On the push&pull protocol for rumour spreading

Hüseyin Acan (Rutgers), Andrea Collevecchio (Monash), Abbas Mehrabian (UBC), and Nick Wormald (Monash)

The model

1. The ground is a simple connected n-vertex graph.
2. Initially, one vertex knows a rumour.
3. Every informed vertex sends the rumour to a random neighbour (PUSH);
   and every uninformed vertex queries a random neighbour about the rumour (PULL).

synchronous variant. at each time-step $1, 2, \ldots$, each vertex performs an operation (PUSH or PULL)[Demers, Gealy, Greene, Hauser, Irish, Larson, Manning, Shenker, Sturgis, Swinehart, Terry, Woods'87].

asynchronous variant. at each time-step $1/n, 2/n, \ldots$, one random vertex performs an operation [Boyd, Ghosh, Prabhakar, Shah'06].

$s(G)$ and $a(G)$: expected time to broadcast the rumour.

Applications

1. Integrity of Replicated databases
2. News propagation in social networks
3. Spread of viruses on the Internet
4. First-passage-percolation with i.i.d. exponential weights
5. Richardson's model for disease spread

Known results

Graph G	$s(G)$	$a(G)$
Star	2	$\ln n + O(1)$
Path	$(4/3)n + O(1)$	$n + O(1)$
Double star	$(1+o(1))n/4$	$(1+o(1))n/4$
Complete	$(1+o(1))\log_3 n$ [Karp,Schindelhauer,Shenker,Vöcking'00]	$\ln n + o(1)$
Hypercube graph	$\Theta(\ln n)$ [Feige-Peleg-Raghavan-Upfal'90]	$\Theta(\ln n)$ [Fill,Pemantle'93]
$\mathcal{G}(n,p)$ (connected)	$\Theta(\ln n)$ [Feige-Peleg-Raghavan-Upfal'90]	$(1+o(1))\ln n$ [Panagiotou,Speidel'13]
General	$O(n \ln n)$ [Feige-Peleg-Raghavan-Upfal'90]	$O(n \ln n)$ [Feige-Peleg-Raghavan-Upfal'90]

- Random regular graphs, expander graphs, Barabási-Albert graphs, Chung-Lu graphs: $s(G), a(G) = \Theta(\log n)$.
- Tight upper bounds for $s(G)$ in terms of expansion profile [Giakkoupis'11,'14].

OUR RESULTS

http://arxiv.org/abs/1411.0948, Proceedings of PODC 2015

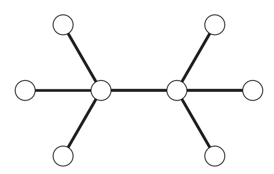
For any connected G on n vertices,

$$s(G) < 5n,$$
$$\ln(n)/5 < a(G) < 4n,$$
$$\frac{1}{\ln n} < \frac{s(G)}{a(G)} < 200 n^{2/3} \ln n,$$

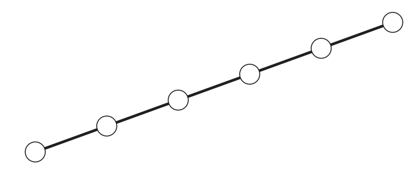
and for infinitely many graphs this ratio is $\widetilde{\Omega}\left(n^{1/3}\right)$.

Remarks

- Giakkoupis, Nazari, Woelfel'16 improved upper bound $O\left(n^{1/2}\right)$.
- Asymptotic tightness of linear upper bounds for $s(G), a(G)$:



$s(G), a(G) \sim n/4$

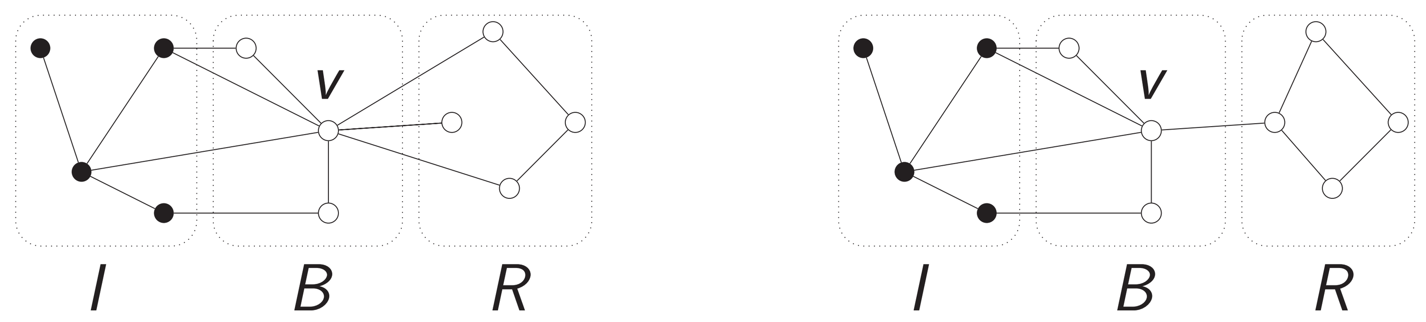


$s(G) \sim 4n/3, a(G) \sim n$

- An alternative viewpoint of the asynchronous variant: every vertex has an independent rate-1 Poisson process, and at times of process performs an operation (PUSH or PULL).

Proof idea for linear upper bound $a(G) < 4n$

Only pull operations are needed!

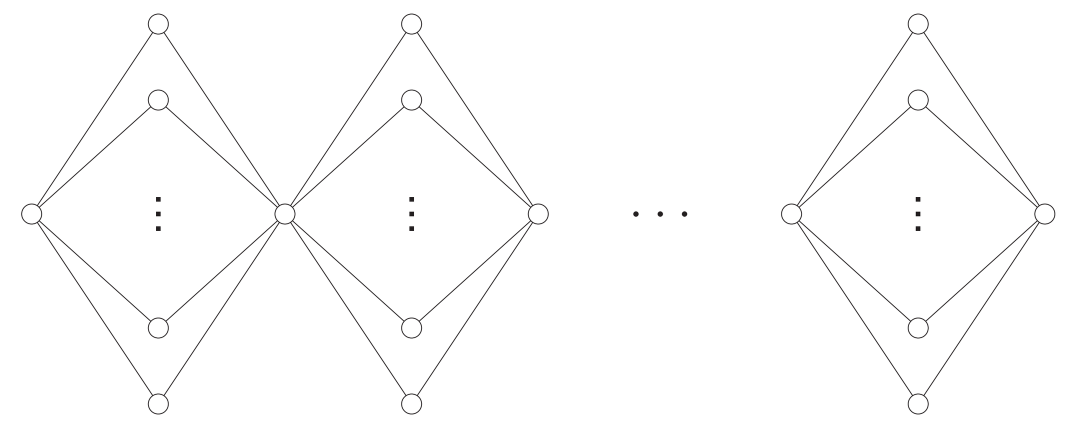


Black vertices are informed, white ones are uninformed. We show inductively the expected remaining time $\le 2|B| + 4|R|$.

Left: there is some boundary vertex v with $\deg_R(v) > \deg_B(v)$: it may take a lot of time to inform v, but once it is informed, $R \downdownarrows$ and $B \upuparrows$.

Right: otherwise, each boundary vertex has pulling rate $\ge 1/2|B|$, and the B boundary vertices work together "in parallel" and average time for one of them to pull the rumour is 2.

Example with $a(G) \ll s(G)$



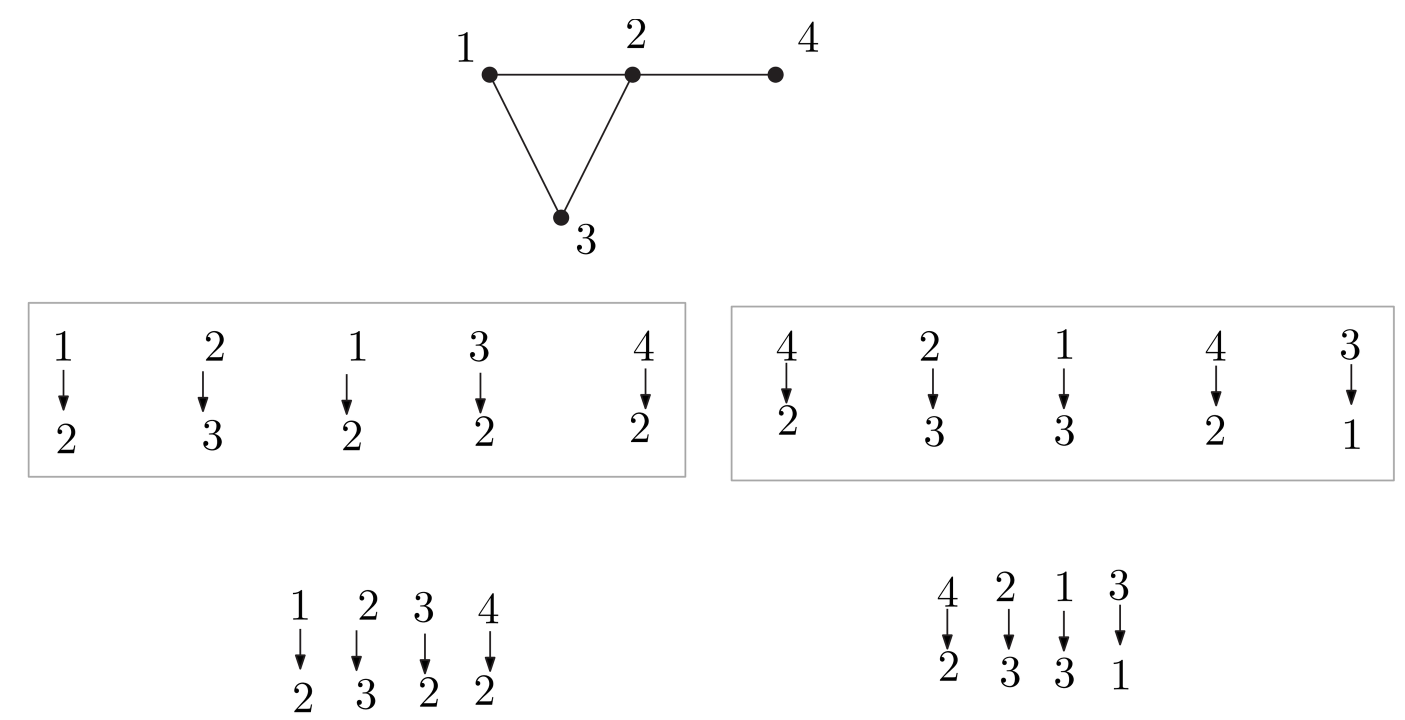
$n^{1/3}$ diamonds, each consisting of $n^{2/3}$ paths.

$$a(G) \le n^{1/3} \times \frac{5}{\sqrt{n^{2/3}}} + \ln n \ll 2n^{1/3} \le s(G),$$

using a birthday paradox argument.

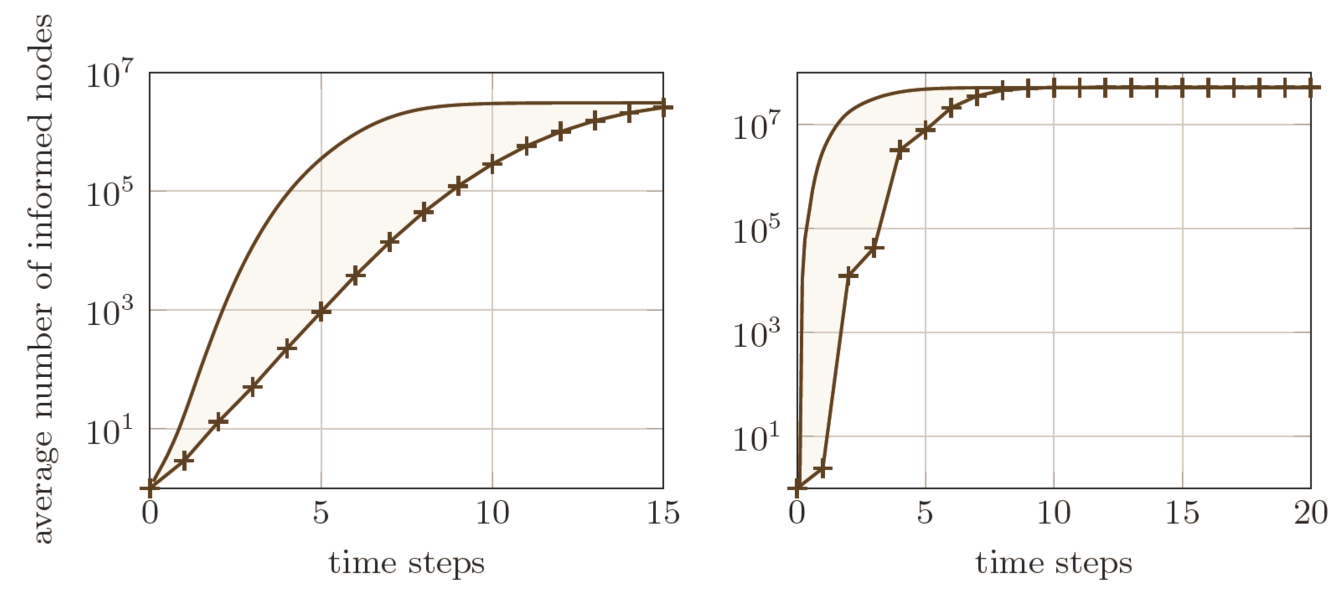
Proof idea for $a(G) < s(G) \times \ln n$

Consider an arbitrary calling sequence:



Using a coupon collector argument, the average length of each block is $n \ln n$, and each step needs time $1/n$.

Experimental comparison of two variants



(a) Orkut network (b) Twitter network

Plots from: Doerr, Fouz, and Friedrich. MedAlg 2012.