}^N \mathbf{1}_{RG_n \leq \varepsilon} \cdot \|\mathbf{G}_n^{\text{ME}} - \mathbf{H}_n^{\text{ME}}\mathbf{F}^T\|_F^2 + \frac{\lambda}{2} \sum_{i,j} \|\mathbf{H}_i^{\text{ME}} - \mathbf{H}_j^{\text{ME}}\|_F^2 G_{ij}. \quad (8)$$

Setting the derivative of problem (8) equal to zeros, we have the following equation:

$$2(\mathbf{H}^{\text{ME}} - \Lambda_\varepsilon \mathbf{G}^{\text{ME}}\mathbf{F}) + 4\lambda \mathbf{L}^{\text{ME}} \mathbf{H}^{\text{ME}} = 0, \quad (9)$$

where Λ_ε is a diagonal matrix with n th diagonal element $\mathbf{1}_{RG_n \leq \varepsilon}$ and \mathbf{L}^{ME} is the Laplacian matrix of \mathbf{G}^{ME} . Therefore, the solution of problem (8) is

$$\mathbf{H}^{\text{ME}} = \max\left(0, (\Lambda_\varepsilon + 2\lambda \mathbf{L}^{\text{ME}})^{-1} (\Lambda_\varepsilon \mathbf{G}^{\text{ME}}\mathbf{F})\right). \quad (10)$$

3.1.2. *Fix \mathbf{H}^{ME} and Update \mathbf{F} .* If \mathbf{H}^{ME} is fixed, problem (6) will become

$$\min_{\mathbf{F}^T \mathbf{F} = \mathbf{I}} \sum_{n=1}^N \mathbf{1}_{RG_n \leq \varepsilon} \cdot \|\mathbf{G}_n^{\text{ME}} - \mathbf{H}_n^{\text{ME}}\mathbf{F}^T\|_F^2. \quad (11)$$

Due to $\mathbf{1}_{RG_n \leq \varepsilon} \in \{0, 1\}$ and $\mathbf{F}^T \mathbf{F} = \mathbf{I}$, problem (11) can be written as

$$\begin{aligned} (11) &\Leftrightarrow \min_{\mathbf{F}^T \mathbf{F} = \mathbf{I}} \left\| \Lambda_\varepsilon^{1/2} (\mathbf{G}^{\text{ME}} - \mathbf{H}^{\text{ME}}\mathbf{F}^T) \right\|_F^2 \\ &\Leftrightarrow \min_{\mathbf{F}^T \mathbf{F} = \mathbf{I}} \left\| \Lambda_\varepsilon (\mathbf{G}^{\text{ME}} - \mathbf{H}^{\text{ME}}\mathbf{F}^T) \right\|_F^2 \\ &\Leftrightarrow \min_{\mathbf{F}^T \mathbf{F} = \mathbf{I}} \text{Tr}(\mathbf{G}^{\text{MET}} \Lambda_\varepsilon \mathbf{H}^{\text{ME}}\mathbf{F}^T), \end{aligned} \quad (12)$$

which is the standard orthogonal procrustes problem. Defining the singular value decomposition of $\mathbf{G}^{\text{MET}} \Lambda_\varepsilon \mathbf{H}^{\text{ME}}$ as \mathbf{USV}^T , the solution of problem (11) is obtained:

$$\mathbf{F} = \mathbf{UV}^T. \quad (13)$$

3.2. *Learning Procedure of RGF-MEC $_{\ell_{2,1}}$.* Similar to RGF-MEC $_{F_2}$, RGF-MEC $_{\ell_{2,1}}$ can make use of the coordinate descent method to solve problem (7), too.

3.2.1. *Fix \mathbf{F} and Update \mathbf{H}^{ME} .* If \mathbf{F} is fixed, problem (7) will become

$$\min_{\mathbf{H}^{\text{ME}} \geq 0} \sum_{n=1}^N \mathbf{1}_{RG_n \leq \varepsilon} \cdot \|\mathbf{G}_n^{\text{ME}} - \mathbf{H}_n^{\text{ME}}\mathbf{F}^T\|_{2,1} + \frac{\lambda}{2} \sum_{i,j} \|\mathbf{H}_i^{\text{ME}} - \mathbf{H}_j^{\text{ME}}\|_F^2 G_{ij}. \quad (14)$$

Setting the derivative of problem (14) equal to zeros, we have the following equation:

$$\sum_{n=1}^N \mathbf{1}_{RG_n \leq \varepsilon} \cdot a_n \frac{\partial \left(\|\mathbf{G}_n^{\text{ME}} - \mathbf{H}_n^{\text{ME}}\mathbf{F}^T\|_F^2 \right)}{\partial \mathbf{H}_n^{\text{ME}}} + 4\lambda \mathbf{L}^{\text{ME}} \mathbf{H}^{\text{ME}} = 0, \quad (15)$$

where $a_n = 1/2 \|\mathbf{G}_n^{\text{ME}} - \mathbf{H}_n^{\text{ME}}\mathbf{F}^T\|_F$. If a_n is stationary, problem (14) can be written as

$$\min_{\mathbf{H}^{\text{ME}} \geq 0} \sum_{n=1}^N a_n \mathbf{1}_{RG_n \leq \varepsilon} \cdot \|\mathbf{G}_n^{\text{ME}} - \mathbf{H}_n^{\text{ME}}\mathbf{F}^T\|_F^2 + \frac{\lambda}{2} \sum_{i,j} \|\mathbf{H}_i^{\text{ME}} - \mathbf{H}_j^{\text{ME}}\|_F^2 G_{ij}. \quad (16)$$

Setting derivatives of problem (16) equal to zeros, we have the following equation:

$$2(\mathbf{H}^{\text{ME}} - \Lambda_{\varepsilon a} \mathbf{G}^{\text{ME}}\mathbf{F}) + 4\lambda \mathbf{L}^{\text{ME}} \mathbf{H}^{\text{ME}} = 0, \quad (17)$$

where $\Lambda_{\varepsilon a}$ is a diagonal matrix with n th diagonal element $a_n \mathbf{1}_{RG_n \leq \varepsilon}$.

Therefore, a solution to problem (14) is

$$\mathbf{H}^{\text{ME}} = \max\left(0, (\Lambda_{\varepsilon a} + 2\lambda \mathbf{L}^{\text{ME}})^{-1} (\Lambda_{\varepsilon a} \mathbf{G}^{\text{ME}}\mathbf{F})\right). \quad (18)$$

3.2.2. *Fix \mathbf{H}^{ME} and Update \mathbf{F} .* If \mathbf{H}^{ME} is fixed, problem (7) will become

$$\min_{\mathbf{F}^T \mathbf{F} = \mathbf{I}} \sum_{n=1}^N \mathbf{1}_{RG_n \leq \varepsilon} \cdot \|\mathbf{G}_n^{\text{ME}} - \mathbf{H}_n^{\text{ME}}\mathbf{F}^T\|_{2,1}. \quad (19)$$

Similarly, problem (19) can be written as

$$\begin{aligned}
(19) &\Leftrightarrow \sum_{n=1}^N a_n \mathbf{1}_{RG_n \leq \varepsilon} \cdot \|\mathbf{G}_n^{\text{ME}} - \mathbf{H}_n^{\text{ME}} \mathbf{F}^T\|_F^2 \\
&\Leftrightarrow \min_{\mathbf{F}^T \mathbf{F} = \mathbf{I}} \|\Lambda_{a\varepsilon}^{1/2} (\mathbf{G}^{\text{ME}} - \mathbf{H}^{\text{ME}} \mathbf{F}^T)\|_F^2 \\
&\Leftrightarrow \min_{\mathbf{F}^T \mathbf{F} = \mathbf{I}} \text{Tr}(\mathbf{G}^{\text{MET}} \Lambda_{a\varepsilon} \mathbf{H}^{\text{ME}} \mathbf{F}^T),
\end{aligned} \tag{20}$$

which is the standard orthogonal procrustes problem. Defining the singular value decomposition of $\mathbf{G}^{\text{MET}} \Lambda_{a\varepsilon} \mathbf{H}^{\text{ME}}$ as $\mathbf{U}' \mathbf{S}' \mathbf{V}'^T$, a solution to problem (19) is obtained:

$$\mathbf{F} = \mathbf{U}' \mathbf{V}'^T. \tag{21}$$

3.3. Convergence Analysis. In RGF-MEC_{F2}, two embeddings \mathbf{H}^{ME} and \mathbf{F} are updated in turn. When updating \mathbf{H}^{ME} and \mathbf{F} , we know that (10) (i.e., the calculation of \mathbf{H}^{ME}) is the local optimal solution of (8) and (13) (i.e., the calculation of \mathbf{F}) is the local optimal solution of (11). Therefore, RGF-MEC_{F2} converges to its local optimal solution.

In RGF-MEC_{ℓ2,1}, two embeddings \mathbf{H}^{ME} and \mathbf{F} are updated in turn, too. It can be seen that (18) (i.e., the calculation of \mathbf{H}^{ME}) is the local optimal solution of (16) instead of (14).

Theorem 1. *If (18) (i.e., the calculation of \mathbf{H}^{ME}) is a solution to (14), the objective value of (14) will not increase.*

Proof. Suppose $e_n = \|(\mathbf{G}_n^{\text{ME}} - \mathbf{H}_n^{\text{ME}} \mathbf{F}^T)\|_F^2$ and $a_n = 1/\sqrt{e_n}$. In RGF-MEC_{ℓ2,1}, the minimizing problem of $\sqrt{e_n}$ in (14) is transformed into the minimizing problem of $a_n e_n$ in (16). Suppose that a_n is stationary and $g(\mathbf{H}) = \lambda/2 \sum_{i,j} \|\mathbf{H}_i^{\text{ME}} - \mathbf{H}_j^{\text{ME}}\|_F^2 G_{ij}$, (18) is a local optimal solution of (16). That is,

$$\begin{aligned}
\sum_n a_n^t e_n^{t+1} \mathbf{1}_{RG_n \leq \varepsilon} + g(\mathbf{H}^{t+1}) &\leq \sum_n a_n^t e_n^t \mathbf{1}_{RG_n \leq \varepsilon} + g(\mathbf{H}^t), \\
\Leftrightarrow \sum_n e_n^{t+1} \mathbf{1}_{RG_n \leq \varepsilon} + g(\mathbf{H}^{t+1}) &\leq \sum_n e_n^t \mathbf{1}_{RG_n \leq \varepsilon} + g(\mathbf{H}^t) \\
\Leftrightarrow \sum_n \sqrt{e_n^{t+1}} \mathbf{1}_{RG_n \leq \varepsilon} + (g \mathbf{H}^{t+1}) &\leq \sum_n \sqrt{e_n^t} \mathbf{1}_{RG_n \leq \varepsilon} + (g \mathbf{H}^t).
\end{aligned} \tag{22}$$

Therefore, the objective value of (14) will not increase and Theorem 1 is proven.

It can be also seen that (21) (i.e., the calculation of \mathbf{F}) is a local optimal solution of (19). In other words, RGF-MEC_{ℓ2,1} converges to its local optimal solution, too. \square

4. Experiments

In this section, experimental results on multiple multivariate time-series datasets and a multivariate electricity consumption series dataset are used to validate the effectiveness of the proposed RGF-MEC framework.

4.1. Experiments on Multiple Multivariate Time-Series Datasets

4.1.1. Multivariate Time-Series Datasets. Multiple realistic multivariate time-series datasets are used to evaluate the proposed RGF-MEC framework, including ArticularyWordRecognition [29], BasicMotions [30], Epilepsy [31], ERing [32], Libras [33], NATOPS [34], StandWalkJump [35], and UWaveGestureLibrary [36]. Among these multivariate time-series datasets, the maximum number of sequences is up to 24, the maximum length of these sequences is up to 2500, the maximum size of instances is 575, and the maximum number of classes is up to 25.

4.1.2. Contrast Algorithms. We compare RGF-MEC with seven representative methods, including m-kShape, m-kDBA, m-kAVG + ED, m-KSC, multiview clustering with adaptive neighbors (MLAN) [37], and multiview spectral clustering via integrating nonnegative embedding and spectral embedding (NESE) [38]. m-kShape, m-kDBA, m-kAVG + ED, and m-KSC are effective distance-based multivariate time-series clustering algorithms [16], which consider the relationship among multiple series and extend traditional univariate time-series clustering methods (i.e., k-shape, KDBA, kAVG + ED, and KSC) for multivariate time-series clustering. MLAN and NESE are multiview clustering algorithms, where each series of a multivariate time series can be seen as a view.

In RGF-MEC, λ is selected from $\{10^{-5}, 10^{-4}, 10^{-3}, 10^{-2}, 10^1, 10^0, 10^1, 10^2, 10^3\}$, and ε is determined according to the reconstruction error and a parameter r that is a hypothetical ratio of outliers in the entire data. r is selected from $\{0\%, 1\%, 2\%, 3\%, 5\%, 10\%, 15\%, 20\%\}$. In addition, parameters of these contrast algorithms are determined according to descriptions in their original papers. We utilize two widely used measurements (i.e., rand index (RI) and normalized mutual information (NMI)) for evaluation. For RI and NMI, a larger value indicates a better result.

4.1.3. Performance Comparison. We first compare the performance of RGF-MEC and other contrast algorithms on real-world multivariate time-series datasets. Performance comparisons between RGF-MEC and other contrast algorithms are reported in Tables 1 and 2, where best results are highlighted in bold. It can be seen that RGF-MEC outperforms other contrast algorithms on most real-world multivariate time-series datasets, where RGF-MEC achieves five best performances on eight real-world datasets in terms of RI or NMI. We can also make the following conclusions from Tables 1 and 2. (1) RGF-MEC is clearly better than other contrast algorithms in terms of “Mean \pm Std,” which demonstrates that RGF-MEC is an effective multivariate time-series clustering method. (2) The performance of NESE is best on the epilepsy dataset, which indicates that each series contains complementary information to each other. If each series is sufficient to represent the similarity between multivariate time series, then the multiview model can make

TABLE 1: Performance comparisons between RGF-MEC and other contrast algorithms on real-world multivariate time-series datasets in terms of RI.

Datasets	m-kAVG + ED	m-kDBA	m-kShape	m-kSC	MLAN	NESE	RGF-MEC
ArticulatoryWordRecognition	0.9522	0.9336	0.7582	0.9510	0.9899	0.9756	0.9960
BasicMotions	0.7718	0.7487	0.5244	0.7718	0.7628	0.7449	0.8090
Epilepsy	0.7684	0.7771	0.5136	0.6044	0.7330	0.8897	0.7269
ERing	0.8046	0.7747	0.7701	0.7494	0.7471	0.7540	0.7747
Libras	0.9111	0.9133	0.6605	0.9227	0.9161	0.9087	0.9255
NATOPS	0.8525	0.8755	0.6534	0.8348	0.7992	0.7637	0.8631
StandWalkJump	0.7333	0.6952	0.3485	0.6571	0.7048	0.6476	0.7429
UWaveGestureLibrary	0.9204	0.8934	0.8015	0.9259	0.9434	0.8553	0.9524
Mean \pm std	0.8393 \pm 0.0764	0.8264 \pm 0.0824	0.6288 \pm 0.1462	0.8021 \pm 0.1209	0.8245 \pm 0.1019	0.8174 \pm 0.1006	0.8488 \pm 0.0946

TABLE 2: Performance comparisons between RGF-MEC and other contrast algorithms on real-world multivariate time-series datasets in terms of NMI.

Datasets	m-kAVG + ED	m-kDBA	m-kShape	m-kSC	MLAN	NESE	RGF-MEC
ArticulatoryWordRecognition	0.8336	0.7409	0.3435	0.7922	0.9435	0.8486	0.9762
BasicMotions	0.5432	0.6388	0.3405	0.5539	0.5648	0.5250	0.6592
Epilepsy	0.4089	0.4707	0.1632	0.3811	0.5318	0.7600	0.3961
ERing	0.4003	0.4058	0.2683	0.3478	0.3185	0.3778	0.4040
Libras	0.6220	0.6224	0.4474	0.7244	0.6141	0.5418	0.6724
NATOPS	0.6430	0.6427	0.3392	0.5998	0.8247	0.3139	0.7755
StandWalkJump	0.5587	0.4658	0.1163	0.4608	0.5545	0.3985	0.5555
UWaveGestureLibrary	0.7134	0.5824	0.4194	0.7582	0.8385	0.5588	0.8395
Mean ± std	0.5904 ± 0.1371	0.5712 ± 0.1059	0.3047 ± 0.1086	0.5773 ± 0.1606	0.6488 ± 0.1916	0.5405 ± 0.1738	0.6598 ± 0.1910

TABLE 3: Performance comparisons between RGF-MEC_{F2} and RGF-MEC_{ℓ2,1} on real-world multivariate time-series datasets.

Datasets	RI		NMI	
	RGF-MEC _{F2}	RGF-MEC _{ℓ2,1}	RGF-MEC _{F2}	RGF-MEC _{ℓ2,1}
ArticularyWordRecognition	0.9960	0.9960	0.9762	0.9762
BasicMotions	0.7910	0.8090	0.6592	0.6112
Epilepsy	0.7269	0.7252	0.3856	0.3961
ERing	0.7747	0.7747	0.4040	0.4040
Libras	0.9255	0.9241	0.6724	0.6651
NATOPS	0.8631	0.8616	0.7155	0.7755
StandWalkJump	0.7048	0.7429	0.4489	0.5555
UWaveGestureLibrary	0.9524	0.9507	0.8395	0.8359
Mean ± Std	0.8418 ± 0.1016	0.8480 ± 0.0945	0.6377 ± 0.1990	0.6524 ± 0.1912

TABLE 4: Performance comparisons between ablation models and RGF-MEC on real-world multivariate time-series datasets.

Datasets	RI				NMI			
	Ablation _{F2}	RGF-MEC _{F2}	Ablation _{ℓ2,1}	RGF-MEC _{ℓ2,1}	Ablation _{F2}	RGF-MEC _{F2}	Ablation _{ℓ2,1}	RGF-MEC _{ℓ2,1}
ArticularyWordRecognition	0.9960	0.9960	0.9960	0.9960	0.9762	0.9762	0.9762	0.9762
BasicMotions	0.7782	0.7910	0.7885	0.8090	0.5832	0.6592	0.5961	0.6112
Epilepsy	0.7269	0.7269	0.7074	0.7252	0.3856	0.3856	0.3248	0.3961
ERing	0.7747	0.7747	0.7724	0.7747	0.4040	0.4040	0.3802	0.4040
Libras	0.9240	0.9255	0.9222	0.9241	0.6724	0.6724	0.6575	0.6651
NATOPS	0.8624	0.8631	0.8567	0.8616	0.7147	0.7155	0.7055	0.7755
StandWalkJump	0.7048	0.7048	0.7429	0.7429	0.4489	0.4489	0.5555	0.5555
UWaveGestureLibrary	0.9524	0.9524	0.9404	0.9507	0.8368	0.8395	0.8031	0.8359

full use of information of each series and achieve good performance.

We then compare the performance of RGF-MEC_{F2} and RGF-MEC_{ℓ2,1} on real-world multivariate time-series datasets. Table 3 gives performance comparisons between RGF-MEC_{F2} and RGF-MEC_{ℓ2,1}. It can be seen that RGF-MEC_{ℓ2,1} is clearly better than RGF-MEC_{F2} in terms of “Mean ± Std.” In other words, RGF-MEC_{ℓ2,1} achieves mean improvements of 0.74% RI and 2.31% NMI compared to RGF-MEC_{F2}. Compared with RGF-MEC_{F2}, RGF-MEC_{ℓ2,1} needs to compute weights $\{a_n = 1/2\|(\mathbf{G}_n^{\text{ME}} - \mathbf{H}_n^{\text{ME}}\mathbf{F}^T)\|_F\}_{n=1}^N$ when updating the non-negative class-related embedding \mathbf{H}^{ME} , where weights $\{a_n\}_{n=1}^N$ can be considered as parameters that penalize larger reconstruction errors. This may make RGF-MEC_{ℓ2,1} more robust on most multivariate time-series datasets.

We also perform ablation studies on RGF-MEC. The objective function of RGF-MEC consists of two parts, the robust graph factorization term and the graph regularization term. Table 4 gives the performance comparisons between the ablation model and RGF-MEC, where Ablation_{F2} denotes RGF-MEC_{F2} without the graph regularization term and Ablation_{ℓ2,1} denotes RGF-MEC_{ℓ2,1} without the graph regularization term. It can be seen that RGF-MEC_{ℓ2,1} (resp. RGF-MEC_{F2}) outperforms Ablation_{ℓ2,1} (resp. Ablation_{F2}) on most multivariate time-series datasets, which prove that the graph regularization term enforces the nonnegative embedding \mathbf{H}^{ME} to satisfy the graph constraint and retain more class-related information.

Finally, we analyze the impact of parameters r and λ on the algorithm performance. Figure 1 shows the performance (i.e., NMI) change process of RGF-MEC_{ℓ2,1} on four

multivariate time-series datasets as r increases. It can be seen that RGF-MEC_{ℓ2,1} takes outliers into consideration and achieves better clustering results on four multivariate time-series datasets. Figure 2 shows the performance (i.e., NMI) change process of RGF-MEC_{ℓ2,1} on four multivariate time-series datasets as λ increases. It can be seen that RGF-MEC_{ℓ2,1} takes the graph constraint of the nonnegative embedding (i.e., $\lambda/2\sum_{i,j}\|\mathbf{H}_i^{\text{ME}} - \mathbf{H}_j^{\text{ME}}\|_F^2 G_{ij}$) into consideration and achieves better clustering results on most multivariate time-series datasets, which is consistent with the results of the ablation study. Next, as shown in Figure 3, we use t -SNE [39] to visualize the class-related embedding \mathbf{H}^{ME} of RGF-MEC_{ℓ2,1} on UWaveGestureLibrary. It can be seen that some classes have obvious outliers, such as class-4, class-6, and class-8. RGF-MEC_{ℓ2,1} takes outliers into consideration and performs better than other contrast algorithms.

4.2. Experiments on Multivariate Electricity Consumption Series Dataset

4.2.1. Multivariate Electricity Consumption Series Dataset.

A multivariate electricity consumption series dataset of a region in China Southern Power Grid [5], denoted as the CSPG dataset, is used to evaluate the proposed RGF-MEC framework. CSPG contains four series, global-active power (household global minute-averaged active power), global-reactive power (household global minute-averaged reactive power), voltage (minute-averaged voltage), and global intensity (household global minute-averaged current intensity). In addition, it contains multivariate series for two

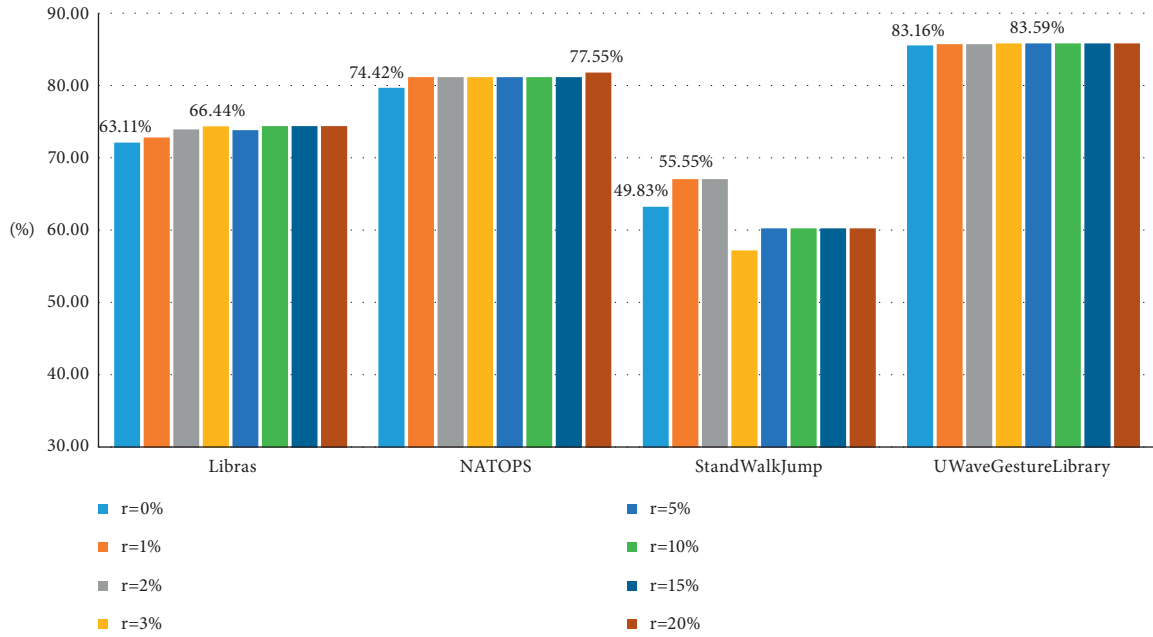


FIGURE 1: Performance (i.e., NMI) change process of RGF-MEC_{ℓ_{2,1}} on four multivariate time-series datasets as r increases.

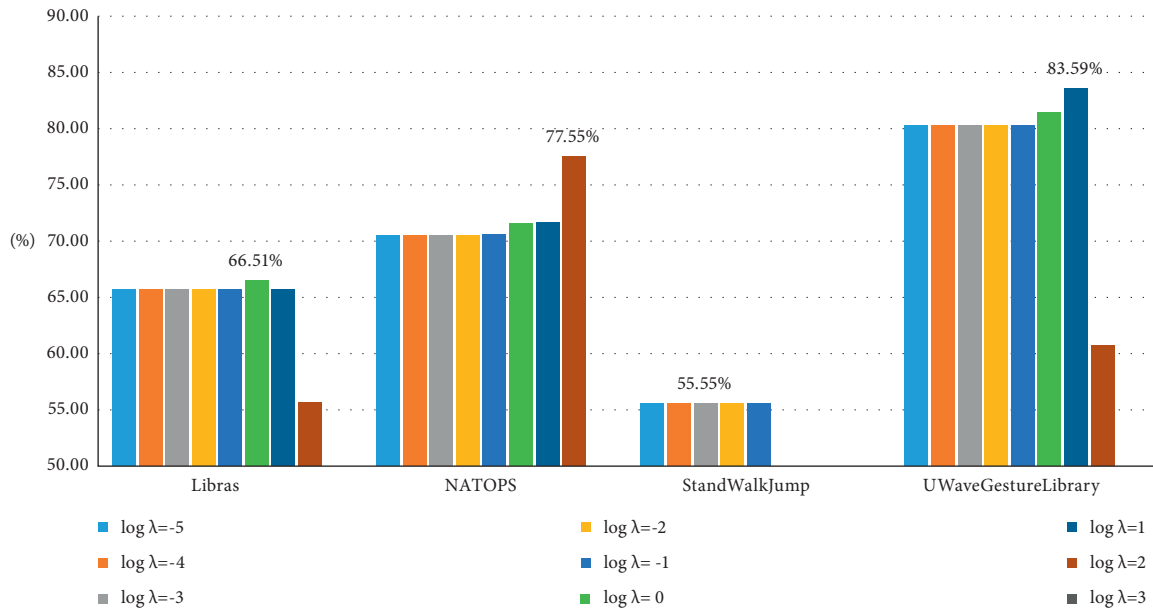


FIGURE 2: Performance (i.e., NMI) change process of RGF-MEC_{ℓ_{2,1}} on four multivariate time-series datasets as λ increases, where algorithms with NMI less than 50% cannot be displayed in this figure.

different time periods, where the length of each series is 180 and the number of instances is 2780.

4.2.2. *Performance Comparison.* Since the effectiveness of RGF-MEC has been verified in Section 4-4.1, we first give results obtained by RGF-MEC in different series combinations on CSPG. Table 5 shows performance comparisons between the combination containing four series and combinations containing arbitrary two series. Compared with the combinations containing arbitrary two series, RGF-MEC achieves best results on the combination

containing four series. Next, we analyze the impact of parameters r, λ on CSPG. Figure 4 shows the performance (i.e., RI and NMI) change process of RGF-MEC on CSPG as r increases. It can be seen that RGF-MEC takes outliers into consideration and achieves better clustering results on CSPG. Figure 5 shows the performance (i.e., RI and NMI) change process of RGF-MEC on CSPG as λ increases. It can be seen that RGF-MEC takes the graph constraint of the nonnegative embedding (i.e., $\lambda/2 \sum_{i,j} \|\mathbf{H}_i^{\text{ME}} - \mathbf{H}_j^{\text{ME}}\|_F^2 G_{ij}$) into consideration and achieves better clustering results on CSPG.

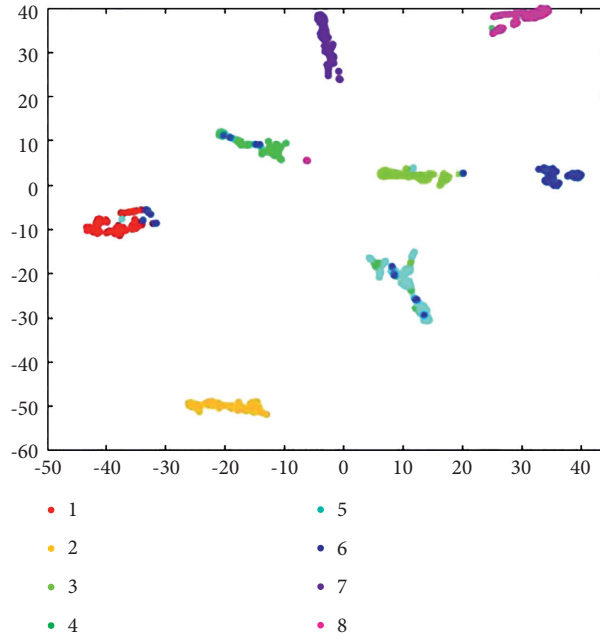


FIGURE 3: Visualization of the class-related embedding H^{ME} of $RGF-MEC_{\ell_{2,1}}$ on UWaveGestureLibrary.

TABLE 5: Performance comparisons among different series combinations on the CSPG dataset.

Different series combinations				RI		NMI	
Global-active power	Global-reactive power	Voltage	Global intensity	$RGF-MEC_{F_2}$	$RGF-MEC_{\ell_{2,1}}$	$RGF-MEC_{F_2}$	$RGF-MEC_{\ell_{2,1}}$
✓	✓	✗	✗	0.7394	0.7394	0.4220	0.4158
✓	✗	✓	✗	0.7394	0.7485	0.3974	0.4218
✓	✗	✗	✓	0.6554	0.6508	0.3452	0.3290
✗	✓	✓	✗	0.6490	0.6351	0.2477	0.2400
✓	✓	✗	✓	0.7349	0.7394	0.4125	0.4138
✗	✗	✓	✓	0.7416	0.7349	0.4043	0.4153
✓	✓	✓	✓	0.7624	0.7600	0.4462	0.4585

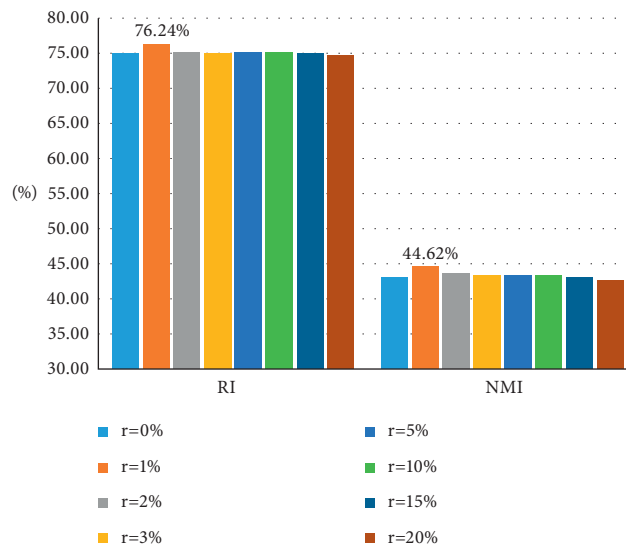


FIGURE 4: Performance (i.e., RI and NMI) change process of RGF-MEC on the CSPG dataset as r increases.

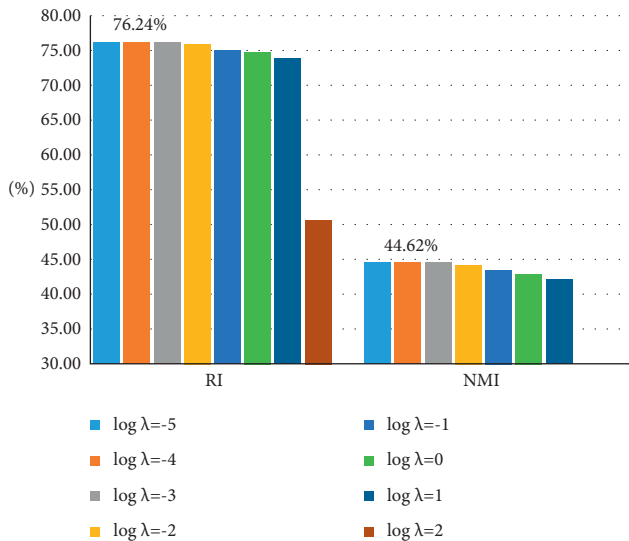


FIGURE 5: Performance (i.e., RI and NMI) change process of RGF-MEC on the CSPG dataset as λ increases, where algorithms with NMI less than 30% cannot be displayed in this figure.

5. Conclusion

This paper builds an effective robust graph factorization framework for multivariate electricity consumption clustering, which performs robust matrix factorization on the similarity of multivariate electricity consumption series and obtains nonnegative class-related representations. The proposed RGF-MEC framework also enforces the nonnegative embedding to satisfy a graph constraint and retain more class-related information. In addition, RGF-MEC utilizes the squared F-norm and the $\ell_{2,1}$ -norm to constrain the robust graph factorization term, named RGF-MEC $_{F_2}$ and RGF-MEC $_{\ell_{2,1}}$. Experimental results demonstrate that RGF-MEC $_{F_2}$ and RGF-MEC $_{\ell_{2,1}}$ achieve competitive results on multiple multivariate time series datasets and a multivariate electricity consumption series dataset.

Data Availability

Eight realistic multivariate time-series datasets are deposited in a public repository (<http://www.timeseriesclassification.com/dataset.php>). The multivariate electricity consumption series dataset used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work was supported by China Southern Power Grid Co., Ltd. and China Southern Power Grid Digital Grid Research Institute Co., Ltd.

References

- [1] J. Xu, X. Kang, Z. Chen et al., "Clustering-based probability distribution model for monthly residential building electricity consumption analysis," *Building Simulation*, vol. 14, no. 1, pp. 149–164, 2021.
- [2] L. Wen, K. Zhou, and S. Yang, "A shape-based clustering method for pattern recognition of residential electricity consumption," *Journal of Cleaner Production*, vol. 212, pp. 475–488, 2019.
- [3] Y. Wang, Q. Chen, C. Kang, and Q. Xia, "Clustering of electricity consumption behavior dynamics toward big data applications," *IEEE Transactions on Smart Grid*, vol. 7, no. 5, pp. 2437–2447, 2016.
- [4] M. M. Rahman, "Environmental degradation: the role of electricity consumption, economic growth and globalisation," *Journal of Environmental Management*, vol. 253, Article ID 109742, 2020.
- [5] K. Zheng, P. Li, S. Zhou et al., "A multi-scale electricity consumption prediction algorithm based on time-frequency variational autoencoder," *IEEE Access*, vol. 35, no. 1–1, 2021.
- [6] K. Gajowniczek, M. Bator, and T. Ząbkowski, "Whole time series data streams clustering: dynamic profiling of the electricity consumption," *Entropy*, vol. 22, no. 12, Article ID 1414, 2020.
- [7] L. Li, W. Li, J. Liao, and X. Hu, "Adaptive state continuity-based sparse inverse covariance clustering for multivariate time series," in *Proceedings of International Conference on Security, Pattern Analysis, and Cybernetics*, pp. 68–74, Guangzhou, China, December 2019.
- [8] Y. Tang, Y. Xie, X. Yang, J. Niu, and W. Zhang, "A multi-scale electricity consumption prediction algorithm based on time-frequency variational autoencoder," *IEEE Access*, vol. 33, no. 3, pp. 1223–1237, 2021.
- [9] Y. Zhao, L. Lin, W. Lu, and Y. Meng, "Landsat time series clustering under modified dynamic time warping," in *Proceedings of International Workshop on Earth Observation and Remote Sensing Applications*, pp. 62–66, Guangzhou, China, July 2016.
- [10] Z. Yue and V. Solo, "Large-scale time series clustering with k-ars," in *Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing*, pp. 6044–6048, Barcelona, Spain, May 2020.
- [11] A. Abdella and I. Uysal, "Sense2vec: representation and visualization of multivariate sensory time series data," *IEEE Sensors Journal*, vol. 21, no. 6, pp. 7972–7988, 2021.
- [12] J. Paparrizos and L. Gravano, "k-shape: efficient and accurate clustering of time series," *Proceedings of ACM SIGMOD Record*, vol. 45, no. 1, pp. 1855–1870, 2015.
- [13] F. Petitjean, A. Ketterlin, and P. Gançarski, "A global averaging method for dynamic time warping, with applications to clustering," *Pattern Recognition*, vol. 44, no. 3, pp. 678–693, 2011.
- [14] C. Faloutsos, M. Ranganathan, and Y. Manolopoulos, "Fast subsequence matching in time-series databases," *ACM SIGMOD Record*, vol. 23, no. 2, pp. 419–429, 1994.
- [15] J. Yang and J. Leskovec, "Patterns of temporal variation in online media," in *Proceedings of ACM International Conference on Web Search and Data Mining*, pp. 177–186, Hong Kong China, February 2011.
- [16] M. Ozer, A. Sapienza, A. Abeliuk, G. Muric, and E. Ferrara, "Discovering patterns of online popularity from time series," *Expert Systems with Applications*, vol. 151, Article ID 113337, 2020.

- [17] H. He and Y. Tan, "Unsupervised classification of multivariate time series using VPCA and fuzzy clustering with spatial weighted matrix distance," *IEEE Transactions on Cybernetics*, vol. 50, no. 3, pp. 1096–1105, 2020.
- [18] H. Li, "Accurate and efficient classification based on common principal components analysis for multivariate time series," *Neurocomputing*, vol. 171, pp. 744–753, 2016.
- [19] Z. X. Li, J. S. Guo, X. B. Hui, and F. F. Song, "Dimension reduction method for multivariate time series based on common principal component," *Control and Decision*, vol. 28, no. 4, pp. 531–536, 2013.
- [20] H. Li, "Multivariate time series clustering based on common principal component analysis," *Neurocomputing*, vol. 349, pp. 239–247, 2019.
- [21] J. Y. Franceschi, A. Dieuleveut, and M. Jaggi, "Unsupervised scalable representation learning for multivariate time series," in *Thirty-third Conference on Neural Information Processing Systems, Neural Information Processing Systems Foundation*, pp. 1–12, Vancouver, Canada, December 2019.
- [22] F. Liu, Y. Lu, and M. Cai, "A hybrid method with adaptive sub-series clustering and attention-based stacked residual lstms for multivariate time series forecasting," *IEEE Access*, vol. 8, pp. 62423–62438, 2020.
- [23] L. Hagen and A. B. Kahng, "New spectral methods for ratio cut partitioning and clustering," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 11, no. 9, pp. 1074–1085, 1992.
- [24] D. Kuang, C. Ding, and H. Park, "Symmetric nonnegative matrix factorization for graph clustering," in *Proceedings of SIAM International Conference on Data Mining*, pp. 106–117, Anaheim, CA, USA, 2012.
- [25] J. Han, K. Xiong, and F. Nie, "Orthogonal and nonnegative graph reconstruction for large scale clustering," in *Proceedings of International Joint Conference on Artificial Intelligence*, pp. 1809–1815, Melbourne, Australia, August 2017.
- [26] L. Zhang, Q. Zhang, B. Du, D. Tao, and J. You, "Robust manifold matrix factorization for joint clustering and feature extraction," in *Proceedings of the 31st AAAI Conference on Artificial Intelligence*, pp. 1662–1668, San Francisco, CA, USA, February 2017.
- [27] Z. Li, J. Tang, and X. He, "Robust structured nonnegative matrix factorization for image representation," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 29, no. 5, pp. 1947–1960, 2018.
- [28] M. Chen and X. Li, "Robust matrix factorization with spectral embedding," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 15, pp. 1–10, 2020.
- [29] M. Shokoohi-Yekta, B. Hu, H. Jin, J. Wang, and E. Keogh, "Generalizing DTW to the multi-dimensional case requires an adaptive approach," *Data Mining and Knowledge Discovery*, vol. 31, no. 1, pp. 1–31, 2017.
- [30] A. Bagnall, H. A. Dau, J. Lines et al., "The Uea multivariate time series classification archive," 2018, <https://arxiv.org/abs/1811.00075>.
- [31] J. R. Villar, P. Vergara, M. Menéndez, E. De la Cal, V. M. González, and J. Sedano, "Generalized models for the classification of abnormal movements in daily life and its applicability to epilepsy convulsion recognition," *International Journal of Neural Systems*, vol. 26, no. 6, Article ID 1650037, 2016.
- [32] M. Wilhelm, D. Krakowczyk, F. Trollmann, and S. Albayrak, "Ering: multiple finger gesture recognition with one ring using an electric field," in *Proceedings of International Workshop on Sensor-Based Activity Recognition and Interaction*, p. 7, Rostock Germany, June 2015.
- [33] D. B. Dias, R. C. B. Madeo, T. Rocha, and H. H. Biscaro, "Hand movement recognition for brazilian sign language: a study using distance-based neural networks," in *Proceedings of International Joint Conference on Neural Networks*, pp. 697–704, Atlanta, GA, USA, June 2009.
- [34] N. Ghouaiel, P. F. Marteau, and M. Dupont, "Continuous pattern detection and recognition in stream - a benchmark for online gesture recognition," *International Journal of Applied Pattern Recognition*, vol. 4, no. 2, pp. 146–160, 2017.
- [35] V. Behravan, N. E. Glover, R. Farry, P. Y. Chiang, and M. Shoaib, "Rateadaptive compressed-sensing and sparsity variance of biomedical signals," in *Proceedings of the 2015 IEEE 12th International Conference on Wearable and Implantable Body Sensor Networks (BSN)*, pp. 1–6, Cambridge, MA, USA, June 2015.
- [36] J. Liu, L. Zhong, J. Wickramasuriya, and V. Vasudevan, "uwave: accelerometer-based personalized gesture recognition and its applications," *Pervasive and Mobile Computing*, vol. 5, no. 6, pp. 657–675, 2009.
- [37] F. Nie, G. Cai, and X. Li, "Multi-view clustering and semi-supervised classification with adaptive neighbours," in *Proceedings of the 31st AAAI Conference on Artificial Intelligence*, pp. 2408–2414, San Francisco, CA, USA, February 2017.
- [38] Z. Hu, F. Nie, R. Wang, and X. Li, "Multi-view spectral clustering via integrating nonnegative embedding and spectral embedding," *Information Fusion*, vol. 55, pp. 251–259, 2020.
- [39] L. van der Maaten and G. Hinton, "Visualizing data using t-sne," *Journal of Machine Learning Research*, vol. 9, no. 86, pp. 2579–2605, 2008.