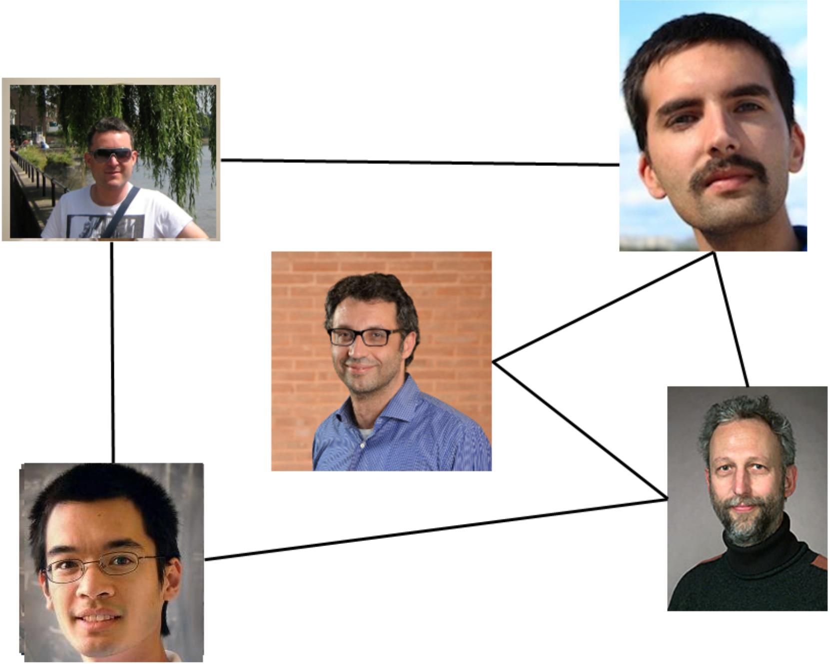
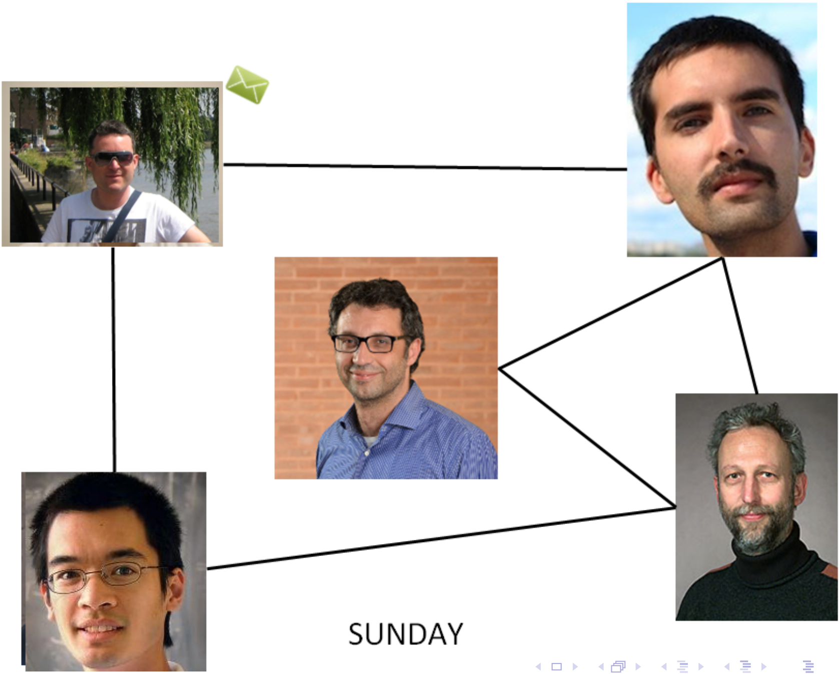
Randomized Rumour Spreading on Random k-trees

Abbas Mehrabian
amehrabi@uwaterloo.ca

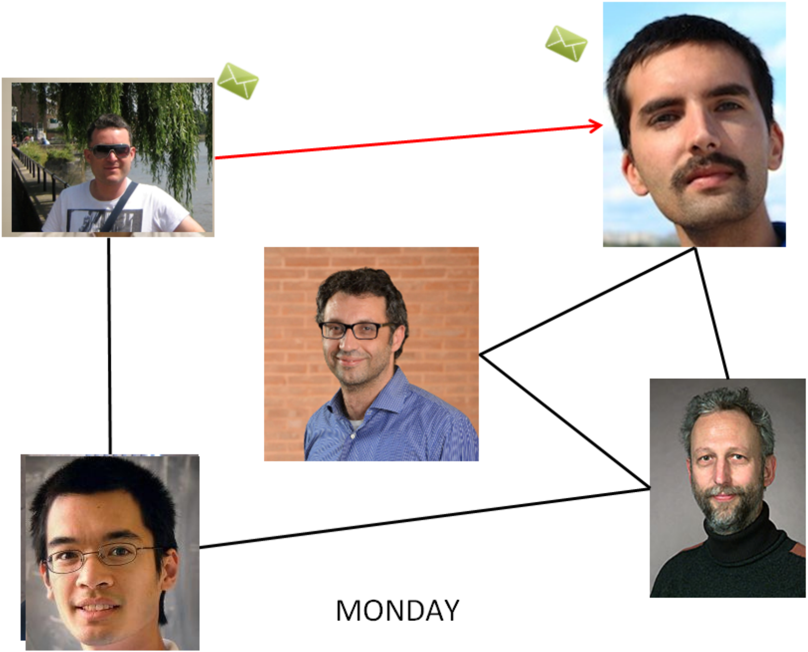
University of Waterloo

13 May 2014
University of Melbourne

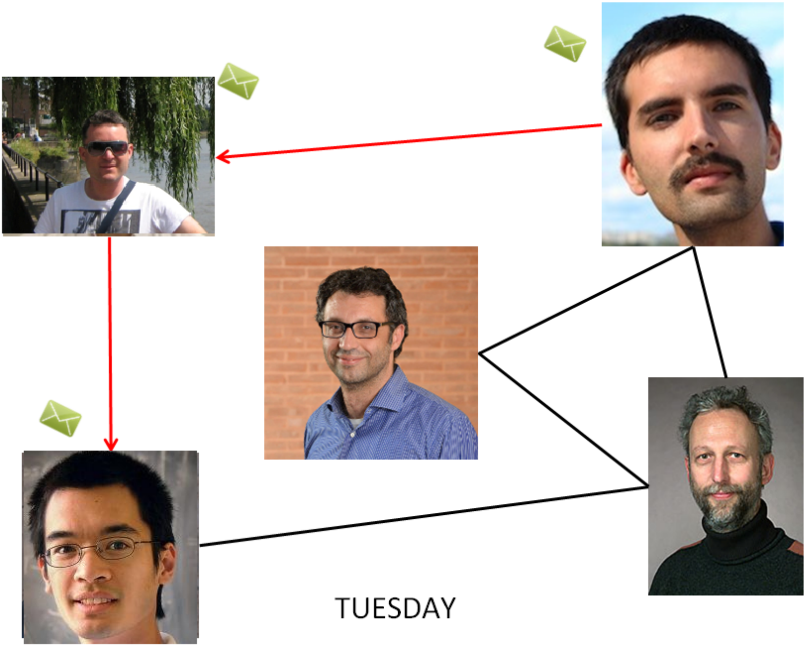




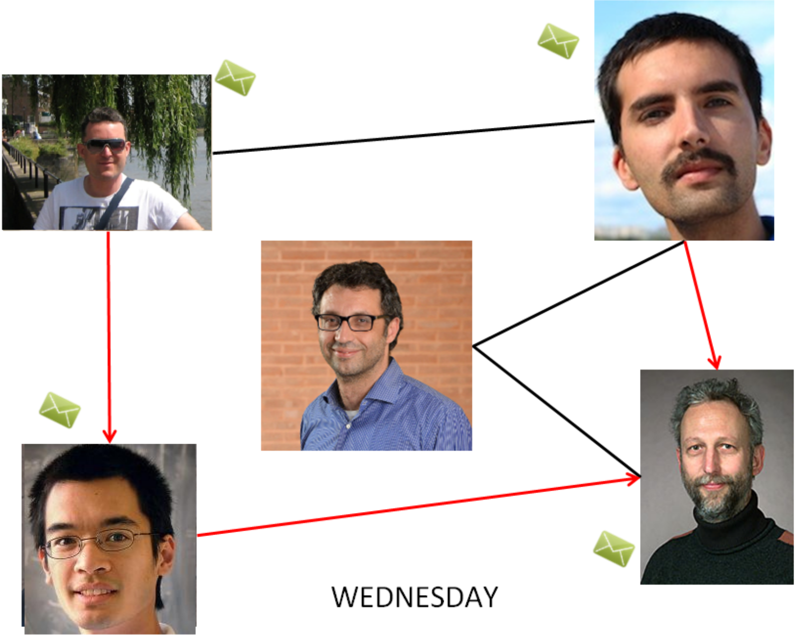
SUNDAY



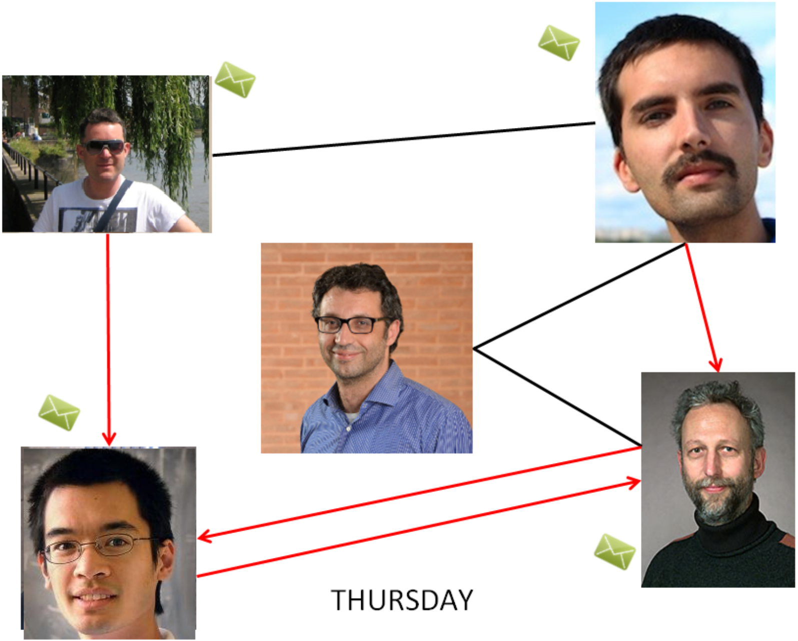
MONDAY

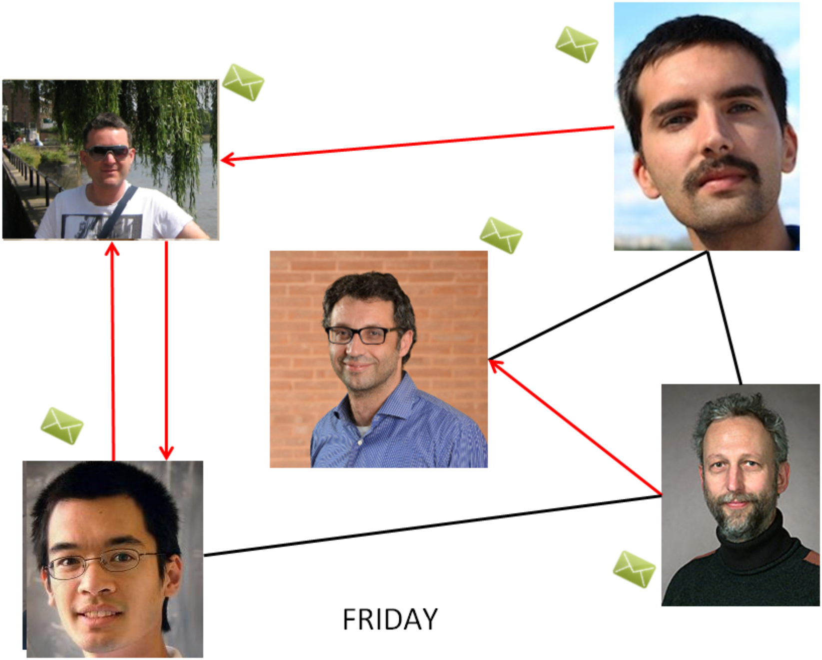


TUESDAY



WEDNESDAY


THURSDAY


FRIDAY

Protocol definition

Demers, Gealy, Greene, Hauser, Irish, Larson, Manning, Shenker, Sturgis, Swinehart, Terry, Woods'87

1. The ground is a simple connected graph.
2. At time 0, one vertex knows a rumour.
3. At each time-step $1, 2, \ldots,$ every informed vertex tells the rumour to a random neighbour.

Protocol definition

Demers, Gealy, Greene, Hauser, Irish, Larson, Manning, Shenker, Sturgis, Swinehart, Terry, Woods'87

1. The ground is a simple connected graph.
2. At time 0, one vertex knows a rumour.
3. At each time-step $1, 2, \ldots,$ every informed vertex tells the rumour to a random neighbour.

Remark 1. Informed vertex may call a neighbour in consecutive steps.
Remark 2. If a vertex receives the rumour at time t, it starts passing it from time $t + 1$.

Protocol definition

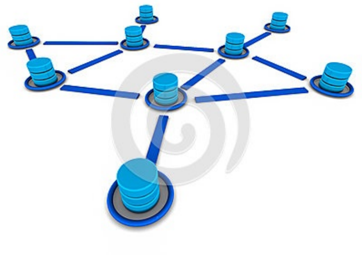
Demers, Gealy, Greene, Hauser, Irish, Larson, Manning, Shenker, Sturgis, Swinehart, Terry, Woods'87

1. The ground is a simple connected graph.
2. At time 0, one vertex knows a rumour.
3. At each time-step $1, 2, \ldots,$ every informed vertex tells the rumour to a random neighbour.

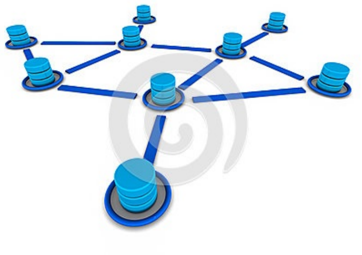
Remark 1. Informed vertex may call a neighbour in consecutive steps.
Remark 2. If a vertex receives the rumour at time t, it starts passing it from time $t + 1$.

inform-time(v): the first time v learns the rumour.
Spread Time: the first time everyone knows the rumour.

Application: distributed computing



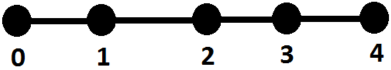
Application: distributed computing



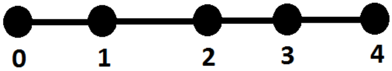
Rumour spreading advantages:

- ✓ Simplicity, locality, no memory
- ✓ Scalability, reasonable link loads
- ✓ Robustness

Example: a path

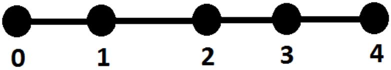


Example: a path



$$\mathrm{inform}-\mathrm{time}(0)=0$$

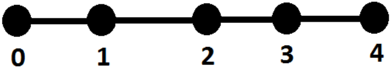
Example: a path



$$\text{inform} - \text{time}(0) = 0$$

$$\text{inform} - \text{time}(1) = 1$$

Example: a path

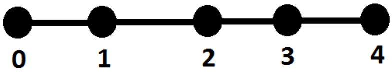


$$\mathrm{inform} - \mathrm{time}(0) = 0$$

$$\mathrm{inform} - \mathrm{time}(1) = 1$$

$$\mathrm{inform} - \mathrm{time}(2) = 1 + \mathrm{Geo}(1/2)$$

Example: a path



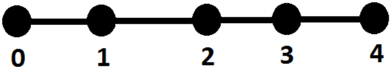
$$\text{inform} - \text{time}(0) = 0$$

$$\text{inform} - \text{time}(1) = 1$$

$$\text{inform} - \text{time}(2) = 1 + \text{Geo}(1/2)$$

$$\text{inform} - \text{time}(3) = 1 + \text{Geo}(1/2) + \text{Geo}(1/2)$$

Example: a path



$$\text{inform} - \text{time}(0) = 0$$

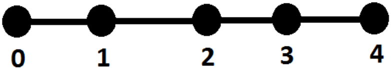
$$\text{inform} - \text{time}(1) = 1$$

$$\text{inform} - \text{time}(2) = 1 + \text{Geo}(1/2)$$

$$\text{inform} - \text{time}(3) = 1 + \text{Geo}(1/2) + \text{Geo}(1/2)$$

$$\text{inform} - \text{time}(4) = 1 + \text{Geo}(1/2) + \text{Geo}(1/2) + \text{Geo}(1/2)$$

Example: a path



$$\mathrm{inform-time}(0) = 0$$

$$\mathrm{inform-time}(1) = 1$$

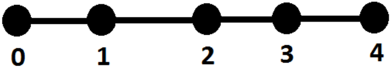
$$\mathrm{inform-time}(2) = 1 + \mathrm{Geo}(1/2)$$

$$\mathrm{inform-time}(3) = 1 + \mathrm{Geo}(1/2) + \mathrm{Geo}(1/2)$$

$$\mathrm{inform-time}(4) = 1 + \mathrm{Geo}(1/2) + \mathrm{Geo}(1/2) + \mathrm{Geo}(1/2)$$

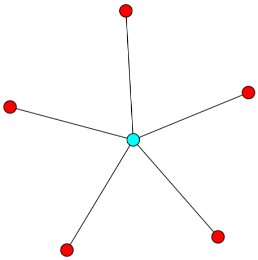
$$\mathbb{E}[\text{Spread Time}] = 1 + 3 \times 2 = 7$$

Example: a path

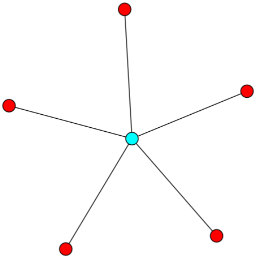


$$\begin{aligned}
\text{inform} - \text{time}(0) &= 0 \\
\text{inform} - \text{time}(1) &= 1 \\
\text{inform} - \text{time}(2) &= 1 + \text{Geo}(1/2) \\
\text{inform} - \text{time}(3) &= 1 + \text{Geo}(1/2) + \text{Geo}(1/2) \\
\text{inform} - \text{time}(4) &= 1 + \text{Geo}(1/2) + \text{Geo}(1/2) + \text{Geo}(1/2) \\
\mathbb{E}[\text{Spread Time}] &= 1 + 3 \times 2 = 7 \\
&= 2n - 3
\end{aligned}$$

Example: a star

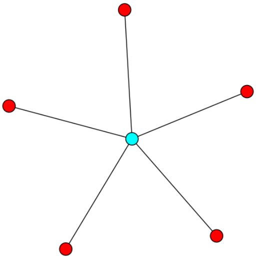


Example: a star



When $k + 1$ vertices are informed and $n - 1 - k$ uninformed, after $\frac{n-1}{n-1-k}$ more rounds a new vertex will be informed.

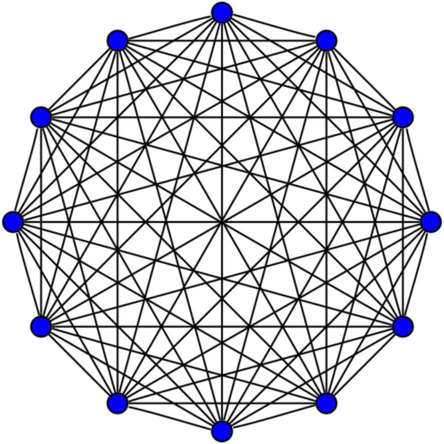
Example: a star



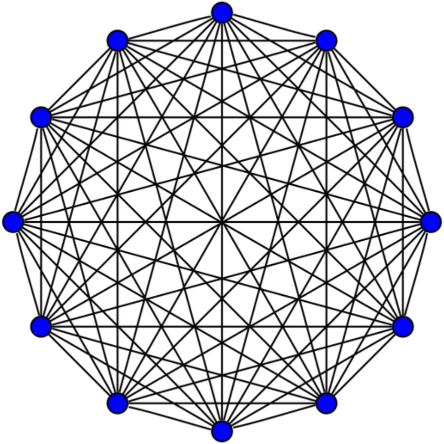
When $k+1$ vertices are informed and $n-1-k$ uninformed, after $\frac{n-1}{n-1-k}$ more rounds a new vertex will be informed.

$$\mathbb{E}[\text{Spread Time}] = \frac{n-1}{n-1} + \frac{n-1}{n-2} + \cdots + \frac{n-1}{2} + \frac{n-1}{1} \approx n \ln n$$

Example: a complete graph



Example: a complete graph



$$\mathbb{E}[\text{Spread Time}] \approx \log_2 n + \ln n \qquad [Pittel'87]$$

Other known results

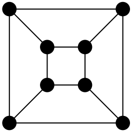
With probability approaching 1, for any starting vertex,

1. $\max\{\text{diameter}(G), \log_2 n\} \leq \text{Spread Time} \leq (1 + o(1))n \ln n$
   [Elsässer and Sauerwald'06]

Other known results

With probability approaching 1, for any starting vertex,

1. $\max\{\text{diameter}(G), \log_2 n\} \leq$ Spread Time $\leq (1 + o(1))n \ln n$ [Elsässer and Sauerwald'06]
2. Spread Time of $\mathcal{H}_d = \Theta(d)$ [Feige, Peleg, Raghavan, Upfal'90]

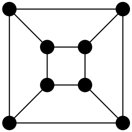


$\mathcal{H}_3$

Other known results

With probability approaching 1, for any starting vertex,

1. $\max\{\text{diameter}(G), \log_2 n\} \leq$ Spread Time $\leq (1 + o(1))n \ln n$ [Elsässer and Sauerwald'06]
2. Spread Time of $\mathcal{H}_d = \Theta(d)$ [Feige, Peleg, Raghavan, Upfal'90]



$\mathcal{H}_3$

3. If $pn \geq (1 + \varepsilon) \ln n$ then Spread Time of $G(n, p) = \Theta(\log n)$ [Feige et al.'90]

Improving the protocol

Uninformed vertices ask the informed ones...

The push-pull protocol

Demers, Gealy, Greene, Hauser, Irish, Larson, Manning, Shenker, Sturgis, Swinehart, Terry, Woods'87

1. The ground is a simple connected graph.
2. At time 0, one vertex knows a rumour.
3. At each time-step $1, 2, \ldots,$
   every informed vertex sends the rumour to a random neighbour (PUSH);
   and every uninformed vertex queries a random neighbour about the rumour (PULL).

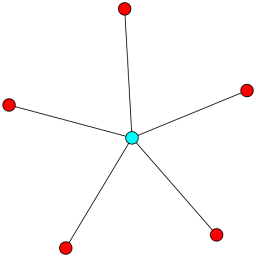
The push-pull protocol

Demers, Gealy, Greene, Hauser, Irish, Larson, Manning, Shenker, Sturgis, Swinehart, Terry, Woods'87

1. The ground is a simple connected graph.
2. At time 0, one vertex knows a rumour.
3. At each time-step $1, 2, \ldots,$
   every informed vertex sends the rumour to a random neighbour (PUSH);
   and every uninformed vertex queries a random neighbour about the rumour (PULL).

Remark 1. Vertices may call the same neighbour in consecutive steps.
Remark 2. If a vertex receives the rumour at time t, it starts passing it from time $t + 1$.
Spread Time: the first time everyone knows the rumour.

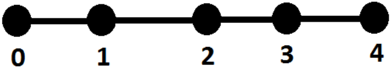
Example: a star



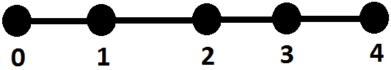
push protocol: $n \ln n$ rounds

push-pull protocol: 1 or 2 rounds

Example: a path

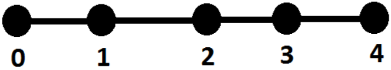


Example: a path



$$\text{inform} - \text{time}(0) = 0$$

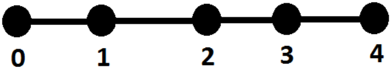
Example: a path



$$\text{inform} - \text{time}(0) = 0$$

$$\text{inform} - \text{time}(1) = 1$$

Example: a path

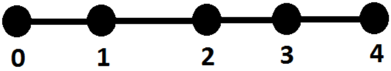


$$\mathrm{inform}-\mathrm{time}(0) = 0$$

$$\mathrm{inform}-\mathrm{time}(1) = 1$$

$$\begin{aligned}\mathrm{inform}-\mathrm{time}(2) &= 1+\min\{\mathrm{Geo}(1/2),\mathrm{Geo}(1/2)\}\\ &= 1+\mathrm{Geo}(3/4)\end{aligned}$$

Example: a path



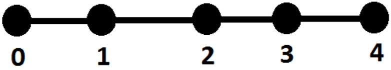
$$\mathrm{inform-time}(0) = 0$$

$$\mathrm{inform-time}(1) = 1$$

$$\mathrm{inform-time}(2) = 1 + \mathrm{Geo}(3/4)$$

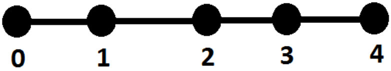
$$\mathrm{inform-time}(3) = 1 + \mathrm{Geo}(3/4) + \mathrm{Geo}(3/4)$$

Example: a path



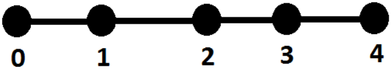
$$\mathrm{inform}-\mathrm{time}(0)=0$$
$$\mathrm{inform}-\mathrm{time}(1)=1$$
$$\mathrm{inform}-\mathrm{time}(2)=1+\mathrm{Geo}(3/4)$$
$$\mathrm{inform}-\mathrm{time}(3)=1+\mathrm{Geo}(3/4)+\mathrm{Geo}(3/4)$$
$$\mathrm{inform}-\mathrm{time}(4)=1+\mathrm{Geo}(3/4)+\mathrm{Geo}(3/4)+1$$

Example: a path



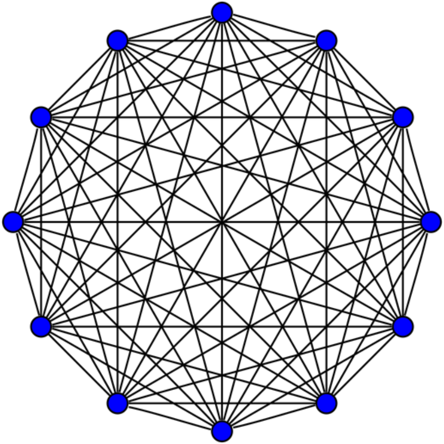
$$\mathrm{inform} - \mathrm{time}(0) = 0$$
$$\mathrm{inform} - \mathrm{time}(1) = 1$$
$$\mathrm{inform} - \mathrm{time}(2) = 1 + \mathrm{Geo}(3/4)$$
$$\mathrm{inform} - \mathrm{time}(3) = 1 + \mathrm{Geo}(3/4) + \mathrm{Geo}(3/4)$$
$$\mathrm{inform} - \mathrm{time}(4) = 1 + \mathrm{Geo}(3/4) + \mathrm{Geo}(3/4) + 1$$
$$\mathbb{E}[\text{Spread Time}] = 2 + 2 \times 4/3 = 14/3$$

Example: a path



$$
\begin{aligned}
\text{inform} - \text{time}(0) &= 0 \\
\text{inform} - \text{time}(1) &= 1 \\
\text{inform} - \text{time}(2) &= 1 + \text{Geo}(3/4) \\
\text{inform} - \text{time}(3) &= 1 + \text{Geo}(3/4) + \text{Geo}(3/4) \\
\text{inform} - \text{time}(4) &= 1 + \text{Geo}(3/4) + \text{Geo}(3/4) + 1 \\
\mathbb{E}[\text{Spread Time}] &= 2 + 2 \times 4/3 = 14/3 \\
&= \frac{4}{3}n - 2 \qquad (\text{versus } 2n - 3 \text{ for push})
\end{aligned}
$$

Example: a complete graph



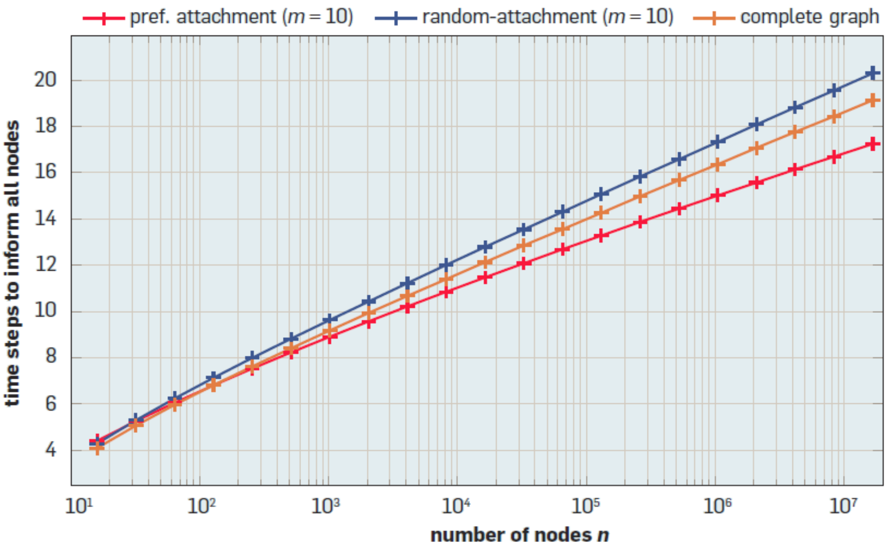
push: $\log_2 n + \ln n + o(\log n)$ [Pittel'87]

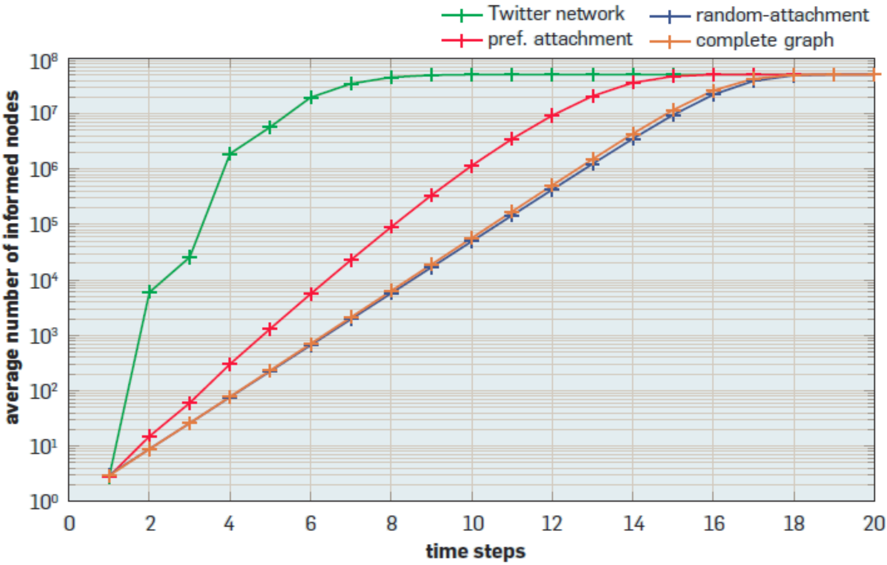
push-pull: $\log_3 n + o(\log n)$ [Karp, Schindelhauer, Shenker, Vöcking'00]

Other results on push-pull protocol

1. Barabasi-Albert preferential attachment graph has
   Spread Time $\Theta(\log n)$,
   PUSH alone has Spread Time $poly(n)$.
2. Random graphs with power-law expected degrees
   (a.k.a. the Chung-Lu model) with exponent $\in (2,3)$ has
   Spread Time $\Theta(\log n)$.
3. If Φ is Cheeger constant (conductance)
   and α is the vertex expansion (vertex isoperimetric number),
   Spread Time $\leq C \max\{\Phi^{-1} \log n, \alpha^{-1} \log^2 n\}$.

Let's see some simulation results...


pref. attachment ($m = 10$)
random-attachment ($m = 10$)
complete graph
time steps to inform all nodes
20
18
16
14
12
10
8
6
4
10^1
10^2
10^3
10^4
10^5
10^6
10^7
number of nodes n



$$n \approx 5 \times 10^7, \quad 64n \text{ edges}$$

Reference

Graphs in the previous slides were taken from:
Doerr, Fouz, Friedrich,
"Why rumors spread so quickly in social networks,"
Communications of the ACM, 2012

Some open problems

1. Design a (deterministic) approximation algorithm for finding the ‘average’ Spread Time of a given graph.

Some open problems

1. Design a (deterministic) approximation algorithm for finding the ‘average’ Spread Time of a given graph.
2. These questions may be asked about the ‘asynchronous’ model.

Some open problems

1. Design a (deterministic) approximation algorithm for finding the ‘average’ Spread Time of a given graph.
2. These questions may be asked about the ‘asynchronous’ model.

Part II: Push-Pull on Random k-trees

joint work with Ali Pourmiri

Random k-trees

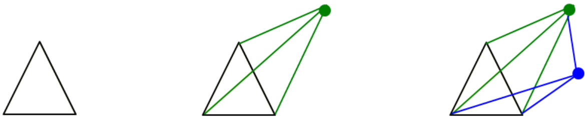
Example ($k = 3$)

Random k-trees

Example ($k = 3$)

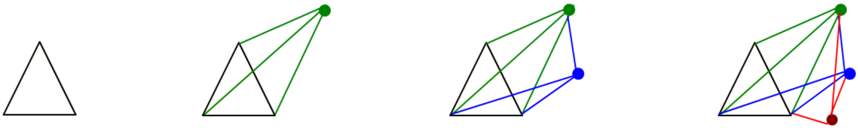
Random k-trees

Example ($k = 3$)

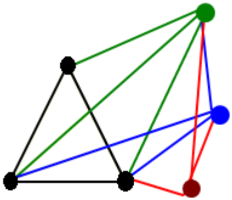


Random k-trees

Example ($k = 3$)



Some properties of random k-trees



- ✓ Logarithmic diameter
- ✓ Power law degree sequence:
  fraction of vertices with degree $d \approx d^{-2-\frac{1}{k-1}}$
- ✓ Constant tree-width k
- ✓ Clustering coefficient $\Omega(1)$
- ✓ Conductance and vertex expansion $o(1)$

Our results

Push-Pull protocol on random k-trees ($k > 1$ fixed):

Theorem (M, Pourmiri'14+)

If initially a random vertex knows the rumour,
a.a.s. after $\ln^{1+3/k} n$ *rounds,* $n - o(n)$ *vertices will know it.*

Our results

Push-Pull protocol on random k-trees ($k > 1$ fixed):

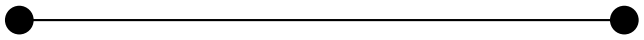
Theorem (M, Pourmiri'14+)

If initially a random vertex knows the rumour, a.a.s. after $\ln^{1+3/k} n$ *rounds,* $n - o(n)$ *vertices will know it.*

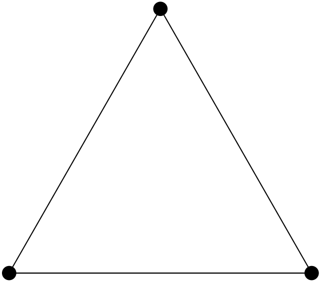
Theorem (M, Pourmiri'14+)

If initially a vertex knows the rumour, a.a.s. the average Spread Time is $> n^{1/3k}$

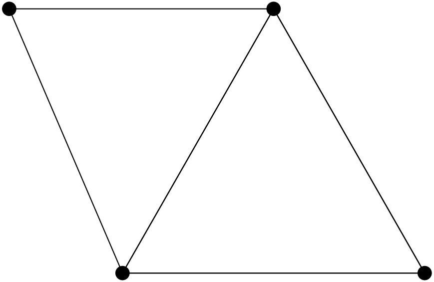
Self-similarity of random k-trees



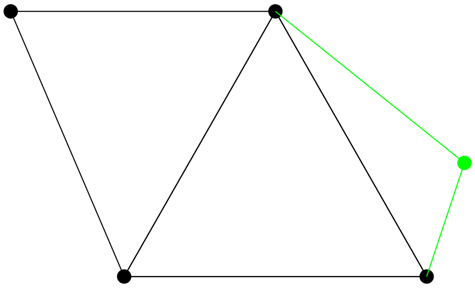
Self-similarity of random k-trees



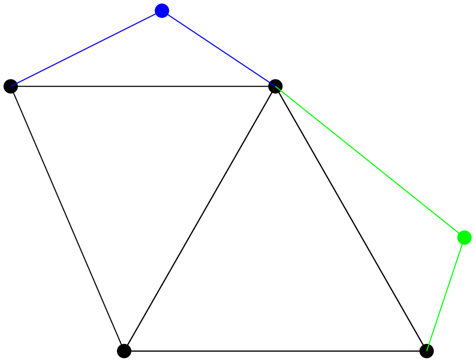
Self-similarity of random k-trees



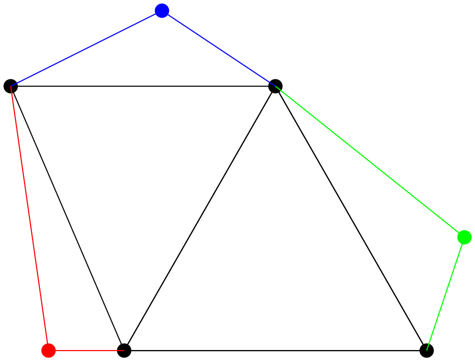
Self-similarity of random k-trees



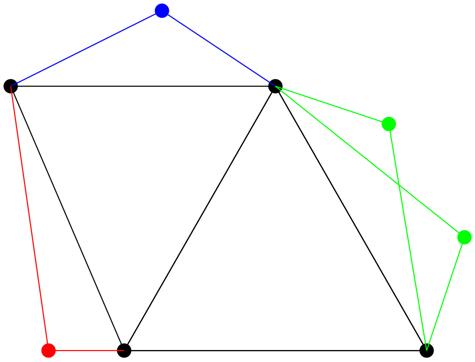
Self-similarity of random k-trees



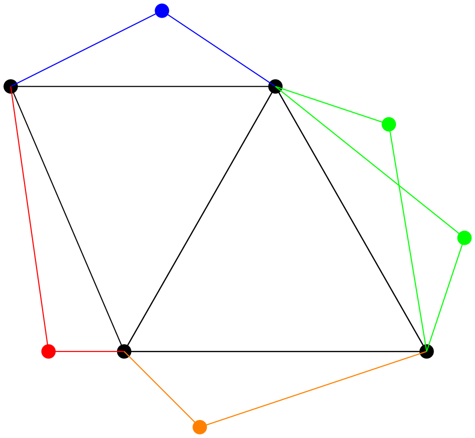
Self-similarity of random k-trees



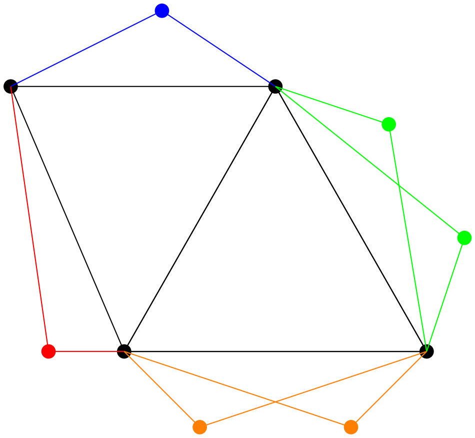
Self-similarity of random k-trees



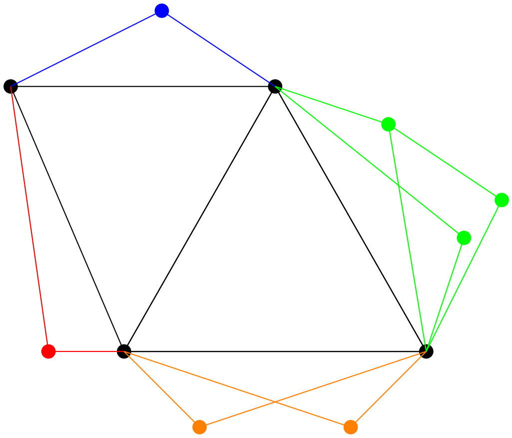
Self-similarity of random k-trees



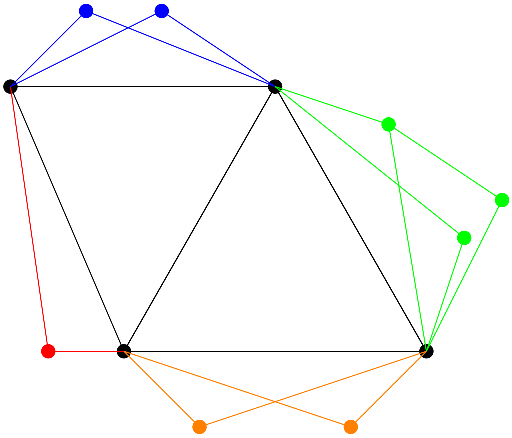
Self-similarity of random k-trees



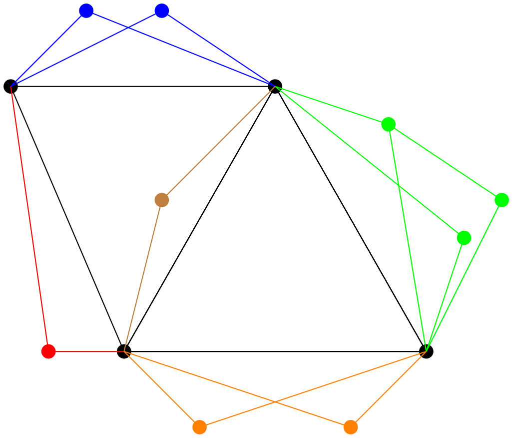
Self-similarity of random k-trees



Self-similarity of random k-trees



Self-similarity of random k-trees



Proof of upper bound

Lemma

Suppose s knows the rumour at time 0, and $\exists(s, v)$-path $s = u_0, u_1, \ldots, u_{l-1}, u_l = v$ s.t. $\min\{\deg(u_i), \deg(u_{i+1})\} \leq d$. Then with prob. $\geq 1 - o(n^{-2})$, inform-time$(v) \leq 6d(l + \ln n)$.

Proof of upper bound

Lemma

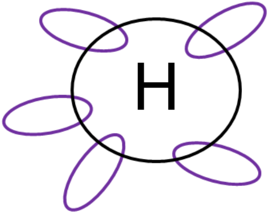
Suppose s knows the rumour at time 0, and $\exists(s,v)$-path $s = u_0, u_1, \ldots, u_{l-1}, u_l = v$ s.t. $\min\{\deg(u_i), \deg(u_{i+1})\} \leq d$. Then with prob. $\geq 1 - o(n^{-2})$, inform-time$(v) \leq 6d(l + \ln n)$.

Proof.

$$
\begin{aligned}
&\mathrm{inform-time}(v) \\
&\leq \min\{\mathrm{Geo}(\frac{1}{d_0}), \mathrm{Geo}(\frac{1}{d_1})\} + \cdots + \min\{\mathrm{Geo}(\frac{1}{d_{l-1}}), \mathrm{Geo}(\frac{1}{d_l})\} \\
&\leq \text{sum of } l \text{ independent } \mathrm{Geo}(1/d) \text{ random variables}
\end{aligned}
$$

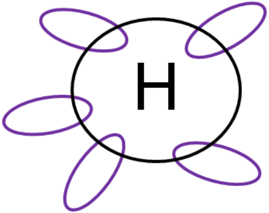
□

Proof of upper bound



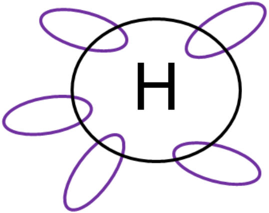
1. H = graph at round $\approx n \ln^{-2/k} n$

Proof of upper bound



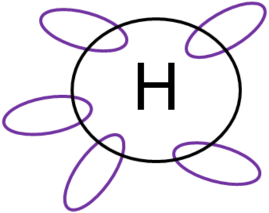
1. H = graph at round $\approx n \ln^{-2/k} n$
2. Almost all vertices in small pieces have degrees $\leq d = \ln^{3/k} n$

Proof of upper bound



1. $H =$ graph at round $\approx n \ln^{-2/k} n$
2. Almost all vertices in small pieces have degrees $\leq d = \ln^{3/k} n$
3. An edge $uv \in E(H)$ is fast if $\deg(u) \leq d$ or $\deg(v) \leq d$ or u and v have a common neighbour with degree $\leq d$.

Proof of upper bound



1. H = graph at round $\approx n \ln^{-2/k} n$
2. Almost all vertices in small pieces have degrees $\leq d = \ln^{3/k} n$
3. An edge $uv \in E(H)$ is fast if $\deg(u) \leq d$ or $\deg(v) \leq d$ or u and v have a common neighbour with degree $\leq d$.
4. $\exists$ an almost-spanning tree of H of height $O(\ln n)$ consisting of fast edges.

The upper bound

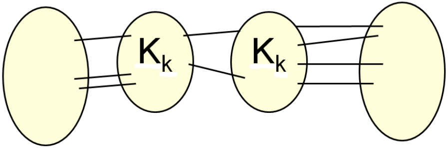
Push-Pull protocol on random k-trees ($k > 1$ fixed):

Theorem (M, Pourmiri'14+)

If initially a random vertex knows the rumour,
a.a.s. after $\ln^{1+3/k} n$ *rounds,* $n - o(n)$ *vertices will know it.*

Proof of lower bound

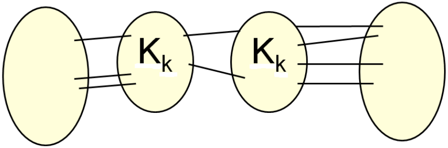
Definition (D-barrier)



Vertices in the two k-cliques have degrees $\geq D$.

Proof of lower bound

Definition (D-barrier)

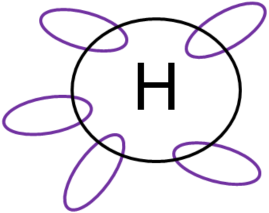


Vertices in the two k-cliques have degrees $\geq D$.

Lemma

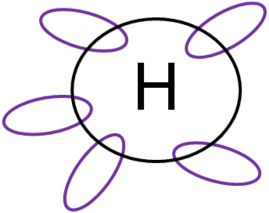
A random k-tree has a $\Omega\left(n^{1-1/k}\right)$-barrier with prob. $\geq \Omega\left(n^{-k}\right)$

Proof of lower bound



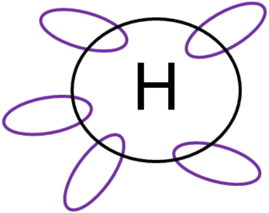
1. H = graph at round $m \approx n^{\frac{k}{k+1}}$

Proof of lower bound



1. H = graph at round $m \approx n^{\frac{k}{k+1}}$
2. Each small piece has a $(n/km)^{1-1/k}$-barrier with prob. $\Omega\left((n/km)^{-k}\right)$.

Proof of lower bound



1. H = graph at round $m \approx n^{\frac{k}{k+1}}$
2. Each small piece has a $(n/km)^{1-1/k}$-barrier with prob. $\Omega\left((n/km)^{-k}\right)$.
3. Since $km\left((n/km)^{-k}\right) \to \infty$ and by independence of pieces, with prob. $1 - o(1)$ there exists a $(n/km)^{1-1/k}$-barrier.

The lower bound

Push-Pull protocol on random k-trees ($k > 1$ fixed):

Theorem (M, Pourmiri'14+)

If initially a vertex knows the rumour,
a.a.s. the average Spread Time is $> n^{1/3k}$

Some open problems

1. Design a (deterministic) approximation algorithm for finding the 'average' Spread Time of a given graph.
2. These questions may be asked about the 'asynchronous' model.