SPREADING OF A LIMITED LIFETIME INFORMATION IN NETWORKS EVOLVING BY PREFERENTIAL ATTACHMENT

Natalia Markovich

•

V.A. Trapeznikov Institute of Control Sciences, Russian Academy of Sciences,
Profsoyuznaya Str. 65, 117997 Moscow Russia
nat.markovich@gmail.com

Abstract

The paper is devoted to the information spreading (propagation) on random graphs evolving by a linear preferential attachment (PA) model. The PA is proposed to play a double role, namely, as the evolution model, i.e. the tool to add new edges and nodes to the network and (or) to remove existing nodes and edges, and as the spreading tool. We assume that a single message is to be propagated within a fixed time interval. In practice, a message may become old and not relevant. A node having a message instantaneously passes on information to one of its neighbour nodes which does not have the message yet. This neighbour may be either a node newly appended to the graph or an existing node. By probabilities of $\alpha-$, $\beta-$ and $\gamma-$schemes of the used PA model a new directed edge is drawn between a new node appended to the graph and an existing node or a new edge is drawn between a pair of existing nodes. By convention the propagation is provided if the new node (or one of the existing nodes) without the message has an incoming edge to an existing node having the information. Distributions of the number of nodes that received the message and the total number of nodes as well as the ratio of the latter random numbers in a fixed time interval with regard to parameters of the PA are obtained.

Keywords: information spreading, directed random graphs, evolution, linear preferential attachment

1. Introduction

Spreading information attracts interest due to many applications like multi-agent systems, internet traffic, parallel computation [1], [2], social networks and spreading of infections [3], [4], percolation [5] and gossip algorithms [6]. We consider the problem of spreading a single message through the directed network evolving by the preferential attachment (PA) model within a fixed time interval. A somewhat similar idea is propagated in [7] where the contact process that is evolving on graphs that are themselves evolving is studied. The PA plays a double role, namely, as the evolution model, i.e. the tool to add new edges and nodes to the network and (or) to remove existing nodes and edges, and as the spreading tool. In [8]-[9] the spreading of a unique message among a fixed number of nodes given beforehand has been considered. In [10] and [9] the PA evolution respectively without and with node and edge deletion was considered. However, a reasonable spreading time may be limited because the information may be outdated. Our objective is to obtain distributions of the number of nodes that received a single message and the total number of nodes in the graph as well as the ratio of the latter random variables (r.v.s) in the fixed time interval $[0, T^*]$ with regard to the parameters of the PA.

Let us denote the graph at evolution step k as $G(k) = (V(k), E(k))$, where $V(k)$ and $E(k)$ are sets of vertices (nodes) and edges, respectively. Let $N(k) = \|V(k)\|$ be the number of nodes in the network at the evolution step k, and let $\|A\|$ denote a cardinality of the set A. The evolution begins with an arbitrary initial directed graph $G(0)$ with at least one node and $\|E(0)\|$ edges. $N(0)$ and $\|E(0)\|$ are assumed to be fixed. We use the linear PA evolution model described in Section 2. Then $N(k)$ is a r.v. since a new node is appended to the graph with some probability.

Let $S(k)$ denote a set of nodes which have the message at the evolution step k. Indeed, $S(k) \subseteq V(k)$ holds. The spreading starts from an initial set of nodes $S(0)$ in which at least one node has the message. The spreading is unsuccesfull at step $k+1$ if a new node $j \in N(k+1)$ joins to node $i \notin S(k)$ by a new edge $(j \to i)$. Hence, $\|S(k)\|$ is a r.v. and $\|S(k)\| \leq N(k)$ holds.

Clocks of nodes are assumed to be asynchronized. As in [1], [2], we assume that new nodes arrive by Poisson time ticks of a global clock, see Fig. 1 (top).

Let $\{\tau_i\}$, $i = 0, 1, 2, ...$, $\tau_0 = 0$, be independent identically distributed (i.i.d.) exponential r.v.s with parameter λ. The sequence $\{\tau_i\}$ means the inter-arrival times of appended new nodes. The message may be propagated not at each tick of the global clock.

Since clock ticks build a Poissonian sequence, then the probability that the number of ticks $\nu(t)$ in time t is equal to $k = 0, 1, 2, ...$ is the following

$$P_k(t) \quad = \quad P\{\nu(t) = k\} = \frac{(\lambda t)^k e^{-\lambda t}}{k!}, \tag{1}$$

and the mean number of ticks in time t (or the renewal function) is $E(\nu(t)) = \lambda t$. By (1) we get $P_k(T^*)$ and $E(\nu(T^*))$ for a fixed time $T^* \geq 0$.

We aim to find the distributions of the number of nodes which have obtained the information by $K^* = \nu(T^*)$ evolution steps, i.e. $\|S(K^*)\|$, as well as of $N(K^*)$ and of the proportion of such nodes $\|S(K^*)\|/N(K^*)$. The information may be delivered to all nodes at some step k, i.e. $\|S(k)\| = N(k)$, but the propagation will be continued until the time T^* which cannot be exceeded. We suppose the evolution is without deletion of nodes and edges. Therefore, the propagated message cannot be lost since the number of nodes with the message does not decrease.

The paper is organized as follows. In Section 2, the PA model for the evolution of directed graphs is recalled. Section 3 contains our main results, namely, probability mass functions (pmfs) of $\|S(K^*)\|$ and $N(K^*)$ as well as the distribution of $\|S(K^*)\|/N(K^*)$. We finalize with conclusions. Proofs are presented in the Appendix.

2. Preferential attachment for directed graphs

By the PA model, networks are built recursively by adding nodes and edges in such a way that new nodes prefer to be connected to existing nodes if they have high node degrees. The PA networks are called scale-free [11]-[12]. It means that the node degree distribution is a power law. A discrete r.v. X exhibits a power-law distribution if

$$\mathrm{P}\left(X = i\right) \sim C i^{-(1+\iota)}, \quad i \to \infty,$$

holds for some positive constants C and ι.

We will use the so-called $\alpha-$, $\beta-$ and $\gamma-$schemes of the linear PA proposed in [11], [13] to model the evolution and information spreading. The latter PA model allows to build evolving random graphs with self-loops and multiple edges generated by the β-scheme. Examples of such evolution and spreading are shown in Fig. 1 (bottom).

Let us recall $\alpha-$, $\beta-$ and $\gamma-$schemes of the PA given in [13]. A type of the new edge is selected by flipping a 3-sided coin with probabilities α, β and γ such that $\alpha + \beta + \gamma = 1$. To this end, the iid trinomial r.v.s with values 1, 2 and 3 and the corresponding probabilities α, β and γ are generated. The parameters $\delta_{in}, \delta_{out} > 0$ allow us to determine the probabilities to select existing nodes when their in- or out-degrees are zero-valued. Let $I_k(w)$ and $O_k(w)$ denote the in- and out-degrees of node w at evolution step k. We assume $I_0(w)$ and $O_0(w)$, the in- and out-degrees in the initial graph $G(0)$, to be fixed.

Let us denote a complement as $A \setminus B$. The graph G_k is obtained from the existing graph G_{k-1} by the following $\alpha-$, $\beta-$ and $\gamma-$schemes.

- By the α-scheme, one appends a new node $v \in V(k) \setminus V(k-1)$ to $G(k-1)$, $k \geq 1$, and a new edge $(v \to w)$ to an existing node $w \in V(k-1)$ with probability α. The node w is chosen with a probability depending on its in-degree in $G(k-1)$

$$P_\alpha(k, w) \quad = \quad \frac{I_{k-1}(w) + \delta_{in}}{k - 1 + \delta_{in} N(k-1)}.$$

- By the β-scheme, one adds a new edge $(v \to w)$ to $E(k-1)$, $k \geq 1$, with probability β, where the existing nodes $v, w \in V(k-1) = V(k)$ are chosen independently from the nodes of $G(k-1)$ with probabilities

$$P_\beta(k, w, v) \quad = \quad \frac{O_{k-1}(v) + \delta_{out}}{k - 1 + \delta_{out} N(k-1)} \cdot \frac{I_{k-1}(w) + \delta_{in}}{k - 1 + \delta_{in} N(k-1)}.$$

- By the γ-scheme, one adds a new node $v \in V(k) \setminus V(k-1)$ to $G(k-1)$, $k \geq 1$, and an edge $(w \to v)$ with probability γ. The existing node $w \in V(k-1)$ is chosen with probability

$$P_\gamma(k, w) \quad = \quad \frac{O_{k-1}(w) + \delta_{out}}{k - 1 + \delta_{out} N(k-1)}.$$

Here, α, β, γ, δ_{in} and δ_{out} are parameters of the PA model. By convention, the information can be spread at step $k+1$ to a new node $v \notin S(k)$ having no information if a new edge is created by the $\alpha-$scheme or if a new edge is created by the $\beta-$scheme between two existing nodes v, w and one of them has the information (i.e. $v \in S(k)$ or $w \in S(k)$). In both cases, the new edge is to be directed to an existing node having the information.

3. Probability mass functions of $\|S(K^*)\|$ and $N(K^*)$

We assume that the nodes and edges are not removed from the graph. Then both $\|S(k)\|$ and $N(k)$ are non-decreasing functions in time. We prove that the success probability depends on the PA parameters α, β, δ_{in}, δ_{out}, and $N(K^*)$ has a Poisson distribution given a graph G_{K^*-1}.

Let the initial graph $G(0)$ contain a unique isolated node that has a message to be spread, i.e. $\|S(0)\| = \|V(0)\| = 1$, $\|E(0)\| = 0$ hold. We denote for brevity $\Omega_k = \{G(0), \cdots, G(k)\} = \{G_0, \cdots, G_k\}$ and the success probabilities for a given G_{k-1} as p_k.

Let $\sum_{c,k,j}$ denote the sum of all $\binom{k}{j} = k!/(j!(k-j)!)$ distinct index combinations among $\{j_1, j_2, ..., j_k\}$ of length j and $\mathbf{1}(A)$ denote the indicator of the event A. For example, for $k = 3$ evolution steps, the combinations with exactly two successes are the following: $(101), (110), (011)$.

Lemma 1. The conditional pmf of $\|S(k)\|$ for a maximum number of evolution steps K^* in the fixed time T^* is the following. For $1 \leq i < K^*$ it holds

$$P\{\|S(K^*)\| \quad = \quad i | G_{K^*-1}\} = e^{-\lambda T^*} \sum_{k=i}^{\infty} \frac{(\lambda T^*)^k}{k!} P\{\|S(k)\| = i | G_{k-1}\}, \tag{2}$$

where

$$P\{\|S(k)\| = i | G_{k-1}\} = \sum_{c,k,i-1} \prod_{n=1}^{i-1} p_{j_n} \prod_{m=i}^{k} (1 - p_{j_m}) \mathbf{1}\{k \geq i \geq 2\}$$
$$+ \quad \prod_{m=1}^{k} (1 - p_m) \mathbf{1}\{k \geq i = 1\} + \prod_{n=1}^{k} p_n \mathbf{1}\{i = k+1\} = \psi(i, j, k), \tag{3}$$

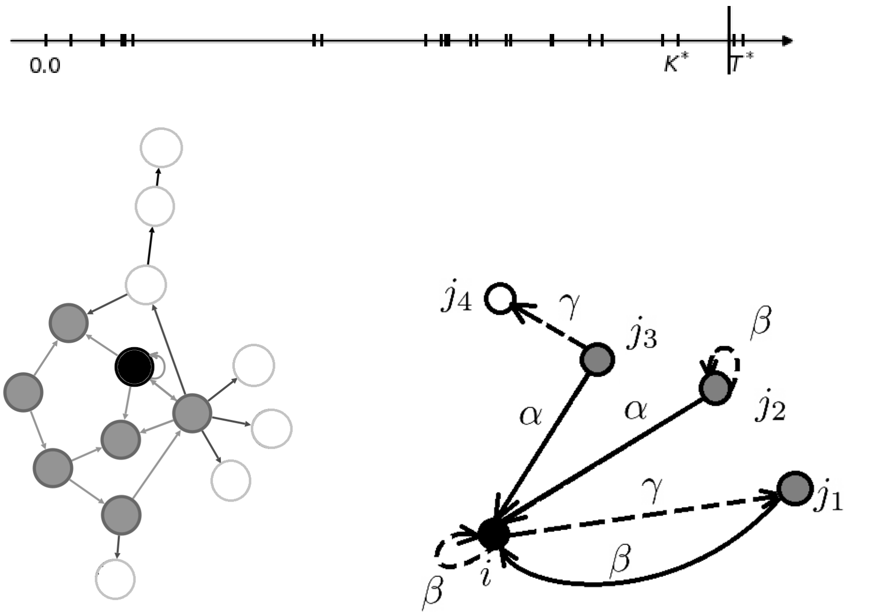


Figure 1: *Poissonian clock ticks with the fixed time T^* and the corresponding maximum number of ticks K^* are shown on top; spreading a message from an initial black node by the PA schemes with parameters $(\alpha, \beta, \gamma) = (0.35, 0.25, 0.4)$ at evolution step $k = 20$, where nodes that obtained the message are shown by grey color and nodes without the message by white color, arrows corresponding to successful spreading are shown by grey and to unsuccessful spreading by black (bottom, left); the scheme of the message spreading from node i to nodes $j_1 - j_4$ (bottom, right) with the same node coloring.*

$$P\{\|S(k)\| = i | G_{k-1}\} = 0, \qquad \textit{otherwise},$$

and success probabilities are

$$p_k = \alpha \sum_{w \in S(k-1)} P_\alpha(k, w) \quad + \quad \beta \sum_{w \in S(k-1)} \sum_{v \in V(k-1) \setminus S(k-1)} P_\beta(k, w, v), \quad k = 1, 2, ... \tag{4}$$

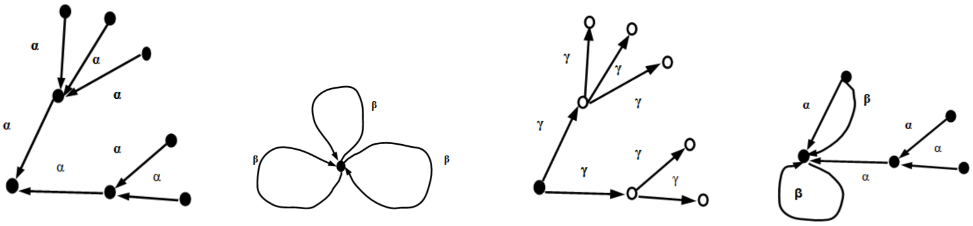
Remark 1. Note that the success probabilities are different at each evolution step and the first term in (3) corresponds to a Poisson binomial distribution.[1]

Remark 2. By (4), $p_1 = \alpha$ holds. It implies the probability that the first new node v is connected to a single node $w \in G(0)$ with the message and w shares the message with v. The second term in (4) equals to zero since the self-loop in the initial node built by the $\beta-$scheme does not lead to the success (sharing the message) or formally, since the set $V(k-1) \setminus S(k-1)$ is empty.

Corollary 1. The collection of conditional probabilities (2) forms a conditional probability distribution.

Example 1. Let us give examples of a full spreading and a full non-spreading of the message to all nodes in the graph. To this end, we consider the PA with parameters $(\alpha, \beta, \gamma) = \{(1, 0, 0), (0, 1, 0), (0, 0, 1)\}$. By (4), if $\alpha = 1$ holds at each step $k \geq 1$, then it implies $p_k = 1$, $\forall k \geq 1$ and the spreading to each new node, see Fig. 2(a); $\beta = 1$, $\forall k \geq 1$ leads to further non-spreading due to self-loops at the initial node, Fig. 2(b); $\gamma = 1$, $\forall k \geq 1$ means $p_k = 0$, $\forall k \geq 1$, i.e. no further spreading, Fig.2(c). The PA model with $\alpha + \beta = 1$, $\alpha \neq 0$, $\beta \neq 0$, $\gamma = 0$ implies $p_k = \alpha$, $\forall k \geq 1$ that means the full spreading. Despite the second term in (4) is equal to zero, $\beta-$scheme with $\beta \neq 0$ contributes to node degrees and thus, to $\sum_{w \in S(k-1)} P_\alpha(k, w)$, Fig. 2(d).

[1] An integer-valued r.v. X is called Poisson binomial and denoted as $X \sim PB(p_1, p_2, ..., p_k)$, if $X =^d \xi_1 + ... + \xi_k$, where $\xi_1, ..., \xi_k$ are independent Bernoulli r.v.s with parameters $p_1, p_2, \ldots, p_k$. The probability distribution of X is $P\{X = k\} = \sum_{A \in [k], \|A\| = j} \left(\prod_{i \in A} p_i \prod_{i \neq A} (1 - p_i) \right)$, where the sum ranges over all subsets of $[k] = \{1, ..., k\}$ of size j [14].



(a) (b) (c) (d)

Figure 2: *Examples of the PA-schemes with $(\alpha,\beta,\gamma) = (1,0,0)$ corresponding to a full spread tree (Fig. 2(a)); self-loops with $(\alpha,\beta,\gamma) = (0,1,0)$ (Fig. 2(b)) and the tree with $(\alpha,\beta,\gamma) = (0,0,1)$ (Fig.2(c)) both as full non-spreading; with (α,β,γ) such that $\alpha+\beta = 1$, $\alpha \neq 0$, $\beta \neq 0$ corresponding to the full spreading (Fig. 2(d)).*

Lemma 2. The pmf of $N(k)$, $k \geq 1$ for a maximum number of evolution steps K^* in a fixed time T^* is the Poissonian:

$$P\{N(K^*) = j|G_{K^*-1}\} = \frac{a^{j-1}}{(j-1)!e^a}, \qquad a = (\alpha+\gamma)\lambda T^*, \quad j = 1,2,...,K^*+1. \tag{5}$$

Corollary 2. If $(\alpha,\beta,\gamma) = (0,0,1)$ holds at each evolution step $k \geq 1$, then

$$P\{\|S(K^*)\| = 1|G_{K^*-1}\} = 1, \quad P\{N(K^*) = K^*+1|G_{K^*-1}\} = 1. \tag{6}$$

If $(\alpha,\beta,\gamma) = (0,1,0)$ holds at each evolution step $k \geq 1$, then

$$P\{\|S(K^*)\| = N(K^*) = 1|G_{K^*-1}\} = 1. \tag{7}$$

If $(\alpha,\beta,\gamma) = (1,0,0)$ holds at each evolution step $k \geq 1$, then

$$P\{\|S(K^*)\| = N(K^*) = K^*+1|G_{K^*-1}\} = 1. \tag{8}$$

Let us consider now the proportion of nodes which obtained the message to the total number of nodes $\|S(K^*)\|/N(K^*)$ in time T^*.

Lemma 3. Let K^* be a Poisson r.v. with the pmf $P\{K^* = k\} = (\lambda^*)^k e^{-\lambda^*}/k!$, $\lambda^* = \lambda T^*$, $k = 0,1,2,....$ Then it holds

$$\begin{aligned} 1 &\leq e^{\lambda^*} P\{\|S(K^*)\|/N(K^*) \leq x|G_{K^*-1}\} \\ &\leq e^{-(\alpha+\gamma)\lambda^*} \sum_{k=1}^{\infty} \frac{(\lambda^*)^k}{k!} \sum_{j=1}^{k+1} \frac{(\alpha+\gamma)^{j-1}(\lambda^*)^{j-1}}{(j-1)!} \sum_{i=1}^{\lfloor xj \rfloor} \psi(i,j,k) + 1, \end{aligned} \tag{9}$$

where $\psi(i,j,k)$ is determined by (3).

Lemma 4. $P\{\|S(K^*)\|/N(K^*) \leq x|G_{K^*-1}\}$, $0 < x \leq 1$, is linear in each $0 \leq p_j \leq 1$, $j = 1,...,k$ for each fixed x.

The maximum of $P\{\|S(K^*)\|/N(K^*) \leq x|G_{K^*-1}\}$ is achieved if at least one of the $\{p_j\}$ is equal to 1. The linearity does not depend on the way to select K^*.

4. Conclusions

We study the propagation of one message among nodes in an evolving network within a fixed time interval T^*. The propagation starts from the initial graph containing a single node with the message to be spread. Considering the network evolution we assume that it can be modeled by the $\alpha-$, $\beta-$ and $\gamma-$ schemes of the linear PA model proposed in [13]. It implies that a new node appended to the network at some evolution step will likely link to an existing node with a large degree. The message may be spread not at each step of the evolution but only when a new edge is directed from a new node (or an existing node) without the message to the existing node with the message. These cases correspond to the $\alpha-$ and $\beta-$ PA schemes, respectively.

The pmfs of the number of nodes obtaining the message and the total number of nodes as well as the distribution function of their ratio in the PA evolved network at time T^* are derived.

5. Acknowledgements

The author was supported by the Russian Science Foundation RSF, project number 24-21-00183.

References

[1] Damon Mosk-Aoyama and Devavrat Shah. "Computing separable functions via gossip". In: *Proceedings of the twenty-fifth annual ACM symposium on Principles of distributed computing*. PODC06. ACM, July 2006. DOI: 10.1145/1146381.1146401.

[2] Keren Censor Hillel and Hadas Shachnai. "Partial information spreading with application to distributed maximum coverage". In: *Proceedings of the 29th ACM SIGACT-SIGOPS symposium on Principles of distributed computing*. PODC '10. ACM, July 2010. DOI: 10.1145/1835698.1835739.

[3] Siddharth Patwardhan et al. "Epidemic Spreading in Group-Structured Populations". In: *Physical Review X* 13.4 (Dec. 2023). ISSN: 2160-3308. DOI: 10.1103/physrevx.13.041054.

[4] Lang Zeng et al. "Contagion dynamics in time-varying metapopulation networks with node's activity and attractiveness". In: *Chaos: An Interdisciplinary Journal of Nonlinear Science* 34.5 (May 2024). ISSN: 1089-7682. DOI: 10.1063/5.0204497.

[5] Peter Gracar and Arne Grauer. *Geometric scale-free random graphs on mobile vertices: broadcast and percolation times*. 2024. DOI: 10.48550/ARXIV.2404.15124.

[6] Devavrat Shah. "Gossip Algorithms". In: *Foundations and Trends® in Networking* 3.1 (2007), pp. 1–125. ISSN: 1554-0588. DOI: 10.1561/1300000014.

[7] Emmanuel Jacob, Amitai Linker, and Peter Mörters. *The contact process on dynamical scale-free networks*. 2022. DOI: 10.48550/ARXIV.2206.01073.

[8] Natalia M. Markovich and Maksim S. Ryzhov. "Information Spreading with Application to Non-homogeneous Evolving Networks". In: *Distributed Computer and Communication Networks*. Springer International Publishing, 2022, pp. 284–292. ISBN: 9783030971106. DOI: 10.1007/978-3-030-97110-6_22.

[9] Natalia M. Markovich and Maksim S. Ryzhov. "Information Spreading in Non-homogeneous Evolving Networks with Node and Edge Deletion". In: *Distributed Computer and Communication Networks: Control, Computation, Communications*. Springer Nature Switzerland, 2024, pp. 119–128. ISBN: 9783031504822. DOI: 10.1007/978-3-031-50482-2_10.

[10] Natalia M. Markovich and Maksim S. Ryzhov. "Information Spreading and Evolution of Non-Homogeneous Networks". In: *Advances in Systems Science and Applications* 22.2 (June 2022), pp. 21–33. DOI: 10.25728/assa.2022.22.2.1186.

[11] Béla Bollobás et al. "Directed scale-free graphs". In: Jan. 2003, pp. 132–139. DOI: 10.1145/644108.644133.

[12] Natalia Markovich and Marijus Vaičiulis. "Extreme Value Statistics for Evolving Random Networks". In: *Mathematics* 11.9 (May 2023), p. 2171. ISSN: 2227-7390. DOI: 10.3390/math11092171.

[13] Phyllis Wan et al. "Are extreme value estimation methods useful for network data?" In: *Extremes* 23.1 (Aug. 2019), pp. 171–195. ISSN: 1572-915X. DOI: 10.1007/s10687-019-00359-x.

[14] Wenpin Tang and Fengmin Tang. "The Poisson Binomial Distribution – Old & New". In: *Statistical Science* 38.1 (Feb. 2023). ISSN: 0883-4237. DOI: 10.1214/22-sts852.

Appendix

5.1. Proof of Lemma 1

Proof.

After the clock tick k, we have either $\|S(k)\| = \|S(k-1)\|$ or $\|S(k)\| = \|S(k-1)\| + 1$, $k = 1, 2, ...$, where $\|S(0)\| = 1$. The size increases, i.e. $\|S(k)\| = \|S(k-1)\| + 1$ holds, if a node $v \notin S(k-1)$ contacts a node $w \in S(k-1)$ and if the edge $(v \to w)$ is directed from v to w. The new edge $(v \to w)$ leading to the success (the increase of $\|S(k)\|$) may be created between two existing nodes $v, w \in V(k-1) = V(k)$ with probability $\beta P_\beta(k, w, v)$ if $v \in V(k-1) \setminus S(k-1) = V(k) \setminus S(k-1)$, $w \in S(k-1)$, or between a newly appended node $v \in V(k) \setminus V(k-1)$ and an existed node $w \in S(k-1)$ with probability $\alpha P_\alpha(k, w)$. Otherwise, the size does not increase, i.e. $\|S(k)\| = \|S(k-1)\|$. The $\gamma-$scheme does not lead to the information spreading from node w to node v by convention. Here, α, β and γ are defined as in Section 2.

The conditional success probability p_k to increase the number of nodes with the message at step k is the following

$$\begin{aligned} & p_k = E\{S(k) - S(k-1)|\Omega_{k-1}\} & (10)\\ = \; & \alpha \sum_{w \in S(k-1)} P_\alpha(k, w) + \beta \sum_{w \in S(k-1)} \sum_{v \in V(k-1) \setminus S(k-1)} P_\beta(k, w, v), \quad \text{for } k = 2, 3, ..., \\ p_1 \; = \; & \alpha P_\alpha(1, w \in S(0)) = \alpha, \qquad \text{for} \qquad k = 1. & (11) \end{aligned}$$

Since $G(0)$ contains a unique node, $V(0) \setminus S(0) = \emptyset$, the $\beta-$scheme can be applied at step $k = 1$, but it does not lead to the "success". The $\beta-$scheme may provide self-loops.

Let us consider the case $k \geq i \geq 2$. The creation of a direction of a new edge may be taken as the experiment. The experiments are independent since they are generated by the iid trinomial r.v.s.. We obtain the Poisson binomial pmf

$$P\{\|S(k)\| = i|\Omega_{k-1}\} = \sum_{c,k,i-1} \prod_{n=1}^{i-1} p_{j_n} \prod_{m=i}^{k} (1 - p_{j_m}), \tag{12}$$

where the sequence $\{p_{j_n}\}$ is determined by (10), (11). By the Markov property one can substitute Ω_{k-1} by G_{k-1} in the latter equality and further.
For $k = i = 1$ it holds

$$P\{\|S(1)\| = 1|G_0\} = \gamma P_\gamma(1, w \in S(0)) + \beta P_\beta(1, w \in S(0)) = \gamma + \beta = 1 - p_1 \tag{13}$$

due to $P_\gamma(1, w \in S(0)) = P_\beta(1, w \in S(0) = 1$.
For $k > i = 1$ it follows

$$\begin{aligned} P\{\|S(k)\| = 1|G_{k-1}\} &= \left(\gamma P_\gamma(1, w \in S(0)) + \beta P_\beta(1, w \in S(0))\right) \prod_{m=2}^{k} (1 - p_m) \\ &= (\gamma + \beta) \prod_{m=2}^{k} (1 - p_m) = \prod_{m=1}^{k} (1 - p_m), & (14) \end{aligned}$$

The event $\{\|S(k)\| = 1|G_{k-1}\}$ means that the set $S(k)$ contains only the initial node with the message.

Let us consider the case $0 \leq k < i$. For $k = 0$ it trivially follows

$$P\{\|S(0)\| = 1\} = 1, \qquad P\{\|S(0)\| = j\} = 0, \qquad j > 1. \tag{15}$$

RT&A, Special Issue, No 8 (85)
Volume 20, July 2025

For $i > k \geq 1$ we have

$$P\{\|S(k)\| = i|G_{k-1}\} = \begin{cases} \alpha^k \prod_{j=1}^k \sum_{\omega \in S(j-1)} P_\alpha(j,\omega), & i = k+1, \\ 0, & i > k+1. \end{cases} \tag{16}$$

Note that it holds

$$\alpha^k \prod_{j=1}^{k} \sum_{\omega \in S(j-1)} P_\alpha(j,\omega) = \prod_{n=1}^{k} p_n. \tag{17}$$

Summarizing (12)-(17) we get (3). Then it holds

$$\begin{aligned} P\{\|S(K^*)\| = i|G_{K^*-1}\} &= \sum_{k=i}^{\infty} P\{\|S(k)\| = i|K^* = k, G_{k-1}\} P\{K^* = k\} \\ &= \sum_{k=i}^{\infty} P\{\|S(k)\| = i|G_{k-1}\} P_k(T^*) \end{aligned}$$

since $S(k)$ and K^* are independent. Using $P_k(T^*)$ in (1) we get (2).

■

5.2. Proof of Corollary 1

Proof.
Since $\|S(k)\|$ has a Poisson binomial distribution and $P\{\|S(k)\| = 0\} = 0$, the collection of probabilities $P\{\|S(K^*)\| = i|\Omega_{K^*-1}\} = P\{\|S(K^*)\| = i|G_{K^*-1}\}$ forms a probability distribution:

$$\sum_{i=0}^{\infty} P\{\|S(K^*)\| = i|G_{K^*-1}\} = 1.$$

■

5.3. Proof of Lemma 2

Proof. It holds $N(k+1) = N(k) + 1$ with probability $\alpha + \gamma$ and $N(k+1) = N(k)$ with probability β. Hence, for a fixed $k \geq 1$ and $N(0) = 1$ we get

$$P\{N(k) = i+1|G_{k-1}\} = \binom{k}{i} (\alpha+\gamma)^i \beta^{k-i}, \qquad i = 0, 1, 2, ..., k.$$

Since $\sum_{i=0}^{\infty} a^i / i! = e^a$, $a > 0$ holds and due to independence of the r.v.s K^* and $N(k)$, $N(k) \leq k+1$, it follows

$$\begin{aligned} P\{N(K^*) = i+1|G_{K^*-1}\} &= \sum_{k=i}^{\infty} P\{N(k) = i+1|K^* = k, G_{K^*-1}\} P\{K^* = k\} \\ &= \sum_{k=i}^{\infty} \binom{k}{i} (\alpha+\gamma)^i \beta^{k-i} P_k(T^*) = \sum_{k=i}^{\infty} \frac{(\alpha+\gamma)^i (1-\alpha-\gamma)^{k-i}}{i!(k-i)!} (\lambda T^*)^k e^{-\lambda T^*} \\ &= \frac{a^i}{i! e^a}, \qquad \text{where} \quad a = (\alpha+\gamma)\lambda T^*. \end{aligned}$$

■

5.4. Proof of Corollary 2

Proof. Let $(\alpha,\beta,\gamma)=(0,0,1)$ hold. By (4), we have $p_k=0$, $k=1,2,....$ By (3) it follows

$$P\{\|S(k)\|=i|G_{k-1}\}=\prod_{m=1}^{k}(1-p_m)\mathbf{1}\{k\geq i=1\}=1$$

and hence, the first statement in (6) holds. Furthermore, we get

$$\begin{aligned}&P\{N(K^*)=K^*+1|G_{K^*-1}\}\\ =&\sum_{k=1}^{\infty}P\{N(k)=k+1|K^*=k,G_{K^*-1}\}P\{K^*=k\}+P\{N(0)=1\}P\{K^*=0\}\\ =&\sum_{k=0}^{\infty}P\{K^*=k\}=1.\end{aligned}$$

For $(\alpha,\beta,\gamma)=(0,1,0)$ (7) follows by (3) since the $\beta-$scheme forms only self-loops in the initial node and $p_k=0$, $k=1,2,....$ For $(\alpha,\beta,\gamma)=(1,0,0)$ (8) follows since each new node gets the message and $p_k=1$, $k=1,2,....$ ■

5.5. Proof of Lemma 3

Proof.

We have by (1)

$$\begin{aligned}&P\{\|S(K^*)\|/N(K^*)\leq x|\Omega_{K^*-1}\}\\ =&\sum_{j=1}^{\infty}P\{\|S(K^*)\|\leq xj|N(K^*)=j,\Omega_{K^*-1}\}P\{N(K^*)=j|\Omega_{K^*-1}\}\\ =&\sum_{j=1}^{\infty}\sum_{i=1}^{\lfloor xj\rfloor}P\{\|S(K^*)\|=i,N(K^*)=j|\Omega_{K^*-1}\}\\ =&\sum_{k=1}^{\infty}\sum_{j=1}^{k+1}\sum_{i=1}^{\lfloor xj\rfloor}P\{\|S(k)\|=i,N(k)=j|K^*=k,\Omega_{k-1}\}P\{K^*=k\}\\ +&\mathbf{1}\{i=j=1,k=0\}P\{K^*=0\}\\ \geq&\mathbf{1}\{\|S(0)\|=N(0)=1\}P\{K^*=0\}=e^{-\lambda^*},\end{aligned}\tag{18}$$

where $x\in(0,1]$. Using the Markov property, we can substitute Ω_k by G_k in expressions above. Note that the r.v.s $N(k)$ and $\|S(k)\|$ are dependent, $1\leq\|S(k)\|\leq N(k)\leq k+1$ for any $k\geq 0$ holds. Then by (1), (3) and (5) we get

$$\begin{aligned}&P\{\|S(K^*)\|/N(K^*)\leq x|G_{K^*-1}\}\\ \leq&\sum_{k=1}^{\infty}\sum_{j=1}^{k+1}\sum_{i=1}^{\lfloor xj\rfloor}P\{\|S(k)\|=i|K^*=k,G_{k-1}\}P\{N(k)=j|K^*=k,G_{k-1}\}P\{K^*=k\}+e^{-\lambda^*}\\ =&\sum_{k=1}^{\infty}\sum_{j=1}^{k+1}\sum_{i=1}^{\lfloor xj\rfloor}\psi(i,j,k)\frac{a^{j-1}}{(j-1)!e^a}\frac{(\lambda^*)^k e^{-\lambda^*}}{k!}+e^{-\lambda^*}\\ =&e^{-(1+\alpha+\gamma)\lambda^*}\sum_{k=1}^{\infty}\sum_{j=1}^{k+1}\sum_{i=1}^{\lfloor xj\rfloor}\psi(i,j,k)\frac{(\alpha+\gamma)^{j-1}(\lambda^*)^{j-1+k}}{(j-1)!k!}+e^{-\lambda^*}.\end{aligned}\tag{19}$$

By (18) and (19) the statement (9) follows. ■

5.6. Proof of Lemma 4

Proof.

Let us substitute the Poisson binomial r.v. $\|S(k)\|$, $k \geq 1$ in (18) by the sum of independent Bernoulli r.v.s $X_1, ..., X_k$ with success probabilities $p_1, ..., p_k$, respectively. Denoting

$$b_{i,j,k} = P\{X_2 + ... + X_k = i, N(k) = j | K^* = k, G_{k-1}\},$$
$$a_{i,j,k} = P\{X_2 + ... + X_k = i-1, N(k) = j | K^* = k, G_{k-1}\} - b_{i,j,k},$$

we get that the next probability is linear in p_1:

$$\begin{aligned} & P\{\|S(K^*)\|/N(K^*) \leq x | G_{K^*-1}\} \\ = & \sum_{k=1}^{\infty} \sum_{j=1}^{k+1} \sum_{i=1}^{\lceil xj \rceil} P\{X_1 + ... + X_k = i, N(k) = j | K^* = k, G_{k-1}\} P_k(T^*) \\ + & \mathbf{1}\{i = j = 1, k = 0\} P_0(T^*) \\ = & \sum_{k=1}^{\infty} \sum_{j=1}^{k+1} \sum_{i=1}^{\lceil xj \rceil} \Big(P\{X_2 + ... + X_k = i, N(k) = j | K^* = k, G_{k-1}\}(1 - p_1) + \mathbf{1}\{i = j = 1, k = 0\} e^{-\lambda T^*} \\ + & P\{X_2 + ... + X_k = i-1, N(k) = j | K^* = k, G_{k-1}\} p_1 \Big) P_k(T^*) + \mathbf{1}\{i = j = 1, k = 0\} e^{-\lambda T^*} \\ = & \sum_{k=1}^{\infty} \sum_{j=1}^{k+1} \sum_{i=1}^{\lceil xj \rceil} \left(b_{i,j,k} + p_1 a_{i,j,k} \right) P_k(T^*) + \mathbf{1}\{i = j = 1, k = 0\} e^{-\lambda T^*}. \end{aligned}$$

One can rewrite the latter expression in the linear form regarding p_1:

$$\phi(x, p_1) \quad = \quad f_1(x) p_1 + f_2(x),$$

where

$$\phi(x, p_1) = P\{\|S(K^*)\|/N(K^*) \leq x | G_{K^*-1}\},$$
$$f_1(x) = \sum_{k=1}^{\infty} \sum_{j=1}^{k=1} \sum_{i=1}^{\lceil xj \rceil} a_{i,j,k} P_k(T^*), \qquad f_2(x) = \sum_{k=1}^{\infty} \sum_{j=1}^{k+1} \sum_{i=1}^{\lceil xj \rceil} b_{i,j,k} P_k(T^*) + e^{-\lambda T^*}.$$

Due to symmetry, the linearity can be shown in any p_j, $1 \leq j \leq k+1$. ■