Weighted Laplacian Method and Its Theoretical Applications

Shijie Xu¹, Jiayan Fang¹ and Xiangyang Li¹,*

¹School of Computer Science and Technology, University of Science and Technology of China, Hefei, China

*Corresponding author e-mail: xiangyangli@ustc.edu.cn

Abstract. Recently, a class of multilevel graph clustering (graph partitioning) algorithms have been extensively studied due to their practical utility. Although these newly proposed algorithms work well, there is still not many theoretical guarantees for the quality of partition in these existing methods due to their intrinsic heuristic properties to a great extent. In this paper, we propose a novel weighted Laplacian method for the multilevel graph clustering problem with more powerful theoretical background to further improve clustering performance mainly in terms of accuracy. Since our algorithm inherits the virtues of spectral methods, it possesses a friendly optimization property since it can produce the global optimal solution of a relaxation to the weighted cut on the coarsened graph in the middle stage. Meanwhile, the multilevel strategy can make it possible for our method to produce final clustering results with reasonable time range. Additionally, the weighted graph Laplacian is also suitable for doubly-weighted graph, which will endow our algorithm with a potential wide range of applications. The experimental results verify that our weighted Laplacian methods is indeed superior over existing algorithms in terms of clustering accuracy while also maintains comparable clustering speed.

1. Introduction

In graph theory, the traditional weighted graph consists only of weights on its edges. Although weighting edges has several practical uses such as scientific simulation, social networks and integrated circuit design, weighting vertices can also serve as an important role for many purposes, e.g. the Hosoya polynomial and Wiener index [34] of vertex-weighted graph has been well studied since its extensive applications in chemical graph theory. The doubly-weighted graph — a graph in which both edges and vertices are weighted — however, is very different from the above two versions of graph and proves to have vital potentials in solving some detailed issues which have not yet studied clearly in some practical problems.

Partial differential equations (PDEs) in graph analysis has been well studied over decades and has gradually developed into a powerful framework to help us understand various phenomena that might occur on graphs. One of particular success lies in the field of spectral graph theory, in which the Laplacian operator has been always playing a crucial role in dealing with some popular problems about graphs. Moreover, the Laplacian operator, acting as a core opera- tor for many second-order PDEs such as Laplace’s equations, Poisson’s equations, and evolution equations, has al- ready given birth to the emergence of various mighty techniques in the area of graph analysis. In some applications to graph processing, these equations have been discretized by resorting to several approximate methods including finite elements, finite volumes and Monte Carlo simulation [31]. Such kind of
Laplacian operators capable of coping with graph problems are uniformly termed as graph Laplacian. Some notions like p-Laplacian and \( \infty \)-Laplacian have been used to describe some important process in physics, biology, or economy [23]. Additionally, the nonlocal p-Laplacian has gained growing interest in various areas like mathematical biology, peridynamics, and image processing [11].

As a successful application to graph Laplacian, spectral graph theory has evolved to be a sort of powerful techniques in the topic of clustering and has been developed systematically over decades [7]. It was deemed to build originally from the minimum cut problem on the graph, and another motivation might come from the graph energy problem, in which the cut of the graph partition can be regarded as the energy of partition [18]. Since most applications to graph analysis are based on discretization methods, spectral methods based on graph Laplacian have been sufficiently exploited for minimizing some particular graph partitioning objective functions, such as the ratio cut [12], the cheeger cut [7] and the normalized cut [29], and can always return the global optimal solution to relaxation of these objective functions. These objective functions also fulfill to establish significant connections among various practical problems due to their strong mathematical background. For example, there exists several important equivalence relations between the total variation problem and the ratio cut problem [32], the weighted kernel k-means problem and the normalized cut problem [9], and the normalized cut problem and the modularity maximization problem based on spectral methods in community detection [22], respectively.

Considering that most existing graph Laplacians are built on the form of either edge-only-weighted graphs or vertex-only-weighted graphs [8][17], techniques developed from these graph Laplacians may suffer from limited applications when confronted with more complex situations. On account of vital potentials of doubly-weighted graphs in coping with some recently important unsolved problems [4][16], we propose the novel weighted Laplacian method inspired by existing theory of graph Laplacian to make it possible for the Laplacian operator suitable for doubly-weighted graphs, for the purpose of extending the practical utility of graph Laplacians. In order to further demonstrate advantages and wide potential use of our proposed strategy, we provide two theoretical applications of the weighted Laplacian method respectively to the design of clustering algorithms for the multilevel graph partitioning problem and the balanced minimum cut problem. The main contributions of this paper are listed as follows:

1. We create the notion of weighted Laplacian in the context of doubly-weighted graphs by resorting to the theory of partial differential equations on graphs, and further propose the weighted Laplacian method by the inspiration of existing spectral methods.

2. For the purpose of demonstrating the theory value of our idea, we give the first application of the weighted Laplacian method to the balanced minimum cut problem. We provide rigorous proof from the perspective of PDEs about the equivalence between the weighted cut problem and the balanced minimum cut problem both in their relaxed versions, thus revealing the possibility that the mature research of numerical solutions of PDEs will be of great assistance in recent studies of graph problems.

3. As a second application, we also show how to embed our weighted Laplacian method into the design of algorithms for multilevel graph partitioning problems. In order to illustrate the practical power of the proposed idea, we prove that the weighted cut problem is essentially equivalent to the initial clustering problem that appears in the second phase of those graph partitioning algorithms with multi-level structure. Furthermore, we post the description of the multilevel graph partitioning algorithm based on the proposed strategy.

The rest of the paper is organized as follows. Section 2 provides a brief overview on some related work including the balanced minimum cut problem and multilevel structure of the graph partitioning problem. In section 3, we give the definition of weighted Laplacian and its important properties, then the weighted Laplacian method is proposed later. Section 4 will mainly dwell on providing two theoretical applications of our weighted Laplacian method and the detailed practical algorithm of one application will be posted. Section 5 reports the performance of our algorithm. Finally, we conclude in section 6.
2. Related Work

Before diving into the details of our work, we first take a brief overview on some excellent research that are closely related to our method.

2.1. Graph Laplacian and Balanced Minimum Cut Problem

The theoretical foundations of spectral graph theory stem originally from the work of [13], and were further developed in decades. Graph Laplacians lie in the heart of major spectral methods. There are several different Laplacians based on edge-only-weighted graphs given in related literatures, e.g. the unnormalized Laplacian

\[ L := D - W \]

the normalized Laplacian

\[ L_N := I - D^{-1/2}WD^{-1/2} \]

and the random walk Laplacian

\[ L_{rw} := I - D^{-1}W \]

Where \( W = \{W_{ij}\} \) and \( D = \text{diag}\{d_i\} \) are the weight matrix and the degree matrix of the graph respectively. Moreover, the vertex-only-weighted graph Laplacians have also been continually studied [16][30]. All of the practical Laplacians mentioned above, however, consider either the edge-only-weighted or vertex-only-weighted graphs. Considering that the minimum cut problem becomes increasingly important in the context of doubly-weighted graphs, to which most existing algorithms are unfortunately difficult to be extended [20], it is worthwhile to develop a new kind of graph Laplacian suitable for doubly-weighted graphs.

In recent years, the balanced minimum cut problem also constantly serves as an important role in various practical situations. There exists a number of different ways to define a series of balance conditions which act as a class of constraints in the balanced minimum cut optimization problem. Based on some particular definitions of the balance conditions, a balanced partition can be produced as a solution of the balanced minimum cut problem via some certain algorithm. See the work of [5][21] for further references. The notion of balanced minimum cut is actually ubiquitous in many graph problems for the reason that some graph partitioning objective functions in spectral graph theory can be essentially regarded as a sort of “balanced” minimum cut, in whose corresponding optimization problem, each cut term expressed in the ratio form respectively has a predefined balance condition as its denominator, such as the cardinality of each partition (ratio cut) or the volume of each partition (normalized cut). Besides, the definition of the perfectly balanced minimum cut can be found in [1].

2.2. Multilevel Graph Partitioning

Under the background of multilevel graph partitioning (also called V-cycle), a plenty of heuristics with different nature have recently been successively developed due to its practical significance [33] in order to de- sign efficient approximation algorithms for graph partitioning problems with reasonable computational time. However, not until the general-purpose multilevel methods were put forward, had not the field of graph partitioning undertaken a truly breakthrough in aspects of both efficiency and partition quality [26]. A specific multilevel graph partitioning algorithm consists of three phases: coarsening — where the problem instance is gradually mapped to a smaller one to reduce the original complexity, initial clustering — where the coarsening graph is partitioned by some specific clustering algorithm, and refining — where the partition for original graph is inversely refined from coarsened partitioning, as shown in figure 1.
Coarsening phase We denote $G_i$ as the $i$-th graph and $V_i$ as the corresponding vertex set of the $i$-th graph. Starting with one original graph $G_0$, a specific coarsening algorithm repeatedly transforms the graph into smaller and smaller graphs $G_1, G_2, \ldots, G_m$ such that $|V_0| > |V_1| > \ldots > |V_m|$. In order to obtain a coarser graph from $G_i$ to $G_{i+1}$, all nodes in $G_i$ will be partitioned by some rules into several groups, each of which can be regarded as a supernode in $G_{i+1}$. Several criteria for grouping the nodes are proposed in [9][14]. In our settings, when combining a set of nodes into a single supernode, the edge weight of the supernode is taken to be the sum of the edge weights of the original nodes that comprise this supernode. Similarly, the degree of a supernode is taken to be the sum of the degrees of all original nodes that are contained in it. See more details in [26].

Initial clustering phase Several initial clustering algorithms have been proposed and further developed, e.g. the region-growing algorithm [15][27][24], the recursive bisection algorithm [3][2], spectral clustering algorithm and weighted kernel algorithm [9][10]. However, all of the above initial clustering algorithms are heuristic, thus failing to provide qualitative guarantee for partitioning results of coarsened graphs. Then, a better initial method is immediately desired, considering that the coarsened graph partition produced by the initial clustering phase, as an intermediate result in the whole multilevel algorithmic framework, can certainly exert significant impacts on the subsequent refining phase in aspects of both time complexity and partition accuracy.

Refining phase As the final step of the multilevel framework, given a partition of graph $G_i$, the refining phase forms a finer partition of the graph $G_{i-1}$ where $G_i$ is a coarsened version of $G_{i-1}$. If there is a partition of $G_i$, which naturally yields a partition of $G_{i-1}$ by projecting, then we run the refining algorithm on the $G_{i-1}$ to get a finer partition. The Kernighan-Lin objective [19] is used to search the local minima by swapping points between different partitions. [27] also designed a local search algorithm to refine the coarsened partition based on negative cycle detection, where a negative cycle corresponds to a set of node movements that will not only decrease the overall cut but also maintain the balance of a partition as well. More popular refining algorithms can be found in [26].

3. Weighted Laplacian

3.1. Some Definitions and Lemmas

**Definition 1** (Doubly-weighted graph). Let $G = (V, M, W)$ be a connected undirected graph with doubly weight where $V = \{1, 2, \ldots, n\}$ is the vertex set of $G$, and vertex-weight matrix $M = \text{diag}\{m_1, m_2, \ldots, m_n\}$ s.t. $m_i > 0$ and edge-weight matrix $W = \{W_{ij}\}$. Let $D = \{d_1, d_2, \ldots, d_n\}$ be the degree matrix of the graph where $d_i = \sum_j W_{ij}$. We say $G = (V, M, W)$ is a doubly-weighted graph.
Definition 2 (Weighted cut). Suppose \( A, B \) are two disjoint subsets of \( V \), the cut of \( A, B \) on the graph \( G \) is 
\[
\text{Cut}(A, B) = \sum_{i \in A, j \in B} W_{ij}.
\]
Let \( \pi = \{C_1, C_2, ..., C_k\} \) be a k-partition of the graph \( G \), i.e. \( V = C_1 \cup C_2 \cup ... \cup C_k \) and \( C_i \cap C_j = \emptyset \) for any \( i \neq j \). We define the weighted cut of the partition \( \pi \) as follows
\[
W\text{cut}(\pi) := \sum_{i=1}^{k} \frac{\text{Cut}(C_i, C_j)}{\text{mvol}(C_i)}.
\]
where \( \text{mvol}(C_i) := \sum_{x \in C_i} m_x \). Besides, we recall the definition of the normalized cut as follows,
\[
N\text{cut}(\pi) := \sum_{i=1}^{k} \frac{\text{Cut}(C_i, C_j)}{\text{vol}(C_i)}.
\]
Where \( \text{vol}(C_i) := \sum_{x \in C_i} d_x \).

The weighted cut problem is an optimization problem that minimizes the above weighted cut for all partition on the doubly-weighted graph. In order to find the optimal solution of this minimization problem, we employ the theory of partial differential equations on graphs. We first give some related definitions and lemmas in the following, then naturally introduce our weighted Laplacian method and show how to apply it in the weighted cut problem.

Definition 3 (Weighted Laplacian). Suppose \( G = (V, M, W) \) is a doubly-weighted graph. Let \( \mathcal{G} \) be the linear space of all functions \( f : V \rightarrow \mathbb{R} \), we define the gradient of \( f \) as a vector
\[
\nabla f := ((f(y) - f(x)) \sqrt{\frac{W_{xy}}{m_x}})_{y \in V},
\]
and the weighted Laplacian \( \Delta \) is an operator in \( \mathcal{G} \) defined as
\[
\Delta f := \sum_{y \in V} (f(x) - f(y)) \frac{W_{xy}}{m_x}.
\]
The integral of \( f \) is defined as \( \int f := \sum_{x \in V} f(x) m_x \), and the inner product in \( \mathcal{G} \) is defined as \( \langle f, g \rangle := \int fg \) for all \( f, g \in \mathcal{G} \).

The following lemma gives an important property of the weighted Laplacian.

Lemma 1. \( \Delta \) is equivalent to the weighted Laplacian matrix
\[
L_{M} = M^{-1/2} (D - W) M^{-1/2} \in \mathbb{R}^{n \times n}.
\]

Proof. Consider a group of normalized orthogonal bases \( \{\delta_i / \sqrt{m_i}\} \) of \( \mathcal{G} \) where \( \delta_i(x) = 1 \) if \( x = i \) and zero otherwise. We have
\[
\Delta(i, j) = \langle \frac{\delta_i}{\sqrt{m_i}}, \Delta(\frac{\delta_j}{\sqrt{m_j}}) \rangle
\]
\[
= \frac{1}{\sqrt{m_i m_j}} (\delta_j(i) \sum_y W_{iy} - \sum_y \delta_j(y) W_{iy})
\]
\[
= \frac{1}{\sqrt{m_i m_j}} (\delta_j(i) d_i - W_{iy})
\]
\[
= L_{M}(i, j)
\]
therefore \( \Delta \) is equivalent to \( L_{M} \). Notice that when \( M = I \) or \( M = D \), the weighted Laplacian becomes the unnormalized Laplacian or the normalized Laplacian, respectively.

Lemma 2. Suppose \( 1_{C_i} \in \mathcal{G} \) be the indicating function of \( C_i \), i.e. \( 1_{C_i}(x) = 1 \) if \( x \in C_i \) and zero otherwise. We have
\[
\text{Cut}(C_i, \overline{C_i}) = \int |\nabla 1_{C_i}|^2, \quad \text{mvol}(C_i) = \int 1_{C_i}^2, \\
\langle f, \Delta f \rangle = \int |\nabla f|^2, \quad \langle f, f \rangle = \int |f|^2 \text{ for all } f \in \mathcal{G}.
\]

Therefore, we have the following equation
\[
\sum_{i=1}^{k} \text{Cut}(C_i, \overline{C_i}) = \sum_{i=1}^{k} \int |\nabla 1_{C_i}|^2 = \sum_{i=1}^{k} \langle 1_{C_i}, \Delta 1_{C_i} \rangle (1)
\]

**Proof.** The equations in this lemma are trivial to verify and we omit the proof here to save space. We also denote \( f_i = 1_{C_i} \) for convenience in later.

### 3.2. Weighted Laplacian Method

We are inspired by the spectral methods, which first relaxes the discreteness condition in the optimization problem, such as in the normalized cut and the ratio cut, then re-converts the real-valued solution into the partition of the graph. Accordingly, the weighted Laplacian method also firstly aims to solve the relaxed version of the weighted cut problem, and then its global optimal solution is obtained. We discuss the condition that is required to achieve its minimum values using the variational approach, based on the Rayleigh quotient shown below,
\[
\mathfrak{R}(f) = \frac{\langle f, \Delta f \rangle}{\langle f, f \rangle} \text{ for all } f \in \mathcal{G},
\]
and the weighted cut can be presented as a functional \( \mathcal{L} \) of indicating functions \( f_1, f_2, \ldots, f_k \) shown as following
\[
\mathcal{L}(f_1, f_2, \ldots, f_k) = \sum_{i=1}^{k} \mathfrak{R}(f_i)
\]

After relaxing the discreteness condition, we derive the relaxed version of the weighted cut problem. In order to minimize the functional \( \mathcal{L} \), by Euler-Lagrange equation with several functions, we can obtain that
\[
\frac{\partial \mathcal{L}}{\partial f_i} = \frac{\partial \mathfrak{R}(f_i)}{\partial f_i} = 0 \text{ for all } i = 1, 2, \ldots, k,
\]

Therefore, we derive the gradient of Rayleigh quotient. Briefly, we regarded the functional \( \mathcal{I}_f := g \mapsto \langle f, g \rangle \) derived by function \( f \) as function \( f \) itself. For all \( f \in \mathcal{G} \), we denote \( P = \langle f, \Delta f \rangle, Q = \langle f, f \rangle \), then we have
\[
\frac{\partial \mathfrak{R}(f)}{\partial f} = \frac{P'Q - PQ'}{Q^2}.
\]

Since
\[
P' = \frac{\partial \langle f, \Delta f \rangle}{\partial f} = 2\Delta f, Q' = \frac{\partial \langle f, f \rangle}{\partial f} = 2f,
\]
then we have
\[
\frac{\partial \mathfrak{R}(f)}{\partial f} = \frac{2\Delta f \cdot \langle f, f \rangle - 2\langle f, \Delta f \rangle \cdot f}{\langle f, f \rangle^2} \nonumber \\
= (\Delta - \frac{\langle f, \Delta f \rangle}{\langle f, f \rangle \cdot I_\beta}) \frac{2f}{\langle f, f \rangle} = 0
\]
where $I_d$ is the identity operator in $\mathcal{G}$, thus \( \frac{2f}{\langle f, f \rangle} \) is the eigenfunction of $\Delta$, with the corresponding eigenvalue $\langle f, \Delta f \rangle / \langle f, f \rangle$. Therefore, $f$ is the eigenfunction of $\Delta$.

Since the weighted Laplacian $\Delta$ is equivalent to the weighted Laplacian matrix $L_M$ (see lemma 1), we only need to compute the eigenvectors of the weighted Laplacian matrix. Moreover, since $\Delta$ (and $L_M$) is Hermitian and the eigenfunctions (and eigenvectors) are mutually orthogonal to each other, the first $k$ smallest eigenfunctions exactly satisfy the orthogonality of indicating functions, thus we are capable of solving the relaxation of the weighted cut problem. Finally, in the practical situation, we need to re-convert the real-valued solution into a partition of the graph. For example, the weighted spectral algorithm presented in next section.

4. Two Theoretical Applications

We first show the equivalence between relaxed balanced minimum cut problem and relaxed weighted cut problem, as well as the equivalence between weighted cut problem and initial clustering problem that arises in the middle stage of graph partitioning algorithms under a multilevel graph structure. The initial clustering algorithm based on weighted Laplacian method is proposed in later.

4.1. Equivalence Between Balanced Minimum Cut Problem and Weighted Cut Problem

As the special cases of weighted cut, existing objective functions like the ratio cut and the normalized cut, in graph partitioning problems have been proposed decades ago and their practical development has been so successful that various application areas rely heavily on them [28]. One motivation for constructing these objective functions is that the optimal solution of the traditional minimum cut problem may lead to an unnatural bias especially for partitioning out small sets of points [29]. Although a strong and natural relation between the weighted cut and the balanced minimum cut has not been revealed up till now, it is still meaningful in theory to provide a deep understanding for the weighted cut problem in variational language via establishing an equivalence relation from the PDEs point of view.

We first give the balance condition for the indicating function, shown as follows,

\[
\int 1_{C_i}^2 = \frac{1}{k} \quad \text{for all } i = 1, 2, ..., k, \quad (2)
\]

which means the integral average of $1_{C_i}^2$ is $\frac{1}{k}$ for each $i$. Since $1_{C_i}^2 = 1_{C_i}$, the partition $\{C_1, C_2, ..., C_k\}$ can be regarded as a balanced partition in a sense. By relaxing the discreteness condition, we consider the balanced minimum cut problem defined below

\[
\min_{f \in \mathcal{F}} \sum_{i=1}^{k} \langle f_i, \Delta f_i \rangle \quad \text{s.t.} \quad \langle kf_i - 1, 1 \rangle = 0. \quad (3)
\]

**Theorem 3.** The balanced minimum cut problem is equivalent to the weighted cut problem.

We define the functional $\mathcal{R}$ on $\mathcal{G}^k$ below

\[
\mathcal{R}(f_1, f_2, ..., f_k) = \sum_{i=1}^{k} \langle f_i, \Delta f_i \rangle,
\]

from the Euler-Lagrange equation with constraints, there exists a constant $\lambda_i$ for each $i$ such that

\[
\frac{\partial \mathcal{R}}{\partial f_i} + \frac{\partial}{\partial f_i} \sum_{i=1}^{k} \lambda_i \langle kf_i^2 - 1, 1 \rangle = 0, \quad (4)
\]

since we have
by equation 4, we know that $\Delta f_i = \lambda_i f_i$ holds for all $i = 1, 2, \ldots, k$ and for some constant $\lambda_i$. Furthermore, the first $k$ smallest eigenfunctions of $\Delta$ solve the balanced minimum cut problem in equation 3 for the reasons that solving the eigenfunctions of $\Delta$ can reduce to computing the eigenvectors of weighted Laplacian matrix $L_M$ and $L_M$ is a Hermitian matrix with its eigenvectors mutually orthogonal, thus exactly satisfying the orthogonality of the indicating functions. We have also presented the weighted Laplacian method in section 3, which is similarly based on solving the eigenfunctions of weighted Laplacian, thus proving that solving the balanced minimum cut problem is equivalent to solving the weighted cut problem.

A remaining question might be why we choose to define the balance condition as $\int 1_c^2 = \int \frac{1}{k}$ in equation 2 rather than the simpler form $\int 1_c = \int \frac{1}{k}$. The main reason is that the later one would produce Laplace’s equation $\Delta f = c$ for some constant $c$, with its trivial solution $f$ be a constant function when $c = 0$ [6], and cannot be re-converted into a meaningful partition of a graph. In a word, the existence of the equivalence relation, which has been proved in this subsection via the weighted Laplacian method, can also shed interesting light on the relation between the research of PDEs theory and recent studies of popular graph problems, thus setting up a bridge to connect these two seemingly unrelated fields.

4.2. Equivalence Between Initial Clustering and Problem and Weighted Cut Problem

The key step in the multilevel graph partitioning framework is to construct an initial partition on a coarsened graph obtained from the first coarsening step. As has been described in the related work, state-of-art graph partition algorithms with multilevel structure are totally heuristic, i.e., they fail to provide any qualitative guarantees for a partition result especially produced by the initial clustering phase. Moreover, although these existing algorithms using the multilevel strategy can indeed largely reduce time complexity of deriving a final partition, comparing with spectral clustering, they always lack in partition accuracy due to their intrinsic heuristic property that appears in the second stage. Additionally, considering that traditional spectral methods with the merit of high accuracy cannot be universally transplanted into the initial clustering phase, it is necessary to apply our weighted Laplacian method to the initial clustering phase to make a good compromise between time efficiency and partition accuracy. Meanwhile, it is also worthwhile to mention that the initial clustering algorithm designed with our proposed method possesses robust theoretical support for their partition result due to the inspiration by graph Laplacian theory, and spectral clustering naturally degenerates to a special case of our weighted Laplacian method. In this subsection, we exhibit a rigorous proof about the equivalence between the weighted cut problem and the initial clustering problem, which can set up a bridge to guide us to design an initial clustering algorithm based on weighted Laplacian method.

Firstly, we introduce the definition of the initial clustering problem. Suppose $\tilde{G}$ is a coarsened graph of $G$, i.e. the vertex set of $\tilde{G}$ is a set of supernodes $V_\tilde{G} = \{C_i\}_{i=1,\ldots,m}$ where $C_i = \{v^i_j\}_{j=1,\ldots,\tilde{n}}$ is a supernode. Therefore, we know $V_\tilde{G}$ forms a partition $\pi = \{C_i\}_{i=1,\ldots,m}$ of the original graph $G$. For any initial clustering $\tilde{\pi}$ of $\tilde{G}$ in the middle stage of multilevel graph partitioning algorithms, its corresponding original partition of $G$ can be recovered by projecting the cluster of a supernode into the nodes it contained, and this recovered partition can be denoted as $\pi'$. Since $\tilde{\pi}$ is an initial clustering on $\tilde{G}$, we define that $\pi'$ is a coarser partition of $\pi$, i.e. if for all $P \in \pi$, there exists $P' \in \pi'$ such that $P \subset P'$, then we denote it as $\pi' \leq \pi$. Thus the initial clustering problem is defined as following
\[
\min_{\pi \in \Pi} \text{Ncut}(\pi).
\]

Now we recall the weighted cut problem mentioned in Section 3. We regarded the coarsened graph \( \tilde{G} \) as a doubly-weighted graph defined in 1. We denote the coarsened graph as \( \tilde{G} = (\tilde{V}, \tilde{M}, \tilde{W}) \). The weighted cut problem on \( \tilde{G} \) is shown as follows
\[
\min_{\tilde{\pi}} \text{Wcut}(\tilde{\pi}).
\]

**Theorem 4.** The weighted cut problem is equivalent to the initial clustering problem.

**Equivalence** For all partitions \( \tilde{\pi} = \{B_1, B_2, ..., B_k\} \) on the coarsened graph \( \tilde{G} \) where \( B_i = \{C_i'\} \), we project the partition \( \tilde{\pi} \) into the original graph, denoted as \( \pi' \), i.e. \( \pi' = \{B'_i\} \) where \( B'_i = \bigcup_{t=1}^{l_i} C'_t \), then we can derive that
\[
\text{Ncut}(\pi') = \sum_{p=1}^{k} \frac{\text{Cut}(B'_p, B'_q)}{\text{vol}(B'_p)} = \sum_{p=1}^{k} \frac{\sum_{x \in B'_p, y \in B'_q} W_{xy}}{\sum_{x \in B'_p} d_x} = \sum_{p=1}^{k} \frac{\sum_{t=1}^{l_p} \sum_{s=1}^{l_q} W_{xy}}{\sum_{t=1}^{l_p} d_x} = \sum_{p=1}^{k} \frac{\sum_{t=1}^{l_p} \sum_{s=1}^{l_q} \text{Cut}(C'_p, C'_q)}{\sum_{t=1}^{l_p} \text{vol}(C'_p)} = \sum_{p=1}^{k} \frac{\text{Wcut}(\tilde{\pi})}{\text{mvol}(B'_p)} = \text{Wcut}(\pi').
\]

Hence we have proved that the initial clustering problem is equivalent to the weighted cut problem, which can be solved by the weighted Laplacian method.

**4.3. Weighted Spectral Algorithm**

We describe our proposed weighted spectral algorithm based on weighted Laplacian method via the above equivalence relation. Algorithm 1 takes as its inputs the original graph \( G \), a partition \( \pi \) and the number of clusters \( k \), then outputs the \( k \)-clustering result of initial clustering. Algorithm 1 gives the pseudo-code of our algorithm. Line 1-8 performs the summation operation on rows and columns of the matrix, which can be executed very efficiently in a parallel system. Another part of algorithm 1 (corresponds the line 9-14) produces an initial clustering of the coarsened graph with \( m \) vertices, based on weighted Laplacian methods. Like traditional spectral clustering, we compute the first \( k \) smallest eigenvectors of the weighted Laplacian matrix \( L_M \) to obtain a matrix \( U \in \mathbb{R}^{m \times k} \) which takes these eigenvectors as its columns, then run \( k \)-means algorithm on the rows of \( U \) to obtain a \( k \)-clustering result.
Algorithm 1 Weighted Spectral Algorithm for Initial Clustering Based on Weighted Laplacian Method

Require: \( G, \pi, k \)
Ensure: an initial clustering result \( \{C_i\}_{i=1,...,k} \)
1: \( (V, M, W) \leftarrow G \)
2: Compute the degree matrix \( D \) of \( W \)
3: for each element \( e_i \) in the partition \( \pi \) do
4: \( \{v_1, v_2, ..., v_n\} \leftarrow e_i \)
5: \( W(v_1, v_2) \leftarrow \sum_{j=1}^{n} W(v_1, v_j) \)
6: \( W(v_1, v_2) \leftarrow \sum_{j=1}^{n} W(v_j, v_2) \)
7: \( M(i, j) \leftarrow \sum_{e_i} D(v_1, v_2) \)
8: end for
9: Compute the degree matrix \( \hat{D} \) of \( W \)
10: \( M \leftarrow \hat{D}^{-1/2}(D - W)\hat{D}^{-1/2} \)
11: \( m \leftarrow \text{dim}(W) \)
12: Make the eigendecomposition \( L_{m \times m} = \lambda x \), the eigenvectors corresponding to the \( k \) smallest eigenvalues are \( x_{i} \) \( i=1,...,k \), let \( U := (x_{i}) \in \mathbb{R}^{m \times k} \)
13: Execute the \( k \)-means algorithm on the rows of \( U = (u_{j}) \) to obtain a clustering \( \{A_{j}\} \)
14: return \( \{C_i\}_{i=1,...,k} \)

5. Experiments

Figure 2 Clustering results on the real-world datasets (lower cut values are better). Our weighted spectral (wspe.) outperforms existing algorithms. Each result in the bar charts takes the average of 10 repeated experiments.

In this section, we present experimental results of the weighted spectral algorithm in our second application of the weighted Laplacian method to validate the practical power of the proposed algorithm for the initial clustering stage. Our experiment framework is modified based on KaHIP [27], which is an open-source framework written in C++ for multilevel graph partitioning.

In order to highlight the obvious advantage of our algorithm in terms of partition accuracy, we conduct several comparative experiments on different datasets. We select three other algorithms for comparison. Firstly, we choose region-growing (r.g.), which performs well than recursive bisection in practice, based on some precious reports [25]. Besides, we also choose spectral clustering (spec.) as an initial clustering algorithm to reflect the superior strength of our weighted Laplacian method (wspe.). Also, as an extremely case, we take as one of the comparative experiment the random initial clustering (rand.), in which every supernode is labeled with a random label ranging from 1 to \( k \), where \( k \) is the number of clusters. All experiments were compiled with g++ 7.4.0, each of which is repeated 10 times on a Linux machine with Intel (R) Xeon(R) Bronze 3104 CPU@1.70 GHz*12 and 32 Gbytes of main memory, to obtain more stable averaged partition results. We perform the partitioning tasks with different scales ranging from 8 to 512 on the dataset add32, and related results are reported in table 1, which indicates that our weighted spectral algorithm returns the lowest normalized cut value, thus always obtaining the highest partition accuracy. In addition, we run the partitioning tasks on different
real-world datasets (see table 2). Each result is taken as the average of the 10 repeated experiments and is plotted as bar charts shown in figure 2. The conclusion drawn above still holds.

**Table 1.** Normalized cut values with different initial clustering algorithms (lower cut values are better).

Our weighted spectral algorithm always obtains the best results on different scales of cluster tasks. Each result is the average of 10 repeated experiments.

| #clusters | 4  | 8  | 16 | 32 | 64 | 128 | 256 | 512 |
|-----------|----|----|----|----|----|-----|-----|-----|
| random (rand.) | .393 | 2.44 | 3.96 | 4.00 | 15.7 | 37.3 | 94.3 | 251 |
| region-growing (r.g.) | .162 | .07 | 5.90 | 7.00 | 18.1 | 34.0 | 85.0 | 186 |
| spectral (spec.) | .345 | 1.40 | 2.68 | 2.94 | 10.5 | 30.0 | 80.5 | 181 |
| weighted spectral (wspe.) | .158 | .954 | 1.17 | 1.69 | 8.70 | 22.6 | 70.0 | 177 |

**Table 2.** characteristics of test graphs

| No. | Graph name  | No. of nodes | No. of edges |
|-----|-------------|--------------|--------------|
| 1   | Add32       | 4960         | 9462         |
| 2   | finance256  | 37376        | 130560       |
| 3   | gupta2      | 62064        | 2093111      |
| 4   | memplus     | 17758        | 54196        |
| 5   | pcrystk02   | 13965        | 477309       |
| 6   | rajat10     | 30202        | 50101        |
| 7   | ramage02    | 16830        | 1424761      |

6. Conclusion

In this paper, we propose the weighted Laplacian method. Developed from the theory of graph Laplacian, our proposed method inherits some intrinsic merits of spectral methods, thus always providing more robust theoretical support for any graph algorithm equipped with the idea of weighted Laplacian. In order to further demonstrate both the theory value and the practical significance of the weighted Laplacian method, we introduce two theoretical applications of this method respectively to the balanced minimum cut problem and the multilevel graph partitioning problem. By resorting to variational methods and the theory of PDEs on graphs, we successfully establish the equivalence between the weighted cut problem and the balanced minimum cut problem in their relaxed versions, which can not only deepen our understanding of a specific popular graph problem, but also shed light upon some important connections between two seemingly unrelated areas, thus providing one possible powerful approach from PDEs’ viewpoint to studying recent popular graph problems. Meanwhile, we also provide a second application to study the multilevel graph partition problem and design a graph partition algorithm using the idea of weighted Laplacian method in order to overcome weaknesses of both traditional spectral methods and existing heuristic graph partitioning algorithms under the multilevel framework. Additionally, several comparative experiments also verify the validity of the proposed weighted Laplacian method in algorithmic design.

Inspired by the theory value of the first application, we would like to attempt to utilize the PDEs methods such as the finite element method and Monte Carlo method to handle some important graph partitioning problems in graph analysis and this will become our future work.

7. Acknowledgments

The work of this article is supported by Cyber Security Project of Ministry of Science and Technology of China, Research on Key Technologies of Internet of Things and Security Guarantee of Smart City (2018YFB080340), National Natural Science Fund for distinguished young scholars (No. 61625205), Key Research Projects in Frontier Science of the Chinese Academy of Sciences (QYZDY-SSW-JSC002).
References

[1] Andreev, K., and Racke, H. 2006. Balanced graph partitioning. Theory of Computing Systems 39(6):929–939.

[2] Aykanat, C.; Cambazoglu, B. B.; and Uçar, B. Multilevel direct k-way hypergraph partitioning with multiple constraints and fixed vertices. Journal of Parallel and Distributed Computing, 2008, 68(5):609–625.

[3] Barnard, S. T., and Simon, H. D. Fast multilevel implementation of recursive spectral bisection for partitioning unstructured problems. Concurrency: Practice and experience, 1994, 6(2):101–117.

[4] Carron, A.; Patel, R.; and Bullo, F. Hitting time for doubly-weighted graphs with application to robotic surveillance. In 2016 European Control Conference (ECC), 2016, 661–665. IEEE.

[5] Chen, X.; Zhexue Haung, J.; Nie, F.; Chen, R.; and Wu, Q. A self-balanced min-cut algorithm for image clustering. In Proceedings of the IEEE International Conference on Computer Vision, 2017, 2061–2069.

[6] Chung, S.-Y., and Berenstein, C. A. ω-harmonic functions and inverse conductivity problems on networks. SIAM Journal on Applied Mathematics, 2005, 65(4):1200–1226.

[7] Chung, F. R., and Graham, F. C. Spectral graph theory. Number 92. American Mathematical Soc. 1997.

[8] Chung, F. R., and Langlands, R. P. A combinatorial laplacian with vertex weights. Journal of combinatorial theory, 1996, Series A 75(2):316–327.

[9] Dhillon, I. S.; Guan, Y.; and Kulis, B. A unified view of kernel k-means, spectral clustering and graph cuts. Citeseer, 2004.

[10] Dhillon, I. S.; Guan, Y.; and Kulis, B. Weighted graph cuts without eigenvectors a multilevel approach. IEEE transactions on pattern analysis and machine intelligence, 2007, 29(11):1944–1957.

[11] Elmoataz, A.; Toutain, M.; and Tenbrinck, D. On the p-laplacian and l1-laplacian on graphs with applications in image and data processing. SIAM Journal on Imaging Sciences, 2015, 8(4):2412–2451.

[12] Hagen, L., and Kahng, A. B. New spectral methods for ratio cut partitioning and clustering. IEEE transactions on computer-aided design of integrated circuits and systems, 1992, 11(9):1074–1085.

[13] Hall, K. M. An r-dimensional quadratic placement algorithm. Management science, 1970, 17(3):219–229.

[14] Holtgrewe, M.; Sanders, P.; and Schulz, C. Engineering a scalable high quality graph partitioner. In 2010 IEEE International Symposium on Parallel & Distributed Processing (IPDPS), 2010, 1–12. IEEE.

[15] Karypis, G., and Kumar, V. A fast and high quality multilevel scheme for partitioning irregular graphs. SIAM Journal on scientific Computing, 1998, 20(1):359–392.

[16] Kim, Y.-J., and Kook, W. Harmonic cycles for graphs. Linear and Multilinear Algebra, 2019, 67(5):965–975.

[17] Knisley, J. R., and Knisley, D. J. Vertex-weighted graphs and their applications. Utilitas Mathematica, 2014, 94:237–249.

[18] Kolmogorov, V., and Zabih, R. What energy functions can be minimized via graph cuts? In European conference on computer vision, 2002, 65–81. Springer.

[19] Lin, S., and Kernighan, B. W. An effective heuristic algorithm for the traveling-salesman problem. Operations research, 1973, 21(2):498–516.

[20] Liu, R.; Feng, S.; Shi, R.; and Guo, W. Weighted graph clustering for community detection of large social networks. Procedia Computer Science, 2014, 31:85–94.

[21] Liu, H.; Han, J.; Nie, F.; and Li, X. Balanced clustering with least square regression. In Thirty-First AAAI Conference on Artificial Intelligence, 2017.

[22] Newman, M. E. Community detection and graph partitioning. EPL (Europhysics Letters), 2013, 103(2):28003.

[23] Oberman, A. M. Finite difference methods for the infinity laplace and p-laplace equations. Journal of Computational and Applied Mathematics, 2013, 254:65–80.
[24] Predari, M., and Esnard, A. A k-way greedy graph partitioning with initial fixed vertices for parallel applications. In 2016 24th Euromicro International Conference on Parallel, Distributed, and Network-Based Processing (PDP), 2016, 280–287. IEEE.

[25] Predari, M.; Esnard, A.; and Roman, J. Comparison of initial partitioning methods for multilevel direct k-way graph partitioning with fixed vertices. Parallel Computing, 2017, 66:22–39.

[26] Safro, I.; Sanders, P.; and Schulz, C. Advanced coarsening schemes for graph partitioning. Journal of Experimental Algorithmics (JEA), 2015, 19:2–2.

[27] Sanders, P., and Schulz, C. Think locally, act globally: Highly balanced graph partitioning. In International Symposium on Experimental Algorithms, 2013, 164–175. Springer.

[28] Saxena, A.; Prasad, M.; Gupta, A.; Bharill, N.; Patel, O. P.; Tiwari, A.; Er, M. J.; Ding, W.; and Lin, C.-T. 2017. A review of clustering techniques and developments. Neurocomputing 267:664–681.

[29] Shi, J., and Malik, J. Normalized cuts and image segmentation. Departmental Papers (CIS), 2000, 107.

[30] Shi, Z.; Osher, S.; and Zhu, W. Weighted graph laplacian and image inpainting. J. Sci. Comput, 2016, 577.

[31] Sun, J.; Ovsjanikov, M.; and Guibas, L. A concise and provably informative multi-scale signature based on heat diffusion. In Computer graphics forum, 2009, volume 28, 1383–1392. Wiley Online Library.

[32] Szlam, A., and Bresson, X. Total variation, cheeger cuts. In ICM, 2010, 1039–1046.

[33] Ubaru, S., and Saad, Y. Sampling and multilevel coarsening algorithms for fast matrix approximations. Numerical Linear Algebra with Applications, 2019, 26(3):e2234.

[34] Wiener, H. Structural determination of paraffin boiling points. Journal of the American Chemical Society, 1947, 69(1):17–20.