Analysis of an on-line algorithm for solving large Markov chains

[Extended Abstract]

Nelly Litvak
University of Twente
P.O. Box 217
7500 AE, Enschede, The Netherlands
N.Litvak@ewi.utwente.nl

Philippe Robert
INRIA Paris — Rocquencourt
Domaine de Voluceau,
78153 Le Chesnay, France
Philippe.Robert@inria.fr

ABSTRACT

Algorithms for ranking of web pages such as Google PageRank assign importance scores according to a stationary distribution of a Markov random walk on the web graph. Although in the classical search scheme the ranking scores are pre-computed off-line, several challenging problems in contemporary web search, such as personalized search and search in entity graphs, require on-line PageRank computation. In this work we present a probabilistic point of view for an original on-line algorithm proposed by Abiteboul, Preda and Cobena [1]. According to this algorithm, at the beginning, each page receives an equal amount of 'cash', and every time when a page is visited by a random walk, it distributes its cash among its outgoing links. The PageRank score of a page is then proportional to the amount of cash transferred from this page. In this paper, instead of dealing with the variable 'cash', which is continuous, we create a two-dimensional discrete 'cat and mouse' Markov chain such that the amount of cash on each page can be expressed via probabilities for this new Markov chain. We also indicate further research directions, such as the analysis of the cat and mouse chain in the case when the cat's movements are described by a classical stochastic process such as the $M/M/1$ random walk.

Categories and Subject Descriptors

G.3 [**Probability and Statistics**]: Markov processes

General Terms

Algorithms,Theory

Keywords

Convergence to equilibrium, Directed graphs, Scaled processes

1. INTRODUCTION

Various influential algorithms for ranking of web pages such as Google PageRank assign importance scores according to a stationary distribution of a Markov random walk on the web graph. The PageRank algorithm, as introduced by Brin and Page [5] in 1998, is formulated as follows. Consider a surfer hopping from one page to another. With probability α, the surfer follows a randomly chosen outgoing link of a current page. Otherwise, with probability $(1-\alpha)$, the surfer jumps to a randomly chosen page. Originally, the parameter α was set equal to 0.85. Such surfing process and its various modifications are modelled as a Markov chain. The PageRank score of page x equals to the stationary probability corresponding to this page, and the pages are ranked according to these scores. The random jump to an arbitrary page ensures the existence and uniqueness of the stationary distribution. It is easy to see that the PageRank ranking scheme implies that a page is ranked high if many important pages link to it.

In the basic web search scheme, the PageRank ranking scores are pre-computed off-line, which requires efficient numerical techniques. The new application in web search triggered a rapid progress in the traditional area of algorithmic solutions for large Markov chains (see [10] for the classical treatment of the subject). The challenges of solving the 'world largest Markov chain' of several billion nodes include speeding up known algorithms, such as power iterations and other linear algebraic methods (see e.g. [4, 7]), as well as developing and testing alternative techniques [1, 3]. Furthermore, in practice, computational efficiency greatly depends on implementation factors such as fast access to link data, which must be stored in a distributed way, and appropriate pruning of links [4]. This causes considerable difficulties in estimation of realistic computational costs and in comparative analysis of various algorithms. These issues are however beyond the scope of this paper. For more detail on PageRank computation we refer to excellent surveys [4, 7] and references therein.

Although the off-line linear-algebraic techniques are dominating in PageRank computations, the on-line algorithms that update the ranking scores while crawling the graph, may be useful for solving a number of challenging problems in contemporary web search. Such challenges include personalized search and search in entity-relation graphs, which requires on-line computation of ranking scores (see e.g. [6]).

The goal of the present work is to present a probabilistic point of view for the algorithm proposed by Abiteboul *et al.* [1]. This original algorithm can be actually applied for computing a stationary distribution of any finite recurrent Markov chain.

According to the algorithm in [1], at the beginning, each page receives an equal amount of 'cash', and every time when a page is visited by a random walk, it distributes its cash among its outgoing links. The PageRank score of a page is then proportional to the amount of cash transferred from this page. The dynamics of the algorithm can be described in terms of a two-dimensional cat and mouse *discrete* Markov chain as follows. Assume that the movements of the crawler over web pages constitute a Markov chain, and we are interested in determining its stationary distribution. In our cat and mouse model, the movement of the cat follow the same transition probabilities as the crawler, whereas the mouse remains in the same state, except when the cat occupies this particular state, at which time the mouse moves with the transition probabilities of the crawler. The distribution of cash among pages is expressed in terms of the probability distribution of location of the mouse given the successive positions of the cat. The probabilistic properties of the cat and mouse Markov chain are of independent mathematical interest.

2. PROBLEM FORMULATION

Consider an irreducible Markov chain with a finite state space $\mathcal{S} = \{1, 2, \ldots, N\}$ and transition matrix $P = (p(\cdot, \cdot))$, its stationary distribution is denoted by $\underline{\pi} = (\pi(1), \ldots, \pi(N))$. We assume that there is no self-loop, that is, $p(x, x) = 0$ for all $x \in \mathcal{S}$. The algorithm suggested in [1] can be used to compute the stationary distribution $\underline{\pi}$ of the Markov chain in the following way. Each node initially gets $1/N$ of cash. Then a crawler starts crawling the pages (nodes) one after another in some order, deterministic or stochastic.

Let C_t be the position of the crawler at time $t = 0, 1, \ldots$. In the original paper [1] several crawling strategies were considered, such as random (crawl a randomly chosen page), greedy (crawl the page with highest cash), and cyclic (crawl pages in cyclic order). In this article, it is assumed that the position of the crawler is a Markov chain with transition matrix P.

For each $x \in \mathcal{S}$, let the *cash* value $V_t(x)$ be the amount of cash stored at node x at the beginning of the t-th step of the algorithm. The initial condition is then $V_0(x) = 1/N$ for all $x \in \mathcal{S}$. At the subsequent steps, when a node x is crawled, it distributes all its cash among its descendants proportionally to the transition probabilities $p(x, y)$, $y \in \mathcal{S}$. In the PageRank context, this means distributing the cash among the outgoing links of page x. Such 'transaction' by node x is added to its *credit history.* Here the credit history of node x at time t is the total amount of cash that the node x has *given away* on the time interval $[0, t)$. We denote this quantity by $H_t(x)$. Formally, we write

$$H_t(x) = \sum_{s=0}^{t-1} V_s(x) \mathbf{1}\{\text{crawler visits node } x \text{ at time } s\},$$

where $\mathbf{1}\{\cdot\}$ is an indicator function. Note that $H_t(x)$ and $V_t(x)$ are, respectively, the amounts of cash received by the node x before and after the last time it was crawled on $[0, t)$. We also define the total history

$$H_t = \sum_{x \in \mathcal{S}} H_t(x).$$

Now, according to [1], at each step t, the estimator of $\pi(x)$ is given by

$$\pi_t(x) = (H_t(x) + V_t(x)) / (1 + H_t), \; x \in \mathcal{S}. \quad (1)$$

In words, $\pi_t(x)$ is the fraction of cash received by node x on $[0, t)$, compared to the total amount of cash received on $[0, t)$ by all nodes together. The algorithm can be used for computation of $\underline{\pi}$ in crawling time and for updating $\underline{\pi}$ in dynamic Markov chains, where the transition matrix P changes in time, adjusting to the changes in the underlying graph structure.

This article reports the work in progress on convergence and properties of the described algorithm. In Section 3 we show that the algorithm converges to the correct stationary distribution. Further, in Section 4 we focus on the cash distribution, and we interpret the continuous cash process via a discrete two-dimensional Markov chain that we call the *cat and mouse* chain. In Section 5 we derive a stationary distribution of the cat and mouse chain, and we use these results in Section 6 to deduce the speed of convergence of the algorithm. Finally, in Section 7 we discuss further research that concerns scaling properties of the process describing the mouse position in case when the cat follows a well-known stochastic process such as the $M/M/1$ random walk. Throughout the paper, we present several open questions and mention possible extensions of the results. These extensions, along with more details, proofs and further results will be included in our upcoming paper [8].

3. CONVERGENCE RESULT

For Markov chains of a special structure, Abiteboul et al. [1] proved that $\pi_t(x)$ in (1) converges to $\pi(x)$ as $t \to \infty$, as long as the crawling process visits each node infinitely often. Their proof can be simplified and generalized to an arbitrary finite irreducible transition matrix P. In Proposition 1 we present the result, and we provide our version of the proof since the argument used there is important in what follows.

Define a vector $\underline{\pi}_t = (\pi_t(1), \pi_t(2), \ldots, \pi_t(N))$. Then the next convergence result holds.

PROPOSITION 1. *If the crawling process (C_t) is a Markov chain with transition matrix P then*

$$\lim_{t \to \infty} \underline{\pi}_t \overset{a.s.}{=} \underline{\pi}.$$

PROOF. First, we first prove that

$$\lim_{t \to \infty} H_t \overset{a.s.}{=} \infty. \quad (2)$$

Indeed, assume that at time $t = 0$, the Markov chain (C_t) is at state x. Consider a cover time defined as a period of time that is needed for the crawling process (C_t) to visit all pages and come back to x. Then (2) follows from the second Borel-Cantelli lemma because a cover time is finite with probability 1, subsequent cover times are independent, and the amount of cash redistributed during a cover time is not less than one.

Now, we use the following easily verified equation from [1]:

$$H_t(x) + V_t(x) = 1/N + \sum_y p(y, x) H_t(y), \quad x \in \mathcal{S}. \quad (3)$$

Equation (3) reflects the fact that all the cash that the page x ever had, consists of the initial amount $1/N$ and the cash received by time t from the pages linking to x. Dividing both sides of (3) by $(H_t + 1)$, we obtain

$$\pi_t(x) = [(1/N) - V_t(x)]/(H_t+1) + \sum_y p(y,x)\pi_t(y), \quad x \in \mathcal{S}. \tag{4}$$

Define $\underline{V}_t = (V_t(1), \ldots, V_t(N))$ and let $\underline{1}$ be the row-vector of ones. Further, let Π be the $N \times N$ matrix whose all rows equal $\underline{\pi}$. Then it is well known and easy to verify that the solution for the linear system (4) is

$$\underline{\pi}_t = \underline{\pi} + [1/(H_t+1)][\underline{V}_t - (1/N)\underline{1}] \sum_{n=0}^{\infty} [P^n - \Pi]. \tag{5}$$

Note that in the right-hand side, the vector $[\underline{V}_t - (1/N)\underline{1}]$ is bounded. Moreover, it is well known that the matrix summation here is finite since it represents the *deviation matrix* of the Markov chain P. Thus, the result follows from the convergence of H_t to infinity a.s. as $t \to \infty$. □

Two remarks can be made about the proof. First, the proof relies heavily on the fact that the Markov chain is finite. Indeed, otherwise, the cover times are infinite, and we can not show that the total credit history goes to infinity. The convergence of the algorithm for infinite recurrent chains remains an open question. Second, the transition matrix of the process (C_t) did not play any role in the proof. Hence, in the statement of the proposition, (C_t) can be actually an arbitrary process with finite cover times. We also note that for a process on a finite state space, the finiteness of cover times is equivalent to the condition that each state is visited infinitely often.

4. CASH DYNAMICS DESCRIBED BY THE CAT AND MOUSE MARKOV CHAIN

Having verified that the algorithm converges to the correct stationary distribution, we would like to investigate the speed of convergence. From Proposition 1 we know that the speed of convergence depends on the factor $1/H_t$ where H_t is the total amount of all transactions on $[0, t)$. Thus, it is essential to have information about the (average) growth of the history in a 'stationary regime'. To this end, we would like to investigate the properties of the cash process $(V_t(x))$ for $x \in \mathcal{S}$. From the description of the algorithm it follows that for all $x \in \mathcal{S}$, $t \geq 0$ the dynamics of the cash satisfies a recurrent equation

$$V_{t+1}(x) = \sum_{y \neq x} \mathbf{1}\{C_t = y\} p(y,x) V_t(y) + \mathbf{1}\{C_t \neq x\} V_t(x). \tag{6}$$

Indeed, if a node $y \neq x$ is crawled then the node x keeps all its cash $V_t(x)$ plus it receives the amount $p(y,x)V_t(y)$ of cash from the node y. Otherwise, if the node x is crawled then it distributes all its cash, and in this case, $V_{t+1}(x)$ is zero. Equation (6) reflects a complex structure of the cash process, and, in particular, its 'semi-discrete-semi-continuous' nature, which leads to great difficulties in analyzing its evolution and even its convergence to stationarity. In order to tackle these problems, we propose to interpret the inconvenient process $(C_t, \underline{V}_t)_{t\geq 0}$ via a much simpler and intuitively more appealing two-dimensional discrete Markov process that we call a *cat and mouse* Markov chain.

Suppose that the cat visits the nodes from $\mathcal{S}$ according to the Markov chain with transition matrix P, thus the cat random walk coincides with the crawling process (C_t) with stationary distribution $\underline{\pi}$. The mouse is 'hiding' at some node until this node is visited by the cat. Once 'found' by the cat, the mouse makes one step according to the same transition matrix P and moves to its new position until again discovered by the cat. We note that several other versions of the cat and mouse games have been studied in the literature, and some of them are presented in the book of Aldous and Fill [2].

Let C_t and M_t be the position of the cat, respectively the mouse at time t. The Markov chain on $\mathcal{S} \times \mathcal{S}$ associated with the cat and mouse game $(C_t, M_t)_{t\geq 0}$ has a transition matrix $Q = (q(\cdot,\cdot))$, where $q[(x,y),(z,w)]$ is the probability that at time $t+1$ the cat's position is z and mouse's position is w provided that at time t the cat's position is x and mouse's position is y. According to the rules of the game, we obtain

$$\begin{aligned} q[(x,y),(z,y)] &= p(x,z) && \text{if } x \neq y; \\ q[(y,y),(z,w)] &= p(y,z)p(y,w). \end{aligned}$$

Now we claim that $V_t(x)$, the amount of cash at node x, can be described through the (conditional) probability that the mouse is at node x. Let us illustrate that the dynamics of the cash and the probability that a node contains the mouse is indeed the same. First of all, since the total amount of cash is 1 we can always choose the initial distribution of the cat and mouse Markov chain in such a way that $\mathbb{P}(M_0 = x) = V_0(x)$ for all $x \in \mathcal{S}$. Now assume that the node y has just been visited by the cat. Then the amount of cash at this node becomes zero, and the probability that the mouse is at this node also becomes zero. Furthermore, for any node $x \neq y$, the amount of cash becomes $V_0(x) + p(y,x)V_0(y)$, and this is exactly the probability that the mouse is at such node since, with probability $V_0(x)$ the mouse already was there, and with probability $p(y,x)V_0(y)$ the mouse came to x after meeting the cat at y. We conclude that for each node x, the probability that the mouse is at x and the amount of cash at x obey the same recurrence relation, given by (6).

Now, formally, let $\mathcal{F}_t$ denote the history of the motion of the cat. In other words, $\mathcal{F}_t$ is the σ-field generated by the variables C_0, C_1, $\ldots$, C_t. Then the cash process can be described as in the next theorem that we state here without the proof.

THEOREM 1. *For $t \geq 0$, the identity*

$$(V_t(x), x \in \mathcal{S}) \overset{dist.}{=} (\mathbb{P}[M_t = x \mid \mathcal{F}_{t-1}], x \in \mathcal{S}) \tag{7}$$

holds in distribution. In particular, for $x \in \mathcal{S}$,

$$\mathbb{E}(V_t(x)) = \mathbb{P}(M_t = x).$$

Theorem 1 transforms the complicated crawling-and-cash process into a relatively simple two-dimensional Markov chain. This enables us, for instance, study the properties of the algorithm in stationarity. For example, since the chain (C_t, M_t) is recurrent, we easily obtain the following corollary.

COROLLARY 1. *For each $x \in \mathcal{S}$, there exists a limit*

$$\lim_{t\to\infty} \mathbb{E}(V_t(x)) = \lim_{t\to\infty} \mathbb{P}(M_t = x).$$

5. STATIONARY DISTRIBUTION OF THE CAT AND MOUSE CHAIN.

If $\nu = \mathbb{P}(C_\infty = \cdot M_\infty = \cdot)$ denotes the invariant distribution of the cat and mouse Markov chain, then from the transition matrix of the cat and mouse game, we obtain

$$\nu(x,y) = \sum_{z \neq y} \nu(z,y)p(z,x) + \sum_z \nu(z,z)p(z,x)p(z,y), \quad x,y \in \mathcal{S}. \tag{8}$$

Since the first coordinate (C_t) is a Markov chain with transition matrix $(p(\cdot,\cdot))$, then obviously

$$\sum_y \nu(x,y) = \pi(x), \qquad x \in \mathcal{S}.$$

By summing up equation (8) with respect to x, one gets that

$$\sum_{z \in \mathcal{S}} \nu(z,z)p(z,y) = \nu(y,y),$$

and therefore that there exists some constant c such that,

$$\nu(y,y) = c\pi(y), \quad y \in \mathcal{S}. \tag{9}$$

Note that necessarily $c < 1$.

The equation (9) reveals a remarkable property of the cat and mouse game. It turns out that in a stationary regime, given that the cat is at node x, the probability to find a mouse at this node equals c, and is the same for each $x \in \mathcal{S}$.

Let us now derive the value of c and the stationary distribution of the mouse component, $\mathbb{P}(M_\infty = y)$, $y \in \mathcal{S}$. This can be done, for instance, using a renewal theory argument. Define as a cycle the time between two successive time instants when the cat and the mouse met at node $y \in \mathcal{S}$. According to (9), the average cycle length is $(c\pi(y))^{-1}$. Furthermore, after the mouse meets the cat at y and leaves this node, it comes back to y from some node x with probability $\pi(x)p(x,y)/\pi(y)$. Finally, the expected time that the mouse has to wait at y till its next move is $\mathbb{E}_x(T_y)$. Therefore, the stationary probability that the mouse is at y equals

$$c\sum_x \pi(x)p(x,y)\mathbb{E}_x(T_y) = c\mathbb{E}_\pi\left(p(C_0,y)T_y\right).$$

Now, the expression for c can be derived from the normalizing condition by summing the right-hand side and equating it to one. As a result, one gets that

$$c^{-1} = \sum_z \mathbb{E}_\pi\left(p(C_0,z)\,T_z\right), \tag{10}$$

$$\mathbb{P}(M_\infty = y) = \frac{\mathbb{E}_\pi\left(p(C_0,y)T_y\right)}{\sum_z \mathbb{E}_\pi\left(p(C_0,z)\,T_z\right)}. \tag{11}$$

5.1 The case of reversible Markov chains

An interesting special case arises if we assume that (C_t) is a reversible Markov chain, that is,

$$\pi(x)p(x,y) = \pi(y)p(y,x), \quad x,y \in \mathcal{S}.$$

Then we obtain

$$\begin{aligned}\mathbb{E}_\pi\left(p(C_0,y)\,T_y\right) &= \sum_x \pi(x)p(x,y)\mathbb{E}_x(T_y)\\ &= \sum_x \pi(y)p(y,x)\mathbb{E}_x(T_y) = \pi(y)\mathbb{E}_y(T_y - 1)\\ &= 1 - \pi(y).\end{aligned}$$

Consequently, if N is the cardinality of $\mathcal{S}$, then from (10) and (11) we get simple formulas

$$c = \frac{1}{N-1} \text{ and } \mathbb{P}(M_\infty = y) = \frac{1-\pi(y)}{N-1}. \tag{12}$$

Tetali [11] showed by linear algebraic methods that if (C_n) is a general recurrent Markov chain, then

$$\sum_z \mathbb{E}_\pi\left(p(C_0,z)\,T_z\right) \leq N - 1. \tag{13}$$

It follows that the value $c = 1/(N-1)$ obtained for reversible chains, is the minimal possible value of c. As mentioned in [2], it would be nice to provide a probabilistic proof of (13). We believe that such proof can be found in the framework of the cat and mouse game. However, so far it remains an open problem.

5.2 Where does the mouse stay?

It follows from equation (12) that in case of reversible chains, the mouse stays more often in nodes y with a smaller value of $\pi(y)$. It is tempting to deduce that this property holds for any Markov chain. This conjecture however is false in general. Formula (11) merely suggests that the mouse hides more often at node y which is 'difficult to find' if we make a step back from y (following the reversed chain (C_t^*)), and then try to reach y again.

For better intuition on this regard, consider a Markov chain that consists of r cycles with one common node 0. The state space is given by

$$\mathcal{S} = \{0, (1,1), \ldots, (1,m_1), \cdots, (r,1), \cdots, (r,m_r)\},$$

and for all $k = 1,\ldots,r$ the transition probabilities are: $p(x,y) = 1$ if $x = (k,i)$, $y = (k,i-1)$, $i = 2,\ldots,m_k$; $p(x,0) = 1$ if $x = (k,1)$; $p(0,y) = 1/r$ if $y = (k,m_k)$. Define $m = m_1 + m_2 + \cdots + m_r$. It is easy to see that

$$\pi(0) = \frac{r}{m+r} \text{ and } \pi(y) = \frac{1}{m+r},\ y \neq 0.$$

Further, the value $E_\pi(p(C_0,y)T_y)$ equals $\pi(y)(m - m_k + r)$ if $y = (k,m_k)$, $k = 1,\ldots,r$, and it equals $\pi(y)$ for all other values of $y \in \mathcal{S}$. Observe that for any y distinct from 0 and (k,m_k), we have $\pi(0) > \pi(y)$ but the probability to find a mouse in 0 is larger than in y. This is because if the mouse is at one of the nodes distinct from (k,m_k), then the cat finds the mouse immediately, so the mouse has to move together with the cat, and thus it visits node 0 relatively often. The situation is different when the mouse is at node (k,m_k) for some $k = 1,\ldots,r$. Here it will take on average longer for the cat to find the mouse. So these are the mouse's 'favorite' nodes. Note that in this example one easily obtains $c = 1/r$.

6. ON THE CONVERGENCE SPEED OF THE ALGORITHM

The analysis of the cat and mouse game sheds some light on the convergence of the on-line algorithm described in Section 2. It follows from (5) that the difference between $\underline{\pi}_t$ and $\underline{\pi}$ is proportional to $1/H_t$, where the history H_t is the sum of all transactions on $[0,t)$. Now, according to Theorem 1, the average amount of cash at node x is equal to the probability that the mouse is at this node. Furthermore, (9) implies that whichever node is crawled, this value remains the same. Hence, the average size of each transaction is the same and equals c. We conclude that, in stationarity, the average history H_t grows exactly linearly in time, as ct. We observed such linearity in simulations, and similar phenomenon was also mentioned in [1] for various crawling strategies.

As we know that the average total history in a stationary regime is ct, and the convergence is reversely proportional to the history, we deduce that reversibility implies a relatively slow convergence of the algorithm. It is unlikely that larger c can be compensated by the factor $[\underline{V}_t - (1/N)\underline{1}]\sum_{n=0}^{\infty}[P^n - \Pi]$ in (5), since the deviation matrix is related to mixing times (see Chapter 2 of [2]), and, as was conjectured in Chapter 9 of [2], mixing times tend to be larger for reversible chains.

Finally, we note that the results presented here can be generalized to other crawling strategies, when (C_t) is a Markov chain governed by the matrix different from P, while the algorithm computes the stationary distribution of P. This will generalize the random and cyclic crawling strategies from [1] and allow to mathematically show which strategy gives the best convergence.

7. FURTHER RESEARCH

Motivated by the studies of the on-line algorithm, the cat and mouse process is actually interesting on its own right. Given the walk of the cat (C_t), which behavior of the mouse can we expect? In particular, we are interested in the cat and mouse game on infinite state space, and we would like to find a proper scaling for the process describing the mouse's position.

For instance, consider the case when (C_t) is the classical continuous-time $M/M/1$ Markov process on $\mathbb{Z}_+$. Such process stays at each state for an exponential time. The transition rate to $x+1$ is λ for all $x \geq 0$, and the transition rate to $x-1$ is μ for $x > 0$. All other transition rates are zero. In terms of transition probabilities, if the current state is 0, then the next state is 1 with probability one; for $x > 0$, the transition probability to $x+1$ is $p = \lambda/(\lambda+\mu)$, and the transition probability to $x-1$ is $q = \mu/(\lambda+\mu)$. We assume that (C_t) is recurrent, that is, $\rho = \lambda/\mu = p/q < 1$.

When the cat visits the state $x > 0$ where the mouse is hiding, the mouse moves instantly to $x+1$ with probability p or to $x-1$ with probability q. If cat and mouse meet at $x = 0$, the mouse immediately moves to $x = 1$. Let M_t be the position of the mouse at time $t > 0$. We would like study how the process (M_t) evolves if the cat and the mouse start together from some remote position $x \to \infty$.

After the cat and the mouse meet at x, there are several possibilities. First, with probability p^2, the cat and the mouse both make a step up and meet again at position $x+1$ after an exponential time with parameter $\mu+\lambda$. Analogously, with probability q^2, they meet at the next step at $x-1$. With probability pq, the mouse makes a step down, and the cat makes a step up. In this case, they will meet again at $x-1$ after a small number of exponentially distributed steps. Finally, with the same probability pq, the mouse will go up and the cat will go down. In this case, with probability $(1-\rho^2)/(1-\rho^{x-1}) \approx (1-\rho^2)$, the cat will reach zero earlier than $x+1$.

It follows that, as $x \to \infty$, the mouse makes a geometric number G of steps before the cat leaves to zero, where

$$\mathbb{P}(G = k) = [1 - pq(1-\rho^2)]^k pq(1-\rho^2), \quad k = 0, 1, \ldots.$$

Given that the cat did not leave towards zero, the mouse makes a step X, where

$$\mathbb{P}(X = +1) = \frac{p^2 + pq\rho^2}{1 - pq(1-\rho^2)} = \frac{p^2(1+\rho)}{1 - pq(1-\rho^2)}$$

and

$$P(X = -1) = \frac{q^2 + pq}{1 - pq(1-\rho^2)} = \frac{q}{1 - pq(1-\rho^2)}.$$

The last step of the mouse is always of size $+1$, after which the cat heads towards zero. Further, the cat will return to $(x+1)$ after a random time T_{x+1}, and it is known that, as $x \to \infty$, the random variable $T_{x+1}\rho^{x+1}$ converges to an exponential random variable with parameter $(\mu-\lambda)^2/\mu$ (see p.120 of [9]). Note that, as $x \to \infty$, the time between the moment when the cat and the mouse meet at x and the moment when the cat leaves to zero, is a finite random variable independent of x, while the time needed for the cat to reach zero is linear in x, and the time it takes to return to the neighborhood of x is of the order ρ^{-x}.

In this set up, the following 'free' continuous-time Markov process $(\tilde{M}_t)$ can serve as an approximation of the mouse's behavior provided that a current state of the mouse is far away from zero. At state x, the transition rate of $(\tilde{M}_t)$ is $\rho^x(\mu-\lambda)^2/\mu$. At each transition, the process $(\tilde{M}_t)$ makes a jump of a random size

$$\Delta = 1 + \sum_{i=1}^{G} X_i,$$

where $X_1, X_2, \ldots$ are i.i.d.'s distributed as X. From the argument above it follows that, as $x \to \infty$, the asymptotic and scaling properties of $(\tilde{M}_t)$ are similar to the ones of (M_t). Therefore, we may concentrate on $(\tilde{M}_t)$.

First, we find

$$\mathbb{E}(\Delta) = \mathbb{E}(G)\mathbb{E}(X) + 1 = -\rho^{-1},$$

Thus, the random walk of the mouse has a negative drift. Now, assume that the process $(\tilde{M}_t)$ starts at x. Then after the first transition of size Δ, the transition rate becomes $\rho^{x+\Delta}(\mu-\lambda)^2/\mu$. Thus, the expected time until the next transition is

$$\mathbb{E}\left(\rho^{-(x+\Delta)}\mu/(\mu-\lambda)^2\right) = \rho^{-x}\mu\mathbb{E}\left(\rho^{-\Delta}\right)/(\mu-\lambda)^2.$$

From the definition of Δ, we obtain a surprising result $\mathbb{E}(\rho^{-\Delta}) = 1$ and thus,

$$\mathbb{E}\left(\rho^{-(x+\Delta)}\mu/(\mu-\lambda)^2\right) = \rho^{-x}\mu/(\mu-\lambda)^2.$$

That is, after starting at x, the average time between transitions does not change. These observations help in choosing the right scaling for $(\tilde{M}_t)$, leading to the following scaled process:

$$\bar{M}_t(x) = \frac{\tilde{M}_{[\rho^{-x}\mu/(\mu-\lambda)^2]t}}{x}, \quad t \geq 0.$$

Simulations confirm that this process has a non-trivial limiting behavior as $x \to \infty$. An example of one realization for $\lambda = 0.3$ and $\mu = 0.7$, is given in Figure 1.

We observe that the process $\bar{M}_t(x)$ is equal one for initial period of time, after which it instantaneously drops to zero. In the realization in Figure 1, the transition from one to zero happens at time $t \approx 17.75$. However, in other realizations, we have seen that the scaled process stays at state one for a

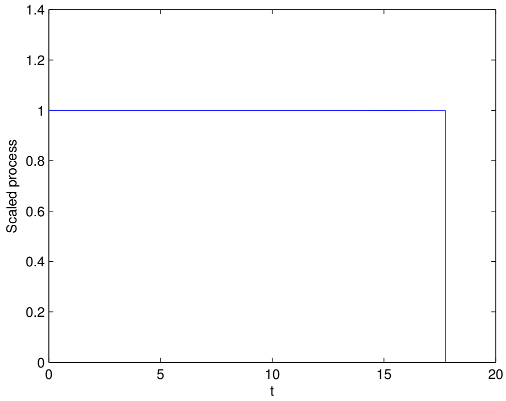

Figure 1: A realization of $\bar{M}_t(x)$, $x = 10^4$

random time, varying, in a dozen of experiments, from 0.2 to as much as 200.

Let us try to provide some intuition for a sudden drop of the scaled process from one to zero. Consider some function $f(x)$ such that

$$\lim_{x\to\infty} f(x) = \infty, \quad \lim_{x\to\infty} f(x)/x \to 0.$$

As we saw above, the random walk of a mouse has a negative drift. Thus, with probability one, at some point, the mouse will leave x heading to zero. Moreover, with probability one, the mouse will reach a point $y = x - f(x)$, where we assume that x is large. The average time that the scaled process spends at y is then $\rho^{-y}/\rho^{-x} = \rho^{f(x)} \to 0$ as $x \to \infty$. Thus, once the mouse has drifted from x towards zero and reached y, the encounters of the cat and the mouse become considerably more frequent, and for the scaled process, the movement to zero happens 'in no time'. We note also that the scaled process can not go beyond 1 because for that the mouse has to reach a level $(1+\varepsilon)x$ for some $\varepsilon > 0$ before it reaches zero. When $x \to \infty$, the probability of such event goes to zero for any positive ε.

Now, the question is how long the scaled process stays in 1. This is equivalent to the scaled time interval that the mouse spends in the 'neighborhood' of x. Thus, consider a random walk with i.i.d. steps $\Delta_1, \Delta_2, \dots$ distributed as Δ. We define

$$S_0 = 0, \quad S_n = \sum_{i=1}^{n} \Delta_i.$$

Then the mouse position after n steps is $x + S_n$, and since Δ is finite and independent of x, we have $(x + S_n)/x \to 1$ as $x \to \infty$ for all finite n. Thus, the scaled process equals one for all finite n. It follows that the time that the scaled process spends at state 1 should be equal to

$$W = \sum_{n=0}^{\infty} [\rho^{-(x+S_n)}/\rho^{-x}] E_n = \sum_{n=0}^{\infty} \rho^{-S_n} E_n, \qquad (14)$$

where $E_1, E_2, \dots$ are i.i.d. exponential random variables with mean 1, independent of the Δ_i's. One can prove that W is finite with probability 1. Then, formally, we can state the following proposition.

Proposition 2. *For any $t \geq 0$,*

$$\lim_{x\to\infty} \bar{M}_t(x) = \begin{cases} 1, & t < W; \\ -\infty, & t \geq W. \end{cases}$$

By Fubini's theorem and the identity $\mathbb{E}(\rho^{-\Delta}) = 1$, for $\mathbb{E}(W)$ we obtain:

$$\mathbb{E}(W) = \sum_{n=0}^{\infty} \mathbb{E}(\rho^{-S_n}) = \sum_{n=0}^{\infty} \left(\mathbb{E}(\rho^{-\Delta})\right)^n = \sum_{n=0}^{\infty} 1^n = +\infty.$$

Thus, W is extremely heavy-tailed: even its expectation does not exist. This is exactly what we observed in simulations, where the values of W varied considerably. In [8], we will analyze the properties of W in more detail, and we will present other scaling results, under various assumptions on the random walk (C_t).

Acknowledgments

Nelly Litvak gratefully acknowledges the support of The Netherlands Organisation for Scientific Research (NWO) under Meervoud grant no. 632.002.401.

8. REFERENCES

[1] S. Abiteboul, M. Preda, and G. Cobena, *Adaptive on-line page importance computation*, WWW' 03: Proceedings of the 12th international conference on World Wide Web, ACM Press New York, NY, USA, 2003, pp. 280–290.

[2] D. Aldous and J. Fill, *Reversible Markov chains and random walks on graphs*, Available at http://www.stat.berkeley.edu/users/aldous/RWG/book.html, 2002.

[3] K. Avrachenkov, N. Litvak, D. Nemirovsky, and N. Osipova, *Monte Carlo methods in PageRank computation: When one iteration is sufficient*, SIAM J. Numer. Anal. (2007).

[4] P. Berkhin, *A survey on PageRank computing*, Internet Math. **2** (2005), 73–120.

[5] S. Brin and L. Page, *The anatomy of a large-scale hypertextual web search engine*, Computer Networks and ISDN Systems **33** (1998), 107–117.

[6] S. Chakrabarti, *Dynamic personalized pagerank in entity-relation graphs*, WWW '07: Proceedings of the 16th international conference on World Wide Web, ACM Press New York, NY, USA, 2007, pp. 571–580.

[7] A. N. Langville and C. D. Meyer, *Deeper inside PageRank*, Internet Math. **1** (2003), 335–380.

[8] N. Litvak and Ph. Robert, *Cat-and mouse games and on-line algorithms for solving large Markov chains*, Tech. report, 2008, in preparation.

[9] Ph. Robert, *Stochastic networks and queues*, French ed., Applications of Mathematics (New York), vol. 52, Springer-Verlag, Berlin, 2003.

[10] W. J. Stewart, *Introduction to the numerical solution of Markov chains*, Princeton University Press, Princeton, NJ, 1994.

[11] P. Tetali, *Design of on-line algorithms using hitting times*, Proceedings of the 5th annual ACM-SIAM symposium on Discrete algorithms, Society for Industrial and Applied Mathematics Philadelphia, PA, USA, 1994, pp. 402–411.