

سرعت پخش شایعات در شبکه‌های اجتماعی

عبّاس محرابیان*

دانشگاه واترلو، دانشکده ترکیبیات و بهینه‌سازی

۳ اسفند ۱۳۹۳

چکیده

یک شبکه اجتماعی را می‌توان به صورت یک گراف مدل کرد که رئوس آن اشخاص هستند و یال‌های آن روابط دوستی را نشان می‌دهند. در این مقاله چند تا از مدل‌های پخش شایعات در شبکه‌های اجتماعی^۱ را معرفی می‌کنیم و مهم‌ترین نتایج مربوط به سرعت پخش شایعات را بیان می‌کنیم.

۱ پروتکل ردگن‌بره

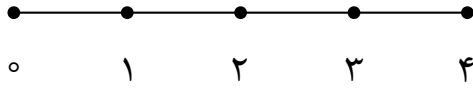
در این مقاله، فقط با گراف‌های بدون جهت، ساده و همبند سروکار خواهیم داشت و نماد n همواره تعداد رئوس را نشان می‌دهد. اولین پروتکلی که بررسی می‌کنیم، پروتکل ردگن‌بره^۲ نام دارد که در زیر تعریف می‌شود.

تعریف. فرض کنید G گرافی بدون جهت، ساده و همبند باشد. در ابتدا یکی از رئوس گراف مانند v خبری را می‌داند. در هر یک از گام‌های بعدی $t = 1, 2, \dots$ ، هر رأسی که خبر را می‌داند، یکی از همسایه‌های خود را به صورت تصادفی و با توزیع یکنواخت انتخاب می‌کند، و خبر را برای او ارسال می‌کند. اولین زمانی که همه رئوس باخبر می‌شوند را زمان پخش^۳ می‌نامیم و آن را با $ST(G, v)$ نشان می‌دهیم، که یک متغیر تصادفی است.

دقت کنید که بنابر تعریف بالا، یک رأس ممکن است دو بار متوالی خبر را برای یک همسایه خاص ارسال کند. به علاوه، اگر رأسی درست در زمان t از خبر آگاه شد، پخش کردن آن را از زمان $t + 1$ شروع می‌کند.

دو مثال ساده ببینیم.

abbas.mehrabian@gmail.com*
rumor spreading in social networks^۱
the push protocol^۲
spread time^۳



شکل ۱: گراف P_5

مثال ۱.۱ (گراف مسیر). فرض کنید P_n گراف مسیر با رئوس $0, 1, \dots, n-1$ باشد و در ابتدا رأس 0 خبر را بداند (شکل ۱ را ببینید). می‌خواهیم امید ریاضی^۴ زمان پخش را محاسبه کنیم. در زمان $t = 1$ رأس 0 خبر را به تنها همسایه خود یعنی رأس 1 می‌فرستد. حال رأس 1 خبر را دارد و دو همسایه دارد. می‌توان فرض کرد که این رأس یک سکه دارد، و در هر گام آن را می‌اندازد: اگر شیر آمد خبر را برای رأس 0 می‌فرستد و اگر خط آمد آن را برای رأس 2 می‌فرستد. بنابراین اگر زمانی را که طول می‌کشد تا این رأس خبر را برای رأس 2 بفرستد با X_1 نشان دهیم، برای هر $k = 1, 2, \dots$ داریم

$$\mathbb{P}[X_1 = k] = \left(\frac{1}{2}\right)^k.$$

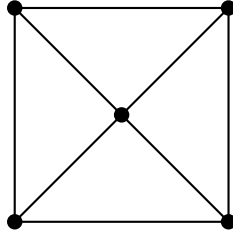
به عبارت دیگر، X_1 متغیری تصادفی با توزیع هندسی با پارامتر $\frac{1}{2}$ است. با همین استدلال می‌توان نشان داد که اگر رأس 2 خبر را برای رأس 3 پس از گذشت زمان X_2 بفرستد، X_2 متغیری تصادفی با توزیع هندسی با پارامتر $\frac{1}{2}$ است. به علاوه، دو متغیر X_1 و X_2 از هم مستقلند. با تکرار این استدلال و استفاده از استقرا، و با توجه به این که هنگامی که رأس $n-1$ آگاه شود همه رئوس آگاه شده‌اند، داریم

$$ST(P_n, 0) = 1 + X_1 + X_2 + \dots + X_{n-2},$$

که X_1, X_2, \dots, X_{n-2} متغیرهای مستقل هندسی با پارامتر $\frac{1}{2}$ هستند. چون امید ریاضی چنین متغیری 2 است و به خاطر خطی بودن امید ریاضی، $\mathbb{E}[ST(P_n, 0)] = 2n - 3$. از طرف دیگر، طبق قانون اعداد بزرگ، برای n های بزرگ $ST(P_n, 0)$ بسیار به مقدار متوسط خود یعنی $2n - 3$ نزدیک است.

مثال ۲.۱ (گراف ستاره). فرض کنید S_n گراف ستاره با n رأس $0, 1, \dots, n-1$ باشد: در این گراف رأس 0 به تمام رئوس دیگر متصل است، و هیچ یال دیگری وجود ندارد. فرض کنید در ابتدا رأس 0 خبر را بداند، و می‌خواهیم امید ریاضی زمان پخش را محاسبه کنیم. در گام اول رأس 0 خبر را به یکی از رئوس دیگر می‌رساند. حال رأس 0 یک همسایه دارد که خبر را می‌داند، و $n-2$ همسایه دارد که خبر را نمی‌دانند. تعداد گام‌هایی که طول می‌کشد رأس 0 خبر را به یک رأس بی‌خبر برساند، یک متغیر تصادفی هندسی با پارامتر $\frac{n-2}{n-1}$ است. به طور کلی، اگر k تا از همسایه‌های رأس 0 خبر را بدانند ($0 \leq k \leq n-2$) و $1-k$ تا بی‌خبر باشند، آن‌گاه تعداد گام‌هایی که طول می‌کشد که رأس تازه‌ای با خبر شود، متغیری هندسی با پارامتر $\frac{n-k-1}{n-1}$ است. به علاوه، این متغیرهای تصادفی از هم مستقلند. چون امید ریاضی چنین

^۴expected value



شکل ۲: یک گراف کامل با ۵ رأس

متغیری $\frac{n-1}{n-k-1}$ است، داریم

$$\mathbb{E} [\text{ST}(S_n, \circ)] = \sum_{k=\circ}^{n-2} \frac{n-1}{n-k-1} = (n-1) \sum_{k=1}^{n-1} \frac{1}{k} \sim n \ln n.$$

در این جا و در ادامه مقاله، برای دو تابع $f(n), g(n)$ ، منظور از $f(n) \sim g(n)$ این است که

$$\lim_{n \rightarrow \infty} \frac{f(n)}{g(n)} = \lim_{n \rightarrow \infty} \frac{g(n)}{f(n)} = 1.$$

از قانون اعداد بزرگ نتیجه می شود که برای n های بزرگ، $\text{ST}(S_n, \circ)$ بسیار به مقدار متوسط خود نزدیک است.

یک گراف کامل، گرافی است که در آن همه رئوس به هم متصل باشند (شکل ۲ را ببینید). در سال ۱۹۸۷ رفتار این پروتکل در یک گراف کامل تحلیل شد و قضیه زیر ثابت شد.

قضیه ۳.۱ ([۴۱]). فرض کنید K_n یک گراف کامل n رأسی باشد و v یکی از رئوس آن باشد. در این صورت داریم $\mathbb{E} [\text{ST}(K_n, v)] \sim \log_2 n + \ln n$ و برای n های بزرگ، متغیر تصادفی $\text{ST}(K_n, v)$ به مقدار متوسط خود نزدیک است.

۱.۱ کران های عمومی. زمان پخش پروتکل ردکن بره را روی سه گراف پایه بیان کردیم. دیدیم که این زمان می تواند بر حسب تعداد رئوس گراف لگاریتمی باشد، خطی باشد، یا از مرتبه $n \ln n$ باشد. می توان پرسید که حداقل و حداکثر زمان پخش روی یک گراف n رأسی چقدر است؟ قضیه زیر بیان می کند که چنانچه ضرایب ثابت مستقل از n برایمان مهم نباشد، گراف های پایه ای که بررسی کردیم گراف های حدی هستند. برای یک گراف داده شده، فاصله دو رأس را برابر تعداد یال های کوتاه ترین مسیری که آن دو رأس را به هم متصل می کند تعریف می کنیم، و بیشترین فاصله ممکن بین دو رأس G را قطر G می نامیم و با $\text{diam}(G)$ نشان می دهیم.

قضیه ۴.۱ ([۲۰]). الف) برای هر گراف n رأسی G و هر رأس v از آن، قطعاً داریم

$$\max \left\{ \log_2 n, \frac{1}{4} \text{diam}(G) \right\} \leq \text{ST}(G, v).$$

(ب) فرض کنید G گرافی n رأسی و v رأسی دلخواه از آن باشد. در این صورت داریم

$$\mathbb{P}[\text{ST}(G, v) \geq \lceil 2n \log_2 n \rceil] \leq \frac{1}{n}.$$

اثبات قسمت الف). با انجام هر گام تعداد رئوسی که خبر را می‌دانند حداکثر دو برابر می‌شود، پس $\log_2 n \leq \text{ST}(G, v)$. فرض کنید u رأسی از G باشد که بیشترین فاصله را از v دارد. این فاصله را با D نشان می‌دهیم. در این صورت، لااقل D گام برای انتقال خبر از v به u لازم است، در نتیجه $D \leq \text{ST}(G, v)$. از طرف دیگر برای هر دو رأس دلخواه x و y از گراف، مسیری از x به v به طول حداکثر D و مسیری از v به y به طول حداکثر D وجود دارد. در نتیجه فاصله بین x و y در گراف از $2D$ بیشتر نیست. با توجه به تعریف قطر، داریم $\text{diam}(G) \leq 2D$. لذا، $\frac{1}{4} \text{diam}(G) \leq D \leq \text{ST}(G, v)$. \square

برای اثبات قسمت (ب) قضیه بالا نیاز به دو گزاره احتمالاتی زیر داریم.

گزاره ۵.۱ (کران اجتماع^۵). فرض کنید m عددی طبیعی و A_1, A_2, \dots, A_m چند پیشامد باشند. در این صورت داریم

$$\mathbb{P}[A_1 \cup A_2 \cup \dots \cup A_m] \leq \mathbb{P}[A_1] + \mathbb{P}[A_2] + \dots + \mathbb{P}[A_m].$$

گزاره ۶.۱ (نامساوی مارکوف^۶). فرض کنید X یک متغیر تصادفی نامنفی و a عددی مثبت باشد. در این صورت داریم

$$\mathbb{P}[X \geq a] \leq \mathbb{E}[X] / a.$$

اثبات قضیه ۴.۱ قسمت ب). فرض کنید u رأسی دلخواه از گراف باشد، و X_u زمانی باشد که رأس u خبر را می‌فهمد. فرض کنید x_0, x_1, \dots, x_k یک کوتاه‌ترین مسیر از v به u باشد، و $x_0 = v$ و $x_k = u$. دقت کنید که هر رأس w از گراف، حداکثر سه تا از همسایه‌هایش در این مسیر قرار دارند. چرا که اگر مثلاً w با چهار رأس x_2, x_3, x_4, x_5 در این مسیر مجاور باشد، می‌توان قسمت x_2, x_3, x_4, x_5 را با مسیر کوتاه‌تر x_2, w, x_5 جایگزین کرد و از طول مسیر کم کرد، که این متناقض با کوتاه‌ترین بودن x_0, \dots, x_k است. نتیجه می‌گیریم که اگر همه همسایه‌های رئوس x_0, \dots, x_k را فهرست کنیم، هر رأس گراف را حداکثر سه بار آورده‌ایم، پس مجموع درجات رئوس x_0, x_1, \dots, x_k حداکثر $3n$ است:

$$\deg(x_0) + \deg(x_1) + \dots + \deg(x_k) \leq 3n.$$

از طرف دیگر، برای هر $0 \leq i \leq k-1$ ، وقتی x_i خبر را فهمید، مدت زمانی که طول می‌کشد خبر را برای رأس x_{i+1} بفرستد، یک متغیر تصادفی هندسی با پارامتر $\deg(x_i)^{-1}$ است (مثال ۱.۱ را به خاطر بیاورید). امید ریاضی این متغیر تصادفی $\deg(x_i)$ است، لذا به دلیل خطی بودن امید ریاضی، داریم

$$\mathbb{E}[X_u] \leq \deg(x_0) + \deg(x_1) + \dots + \deg(x_{k-1}).$$

the union bound^۵
Markov's inequality^۶

دقت کنید که در این جا نمی توان تساوی قرار داد، چون ممکن است مسیره های دیگری هم از v به u وجود داشته باشند، و ما فقط مدتی که طول می کشد خبر از طریق یک مسیر خاص منتقل شود را تحلیل کردیم.

از ترکیب نامساوی های بالا نامساوی $\mathbb{E}[X_u] \leq 3n$ به دست می آید. نامساوی مارکوف نتیجه می دهد

$$\mathbb{P}[X_u \geq 6n] \leq \frac{1}{4}.$$

پس از گذشت زمان $6n$ ، به احتمال لااقل یک دوم رأس u خبر را می داند. بنابراین پس از $12n$ گام، به احتمال لااقل سه چهارم رأس u خبر را می داند، زیرا می توان این $12n$ گام را به صورت دو بخش $6n$ گامی تصور کرد، و احتمال این که نه در بخش اول رأس u مطلع شود و نه در بخش دوم، حداکثر یک چهارم است. به همین ترتیب، داریم:

$$\mathbb{P}[X_u \geq (2 \lceil \log_2 n \rceil) \times 6n] \leq \left(\frac{1}{4}\right)^{\lceil \log_2 n \rceil} = n^{-2}.$$

به کمک کران اجتماع داریم

$$\begin{aligned} \mathbb{P}[\text{ST}(G, v) \geq 12n \lceil \log_2 n \rceil] &= \mathbb{P}\left[\bigcup_{w \in V(G)} X_w \geq 12n \lceil \log_2 n \rceil\right] \\ &\leq \sum_{w \in V(G)} \mathbb{P}[X_w \geq 12n \lceil \log_2 n \rceil] \leq n \times n^{-2} = \frac{1}{n}. \quad \square \end{aligned}$$

تعریف. برای یک گراف n رأسی G ، زمان تضمینی پخش^v را کوچک ترین عددی مثل τ می گیریم به طوری که بعد از τ گام، احتمال این که همه^v رؤوس خبر را بدانند لااقل $1 - \frac{1}{n}$ باشد، مستقل از این که رأس شروع کننده چه باشد. این عدد که یک پارامتر گراف G است را با $\text{gst}(G)$ نشان می دهیم. به زبان ریاضی،

$$\text{gst}(G) := \inf \left\{ \tau : \max_{v \in V(G)} \{\mathbb{P}[\text{ST}(G, v) > \tau]\} \leq \frac{1}{n} \right\}.$$

بنابراین قضیه^v 4.1 بیان می کند که برای هر گراف G داریم

$$\max \left\{ \log_2 n, \frac{1}{4} \text{diam}(G) \right\} \leq \text{gst}(G) \leq 12n \lceil \log_2 n \rceil.$$

کران های بالا و پایین بهتری در مقاله^v [19] اثبات شده اند.

قضیه^v پسین برای بسیاری از گراف ها کران بالای بهتری می دهد. ایده اثبات آن شبیه قضیه^v پیشین است، ولی در اثبات آن از قضیه^v چرنوف^a استفاده می شود.

قضیه^v 7.1 ([20]). فرض کنید G یک گراف باشد و Δ بزرگ ترین درجه^v رؤوسش باشد. در این صورت،

$$\text{gst}(G) \leq 6\Delta(\text{diam}(G) + \log_2 n).$$

^v guaranteed spread time
^a the Chernoff bound

۲.۱ کمی درباره تاریخچه و کاربردها. پیش از ادامه کار خوب است کمی درباره تاریخچه و کاربردهای این پروتکل بدانیم. اولین مقاله‌ای که این پروتکل را به صورت ریاضی تحلیل کرد در سال ۱۹۸۷ منتشر شد [۴۱].

در همان سال و به صورت مستقل، این پروتکل در یک کنفرانس علوم کامپیوتری معرفی شد [۹]. یک پایگاه داده را در نظر بگیرید که چند نسخه همانند از آن وجود دارد که از طریق یک شبکه به هم وصلند^۹. هنگامی که کاربری یکی از نسخه‌ها را تغییر می‌دهد، این تغییر باید در تمام نسخه‌ها اعمال شود. برای این کار می‌توان از این پروتکل یا پروتکل‌های مشابه استفاده کرد. برای جزئیات بیشتر [۹] را ببینید.

سه سال بعد، کاربرد دیگری از این پروتکل مطرح شد [۲۰]. یک عملیات پردازشی را در نظر بگیرید که انجام آن بین چند پردازنده پخش شده است^{۱۰}. این پردازنده‌ها یک شبکه تشکیل می‌دهند که یال‌های آن، کانال‌های ارتباطی بین پردازنده‌ها را نشان می‌دهند. اگر پردازنده‌ای بخواهد اطلاعاتی را برای همه پردازنده‌های دیگر بفرستد، می‌تواند از این پروتکل استفاده کند. این پروتکل سه خوبی دارد: (۱) ساده است (هر پردازنده کافی است فقط همسایه‌های خودش را بشناسد و لازم نیست اطلاعاتی در مورد ساختار شبکه داشته باشد)، (۲) با بزرگ شدن شبکه، الگوریتم تغییر نمی‌کند، و (۳) در برابر برخی از خطاها مانند خرابی پردازنده‌ها و لینک‌ها تا حدی مقاوم است. برای جزئیات بیشتر [۲۰] را ببینید.

طی ده سال اخیر از چنین پروتکل‌هایی برای مدل‌سازی پخش اخبار در شبکه‌های اجتماعی نیز استفاده شده است، مثلاً [۱۳] را ببینید. شایان ذکر است که سال‌ها پیش از آن که علوم کامپیوتردانان به این پروتکل علاقمند شوند، زیست‌شناسان پروتکل‌های مشابهی برای مدل‌سازی پخش بیماری‌ها^{۱۱} در شبکه‌های انسانی تعریف کرده بودند؛ مثلاً [۱۵] را ببینید.

۳.۱ کران‌هایی برای گراف‌های خاص. قضایای ۴.۱ و ۷.۱ برای کلیه گراف‌ها درست هستند. در گراف‌های خاص کران‌های بالا و پایینی مختلفی برای زمان تضمینی پخش اثبات شده است که در ادامه برخی از آن‌ها را می‌آوریم. فرض کنید d یک عدد طبیعی باشد. گراف ابرمکعب d -بعدی^{۱۲} که به طریق زیر تعریف می‌شود را با \mathcal{H}_d نشان می‌دهیم: مجموعه رئوس \mathcal{H}_d همه رشته‌های به طول d از 0 و 1 هستند، و دو رأس آن به هم وصلند اگر تنها در یک مؤلفه با هم تفاوت داشته باشند. این گراف $n = 2^d$ رأس دارد و درجه هر رأس آن d است، و قطر آن هم d است.

قضیه ۸.۱ ([۲۰]). اعداد ثابت $0 < C$ و C' وجود دارند به طوری که برای هر d ,

$$Cd \leq \text{gst}(\mathcal{H}_d) \leq C'd.$$

دقت کنید که $d = \log_2 n$ ، بنابراین قضیه بالا بیان می‌کند که زمان پخش این گراف عملاً لگاریتمی است، با این که تعداد یال‌هایش خیلی زیاد نیست.

a replicated database^۹
a distributed computation^{۱۰}
epidemics^{۱۱}
the d -dimensional hypercube^{۱۲}

یک متحرک تصادفی^{۱۳}، در واقع یک زنجیر مارکوف^{۱۴} است که مجموعه حالت‌های^{۱۵} آن همان مجموعه رئوس گراف است. این متحرک از رأسی از گراف شروع به حرکت می‌کند و در هر گام به صورت تصادفی به یکی از رئوس مجاور می‌رود که با توزیع یکنواخت از بین همه همسایه‌ها انتخاب شده است. در مقاله [۱۸] ارتباطات جالبی بین زمان پخش خبر در یک گراف و زمانی که طول می‌کشد یک متحرک تصادفی همه رئوس آن گراف را ببیند اثبات شده است. در زیر یکی از نتایج این مقاله را می‌آوریم.

قضیه ۹.۱ ([۱۸]). فرض کنید G گرافی با n رأس و m یال باشد و δ و Δ به ترتیب کم‌ترین و بیش‌ترین درجات رئوس باشند. فرض کنید $\text{cov}(G)$ امید ریاضی زمانی باشد که طول می‌کشد تا یک متحرک تصادفی کلیه رئوس گراف را ببیند، با فرض این که رأس آغازین طوری انتخاب شده باشد که این امید ریاضی بیشینه شود. به علاوه، فرض کنید $\text{ast}(G)$ امید ریاضی زمان پخش پروتکل ردکن‌بره باشد، با فرض این که رأس آغازین طوری انتخاب شده باشد که این امید ریاضی بیشینه شود. در این صورت، اعداد ثابت C و C' وجود دارند که

$$C \frac{\sqrt{n \ln n}}{\Delta} \leq \frac{\text{cov}(G)}{\text{ast}(G)} \leq C' \frac{m \ln n}{\delta}.$$

۲ پروتکل بده‌بستان

پروتکل بده‌بستان^{۱۶} به صورت زیر تعریف می‌شود.

تعریف. فرض کنید G گرافی بدون جهت، ساده و همبند باشد. در ابتدا یکی از رئوس گراف مانند v خبری را می‌داند. در هر یک از گام‌های بعدی $t = 1, 2, \dots$ ، هر رأسی که خبر را می‌داند، یکی از همسایه‌های خود را به صورت تصادفی انتخاب می‌کند، و خبر را برای او ارسال می‌کند؛ و هر رأسی که خبر را نمی‌داند، یکی از همسایه‌های خود را به صورت تصادفی انتخاب می‌کند و خبر را از او می‌پرسد (اگر همسایه خبر را بداند به او اطلاع می‌دهد وگرنه انفاقی نمی‌افتد). اولین زمانی که همه رئوس باخبر می‌شوند را با $\text{ST}_{\text{pp}}(G, v)$ نشان می‌دهیم، که یک متغیر تصادفی است.

این پروتکل در مقاله [۹] معرفی شد و نخستین بار در [۳۶] به صورت ریاضی تحلیل شد. دو مثال ساده ببینیم.

مثال ۱.۲ (گراف ستاره). فرض کنید S_n گراف ستاره باشد که رأس مرکزی آن \circ و بقیه رئوس آن $1, 2, \dots, n-1$ نام‌گذاری شده‌اند (شکل ۳ را ببینید). اگر در ابتدا رأس مرکزی خبر را بدانند، در همان اولین گام همه رئوس خبر را از او می‌پرسند و باخبر می‌شوند. اگر در ابتدا رأس دیگری خبر را بدانند، در گام اول خبر را به رأس مرکزی می‌گویند، و در گام دوم سایر رئوس خبر را از رأس مرکزی می‌پرسند. در نتیجه برای هر i بین 1 و $n-1$ داریم $\text{ST}_{\text{pp}}(S_n, \circ) = 1$ و $\text{ST}_{\text{pp}}(S_n, i) = 2$. پس زمان پخش برای پروتکل بده‌بستان در گراف ستاره یک متغیر قطعی^{۱۷} است.

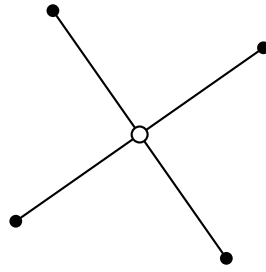
^{۱۳}random walk

^{۱۴}Markov chain

^{۱۵}state space

^{۱۶}the push&pull protocol

^{۱۷}deterministic



شکل ۳: گراف S_5

مثال ۲.۲ (گراف مسیر). فرض کنید P_n گراف مسیر بارئوس $1, \dots, n-1, \circ$ باشد و در ابتدا رأس \circ خبر را بداند. می‌خواهیم امید ریاضی زمان پخش را محاسبه کنیم. در گام نخست رأس \circ خبر را به تنها همسایه خود یعنی رأس ۱ می‌فرستد. فرض کنیم X_2 مدت زمانی باشد که طول می‌کشد تا رأس ۲ با خبر شود. در هر یک از گام‌های بعدی، رأس ۱ خبر را به احتمال یک‌دوم به رأس ۲ می‌فرستد، و رأس ۲ به احتمال یک‌دوم خبر را از رأس ۱ می‌پرسد. بنابراین در هر یک از این گام‌ها، احتمال این که رأس ۲ با خبر نشود یک‌چهارم است. چون این گام‌ها مستقلند، برای هر $k = 1, 2, \dots$ داریم

$$\mathbb{P}[X_1 = k] = \frac{3}{4} \left(\frac{1}{4}\right)^{k-1}.$$

به عبارت دیگر، X_2 متغیری هندسی با پارامتر $\frac{3}{4}$ است. با همین استدلال می‌توان نشان داد که اگر از لحظه‌ای که رأس ۲ خبر را فهمید، مدت زمان X_3 طول بکشد تا رأس ۳ خبر را بفهمد، X_3 متغیری هندسی با پارامتر $\frac{3}{4}$ است. به علاوه، دو متغیر X_2 و X_3 از هم مستقلند. با تکرار این استدلال و استفاده از استقرا، و با توجه به این که هنگامی که رأس $n-1$ آگاه شود همه رئوس آگاه شده‌اند، داریم

$$ST_{pp}(P_n, \circ) = 1 + X_2 + X_3 + \dots + X_{n-2} + 1,$$

که X_2, X_3, \dots, X_{n-2} متغیرهای مستقل هندسی با پارامتر $\frac{3}{4}$ هستند. (دقت کنید که هر زمانی که رأس $n-2$ با خبر شود، رأس $n-1$ در گام بعدی خبر را از او می‌پرسد و با خبر می‌شود.) چون امید ریاضی چنین متغیری $\frac{4}{3}$ است و به خاطر خطی بودن امید ریاضی، $\mathbb{E}[ST_{pp}(P_n, \circ)] = \frac{4}{3}n - 2$. از طرف دیگر، طبق قانون اعداد بزرگ، برای n های بزرگ $ST_{pp}(P_n, \circ)$ بسیار به مقدار متوسط خود نزدیک است.

یک گراف مهم دیگر گراف کامل است. در سال ۲۰۰۰ رفتار این پروتکل در یک گراف کامل تحلیل شد و قضیه زیر ثابت شد.

قضیه ۳.۲ ([۳۶]). فرض کنید K_n یک گراف کامل n رأسی باشد و v یکی از رئوس آن باشد. در این صورت داریم $\mathbb{E}[ST_{pp}(K_n, v)] \sim \log_3 n$ و برای n های بزرگ، متغیر تصادفی $ST_{pp}(K_n, v)$ به مقدار متوسط خود نزدیک است.

۱.۲ کران‌های عمومی. در مثال ۲.۱ دیدیم که در پروتکل ردکن بره، امیدریاضی زمان پخش یک گراف n رأسی، مثلاً گراف ستاره، می‌تواند از مرتبه $n \ln n$ باشد. قضیه زیر، که اثبات آن به کمک استقرا چندان پیچیده نیست، نشان می‌دهد که چنین وضعیتی در مورد پروتکل بده‌بستان رخ نمی‌دهد.

قضیه ۴.۲ ([۱]). برای هر گراف همبند n رأسی مثل G و هر رأس v از آن، داریم

$$\mathbb{E} [\text{ST}_{\text{pp}}(G, v)] < 5n.$$

برای اثبات این قضیه نیاز به گزاره احتمالاتی زیر داریم که اثباتش ساده است و به خواننده واگذار می‌شود.

گزاره ۵.۲. فرض کنید X یک متغیر تصادفی نامنفی باشد. در این صورت داریم:

$$\mathbb{E} [X] = \int_0^{\infty} \mathbb{P} [X > t] dt.$$

اگر مقدار X همواره عددی طبیعی باشد، فرمول ساده‌تر زیر را داریم:

$$\mathbb{E} [X] = \sum_{t=0}^{\infty} \mathbb{P} [X > t].$$

اثبات قضیه ۴.۲. در این اثبات فقط از عمل «خبر پرسیدن» رئوس استفاده می‌کنیم، یعنی فرض می‌کنیم هر رأسی که خبر ندارد در هر گام خبر را از یک همسایه تصادفی خود می‌پرسد (و اگر همسایه خبر را بداند به او اطلاع می‌دهد)، ولی رئوسی که خبر را می‌دانند کاری انجام نمی‌دهند. دقت کنید که چون داریم یک کران بالا برای زمان پخش ثابت می‌کنیم، می‌توانیم این فرض را بکنیم.

قرار دهید $\alpha = \sqrt{e}/(\sqrt{e} - 1) \approx 2/54$. یک لحظه را در طول اجرای پروتکل در نظر بگیرید و فرض کنید در مجموع m رأس هستند که خبر را نمی‌دانند، و b رأس هستند که خبر را نمی‌دانند ولی همسایه‌ای دارند که خبر را می‌داند (این رئوس را رئوس مرزی می‌نامیم). در این صورت ادعا می‌کنیم امیدریاضی زمانی که طول می‌کشد تا همه رئوس باخبر بشوند حداکثر $2b + \alpha m$ است. اگر این ادعا درست باشد، نتیجه می‌گیریم

$$\mathbb{E} [\text{ST}_{\text{pp}}(G, v)] < (2 + \alpha)n < 5n.$$

در ادامه ادعای بالا را ثابت می‌کنیم. فرض کنید I مجموعه رئوس باخبر باشد، B مجموعه رئوس مرزی باشد، و R مجموعه بقیه رئوس باشد. قرار دهید $b = |B|$ و $m = |B| + |R|$. فرض کنید $d(v)$ درجه رأس v در گراف G باشد و برای مجموعه S از رئوس، تعداد همسایه‌های v در S را با $d_S(v)$ نشان دهید.

ادعای مذکور را با استقرای روی m ثابت می‌کنیم. وقتی $m = 1$ باشد، تنها رأس باقی‌مانده خبر را در گام بعدی از یک همسایه باخبر می‌پرسد پس دقیقاً یک گام طول می‌کشد تا همه رئوس باخبر شوند. چون

$$1 < (2 + \alpha) \times 1 - 2 \times 1$$

پس پایه استقرا ثابت می شود.

حال فرض کنید $m < 1$ و یک لحظه از اجرای پروتکل را در نظر بگیرید. در این لحظه دو وضعیت ممکن است وجود داشته باشد. در وضعیت اول، رأسی مرزی مثل v وجود دارد که $d_R(v) \geq d_B(v)$ (شکل ۴ (چپ) را ببینید). در این حالت، انقدر صبر می کنیم تا v خبر را بفهمد، یعنی همه تماس هایی که رئوسی به غیر از v انجام می دهند را نادیده می گیریم. بدیهی است می توان این فرض را کرد زیرا داریم برای زمان پخش یک کران بالا می دهیم. مدت زمانی که طول می کشد تا v بالاخره با یک همسایه باخبر تماس بگیرد و خبر را بپرسد یک متغیر تصادفی هندسی با پارامتر $\frac{d_I(v)}{d(v)}$ است و میانگین آن برابرست با $\frac{d(v)}{d_I(v)}$. داریم

$$\frac{d(v)}{d_I(v)} = \frac{d_I(v) + d_R(v) + d_B(v)}{d_I(v)} \leq 1 + \frac{2d_R(v)}{d_I(v)} \leq 2d_R(v) + 1.$$

پس مدت زمانی که به طور متوسط طول می کشد تا v خبر را بفهمد حداکثر $2d_R(v) + 1$ است. وقتی v خبر را فهمید، تعداد رئوس بی خبر یکی کم می شود، و تعداد رئوس مرزی $1 - d_R(v)$ تا زیاد می شود. با استفاده از فرض استقرادری می یابیم که در این حالت امیدریاضی زمانی که طول می کشد تا همه رئوس باخبر بشوند حداکثر برابرست با

$$2d_R(v) + 1 + (\alpha + 2)(m - 1) - 2(b + d_R(v) - 1) < (\alpha + 2)m - 2b,$$

و حکم استقرا ثابت می شود.

در وضعیت دوم، برای همه رئوس مرزی v داریم $d_R(v) < d_B(v)$ (شکل ۴ (راست) را ببینید). بنابراین برای همه این رئوس داریم

$$\frac{d_I(v)}{d(v)} = \frac{d_I(v)}{d_I(v) + d_R(v) + d_B(v)} \geq \frac{1}{1 + d_R(v) + d_B(v)} \geq \frac{1}{2d_B(v)} \geq \frac{1}{2b}. \quad (1)$$

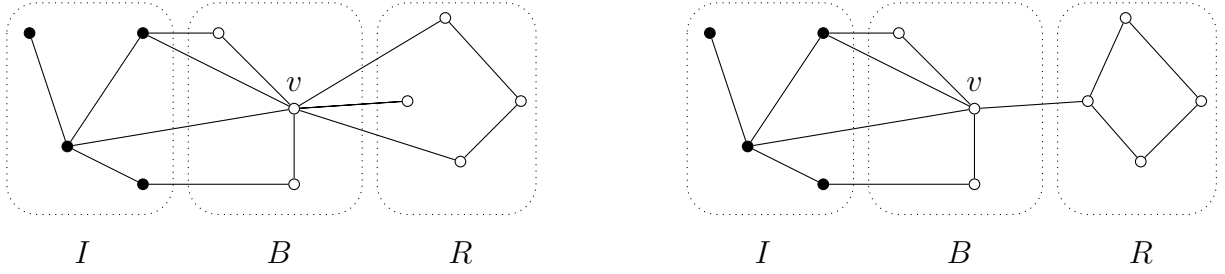
زمانی که طول می کشد تا رأس جدیدی باخبر بشود را با X نشان می دهیم. در این صورت،

$$X = \min\{X_1, X_2, \dots, X_b\},$$

که در آن X_i زمانی است که طول می کشد تا رأس مرزی i ام باخبر بشود. دقت کنید که X_i ها متغیرهای تصادفی هندسی هستند، و با توجه به نامساوی (۱)، پارامتر هر یک از آنها لااقل $1/2b$ است. به علاوه این متغیرها از هم مستقلند زیرا رئوس مستقل از هم عمل می کنند. با استفاده از گزاره ۵.۲، به دست می آوریم

$$\mathbb{E}[X] = \sum_{t=0}^{\infty} \mathbb{P}[X > t] = \sum_{t=0}^{\infty} \prod_{i=1}^b \mathbb{P}[X_i > t] \leq \sum_{t=0}^{\infty} \left(1 - \frac{1}{2b}\right)^{tb} \leq \sum_{t=0}^{\infty} (e^{-1/2b})^{tb} = \sum_{t=0}^{\infty} e^{-t/2} = \alpha.$$

در محاسبه بالا از نامساوی $1 + x \leq e^x$ استفاده کردیم که برای هر x حقیقی درست است. بنابراین بعد از به طور متوسط α گام، رأسی مرزی خبر را می فهمد. وقتی این اتفاق افتاد، تعداد رئوس بی خبر یکی کم می شود، و تعداد رئوس مرزی یا کم



شکل ۴: چپ: وضعیت اول در اثبات قضیه ۴.۲: رأس مرزی v وجود دارد که $d_R(v) \geq d_B(v) = ۲$. راست: وضعیت دوم: برای همه رؤوس مرزی داریم $d_R(v) < d_B(v)$. در شکل، رؤوس باخبر سیاه و رؤوس بی خبر سفیدند.

نمی‌شود یا دقیقاً یکی کم می‌شود. بنابراین در این حالت با استفاده از فرض استقرا، امید ریاضی زمانی که طول می‌کشد تا همه رؤوس باخبر بشوند حداکثر برابرست با

$$\alpha + (\alpha + ۲)(m - ۱) - ۲(b - ۱) = (\alpha + ۲)m - ۲b,$$

□

و حکم استقرا ثابت می‌شود.

تعریف. زمان متوسط پخش^{۱۸} گراف G را با $\text{ast}_{\text{pp}}(G)$ نشان می‌دهیم و به صورت زیر تعریف می‌کنیم:

$$\text{ast}_{\text{pp}}(G) := \max\{\mathbb{E}[\text{ST}_{\text{pp}}(G, v)] : v \in V(G)\}.$$

بنابراین قضیه ۴.۲ بیان می‌کند که $\text{ast}_{\text{pp}}(G)$ حداکثر خطی است.

زمان تضمینی پخش را می‌توان مشابه قبل برای این پروتکل نیز تعریف کرد؛ این پارامتر را با $\text{gst}_{\text{pp}}(G)$ نشان می‌دهیم. به عنوان تمرین، گرافی را بسازید که زمان تضمینی پخش آن لااقل از مرتبه $\Omega(n \ln n)$ باشد. (راهنمایی: مرکز دو ستاره را به هم وصل کنید.)

قضیه ۶.۲ ([۱]). برای هر گراف G داریم

$$\text{gst}_{\text{pp}}(G) \leq e \lceil \ln n \rceil \times \text{ast}_{\text{pp}}(G).$$

اثبات. فرض کنید t عددی حقیقی، k عددی طبیعی، و v رأسی دلخواه باشد. ابتدا نشان می‌دهیم

$$\mathbb{P}[\text{ST}_{\text{pp}}(G, v) > kt] \leq \mathbb{P}[\text{ST}_{\text{pp}}(G, v) > t]^k. \quad (۲)$$

بازه زمانی $[0, kt]$ را به k بازه مساوی به طول t تقسیم کنید. الگوریتمی را در نظر بگیرید که مشابه بده‌بستان عمل می‌کند با این تفاوت که فقط به اندازه زمان t فرصت دارد تا همه رؤوس را باخبر کند، و اگر نتوانست در این زمان رؤوس را باخبر کند،

^{۱۸}average spread time

دوباره از اول شروع می‌کنیم. دقت کنید که اگر k بار این الگوریتم را اجرا کنیم، احتمال این که هرگز موفق نشود که رئوس را باخبر کند $\mathbb{P}[\text{ST}_{\text{pp}}(G, v) > t]^k$ است. در عین حال، اگر به پروتکل بده‌بستان به اندازه kt فرصت بدهیم، احتمال موفقیت آن از k بار انجام این الگوریتم بیشتر است، در نتیجه (۲) درست است.

از طرف دیگر، از نامساوی مارکوف داریم

$$\mathbb{P}[\text{ST}_{\text{pp}}(G, v) > e\mathbb{E}[\text{ST}_{\text{pp}}(G, v)]] < 1/e$$

از ترکیب این نامساوی و (۲) به دست می‌آوریم:

$$\mathbb{P}[\text{ST}_{\text{pp}}(G, v) > e[\ln n] \times \mathbb{E}[\text{ST}_{\text{pp}}(G, v)]] < 1/n.$$

□

چون این نامساوی برای هر رأس ابتدایی v درست است، قضیه ثابت می‌شود.

به عنوان تمرین و با استفاده از نامساوی (۲) و گزاره ۵.۲، نشان دهید برای هر گراف G داریم

$$(1 - 1/n) \text{ast}_{\text{pp}}(G) \leq \text{gst}_{\text{pp}}(G).$$

شهود ما می‌گوید هر چقدر ارتباطات بین بخش‌های مختلف یک گراف بیشتر باشد، خبر سریع‌تر بین بخش‌های مختلف پخش می‌شود. برعکس، اگر مثلاً دو بخش گراف باشند که هر یک درون خود ارتباطات زیادی داشته باشند ولی بین این دو بخش تعداد یال‌های کمی وجود داشته باشد، مدت بیشتری طول می‌کشد تا خبر از یک بخش به بخش دیگر منتقل شود. با دقت کردن این شهود به کمک دو پارامتر مختلف گراف، دو قضیه زیر اثبات شده‌اند.

تعریف (رسانایی گراف^{۱۹}). برای یک مجموعه S از رئوس گراف G ، تعریف می‌کنیم $\text{vol}(S) = \sum_{v \in S} \deg(v)$ ، و تعداد یال‌هایی را که یک سرشان در S است و سر دیگرشان در $V \setminus S$ است با $\partial(S)$ نشان می‌دهیم. رسانایی گراف G را به صورت زیر تعریف می‌کنیم:

$$\phi(G) := \min \left\{ \frac{\partial(S)}{\text{vol}(S)} : S \subseteq V, \text{vol}(S) \leq \text{vol}(V)/2 \right\}.$$

دقت کنید که $\text{vol}(V)$ دقیقاً دو برابر تعداد یال‌های G است. رسانایی گراف همیشه عددی بین ۰ و ۱ است، و برابر ۰ است اگر و تنها اگر گراف ناهمبند باشد. به صورت شهودی، هر چه ارتباطات بین بخش‌های مختلف یک گراف بیشتر باشد، رسانایی آن هم بیشتر است. قضیه زیر که ارتباط جالبی بین این پارامتر و سرعت پخش پروتکل بده‌بستان را بیان می‌کند در سال ۲۰۱۱ به اثبات رسید.

قضیه ۷.۲ ([۲۶]). عدد ثابت C وجود دارد که برای هر گراف n رأسی G داریم $\text{gst}_{\text{pp}}(G) \leq C(\ln n)/\phi(G)$.

^{۱۹}graph conductance

همچنین ثابت شده که این کران قابل بهتر شدن نیست (به جز این که ثابت C امکان بهبود دارد)؛ در مقاله [۷] گراف‌هایی ساخته شده‌اند که زمان تضمینی پخش آن‌ها لااقل از مرتبه $\Omega((\ln n)/\phi(G))$ است.

تعریف (عدد برابر محیطی رأسی^{۲۰}). برای مجموعه S از رئوس گراف G ، مجموعه رئوسی از $V \setminus S$ که همسایه‌ای در S دارند را با $N(S)$ نشان می‌دهیم. عدد برابر محیطی رأسی گراف G را به صورت زیر تعریف می‌کنیم:

$$\alpha(G) := \min \left\{ \frac{|N(S)|}{|S|} : S \subseteq V, |S| \leq |V|/2 \right\}.$$

دقت کنید که این پارامتر هم همیشه عددی بین ۰ و ۱ است، و برابر ۰ است اگر و تنها اگر گراف ناهمبند باشد. قضیه زیر که ارتباط جالبی بین این پارامتر و سرعت پخش پروتکل بده‌بستان را بیان می‌کند در سال ۲۰۱۴ به اثبات رسید.

قضیه ۸.۲ ([۲۷]). ثابت C وجود دارد که برای هر گرافی G رأسی با بزرگ‌ترین درجه Δ داریم

$$\text{gst}_{\text{pp}}(G) \leq C(\ln \Delta)(\ln n)/\alpha(G).$$

همچنین ثابت شده که این کران قابل بهتر شدن نیست (به جز این که ثابت C امکان بهبود دارد)؛ در مقاله [۲۸] گراف‌هایی ساخته شده‌اند که زمان تضمینی پخش آن‌ها لااقل از مرتبه $\Omega(\ln \Delta \times (\ln n)/\alpha(G))$ است.

پارامتر مهم دیگری که میزان همبستگی یک گراف را نشان می‌دهد، عدد برابر محیطی یالی^{۲۱} گراف نام دارد و به صورت زیر تعریف می‌شود:

$$\iota(G) := \min \left\{ \frac{\partial(S)}{|S|} : S \subseteq V, |S| \leq |V|/2 \right\}.$$

قضیه‌های ۷.۲ و ۸.۲ بیان می‌کنند که اگر گرافی رسانایی اش از عدد ثابتی بیشتر باشد، یا عدد برابر محیطی رأسی اش از عدد ثابتی بیشتر باشد، آن‌گاه زمان تضمینی پخش اش حداکثر یک چندجمله‌ای برحسب $\ln n$ است. این گزاره درباره عدد برابر محیطی یالی درست نیست: در مقاله [۸] خانواده‌ای از گراف‌ها معرفی شده است که عدد برابر محیطی یالی شان از عدد ثابتی بیشتر است ولی زمان تضمینی پخش شان لااقل از مرتبه $\Omega(n)$ است.

۳ پروتکل بده‌بستان ناهمزمان

تعریف. منظور از یک ساعت پواسون^{۲۲} با پارامتر λ ، ساعتی زنگ‌دار است که لحظات زنگ زدن آن متناظر با یک فرایند پواسون با پارامتر λ است. به عبارت دیگر، این ساعت در لحظات $X_1 \leq X_2 \leq X_3 \leq \dots$ زنگ می‌زند، که $X_1, X_2 - X_1, X_3 - X_2, \dots$ متغیرهای مستقل نمایی با نرخ λ هستند.

یادآوری می‌کنیم که یک متغیر تصادفی با توزیع نمایی (که برای سادگی آن را یک متغیر نمایی می‌نامیم) با نرخ (یا پارامتر) λ متغیری پیوسته و نامنفی مثل X است به طوری که برای هر $t \geq 0$ داریم $\mathbb{P}[X > t] = e^{-\lambda t}$. امید ریاضی

^{۲۰}vertex isoperimetric number

^{۲۱}edge isoperimetric number

^{۲۲}Poisson clock

X برابر λ^{-1} است. متغیرهای نمایی خواص جالبی دارند که در این بخش از آن‌ها استفاده می‌کنیم. دو خاصیت زیر از تعریف تابع توزیع آن‌ها به دست می‌آید:

۱. فرض کنید X متغیری نمایی باشد. برای هر $s > t$ داریم $\mathbb{P}[X > s | X > t] = \mathbb{P}[X > s - t]$. به عبارت دیگر، متغیر نمایی بدون حافظه است. خوب است کمی درباره مفهوم بی حافظه بودن متغیر نمایی توضیح دهیم. فرض کنید به ماهی گیری به برکه ای رفته‌اید که می‌دانید ماهی‌های زیادی دارد و ساعت ۴ قلاب خود را در آب انداخته‌اید. فرض کنید ساعت ۴:۳۰ شده و تا این لحظه هیچ ماهی ای نگرفته‌اید. از این لحظه به بعد، چقدر طول می‌کشد تا ماهی بگیرید؟ آیا این موضوع که تا این لحظه ماهی نگرفته‌اید باعث می‌شود فکر کنید به زودی ماهی خواهید گرفت؟ شهوداً می‌دانیم که مدتی که از همین لحظه طول می‌کشد تا ماهی اول را بگیریم مستقل از اینست که تا الان چقدر صبر کرده‌ایم. یک متغیر نمایی در چنین وضعیتی می‌تواند مدل ریاضی مناسبی برای زمانی که طول می‌کشد تا اولین ماهی را بگیرید باشد، چرا که این متغیر هم بدون حافظه است و اگر بدانیم $X > t$ ، آن‌گاه توزیع $X - t$ دقیقاً مثل خود توزیع اولیه X است، یعنی این متغیر گذشته را «فراموش می‌کند» و انگار که لحظه t درست مثل همان لحظه آغاز است.

می‌توان نشان داد که متغیر نمایی تنها متغیر تصادفی پیوسته بدون حافظه است، و متغیر هندسی هم تنها متغیر گسسته بدون حافظه است. بنابراین این دو خانواده از متغیرها از این نظر مشابه یکدیگرند.

۲. فرض کنید X و Y متغیرهای نمایی مستقل باشند. در این صورت $\min\{X, Y\}$ هم متغیر نمایی است که نرخش مجموع نرخ X و Y است.

پروتکل بده‌بستان ناهم‌زمان^{۲۳} به صورت زیر تعریف می‌شود.

تعریف. فرض کنید G گرافی بدون جهت، ساده و همبند باشد. در هر رأس G یک ساعت پواسون با پارامتر ۱ قرار دارد، و این ساعت‌ها از هم مستقلند. در ابتدا یکی از رئوس گراف مانند v خبری را می‌داند. هر بار که ساعت یک رأس مانند u زنگ می‌زند، یکی از همسایه‌های خود مثل w را به تصادف انتخاب می‌کند و با او تماس می‌گیرد: اگر u خبر را بداند آن را برای w می‌فرستد و اگر u خبر را نداند خبر را از w می‌پرسد (اگر w خبر را بداند u هم باخبر می‌شود وگرنه اتفاقی نمی‌افتد). اولین زمانی که همه رئوس باخبر می‌شوند را با $\text{ST}_{\text{app}}(G, v)$ نشان می‌دهیم، که یک متغیر تصادفی است.

این پروتکل در مقاله [۵] معرفی شد. به راحتی می‌توان دید که بین زمان پخش در این پروتکل و مسئله احتمالاتی first-passage-percolation وقتی وزن یال‌ها متغیرهای نمایی مستقل باشند ارتباطات جالبی وجود دارد و به خصوص برای گراف‌های منظم، این دو مسئله یکی هستند. مسئله دوم بیشتر توسط احتمال‌دان‌ها بررسی شده است، مثلاً [۳۴] را ببینید.

برای تحلیل این پروتکل تعریف زیر کارگشا خواهد بود.

^{۲۳}the asynchronous push&pull protocol

تعریف. فرض کنید $e = uv$ یالی در گراف باشد. تأخیر یال e را با $T(e)$ نشان می‌دهیم و به صورت زیر تعریف می‌کنیم. فرض کنیم τ اولین زمانی باشد که یکی از u و v باخبر می‌شود، و ρ اولین زمانی بعد از τ باشد که یکی از u و v با دیگری تماس می‌گیرد. در این صورت قرار می‌دهیم $T(e) = \rho - \tau$ ، که یک متغیر تصادفی نامنفی است. توجه کنید که پس از زمان ρ هر دوی u و v خبر را می‌دانند.

لم ۱.۳. فرض کنید $e = uv$ یالی در گراف باشد. در این صورت $T(e)$ متغیری تصادفی با توزیع نمایی با نرخ $1/\deg(u) + 1/\deg(v)$ است.

اثبات. فرض کنیم τ اولین زمانی باشد که یکی از u و v باخبر می‌شود. به دلیل بی‌حافظه بودن متغیر نمایی، هر بار که ساعتی زنگ می‌زند و پس از آن که ارتباطی بین دو رأس برقرار می‌شود، می‌توان فرض کرد که ساعت‌های رؤس دوباره از اول شروع به کار می‌کنند. بنابراین، اگر Y_u اولین زمانی بعد از τ باشد که ساعت u زنگ می‌زند، آن‌گاه $Y_u - \tau$ متغیری نمایی با پارامتر ۱ است. از طرف دیگر، هر بار که ساعت u زنگ می‌زند، این رأس به احتمال $1/\deg(u)$ با v تماس می‌گیرد. بنابراین اگر Z_u اولین زمانی بعد از τ باشد که u با v تماس می‌گیرد، طبق ویژگی‌های متغیرهای تصادفی نمایی، $Z_u - \tau$ هم یک متغیر نمایی است ولی با نرخ $1/\deg(u)$. به همین ترتیب، اگر Z_v اولین زمانی بعد از τ باشد که v با u تماس می‌گیرد، $Z_v - \tau$ یک متغیر نمایی با نرخ $1/\deg(v)$ است، و Z_u و Z_v از هم مستقلند. در نتیجه داریم

$$T(e) = \min\{Z_u - \tau, Z_v - \tau\}.$$

ولی می‌دانیم که مینیمم دو متغیر نمایی مستقل، خود یک متغیر نمایی است با نرخ برابر مجموع نرخ آن دو متغیر. بنابراین، $T(e)$ متغیری تصادفی با توزیع نمایی است و نرخ آن هم $1/\deg(u) + 1/\deg(v)$ می‌باشد. \square

دو مثال ببینیم.

مثال ۲.۳ (گراف ستاره). فرض کنید S_n گراف ستاره باشد که رأس مرکزی آن \circ و بقیه رؤس آن $1, 2, \dots, n-1$ نام‌گذاری شده‌اند، و فرض کنید که در ابتدا رأس \circ خبر را بداند. داریم:

$$\text{ST}_{\text{app}}(S_n, \circ) = \max\{T(\circ, 1), T(\circ, 2), \dots, T(\circ, n-1)\}.$$

طبق لم ۱.۳، $T(\circ, k)$ ‌ها متغیرهای نمایی با نرخ $\frac{n}{n-1}$ هستند. به علاوه، این متغیرها از هم مستقلند. در نتیجه، داریم

$$\mathbb{P}[\text{ST}_{\text{app}}(S_n, \circ) \leq t] = (1 - \exp(-nt/(n-1)))^{n-1},$$

و با انجام مقداری محاسبه می‌توان دید که

$$\mathbb{E}[\text{ST}_{\text{app}}(S_n, \circ)] \sim \ln n.$$

هم‌چنین $\text{ST}_{\text{app}}(S_n, \circ)$ برای n ‌های بزرگ به امیدریاضی خود نزدیک است.

اگر در ابتدا رأسی به جز رأس \circ خبر را بداند (مثلاً رأس ۱)، آن گاه داریم

$$\text{ST}_{\text{app}}(S_n, 1) = T(\circ, 1) + \max\{T(\circ, 2), T(\circ, 3), \dots, T(\circ, n-1)\}.$$

به راحتی می توان دید که هم چنان داریم $\mathbb{E}[\text{ST}_{\text{app}}(S_n, 1)] \sim \ln n$. جزئیات محاسبات به خواننده واگذار می شود.

مثال ۳.۳ (گراف مسیر). فرض کنید P_n گراف مسیر بارتوس $1, \dots, n-1, \circ$ باشد و در ابتدا رأس \circ خبر را بداند. می خواهیم امید ریاضی زمان پخش را محاسبه کنیم. دقت کنید که مسیری که خبر طی می کند دقیقاً مشخص است، در نتیجه داریم

$$\text{ST}_{\text{app}}(P_n, \circ) = T(\circ, 1) + T(1, 2) + T(2, 3) + \dots + T(n-2, n-1),$$

و طبق لم ۱.۳، $T(\circ, 1)$ و $T(n-2, n-1)$ نامایی بانرخ $2/3$ هستند و سایر $T(i, i+1)$ ها نامایی بانرخ ۱ هستند. نتیجه می گیریم

$$\mathbb{E}[\text{ST}_{\text{app}}(P_n, \circ)] = 2 \times (2/3) + (n-3) \times 1 = n - 5/3.$$

به علاوه، همه $T(i, i+1)$ ها از هم مستقلند. از قانون اعداد بزرگ نتیجه می گیریم که برای n های بزرگ $\text{ST}_{\text{app}}(P_n, \circ)$ بسیار به مقدار متوسط خود نزدیک است.

یک گراف مهم دیگر گراف کامل است. در سال ۱۹۹۹ رفتار این پروتکل در یک گراف کامل تحلیل شد و قضیه زیر ثابت شد. در حقیقت این قضیه توسط یک احتمال دان و برای مسئله first-passage-percolation ثابت شد، ولی چون گراف کامل منظم است، نتیجه آن برای پروتکل بدهستان ناهمزمان هم صدق می کند.

قضیه ۴.۳ ([۳۵]). فرض کنید K_n یک گراف کامل n رأسی باشد و v یکی از رؤس آن باشد. در این صورت داریم $\mathbb{E}[\text{ST}_{\text{app}}(K_n, v)] \sim \ln n$ و متغیر تصادفی $\text{ST}_{\text{app}}(K_n, v)$ به مقدار متوسط خود نزدیک است.

یک پرسش طبیعی در مورد این پروتکل این است: برای یک گراف n رأسی، زمان پخش چقدر می تواند کوچک یا بزرگ باشد؟ قضیه زیر که سال گذشته اثبات شد نشان می دهد که از نظر مرتبه، گراف های کامل و ستاره کمترین زمان پخش را دارند و گراف مسیر بیشترین زمان پخش را دارد.

قضیه ۵.۳ ([۱]). برای هر گراف همبند n رأسی مثل G و هر رأس v از آن، داریم

$$\frac{\ln n}{3} < \mathbb{E}[\text{ST}_{\text{pp}}(G, v)] < 5n.$$

اثبات این قضیه را به عنوان تمرین به خواننده واگذار می کنیم. اثبات کران بالا بسیار شبیه اثبات قضیه ۴.۲ است. برای اثبات کران پایین آن باید از خواص استاندارد فرایندهای پواسون استفاده کرد.

۱.۳ مقایسه پروتکل‌های بده‌بستان و بده‌بستان ناهمزمان. دیدیم که زمان پخش گراف ستاره در پروتکل بده‌بستان حداکثر ۲ است حال آن که زمان پخش یک گراف n رأسی در پروتکل بده‌بستان ناهمزمان لااقل از مرتبه $\Omega(\ln n)$ است. از طرف دیگر زمان پخش در هر دو پروتکل حداکثر از مرتبه $O(n)$ است و این کران را نمی‌توان در حالت کلی بهتر کرد. حال می‌توان پرسید که برای یک گراف داده شده، بین زمان پخش این دو پروتکل چه رابطه‌ای وجود دارد؟ نویسندگان [۱۲] این دو پروتکل را روی گراف‌های زیادی شبیه‌سازی کردند، و روی همه آن گراف‌ها (که مدل‌هایی برای شبکه‌های اجتماعی بودند و یکی از آن‌ها گراف twitter بود) پروتکل بده‌بستان ناهمزمان سریع‌تر از بده‌بستان خبر را پخش کرد (البته واضح است که شبیه‌سازی را روی گراف ستاره انجام ندادند).

برای این که ابهامی پیش نیاید، در ادامه پروتکل بده‌بستان که در بخش ۲ معرفی شد را پروتکل بده‌بستان همزمان^{۲۴} می‌نامیم. قضیه زیر یک رابطه کلی بین زمان‌های پخش دو پروتکل بیان می‌کند.

قضیه ۶.۳ ([۱]). اعداد ثابت $C, C' > 0$ وجود دارند که برای هر گراف همبند n رأسی مثل G و هر رأس v از آن، داریم

$$\frac{C}{\ln^2 n} < \frac{\mathbb{E}[\text{ST}_{\text{pp}}(G, v)]}{\mathbb{E}[\text{ST}_{\text{app}}(G, v)]} < C' n^{2/3} \ln n.$$

در همه گراف‌هایی که ما بررسی کردیم رابطه قوی‌تر $\mathbb{E}[\text{ST}_{\text{app}}(G, v)] \leq O(\mathbb{E}[\text{ST}_{\text{pp}}(G, v)] + \ln n)$ برقرار بود ولی نتوانستیم آن را اثبات کنیم. شاید شما بتوانید آن را اثبات کنید و ممنون می‌شویم اگر برای ما هم بفرستید.

کران بالای قضیه بالا ممکن است زیادی بزرگ به نظر برسد، ولی در ادامه گرافی را معرفی می‌کنیم که زمان متوسط پخش پروتکل بده‌بستان ناهمزمان در آن لگاریتمی است ولی زمان متوسط پخش پروتکل بده‌بستان همزمان در آن لااقل از مرتبه چند جمله‌ای است.

مثال ۷.۳ (گراف گردن‌بند). فرض کنید k و m اعداد صحیح بزرگ‌تر از ۱ باشند و گراف گردن‌بند که در شکل ۵ نشان داده شده را در نظر بگیرید. این گراف $m + 1$ رأس سفید مثل $v_0, v_1, \dots, v_{m-1}, v_m$ دارد که روی یک خط چیده شده‌اند، و بین هر دو رأس سفید متوالی، k مسیر مجزا به طول ۲ از رئوس سیاه قرار دارد. پس تعداد کل رئوس برابر است با $n = km + m + 1$.

می‌خواهیم امید ریاضی زمان پخش با شروع از رأس v_0 را در دو پروتکل تخمین بزنیم. چون فاصله گرافی بین v_0 و v_m برابر $2m$ است، قطعاً داریم $\text{ST}_{\text{pp}}(G, v_0) \geq 2m$. حال فرض کنید برای هر $0 \leq i \leq m$ ، T_i برابر باشد با اولین لحظه‌ای که رأس v_i خبر را می‌فهمد. پس $T_0 = 0$.

لم ۸.۳. برای هر $0 \leq i \leq m - 1$ داریم

$$\mathbb{E}[T_{i+1} - T_i] \leq \frac{2}{\sqrt{k}}.$$

^{۲۴}synchronous push&pull protocol

اثبات. اثبات را برای $i = 0$ ارائه می‌کنیم، برای سایر i ها اثبات مشابه است. فرض کنید x یکی از همسایه‌های مشترک v_0 و v_1 باشد. طبق لم ۱.۳، $T(v_0, x)$ متغیری نمایی با نرخ $1/k + 1/2$ است و $T(x, v_1)$ متغیری نمایی با نرخ $1/(2k) + 1/2$ است بنابراین برای هر $t \geq 0$

$$\mathbb{P}[T(v_0, x) \leq 2t] = 1 - \exp(-2t(1/k + 1/2)) \geq 1 - e^{-t}$$

و مشابهاً می‌توان دید $\mathbb{P}[T(x, v_1) \leq 2t] \geq 1 - e^{-t}$. در نتیجه

$$\mathbb{P}[T(v_0, x) + T(x, v_1) > 4t] \leq 1 - \mathbb{P}[T(v_0, x) \leq 2t] \mathbb{P}[T(x, v_1) \leq 2t] \leq 1 - (1 - e^{-t})^2.$$

چون بین v_0 و v_1 تعداد k مسیر مجزای مثل هم به طول ۲ وجود دارد،

$$\mathbb{P}[T_\setminus - T_\circ > 4t] = (\mathbb{P}[T(v_0, x) + T(x, v_1) > 4t])^k \leq (1 - (1 - e^{-t})^2)^k = (2e^{-t} - e^{-2t})^k.$$

بنابراین طبق گزاره ۵.۲،

$$\begin{aligned} \mathbb{E}[T_\setminus - T_\circ] &= \int_0^\infty \mathbb{P}[T_\setminus - T_\circ > t] dt \leq \int_0^\infty (2e^{-t/4} - e^{-t/2})^k dt \\ &= \int_0^4 (2e^{-t/4} - e^{-t/2})^k dt + \int_4^\infty (2e^{-t/4} - e^{-t/2})^k dt \end{aligned} \quad (3)$$

برای محدود کردن انتگرال اول از نامساوی $2e^{-y} - e^{-2y} \leq e^{-y^2/4}$ استفاده می‌کنیم که برای هر $y \in [0, 1]$ صادق است (اثبات آن به خواننده واگذار می‌شود) پس

$$\int_0^4 (2e^{-t/4} - e^{-t/2})^k dt \leq \int_0^4 e^{-kt^2/64} dt < \int_0^\infty e^{-kt^2/64} dt$$

با تغییر متغیر $r = t\sqrt{k/32}$ داریم

$$\int_0^\infty e^{-kt^2/64} dt = \int_0^\infty e^{-r^2/2} \sqrt{32/k} dr = \frac{1}{2} \sqrt{\pi/k}.$$

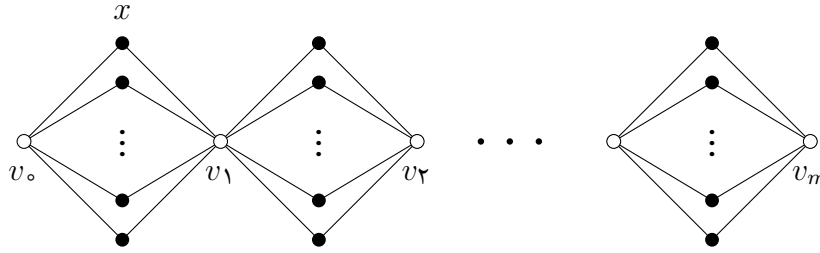
که در تساوی آخر از فرمول $\int_0^\infty e^{-r^2/2} dr = \sqrt{2\pi}$ استفاده کرده‌ایم (تابع چگالی متغیرهای نرمال را به خاطر بیاورید!).

حال انتگرال دوم در (۳) را محدود می‌کنیم:

$$\int_4^\infty (2e^{-t/4} - e^{-t/2})^k dt < \int_4^\infty 2^k e^{-kt/4} dt = \frac{2^{k+2}}{ke^k} < 4/\sqrt{k}.$$

□

چون $20 < 4 + 8\sqrt{\pi}$ لم ثابت می‌شود.



شکل ۵: گراف گردنبنند

از لم ۸.۳ و خطی بودن امیدریاضی نتیجه می‌شود $\mathbb{E}[T_m] \leq 2 \circ m / \sqrt{k}$. حال سؤال اینست که وقتی همهٔ رؤوس سفید باخبر شدند، چقدر دیگر باید صبر کنیم تا رؤوس سیاه هم باخبر شوند؟ توجه کنید که هر رأس سیاه فقط همسایه‌های سفید دارد. بنابراین به محض این که ساعت یک رأس سیاه زنگ بخورد، بلافاصله آن رأس خبر را از یک همسایهٔ سفیدش می‌گیرد. فرض کنید در لحظه‌ای که همهٔ رؤوس سفید باخبر شدند، ℓ رأس سیاه بی‌خبر وجود داشته باشند. مدتی که طول می‌کشد تا ساعت یک رأس سیاه خاص زنگ بخورد یک متغیر نمایی با نرخ ۱ است (این جا از بی‌حافظه بودن متغیر نمایی استفاده می‌کنیم، چرا که فرض می‌کنیم زنگ خوردن ساعت یک رأس از T_m به بعد، به زنگ خوردن‌های قبلی آن وابسته نیست). دقت کنید که اگر ℓ متغیر تصادفی مستقل نمایی با نرخ ۱ داشته باشیم، امیدریاضی ماکسیمم آن‌ها کوچک‌تر از $2 \ln \ell$ است (اثبات کنید!) بنابراین

$$\mathbb{E}[\text{ST}_{\text{app}}(G, v_0)] \leq \mathbb{E}[T_m] + 2 \ln \ell < 2 \circ m / \sqrt{k} + 2 \ln n .$$

برای یک n داده شده می‌توان k را از مرتبهٔ $(n / \ln n)^{2/3}$ و m را از مرتبهٔ $n^{1/3} (\ln n)^{2/3}$ طوری انتخاب کرد که داشته باشیم $n = km + m + 1$. بنابراین برای این گراف به دست می‌آوریم

$$\frac{\mathbb{E}[\text{ST}_{\text{pp}}(G, v_0)]}{\mathbb{E}[\text{ST}_{\text{app}}(G, v_0)]} \geq \frac{2m}{2 \circ m / \sqrt{k} + 2 \ln n} = \Theta((n / \ln n)^{1/3}) .$$

با توجه به بحث بالا یک سؤال جالب به ذهن می‌رسد: بزرگ‌ترین عدد κ کدام است که برای هر گراف همبند n رأسی مثل G و هر رأس v از آن داشته باشیم

$$\frac{\mathbb{E}[\text{ST}_{\text{pp}}(G, v)]}{\mathbb{E}[\text{ST}_{\text{app}}(G, v)]} \leq n^\kappa .$$

از قضیهٔ ۶.۳ و مثال بالا درمی‌یابیم که جواب این سؤال باید عددی در بازهٔ $[1/3, 2/3]$ باشد.

چرا در مثال بالا پروتکل ناهمزمان انقدر سریع‌تر عمل می‌کند؟ به طور شهودی می‌توان چنین پاسخ داد: دلیل اصلی همان لم ۸.۳ هستش که بیان می‌کند پروتکل ناهمزمان به طور متوسط به $O(1/\sqrt{k})$ واحد زمانی نیاز دارد تا خبر را از یک رأس سفید به بعدی برساند، ولی پروتکل همزمان به ۲ واحد نیاز دارد. دو متغیر تصادفی مثل G و E در نظر بگیرید، که G هندسی با پارامتر $1/2$ و E نمایی با نرخ $1/2$ باشد. هر دوی این متغیرها امیدریاضی‌شان ۲ است، ولی E واریانس‌اش

هشت برابر G است. یعنی احتمال این که E خیلی کوچکتر از ۲ باشد خیلی بیشتر از احتمال مشابه برای G است. البته مشابهاً احتمال این که E خیلی بزرگتر از ۲ باشد هم خیلی بیشتر از احتمال مشابه برای G است، ولی در این جا این نکته به چشم نمی آید، چون k مسیر مجزا بین دو رأس سفید داریم، و کافی است یکی از آن ها خیلی سریع عمل کند تا خبر به سرعت بین دو رأس سفید جابجا شود. در حالت ناهمزمان که تأخیر یال ها نمایی است، احتمال این که یالی خیلی سریع عمل کند بیشتر از حالت هندسی است و همین باعث می شود امید ریاضی زمان انتقال بین دو رأس سفید در حالت ناهمزمان بسیار کوچکتر باشد.

۴ سایر جنبه ها

در این نوشتار سه پروتکل برای پخش شایعات در شبکه ها را معرفی کردیم و زمان پخش آن ها روی گراف های مختلفی را تحلیل کردیم. این مسئله از جنبه های دیگری نیز قابل بررسی است که در این بخش برخی از آن ها را مختصراً بیان می کنیم.

۱. فرض کنید یک گراف G و رأس v از آن به ما داده شده باشد. چگونه می توانیم تخمین مناسبی از زمان پخش شایعه با شروع از v در یکی از سه پروتکل معرفی شده ارائه دهیم؟ یک راه این است که پروتکل را چندین بار شبیه سازی کنیم و در این صورت میانگین زمان های پخش تخمین مناسبی از امید ریاضی زمان پخش می باشد. با افزایش تعداد شبیه سازی ها می توان دقت این فرایند را به دلخواه زیاد کرد. یک مسئله جالب اینست: یک الگوریتم تقریبی غیر تصادفی^{۲۵} برای امید ریاضی زمان پخش در یکی از سه پروتکل معرفی شده ارائه کنید.

۲. زمان پخش پروتکل های شایعه پراکنی روی گراف های تصادفی مختلفی تحلیل شده است (حتی در یکی از اولین مقاله های مربوط به پروتکل ردکن بره یعنی [۲۰]). برای نتایج به دست آمده روی گراف های تصادفی کلاسیک^{۲۶} مقاله [۳۸] (برای پروتکل ردکن بره) و [۳۹] (برای پروتکل بدهستان ناهمزمان) را ببینید. برای نتایج به دست آمده روی گراف های منظم تصادفی^{۲۷} مقالات [۲۲، ۳۸] (برای پروتکل ردکن بره) و [۳] (برای پروتکل بدهستان ناهمزمان) را ببینید. برای نتایج به دست آمده برای پروتکل ردکن بره روی گراف های هندسی تصادفی^{۲۸} مقاله [۲۵] را ببینید. برای نتایج به دست آمده روی گراف های باراباسی-آلبرت^{۲۹} مقاله [۱۰] (برای پروتکل بدهستان همزمان) و [۱۱] (برای پروتکل بدهستان ناهمزمان) را ببینید. برای نتایج به دست آمده روی گراف های چانگ-لو^{۳۰} مقاله [۲۳] (برای پروتکل های بدهستان) را ببینید. برای نتایج به دست آمده روی k -درخت های تصادفی^{۳۱} مقاله [۳۷] (برای پروتکل بدهستان همزمان) را ببینید.

۳. یک شبکه اجتماعی واقعی در طول زمان تغییر می کند. مقالات [۱۶، ۲۹] عملکرد پروتکل های شایعه پراکنی را روی

deterministic approximation algorithm^{۲۵}

Erdős-Rényi random graphs^{۲۶}

random regular graphs^{۲۷}

random geometric graphs^{۲۸}

Barabási-Albert preferential attachment graphs^{۲۹}

Chung-Lu random graphs with given expected degrees^{۳۰}

random k -trees^{۳۱}

شبکه‌های در حال تغییر تحلیل کرده‌اند.

۴. یک پروتکل شایعه‌پراکنی خوب، نه تنها سریع خبر را پخش می‌کند بلکه تعداد پیام‌هایی که ردوبدل می‌شود هم در آن کم است. در مقاله [۲۴] تعداد پیام‌ها در برخی پروتکل‌ها تحلیل شده است.

۵. در یک شبکه واقعی که از پردازنده‌ها تشکیل شده ممکن است گاهی پردازنده‌ها درست کار نکنند و پیام را صحیح انتقال ندهند. در مقاله [۲۱] مقاومت این پروتکل‌ها در برابر چنین خطاهایی ارزیابی شده است. در مقاله [۲] پروتکل جدیدی طراحی شده با این هدف که در برابر چنین خطاهایی مقاوم‌تر باشد.

۶. برخی مقالات این پروتکل‌ها را روی گراف‌های مختلفی شبیه‌سازی کرده‌اند و عملکرد آن‌ها را به صورت تجربی و نه ریاضی تحلیل کرده‌اند، از جمله [۱۲، ۴۰].

۷. همه پروتکل‌هایی که در این جا معرفی کردیم تصادفی بودند ولی می‌دانیم دستگاه‌های محاسباتی امروز نمی‌توانند اعداد واقعاً تصادفی را تولید کنند بنابراین می‌توان سؤال کرد که آیا می‌توان پروتکل‌های قطعی خوبی طراحی کرد یا لااقل پروتکلی که از تعداد بیت‌های تصادفی کمتری استفاده می‌کند؟ در مقاله [۳۳] یک پروتکل قطعی معرفی شده و در [۳۱، ۱۴، ۳۰] پروتکل‌هایی که از بیت‌های تصادفی کمتری استفاده می‌کنند.

۸. در پروتکل‌هایی که معرفی کردیم فرض کردیم رئیس حافظه‌ای ندارند. آیا می‌توان با اضافه کردن حافظه به رئیس این پروتکل‌ها را بهتر کرد؟ مثلاً اگر یک رأس خبر را دو بار متوالی برای یک همسایه نفرستد، آیا پروتکل بهتر می‌شود؟ در مقاله [۱۷] این دست سؤالات بررسی شده‌اند.

۹. روش‌های جبری برای طراحی پروتکل‌های شایعه‌پراکنی در [۳۲] مورد بررسی قرار گرفته‌اند.

۱۰. مسئله‌ای که در این مقاله بررسی کردیم این بود که یک رأس خبری را می‌دانست و می‌خواست آن را به بقیه هم انتقال دهد. حال فرض کنید هر رأسی یک خبر متمایزی دارد و می‌خواهد آن را به بقیه رئیس اطلاع دهد. مسئله طراحی پروتکل‌های مناسب برای حل این مسئله در [۴، ۶] بررسی شده است.

مراجع

- [1] Acan, H., Collevocchio, A., Mehrabian, A., and Wormald, N. On the push&pull protocol for rumour spreading. submitted, 2014.
- [2] Alistarh, D., Gilbert, S., Guerraoui, R., and Zadimoghaddam, M. How efficient can gossip be? (On the cost of resilient information exchange). in *Automata, languages and programming (ICALP'10). Part II*, vol. 6199 of *Lecture Notes in Comput. Sci.*, pp. 115–126. Springer, Berlin, 2010.

- [3] Amini, H., Draief, M., and Lelarge, M. Flooding in weighted sparse random graphs. *SIAM J. Discrete Math.*, 27(1):1–26, 2013.
- [4] Berenbrink, P., Czyzowicz, J., Elsässer, R., and Gasieniec, L. Efficient information exchange in the random phone-call model. in *Automata, languages and programming (ICALP’10). Part II*, vol. 6199 of *Lecture Notes in Comput. Sci.*, pp. 127–138. Springer, Berlin, 2010.
- [5] Boyd, S., Ghosh, A., Prabhakar, B., and Shah, D. Randomized gossip algorithms. *IEEE Transactions on Information Theory*, 52(6):2508–2530, 2006.
- [6] Censor-Hillel, K., Haeupler, B., Kelner, J. A., and Maymounkov, P. Global computation in a poorly connected world: fast rumor spreading with no dependence on conductance. in *44th Symp. Theory of Computing (STOC)*, pp. 961–970, 2012.
- [7] Chierichetti, F., Lattanzi, S., and Panconesi, A. Almost tight bounds for rumour spreading with conductance. in *Proc. 42nd Symp. Theory of Computing (STOC)*, pp. 399–408, 2010.
- [8] Chierichetti, F., Lattanzi, S., and Panconesi, A. Rumor spreading in social networks. *Theoretical Computer Science*, 412(24):2602 – 2610, 2011. Selected Papers from 36th International Colloquium on Automata, Languages and Programming (ICALP 2009).
- [9] Demers, A., Greene, D., Hauser, C., Irish, W., Larson, J., Shenker, S., Sturgis, H., Swinehart, D., and Terry, D. Epidemic algorithms for replicated database maintenance. in *Proc. 6th Symp. Principles of Distributed Computing (PODC)*, pp. 1–12, 1987.
- [10] Doerr, B., Fouz, M., and Friedrich, T. Social networks spread rumors in sublogarithmic time. in *Proc. 43th Symp. Theory of Computing (STOC)*, pp. 21–30, 2011.
- [11] Doerr, B., Fouz, M., and Friedrich, T. Asynchronous rumor spreading in preferential attachment graphs. in *Proc. 13th Scandinavian Workshop Algorithm Theory (SWAT)*, pp. 307–315, 2012.
- [12] Doerr, B., Fouz, M., and Friedrich, T. Experimental analysis of rumor spreading in social networks. in *Design and analysis of algorithms*, vol. 7659 of *Lecture Notes in Comput. Sci.*, pp. 159–173. Springer, Heidelberg, 2012.
- [13] Doerr, B., Fouz, M., and Friedrich, T. Why rumors spread so quickly in social networks. *Commun. ACM*, 55(6):70–75, 2012.

- [14] Doerr, B., Friedrich, T., and Sauerwald, T. Quasirandom rumor spreading. in *Proceedings of the Nineteenth Annual ACM-SIAM Symposium on Discrete Algorithms (SODA'08)*, pp. 773–781. ACM, New York, 2008.
- [15] Durrett, R. Stochastic growth models: recent results and open problems. in *Mathematical approaches to problems in resource management and epidemiology (Ithaca, NY, 1987)*, vol. 81 of *Lecture Notes in Biomath.*, pp. 308–312. Springer, Berlin, 1989.
- [16] Dutta, C., Pandurangan, G., Rajaraman, R., Sun, Z., and Viola, E. On the complexity of information spreading in dynamic networks. in *Proceedings of the Twenty-Fourth Annual ACM-SIAM Symposium on Discrete Algorithms, SODA '13*, pp. 717–736. SIAM, 2013.
- [17] Elsässer, R. and Sauerwald, T. The power of memory in randomized broadcasting. in *Proc. 19th Symp. Discrete Algorithms (SODA)*, pp. 218–227, 2008.
- [18] Elsässer, R. and Sauerwald, T. Cover time and broadcast time. in *STACS 2009: 26th International Symposium on Theoretical Aspects of Computer Science*, vol. 3 of *LIPICs. Leibniz Int. Proc. Inform.*, pp. 373–384. Schloss Dagstuhl. Leibniz-Zent. Inform., Wadern, 2009.
- [19] Elsässer, R. and Sauerwald, T. On the runtime and robustness of randomized broadcasting. *Theoret. Comput. Sci.*, 410(36):3414–3427, 2009.
- [20] Feige, U., Peleg, D., Raghavan, P., and Upfal, E. Randomized broadcast in networks. *Random Struct. Algorithms*, 1(4):447–460, 1990.
- [21] Fountoulakis, N., Huber, A., and Panagiotou, K. Reliable broadcasting in random networks and the effect of density. in *Proc. 29th IEEE Conf. Computer Communications (INFOCOM)*, pp. 2552–2560, 2010.
- [22] Fountoulakis, N. and Panagiotou, K. Rumor spreading on random regular graphs and expanders. in *Proc. 14th Intl. Workshop on Randomization and Comput. (RANDOM)*, pp. 560–573, 2010.
- [23] Fountoulakis, N., Panagiotou, K., and Sauerwald, T. Ultra-fast rumor spreading in social networks. in *23th Symp. Discrete Algorithms (SODA)*, pp. 1642–1660, 2012.
- [24] Fraigniaud, P. and Giakkoupis, G. On the bit communication complexity of randomized rumor spreading. in *Proceedings of the Twenty-second Annual ACM Symposium on Parallelism in Algorithms and Architectures, SPAA '10*, pp. 134–143, New York, NY, USA, 2010. ACM.

- [25] Friedrich, T., Sauerwald, T., and Stauffer, A. Diameter and broadcast time of random geometric graphs in arbitrary dimensions. *Algorithmica*, 67(1):65–88, 2013.
- [26] Giakkoupis, G. Tight bounds for rumor spreading in graphs of a given conductance. in *Proc. 28th Symp. Theoretical Aspects of Computer Science (STACS)*, pp. 57–68, 2011.
- [27] Giakkoupis, G. Tight bounds for rumor spreading with vertex expansion. in *Proc. 25th Symp. Discrete Algorithms (SODA)*, pp. 801–815, 2014.
- [28] Giakkoupis, G. and Sauerwald, T. Rumor spreading and vertex expansion. in *Proc. 23th Symp. Discrete Algorithms (SODA)*, pp. 1623–1641, 2012.
- [29] Giakkoupis, G., Sauerwald, T., and Stauffer, A. Randomized rumor spreading in dynamic graphs. in Esparza, Fraigniaud, Husfeldt, and Koutsoupias, eds. , *Automata, Languages, and Programming (ICALP'14)*, vol. 8573 of *Lecture Notes in Computer Science*, pp. 495–507. Springer Berlin Heidelberg, 2014.
- [30] Giakkoupis, G. and Woelfel, P. On the randomness requirements of rumor spreading. in *Proceedings of the Twenty-Second Annual ACM-SIAM Symposium on Discrete Algorithms (SODA '11)*, pp. 449–461. SIAM, Philadelphia, PA, 2011.
- [31] Guo, Z. and Sun, H. Gossip vs. markov chains, and randomness-efficient rumor spreading. in *Proc. 26th Symp. Discrete Algorithms (SODA)*, 2015.
- [32] Haeupler, B. Analyzing network coding gossip made easy. in *STOC'11—Proceedings of the 43rd ACM Symposium on Theory of Computing*, pp. 293–302. ACM, New York, 2011.
- [33] Haeupler, B. Simple, fast and deterministic gossip and rumor spreading. in *Proceedings of the Twenty-Fourth Annual ACM-SIAM Symposium on Discrete Algorithms, SODA '13*, pp. 705–716. SIAM, 2013.
- [34] Howard, C. D. Models of first-passage percolation. in *Probability on discrete structures*, vol. 110 of *Encyclopaedia Math. Sci.*, pp. 125–173. Springer, Berlin, 2004.
- [35] Janson, S. One, two and three times $\log n/n$ for paths in a complete graph with random weights. *Combin. Probab. Comput.*, 8(4):347–361, 1999.
- [36] Karp, R., Schindelhauer, C., Shenker, S., and Vöcking, B. Randomized Rumor Spreading. in *41st Symp. Foundations of Computer Science (FOCS)*, pp. 565–574, 2000.

- [37] Mehrabian, A. and Pourmiri, A. Randomized rumor spreading in poorly connected small-world networks. in Kuhn, F., ed. , *Distributed Computing (DISC '14)*, vol. 8784 of *Lecture Notes in Computer Science*, pp. 346–360. Springer Berlin Heidelberg, 2014. full version available on <http://arxiv.org/abs/1410.8175>.
- [38] Panagiotou, K., Pérez-Giménez, X., Sauerwald, T., and Sun, H. Randomized rumour spreading: The effect of the network topology. *Combinatorics, Probability and Computing*, 24:457–479, March 2015.
- [39] Panagiotou, K. and Speidel, L. Asynchronous rumor spreading on random graphs. in Cai, Leizhen, Cheng, Siu-Wing, and Lam, Tak-Wah, eds. , *Algorithms and Computation*, vol. 8283 of *Lecture Notes in Computer Science (ISAAC'13)*, pp. 424–434. Springer Berlin Heidelberg, 2013.
- [40] Patsonakis, C. and Roussopoulos, M. Asynchronous rumour spreading in social and signed network. *arXiv*, 1310.6119v3 [cs.SI], 2015.
- [41] Pittel, B. On spreading a rumor. *SIAM Journal on Applied Mathematics*, 47(1):213–223, 1987.