Spectral Methods, Graph Partitioning, and Clustering

Shang-Hua Teng
Boston University/Akamai

Joint work with Daniel Spielman (MIT)
Ratio = \frac{\text{edges cut}}{\text{vertices removed}}
Eigenvector tries for good ratio cut
Planar graph eigenvalue bound

$$\lambda < \frac{8 \Delta}{n}$$

$$\Delta = \text{max degree}$$
\[ \lambda \left( \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \right) = \left( \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \right) \left( \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \right) \]

\[ \lambda = \frac{\sum_{(i,j) \in E} (x_i - x_j)^2}{\sum_{i} x_i^2} \]
Graph Partitioning

Bisection
Motivation

- Classical idea (Donath-Hoffman; 1972)
- Works well experimentally
- WHY? And ALWAYS?
- Graphs that arise in practice:
  - planar graphs
  - meshes, N-body graphs,
  - nearest neighbor graphs
- Other Applications: Data Clustering
Future Research and Open Questions

- Constant-factor approximation of bisection
- Eigenvectors and multicommodity flow
- Spectral methods for combinatorial problems: coloring, clustering, ordering, independent sets
- Graph embedding and geometry of graphs
# Small Separator Theorems

<table>
<thead>
<tr>
<th>Category</th>
<th>Bound</th>
<th>Authors</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trees</td>
<td>$(1, 1:2)$</td>
<td>Jordan</td>
</tr>
<tr>
<td>Planar Graphs</td>
<td>$(\sqrt{n}, 1:2)$</td>
<td>Lipton–Tarjan</td>
</tr>
<tr>
<td>Bounded–Genus Graphs</td>
<td>$(\sqrt{gn}, 1:2)$</td>
<td>Gilbert–Hutchinson–Tarjan</td>
</tr>
<tr>
<td>Bounded–Minor Graphs</td>
<td>$(h^{3/2}\sqrt{n}, 1:2)$</td>
<td>Alon–Seymour–Thomas</td>
</tr>
<tr>
<td>Nearest Neighbor Graphs</td>
<td>$(n^{1-1/d}, 1:d+1)$</td>
<td>Miller–Teng–Thurston–Vavasis</td>
</tr>
<tr>
<td>Finite–Element Meshes</td>
<td>$(n^{1-1/d}, 1:d+1)$</td>
<td>Miller–Teng–Thurston–Vavasis</td>
</tr>
<tr>
<td>N–body Graphs</td>
<td>$(n^{1-1/d} \lg n, 1:d+1)$</td>
<td>Teng</td>
</tr>
</tbody>
</table>
Spectral Separator Theorems

- Planar Graphs (Bounded Degree)
- Well-Shaped Meshes
- N-Body Graphs
- Nearest Neighbor Graphs
Well Shaped Mesh

$R : r$ is bounded
Convergence of Kleinberg Algorithms

• Eigenvalues and Eigenvectors
  \[ Ax = \lambda x \]

• Related with Spectral method for graph partitioning (Spielman-Teng)
  - Principle eigenvector projects good localities.
  - Eigenvector can be used for partitioning and clustering
Kleinberg’s Algorithm

- **Hubs**: pages with links to many quality authorities
- **Authorities**: pages with links from many quality hubs
- **Hubs** (imagine a good textbook and survey paper) and authorities (imagine Karp’s first paper on NP complete problem).
  - $A(q) \sim \sum_{p \in \text{IN}(q)} H(p)$
  - $H(q) \sim \sum_{p \in \text{OUT}(q)} A(p)$
Ranking Relevant Web-pages

- Use Link structures (Web-Graph)
  - Pages with high in-degree are important
  - Pages have links from important pages are important
- Model Web-graph as Markov Chains
  - Model random surfers
  - Roughly, let $R(q)$ be the rank of a page $q$ and let $IN(q)$ be the set of pages that refer $q$, then

  $$R(q) \sim \sum_{p \in IN(q)} \frac{R(p)}{N_p}$$

- The rank is related with the singular vector of the web-matrix.
Challenging Problems

- Searching Relevant Information
- Fast Delivery of Contents
- Secure Communication and Transaction
- Very Very Large Scale
- User Pattern Detection, and Profile Generation

Akamai
Clustering and Hierarchy
Clustering and Hierarchy

Subclusters. (Relative sizes of clusters are preserved)

Parent Cluster
Embedding Lemma:

\[
\lambda = \min_{\sum \mathbf{v}_i = 0} \frac{\sum \text{dist}(\mathbf{v}_i, \mathbf{v}_j)^2}{\sum \| \mathbf{v}_i \|^2}
\]
Donath-Hoffman

Fiedler

Cheeger, Alon, Sinclair-Jerrum

Pothen-Simon-Liou

Guattery-Miller
<table>
<thead>
<tr>
<th></th>
<th>$\lambda$</th>
<th>cut size</th>
</tr>
</thead>
<tbody>
<tr>
<td>2D mesh</td>
<td>$O(1/n)$</td>
<td>$O(n^{1/2})$</td>
</tr>
<tr>
<td>3D mesh</td>
<td>$O(1/n^{2/3})$</td>
<td>$O(n^{2/3})$</td>
</tr>
</tbody>
</table>
Well-Shaped Meshes and Sphere-Packings

Miller-Teng-Thurston-Vavasis

Well-Shaped Meshes ↔ Sphere-Packings

Miller-Talmor-Teng
Planar graph eigenvalue bound

\[ \sum_{(i,j) \in E} \text{dist}(v_i, v_j)^2 < 2\Delta \sum r_i^2 < 8\Delta \]

\[ \lambda < \frac{8\Delta}{n} \]

\( \Delta = \text{max degree} \)
\[ \Sigma \pi r_i^2 < 4\pi \]
\[ \text{dist}(v_i, v_j)^2 < 2 r_i^2 + 2 r_j^2 \]
Center of gravity at sphere center
Center of gravity at sphere center

proof: Brouwer’s fixed point theorem.
Use Brouwer's Fixed Point Theorem
Clustering
sphere-preserving map
Koebe-Andreev-Thurston Embedding Theorem:

kissing disks for planar graphs
Proof Outline

1. relate $\lambda$ to quality of embedding

2. prove graphs have good embeddings
Rayleigh Quotient

\[ \lambda_x = \frac{x^T L x}{x^T x} \]

cut ratio \( \leq (2 \Delta \lambda_x)^{1/2} \)

[Cheeger, Alon, Sinclair-Jerrum]
[Mihail]
### Results:

<table>
<thead>
<tr>
<th></th>
<th>$\lambda$</th>
<th>Ratio</th>
<th>Cut Size</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>planar</strong></td>
<td>$O(1/n)$</td>
<td>$O(1/n^{1/2})$</td>
<td>$O(n^{1/2})$</td>
</tr>
<tr>
<td><strong>2D mesh</strong></td>
<td>$O(1/n)$</td>
<td>$O(1/n^{1/2})$</td>
<td>$O(n^{1/2})$</td>
</tr>
<tr>
<td><strong>3D mesh</strong></td>
<td>$O(1/n^{2/3})$</td>
<td>$O(1/n^{1/3})$</td>
<td>$O(n^{2/3})$</td>
</tr>
</tbody>
</table>
Spectral Partitioning
Spectral Methods Always Work

Myth
Bisection may fail: eigenvector tries for good ratio cut
\( \lambda \) small \rightarrow \text{cut of small ratio}

[Cheeger, Alon, Sinclair-Jerrum]

\[ \text{ratio} \sim \sqrt{\lambda} \]
Spectral Embedding

X: Second Eigenvector

Y: Third Eigenvector

Smaller Eigenvalue, better locality
Rayleigh Quotient

$$\lambda_x = \frac{x^T L x}{x^T x} = \frac{\sum_{(i,j) \in E} (x_i - x_j)^2}{\sum x_i^2}$$

$$\lambda_2 = \min_{x \perp 1} \lambda_x$$

Small eigenvalues imply locality
Properties of Laplacian
(Assume G is connected)

- Symmetric
- \( \lambda_1 = 0 \)
- \( \lambda_2, \lambda_3, \ldots, \lambda_n > 0 \)
- \( x^T L x = \sum_{(i,j) \text{ in } E} (x_i - x_j)^2 \)
Laplacian of a Graph

\[
\begin{pmatrix}
1 & -1 & 0 & 0 \\
-1 & 2 & -1 & 0 \\
0 & -1 & 2 & -1 \\
0 & 0 & -1 & 1
\end{pmatrix}
\]

\[L\]
Eigenvalue and Eigenvector

\[ A \mathbf{x} = \lambda \mathbf{x} \]
Partitioning Methods

Local Improvement (Kernighan–Lin)
Multicommodity Flows (Leighton–Rao)
Multilevel (Bui–Jones: Chaco; MeTiS)
Geometric (Miller–Teng–Thurston–Vavasis)
Spectral (Eigenvector–Based) (Donath–Hoffman)
From Sparsest Partition to Bisection

If every subgraph of $G$ of size $x$ has a partition of sparsity

$O(1/x^\alpha)$

then $G$ has a bisection of cut size

$O(n^{1-\alpha})$. 
\[ \phi(x) \rightarrow \int_{1}^{n} \phi(x) \, dx \]

\[ O(1/x^\alpha) \rightarrow O(n^{1-\alpha}) \]
From Sparse Cut to Bisection
Sparsity: Reduce Two Parameters to One

Sparsity = \frac{\text{Cut-Size}}{\text{"Volume" of the Smaller Side}}

Surface-to-Volume Ratio  Isoperimetric Number  Cut-Ratio
# Applications

- VLSI Design
- Parallel Processing
- Scientific Computing
- Information Organization
- Efficient Search Structure
Partition and Bisection

$E(\bar{u}\bar{u})$

Cut-Size
Splitting-Ratio

Bisection

Separator

A
B
C
Graph Partitioning
Spectral Methods

- Eigenvector and Eigenvector
- Underlying Matrices
- Classical Method (70's)
- Many Variations and Software
- Great Experimental Results
- Lack of Mathematical Justification
cut ratio \[ \phi(x) \rightarrow \int_1^n \phi(x) \, dx \]

bisection \[ O(n^{-\alpha}) \rightarrow O(n^{1-\alpha}) \]
Eigenvalue of a Graph

\[ \lambda \begin{pmatrix} -.7 \\ -.3 \\ .3 \\ .7 \end{pmatrix} = \begin{pmatrix} 1 & -1 & 0 & 0 \\ -1 & 2 & -1 & 0 \\ 0 & -1 & 2 & -1 \\ 0 & 0 & -1 & 1 \end{pmatrix} \begin{pmatrix} -.7 \\ -.3 \\ .3 \\ .7 \end{pmatrix} \]

Eigenvalue with min non-zero \( \lambda \)