# Speeding up non-Markovian First Passage Percolation with a single extra edge

Alexey Medvedev, CEU, IM SBRAS

joint work with Gabor Pete, BME, Renyi Institute of Mathematics

## SI spreading process (FPP)

Given $G = (V, E, s)$, $|V| = n$ – connected graph with root $s$;

- Attach i.i.d. random lengths $X_e$ to each edge $e \in E$;
- Each vertex $v \in V$ may be in one of the following two states:
  - *susceptible* (S);
  - *infected* (I).

**Initial setting:** $s$ is in state $I$, $V \backslash \{s\}$ are in state $S$;
**Process:** flow of speed 1 through each possible edge from *infected* vertices to *susceptible* ones.

The *SI spreading process* $T = (T_k)_{k=1}^n$ on the rooted graph $G$ is defined as the **minimum** time to infect $k$ vertices in the graph from the root $s$. By definition we let $T_1 = 0$.

## Mathematical description

Let $\mathcal{G}_k$ be the set of connected induced subtrees of $G$ on $k$ vertices with the same root $s$. Denote the *passage time* of the path $\gamma = (e_1, \dots, e_m)$ as

$$T(\gamma) = \sum_{i=1}^{m} X_{e_i},$$

and let $T(s,t)$ be the shortest passage time, or

$$T(s,t) = \inf\{T(\gamma) : \gamma \text{ is a path from } s \text{ to } t\}.$$

Then the infection time of $k$ vertices can be defined as

$$T_k = \min_{H_k \in \mathcal{G}_k} \max_{t \in V(H_k)} |T(s,t)|.$$

## Motivation

The main motivation of this research comes from a question, posed in the work of Kertész and Horváth. The authors considered a computer simulation of the non-Markovian SI spreading on various rooted trees on $n$ vertices with power-law edge lengths $X$ having tail distribution

$$\mathbb{P}(X > t) = (1 \wedge t^{-\alpha})$$

where $\alpha < 1$, or, in other words, having

$$\mathbb{E}(X) = \infty.$$

**Definition.** *Define a* **computer realisation** *of the* **SI** *process as a sampling the random variables on the edges.*

**Definition.** **Spreading curve** *of a computer realisation is the collection of points* $(T_k, k/n)$*, where* $1 \leqslant k \leqslant n$.

**Definition.** **Average spreading curve** *is the average of the collected spreading curves after* $M$ *realisations:*

$$(\langle T_k \rangle, k/n) = \left(\frac{\sum_{i=1}^{M} T_k^{(i)}}{M}, k/n\right).$$

The authors found the appearance of so-called " large plateaux" on average spreading curves of the computer simulation of the SI spreading process on trees, which do not vanish after increasing the number of realisations $M$. These plateaux are given by the following differences:

$$\frac{1}{M}\left(\langle T_{k+1} \rangle - \langle T_k \rangle\right),$$

for some $k < n$.

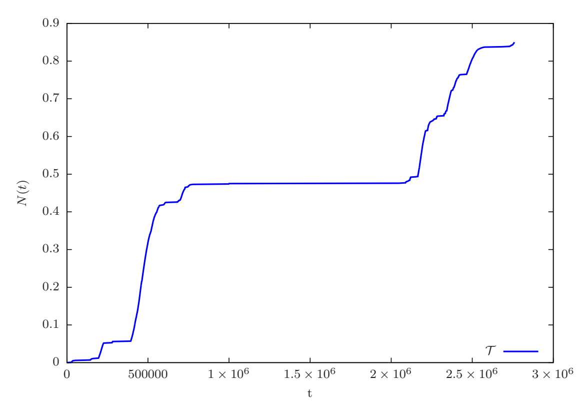


Figure: Example of the computer simulation of $M = 1000$ runs of the SI spreading on the critical Galton-Watson tree with $1000$ vertices and power-law edge lengths $X$ with $\alpha = 0.8$.

**D.X. Horváth, J. Kertész,** *Spreading dynamics on networks: the role of burstiness, topology and non-stationarity.* New Journal of Physics, 16(7):073037, 2014.

## Hypothesis

Kertesz and Horvath have made the following hypothesis, based on the calculation of the first infection time. Let the root $s$ have degree $d \geqslant 1$ and let $X_1, \dots, X_d$ be the random power-law lengths of incident edges. Then,

$$\mathbb{E}(T_2) = \mathbb{E}(\min\{X_1, \dots, X_d\}) < \infty \text{ when } \alpha > \frac{1}{d}.$$

**Hypothesis.** *When the infection has always* $d$ *ways to spread at each time in the process, then there are no large plateaux on the average spreading curve, when* $\alpha \in (1/d, 1)$.

## Current work

**Case $d = 2$:** consider the SI spreading on graphs with power-law edge lengths with $\alpha \in (1/2, 1)$.

**Question:** describe the appearance of the first large plateau on the average spreading curve, or the fraction of vertices that is infected in finite expected time with probability 1. By the Law of Large Numbers it is exactly the $k$, for which the $\mathbb{E}(T_k) < \infty$ and $\mathbb{E}(T_{k+1}) = \infty$.

## Spreading on a cycle

Same computer simulation of the SI process on a cycle $C_n$ with same number of vertices shows no large plateaux after averaging over $M = 1000$ runs:

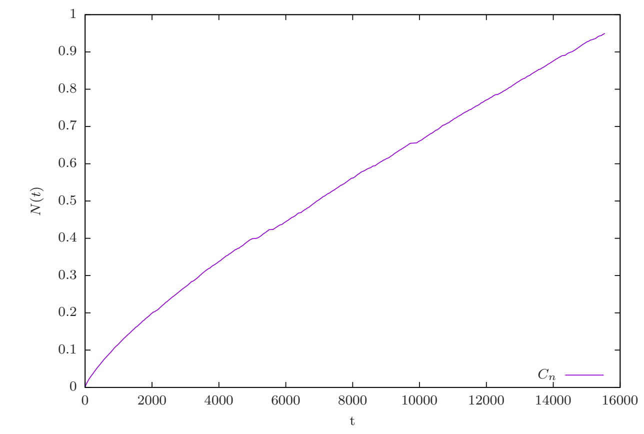


Figure: Example of the same computer simulation of the SI spreading on the cycle $C_n$ with $n = 1000$ vertices.

**Graph limit of the rooted cycle:** doubly infinite line $C_\infty = (V, E, 0)$, where

- $V = \{0, \pm1, \pm2, \dots\}$, with a root at $0$;
- for each $i, j \in V$, the pair $(i, j) \in E$ iff $|i - j| = 1$.

Edge lengths $X_i$:

- $i$ is a label of the greater vertex, if one of the ends is positive;
- $i$ is the label of the smaller otherwise.

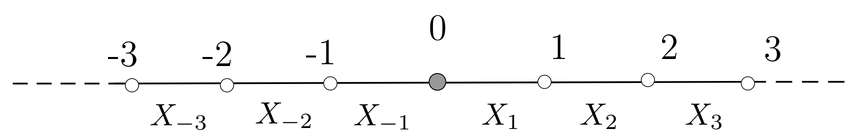



Figure: Graph of a doubly infinite line $C_\infty$.

**Theorem.** *Consider the SI spreading process* $(T_k)_{k=1}^{\infty}$ *on a graph of a doubly infinite line* $C_\infty$ *with root* $0$ *and power-law edge lengths with* $\alpha \in (1/2, 1)$*. Then the expected time to infect* $k \geqslant 1$ *vertices is bounded:*

$$\mathbb{E}(T_k) \asymp k^{1/\alpha}.$$

## Spreading on a star

Consider the graph of a $n$-star $ST_n = (V, E, 0)$, where

- $V = \{0, 1, 2, \dots, n\}$, with a root at $0$;
- for each $i \geqslant 1$, the pair $(0, i) \in E$.

Let $T = (T_k)_{k=1}^{\infty}$ be the SI spreading process on $ST(n)$ with power-law distributed random weights with $\alpha \in (1/2, 1)$, denoted as $X_i$, where $i$ is a label of the greater vertex.

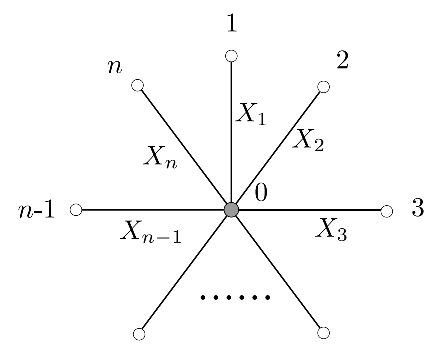



Figure: Graph of a $n$-star $ST_n$.

**Theorem.** *Consider the SI spreading process* $(T_k)_{k=0}^{n}$ *on a graph of a star* $ST_n$ *with root* $0$ *and power-law edge lengths with* $\alpha \in (1/2, 1)$*. Then the expected time to infect* $k$ *vertices is bounded up until* $k \leqslant n - 1$*, and for* $n \geqslant 2$*,*

$$\mathbb{E}(T_k) \leqslant Ck^{1/\alpha}.$$

## General static graphs

Consider the SI process $(T_k)_{k=1}^{n}$ on a finite connected rooted graph $G = (V, E, s)$ with root $s$ and power-law edge lengths with $\alpha \in (1/2, 1)$. For any rooted $G$ there exists the specific number $\kappa(G, s) \in \{1, \dots, n\}$, such that any $k < \kappa(G, s)$ vertices are infected in the finite expected time. This number has the following combinatorial description.

**Lemma.** *Let* $|\mathcal{C}(s, G \backslash e)|$ *be the size of the connected component of vertex* $s$ *in the graph* $G$ *without edge* $e$*. Then,*

$$\kappa(G, s) = \min_{e \in E(G)} |\mathcal{C}(s, G \backslash e)|.$$

The following theorem provides a bound on the expected time of infection.

**Theorem.** *Consider the SI process* $(T_k)_{k=1}^{n}$ *on a connected rooted graph* $G = (V, E, s)$ *on* $n$ *vertices with root* $s$ *and power-law edge lengths with* $\alpha \in (1/2, 1)$*. Then for each* $k$*, where* $1 \leqslant k < \kappa(G, s)$*, the expected time to infect* $k$ *vertices is bounded by*

$$\mathbb{E}(T_k) \leqslant Ck^{1/\alpha},$$

*where* $C > 0$ *is a constant, that depends on* $\alpha$.

Simple corollary provides the description of graphs, which are infected in finite expected time.

**Lemma.** *Let* $G$ *be a finite* **2-edge-connected** *graph on* $n$ *vertices with power-law edge lengths with* $\alpha \in (1/2, 1)$*. Then for any root* $s$ *the SI process* $(T_k)_{k=1}^{n}$ *on* $(G, s)$ *infects any* $k \leqslant n$ *vertices in finite bounded time:*

$$\mathbb{E}(T_k) \leqslant Ck^{1/\alpha},$$

*where* $C > 0$ *is a constant, that depends on* $\alpha$.


e-mail: an.medvedev@yahoo.com

# Speeding up non-Markovian First Passage Percolation with a single extra edge

Alexey Medvedev, CEU, IM SBRAS

joint work with Gabor Pete, BME, Renyi Institute of Mathematics

## Critical Galton-Watson trees

Consider the critical Galton-Watson (CGW) tree $\mathcal{T}$ derived from the CGW process $(Z_k)_{k\geqslant 0}$, where $Z_0 = 1$, with a root $s$ at $Z_0$ and with *integer* offspring distribution $\xi$, such that $Var(\xi) = \sigma^2 < \infty$. Denote as $\mathcal{T}_{+e}$ the CGW tree $\mathcal{T}$ with an *extra edge* attached to the root and a randomly chosen vertex in the tree and consider a computer simulation of the SI spreading process with power-law edge lengths with $\alpha \in (1/2, 1)$. There are no plateaux up to a certain fraction of vertices, which is larger than in the case of no extra edge.

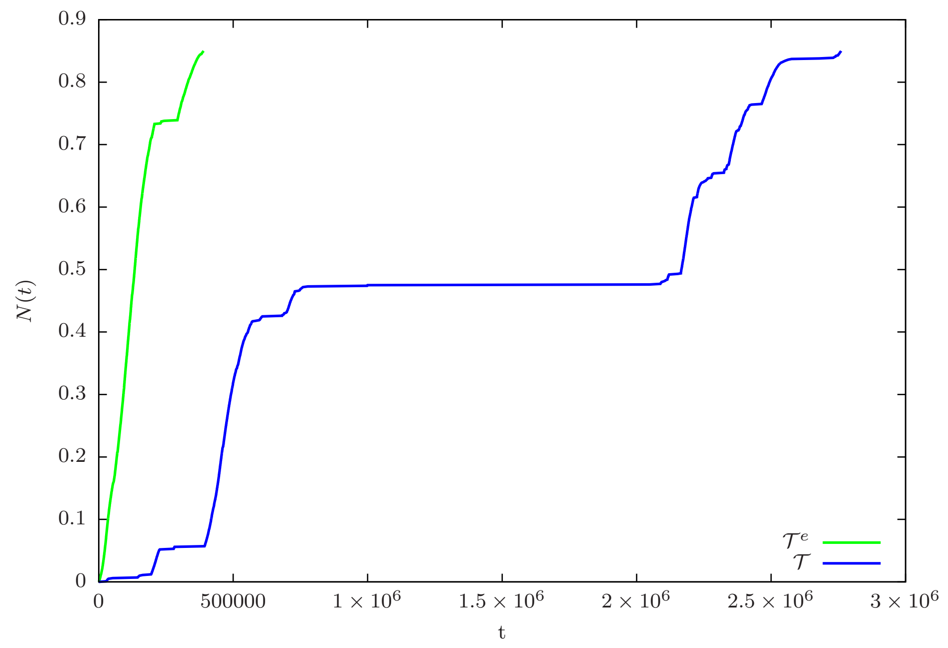



Figure: Example of a computer simulation of $M = 1000$ realisations of the SI spreading on the trees $\mathcal{T}$ and $\mathcal{T}_{+e}$ with $1000$ vertices and power-law edge lengths with $\alpha = 0.8$.

The tree $\mathcal{T}_{+e}$ has the following structure:

- a cycle $C_n$ of length $n$;
- $M$ rooted trees $t_1, t_2, \dots, t_M$ attached to it by edges $e_1, e_2, \dots, e_M$, where $M \geqslant 1$.

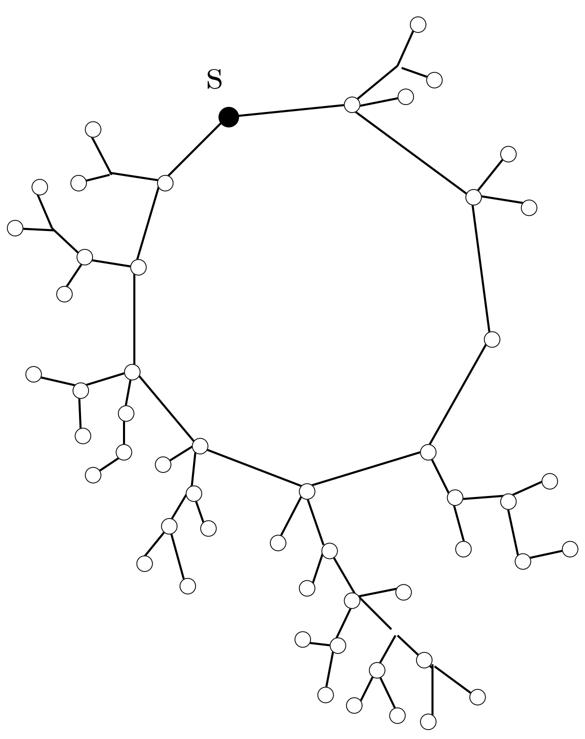



Figure: Schematic structure of a CGW tree with extra edge.

Based on the combinatorial description of $\kappa$ it is easy to find that w.h.p.:

$$\kappa(\mathcal{T}, s) = |\mathcal{T}| - \max_{i=1,\dots,\xi} |t_i|,$$

where $\xi$ is the degree of a root, and

$$\kappa(\mathcal{T}_{+e}, s) = |\mathcal{T}_{+e}| - \max_{i=1,\dots,M} |t_i|.$$

The following theorem provides a bound on the expected time of infection.

**Theorem.** *Consider the SI process $(T_k)_{k=1}^n$ on the conditioned CGW tree $\mathcal{T}^N = (\mathcal{T} | Z_N > 0)$ with root $s$ at $Z_0$ and integer offspring distribution $\xi$, such that $Var(\xi) < \infty$, with power-law edge lengths with $\alpha \in (1/2, 1)$. Then,*

- *as $N \to \infty$, the sequence of r.v. $\kappa(\mathcal{T}^N, s)$ is tight.*

*Let $\mathcal{T}^N_{+e}$ denote the tree $\mathcal{T}^N$ with an* **extra edge** *attached to $s$ and a uniformly chosen vertex. Then,*

- *as $N \to \infty$, for any $\varepsilon > 0$ there exists $\delta > 0$, such that*

$$\mathbb{P}\Big(\frac{\kappa(\mathcal{T}^N_{+e}, s)}{|\mathcal{T}^N_{+e}|} > \delta\Big) > 1 - \varepsilon.$$

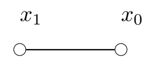
## Acknowledgements

The research has been supported by the European Research Council grant "Limits of discrete structures", 617747, and the Grant 16-01-00499 of the Russian Foundation of Basic Research.

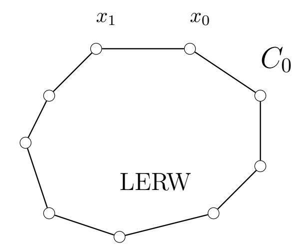
## Uniform spanning tree (UST)

A uniform spanning tree $\mathcal{T}_n^{+e}$ of a **complete graph** $K_n$ with an extra edge can be generated using the *Wilson's algorithm*:

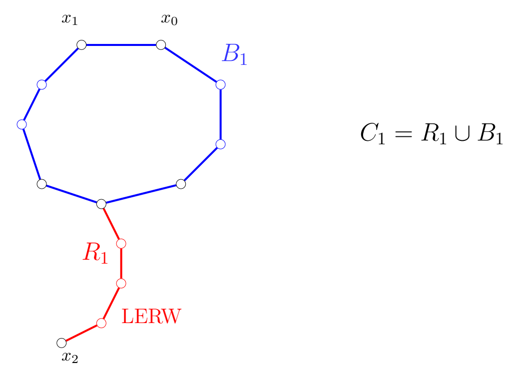
- Take the vertices $x_0$ and $x_1$, and add the edge $(x_0, x_1)$.



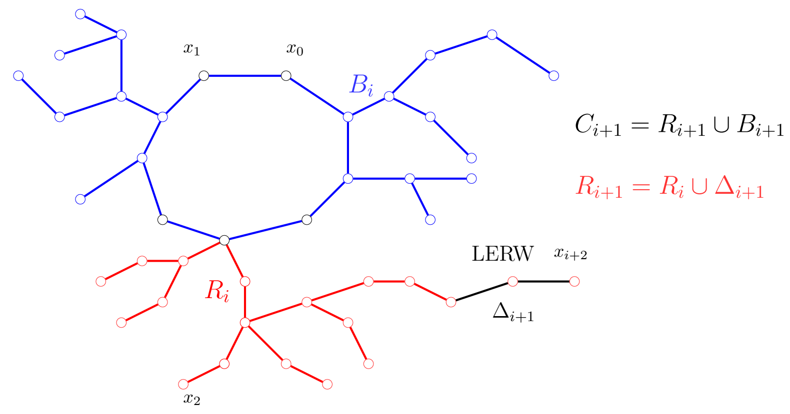

- Run a Loop-Erased Random Walk (LERW) from $x_1$ to $x_0$; the union of this walk with the existing edge will be denoted by $C_0$.




- Run a LERW from $x_i$ until hitting $C_0$. This walk, including its endpoint, will be denoted by $R_1$, while $C_0$ minus the endpoint of $R_1$ will be $B_1$. We also set $C_1 = R_1 \cup B_1$.




- Now, recursively, given $C_i = R_i \cup B_i$ for some $i \geq 1$, run a LERW from $x_{i+2}$ until hitting $C_i$, denoting the walk, excluding its endpoint, by $\boldsymbol{\Delta}_{i+1}$. If the endpoint is in $R_i$, then we let $R_{i+1} = R_i \cup \boldsymbol{\Delta}_{i+1}$; otherwise, we let $B_{i+1} = B_i \cup \boldsymbol{\Delta}_{i+1}$. We will sometimes refer to $C_i$ as the set of colored vertices, red in $R_i$, blue in $B_i$.




- We iterate this procedure until we reach $C_l = \{x_0, x_1, \dots, x_{n-1}\}$. In retrospect, we also set $\boldsymbol{\Delta}_0 = C_0 \setminus \{x_0\}$ and $\boldsymbol{\Delta}_1 = C_1 \setminus C_0$.

The process $\big(|R_i|, |B_i|\big)_{i\geqslant 2}$ is a Pólya urn with random, symmetric, time-inhomogeneous increments: at each step $i \geqslant 2$ we have

$$\mathbb{P}(|\boldsymbol{\Delta}_{i+1}| \text{ attach to red}) = \frac{|R_i|}{|R_i| + |B_i|},$$

$$\mathbb{P}(|\boldsymbol{\Delta}_{i+1}| \text{ attach to blue}) = \frac{|B_i|}{|R_i| + |B_i|}.$$

A straightforward formula for the $|\boldsymbol{\Delta}_{i+1}|$ is given

$$\mathbb{P}(|\boldsymbol{\Delta}_{i+1}| = k \mid C_i) = \frac{|C_i| + k}{n} \prod_{j=1}^{k-1} \Big(1 - \frac{|C_i| + j}{n}\Big).$$

In particular, for $i = 0$, the distribution of $|\boldsymbol{\Delta}_0|/\sqrt{n} = |B_0|/\sqrt{n}$ has a limit: the Rayleigh distribution $\rho_0$, given by

$$\mathbb{P}(\rho_0 > t) = \exp(-t^2/2)\,, \qquad t \geq 0\,.$$

Also, the distribution of $|\boldsymbol{\Delta}_1|/\sqrt{n} = |R_0|/\sqrt{n}$, conditioned on $|C_0| = |\boldsymbol{\Delta}_0| \sim \sqrt{n}\rho_0$, has a limit, denoted by $\rho_1$:

$$\mathbb{P}(\rho_1 > t \mid \rho_0) = \exp(-(t + \rho_0)^2/2)\,\exp(\rho_0^2/2)\,, \qquad t \geq 0\,.$$

Thus, $\big(|R_0|, |B_0|\big)/\sqrt{n}$ has an absolutely continuous limit law on $[0, \infty)^2$, and $|R_0|/|C_0|$ has an absolutely continuous limit law on $[0, 1]$.

**Theorem.** *The final partition $(R_l, B_l)$ of the set of $n$ vertices, as $n \to \infty$, is asymptotically almost surely non-trivial: for every $\varepsilon > 0$ there exists $\delta > 0$ such that*

$$\mathbb{P}(|R_l|/|C_l| \in (\delta, 1 - \delta)) > 1 - \varepsilon,$$

*for all $n > n_\varepsilon$.*

It is a simple corollary that since limit law a.s. has no atoms at 0 or 1, then for every $\varepsilon > 0$ there exists $\delta > 0$, such that

$$\mathbb{P}(\kappa(\mathcal{T}_n^{+e}, x_0) > \delta n) > 1 - \varepsilon.$$



e-mail: an.medvedev@yahoo.com