Fair Spectral Clustering with Multiple Sensitive Attributes via Shapley Values

Tao Wu[1], Zhijing Yang[2], Junjie Zheng[1], Hui Zhang[1*]
[1]Southwest University of Science and Technology, Mianyang, China
[2]Southeast University, Nanjing, China
Email: wutao@mails.swust.edu.cn; yangzj@seu.edu.cn;
zhengjunjie@mails.swust.edu.cn; zhanghui@swust.edu.cn

***Abstract*—Fair spectral clustering has attracted increasing attention in recent years. Although numerous methods have been proposed with promising results, most of them focus on a single sensitive attribute, while fair spectral clustering algorithms ensuring multiple sensitive attributes have received relatively little attention. To solve this issue, this paper introduces an innovative Multi-sensitive Attribute Fair Spectral Clustering algorithm (MFSC) by incorporating Shapley value theory. To verify the effectiveness of the proposed MFSC, we conducted experiments on four benchmark datasets, comparing it with spectral clustering and fair spectral clustering methods in terms of fairness and clustering performance. Experimental results demonstrate that MFSC substantially improves fairness while maintaining clustering quality.**

***Index Terms*—spectral clustering, fairness, multiple sensitive attributes**

I. INTRODUCTION

Spectral clustering has garnered significant attention due to its robust theoretical foundation rooted in graph theory [1]. In spectral clustering, the clustering problem is reformulated as a graph partitioning problem, where each sample is represented as a vertex and similarities define edge weights [2]. It typically constructs a similarity matrix, computes the graph Laplacian, and performs eigendecomposition to obtain a low-dimensional spectral embedding, followed by standard algorithms such as k-means to produce the final partitions [3]. However, conventional spectral clustering optimizes graph-cut objectives without explicitly accounting for sensitive attributes, which can lead to unfairness toward protected groups. In critical decision-making areas such as credit approval, medical diagnosis, and the allocation of public resources, the fairness of algorithmic decision-making is particularly important [4]. Algorithmic fairness aims to ensure equitable treatment of individuals or groups in automated decision-making processes and to prevent unfair outcomes resulting from sensitive attributes such as race or gender. It can be broadly categorized into individual fairness and group fairness: individual fairness requires that similar individuals receive similar treatment [5], while group fairness requires comparable representation and treatment across protected groups [6]. For example, in credit approval, group fairness requires that approval rates do not differ significantly across applicants of different genders or races, whereas individual fairness requires that applicants with similar credit profiles receive similar decisions [7].

Fair spectral clustering has been extensively studied as a practical approach to enforcing algorithmic fairness and has demonstrated significant empirical success in recent years [8], [9]. By incorporating fairness constraints into the spectral embedding process, fair spectral clustering can regulate the balance of sensitive-group proportions in the learned low-dimensional representation while preserving the advantages of spectral methods for graph-structured and nonlinearly separable data [10]. This renders spectral clustering a natural and effective framework for fairness-aware clustering, where fairness can be promoted directly during the embedding stage rather than solely after cluster formation.

However, most existing fair spectral clustering methods are designed for a single sensitive attribute only, which limits their applicability in real-world settings where multiple sensitive factors (e.g., gender, race, age) simultaneously influence decisions [11], [12]. When multiple sensitive attributes coexist, enforcing fairness with respect to only one attribute may still result in substantial unfairness regarding others, while naive combinations of multiple constraints may lead to excessive computational demands. These challenges highlight the design for fair spectral clustering methods that can address multiple sensitive attributes.

Therefore, to address this gap, this paper introduces MFSC, a multi-sensitive-attribute fair spectral clustering framework that incorporates Shapley value theory to enhance fairness when multiple sensitive attributes coexist. MFSC adheres to the standard spectral clustering relaxation approach: it learns a low-dimensional spectral embedding and subsequently applies k-means to obtain the final clusters. The key distinction is that MFSC explicitly measures, for each sensitive attribute, the extent to which the sensitive-group proportions within each cluster deviate from the global proportions, and then penalizes such deviations during embedding learning. Importantly, MFSC provides a method for balancing fairness across different sensitive attributes. Experiments conducted on multiple datasets demonstrate that MFSC can simultaneously enhance fairness across multiple sensitive attributes while maintaining clustering quality. The contribution of this work can be summarized as follows:

- MFSC is proposed in this paper, a multi-sensitive-attribute fair spectral clustering framework that incorporates Shapley value theory for simultaneous processing across multiple sensitive attributes.

- This paper conducts extensive experiments on four benchmark datasets with varying numbers of clusters. The experiments demonstrate that MFSC improves fairness with respect to multiple sensitive attributes while maintaining clustering quality comparable to standard spectral clustering and single-sensitive fair spectral clustering.
- To support Shapley-based in fair spectral clustering, this paper designs a comprehensive clustering evaluation criterion that jointly accounts for fairness and clustering quality. This provides a well-defined and comparable coalition value for Shapley computation and enables a joint evaluation of fairness and clustering quality.

II. RELATED WORK

Fair spectral clustering, which incorporates fairness into the spectral embedding, is reformulated as an optimization problem with an objective function and fairness constraints [3]. Given the widespread popularity of spectral clustering for partitioning graph data, Kleindessner et al. [8] studied a version of constrained spectral clustering that incorporates the fairness notion proposed by Chierichetti et al. [13]. According to this notion, a clustering is fair if every demographic group is approximately proportionally represented in each cluster. Kleindessner et al. developed variants of both normalized and unnormalized constrained spectral clustering and demonstrated that these variants yield fairer clusterings on both synthetic and real data.

While Kleindessner et al.'s algorithm is able to find the fairer clustering, it is compromised by high computational costs due to the algorithm's deflation. Wang et al. [9] subsequently proposed an improved algorithm, called SFairSC, which involves only sparse matrix-vector products and is able to fully exploit the sparsity of the fair spectral clustering model.

More recently, Yang et al. [12] proposed a framework of fair spectral clustering algorithms based on fair Laplacian matrices and instantiated three fair spectral clustering algorithms under three different definitions, which include group fairness and individual fairness, as well as another type of fairness problem that exists in spectral clustering, defined as scale fairness.

Despite this progress, existing fair spectral clustering methods are predominantly developed and evaluated under a single sensitive attribute, while fair spectral clustering algorithms for multiple sensitive attributes remain relatively underexplored.

III. METHODOLOGY

A. Single-Sensitive Attribute Fair Spectral Clustering

Ideal fairness is often difficult to achieve. This paper relaxes the fairness requirement by adding a regularization term that penalizes fairness deviations, while keeping a clear bound-based definition of fairness for interpretation and evaluation.

Definition 1 This paper introduces a nonnegative hyperparameter θ to describe an acceptable range of group-proportion deviation. If the distribution of sensitive groups in each cluster lies within these bounds, then the clustering is fair:

$$\left|\frac{|C_j \cap S_q|}{|C_j|} - \frac{|S_q|}{|S|}\right| \le \theta, \quad j \in [k],\ q \in [g],\ \theta \ge 0. \tag{1}$$

where C_j is the j-th cluster, and this paper considers a single sensitive attribute that takes g categorical values (protected groups). Let S_q denote the set of samples whose sensitive value equals the q-th group, and $|S| = n$ be the total number of samples. A smaller θ means a stricter fairness requirement.

Normalized indicator and linear fairness deviation. To express the fairness deviation as a linear form in the normalized cluster-indicator matrix (consistent with the standard RatioCut relaxation), this paper multiplies both sides of (1) by $\sqrt{|C_j|}$, yielding

$$\left|\frac{|C_j \cap S_q|}{\sqrt{|C_j|}} - \frac{|S_q|}{|S|}\sqrt{|C_j|}\right| \le \theta\sqrt{|C_j|}, \quad j \in [k],\ q \in [g]. \tag{2}$$

For $q \in [g]$, define the sensitive-group indicator vector $t_q \in \mathbb{R}^n$ by $(t_q)_i = t_{iq}$, where

$$t_{iq} = \begin{cases} 1 & \text{if } i \in S_q, \\ 0 & \text{otherwise.} \end{cases} \tag{3}$$

For $j \in [k]$, define the normalized cluster-indicator matrix $H = [h_{ij}] \in \mathbb{R}^{n\times k}$ in the discrete setting by

$$h_{ij} = \begin{cases} \frac{1}{\sqrt{|C_j|}} & \text{if } i \in C_j, \\ 0 & \text{otherwise.} \end{cases} \tag{4}$$

Let $\mathbf{1}_n$ be the all-ones vector and define $p_q = \frac{|S_q|}{|S|}$. Construct the constraint matrix $F = [f_1, \ldots, f_g] \in \mathbb{R}^{n\times g}$ with

$$f_q = t_q - p_q \mathbf{1}_n. \tag{5}$$

Then

$$(F^T H)_{qj} = \sum_{i=1}^{n} (t_{iq} - p_q)\, h_{ij} = \frac{|C_j \cap S_q|}{\sqrt{|C_j|}} - p_q\sqrt{|C_j|},$$

which is exactly the fairness deviation appearing in (2).

In the discrete setting, H is determined by the partition $\{C_j\}$. Following the standard spectral relaxation for RatioCut, this paper relaxes H to a real matrix satisfying $H^T H = I_k$. Directly enforcing the element-wise bound constraints in (1) on the relaxed variable H is inconvenient for optimization. Therefore, this paper keeps (1) as an interpretable fairness criterion for evaluation (with tolerance θ), while optimizing a soft penalty (controlled by λ) that encourages small fairness deviations:

$$\begin{aligned} \min_{H} \quad & \operatorname{tr}(H^T L H) + \lambda \|F^T H\|_F^2, \\ \text{s.t.} \quad & H^T H = I_k, \end{aligned} \tag{6}$$

where $L = D - A$ is the (unnormalized) graph Laplacian of a symmetric adjacency matrix A, $\lambda \ge 0$ controls the fairness–utility trade-off, and $\|\cdot\|_F$ is the Frobenius norm.

Since $\|F^T H\|_F^2 = \operatorname{tr}(H^T F F^T H)$, (6) is equivalent to

$$\min_{H:\, H^T H = I_k} \operatorname{tr}\big(H^T (L + \lambda F F^T) H\big). \tag{7}$$

Thus, H^* is obtained by taking the eigenvectors corresponding to the k smallest eigenvalues of

$$L' = L + \lambda F F^T.$$

Algorithm 1 Single-Sensitive Attribute Fair Spectral Clustering

Require: Adjacency matrix A (symmetric); number of clusters k; sensitive groups $\{S_q\}_{q=1}^{g}$; regularization λ

Ensure: Spectral embedding H^* and cluster assignments

1: Compute degree matrix D and Laplacian $L = D - A$.
2: Construct $F = [f_1, \ldots, f_g]$ via (5).
3: Form $L' = L + \lambda F F^T$.
4: Let H^* be the eigenvectors of L' corresponding to the k smallest eigenvalues.
5: Apply k-means on rows of H^* to obtain final clusters.

B. Multi-Sensitive Attributes Fair Spectral Clustering

Given multiple sensitive attributes, it is necessary to balance clustering quality with fairness across attributes. Assume there are m sensitive attributes, indexed by $\mathcal{M} = [m] = \{1, \ldots, m\}$. For sensitive attribute SA_i with g_i categorical values (protected groups), let $S_{i,q} \subseteq S$ denote the set of samples whose SA_i value equals the q-th group, and let $p_{i,q} = |S_{i,q}|/|S|$. Define the group-indicator vector $t_{i,q} \in \mathbb{R}^n$ by $(t_{i,q})_\ell = 1$ if sample $\ell \in S_{i,q}$ and 0 otherwise. We then construct the corresponding fairness constraint matrix $F_i = [f_{i,1}, \ldots, f_{i,g_i}] \in \mathbb{R}^{n \times g_i}$ with $f_{i,q} = t_{i,q} - p_{i,q}\mathbf{1}_n$, which is analogous to (5).

Instead of multiple hard equality constraints or scaling hard constraints by weights, this paper adopts a penalty-based formulation in which attribute weights directly modulate the penalty strength:

$$\begin{aligned} \min_{H} \quad & \operatorname{tr}(H^T L H) + \lambda \sum_{i=1}^{m} \alpha_i \left\| F_i^T H \right\|_F^2, \\ \text{s.t.} \quad & H^T H = I_k, \end{aligned} \tag{8}$$

where $\lambda \geq 0$ controls the overall fairness–utility trade-off and $\alpha_i \geq 0$ is the importance weight for attribute SA_i.

This paper uses Shapley values to quantify the marginal contribution of each sensitive attribute to the overall fairness–utility behavior. For any coalition $\mathcal{S} \subseteq \mathcal{M}$, define the coalition embedding $H_{\mathcal{S}}^*$ as the minimizer of

$$\min_{H:\, H^T H = I_k} \operatorname{tr}(H^T L H) \;+\; \lambda \sum_{i \in \mathcal{S}} \left\| F_i^T H \right\|_F^2. \tag{9}$$

This admits a closed-form spectral solution: $H_{\mathcal{S}}^* \leftarrow$ the eigenvectors corresponding to the k smallest eigenvalues of $L + \lambda \sum_{i \in \mathcal{S}} F_i F_i^T$. This paper then applies k-means to the rows of $H_{\mathcal{S}}^*$ to obtain discrete clusters.

Definition 2 (Comprehensive Clustering Index, CCI). To compute Shapley values, we need to define a coalition value function $v(\mathcal{S})$ for each subset of sensitive attributes. For interpretability, we construct $v(\mathcal{S})$ so that larger values indicate better overall performance. To this end, we design a Comprehensive Clustering Index that combines a fairness score and a clustering-quality score. Let $a := AED(\mathcal{S})$ and $b := \text{Silhouette}(\mathcal{S})$ denote the coalition-level fairness and clustering-quality values, respectively.

Normalize AED:

$$A_{\text{norm}} = \frac{a - a_{\min}}{a_{\max} - a_{\min}}, \quad A_{\text{norm}} \in [0, 1],$$

and convert it to a fairness score (larger is better):

$$A_{\text{fair}} = 1 - A_{\text{norm}}.$$

Normalize Silhouette:

$$B_{\text{norm}} = \frac{b - b_{\min}}{b_{\max} - b_{\min}}, \quad B_{\text{norm}} \in [0, 1].$$

Then

$$CCI = w_A \cdot A_{\text{fair}} + w_B \cdot B_{\text{norm}}, \quad w_A + w_B = 1.$$

Definition 3 (Fairness Performance Metric). This paper uses the Average Euclidean Distance (AED) [14] as fairness metric. Consider one sensitive attribute SA with l protected groups. Let $S_{SA,r} \subseteq S$ denote the set of samples whose SA value equals the r-th group, and let $p_r = \frac{|S_{SA,r}|}{|S|}$ with $\sum_{r=1}^{l} p_r = 1$. Let $P_{SA} = [p_1, \ldots, p_l]^T$ denote the global group distribution of the dataset. For the j-th cluster C_j, let $C_{SA}^{(j)} = [c_{j1}, \ldots, c_{jl}]^T$ denote the group distribution within this cluster, where $c_{jr} = \frac{|C_j \cap S_{SA,r}|}{|C_j|}$ and $\sum_{r=1}^{l} c_{jr} = 1$. The Euclidean distance for cluster j is defined as:

$$ED_j = \left\| C_{SA}^{(j)} - P_{SA} \right\|_2. \tag{10}$$

The AED is the average distance over all clusters:

$$AED = \frac{1}{k} \sum_{j=1}^{k} ED_j. \tag{11}$$

A smaller AED indicates better fairness. Applying this definition to sensitive attribute SA_i yields $AED_i(\cdot)$.

For a non-empty coalition $\mathcal{S}$, this paper defines the coalition-level fairness as the average AED over attributes in $\mathcal{S}$:

$$AED(\mathcal{S}) = \frac{1}{|\mathcal{S}|} \sum_{i \in \mathcal{S}} AED_i(H_{\mathcal{S}}^*), \quad |\mathcal{S}| \geq 1. \tag{12}$$

where $AED_i(\cdot)$ is computed on attribute SA_i (Definition 3). For $\mathcal{S} = \emptyset$, we use vanilla spectral clustering ($\lambda = 0$) to obtain both $AED(\emptyset)$ and the Silhouette score, and then compute $CCI(\emptyset)$. Finally,

$$v(\mathcal{S}) := CCI(\mathcal{S}). \tag{13}$$

The Shapley value of attribute i is

$$\phi_i \;=\; \sum_{\mathcal{S} \subseteq \mathcal{M} \setminus \{i\}} \frac{|\mathcal{S}|!\,(m - |\mathcal{S}| - 1)!}{m!} \Big(v(\mathcal{S} \cup \{i\}) - v(\mathcal{S}) \Big). \tag{14}$$

This paper converts $\{\phi_i\}$ into nonnegative penalty weights. Let $Z = \sum_{r=1}^{m} \max(\phi_r, 0)$. If $Z > 0$, set

$$\alpha_i = \frac{\max(\phi_i, 0)}{Z}. \tag{15}$$

Otherwise (all $\phi_i \leq 0$), this paper uses a uniform fallback $\alpha_i = 1/m$. This step ensures that attributes with larger positive

marginal contributions receive stronger penalty strength in (8), which changes the optimizer.

Define a weighted constraint matrix

$$F_w = \left[\sqrt{\alpha_1}F_1,\ \sqrt{\alpha_2}F_2,\ \ldots,\ \sqrt{\alpha_m}F_m\right]. \tag{16}$$

Then $\sum_{i=1}^{m} \alpha_i \|F_i^T H\|_F^2 = \|F_w^T H\|_F^2$, and (8) is equivalent to

$$\min_{H:\, H^T H = I_k} \operatorname{tr}\left(H^T (L + \lambda F_w F_w^T) H\right). \tag{17}$$

Thus this paper obtains H^* by taking the eigenvectors corresponding to the k smallest eigenvalues of $L' = L + \lambda F_w F_w^T$, followed by k-means on the rows of H^*.

Equation (14) requires evaluating $v(\mathcal{S})$ for coalitions $\mathcal{S} \subseteq \mathcal{M}$, which is $O(2^m)$ coalition solves in the worst case. In our experiments m is small and enumeration is feasible. For larger m, $\{\phi_i\}$ can be approximated via Monte-Carlo Shapley sampling by randomly permuting attributes and averaging marginal gains.

For a fixed dataset and fixed k, the normalization ranges $a_{\min}, a_{\max}$ (AED) and $b_{\min}, b_{\max}$ (Silhouette) are computed over the set of coalitions that are evaluated, so that $CCI(\mathcal{S})$ is comparable across different coalitions.

Algorithm 2 Multi-sensitive Attribute Fair Spectral Clustering

Require: Adjacency matrix A (symmetric); number of clusters k; sensitive attributes $\{SA_1, \ldots, SA_m\}$; regularization λ

Ensure: Spectral embedding H^* and cluster assignments

1: Compute Laplacian $L = D - A$.
2: For each sensitive attribute SA_i, construct the constraint matrix F_i.
3: For coalitions $\mathcal{S} \subseteq \mathcal{M}$ (or sampled coalitions), solve (9), run k-means, compute $CCI(\mathcal{S})$, and obtain Shapley values $\{\phi_i\}_{i=1}^m$ via (14).
4: Normalize $\{\phi_i\}$ into weights $\{\alpha_i\}$ using (15) (with uniform fallback if needed).
5: Construct $F_w = [\sqrt{\alpha_1}F_1, \ldots, \sqrt{\alpha_m}F_m]$ and form $L' = L + \lambda F_w F_w^T$.
6: Let H^* be the eigenvectors of L' corresponding to the k smallest eigenvalues.
7: Apply k-means on rows of H^* to obtain final clusters.

Definition 4 (Clustering Performance Metric). This paper uses the Silhouette Coefficient [15] to evaluate clustering quality in terms of compactness and separation. For a sample x_i, let $m(i)$ be the average distance from x_i to all other samples in its assigned cluster, and let $n(i)$ be the minimum (over all other clusters) of the average distance from x_i to that cluster. The silhouette value of x_i is:

$$sc(i) = \frac{n(i) - m(i)}{\max\{m(i), n(i)\}}. \tag{18}$$

The overall Silhouette Coefficient is the mean of $sc(i)$ over all samples. Larger values indicate better clustering quality.

IV. EVALUATION

A. Datasets and Experimental Design

Dataset: This paper summarize the datasets and sensitive attributes used for evaluation in Table I. The Adult dataset [16] predicts if income exceeds $50K/year using 600 census records with various demographic and employment attributes. The Athlete dataset [17] covers 120 years of Olympic history, detailing 1,000 athletes' demographics, events, and performance. The Obesity dataset [18] has 2,111 records from Mexico, Peru, and Colombia, with 17 attributes related to eating habits and physical condition, classified into six obesity levels. The Census1990 dataset [19] covers diverse demographic and socioeconomic attributes.

TABLE I
DATASETS USED FOR EVALUATIONS.

Dataset	Samples	Sensitive attributes
Adult	48,842	*sex*, *income*, *race*
Athlete	134,732	*medal*, *season*, *sex*
Obesity	2,111	*gender*, *smoke*, *family*
Census1990	2,458,285	*isex*, *ienglish*, *icitizen*

Baseline: Specifically, this paper select the classic spectral clustering (SC) [20] methods as one of the baselines. Since there is limited research on multi-sensitivity attribute fair spectral clustering methods, which mainly focuses on single sensitive attribute fair spectral clustering methods, this paper only compare them with classic fair spectral clustering algorithms. FairSC [8] and SFairSC [9] are used for comparisons.

Implementation Details: For each dataset with three sensitive attributes, this paper compare MFSC against SC and two single-sensitive fair spectral clustering baselines (FairSC and SFairSC). Since the baselines handle only one sensitive attribute at a time, this paper run them separately for each attribute. MFSC is run once with all sensitive attributes jointly optimized. This paper set the number of clusters from 5 to 10 in each experiment to explore MFSC's performance under different granularities, observing its stability and effectiveness. Each experiment was repeated 50 times, and the final results were averaged to eliminate random biases.

B. Analysis

*1) **Fairness Evaluation**:* As shown in Table II, compared with single-sensitive fair spectral clustering methods (FairSC and SFairSC), which optimize fairness with respect to only one sensitive attribute at a time, MFSC incorporates fairness constraints for multiple sensitive attributes within a single optimization and adaptively allocates the strength of fairness regularization across attributes via Shapley-based weights. Although MFSC needs to simultaneously ensure fairness over multiple attributes, the experimental results demonstrate that it can still achieve the best or second-best performance in several key scenarios. For example, on the *race* attribute of the Adult dataset, MFSC consistently attains the best or second-best results for $k \in \{5, \ldots, 10\}$. Moreover, when $k = 7$, MFSC

TABLE II

COMPARISONS OF MFSC WITH BOTH SPECTRAL CLUSTERING AND FAIR SPECTRAL CLUSTERING METHODS ON FOUR BENCHMARKS. THE BEST AND THE SECOND-BEST RESULTS ARE MARKED IN BOLD AND UNDERLINE, RESPECTIVELY

Dataset	k	SC	FairSC	SFairSC	MFSC	SC	FairSC	SFairSC	MFSC	SC	FairSC	SFairSC	MFSC
		sex				*income*				*race*			
Adult	5	49.84	48.66	48.53	**47.60**	36.43	35.48	35.45	**33.57**	16.47	15.41	15.41	**15.41**
	6	43.82	41.94	**41.52**	41.80	41.37	**36.05**	39.59	39.59	17.32	16.07	16.07	**16.07**
	7	43.62	42.74	**42.36**	42.71	39.18	**31.05**	36.28	37.37	32.00	16.52	17.67	**14.23**
	8	40.63	43.38	**38.73**	40.16	37.97	37.30	**36.38**	43.56	28.90	16.85	17.92	**14.64**
	9	50.00	41.03	**38.17**	41.46	42.76	**37.12**	41.35	46.73	27.82	15.20	17.11	**15.15**
	10	50.18	**40.59**	41.16	46.86	45.70	**34.99**	40.89	44.88	26.96	**15.60**	17.14	16.26
		medal				*season*				*sex*			
Athlete	5	33.78	**9.02**	12.10	24.48	52.19	46.61	**37.23**	46.61	27.42	25.32	25.30	**25.06**
	6	32.26	**10.63**	11.81	12.27	53.14	51.54	**35.42**	50.88	26.73	**26.27**	26.30	**26.27**
	7	28.77	11.08	**11.02**	11.37	47.41	47.40	47.39	**34.75**	25.49	25.18	**22.88**	25.20
	8	26.71	11.96	**10.17**	11.68	44.82	44.80	44.83	**33.69**	33.84	25.97	32.71	**25.60**
	9	25.28	12.00	**9.31**	15.17	43.82	42.74	**41.55**	42.74	34.14	**26.24**	32.37	**26.24**
	10	23.98	12.17	**9.86**	10.93	41.10	40.09	38.70	**31.62**	36.14	**26.74**	32.01	32.56
		gender				*smoke*				*family*			
Obesity	5	57.37	47.32	**36.40**	44.25	29.51	20.47	**1.81**	2.25	48.12	**38.72**	45.00	41.08
	6	50.27	47.96	**32.47**	49.98	26.84	24.77	**2.00**	24.79	42.74	**36.59**	41.60	38.71
	7	52.97	51.01	**46.76**	46.79	41.02	21.41	**1.99**	21.56	39.73	**32.52**	37.62	36.95
	8	49.44	53.27	**39.46**	49.76	36.01	19.20	**4.46**	19.26	43.92	**31.70**	37.81	38.02
	9	51.78	51.13	**40.71**	52.01	32.46	32.35	**2.62**	17.55	36.58	**28.47**	33.31	34.64
	10	53.63	53.59	**44.82**	48.16	29.63	29.55	**3.11**	16.04	37.98	**28.78**	33.60	33.33
		isex				*ienglish*				*icitizen*			
Census1990	5	31.18	31.17	**6.63**	31.15	10.78	9.22	**8.90**	10.58	8.92	8.69	**5.01**	8.69
	6	42.04	39.05	**6.54**	28.46	31.03	11.85	**10.76**	12.03	29.36	10.34	**7.19**	10.34
	7	34.62	34.61	**7.28**	24.00	28.44	**11.78**	11.89	12.24	26.22	**8.36**	8.57	9.27
	8	38.26	31.19	**22.39**	31.00	26.87	12.87	**11.47**	12.60	26.92	24.19	**11.12**	24.13
	9	35.83	34.82	**25.78**	28.30	38.17	25.63	**12.16**	26.06	36.47	35.62	**11.69**	22.94
	10	40.61	32.66	30.77	**30.66**	36.06	24.29	14.48	**14.00**	33.39	33.33	**11.62**	13.55

achieves an AED of 14.23, improving the second-best single-sensitive baseline (FairSC at 16.52) by 13.9%; when $k = 8$, it also improvs FairSC (16.85) by approximately 13.1%. Similarly, on the *season* attribute of the Athlete dataset, MFSC achieves the best result under multi-sensitive constraints (e.g., AED$= 31.62$ at $k = 10$), outperforming the strongest single-sensitive method SFairSC (38.70) by 18.3%. These results indicate that MFSC not only enhances fairness across multiple sensitive attributes simultaneously but also maintains—and in some cases surpasses—the best single-attribute fairness performance, thereby demonstrating the advantages of multi-sensitive fairness constraints.

*2) **Clustering Performance Evaluation**:* As shown in Fig. 1, MFSC remains competitive with (and in several cases improves upon) SC, despite simultaneously fairness constraints for multiple sensitive attributes. On the *Adult* dataset, MFSC consistently yields higher Silhouette values than SC for $k \in \{6, \ldots, 9\}$. For example, at $k = 7$, MFSC improves from 0.308 (SC) to 0.395 (improvement 28.2%). Averaged over $k = 5$ to 10, MFSC improves the mean Silhouette score from 0.356 (SC) to 0.383 (improvement 7.6%). On the *Athlete* dataset, MFSC attains the best result at $k = 6$ (0.360), but for larger k values (e.g., $k \geq 8$), SFairSC tends to dominate in Silhouette score. This suggests that preserving clustering quality becomes more challenging when multiple fairness constraints are enforced on this dataset. On the *Obesity* dataset, MFSC achieves strong Silhouette scores across all k values. Notably, at $k = 5$, SC and FairSC produce much lower Silhouette values (0.135), whereas MFSC reaches 0.405 (approximately 200% relative to SC), indicating substantially better compactness and separation under the learned partition. On the *Census1990* dataset, MFSC is comparable to SC for smaller k values.

In summary, these results demonstrate that MFSC improves fairness while maintaining clustering quality. Across multiple datasets and k values, it matches or surpasses SC and remains competitive with the strongest single-sensitive baselines in terms of Silhouette score, while optimizing fairness constraints

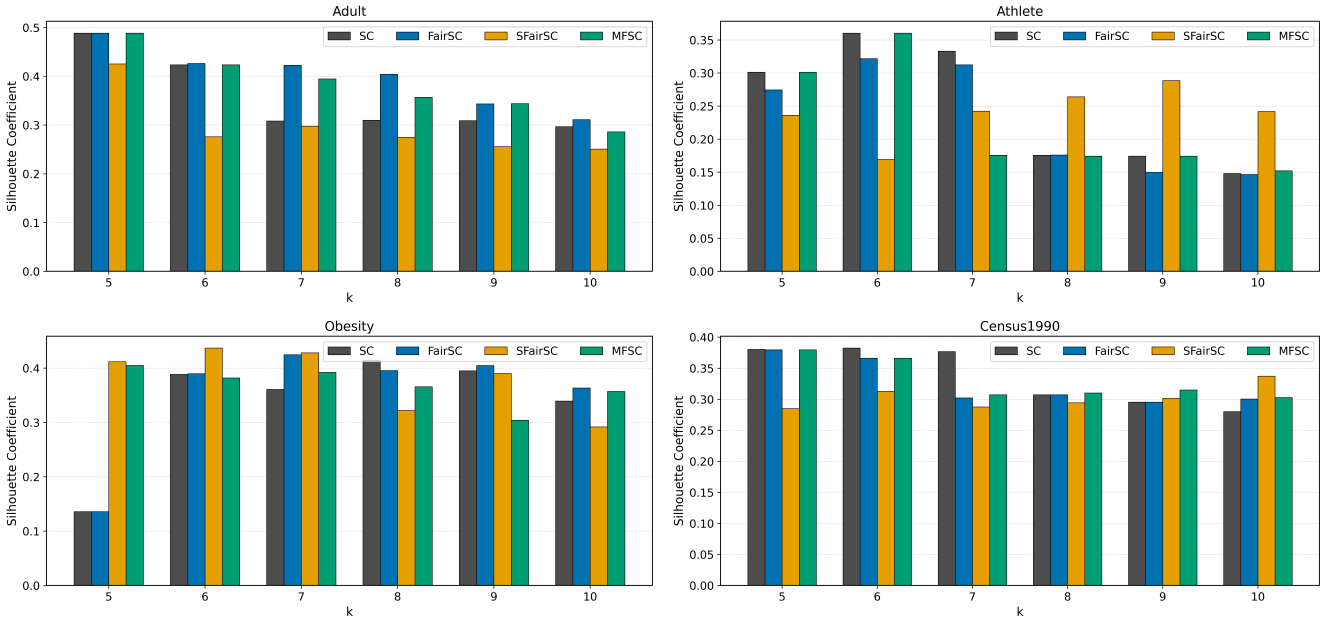

Fig. 1. Comparison of Silhouette Coefficients on four datasets. The X-axis represents the number of clusters, while the Y-axis measures an important indicator of clustering quality—the Silhouette Coefficient. A higher Silhouette Coefficient generally indicates better clustering performance. To clearly distinguish between different clustering algorithms, we use four distinct solid colors.

for multiple sensitive attributes in a single run.

V. CONCLUSIONS

To the best of our knowledge, this paper is the first to propose MFSC, a multi-sensitive-attribute fair spectral clustering method that optimizes fairness constraints across several sensitive attributes in a single run. Extensive experiments on multiple benchmarks and a range of cluster numbers k demonstrate that MFSC improves fairness while preserving clustering quality. Moreover, MFSC remains competitive in terms of clustering quality compared with single-sensitive fair spectral clustering baselines, while providing the advantage of simultaneously ensuring fairness across multiple sensitive attributes.

VI. ACKNOWLEDGMENT

This paper is supported by the Major Project of Sichuan Provincial Natural Science Foundation (2025ZNSFSC0005), the Sichuan Science and Technology Program (2024YFFK0120), the Central Government Guides Local Projects of China (2025ZYDF105) and the National Science Foundation of China (62072419).

REFERENCES

[1] K. Berahmand, M. Mohammadi, A. Faroughi, and R. P. Mohammadiani, "A novel method of spectral clustering in attributed networks by constructing parameter-free affinity matrix," *Cluster Computing*, vol. 25, no. 2, pp. 869–888, 2022.

[2] J. Shi and J. Malik, "Normalized cuts and image segmentation," *IEEE Transactions on pattern analysis and machine intelligence*, vol. 22, no. 8, pp. 888–905, 2000.

[3] U. Von Luxburg, "A tutorial on spectral clustering," *Statistics and computing*, vol. 17, no. 4, pp. 395–416, 2007.

[4] A. Chhabra, K. Masalkovaitė, and P. Mohapatra, "An overview of fairness in clustering," *IEEE Access*, vol. 9, pp. 130 698–130 720, 2021.

[5] S. Mahabadi and A. Vakilian, "Individual fairness for k-clustering," in *International conference on machine learning*. PMLR, 2020, pp. 6586–6596.

[6] X. Chen, B. Fain, L. Lyu, and K. Munagala, "Proportionally fair clustering," in *International conference on machine learning*. PMLR, 2019, pp. 1032–1041.

[7] B. Yan, Z. Yang, J. Zheng, Y. Tang, and H. Zhang, "Individual fair fuzzy c-means clustering via density-adaptive spectral regularization," *Neurocomputing*, vol. 651, p. 130794, 2025.

[8] M. Kleindessner, S. Samadi, P. Awasthi, and J. Morgenstern, "Guarantees for spectral clustering with fairness constraints," in *International conference on machine learning*. PMLR, 2019, pp. 3458–3467.

[9] J. Wang, D. Lu, I. Davidson, and Z. Bai, "Scalable spectral clustering with group fairness constraints," in *International conference on artificial intelligence and statistics*. PMLR, 2023, pp. 6613–6629.

[10] Z. Yang, H. Zhang, C. Yang, B. Li, X. Zhao, and Y. Long, "Spectral clustering with scale fairness constraints," *Knowledge and Information Systems*, vol. 67, no. 1, pp. 273–300, 2025.

[11] S. Bera, D. Chakrabarty, N. Flores, and M. Negahbani, "Fair algorithms for clustering," *Advances in Neural Information Processing Systems*, vol. 32, 2019.

[12] Z. Yang, H. Zhang, C. Yang, B. Li, X. Zhao, and Y. Long, "Fair laplace: A unified framework for fair spectral clustering," *Information Processing & Management*, vol. 62, no. 4, p. 104124, 2025.

[13] F. Chierichetti, R. Kumar, S. Lattanzi, and S. Vassilvitskii, "Fair clustering through fairlets," *Advances in neural information processing systems*, vol. 30, 2017.

[14] S. S. Abraham, S. S. Sundaram *et al.*, "Fairness in clustering with multiple sensitive attributes," *arXiv preprint arXiv:1910.05113*, 2019.

[15] A. M. Bagirov, R. M. Aliguliyev, and N. Sultanova, "Finding compact and well-separated clusters: Clustering using silhouette coefficients," *Pattern Recognition*, vol. 135, p. 109144, 2023.

[16] B. Becker and R. Kohavi, "Adult," UCI Machine Learning Repository, 1996, DOI: https://doi.org/10.24432/C5XW20.

[17] rgriffin, "120 years of olympic history: athletes and results," Kaggle, 2018, url: https://www.kaggle.com/datasets/heesoo37/120-years-of-olympic-history-athletes-and-results/data.

[18] "Estimation of Obesity Levels Based On Eating Habits and Physical Condition ," UCI Machine Learning Repository, 2019, DOI: https://doi.org/10.24432/C5H31Z.

[19] T. B. Meek, Chris and D. Heckerman, "US Census Data (1990)," UCI Machine Learning Repository, 2001, DOI: https://doi.org/10.24432/C5VP42.

[20] A. Ng, M. Jordan, and Y. Weiss, "On spectral clustering: Analysis and an algorithm," *Advances in neural information processing systems*, vol. 14, 2001.