Bull Braz Math Soc, New Series 37(4), 571-580


# Random walks systems on complete graphs

Oswaldo S.M. Alves, Elcio Lebensztayn,
Fábio P. Machado and Mauricio Z. Martinez

**Abstract.** We study two versions of random walks systems on complete graphs. In the first one, the random walks have geometrically distributed lifetimes so we define and identify a non-trivial critical parameter related to the proportion of visited vertices before the process dies out. In the second version, the lifetimes depend on the past of the process in a non-Markovian setup. For that version, we present results obtained from computational analysis, simulations and a mean field approximation. These three approaches match.

**Keywords:** random walks, complete graphs, mean field approximation.

**Mathematical subject classification:** 60K35, 60J05, 60J85.

## 1 Introduction

We study two versions of a model of discrete-time random walk systems on finite graphs. This model, known as frog model, has been considered on infinite graphs, in particular hypercubic lattices and homogeneous trees, for which results as shape theorem and phase transition have been proved. See for instance [2], [3], [4] and the references therein.

Our interest in this paper is to study the behavior of this model on complete graphs. The basic form of the model is described as follows. At time zero there is one active particle in a fixed vertex of the graph. That particle performs a random walk up to the time it dies. In all other vertices there are inactive particles. At each step of the process, active particles disappear with probability $(1 - p)$ or survive with probability $p$, independently of each other. When an active particle survives, it jumps to a neighboring vertex randomly chosen. If an active particle hits a sleeping one, the latter is activated and starts to perform a random walk independently of everything else.

Received 26 April 2006.

In the case of finite graphs, the main object under study is the *coverage* of the graph, that is, the proportion of visited vertices at the end of the process, after all active particles have died. An important reference for random walks on graphs is the book of Aldous and Fill [1]. In section 2, we consider the model as described above and identify, for the class of $n$-complete graphs, the critical value of $p$ below which the sequence of coverage $\{\alpha_n(p)\}$ converges in distribution to 0 as $n \to \infty$. The section 3 is devoted to a version in which the lifetime of an active particle depends on the past of the process. It survives up to the time it hits a vertex which has been visited before by an active particle. We present results obtained from computational analysis, stochastic simulations and a mean field approximation. These approaches agree remarkably.

## 2 Geometric lifetime

In this section, we deal with the basic version of the frog model on complete graphs, whose formal definition follows. For $n \geq 3$, let $K_n$ be the $n$-complete graph (the graph with vertex set $\mathcal{V} = \{1, 2, \ldots, n\}$ and each pair of vertices linked by an edge). Consider one particle at each vertex of $K_n$, all but one being inactive. Let $\left\{\left(S_t^x\right)_{t\in\mathbb{N}} ; x \in \mathcal{V}\right\}$ and $\left\{\left(\tau_p^x\right) ; x \in \mathcal{V}\right\}$ be independent sets of independent identically distributed random objects defined as follows. For each $x \in \mathcal{V}$, $\left(S_t^x\right)_{t\in\mathbb{N}}$ is a discrete time simple random walk on $K_n$ starting from $x$ (it describes the trajectory of the particle placed initially at $x$ when it is activated), and $\tau_p^x$, which stands for the lifetime of that particle, is a random variable whose law is given by $\mathbf{P}\left(\tau_p^x = k\right) = (1-p)p^{k-1}$, $k = 1, 2, \ldots$, where $p \in [0, 1]$ is a fixed parameter. In conclusion, the particle at vertex $x$, in the event it is activated, moves as $\left(S_t^x\right)_{t\in\mathbb{N}}$ but disappears $\tau_p^x$ units of time after being activated.

**Definition 2.1.**

(i) *For a realization of the frog model in $K_n$ with parameter $p$, let $C_n(p)$ be the set of vertices of $K_n$ visited by active particles and $|C_n(p)|$ be the number of elements of $C_n(p)$. We define the* coverage *of $K_n$ by*

$$\alpha_n(p) = |C_n(p)|/n.$$

(ii) *We define the* critical parameter *of the model by*

$$p_c = \sup\ \{p : \alpha_n(p) \Rightarrow 0\ \ as\ \ n \to \infty\},$$

*where $\Rightarrow$ denotes convergence in distribution.*

Next we prove that $p_c = 1/2$, therefore the frog model has phase transition. To show this assertion, we prove that for $p \leq 1/2$ and every $\varepsilon > 0$,

$$\lim_{n\to\infty} \mathbf{P}(\alpha_n(p) \geq \varepsilon) = 0$$

and that for $p > 1/2$, there are constants $\varepsilon = \varepsilon(p) > 0$ and $\delta = \delta(p) > 0$ such that, for all $n$,

$$\mathbf{P}\left(\alpha_n(p) \geq \varepsilon\right) \geq \delta.$$

**Theorem 2.1.** $p_c = \dfrac{1}{2}$.

**Proof.** First let $p \leq 1/2$ and consider the Galton-Watson branching process that starts with a single individual and in which the family size has geometric distribution with parameter $(1-p)$. This branching process dies out with probability 1. Furthermore, if we call $Z_t$ the size of the $t^{th}$ generation of this branching process, we have that, for every $\varepsilon > 0$,

$$\mathbf{P}\left(|C_n(p)| \geq \varepsilon n\right) \leq \mathbf{P}\left(\sum_{t=1}^{\infty} Z_t \geq \varepsilon n\right) \stackrel{n\to\infty}{\longrightarrow} 0.$$

Now let $p > 1/2$. In this case, we choose $\varepsilon = \varepsilon(p) > 0$ such that $(1-\varepsilon)$ $p > 1/2$. We let the process develop up to the time when there are $\varepsilon n$ vertices visited by active particles, pointing that this event has probability bounded away from zero. Observe that up to this time the frog model dominates the following supercritical branching process. Each individual generates two descendants with probability $(1-\varepsilon)p$ and no descendant with probability $1-(1-\varepsilon)p$. Observe that at each step before reaching $\varepsilon n$ visited vertices, each active particle in the frog model hits an inactive particle (activating it) with probability larger than $(1-\varepsilon)p$. In order to avoid correlation considerations, one could consider that each active particle of the set of active particle at any time, move in its turn, being the probability of hitting the inactive set of vertices updated, according to what happened to the last jump. All the computed probabilities would be larger that $(1-\varepsilon)p$. So, with bounded away from zero probability, the number of visited vertices in frog model is at least $\varepsilon n$, for all $n$. $\square$

Note that, from the first part of the proof, $\alpha_n(1/2) \Rightarrow 0$ as $n \to \infty$.

## 3 Non-geometric lifetime

In the model considered so far, the lack of memory of the geometric distribution plays an important role in the Markovian behavior of the process. In this section, we study a model in which the lifetime of an active particle depends on the past of the process. As far as we know, there are no references about this kind of model.

We work with the following version of frog model on $K_n$. Initially, there is one particle at each vertex of $K_n$; only one is active, the others are inactive. In the event it is activated, the particle at a vertex $x$ follows an independent simple random walk on $K_n$ and activates the inactive particles that encounters along its way. However, each active particle dies at the first time it jumps on a vertex which has been visited before.

### 3.1 Computational analysis

We define $A_t$, $D_t$ and $I_t$ as the number of active particles at time $t$, the number of vertices whose original particles have already died up to time $t$ and the number of particles still inactive at time $t$, respectively. Note that $\{(A_t, D_t, I_t)\}_{t\geq 0}$ is a Markov chain (in fact, $\{(A_t, I_t)\}_{t\geq 0}$ is) such that $A_1 = 2$, $D_1 = 0$, and $I_1 = n-2$ with probability 1. Moreover, $A_t + D_t + I_t = n$ for all $t \geq 0$.

We underline two important features of this Markov chain: first, it has absorbing states, so that it stops at the time $T = \min\{t > 0 : A_t = 0\}$. Second, each state $(a, d, i)$ is achieved at most once. For $(a, d, i)$ such that $a + d + i = n$, we define

$$\begin{aligned} P(a, d, i) &= \mathbf{P}(A_t = a, D_t = d, I_t = i \ \text{ for some } \ t \geq 0) \\ &= \sum_{t\geq 0} \mathbf{P}(A_t = a, D_t = d, I_t = i). \end{aligned}$$

We have that $P(2, 0, n-2) = 1$, $P(1, 1, n-2) = 0$ and $P(0, d, 0) = 0$. We denote by $V_t = A_t + D_t = n - I_t$ the number of visited vertices at time $t$, so that the *coverage* of the process is $\alpha_n = V_T/n$.

Next we show some relations which are helpful in the computational analysis we do of the probability of the total coverage event ($\alpha_n = 1$). First observe that the probability of reaching a number $a$ of active particles without occurring any death of particles is

$$P(a, 0, i) = \sum_{j=1}^{[a/2]} \binom{i+j}{j} \left(\frac{1}{n-1}\right)^{a-j} f(a-j, j) P(a-j, 0, i+j) \tag{3.1}$$

where $a + i = n$ and

$$f(s, j) = \sum_{r=0}^{j-1} (-1)^r \binom{j}{r} (j-r)^s$$

can be though as the number of ways of displacing $s$ balls into $j$ urns leaving no empty urns. The last display is an immediate consequence of the inclusion-exclusion formula. To understand (3.1), observe that one only could reach the state $(a, 0, i)$ from a state $(a', d', i')$ with $a' = a - j$, $d' = 0$ and $i' = i + j$. For that to happen, $a - j$ active particles must wake up $j$ particles among $i + j$ inactive ones. Besides $j$ cannot be larger than $a/2$. In conclusion, (3.1) is obtained by conditioning on the previous situation of the system.

Our next result shows a general relation for $P(a, d, i)$.

**Theorem 3.1.** *For* $(a, d, i)$ *such that* $a + d + i = n$ *and* $a > 0$

$$P(a,d,i) \;=\; \sum_{d'=0}^{d} \sum_{j=1}^{[a/2]} \binom{d-d'+a-j}{d-d'} \binom{i+j}{i} \left(\frac{1}{n-1}\right)^{d+a-d'-j}$$
$$(d+a-j-1)^{d-d'} f(a-j,j) P(a+d-d'-j, d', i+j).$$

**Proof.** Again we condition on the previous state. From state $(a', d', i')$ one can visit state $(a, d, i)$ if and only if $d' \leq d$ and $i < i'$. As a consequence $i' = i + j$ and $a' = a + (d - d') - j$. Observe that $d - d'$ particles will die among the $a + (d - d') - j$ active ones. Moreover $a - j$ active particles must wake up $j$ inactive particles among the $i + j$ ones. Of course $j$ must be smaller than $a/2$. $\square$

The last relation deals with the absorbing state since it is the target of the computational analysis developed next. We claim that if $d + i = n$

$$P(0,d,i) \;=\; \sum_{a=2}^{d} (d-1n-1)^a \, P(a, d-a, i)$$

as from state $(a', d', i')$ one can reach $(0, d, i)$ if and only if $d = a' + d'$ and $i = i'$ which means that all active particles die suddenly.

Finally, note that the probability of total coverage of $K_n$ is given by

$$\rho_n \;=\; 1 - \sum_{d=2}^{n-1} P(0, d, n-d).$$

Using the formulas given above, we compute the values of $\rho_n$ for various values of $n \leq 200$ (see Table 1). The quite intensive computational task points to the fact that $\rho_n$ is decreasing in $n$ and $\rho_{200}$ is of order $10^{-14}$, so we conjecture that

$$\rho_n \downarrow 0 \text{ as } n \to \infty.$$

| $n$ | $\rho_n$ |
|---|---|
| 10 | $2.38561 \times 10^{-1}$ |
| 25 | $2.43941 \times 10^{-2}$ |
| 50 | $5.36173 \times 10^{-4}$ |
| 75 | $1.17556 \times 10^{-5}$ |
| 100 | $2.57646 \times 10^{-7}$ |
| 200 | $5.95080 \times 10^{-14}$ |

Table 1: Values of $\rho_n$.

### 3.2 Simulations

We perform simulations of the model for $n = 500$, 1000, 2000 and 5000, keeping track of the evolution of the values of $A_t/n$ and $V_t/n$. We perceive that, for each $n$, the global behavior of the process does not have significant differences among these simulations. Typical illustrations of this behavior are shown in Figure 1. Simulations indicate, therefore, the following behavior: up to the time in which $A_t/n$ achieves a value close to 0.36, $V_t$ grows fast and $A_t$ is less than $V_t$, but very close to it. That characteristic comes from the fact that, up to that time, there is enough room for the active particles to jump on the set of the unvisited vertices. From that instant on, $V_t$ begins to grow more slowly, achieving a final value very close to $0.83\,n$. Hence, it seems clear that the sequence of coverage $\{\alpha_n\}$ converges in distribution as $n \to \infty$.

### 3.3 Mean field approximation

In this section we study the model described on section 3 through a mean field approximation approach. The goal is to obtain rigorous results, consistent with the simulation performed with the original model, by an analytic and combinatorial relatively small effort. Such approach must be taken carefully as it does not consider fundamental correlations presented in the original model. Anyway it agrees nicely with what is presented in the two precedent subsections.

The Markov chain $\{(A_t, D_t, I_t)\}_{t\geq 0}$ defined at the beginning of subsection 3.1, such that $A_t$ is the number of active particles at time $t$, $D_t$ is the number of vertices whose original particles have already died up to time $t$, and $I_t$ is the number of particles still inactive at time $t$, is still under consideration. In this section, we also consider the system $\{(a_t, d_t, i_t)\}_{t\geq 0}$ which is the mean field approximation for $\{(A_t/n, D_t/n, I_t/n)\}_{t\geq 0}$. By $v_t$ we denote the sum $a_t + d_t$. Remember that $\{(A_t, D_t, I_t)\}_{t\geq 0}$ is a Markov chain (in fact, $\{(A_t, I_t)\}_{t\geq 0}$ is) such

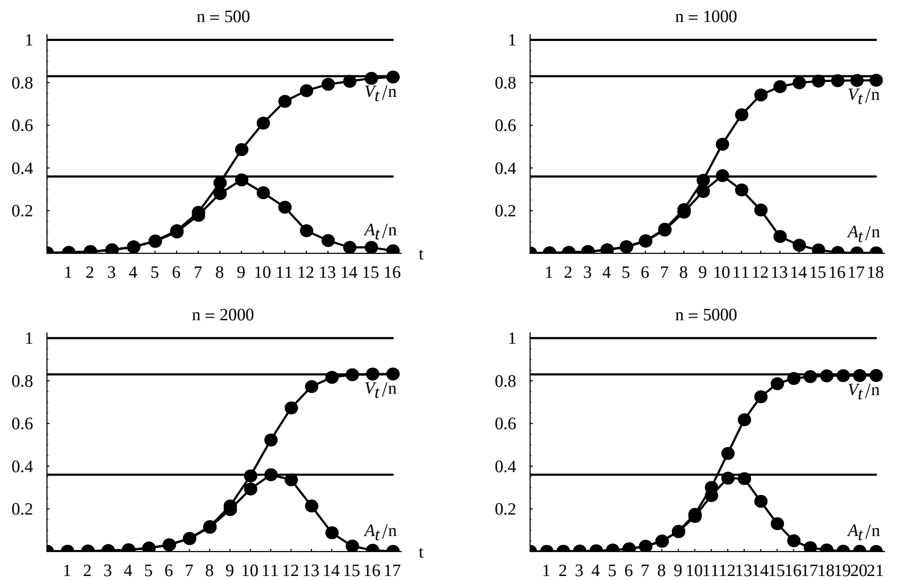


Figure 1: Evolution of $A_t/n$ and $V_t/n$ in simulations for $n = 500$, 1000, 2000 and 5000. Levels of 0.36 and 0.83 are indicated.

that $A_1 = 2$, $D_1 = 0$, $I_1 = n-2$ with probability 1. Moreover, $A_t + D_t + I_t = n$ for all $t \geq 0$.

The evolution of the process $\{(a_t, d_t, i_t)\}_{t\geq 0}$ is the following:

$$d_{t+1} \;=\; d_t + (a_t + d_t) \cdot a_t$$

as in the original model, particles which are active at time $t$ die if they jump into sites which have been visited up to time $t$ . Besides, a proportion $(1 - (a_t + d_t))$ of the amount of the set of active particles stay active, jumping on sites which have not been visited by active particles so far. Each one of the inactive particles has probability

$$\left(1 - \frac{1}{i_t \cdot n}\right)^{a_t \cdot i_t \cdot n} \;\sim\; \exp(-a_t)$$

of not being hit by an active particle. This justifies the following dynamic

$$i_{t+1} \;=\; i_t - i_t \cdot (1 - \exp(-a_t))$$

which implies that

$$a_{t+1} \;=\; a_t - (a_t + d_t) \cdot a_t + i_t \cdot (1 - \exp(-a_t)).$$

The evolution of this system of equations could be studied under any initial input but our interest is study its behavior for $a_0 = 2/n$ and $d_0 = 0$ as this is analogous to what happens in the first step of the original process.

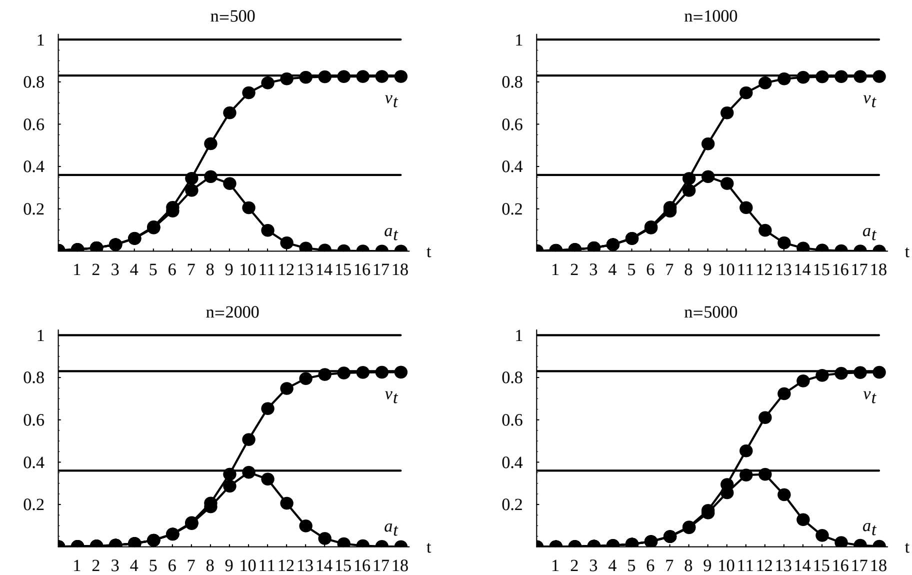


Figure 2: Evolution of $a_t$ and $v_t$ in the mean field approximation for $n = 500$, 1000, 2000 and 5000. Levels of 0.36 and 0.83 are indicated.

**Theorem 3.2.**

(i) *For all* $t, a_t > 0$ *and* $\lim_{t\to\infty} a_t = 0$;

(ii) $\exists M = M(n)$ *such that*

$$a_0 < a_1 < \cdots < a_{M-1} \leq a_M > a_{M+1} > a_{M+2} > \cdots;$$

(iii) $\ln(0.0005n)/\ln 2 < M < \ln(0.001n)/\ln 1.994 + 30$ *for* $n > 1000$;

(iv) $\lim_{t\to\infty}(a_t + d_t) \in (0,82; 0,83)$; *besides* $a_M \in (0,35; 0,36)$.

**Proof.** We present a sketch of the proof as most of it demands ordinary computations. By induction one shows that $a_t > 0$, $d_t \geq 0$ and that $0 < v_t < 1$ for all possible values of $t$. As a consequence of these facts, one proves the existence for each value of $n$ (the number of vertices of the graph), of $\lim_{t\to\infty} v_t$ and

$\lim_{t\to\infty} d_t$. Observing that $d_{t+1} = d_t + (a_t + d_t)a_t$ and using the later facts one proves that $\lim_{t\to\infty} a_t = 0$ and from this follows that $\lim_{t\to\infty} v_t = \lim_{t\to\infty} d_t$. Another consequence is that the sequence $\{a_t\}$ follows the pattern

$$a_0 < a_1 < \cdot < a_{M-1} \leq a_M > a_{M+1} > a_{M+2} > \cdots .$$

Besides, $v = \lim_{t\to\infty} v_t \geq 1/2$, and from this follows that $a_M > 0.15$. By its turn, this fact together with $a_{t+1} < 2a_t$ can be used to prove that there exists $t_1$ such that $0.001 < a_{t_1} < 0.002$. By using two subsequent terms of the Taylor's expansion for $e^x$, together with the fact that $a_t < 0.002$ implies that $d_t < 0.003a_t$ (and that happens as $a_0 = 2/n$), one gets that $0.0005n \leq 2^{t_1}$ and $1.994^{t_1} < 0.001n$. From this $(iii)$ follows. Finally one proves that $a_M \in (0,35; 0,36)$ and that $\lim_{t\to\infty}(a_t + d_t) \in (0,82; 0,83)$. $\square$

Figure 2 shows the evolution of $a_t$ and $v_t$, quantities that came from the mean field approximation, for $n = 500$, 1000, 2000 and 5000. The remarkable resemblance between Figures 1 and 2 comes from the fact that the the correlations, even though presents, are not able to mess up with the global behavior of the process. Both also agree with Table 1 as they show that the set of active particles are not able of covering the whole graph as its size increase.

**Acknowledgements.** Elcio Lebensztayn is thankful to FAPESP (05/04001–5), Mauricio Z. Martinez to FAPESP (02/07705–5) and Fábio P. Machado to CNPq (304088/2004–8), for financial support. We thank Alexandre Leichsenring and Pablo Rodriguez for helpful discussions. We would like to take this opportunity to thank our friend and co-author Oswaldo Alves for the nice time we spent together. He will be always a source of inspiration for us.

## References

[1] D. Aldous and J.A. Fill, *Reversible Markov Chains and Random Walks on Graphs*. Available at http://www.stat.berkeley.edu/users/aldous/RWG/book.html.

[2] O.S.M. Alves, *Phase transition and shape theorem for a chain reaction model*. PhD thesis, Universidade de São Paulo (2001).

[3] O.S.M. Alves, F.P. Machado and S.Yu. Popov, Phase transition for the frog model. *Electron. J. Probab.*, **7** (16) (2002), 1–25.

[4] E. Lebensztayn, F.P. Machado and S. Popov, An improved upper bound for the critical probability of the frog model on homogeneous trees. *J. Statist. Phys.* **119** (2) (2005), 331–345.

**Oswaldo S.M. Alves**
Universidade Federal de Goiás
BRASIL

**Elcio Lebensztayn**
Universidade de São Paulo
BRASIL

E-mail: elcio@ime.usp.br

**Fábio P. Machado**
Universidade de São Paulo
BRASIL

E-mail: fmachado@ime.usp.br

**Mauricio Z. Martinez**
Universidade de São Paulo
BRASIL

E-mail: zuluaga@ime.usp.br