# Adaptive Rumor Spreading

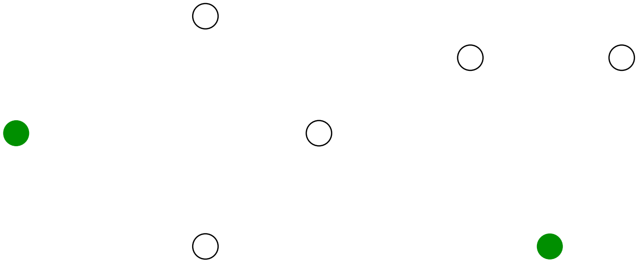
joint work with
J.R. Correa (U. Chile)
N. Olver (VU Amsterdam)
A. Vera (U. Chile / Cornell U.)

Marcos Kiwi

U. Chile

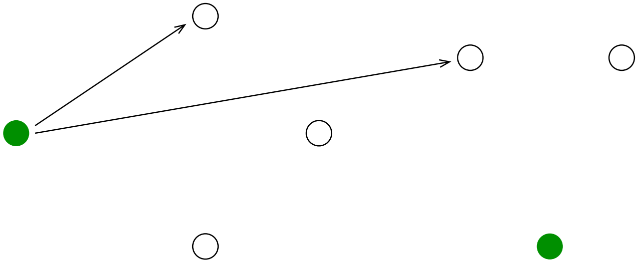
# Context

Inter-communicating agents, population, clients, devices, etc.



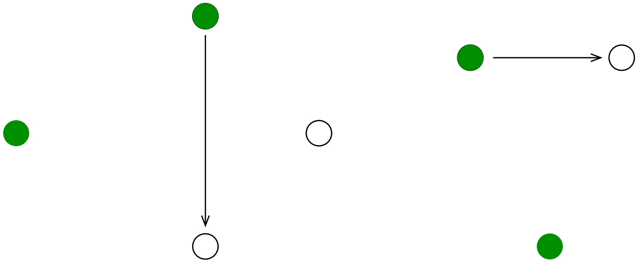
# Context

Inter-communicating agents, population, clients, devices, etc.



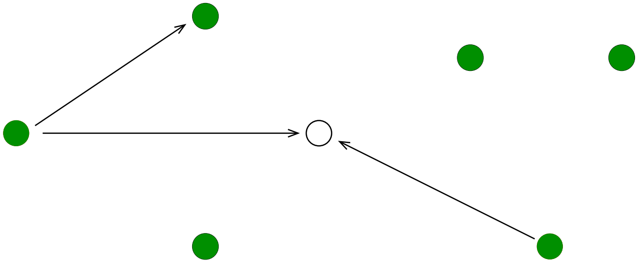
# Context

Inter-communicating agents, population, clients, devices, etc.



# Context

Inter-communicating agents, population, clients, devices, etc.



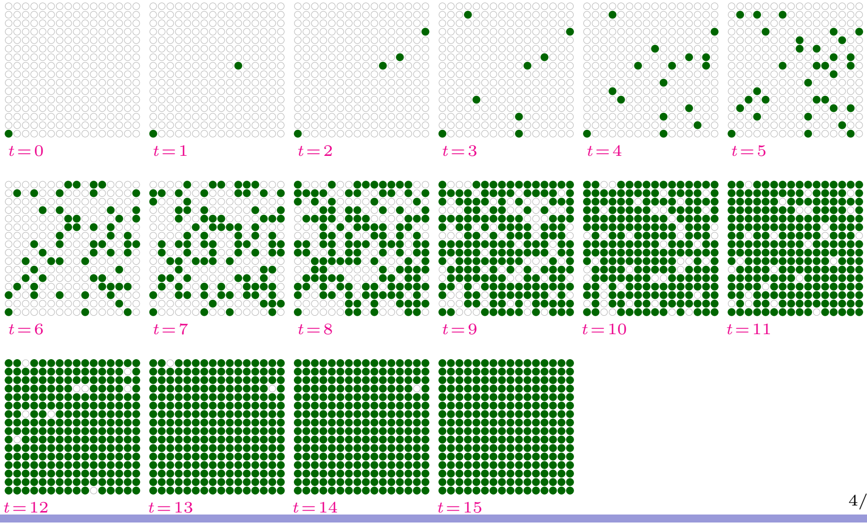
# Introduction

Studied as:

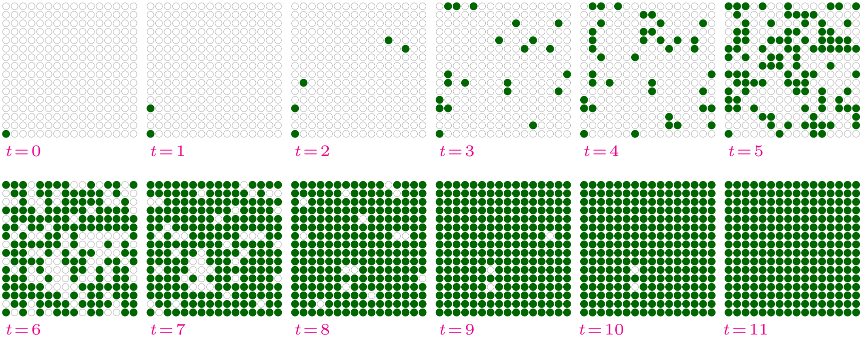
- Rumor spreading in social networks: content, updates, new technology ...
- Stochastic epidemic models.
- Viral marketing (where the selection of starting nodes is crucial). Domingos and Richardson (2001), ...
- Communication off-loading in opportunistic networks. Whitbeck at al. (2011), Sciancalepore et al. (2014)

**Note:** Models differ in way time evolves (a/synchronous) and communication protocol. Demers et al. (1987), Boyd et al. (2006), ...

# Synchronous discrete time rumor spreading in $K_n$

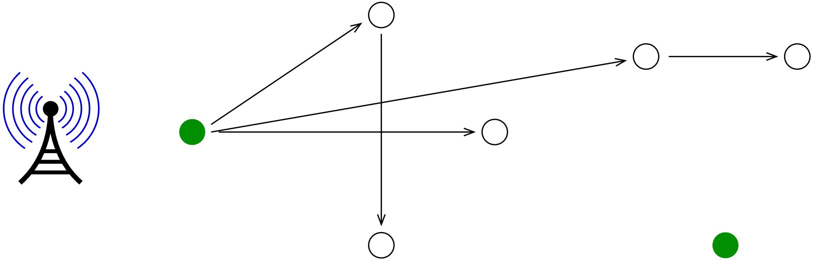


# Asynchronous continuous time rumor spreading in $K_n$



# Opportunistic networks

Proposal to address overload problem in data networks:
exploit opportunistic communications.



# Our focuss: Fix deadline scenario

**Motivation:**

- QoS restrictions in content delivery of news, traffic updates, etc.
- Marketing products such as flights, event tickets, etc.
- Distribution of alerts, say a tsunami warning.
- Sensing/monitoring environment within time windows.

# Our focuss: Fix deadline scenario

**Motivation:**

- QoS restrictions in content delivery of news, traffic updates, etc.
- Marketing products such as flights, event tickets, etc.
- Distribution of alerts, say a tsunami warning.
- Sensing/monitoring environment within time windows.

**Previous work:**

- Studied heuristically for real data. Whitbeck et al. (2011)
- Control theory based algorithms greatly outperform static ones. Sciancalepore et al. (2014)

Introduction
Motivation

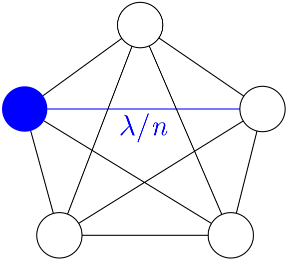
Model
Description

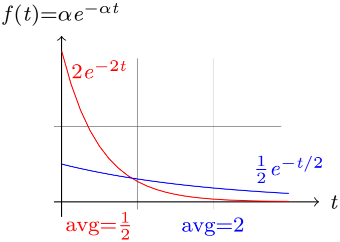
Analysis
Non-adaptive case
Adaptive case
Large deadline
Small deadline
Moderate deadline

Conclusion
Additional results
Further work

# Model

- $n$ nodes.
- Each pair of nodes can meet and gossip. If any one of them is active, then both become active.
- Continuous time.
- Pairs meet according to a Poisson process of rate $\lambda/n$, i.e., time between two consecutive meetings are exponentials of rate $\lambda/n$.



# The problem

- Pushes (external activations) have unit cost.
- Opportunistic communications (intra-network activations) have no cost.
- Nodes send an acknowledgement when they become active.
- At a given horizon/deadline $\tau$ all nodes must be active.

# The problem

- Pushes (external activations) have unit cost.
- Opportunistic communications (intra-network activations) have no cost.
- Nodes send an acknowledgement when they become active.
- At a given horizon/deadline $\tau$ all nodes must be active.

Two possible scenarios: Adaptive and non-adaptive.

# The problem

- Pushes (external activations) have unit cost.
- Opportunistic communications (intra-network activations) have no cost.
- Nodes send an acknowledgement when they become active.
- At a given horizon/deadline $\tau$ all nodes must be active.

Two possible scenarios: Adaptive and non-adaptive.

Natural question:
Compare best adaptive and non-adaptive strategies that minimize the expected overall number of pushes.

# Main result

## Theorem

*The adaptivity ratio, i.e., the ratio $\rho(\tau)$ between the expectated costs of the optimal non-adaptive vs the expected cost of the optimal adaptive strategies, is $O(1)$.*

In plain English:

No significant gain in being adaptive!

Introduction
Motivation

Model
Description

Analysis
Non-adaptive case
Adaptive case
Large deadline
Small deadline
Moderate deadline

Conclusion
Additional results
Further work

# Non-adaptive

When $i$ nodes are active, the time $T_i$ until $i+1$ nodes become active is the min of $i(n-i)$ i.i.d. $\mathrm{Exp}(\lambda/n)$, i.e.

$$T_i \sim \mathrm{Exp}(\lambda_i), \quad \text{for } \lambda_i := \frac{\lambda}{n} i(n-i).$$

**Recall:** $\mathbb{E}(T_i) = \sigma(T_i) = \frac{1}{\lambda_i}$.

# Non-adaptive

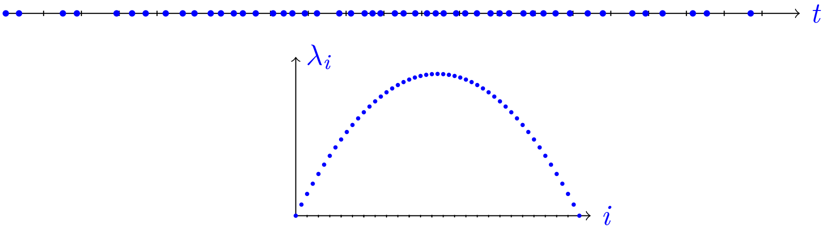
When $i$ nodes are active, the time $T_i$ until $i+1$ nodes become active is the min of $i(n-i)$ i.i.d. $\mathrm{Exp}(\lambda/n)$, i.e.

$$T_i \sim \mathrm{Exp}(\lambda_i), \quad \text{for } \lambda_i := \frac{\lambda}{n} i(n-i).$$

**Recall:** $\mathbb{E}(T_i) = \sigma(T_i) = \frac{1}{\lambda_i}$.

Henceforth $\lambda = 1$.

Example ($n = 50$, $\lambda = 1$)




Stochastic domination argument implies that non-adaptive does not perform more than $n/2$ pushes (therefore, neither adaptive).

# Optimal non-adaptive

- Say $K_N(t)$ is the number of active nodes at $t$ (right cont.)
- We are interested in determining

$$\textsc{Cost}_N := \min_{k=1,\ldots,n} (k + u_k(0))$$

where $u_k(t) := \mathbb{E}(n - K_N(\tau^-) | K_N(t) = k)$ is the expected number of inactive nodes just before the deadline given that at time $t$ there are $k$ active nodes.

# Some observations

- $u_k(t)$ is a convex function of $k$ (via submodularity arguments).
- $k + u_k(t)$ has a "unique" minima.
- $u_k(t)$ is essentially solution to the classical epidemic/diffusion model $\frac{dx}{dt} = -\lambda x(n-x)$, i.e., the "logistic function"

$$u_k(t) = \frac{n(1+o(1))}{1 + \frac{k}{n-k} e^{\tau - t}} + o(1).$$

- Optimal $k$ is $k_N = \frac{n(1+o(1))}{1 + e^{\tau/2}}$ and $\textsc{Cost}_N = 2k_N(1+o(1))$.

# Adaptive: Large deadline $(\tau \geq (2+\delta)\ln n)$

If at $t = 0$ there is just $1$ active node, then the time until every node is active is concentrated around $2\ln n + \mathcal{O}(1)$

Janson (1999)

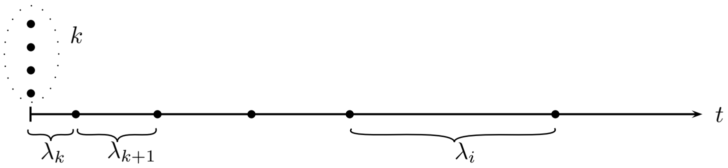
Thus, the adaptivity ratio is $1 + o(1)$.

# Adaptive: Small deadline $(\tau \leq 2 \ln \ln n)$

- Consider a Poisson process of unit rate in the line.

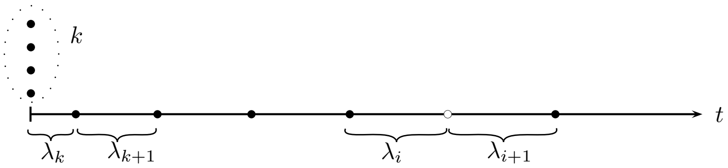
# Adaptive: Small deadline ($\tau \leq 2 \ln \ln n$)

- Consider a Poisson process of unit rate in the line.
- Given points $S_i, S_{i+1}$, a rescaling gives $T_i \sim \frac{S_{i+1} - S_i}{\lambda_i}$



# Adaptive: Small deadline ($\tau \leq 2 \ln \ln n$)

- Consider a Poisson process of unit rate in the line.
- Given points $S_i, S_{i+1}$, a rescaling gives $T_i \sim \frac{S_{i+1} - S_i}{\lambda_i}$
- A push can be thought of as adding a point.



# Adaptive: Small deadline (cont.)

### Definition (Clairvoyant strategy)

Chooses when to push knowing point process realization.

Clearly:

- Outperforms adaptive.
- Adds points only at the beginning.

# Adaptive: Small deadline (cont.)

## Definition (Clairvoyant strategy)

Chooses when to push knowing point process realization.

Clearly:

- Outperforms adaptive.
- Adds points only at the beginning.

## Claim

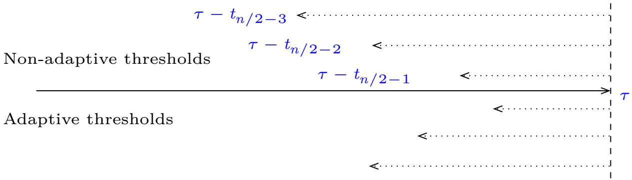
$$\textsc{Cost}_A \geq \textsc{Cost}_{CV} = \textsc{Cost}_N - \mathcal{O}(\sqrt{n}\ln^2 n)$$

and

$$\textsc{Cost}_N = \Omega(n/\ln n).$$

# Adaptive: Moderate deadlines $(2 \ln \ln n < \tau < 2(1+\delta) \ln n)$

Insight: optimal adaptive/non-adaptive are given by thresholds:



# Relaxed adaptive

A relaxed adaptive is an adaptive strategy that:

- Pushes for free.
- Pushes if there are less than $n/2$ active nodes or $\tau$ is reached.
- Pushes at $t$ only if $t \geq \tau - t_{K^*(t)}$ where $K^*(t)$ is the number of active nodes at time $t$ when following the optimal adaptive strategy.

# Relaxed adaptive

A relaxed adaptive is an adaptive strategy that:

- Pushes for free.
- Pushes if there are less than $n/2$ active nodes or $\tau$ is reached.
- Pushes at $t$ only if $t \geq \tau - t_{K^*(t)}$ where $K^*(t)$ is the number of active nodes at time $t$ when following the optimal adaptive strategy.

Then, associate a Doob sub-martingale, re-scale, transform, and show is dominated by a negative drift transformed Poisson process.

Introduction
Motivation

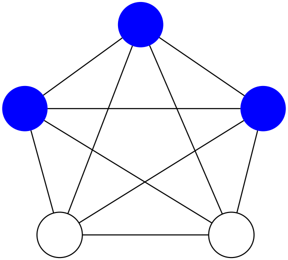
Model
Description

Analysis
Non-adaptive case
Adaptive case
Large deadline
Small deadline
Moderate deadline

Conclusion
Additional results
Further work

# Additional results

- Results hold also for target set version of the problem.



- Unbounded adaptivity ratio for inhomogenous networks.

# Some questions

- What is the exact adaptivity ratio? Could it be $2 + o(1)$?

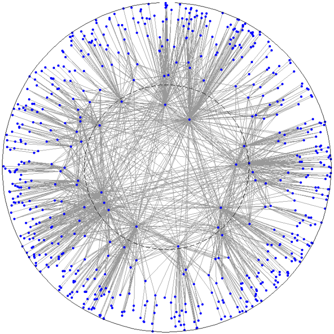
  There is strong numerical evidence that $2$ is the right constant factor.

- Is $\textsc{Cost}_N \leq \textsc{Cost}_A + \mathcal{O}(1)$?

- Is there a broader class of graphs maintaining the constant gap result? Maybe high conductance graphs?

# Some questions (cont.)

- What about the inhomogenous setting? Say
  - rates satisfy triangle inequality: $\lambda_{i,j}^{-1} \leq \lambda_{i,k}^{-1} + \lambda_{k,j}^{-1}$ for all $i, j, k$.
  - there is an underlying geometric space so $\lambda_{i,j}^{-1} \sim \frac{1}{(\mathrm{dist}(i,j))^{\beta}}$.
  - rates indicate whether or not $ij$ is an edge of a HRG.



How should fairness issues be dealt with?

The End!