Expansion and lack thereof in randomly perturbed graphs

Abraham D. Flaxman

Theory Group,
Microsoft Research

March 20, 2007

Outline

1. Introduction
   - Random Graphs
   - Randomly perturbed graphs

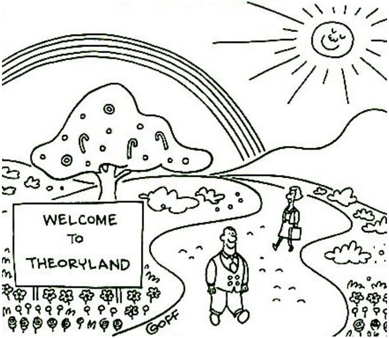
2. Expansion
   - Expansion in the real world
   - Expansion in randomly perturbed graphs

Random Graphs

- Started out as pure math
- Didn't have to answer to experiments

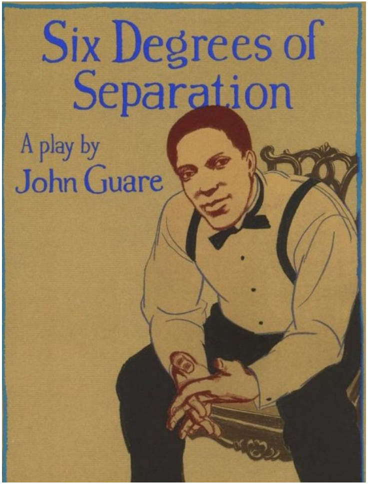
Random Graphs

- Started out as pure math
- Didn't have to answer to experiments



Experiments

Experiments



Theory

Theory

Attempts to keep up with the experiments

Theory

Attempts to keep up with the experiments

- Watts-Strogatz model gets clustering coefficient

Theory

Attempts to keep up with the experiments

- Watts-Strogatz model gets clustering coefficient
  No heavy tail

Theory

Attempts to keep up with the experiments

- Watts-Strogatz model gets clustering coefficient
  No heavy tail
- Barabasi-Albert model gets heavy tail

Theory

Attempts to keep up with the experiments

- Watts-Strogatz model gets clustering coefficient
  No heavy tail
- Barabasi-Albert model gets heavy tail
  Does not densify

Theory

Attempts to keep up with the experiments

- Watts-Strogatz model gets clustering coefficient
  No heavy tail
- Barabasi-Albert model gets heavy tail
  Does not densify
- Next?

Theory

Attempts to keep up with the experiments

- Watts-Strogatz model gets clustering coefficient
  No heavy tail
- Barabasi-Albert model gets heavy tail
  Does not densify
- Next? Dorogovtsev-Mendes model,

Theory

Attempts to keep up with the experiments

- Watts-Strogatz model gets clustering coefficient
  No heavy tail
- Barabasi-Albert model gets heavy tail
  Does not densify
- Next? Dorogovtsev-Mendes model, Kleinberg's Model,

Theory

Attempts to keep up with the experiments

- Watts-Strogatz model gets clustering coefficient
  No heavy tail
- Barabasi-Albert model gets heavy tail
  Does not densify
- Next? Dorogovtsev-Mendes model, Kleinberg's Model, Copying Model,

Theory

Attempts to keep up with the experiments

- Watts-Strogatz model gets clustering coefficient
  No heavy tail
- Barabasi-Albert model gets heavy tail
  Does not densify
- Next? Dorogovtsev-Mendes model, Kleinberg's Model, Copying Model, Random surfer model,

Theory

Attempts to keep up with the experiments

- Watts-Strogatz model gets clustering coefficient
  No heavy tail
- Barabasi-Albert model gets heavy tail
  Does not densify
- Next? Dorogovtsev-Mendes model, Kleinberg's Model, Copying Model, Random surfer model, Forest fire model,

Theory

Attempts to keep up with the experiments

- Watts-Strogatz model gets clustering coefficient
  No heavy tail
- Barabasi-Albert model gets heavy tail
  Does not densify
- Next? Dorogovtsev-Mendes model, Kleinberg's Model, Copying Model, Random surfer model, Forest fire model, Geometric Preferential Attachment,

Theory

Attempts to keep up with the experiments

- Watts-Strogatz model gets clustering coefficient
  No heavy tail
- Barabasi-Albert model gets heavy tail
  Does not densify
- Next? Dorogovtsev-Mendes model, Kleinberg's Model, Copying Model, Random surfer model, Forest fire model, Geometric Preferential Attachment, HOTS,

Theory

Attempts to keep up with the experiments

- Watts-Strogatz model gets clustering coefficient
  No heavy tail
- Barabasi-Albert model gets heavy tail
  Does not densify
- Next? Dorogovtsev-Mendes model, Kleinberg's Model, Copying Model, Random surfer model, Forest fire model, Geometric Preferential Attachment, HOTS, CIPA,

Theory

Attempts to keep up with the experiments

- Watts-Strogatz model gets clustering coefficient
  No heavy tail
- Barabasi-Albert model gets heavy tail
  Does not densify
- Next? Dorogovtsev-Mendes model, Kleinberg's Model, Copying Model, Random surfer model, Forest fire model, Geometric Preferential Attachment, HOTS, CIPA, etc.

Randomly perturbed graphs

Randomly perturbed graphs

Start with a pretty arbitrary graph $\overline{G}$, and perturb it by adding sparse random graph R, to obtain

$$G = \overline{G} + R.$$

Randomly perturbed graphs

Randomly perturbed graphs

Based on

- Smoothed analysis [Spielman and Teng]

Randomly perturbed graphs

Based on

- Smoothed analysis
  [Spielman and Teng]

- Diameter of a cycle plus a random matching
  [Bollobás and Chung]
- How many random edges make a dense graph Hamiltonian?
  [Bohman, Frieze, and Martin]

A proposed approach for real-world graphs

Theorems that hold for

a sufficiently arbitrary graph
and a sufficiently small perturbation

should be
valid predictions for real-world networks.

Example

Theorem

Let $\overline{G}$ be any connected n-graph, and let $R \sim \mathbb{G}_{n,\epsilon/n}$. Then, with high probability, $G = \overline{G} + R$ has

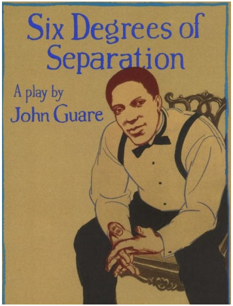
$$\mathrm{diam}(G) = \mathcal{O}\left(\epsilon^{-1} \log n\right).$$

Example

Theorem

Let $\overline{G}$ be any connected n-graph, and let $R \sim \mathbb{G}_{n,\epsilon/n}$. Then, with high probability, $G = \overline{G} + R$ has

$$\operatorname{diam}(G) = \mathcal{O}\left(\epsilon^{-1} \log n\right).$$



A scientific question

Is the randomly perturbed graph a good model for the real world?

Consider property of *Expansion*

Consider property of *Expansion*

- Vertex Expansion: For all $S \subset V$ with $|S| \leq n/2$,

$$|\Gamma(S)| \geq \alpha |S|.$$

Consider property of *Expansion*

- Vertex Expansion: For all $S \subset V$ with $|S| \leq n/2$,

$$|\Gamma(S)| \geq \alpha |S|.$$

- Conductance: For all $S \subset V$ with $2e(S) + e(S, \overline{S}) \leq |E|$,

$$\frac{e(S, \overline{S})}{2e(S) + e(S, \overline{S})} \geq \delta.$$

Consider property of *Expansion*

- Vertex Expansion: For all $S \subset V$ with $|S| \leq n/2$,

$$|\Gamma(S)| \geq \alpha |S|.$$

- Conductance: For all $S \subset V$ with $2e(S) + e(S, \overline{S}) \leq |E|$,

$$\frac{e(S, \overline{S})}{2e(S) + e(S, \overline{S})} \geq \delta.$$

- Eigenvalue gap: For matrix M given by

$$M_{i,j} = \begin{cases} \deg(i), & \text{if } i = j; \\ -1, & \text{if } \{i, j\} \in E; \\ 0, & \text{otherwise}; \end{cases} \qquad \lambda_1(M) \geq \epsilon$$

Consider property of *Expansion*

- Vertex Expansion: For all $S \subset V$ with $|S| \leq n/2$,

$$|\Gamma(S)| \geq \alpha|S|.$$

- Conductance: For all $S \subset V$ with $2e(S) + e(S, \overline{S}) \leq |E|$,

$$\frac{e(S, \overline{S})}{2e(S) + e(S, \overline{S})} \geq \delta.$$

- Eigenvalue gap: For matrix M given by

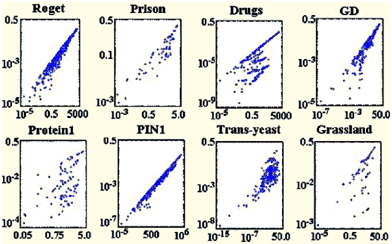
$$M_{i,j} = \begin{cases} \deg(i), & \text{if } i = j; \\ -1, & \text{if } \{i, j\} \in E; \\ 0, & \text{otherwise}; \end{cases} \qquad \lambda_1(M) \geq \epsilon$$

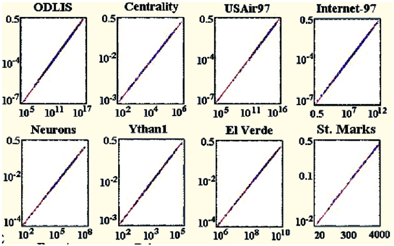
Good to have expansion and good not to have expansion, too.

What sort of expansion should we expect?

What sort of expansion should we expect?

E. Estrada, Spectral scaling and good expansion properties in complex networks, *Europhysics Letters*, 73 (4), pp. 649–655 (2006).





In theory:

In theory:

Depends on the base graph. For $\overline{G}$ connected, and $G = \overline{G} + R$,

In theory:

Depends on the base graph. For $\overline{G}$ connected, and $G = \overline{G} + R$,

Theorem

If $R \sim \mathbb{G}_{n,\epsilon/n}$, then G is not necessarily an expander.

In theory:

Depends on the base graph. For $\overline{G}$ connected, and $G = \overline{G} + R$,

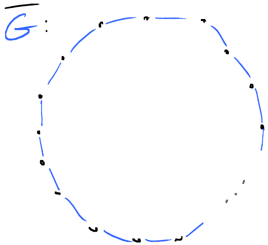
Theorem

If $R \sim \mathbb{G}_{n,\epsilon/n}$, then G is not necessarily an expander.

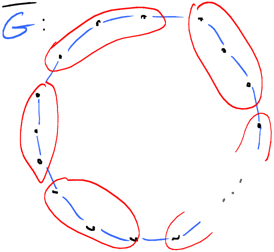
Theorem

If $R \sim \mathbb{G}_{1\text{-}out}$ then G is an expander **whp**.

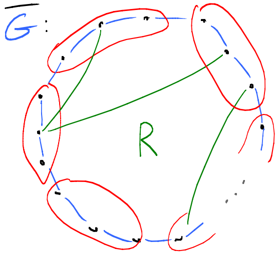
Proof if $R \sim \mathbb{G}_{n,\epsilon/n}$ then G not necess. expander



Proof if $R \sim \mathbb{G}_{n,\epsilon/n}$ then G not necess. expander



Proof if $R \sim \mathbb{G}_{n,\epsilon/n}$ then G not necess. expander



Proof if $R \sim \mathbb{G}_{1\text{-out}}$ then G is expander

Proof if $R \sim \mathbb{G}_{\text{1-out}}$ then G is expander

How do you prove that a k-out is an expander, for large k?

Proof if $R \sim \mathbb{G}_{1\text{-out}}$ then G is expander

How do you prove that a k-out is an expander, for large k?

$$\mathbb{P}\left[\exists S : e(S, \overline{S}) \leq \delta \cdot |S|\right] \leq \sum_{\ell} \binom{n}{\ell} \mathbb{P}\left[e(S, \overline{S}) \leq \delta\ell\right] = o(1).$$

Proof if $R \sim \mathbb{G}_{\text{1-out}}$ then G is expander

How do you prove that a k-out is an expander, for large k?

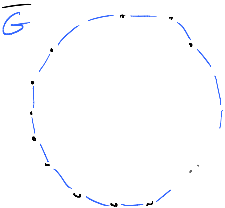
$$\mathbb{P}\left[\exists S : e(S, \overline{S}) \leq \delta \cdot |S|\right] \leq \sum_{\ell} \binom{n}{\ell} \mathbb{P}\left[e(S, \overline{S}) \leq \delta\ell\right] = o(1).$$

This doesn't work unless k is a large enough constant. (And it shouldn't, since it's not true for $k = 1$.)

Proof if $R \sim \mathbb{G}_{\text{1-out}}$ then G is expander

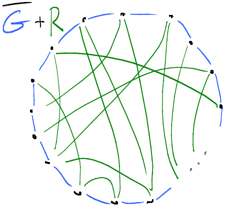
Proof if $R \sim \mathbb{G}_{1\text{-out}}$ then G is expander

What if $\overline{G}$ is a cycle?



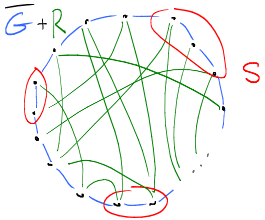
Proof if $R \sim \mathbb{G}_{1\text{-out}}$ then G is expander

What if $\overline{G}$ is a cycle?



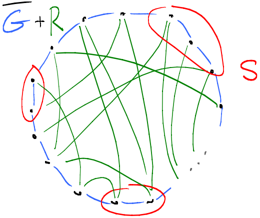
Proof if $R \sim \mathbb{G}_{1\text{-out}}$ then G is expander

What if $\overline{G}$ is a cycle?



Proof if $R \sim \mathbb{G}_{1\text{-out}}$ then G is expander

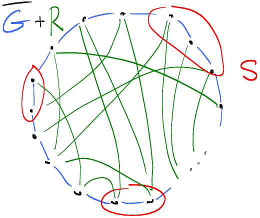
What if $\overline{G}$ is a cycle?



$$\sum_{S:\ |S|=s} \mathbb{P}[e(S,\overline{S}) \leq \delta s] \leq \sum_{k} 2\binom{n}{2k} \mathbb{P}[e_R(S,\overline{S}) \leq \delta s - 2k]$$

Proof if $R \sim \mathbb{G}_{1\text{-out}}$ then G is expander

What if $\overline{G}$ is a cycle?



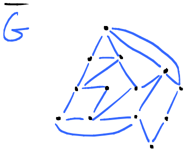
$$\sum_{S:\ |S|=s} \mathbb{P}[e(S,\overline{S}) \leq \delta s] \leq \sum_{k} 2\binom{n}{2k} \mathbb{P}[e_R(S,\overline{S}) \leq \delta s - 2k]$$

✓

Proof if $R \sim \mathbb{G}_{\text{1-out}}$ then G is expander

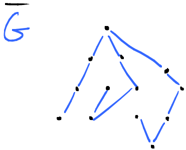
Proof if $R \sim \mathbb{G}_{1\text{-out}}$ then G is expander

For general $\overline{G}$, need something to take the place of the cycle.



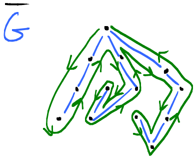
Proof if $R \sim \mathbb{G}_{1\text{-out}}$ then G is expander

For general $\overline{G}$, need something to take the place of the cycle.



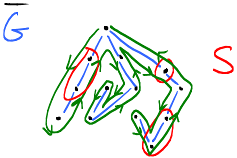
Proof if $R \sim \mathbb{G}_{1\text{-out}}$ then G is expander

For general $\overline{G}$, need something to take the place of the cycle.



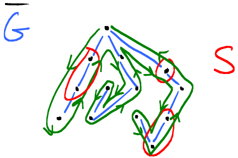
Proof if $R \sim \mathbb{G}_{1\text{-out}}$ then G is expander

For general $\overline{G}$, need something to take the place of the cycle.



Proof if $R \sim \mathbb{G}_{1\text{-out}}$ then G is expander

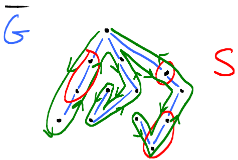
For general $\overline{G}$, need something to take the place of the cycle.



$$\sum_{S:\ |S|=s} \mathbb{P}[e(S,\overline{S}) \leq \delta s] \leq \sum_k 2\binom{2n}{k}\mathbb{P}[e_R(S,\overline{S}) \leq \delta s - k/2]$$

Proof if $R \sim \mathbb{G}_{1\text{-out}}$ then G is expander

For general $\overline{G}$, need something to take the place of the cycle.

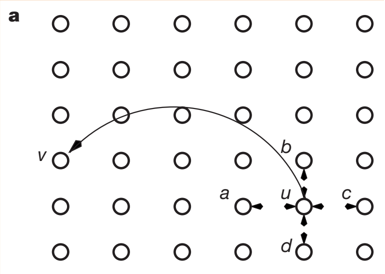


$$\sum_{S:\, |S|=s} \mathbb{P}[e(S,\overline{S}) \leq \delta s] \leq \sum_k 2\binom{2n}{k}\mathbb{P}[e_R(S,\overline{S}) \leq \delta s - k/2]$$ ✓

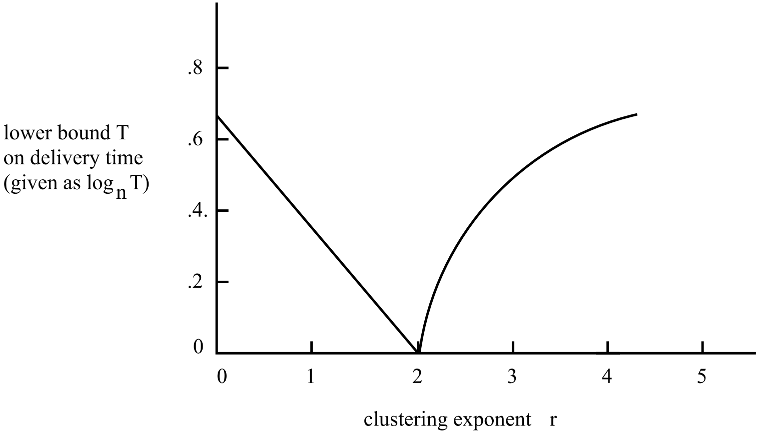
Curious extension of these techniques

Kleinberg's extension of Watts-Strogatz model

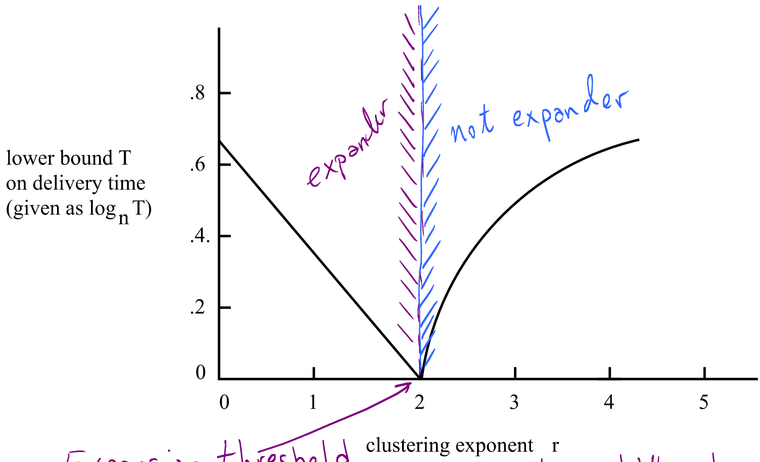
Figure 1 The navigability of small-world networks. **a,** The network model is derived from an $n \times n$ lattice. Each node, *u*, has a short-range connection to its nearest neighbours (*a, b, c* and *d*) and a long-range connection to a randomly chosen node, where node *v* is selected with probability proportional to $r^{-\alpha}$, where *r* is the lattice ('Manhattan') distance between *u* and *v*, and $\alpha \geq 0$ is a fixed clustering exponent. More generally, for $p,q \geq 1$, each node *u* has a short-range connection to all nodes within *p* lattice steps, and *q* long-range connections generated independently from a distribution with clustering exponent α. **b,** Lower bound from my charac-



Curious extension of these techniques



Curious extension of these techniques



Expansion threshold matches "algorithmic small world" value


ENTRANCE
To
MYSTERY SPOT

Conclusion

Conclusion

- Real-world graphs are interesting
- Randomly perturbed graphs can model them
- Doesn't make a prediction on expansion
- This is consistent with the data

Conclusion

- Real-world graphs are interesting
- Randomly perturbed graphs can model them
- Doesn’t make a prediction on expansion
- This is consistent with the data