Talk is based on joint work with Acan, Collevecchio, and Wormald (paper is available on arXiv: Search for ‘On the push&pull protocol for rumour spreading’)

Definitions

action v calls a random neighbour and they exchange info.

synchronous push&pull each vertex performs an action at times $1, 2, 3, \ldots$

asynchronous push&pull each vertex performs an action at times of a Poisson process with rate 1 (continuous time model)

At times $1/n, 2/n/\ldots$, a random vertex performs an action (discrete time model)

$\mathrm{ast}_{\mathsf{s}}(G)$**,** $\mathrm{ast}_{\mathsf{s}}(G)$ average time to spread the rumour.

Remarks, observations and applications

- Choice of starting vertex, we assume worst.
- In synchronous, all act at the same time!
- No memory!
- Replicated database, broadcasting in distributed computing, news propagation in social networks, Sensor networks, etc.
- Connections with first-passage-percolation and Richardson’s model for the spread of a disease

Observation (asynch. continuous): For a cut (S, S^c), expected time for rumour to pass

$$\left(\sum_{(u,v)\in S\times S^c} \left(\frac{1}{\deg(u)} + \frac{1}{\deg(v)} \right) \right)^{-1} = \left(\sum_{u\in S} \frac{\deg(u, S^c)}{\deg(u)} + \sum_{v\in S^c} \frac{\deg(v, S)}{\deg(v)} \right)^{-1}$$

Graph G	$\text{ast}_{\mathsf{s}}(G)$	$\text{ast}_{\mathsf{a}}(G)$
Path	$(4/3)n + O(1)$	$n + O(1)$
Star	2	$\log n + O(1)$
Complete	$\sim \log_3 n$ [Karp,Schindelhauer,Shenker,Vöcking'00]	$\log n + o(1)$
General	$O\left(\min\{\Phi(G)^{-1} \cdot \log n, \alpha(G)^{-1} \cdot \log^2 n\}\right)$ [Giakkoupis'11,14]	
Hypercube	$\Theta(\log n)$ [Feige,Peleg,Raghavan,Upfal'90]	$\Theta(\log n)$ [Fill,Pemantle'93]
$\mathcal{G}(n,p)$ $1 < \frac{np}{\log n}$ fixed	$\Theta(\log n)$ [Feige,Peleg,Raghavan,Upfal'90]	$\sim \log n$ [Panagiotou,Speidel'13]
$\mathcal{G}(n,d)$ $2 < d$ fixed	$\Theta(\log n)$ [Fountoulakis,Panagiotou'10]	$\sim (\log n)(d-1)/(d-2)$ [Amini,Draief,Lelarge'13]
preferential attachment (Barabási-Albert)	$\Theta(\log n)$ [Doerr,Fouz,Friedrich'11]	$\Theta(\log n)$ [Doerr,Fouz,Friedrich'12]
Random geometric graphs in $\left[0, n^{1/d}\right]^d$	$\Theta(n^{1/d}/r + \log n)$ [Friedrich,Sauerwald,Stauffer'13]	
Necklace graph $N_{m,k}$	$\Omega(m)$	$O(\log n + m/\sqrt{k})$

Notation. $\sim$ means equality up to a $1 + o(1)$ factor

$$\alpha(G) = \min\left\{\frac{|\partial S|}{|S|} : S \subseteq V(G), 0 < |S| \leq |V(G)|/2\right\}$$
$$\Phi(G) = \min\left\{\frac{e(S, V(G) \setminus S)}{\mathtt{vol}(S)} : S \subseteq V(G), 0 < \mathtt{vol}(S) \leq \mathtt{vol}(V(G))/2\right\},$$

where given $S \subseteq V(G)$, ∂S is the set of vertices in $V(G) \setminus S$ that have a neighbour in S, $e(S, V(G) \setminus S)$ is number of edges between S and $V(G) \setminus S$, and $\mathtt{vol}(S) = \sum_{u \in S} \deg(u)$.

Remark. For $\mathcal{G}(n,p)$, finer results were proved recently for the push protocol by [Panagiotou, Pérez-Giménez, Sauerwald, Sun'14].

Our results. [Acan, Collevecchio, Mehrabian and Wormald'14] For any simple undirected connected n-vertex graph G,

$$\frac{\log n}{3} \leq \text{ast}_{\mathsf{a}}(G) \leq 4n$$
$$2 \leq \text{ast}_{\mathsf{s}}(G) \leq 4.6n$$
$$\frac{C}{\log^2 n} \leq \frac{\text{ast}_{\mathsf{s}}(G)}{\text{ast}_{\mathsf{a}}(G)} \leq C' n^{2/3} \log n, \text{ and this ratio can be } \Omega\left(\sqrt[3]{\frac{n}{\log n}}\right).$$

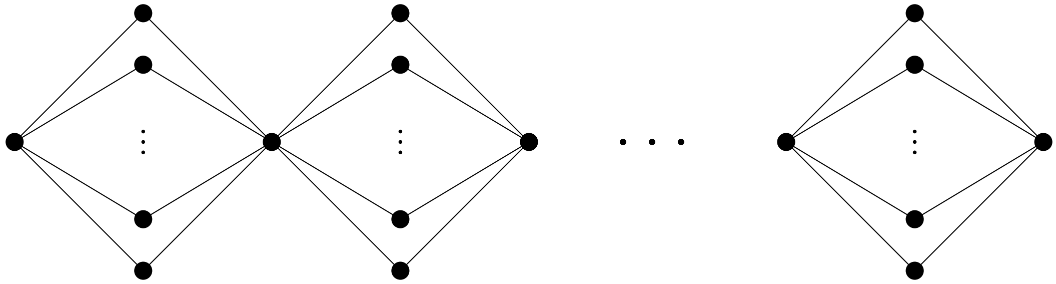
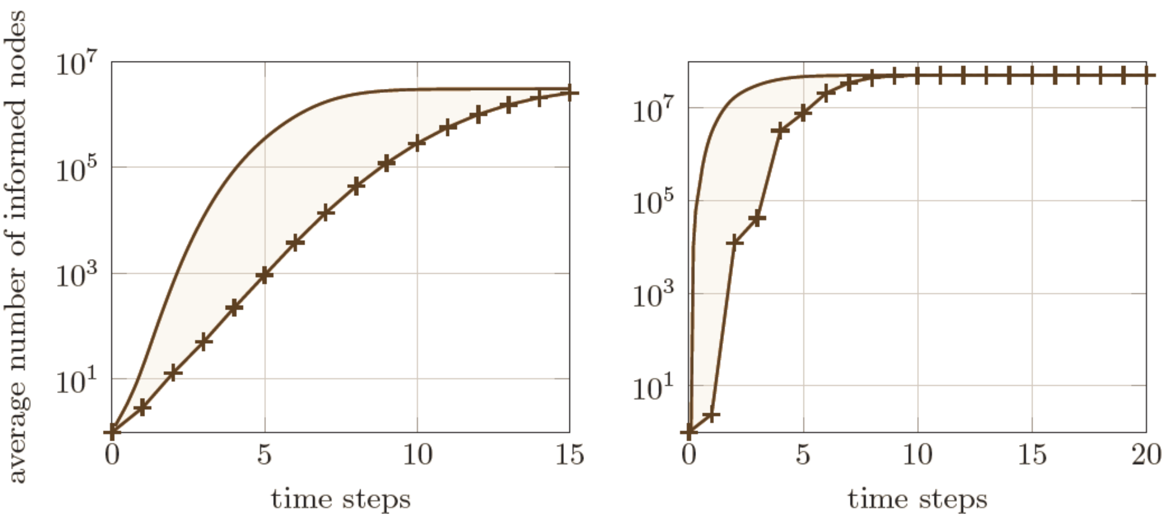


Figure 1: Necklace graph $N_{m,k}$: m diamonds of size k



(a) Orkut network **(b)** Twitter network

Figure 5: Comparison of synchronous (—+—) and asynchronous (——) rumor spreading without memory on two real social networks. The x-axis corresponds to the time steps (in the synchronous setting) or the time (in the asynchronous setting). The y-axis corresponds to the number of informed vertices after this time, averaged over 1000 runs for the Orkut network and 50 runs for the Twitter network.
In both cases, the asynchronous counterparts spread the rumor significantly faster than the synchronous models.

(Figure 5 is from Doerr, Fouz, and Friedrich: Experimental analysis of rumor spreading in social networks. First Mediterranean Conference on Algorithms, MedAlg 2012.)

Examples

Star graph

2 in synchronous, $\sim \log n$ in asynchronous

Path graph

Synchronous: 1 step to pass the boundary edges, Geo(3/4) steps to pass the middle edges... average total is $(4/3)n - 2$

Asynchronous: Exp(3/2) to pass the first edge, Exp(1) to pass the middle edges ... average total is $n - 5/3$.

Necklace graph $N_{m,k}$

The number of vertices is $n = km + m + 1$, there are m hubs

Let $Z :=$ time rumour passes between two hubs $= \min\{Z_1, \ldots, Z_k\}$ and $Z_1 \overset{s}{\leq} \mathrm{Exp}(1/2) + \mathrm{Exp}(1/2)$ so

$$\mathbb{P}\left[Z > t\right] = \mathbb{P}\left[Z_1 > t\right]^k \leq \left(1 - \mathbb{P}\left[\mathrm{Exp}(1/2) \leq t/2\right]^2\right)^k = (2e^{-t/4} - e^{-t/2})^k$$

and integration gives $\mathbb{E}\left[Z\right] = O(1/\sqrt{k})$. So $\mathrm{ast}_{\mathsf{a}}(G) = O(m/\sqrt{k}) + O(\log n)$.

Choosing $k = \Theta\left((n/\log n)^{2/3}\right)$ and $m = \Theta\left(n^{1/3}(\log n)^{2/3}\right)$ gives

$$\frac{\mathrm{ast}_{\mathsf{s}}(G)}{\mathrm{ast}_{\mathsf{a}}(G)} = \Omega\left(\sqrt[3]{n/\log n}\right)$$

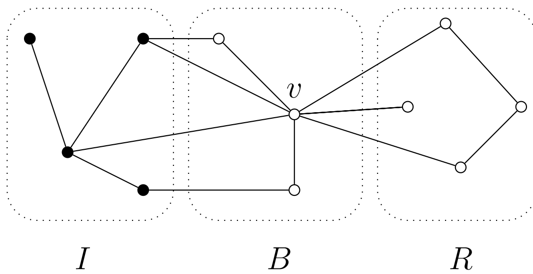


Figure 2: First case

Proofs

Proof of $\mathrm{ast}_{\mathsf{a}}(G) \leq 4n$

We will actually prove this using pull operations only. m uninformed vertices, b of which have informed neighbours (boundary vertices), expected remaining time $\leq 4m - 2b$.

I = informed vertices
B = boundary vertices, $b = |B|$
R = remaining vertices
$|B| + |R| = m$. Let $d(v)$ denote the degree of v in G and, for a set S of vertices, let $d_S(v)$ denote the number of neighbours of v in S. We consider two cases.

Case 1. $\exists v \in B$ with $d_R(v) \geq d_B(v)$. We wait for v to be informed. The expected time taken for v to pull the rumour from vertices in I is

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

Once v is informed, m decreases by 1, b increases by $d_R(v) - 1$.

$$1 + 2d_R(v) + 4(m-1) - 2(b + d_R(v) - 1) < 4m - 2b.$$

Case 2. Any boundary vertex v has a 'pulling rate' of

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

The b together have a pulling rate of at least $1/2$ so the expected time until a boundary vertex is informed is at most 2. Once this happens, m decreases by 1 and b either does not decrease or decreases by at most 1, and the inductive hypothesis concludes the proof.

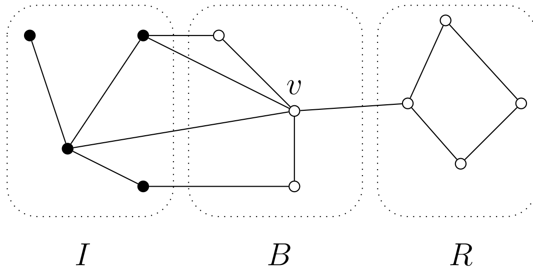


Figure 3: Second case

Proof of last line

What we prove indeed is:

Theorem 1.
$$\frac{C}{\log n} \leq \frac{\mathrm{gst}_{\mathsf{s}}(G)}{\mathrm{gst}_{\mathsf{a}}(G)} \leq C' n^{2/3}$$

where
$$\mathrm{gst}(G) = \inf\{t : \mathbb{P}\left[\textit{spread time} > t\right] < 1/n\}$$

Remark. If G regular, $\mathrm{gst}_{\mathsf{s}}(G) = \Theta(\mathrm{gst}_{\mathsf{a}}(G))$. [Sauerwald'10]

Proof of lower bound

In the aysnch. protocol, in any time-interval of size $4\log n$, w.h.p. all vertices have made a call, i.e. the progress has been like one round of the synch. protocol, hence $\mathrm{gst}_{\mathsf{a}}(G) \leq (8\log n)\,\mathrm{gst}_{\mathsf{s}}(G)$.

Proof of upper bound

For regular graphs G we can prove

$$\mathrm{gst}_{\mathsf{s}}(G) \leq O(\sqrt{n}\,\mathrm{gst}_{\mathsf{a}}(G))$$

Consider the asynchronous version. List the vertices in the order their clocks ring. The list ends once all the vertices are informed. Now consider the natural coupling between the two protocols, the synchronous actions follow the same ordering as in the list. We partition the list into blocks such that in each block no vertex appears twice. the synchronous protocol in each round will inform a superset of the set of vertices informed by the asynchronous variant in any single block. The size of each block is of order $\sqrt{n}$ by the birthday paradox, qed.

For general graphs: Enumerate the callers and callees ... the list has size about $n \times \mathrm{gst}_{\mathsf{a}}(G)$
vertex is *special* : probability a callee $\geq n^{-2/3}$
partition list into blocks, in each block only special vertices can be repeated as callees
The average size of a block $\approx n^{1/3}$, so about $n^{2/3} \times \mathrm{gst}_{\mathsf{a}}(G)$ blocks
The spread time of the synch. protocol $\overset{s}{\leq}$ number of blocks $+n^{2/3}$.

Indeed, we prove

$$\mathrm{gst}_{\mathsf{s}}(G) \leq n^{1-\alpha} + 64\mathrm{gst}_{\mathsf{a}}(G)n^{(1+\alpha)/2} \qquad \forall \alpha \in [0,1)$$