Internet Mathematics Vol. 6, No. 3: 315–330

A Local Graph Partitioning Algorithm Using Heat Kernel Pagerank

Fan Chung

Abstract. We give an improved local partitioning algorithm using heat kernel pagerank, a modified version of PageRank. For a subset S with Cheeger ratio (or conductance) h, we show that at least a quarter of the vertices in S can serve as seeds for heat kernel pagerank that lead to local cuts with Cheeger ratio at most $O(\sqrt{h})$, improving the previous bound by a factor of $\sqrt{\log s}$, where s denotes the volume of S.

1. Introduction

With the emergence of massive information networks, many previous graph algorithms are no longer feasible for various applications. A basic setup for a generic algorithm usually includes a graph as a part of its input. This, however, is no longer possible for dealing with massive graphs with prohibitively large size. Instead, the (host) graph, such as the Web graph or various social networks, is usually meticulously crawled, organized, and stored in some appropriate database. The *local* algorithms that we study here involve only "local access" of the database of the host graph. For example, selecting a neighbor of a specified vertex is considered to be a type of local access. Of course, it is desirable to minimize the number of local accesses needed, preferably independent of n, the number of vertices in the host graph (which may as well be regarded as "infinity"). In this paper, we consider a local algorithm that improves the performance bound of previous local partitioning algorithms.

1542-7951/09 $0.50 per page

Graph partitioning problems have long been studied and used for a wide range of applications, typically through divide-and-conquer approaches. Since the exact solution for graph partitioning is known to be NP-complete [Garey and Johnson 79], various approximation algorithms have been utilized. One of the best known partition algorithms is the spectral algorithm. The vertices are ordered using an eigenvector, and only cuts that are initial segments in such an ordering are considered. The advantage of such a "one-sweep" algorithm is to reduce the number of cuts under consideration from an exponential number in n to a linear number. Still, there is a performance guarantee within a quadratic order by using a relation between eigenvalues and the Cheeger constant, called the Cheeger inequality. However, for a large graph (say with a hundred million vertices), the task of computing an eigenvector is often too costly and not competitive.

A local partitioning algorithm finds a partition that separates a subset of nodes of specified size near specified seeds. In addition, the running time of a local algorithm is required to be proportional to the size of the separated part but independent of the total size of the graph. A local partitioning algorithm using random walks was first given in [Spielman and Teng 04]. An analysis of their algorithm is based on a mixing result of Lovász and Simonovits in their work on approximating the volume of a convex body. The same mixing result was also proved earlier independently in [Mihail 89]. In a previous paper [Andersen et al. 06], a local partitioning algorithm was given using PageRank, a concept first introduced in [Brin and Page 98] that has been widely used for Web search algorithms. The notion of PageRank that can be carried out for any graph is basically an efficient way of organizing random walks in a graph. As seen in the detailed definition given later, PageRank can be expressed as a geometric sum of random walks starting from a seed (or an initial probability distribution), with its speed of propagation controlled by a jumping constant. The usual question in random walks is to determine how many steps are required to get close to a stationary distribution. In the use of PageRank, the problem is reduced to specifying the range for the jumping constant to achieve the desired mixing. The advantage of using PageRank as in [Andersen et al. 06] is to reduce the computational complexity by a factor of $\log n$.

In this paper, we consider a modified version of PageRank called *heat kernel pagerank*. Like PageRank, heat kernel pagerank has two parameters: a seed and a heat constant or temperature. Heat kernel pagerank can be expressed as an exponential sum of random walks from the seed, scaled by the temperature. In addition, heat kernel pagerank satisfies a heat equation that dictates the rate of diffusion. We will examine several useful properties of heat kernel pagerank. In particular, for a given subset of vertices, we consider eigenvalues on the induced subgraph on S satisfying a Dirichlet boundary condition (details to be given in the next section). We will show that for a subset S with Cheeger ratio h, there are many vertices in S (whose volume is at least a quarter of the volume of S) such that the one-sweep algorithm using heat kernel pagerank with such vertices as seeds will find local cuts with Cheeger ratio $O(\sqrt{h})$. This improves the previous bound of $O(\sqrt{h \log s})$ in a similar theorem using PageRank [Andersen et al. 06, Andersen et al. 07]. Here s denotes the volume of S, and a local cut has volume at most s.

2. Preliminaries

In a graph G, the transition probability matrix W of a typical random walk on a graph $G = (V, E)$ is a matrix with columns and rows indexed by V and is defined by

$$W(u, v) = \begin{cases} 1/d_u & \text{if } \{u, v\} \in E, \\ 0 & \text{otherwise,} \end{cases}$$

where the *degree* of v, denoted by d_v, is the number of vertices to which v is adjacent. We can write $W = D^{-1}A$, where A denotes the *adjacency matrix* of G and D is the *diagonal degree matrix*. A random walk on a graph G has a stationary distribution π if G is connected and nonbipartite. The stationary distribution π, if it exists, satisfies $\pi(u) = d_u / \sum_v d_v$.

The version of PageRank we consider here is also called personalized PageRank (see [Haveliwala 03, Jeh and Widom 03]), which generalizes the version first introduced by Brin and Page. PageRank has two parameters: a *preference vector* f (i.e., the probabilistic distribution of the seed(s)) and a *jumping constant* α. Here, the function $f : V \to \mathbb{R}$ is taken to be a row vector so that W can act on f from the right by matrix multiplication. PageRank $\mathrm{pr}_{\alpha,f}$ with scale parameter α and preference vector f satisfies the following recurrence relation:

$$\mathrm{pr}_{\alpha,f} = \alpha f + (1 - \alpha)\mathrm{pr}_{\alpha,f} W. \tag{2.1}$$

An equivalent definition for PageRank is the following:

$$\mathrm{pr}_{\alpha,f} = \alpha \sum_{k=0}^{\infty} (1 - \alpha)^k f W^k. \tag{2.2}$$

For example, if we have one starting seed denoted by vertex u, then f can be written as the $(0, 1)$-indicator function χ_u of u. Another example is to take f to be the constant function with value $1/n$ at every vertex as in the original definition of PageRank in [Brin and Page 98].

Heat kernel pagerank also has two parameters, with (temperature) $t \geq 0$ and a preference vector f defined as follows:

$$\rho_{t,f} = e^{-t} \sum_{k=0}^{\infty} \frac{t^k}{k!} f W^k. \tag{2.3}$$

We see that (2.3) is just an exponential sum, while (2.2) is a geometric sum. For many combinatorial problems, exponential generating functions play a useful role. Heat kernel pagerank as defined in (2.3) satisfies the following heat equation:

$$\frac{\partial}{\partial t} \rho_{t,f} = -\rho_{t,f}(I - W). \tag{2.4}$$

Let us define $L = I - W$. Then the definition of heat kernel pagerank in (2.3) can be rewritten as follows:

$$\rho_{t,f} = f H_t,$$

where H_t is defined by

$$H_t = e^{-t} \sum_{k=0}^{\infty} \frac{t^k}{k!} W^k = e^{-t(I-W)} = e^{-tL} = \sum_{k=0}^{\infty} \frac{(-t)^k}{k!} L^k.$$

From the above definition, we have the following facts for $\rho_{t,f}$, which will be useful later.

Lemma 2.1. *For a graph G, its heat kernel pagerank ρ satisfies the following:*

(i) $\rho_{0,f} = f$.

(ii) $\rho_{t,\pi} = \pi$.

(iii) $\rho_{t,f} \mathbf{1}^* = f \mathbf{1}^* = 1$ *if f satisfies $\sum_v f(v) = 1$, where $\mathbf{1}$ denotes the all-1's function (as a row vector) and $\mathbf{1}^*$ denotes the transpose of $\mathbf{1}$.*

(iv) $W H_t = H_t W$.

(v) $D H_t = H_t^* D$ *and* $H_t = H_{t/2} H_{t/2} = H_{t/2} D^{-1} H_{t/2}^* D$.

Proof. Since $H_0 = I$, (i) follows. Statements (ii) and (iii) can be easily checked, while (iv) follows from the fact that H_t is a polynomial of W. Finally, (v) is a consequence of the fact that $W = D^{-1} A$. □

3. Dirichlet Eigenvalues and the Restricted Heat Kernel

For a subset S of $V(G)$, there are two types of boundary of S: the *vertex boundary* $\delta(S)$ and the *edge boundary* $\partial(S)$. The vertex boundary $\delta(S)$ is defined as follows:

$$\delta(S) = \{u \in V(G) \setminus S : u \sim v \text{ for some } v \in S\}.$$

For a single vertex v, the *degree* of v, denoted by d_v, is equal to $|\delta(v)|$ (which is short for $|\delta(\{v\})|$.

The *closure* of S, denoted by S^*, is the union of S and δS. For a function $f : S^* \to \mathbb{R}$, we say that f satisfies the *Dirichlet boundary condition* if $f(u) = 0$ for all $u \in \delta(S)$. We use the notation $f \in \mathbf{D}_S^*$ to denote that f satisfies the Dirichlet boundary condition and we require that $f \neq 0$.

For $f \in \mathbf{D}_S^*$, we define a Dirichlet Rayleigh quotient

$$R_S(f) = \frac{\sum_{x \sim y}(f(x) - f(y))^2}{\sum_{x \in S} f^2(x) d_x},$$

where the sum is to be taken over all unordered pairs of vertices $x, y \in S^*$ such that $s \sim y$.

The Dirichlet eigenvalue of an induced subgraph on S of a graph G can be defined as follows:

$$\lambda_S = \inf_{f \in \mathbf{D}_S^*} R_S(f) = \inf_{f \in \mathbf{D}_S^*} \frac{\langle f, (D_S - A_S) f \rangle}{\langle f, D_S f \rangle} = \inf_{g \in \mathbf{D}_S^*} \frac{\langle g, \mathcal{L}_S g \rangle}{\langle g, g \rangle},$$

where X_S denotes the submatrix of a matrix X with rows and columns restricted to those indexed by vertices in S, and the Laplacians $\mathcal{L}$ and $\mathcal{L}_S$ are defined by $\mathcal{L} = D^{-1/2}(D - A)D^{-1/2}$ and $\mathcal{L}_S = D_S^{-1/2}(D_S - A_S)D_S^{-1/2}$, respectively. Here we call f the combinatorial Dirichlet eigenfunction if $R_S(f) = \lambda_S$. The Dirichlet eigenfunctions are the eigenfunctions of the matrix $\mathcal{L}_S$. (A detailed proof of these statements can be found in [Chung 97].)

The Dirichlet eigenvalues of S are the eigenvalues of $\mathcal{L}_S$, denoted by

$$\lambda_{S,1} \leq \lambda_{S,2} \leq \cdots \leq \lambda_{S,s},$$

where $s = |S|$. The smallest Dirichlet eigenvalue $\lambda_{S,1}$ is also denoted by λ_S. If the induced subgraph on S is connected, then the eigenvector of $\mathcal{L}_S$ associated with λ_S is all positive (using the Perron–Frobenius theorem [Perron 33] on $I - \mathcal{L}_S$). The reader is referred to [Chung 97] for various properties of Dirichlet eigenvalues.

For a subset S of vertices in G, the edge separator (or the edge boundary) whose removal separates S consists of all edges leaving S. Namely,

$$\partial S = \{\{u, v\} \in E : u \in S \text{ and } v \notin S\}.$$

How good is the edge separator? The answer depends on both the size of the edge separator and the *volume* of S, which is denoted by $\mathrm{vol}(S)$ and is equal to $\sum_{u \in S} d_u$. The volume of a graph G, denoted by $\mathrm{vol}(G)$, is $\sum_u d_u$.

The *Cheeger ratio* of S, denoted by h_S, is defined by

$$h_S = \frac{|\partial S|}{\min\{\mathrm{vol}(S), \mathrm{vol}(\bar{S})\}},$$

where $\bar{S} = V \setminus S$ denotes the complement of S. The *Cheeger constant* of a graph G is $h_G = \min_{S \subseteq V} h_S$. The Cheeger constant of a graph G is often called the conductance of G. For a given subset S, we define the *local Cheeger ratio*, denoted by h_S^*, as follows:

$$h_S^* = \inf_{T \subseteq S} h_T.$$

Note that in general, h_S is not necessarily equal to h_S^*.

The Dirichlet eigenvalue λ_S and the local Cheeger ratio h_S^* are related by the following local Cheeger inequality:

$$h_S^* \geq \lambda_S \geq \frac{(h_S^*)^2}{2}.$$

A proof can be found in [Chung 07a].

In [Chung and Oden 00], the weighted Rayleigh quotient is defined as

$$R_\phi(f) = \sup_c \frac{\sum_{x \sim y} (f(x) - f(y))^2 \phi(x)\phi(y)}{\sum_{x \in S} (f(x) - c)^2 \phi^2(x) d_x}, \tag{3.1}$$

where ϕ is the combinatorial Dirichlet eigenfunction that achieves λ_S . Then we can define

$$\lambda_\phi = \inf_{f \neq 0} R_\phi(f).$$

Here we state several useful facts concerning λ_S and λ_ϕ (see [Chung and Oden 00]).

Lemma 3.1. *For an induced subgraph S of a graph G, the Dirichlet eigenvalues of S satisfy*

$$\lambda_{S,2} - \lambda_S = \lambda_\phi \geq \lambda_S.$$

Theorem 3.2. *For an induced subgraph S of G, the combinatorial Dirichlet eigenfunction ϕ that achieves λ_S satisfies*

$$\left(\sum_{x \in S} \phi(x) d_x \right)^2 \geq \frac{1}{2} \mathrm{vol}(S) \sum_{x \in S} \phi^2(x) d_x.$$

Proof. We note that

$$\lambda_{S,2} = \inf_f \frac{\sum_{x\sim y}(f(x)-f(y))^2}{\sum_{x\in S} f(x)^2 d_x},$$

where f ranges over all functions satisfying $\sum_{x\in S} f(x)\phi(x) = 0$. Now we use the fact that

$$\lambda_S = \frac{\sum_{x\sim y}(\phi(x)-\phi(y))^2}{\sum_{x\in S}\phi(x)^2 d_x}.$$

We then have

$$\lambda_{S,2} \le \frac{\sum_{x\sim y}(\phi(x)-\phi(y))^2}{\sum_{x\in S}(\phi(x)-c)^2 d_x} = \frac{\lambda_S \sum_{x\in S}\phi^2(x)d_x}{\sum_{x\in S}\phi^2(x)d_x - c^2\mathrm{vol}(S)},$$

where

$$c = \frac{\sum_{x\in S}\phi(x)d_x}{\mathrm{vol}(S)}.$$

This implies

$$\frac{\left(\sum_{x\in S}\phi(x)d_x\right)^2}{\mathrm{vol}(S)\sum_{x\in S}\phi^2(x)d_x} \ge \frac{\lambda_{S,2}-\lambda_S}{\lambda_{S,2}} = \frac{\lambda_\phi}{\lambda_\phi+\lambda_S} \ge \frac{1}{2}$$

using Lemma 3.1. □

4. A Lower Bound for Restricted Heat Kernel Pagerank

For a given set S, we consider the distribution f_S with the vertex u chosen with probability $f_S(u) = d_u/\mathrm{vol}(S)$ if $u \in S$, and 0 otherwise. Note that f_S can be written as $\frac{1}{\mathrm{vol}(S)}\chi_S D$, where χ_S is the indicator function for S. For any function $g : V \to \mathbb{R}$, we define $g(S) = \sum_{v\in S} g(v)$.

In this section, we wish to establish a lower bound for the expected value of heat kernel pagerank $\rho_{t,u} = \rho_{t,\chi_u}$ over u in S. We note that

$$\mathbb{E}(\rho_{t,u}) = \sum_{u\in S}\frac{d_u}{\mathrm{vol}(S)}\rho_{t,u} = f_S H_t(S).$$

We consider the restricted heat kernel H_t', defined below, for a fixed subset S. Then the definition of the heat kernel pagerank in (2.3) can be rewritten as follows:

$$\rho'_{t,f} = fH'_t,$$

where H'_t is defined by

$$H'_t = e^{-t}\sum_{k=0}^{\infty}\frac{t^k}{k!}W_S^k = e^{-t(I_S-W_S)} = e^{-tL_S} = \sum_{k=0}^{\infty}\frac{(-t)^k}{k!}L_S^k.$$

Also, H'_t satisfies the following heat equation:

$$\frac{\partial}{\partial t}H'_t = -L_S H'_t. \tag{4.1}$$

From the above definition, we immediately have the following result.

Lemma 4.1.

$$H'_t(x,y) \le H_t(x,y)$$

for all vertices x and y. In particular, for a nonnegative function $f : V \to \mathbb{R}$, we have

$$fH'_t(v) \le fH_t(v)$$

for every vertex v in G.

Therefore, it suffices to establish the desired lower bound for the expected value of restricted heat kernel pagerank.

Following the notation in Section 3, we consider the Dirichlet eigenvalues $\lambda_{S,i}$ of S and their associated Dirichlet combinatorial eigenfunction ϕ_i with $R(\phi_i) = \lambda_{S,i}$, for $i = 1,\dots,|S|$. Clearly, $\phi_i D_S^{1/2}$ are orthogonal eigenfunctions of $\mathcal{L}_S$ and form a basis for functions defined on S. Here we assume that $\sum_{u\in S}\phi_i(u)^2 d_u = 1$. To simplify the notation, in this proof we write $\lambda'_i = \lambda_{S,i}$ and $\lambda'_1 = \lambda_S$.

We express $f = \sqrt{\mathrm{vol}(S)} f_S$ in terms of ϕ_i as follows:

$$fD_S^{-1/2} = \sum_i a_i \phi_i D_S^{1/2},$$

where

$$a_i = \sum_{u\in S}\frac{\phi_i(u)d_u}{\sqrt{\mathrm{vol}(S)}}.$$

Since $\|fD_S^{-1/2}\| = \|fD_S^{-1/2}\|_2 = 1$, we have

$$\sum_i a_i^2 = 1.$$

From the above definitions we have

$$f_S H'_t(S) = \|f\mathcal{H}'_{t/2}\|^2,$$

where $\mathcal{H}'_t = D_S^{1/2} H'_t D_S^{-1/2}$. From Theorem 3.2, we know that

$$a_1^2 = \frac{\left(\sum_{u \in S} \phi_1(u) d_u\right)^2}{\mathrm{vol}(S)} \geq \frac{1}{2} \geq 1 - \sum_{j \neq 1} a_j^2.$$

Since

$$f_S H'_t(S) = \sum_i a_i^2 e^{-\lambda'_i t},$$

we have

$$f_S H'_t(S) \geq a_1^2 e^{-\lambda_S t} \geq \frac{1}{2} e^{-\lambda_S t}.$$

We have proved the following theorem.

Theorem 4.2. *For a subset S, the Dirichlet heat kernel H'_t satisfies*

$$f_S H'_t(S) \geq \frac{1}{2} e^{-\lambda_S t}.$$

As an immediate consequence of Theorem 3.2, Lemmas 3.1 and 4.1, and Theorem 4.2, we have the following corollary.

Corollary 4.3. *In a graph G, for a subset S of vertices, heat kernel pagerank ρ_{t,f_S} satisfies*

$$\mathbb{E}(\rho_{t,u}(S)) = \rho_{t,f_S}(S) \geq \frac{1}{2} e^{-\lambda_S t} \geq \frac{1}{2} e^{-h_S^* t},$$

where u is chosen according to f_S.

5. A Local Lower Bound for Heat Kernel Pagerank

Corollary 4.3 states that the expected value of $\rho_{t,u}$ is at least $e^{-h(S)t}$. Thus, there exists a vertex u in S such that $\rho_{t,u}(S)$ is at least $e^{-h(S)t}$. However, in order to have an efficient local partitioning algorithm, we need to show that there are many vertices v satisfying

$$\rho_{t,v}(S) \geq c\rho_{t,f_S}(S)$$

for some absolute constant c. To do so, we will prove the following theorem.

Theorem 5.1. *In a graph G with a given subset S of vertices, the subset $T = \{u \in S : \rho_{t,u}(S) \geq \frac{1}{4} e^{-t\lambda_S}\}$ satisfies*

$$\mathrm{vol}(T) \geq \frac{1}{4} \mathrm{vol}(S)$$

if $t \geq 1/\lambda_S$.

From Lemma 4.1, Theorem 5.1 follows directly from Theorem 5.2.

Theorem 5.2. *In a graph G with a given subset S of vertices, the subset $T = \{u \in S : \rho'_{t,u}(S) \geq \frac{1}{4}e^{-t\lambda_S}\}$ satisfies*

$$\mathrm{vol}(T) \geq \frac{1}{4}\mathrm{vol}(S)$$

if $t \geq 1/\lambda_S$.

To prove Theorem 5.2, we first prove the following lemma using second-moment methods.

Lemma 5.3. *If $t \geq 1/\lambda_S$, then*

$$\sum_{u\in S} \frac{d_u}{\mathrm{vol}(S)} \left(\rho'_{t,u}(S) - \rho'_{t,f_S}(S)\right)^2 \leq \frac{5}{4}\rho'_{t,f_S}(S)^2.$$

Proof. We note that

$$\begin{aligned}\sum_{u\in S} \frac{d_u}{\mathrm{vol}(S)} \left(\rho'_{t,u}(S) - \rho'_{t,f_S}(S)\right)^2 &= \sum_{u\in S} \frac{d_u}{\mathrm{vol}(S)}\rho'_{t,u}(S)^2 - (\rho'_{t,f_S}(S))^2 \\ &= f_S H'_{2t}(S) - (f_S H'_t(S))^2.\end{aligned}$$

It suffices to show that

$$f_S H'_{2t}(S) \leq \frac{9}{4}(f_S H'_t(S))^2.$$

We consider the Dirichlet eigenvalues $\lambda_{S,i}$ of S and their associated Dirichlet combinatorial eigenfunctions ϕ_i with $R(\phi_i) = \lambda_{S,i} = \lambda'_i$, for $i = 1, \ldots, |S|$. Here $\phi_i D_S^{1/2}$ are orthonormal eigenfunctions of $\mathcal{L}_S$ with $\sum_{u\in S} \phi_i(u)^2 d_u = 1$.

We can write $f = \chi_S D_S \mathrm{vol}(S)^{-1/2} = f_S\sqrt{\mathrm{vol}(S)}$ by

$$f D_S^{-1/2} = \sum_i a_i \phi_i D_S^{1/2}.$$

We have

$$\sum_i a_i^2 = 1,$$

since $\|f D_S^{-1/2}\|_2 = 1$. From Theorem 3.2, we know that

$$a_1^2 \geq 1/2 \geq 1 - \sum_{j\neq 1} a_j^2.$$

Also, we have

$$f_S H'_{2t}(S) = \|f\mathcal{H}'_t\|^2 = \sum_i a_i^2 e^{-2\lambda'_i t},$$

where $\mathcal{H}'_t = D_S^{1/2} H'_t D_S^{-1/2}$.

Therefore we have, for $t \geq 1/\lambda'_1$,

$$\begin{aligned} f_S H'_{2t}(S) &= \sum_i a_i^2 e^{-2\lambda'_i t} \\ &\leq a_1^2 e^{-2\lambda'_1 t} + (1 - a_1^2) e^{-2\lambda'_2 t} \\ &= a_1^2 e^{-2\lambda'_1 t} + (1 - a_1^2) e^{-4\lambda'_1 t} \\ &\leq \frac{9}{8} a_1^2 e^{-2\lambda'_1 t}, \end{aligned}$$

since $\lambda'_2 \geq 2\lambda'_1$ using Lemma 3.1. Therefore we have

$$f_S H'_{2t}(S) \leq \frac{9}{8} a_1^2 e^{-2\lambda_S t} \leq \frac{9}{4} a_1^4 e^{-2\lambda_S t} \leq \frac{9}{4} \left(\sum_i a_i^2 e^{-\lambda'_i t} \right)^2,$$

as desired. □

Now we are ready to prove Theorem 5.2, which then implies Theorem 5.1.

Proof of Theorem 5.2. Suppose $\mathrm{vol}(T) \leq \mathrm{vol}(S)/4$. We wish to show that this leads to a contradiction. Let $T' = S \setminus T$. We consider

$$\begin{aligned} &\sum_{u \in S} \frac{d_u}{\mathrm{vol}(S)} \left(\rho'_{t,u}(S) - \rho'_{t,f_S}(S) \right)^2 \\ &= \sum_{u \in T} \frac{d_u}{\mathrm{vol}(S)} \left(\rho'_{t,u}(S) - \rho'_{t,f_S}(S) \right)^2 + \sum_{u \in T'} \frac{d_u}{\mathrm{vol}(S)} \left(\rho'_{t,u}(S) - \rho'_{t,f_S}(S) \right)^2 \\ &\geq \frac{\left(\sum_{u \in T} \frac{d_u}{\mathrm{vol}(S)} (\rho'_{t,u}(S) - \rho'_{t,f_S}(S)) \right)^2}{\mathrm{vol}(T)/\mathrm{vol}(S)} + \frac{\left(\sum_{u \in T'} \frac{d_u}{\mathrm{vol}(S)} (\rho'_{t,u}(S) - \rho'_{t,f_S}(S)) \right)^2}{\mathrm{vol}(T')/\mathrm{vol}(S)} \\ &\geq \left(\sum_{u \in T'} \frac{d_u}{\mathrm{vol}(S)} (\rho'_{t,u}(S) - \rho'_{t,f_S}(S)) \right)^2 \left(\frac{1}{\mathrm{vol}(T)/\mathrm{vol}(S)} + \frac{1}{\mathrm{vol}(T')/\mathrm{vol}(S)} \right), \end{aligned}$$

since

$$\sum_{u \in S} \frac{d_u}{\mathrm{vol}(S)} (\rho'_{t,u} - \rho'_{t,f_S}) = 0.$$

Therefore we have

$$\sum_{u\in S}\frac{d_u}{\mathrm{vol}(S)}\left(\rho'_{t,u}(S)-\rho'_{t,f_S}(S)\right)^2$$
$$\geq\left(\frac{\mathrm{vol}(T')}{\mathrm{vol}(S)}\left(\frac{3}{4}\right)\rho'_{t,f_S}(S)\right)^2\left(\frac{1}{\mathrm{vol}(T)/\mathrm{vol}(S)}+\frac{1}{\mathrm{vol}(T')/\mathrm{vol}(S)}\right)$$
$$\geq\left(4\left(\frac{3}{4}\right)^4+\left(\frac{3}{4}\right)^3\right)\rho'_{t,f_S}(S)^2\geq\frac{27}{16}\rho'_{t,f_S}(S)^2>\frac{5}{4}\rho'_{t,f_S}(S)^2,$$

which is a contradiction to Lemma 5.3. Theorem 5.2 is proved. □

6. An Upper Bound for Heat Kernel Pagerank

For a function $f : V \to \mathbb{R}$, we can order vertices according to their f values. The ordered list is then called a *sweep*. The segment S_i consists of the first i vertices (by breaking ties arbitrarily). We define an s-local Cheeger ratio of a sweep f, denoted by $h_{f,s}$, to be the minimum Cheeger ratio of the segment S_i with $0 \leq \mathrm{vol}(S_i) \leq 2s$. If no such segment exists, then we set $h_{f,s}$ to be 0. We will establish the following upper bound for heat kernel pagerank in terms of s-local Cheeger ratios. The proof is similar to that in [Chung 07b], but simpler.

Theorem 6.1. *In a graph G with a subset S with volume $s \leq \mathrm{vol}(G)/4$, for any vertex u in G, we have*

$$\rho_{t,u}(S)-\pi(S)\leq\sqrt{\frac{s}{d_u}}e^{-t\kappa^2_{t,u,s}/4},$$

where $\kappa_{t,u,s}$ denotes the minimum s-local Cheeger ratio over a sweep of $\rho_{t,u}$.

Proof. For a function f $V \to \mathbb{R}$, we define $f(u,v) = f(u)/d_u$ if v is adjacent to u and 0 otherwise. For an integer x, $0 \leq x \leq \mathrm{vol}(G)/2$, we define

$$f(x)=\max_{T\subseteq V\times V,|T|=x}\sum_{(u,v)\in T}f(u,v).$$

We can extend f to all real $x = k + r$ with $0 \leq r < 1$ by defining

$$f(x)=(1-r)f(k)+rf(k+1).$$

If $x = \mathrm{vol}(S_i)$, where S_i consists of vertices with the i highest values of $f(u)/d_u$, then it follows from the definition that $f(x) = \sum_{u\in S_i} f(u)$. Also, $f(x)$ is concave in x.

We consider the lazy walk $\mathbf{W} = (I + W)/2$. Then

$$\begin{aligned} f\mathbf{W}(S) &= \frac{1}{2}\left(f(S) + \sum_{u\sim v\in S} f(u,v)\right) \\ &= \frac{1}{2}\left(\sum_{u \text{ or } v\in S} f(u,v) + \sum_{u \text{ and } v\in S} f(u,v)\right) \\ &\leq \frac{1}{2}\left(f(\mathrm{vol}(S) + |\partial S|) + f(\mathrm{vol}(S) - |\partial S|)\right) \\ &= \frac{1}{2}\left(f(\mathrm{vol}(S)(1+h_S)) + f(\mathrm{vol}(S)(1-h_S))\right). \end{aligned}$$

This can be straightforwardly extended to real x with $0 \leq x \leq \mathrm{vol}(G)/2$. In particular, we focus on x satisfying $0 \leq x \leq 2s \leq \mathrm{vol}(G)/2$ and we choose $f_t = \rho_{t,u} - \pi$. Then

$$f_t\mathbf{W}(x) \leq \frac{1}{2}\left(f_t(x(1+\kappa_{t,u,s})) + f_t(x(1-\kappa_{t,u,s}))\right).$$

We now consider for $x \in [0, 2s]$,

$$\begin{aligned} \frac{\partial}{\partial t} f_t(x) &= -\rho_{t,u}(I - W)(x) \\ &= -2\rho_{t,u}(I - \mathbf{W})(x) \\ &= -2f_t(x) + 2f_t\mathbf{W}(x) \qquad\qquad (6.1) \\ &\leq -2f_t(x) + f_t(x(1+\kappa_{t,u,s})) + f_t(x(1-\kappa_{t,u,s})) \\ &\leq 0 \end{aligned}$$

by the concavity of f_t. Suppose $g_t(x)$ is a solution of the equation in (6.1) satisfying $f_0(x) \leq g_0(x)$, $f_t(0) = g_t(0)$, and $\frac{\partial}{\partial t} f_t(x)|_{t=0} \leq \frac{\partial}{\partial t} g_t(x)|_{t=0}$. Then we have $f_t(x) \geq g_t(x)$. It is easy to check that $g_t(x) \leq e^{-t\kappa_{t,u,s}^2/4}\sqrt{x/d_u}$ using $-2 + \sqrt{1+x} + \sqrt{1-x} \leq -x^2/4$. Thus,

$$\rho_{t,u}(S) - \pi(S) \leq \rho_{t,u}(s) - \pi(s) \leq \sqrt{\frac{s}{d_u}} e^{-t\kappa_{t,u,s}^2/4},$$

as desired. □

7. A Local Cheeger Inequality and a Local Partitioning Algorithm

Let h_s denote the minimum Cheeger ratio h_S with $0 \leq \mathrm{vol}(S) \leq 2s$. Also let $\kappa_{t,2s}$ denote the minimum of $\kappa_{t,u,2s}$ over all u. Combining Theorem 5.1 and

Theorem 6.1, we have that the set of u satisfying

$$\frac{1}{2}e^{-t\ h_S^*} \leq \frac{1}{2}e^{-t\lambda_S} \leq \rho_{t,f_S}(s) - \pi(s) \leq \sqrt{s}e^{-t\kappa_{t,u,2s}^2/4}$$

has volume at least $\mathrm{vol}(S)/4$, provided $t \geq 1/h_S^2 \geq 1/\lambda_S$.

As an immediate consequence, we have the following local Cheeger inequality.

Theorem 7.1. *For a subset S of a graph G with $\mathrm{vol}(S) = s \leq \mathrm{vol}(G)/4$ and $t \geq \log s/(h_S^*)^2$, with probability at least $1/4$ a vertex u in S satisfies*

$$h_S^* \geq \lambda_S \geq \frac{\kappa_{t,u}^2}{4} - \frac{2\log s}{t},$$

where the Cheeger ratio $\kappa_{t,u}$ is determined by heat kernel pagerank with seed u.

Corollary 7.2. *For $s \leq \mathrm{vol}(G)/4$, $t \geq 4\log s/(h_S^*)^2$, and a set S of volume s, the Cheeger ratio $\kappa_{t,u}$, determined by heat kernel pagerank with a random seed u in S satisfies*

$$h_S^* \geq \lambda_S \geq \frac{\kappa_{t,u}^2}{8}$$

with probability at least $1/4$.

The above local Cheeger inequalities are closely associated with local partition algorithms. A local partition algorithm has inputs including a vertex as the seed, the volume s of the target set, and a target value ϕ for the Cheeger ratio of the target set. The local Cheeger inequality in Theorem 6.1 suggests the following local partition algorithm. In order to find the set achieving the minimum s-local Cheeger ratio, one can simply consider a sweep of heat kernel pagerank with further restrictions to the cuts with smaller parts of volume between 0 and $2s$.

Theorem 7.1 implies the following.

Theorem 7.3. *In a graph G, for a set S with volume $s \leq \mathrm{vol}(G)/4$ and Cheeger ratio $h_S \leq \phi^2$, there is a subset $S' \subseteq S$ with $\mathrm{vol}(S') \geq \mathrm{vol}(S)/4$ such that for any $u \in S'$, the sweep using heat kernel pagerank $\rho_{t,u}$, with $t = 2\phi^{-2}\log s$, will find a set T with s-local Cheeger ratio at most 2ϕ.*

We note that the performance bound for the Cheeger ratio improves the earlier result in [Andersen et al. 06] by a factor of $\log s$. In fact, the inequality in Theorem 7.1 suggests a whole range of tradeoffs. If we choose t to be $t = 2\phi^{-2}$ instead, then the guaranteed Cheeger ratio as in the above statement will be $2\phi\log s$, and we obtain the same approximation results as in [Andersen et al. 06].

We remark that the computational complexity of the above partitioning algorithm is the same as that of computing heat kernel pagerank. However, the algorithmic design for heat kernel pagerank has not been as extensively studied as that of PageRank. More research is needed in this direction.

References

[Andersen et al. 06] R. Andersen, F. Chung, and K. Lang. "Local Graph Partitioning Using PageRank Vectors." In *Proceedings of the 47th Annual IEEE Symposium on Founation of Computer Science (FOCS'2006)*, pp. 475–486. Washington, DC: IEEE Computer Society, 2006.

[Andersen et al. 07] R. Andersen, F. Chung, and K. Lang. "Detecting Sharp Drops in PageRank and a Simplified Local Partitioning Algorithm." In *Theory and Applications of Models of Computation: 4th International Conference, TAMC 2007, Shanghai, China, May 22–25, 2007, Proceedings*, Lecture Notes in Computer Science 4484, pp. 1–12. Berlin: Springer, 2007.

[Brin and Page 98] S. Brin and L. Page. "The Anatomy of a Large-Scale Hypertextual Web Search Engine." *Computer Networks and ISDN Systems* 30:1–7 (1998), 107–117.

[Cheeger 70] J. Cheeger. "A Lower Bound for the Smallest Eigenvalue of the Laplacian." *In Problems in Analysis*, edited by R. C. Gunning, pp. 195–199. Princeton: Princeton University Press, 1970.

[Chung 97] F. Chung. *Spectral Graph Theory*, CBMS Regional Conference Series in Mathematics 92. Providence: American Mathematical Society, 1997.

[Chung 07a] F. Chung. "Random Walks and Local Cuts in Graphs." *LAA* 423 (2007), 22–32.

[Chung 07b] F. Chung. "The Heat Kernel as the PageRank of a Graph." *PNAS* 105:50 (2007), 19735–19740.

[Chung and Lu 06] F. Chung and L. Lu. *Complex Graphs and Networks*, CBMS Regional Conference Series in Mathematics 107. Providence: American Mathematical Society, 2006.

[Chung and Oden 00] F. Chung and K. Oden. "Weighted Graph Laplacians and Isoperimetric Inequalities." *Pacific Journal of Mathematics* 192 (2000), 257–273.

[Chung and Yau 99] F. Chung and S.-T. Yau. "Coverings, Heat Kernels and Spanning Trees." *Electronic Journal of Combinatorics* 6 (1999), #R12.

[Coppersmith and Winograd 90] D. Coppersmith and S. Winograd. "Matrix Multiplication via Arithmetic Progressions." *J. Symbolic Comput.* 9 (1990), 251–280.

[Garey and Johnson 79] M. R. Garey and D. S. Johnson. *Computers and Intractability: A Guide to the Theory of NP-Completeness*. San Francisco: W. H. Freeman, 1979.

[Haveliwala 03] T. H. Haveliwala. "Topic-Sensitive PageRank: A Context-Sensitive Ranking Algorithm for Web Search." *IEEE Trans. Knowl. Data Eng.* 15:4 (2003), 784–796.

[Jeh and Widom 03] G. Jeh and J. Widom. "Scaling Personalized Web Search." In *Proceedings of the 12th International Conference on World Wide Web*, pp. 271–279. New York: ACM, 2003.

[Jerrum and Sinclair 89] M. Jerrum and A. J. Sinclair. "Approximating the Permanent." *SIAM J. Computing* 18 (1989), 1149–1178.

[Kannan et al. 04] R. Kannan, S. Vempala, and A. Vetta. "On Clusterings: Good, Bad and Spectral." *JACM* 51 (2004), 497–515.

[Lovász and Simonovits 90] L. Lovász and M. Simonovits. "The Mixing Rate of Markov Chains, an Isoperimetric Inequality, and Computing the Volume." In *Proceedings of the 31st Annual Symposium on Foundations of Computer Science*, pp. 346–354. Washington, DC: IEEE Computer Society, 1990.

[Lovász and Simonovits 93] L. Lovász and M. Simonovits. Random Walks in a Convex Body and an Improved Volume Algorithm." *Random Structures and Algorithms* 4 (1993), 359–412.

[Mihail 89] M. Mihail. "Conductance and Convergence of Markov Chains: A Combinatorial Treatment of Expanders." *FOCS* (1989), 526–531.

[Perron 33] O. Perron. *Theorie der algebraischen Gleichungen, II* (zweite Auflage). Berlin: de Gruyter, 1933.

[Schoen and Yau 94] R. M. Schoen and S. T. Yau. *Differential Geometry.* Cambridge, MA: International Press, 1994.

[Spielman and Teng 04] D. Spielman and S.-H. Teng. "Nearly-Linear Time Algorithms for Graph Partitioning, Graph Sparsification, and Solving Linear Systems." In *Proceedings of the 36th Annual ACM Symposium on Theory of Computing*, pp. 81–90. New York: ACM, 2004.

Fan Chung, University of California, San Diego, Department of Mathematics, 9500 Gilman Drive, La Jolla, CA 92093-0112 (fan@math.ucsd.edu)

Received June 30, 2009; accepted June 3, 2010.