Inference of a Rumor's Source in the Independent Cascade Model (Supplementary Material)

A OMITTED PROOFS

A.1 PROOF OF THEOREM 1

Observe that by definition, we have for any $v, w \in V$ that $\mathbb{P}(\boldsymbol{\omega} = v) = \mathbb{P}(\boldsymbol{\omega} = w)$. Thus, by Bayes' rule and the law of total probability we get

$$\begin{aligned}\mathbb{P}(\boldsymbol{\omega} = v \mid \boldsymbol{X}^\star = X) &= \frac{\mathbb{P}(\boldsymbol{X}^\star = X \mid \boldsymbol{\omega} = v)\mathbb{P}(\boldsymbol{\omega} = v)}{\mathbb{P}(\boldsymbol{X}^\star = X)} \\ &= \frac{\mathbb{P}(\boldsymbol{X}^\star = X \mid \boldsymbol{\omega} = v)\mathbb{P}(\boldsymbol{\omega} = v)}{\sum\limits_{\omega \in V} \mathbb{P}(\boldsymbol{X}^\star = X \mid \boldsymbol{\omega} = \omega)\mathbb{P}(\boldsymbol{\omega} = w)} = \frac{\mathbb{P}(\boldsymbol{X}^\star = X \mid \boldsymbol{\omega} = v)}{\sum\limits_{\omega \in V} \mathbb{P}(\boldsymbol{X}^\star = X \mid \boldsymbol{\omega} = \omega)}.\end{aligned}$$

As $\sum_{\omega \in V} \mathbb{P}(\boldsymbol{X}^\star = X \mid \boldsymbol{\omega} = \omega)$ is independent from v, we have

$$\arg\max_{v \in V} \mathbb{P}(\boldsymbol{\omega} = v \mid \boldsymbol{X}^\star = X) = \arg\max_{v \in V} \mathbb{P}(\boldsymbol{X}^\star = X \mid \boldsymbol{\omega} = v)$$

and the theorem follows. □

A.2 PROOF OF PROPOSITION 8

Proof. The recurrence $\bar{x}_t = \exp(-\lambda p(1 - \bar{x}_{t-1}))$ can be easily calculated by the probability generating function of the Poisson distribution. Indeed, let $f_{\mathrm{Po}(\lambda)}(s) = \mathbb{E}\big[s^{\mathrm{Po}(\lambda)}\big]$ be the probability generating function of the Poisson distribution. It is well known that

$$f_{\mathrm{Po}(\lambda)}(s) = \exp(-\lambda(1 - s)).$$

We refer to [Grimmett and Stirzaker, 2020] for a detailed explanation of the connection between the probability generating function and the extinction probability of branching processes.

Now, for brevity, suppose that $v = \boldsymbol{\omega}_c$. Let $\mathcal{V}_0$ be the event that v has exactly $k \leq \boldsymbol{d}_0 \leq d$ children that get activated by v. Similarly as before, $\mathbb{P}(\mathcal{V}_0) = \mathbb{P}(\mathrm{Po}(\lambda p) = \boldsymbol{d}_0)$ and of course, $\boldsymbol{d}_0$ needs to be at least k as differently, the probability of having k active sub-trees was zero.

Given $\mathcal{V}_0$, we again start $\boldsymbol{d}_0$ independent Galton-Watson processes with offspring distribution $\mathrm{Po}(\lambda p)$ in the children. Therefore, the probability of observing exactly k active sub-trees is the probability that exactly k out of $\boldsymbol{d}_0$ of those processes are not extinct after $t_v^{\boldsymbol{X}^\star}$ steps. Of course, the number of such active sub-trees at time t is distributed as $\mathrm{Bin}(\boldsymbol{d}_0, \bar{x}_t)$ given $\mathcal{V}_0$ and the first part of the formula follows.

As in the d-regular case, if on contrary v is not the closest candidate but a node further apart from $\boldsymbol{X}^\star$, we observe that from the originally $1 \leq \boldsymbol{d}_0 \leq d$ Galton-Watson processes originated in the children of v, exactly one process needed to survive and $\boldsymbol{d}_0 - 1$ needed to be extinct at time $t_v^{\boldsymbol{X}^\star}$. □

A.3 PROOF OF THEOREM 3 (I)

Proof of Theorem 3 (i). As in the d-regular case, the first part of Theorem 3 follows by the first part of Proposition 8. If $\lambda p \leq 1$, the smallest fixed-point of $\bar{x} \mapsto \exp -\lambda p(1-\bar{x})$ is $\bar{x} = 1$. Therefore, $\bar{x}_t = 1 - o_t(1)$ describes the probability that the underlying spreading process died out until time-step t. More precisely, by the recurrence equation, we find the following. Suppose that $\varepsilon_t = o_t(1)$ denotes the convergence speed towards 1. Then, by the recurrence equation and a Taylor approximation we have

$$1 - \varepsilon_t = 1 - \lambda p\left(\varepsilon_{t-1} + \frac{\lambda^2 p^2 \varepsilon_{t-1}^2}{2}\right) + O\left(\varepsilon_{t-1}^3\right).$$

If $\lambda p < 1$, we directly find that $\varepsilon_t = O((\lambda p)^t)$ decays exponentially fast in t. If $\lambda p = 1$, this is much more subtle. Indeed, we find

$$\varepsilon_t \leq \left(\varepsilon_{t-1} - \frac{\varepsilon_{t-1}^2}{2}\right) + O\left(\varepsilon_{t-1}^3\right)$$

and therefore, we only get $\varepsilon_t = O\left(t^{-1}\right)$ in this case.

Since we assume p to be a constant, clearly $\lambda = O(1)$ as well. Unfortunately, the Poisson tails are kind of heavy. More precisely, even if λ is a constant, the probability that a $\mathrm{Po}(\lambda)$ variable becomes large is not negligible. We analyze this by a careful application of limits. Recall that we assume that the underlying tree-network is infinite. We model this as follows. Suppose that the tree-network consists of n vertices and we will let $n \to \infty$.

Let $C > 0$, then the probability that the number of neighbors of a specific node v exceeds C is, for large C, given by Chernoff bounds as

$$\mathbb{P}(|\partial v| > C) \leq \exp\left(-\frac{(C-\lambda)^2}{2C}\right) \sim \exp(-C/2).$$

As the number of spawned children is independent for all vertices, the number of vertices of degree at least C is stochastically dominated by $\mathrm{Bin}(n, \exp(-C/2))$. Thus, with probability $1 - o_n(1)$, there are no more than $O(n\sqrt{\ln(n)}\exp(-D))$ vertices of degree $D > 0$ for a sufficiently large constant D (independent of n) if $n \to \infty$.

We denote by $\mathcal{D}$ the event that this is actually true. Thus, conditioned on $\mathcal{D}$, there are only $O(n\sqrt{\ln(n)}\exp(-D))$ vertices of degree at least D. Now, we chose $\boldsymbol{\omega}$ uniformly at random out of all vertices. Therefore, given $\mathcal{D}$, the probability that $\boldsymbol{\omega}$ has small degree is

$$\mathbb{P}(|\partial\boldsymbol{\omega}| > D \mid \mathcal{D}) = 1 - O\left(\frac{\sqrt{\ln(n)}}{\exp(-D)}\right).$$

Clearly, this becomes only a high probability event if $D = \Omega(\ln\ln n)$. In the worst case, we find that a union bound over all activated children of $\boldsymbol{\omega}$ leads only to ultimate extinction of all processes, if $O\left(\frac{\ln\ln n}{t}\right) = o_t(1)$, or, differently, that $t = \omega(\ln\ln n)$. As in the theorem, we only claim the assertion in the limit $t \to \infty$ and we assume the underlying tree-network to be infinite. This proves the claim of the theorem. We remark at this point that the assumption that t depends slightly on n does no harm to applications as, on real networks, $\ln\ln n$ can be seen as a constant. □

B SIMULATION DATA

References

Geoffrey Grimmett and David Stirzaker. *Probability and Random Processes*. Oxford University Press, New York, 4 edition, 2020. ISBN 978-0-198-84760-1.

Table 1: Simulation results for random geometric graphs.

p	number of successes	$\boldsymbol{\omega}_c \neq \boldsymbol{\omega}$	$\boldsymbol{X}^\star = \emptyset$	average distance	maximum distance
0.00	0	0	100		0
0.05	1	1	98	2.25	4
0.10	5	29	66	2.39	5
0.15	25	59	16	2.02	6
0.20	34	63	3	1.68	5
0.25	51	48	1	1.51	4
0.30	62	36	2	1.40	5
0.35	71	29	0	1.24	5
0.40	86	14	0	1.11	4
0.45	94	6	0	1.04	3
0.50	94	6	0	1.13	5
0.55	95	5	0	1.07	5
0.60	100	0	0	0.97	4
0.65	95	5	0	1.03	6
0.70	99	1	0	0.79	3
0.75	99	1	0	1.03	5
0.80	100	0	0	0.97	5
0.85	100	0	0	0.96	6
0.90	100	0	0	0.66	3
0.95	98	2	0	0.87	5
1.00	100	0	0	0.85	6

Table 2: Simulation results for Erdős-Rényi graphs.

p	number of successes	$\boldsymbol{\omega}_c \neq \boldsymbol{\omega}$	$\boldsymbol{X}^\star = \emptyset$	average distance	maximum distance
0.00	0	0	100	-	-
0.05	0	0	100	-	-
0.10	0	0	100	-	-
0.15	0	1	99	6.00	6
0.20	0	6	94	7.50	9
0.25	0	14	86	6.63	8
0.30	2	30	68	7.34	10
0.35	11	35	54	7.23	10
0.40	21	49	30	5.87	9
0.45	33	43	24	6.14	9
0.50	42	33	25	1.15	8
0.55	54	31	15	0.54	3
0.60	63	24	13	0.36	3
0.65	74	19	7	0.30	2
0.70	78	17	5	0.21	2
0.75	78	15	7	0.17	2
0.80	82	12	6	0.18	3
0.85	81	15	4	0.17	2
0.90	86	13	1	0.17	2
0.95	85	10	5	0.10	1
1.00	92	5	3	0.05	1

Table 3: Simulation results for random regular graphs (configuration model).

p	number of successes	$\boldsymbol{\omega}_c \neq \boldsymbol{\omega}$	$\boldsymbol{X}^\star = \emptyset$	average distance	maximum distance
0.00	0	0	100	-	-
0.05	0	0	100	-	-
0.10	0	0	100	-	-
0.15	0	0	100	-	-
0.20	0	0	100	-	-
0.25	0	5	95	7.20	9
0.30	0	16	84	7.53	11
0.35	2	26	72	6.86	12
0.40	19	43	38	5.37	11
0.45	38	40	22	2.70	11
0.50	43	41	16	2.06	9
0.55	70	23	7	0.57	6
0.60	76	21	3	0.34	4
0.65	86	11	3	0.15	3
0.70	87	10	3	0.14	3
0.75	98	2	0	0.02	1
0.80	97	3	0	0.03	1
0.85	99	1	0	0.01	1
0.90	100	0	0	0	0
0.95	100	0	0	0	0
1.00	100	0	0	0	0