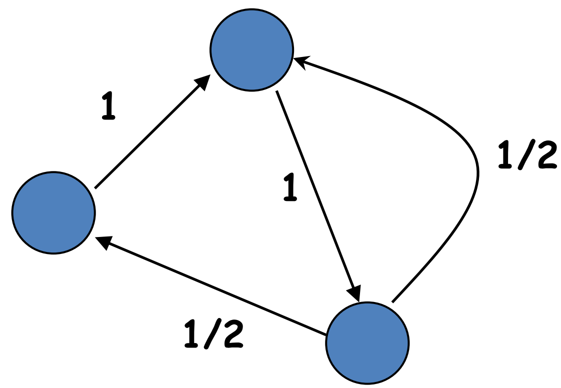
قدم زدن تصادفی
و مجموعه‌های تکمیل‌شونده

ارائه: سپیده آقاملائی

مقاله: چان، کواک، لائو

قدم زدن تصادفی

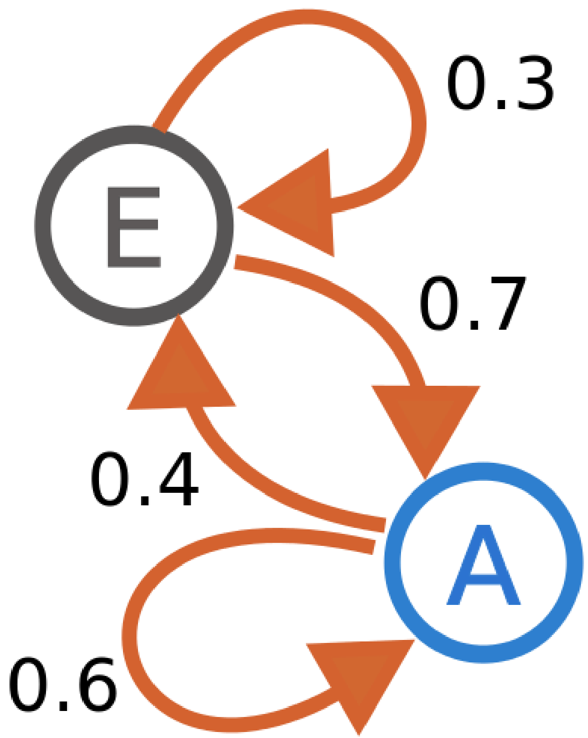
- الگوریتم:
  - با احتمال یکنواخت یکی از همسایه‌هایت را انتخاب کن.
  - تکرار



- متوسط زمان دیدن همه‌ی رأسها در گراف همبند: O(mn)
- Antony Gormley, *Quantum Cloud*, London 2000

زنجیره مارکوف

- قدم زدن تصادفی روی گراف جهت‌دار
  - احتمال دلخواه (غیریکنواخت)



- شرط همگرایی:
  - گراف قویاً همبند باشد.
  - ب.م.م طول دورها ۱ باشد.

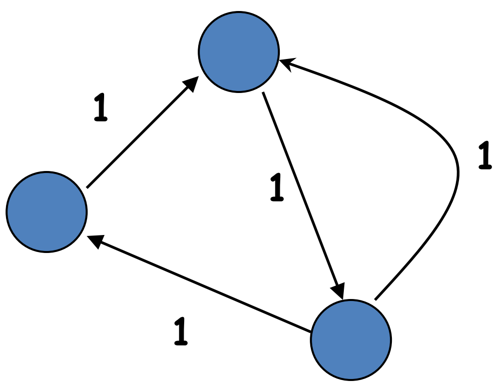
ماتریس قدم‌زدن تصادفی

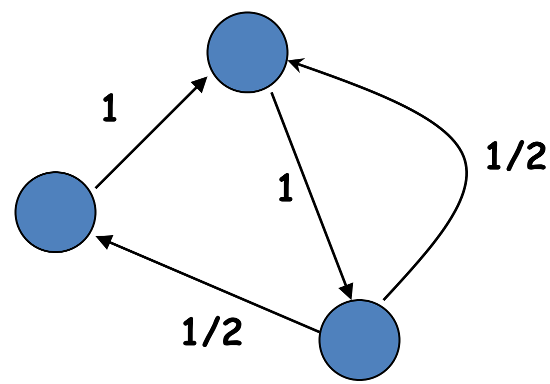
0	1	0
0	0	1
1	1	0

Adjacency matrix A

0	1	0
0	0	1
1/2	1/2	0

Transition matrix P



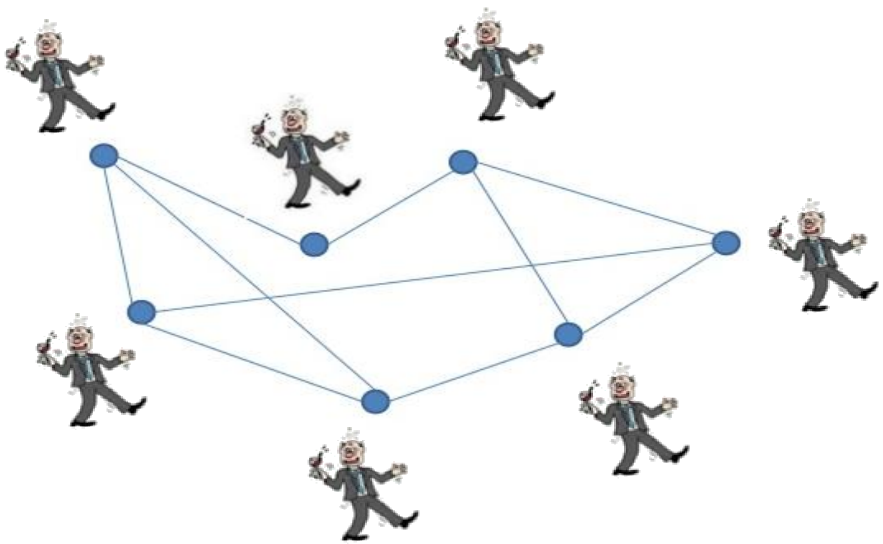


زمان همگرایی (mixing time)

$p^0 = [0\ 0\ \ \dots 0\ 1\ 0 \dots 0]^T$

$p^{t+1} = A^t p^0$

زمان همگرایی:
حداقل زمان لازم برای
اینکه فاصله‌ی p^t از
بردار $[\frac{1}{n}\ \ \dots\ \ \frac{1}{n}]$ از یک
مقدار ثابت کمتر شود.



$$T_{\text{mix}} = \min\{t : \|p_t - \frac{\vec{1}}{n}\|_1 \leqslant 1/4 \text{ for all initial distribution } p_0\}.$$

مقدار ویژه‌های ماتریس قدم زدن تصادفی

$$1 = \mu_1 \geq \ \dots \ \geq \ \mu_n \geq -1 \quad \mu = \max\{\mu_2, |\mu_n|\}$$

- اختلاف مقدار ویژه‌ها (spectral gap) = $1 - \mu$
- کران بالای زمان همگرایی = $O(\frac{\log n}{1-\mu})$

$$T_{mix} = \left\| p^t - \frac{\vec{1}}{n} \right\| \approx O(\sqrt{n}\,\mu^t) = O(1)$$

$$\Rightarrow t = -\frac{\log n}{\log \mu} = \frac{\log n}{\log 1 - \log \mu} \geq \frac{\log n}{1 - \mu}$$

قدم زدن تصادفی تنبل

- نامساوی چیگر:

$$\frac{1}{2}(1-\mu_2) \leq \phi(G) \leq \sqrt{2(1-\mu_2)}$$

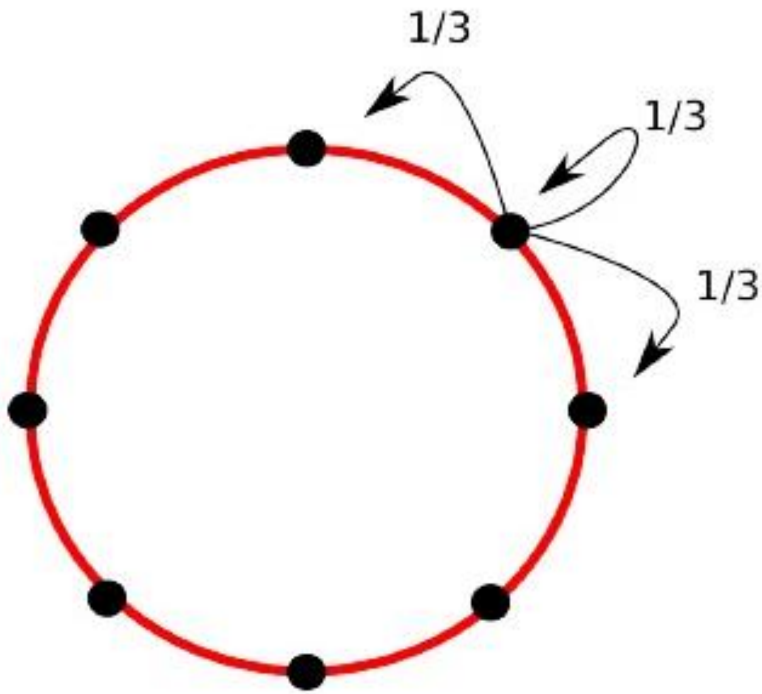
- پس

$$1-\mu_2 = \Omega(\phi^2(G))$$

- حل مشکل μ_n
  - کاری می‌کنیم که همیشه نامنفی باشد.
  - قدم زدن تصادفی تنبل (اضافه کردن طوقه)

چقدر تنبل؟

- مثال: احتمال‌های یکنواخت (قدم زدن تصادفی)
- ایراد:
  - وابستگی به درجه
- راه حل:
  - احتمال ثابت
- بهبود:
  - چه احتمالی؟



قدم زدن تصادفی تنبل

- با احتمال ۱/۲ در همان رأسی که هستی بمان.
- مقدار ویژه‌های ماتریس جدید:

- $A' = \frac{1}{2}A + \frac{1}{2}I$
- $|A - \mu I| = 0$
- $|A' - \mu' I| = \left|\frac{1}{2}A + \frac{1}{2}I - \mu' I\right|$
$= \frac{1}{2}|A - (2\mu' - 1)I| = 0$
- $-1 \leq \mu \leq 1 \Rightarrow \ 0 \leq \mu' \leq 1$

مجموعه‌های تکمیل‌شونده

- زنجیره‌ی مارکوف
- عدد U را به صورت تصادفی یکنواخت از بازه‌ی [0,1] انتخاب کن. مجموعه‌ی بعدی:

$$\bar{S} := \{y : w(y, S) \geq U\}$$

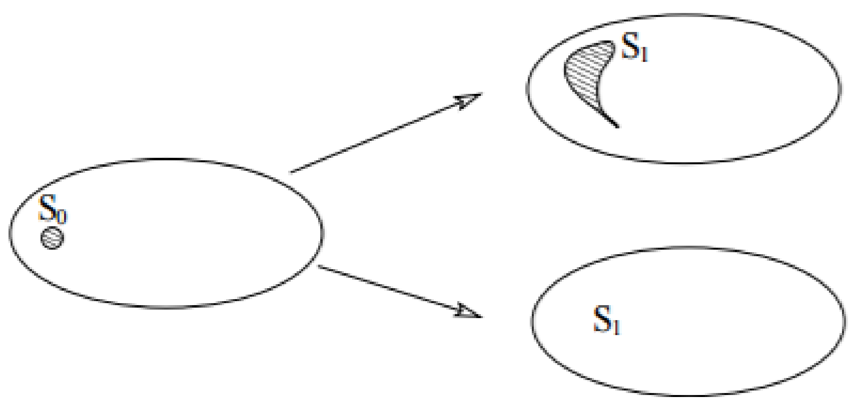


Figure 1: When u is small then the set grows, but when u is big then it shrinks.

مجموعه‌های تکمیل‌شونده

- هر زنجیره مارکوف، متناظر برش روی گراف وزن‌دار با وزن یال‌های متناسب با احتمالها است.

- با فرض اینکه ترتیب دیدن $S_0, S_1, \ldots, S_n$ باشد و pr_A احتمال شروع از A باشد، احتمال قدم‌زدن تصادفی بر حسب مجموعه تکمیل‌شونده:

$$pr(x, y) = \frac{pr(\ vertex = y)}{pr(\ vertex = x)} pr_{\{x\}}(y \in S_n)$$

زمان همگرایی قدم زدن تصادفی

- پیمانه‌ی یک مجموعه و یک گراف

$$\psi(S) := 1 - \mathbb{E}[\frac{|\bar{S}|}{|S|}] \quad \psi(G) := \min_{S:|S| \leq \frac{n}{2}} \psi(S)$$

- زمان همگرایی: در گراف تنبل:

$$O(\frac{\log n}{\psi(G)}) \qquad \psi(G) \geq \Omega(\phi^2(G))$$

اثبات ترکیباتی زمان همگرایی

- به جای نصف رأسها، یک کسر ثابت از آنها را می‌گیریم.
- δ-small set expansion ($0 < \delta \leq \frac{1}{2}$)

$$\phi_\delta(G) \coloneqq \min_{S:|S|\leq\delta n} \phi(S)$$

- زمان همگرایی قدم زدن تصادفی تنبل:

$$T_{mix} \leq \int_{1/n}^{1/2} \frac{d\delta}{\delta \phi_\delta(G)^2}$$

حدس مجموعه‌های کوچک

- به ازای هر اپسیلون، دلتایی وجود دارد که تشخیص دو حالت زیر از هم ان‌پی-سخت است:

1. There is a set S with $\phi(S) \leqslant \epsilon$ and $|S| \leqslant \delta n$;
2. $\phi(S) \geqslant 1 - \epsilon$ for every set S with $|S| \leqslant \delta n$.

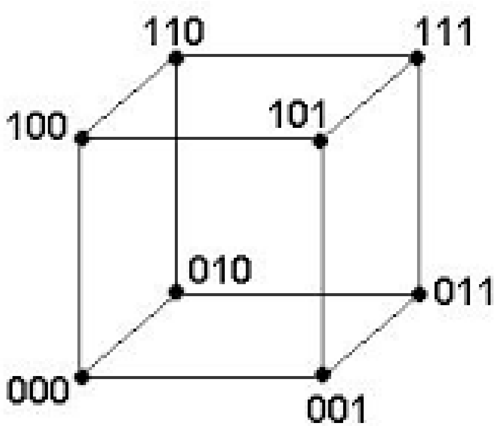
- مسایل مرتبط
  - بهبود زمان پیدا کردن برش چندگانه با قدم‌زدن تصادفی
  - حدس بازی یکتا
  - افراز محلی گراف

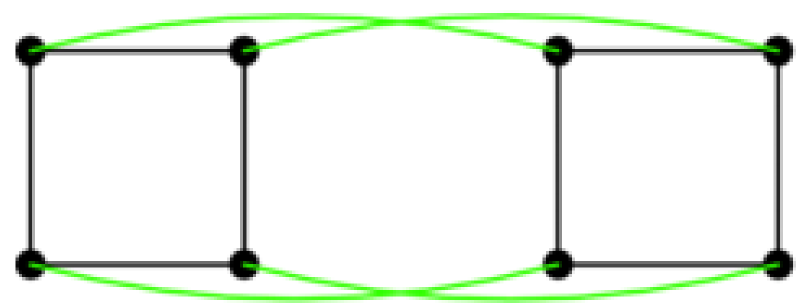
حدس اویس قرن

- روش مبتنی بر مجموعه‌های تکمیل‌شونده
- حدس: همه‌ی مجموعه‌های دیده شده در یک فرایند مجموعه‌ی تکمیل‌شونده اندازه‌ی $O(|S^*|)$ دارند.
- اویس قرن ثابت کرده است که در صورتی که این حدس درست باشد، فرضیه گسترش مجموعه-کوچک رد می‌شود.
- مثال نقض: ابرمکعب نویزی

ابرمکعب (مرور)

- $n=2^d$
- $m=dn/2 = d2^{d-1}$
- گسترش گراف وقتی به دست می‌آید که روی یک بعد ببریم:
- $\phi(G) = \frac{E(S,V-S)}{d|S|} = \frac{2^{d-1}}{d\times 2^{d-1}} = \frac{1}{d}$





گراف همینگ (ابرمکعب غیردودویی)

- حاصل ضرب d گراف کامل k-رأسی
- H(d,k)
- حالت خاص: ابرمکعب=H(d,2)
- تعداد رأسها = k^d
- تعداد یالها = $\frac{k^d \times d \times (k-1)}{۲}$
- (d(k-1))-منتظم
- گسترش یالی = $\frac{k-1}{d}$

ابرمکعب نویزی

- ابرمکعب d-بعدی با الفبای ۱ تا k
- وزن یالها = احتمال رفتن از سر اول به سر دوم
- ساخت گراف:شروع از همه ۰، با احتمال $1-\varepsilon$ یکسان و در غیر این صورت تصادفی یکنواخت
- گراف ۱-منتظم
- برش روی یک بعد = گسترش اپسیلون و اندازه δn
- گسترش مجموعه کوچک حداکثر اپسیلون
- هر مجموعه با اندازه‌ی $\delta_\varepsilon n$ با احتمال ۱ گسترش $1-\varepsilon$ دارد.

نتیجه رد شدن حدس

- تعمیم محدودیت به:
- الگوریتم pagerank
- الگوریتم محلی تقسیم گراف
- الگوریتم heat kernel
- کران پایین برای ضریب تقریب مجموعه‌های کامل‌شونده:
- $O\left(\sqrt{\phi(S)\log(|S|}\right)$
- $\varepsilon = \frac{1}{\log|S|}$

تأثیر روی افراز محلی گراف

- به ازای هر مجموعه‌ی هدف S^* و اپسیلون مثبت، زیرمجموعه‌ی S' با اندازه‌ی حداقل $|S^*|/2$ وجود دارد که اگر از رأس v عضو S فرایند مجموعه‌ی تکمیل‌شونده را اجرا کنیم، با احتمال ثابت به مجموعه‌ی S می‌رسیم که:

$$\phi(S) = O(\sqrt{\phi(S^*)/\epsilon}) \ and \ |S| = O(|S^*|^{1+\epsilon})$$

- زمان الگوریتم:

$$O(|S^*|^{1+2\epsilon}\phi(S)^{-\frac{1}{2}}\log^2 n)$$

چهار اثبات نامساوی چیگر

- graph spectral method — spectral partition algorithm
- random walks
- PageRank
- heat kernel

local partition algorithms

PageRank versus heat kernel

$$p_{\alpha,s} = \alpha \sum_{k=0}^{\infty} (1-\alpha)^k (sW^k)$$

Geometric sum

$$\rho_{t,s} = e^{-t} \sum_{k=0}^{\infty} s \frac{(tW)^k}{k!}$$

Exponential sum

$$p = \alpha + (1-\alpha) pW$$

recurrence

$$\frac{\partial \rho}{\partial t} = -\rho (I - W)$$

Heat equation

منحنی لواژ–سیمینوویتس

- تابع تجمعی p^i

$$C(p,x) = \max_{c\in[0,1]^n : \sum_i c(i) = x} \sum_{i\in V} c(i)\cdot p(i)$$

- تکه‌ای خطی
- شیب = $\rho(e_i)$ = تابعی از رأس مبدأ یال
- مقعر: انتگرال یک تابع نزولی
- همگرایی = رسیدن به خط x/n

منحنی لواژ–سیمینوویتس

- از رابطه زیر t به دست می‌آید:

$$C(A^t p, x) \leqslant \frac{x}{n} + \sqrt{x}(1 - \frac{\phi(G)^2}{8})^t,$$

- اثبات: با اعمال استقرایی نامساوی زیر برای $x \leqslant n/2$
  - نامساوی ینسن (توابع مقعر)

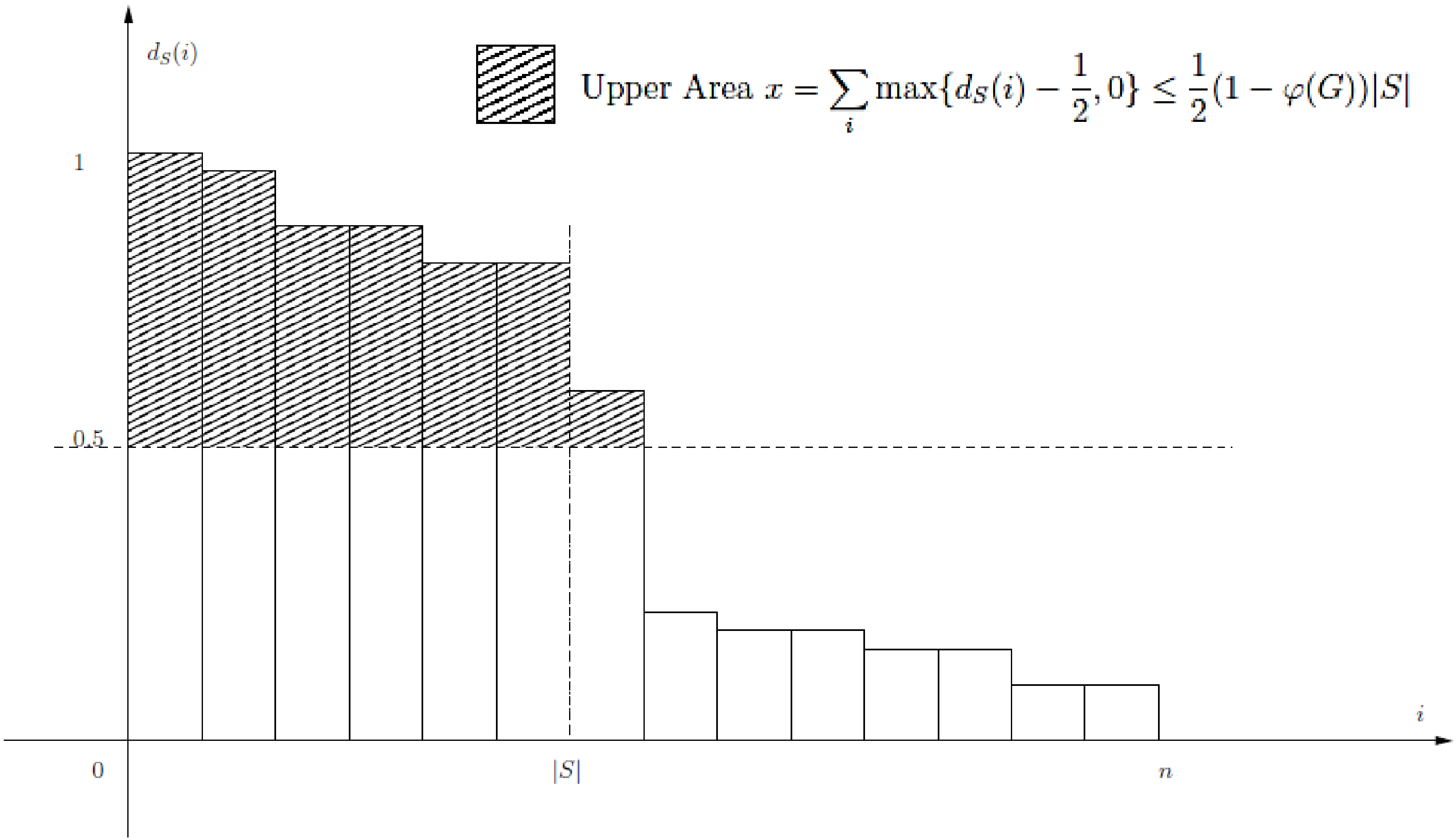
$$C(Ap, x) \leqslant \frac{1}{2}\big(C(p, x(1-\phi(G))) + C(p, x(1+\phi(G)))\big)$$

اثبات نامساوی

- احتمال پس از یک گام از قدم‌زدن تصادفی
- d_S: متغیر مشخصه‌ی حضور رأس i در S
- رأسهای دلخواه S $(Ap)(S) := \sum_{i\in S}(Ap)_i.$
- d_S نزولی مرتب $(Ap)(S) = \sum_{i=1}^{n} d_S(i)\cdot p(i).$
- جمع تلسکوپی (افقی روی نمودار)

$$(Ap)(S) = \sum_{i=1}^{n} d_S(i)\cdot p(i) = \sum_{i=1}^{n}(d_S(i) - d_S(i+1))\sum_{j=1}^{i} p(j)$$
$$\leqslant \sum_{i=1}^{n}(d_S(i) - d_S(i+1)))\cdot C(p,i).$$

منحنی d_S بر حسب i (رأس)



ادامه اثبات

- نامساوی ینسن

$y = \sum_{i=1}^{n} \min\{d_S(i), \frac{1}{2}\}$

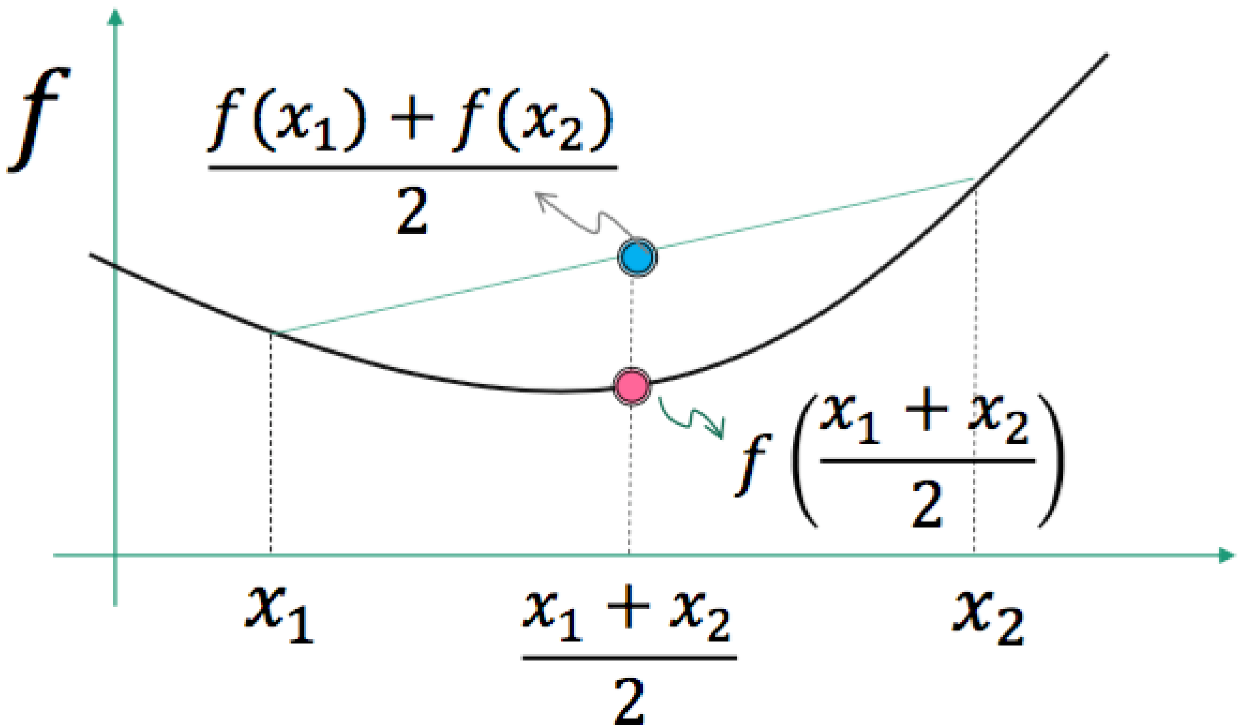
$$
\begin{aligned}
(Ap)(S) &\leqslant \frac{1}{2}(C(p, 2x) + C(p, 2y)) \\
&= \frac{1}{2}(C(p, 2x) + C(p, 2|S| - 2x)).
\end{aligned}
$$

- نتیجه تعریف شکاف طیفی ترکیبیاتی

$$x \leqslant \frac{1}{2}(1 - \varphi(G))|S|.$$

نامساوی ینسن (توابع محدب)

- (مقعر جهت نامساوی برعکس می‌شود.)



بهبود این مقاله

- نسخه‌ی ترکیبیاتی اختلاف مقدار ویژه‌ها
- گسترش یالی:

$$\min_{S\subseteq V,|S|\leqslant n/2} \frac{w(S,V-S)}{|S|} = \min_{S\subseteq V,|S|\leqslant n/2} 1-\frac{w(S,S)}{|S|}.$$

- نسخه‌ی ترکیبیاتی:

$$\varphi(G) := \min_{S\subseteq V,T\subseteq V,|S|=|T|\leqslant n/2} 1-\frac{w(S,T)}{|S|}.$$

- نامساوی

$$w(S,T) \leqslant (1-\varphi(G))|S|$$

بهبود این مقاله

- بدون نیاز به تنبل بودن قدم‌زدن تصادفی، مسئله را حل می‌کنند.
- جایگزین کردن با نسخه ترکیبیاتی

به ازای هر $x > \frac{n}{2}$

$$C(Ap, x) \leqslant \frac{1}{2}\big(C(p, x(1-\varphi(G))) + C(p, x(1+\varphi(G)))\big)$$

توان‌های بالاتر گراف

- طبق نامساوی برای توابع مقعر:
  - قبلاً بر حسب گسترش مجموعه کوچک
  - فقط برای گرافهای تنبل

$$\phi_{\delta/2}(G^t) \geqslant \Omega(\min(\sqrt{t} \cdot \phi_\delta(G), 1)).$$

  - حالا بر حسب نسخه‌ی ترکیبیاتی آن

$$\phi_{\delta/2}(G^t) \geqslant \Omega(\min(\sqrt{t} \cdot \varphi_\delta(G), 1)).$$

توان‌های بالاتر گراف

- بر حسب نسخه‌ی ترکیبیاتی آن

$$\phi_{\delta/2}(G^t) \geqslant \Omega(\min(\sqrt{t} \cdot \varphi_\delta(G), 1)).$$

- نامساوی بین این دو نسخه

$$\varphi_{\delta/2}(G) \geqslant \phi_\delta(G)/2.$$

- نتیجه: تعمیم به گرافهای کلی

$$\phi_{\delta/4}(G^t) \geqslant \Omega(\min(\sqrt{t} \cdot \phi_\delta(G), 1)).$$

- برای یک دور تنبل این نتیجه به طور مجانبی tight است.

تعمیم نتایج به گرافهای کلی

- گسترش

$$\phi(S) := \frac{w(S, V-S)}{\mathrm{vol}(S)} \quad \text{and} \quad \phi(G) := \min_{S:\mathrm{vol}(S)\leqslant \mathrm{vol}(V)/2} \phi(S).$$

- شکاف ترکیبیاتی

$$\varphi(G) := \min_{S,T:\mathrm{vol}(S)=\mathrm{vol}(T)\leqslant \mathrm{vol}(V)/2} 1 - \frac{w(S,T)}{\mathrm{vol}(S)}.$$

- در حالتی که همه حجم‌های متمایز داشته باشند:
  - d بردار درجه گراف

$$\varphi(G) := \min_{\chi_S\in[0,1]^V, \chi_T\in[0,1]^V : \langle \chi_S, \vec{d}\rangle = \langle \chi_T, \vec{d}\rangle \leqslant \mathrm{vol}(V)/2} 1 - \frac{\langle \chi_S, A\chi_T\rangle}{\langle \chi_S, \vec{d}\rangle}.$$

تعمیم نتایج به گرافهای کلی (ادامه)

- منحنی لواژ-سیمینوویتس

$$C(p) : \mathrm{vol}(V) \to \mathbb{R}$$

$$C(p, x) := \max_{c \in [0,1]^n : \langle c, \vec{d} \rangle = x} \langle c, p \rangle.$$

- نقاط مرزی منحنی C(p) برای p نزولی مرتب‌شده

$$\sum_{j=1}^{i} \deg(j)$$

تعمیم نتایج به گرافهای کلی (ادامه)

- منحنی میله‌ای
  - مرتب‌سازی روی d_S/deg
  - عرض هر میله deg
  - ارتفاع هر میله d_S/deg
  - عرض کل = $\mathrm{vol}(V)$

$$(Ap)(S) \leqslant \frac{1}{2}(C(p, 2x) + C(p, 2y)).$$

مسایل حل‌نشده

- گراف داده شده به طور مجانبی بهترین بود، تعمیم به حالت کلی حل نشده است. (خود مقاله)

پایان