Exact Results for the Barabási Model of Human Dynamics

Alexei Vázquez

Department of Physics and Center for Complex Networks Research, University of Notre Dame, Notre Dame, Indiana 46556, USA

Human activity patterns display a bursty dynamics with interevent times following a heavy tailed distribution. This behavior has been recently shown to be rooted in the fact that humans assign their active tasks different priorities, a process that can be modeled as a priority queueing system [A.-L. Barabási, Nature (London) **435**, 207 (2005)]. In this Letter we obtain exact results for the Barabási model with two tasks, calculating the priority and waiting time distribution of active tasks. We demonstrate that the model has a singular behavior in the extremal dynamics limit, when the highest priority task is selected first. We find that independently of the selection protocol, the average waiting time is smaller or equal to the number of active tasks, and discuss the asymptotic behavior of the waiting time distribution. These results have important implications for understanding complex systems with extremal dynamics.

Several problems of practical interest require us to understand human activity patterns [1–3]. Typical examples are the design of telephone systems or web servers, where it is critical to know how many users would use the service simultaneously. The traditional approach to characterize the timing of human activities is based in two assumptions: the execution of each task is independent from the others, and each task is executed at a constant rate [1–4]. A specific task, such as sending Email or making phone calls, is then modeled as a Poisson process [4], characterized by a homogeneous activity pattern. More precisely, the time interval between two consecutive executions of a task follows an exponential distribution. An increasing amount of empirical evidence is indicating, however, that human activity patterns are rather heterogeneous, with short periods of high activity separated by long periods of inactivity [1,5–11]. This heterogeneity is characterized by a heavy tail in the distribution of the time interval between two consecutive executions of the given task [5,10,11].

In practice the execution of one task is not independent of the others. Humans keep track of a list of active tasks from where they decide what to do next, the selection of one task implying the exclusion of the others. This picture lead Barabási to model the task management by a human as a queueing system, where the human plays role of the server [5]. Queueing systems [12] have already received some attention in the physics literature [13–16]. This interest is motivated by the observation of a nonequilibrium phase transition from a noncongested phase with a stationary number of active tasks to a congested phase where the number of active tasks grows in time. In the noncongested phase the mean waiting time before the execution of an active task is finite. When approaching the phase transition point the mean waiting time diverges, while it grows with time in the congested phase.

The Barabási model belongs, however, to a new class of queueing models with a fixed number of active tasks. In this case the behavior of interest comes from the task

selection protocol. In the extremal dynamics limit, when the highest priority task is selected first, numerical simulations and heuristic arguments show that most of the tasks are executed in one step, while the waiting time distribution of tasks waiting more than one step exhibits a heavy tail [5]. Yet, further research is required to obtain the scaling behavior in the vicinity of this singular point.

In this work we obtain exact results for the Barabási model, allowing us to prove previous conjectures based on heuristic arguments and numerical simulations, and creating a solid background for future research. We calculate the priority and waiting time distribution of those tasks remaining in the list for the case of two active tasks. We corroborate the observation of a singular behavior in the limit when the task with the highest priority is selected first, and derive the corresponding scaling behavior. We also obtain an upper bound for the average waiting time, which is independent of the selection protocol. Based on this result we discuss the asymptotic behaviors of the waiting time distribution. All the results presented here were checked by numerical simulations, providing a perfect match with the theoretical curves.

Barabási model.—The Barabási model is defined as follows. A human keeps track of a list with L active tasks that he/she must do. A priority $x \geq 0$ is assigned to each active task when it is added to the list, with a probability density function (PDF) $\rho(x)$. The list is started at $t = 0$ by adding L new tasks to it. At each discrete time step $t > 0$ the task in the list with the highest priority is selected with probability p, and with probability $1 - p$ a task is selected at random. The selected task is executed, removed from the list, and a new task is added. The control parameter p interpolates between the random selection protocol at $p = 0$ and the highest priority first selection protocol at $p = 1$.

The numerical simulations indicate that the case $L = 2$ already exhibits the relevant features of the model [5]. Furthermore, if we focus on a single task, such as sending Email, we can model the active tasks list as a list with two tasks, one corresponding to sending Email and the other to

doing something else. Within this scenario the waiting time coincides with the time between two consecutive executions of the corresponding task. Thus, the $L = 2$ case provides us with a minimal model to study the statistical properties of the time between the consecutive execution of specific tasks.

Consider the Barabási model with $L = 2$. The task that has been just selected and its priority reassigned will be called the new task, while the other task will be called the old task. Let $\rho(x)$ and $R(x) = \int_0^x dx \rho(x)$ be the priority PDF and distribution function of the new task, which are given. In turn, let $\rho_1(x, t)$ and $R_1(x, t) = \int_0^x dx \rho_1(x, t)$ be the priority PDF and distribution function of the old task at the tth step. At the $(t + 1)$th step, there are two tasks on the list, their priorities being distributed according to $R(x)$ and $R_1(x, t)$, respectively. After selecting one task the old task will have the distribution function

$$R_1(x, t + 1) = \int_0^x dx' \rho_1(x', t)q(x') + \int_0^x dx' \rho(x)q_1(x', t),$$ (1)

where

$$q(x) = p[1 - R(x)] + (1 - p)\frac{1}{2}$$ (2)

is the probability that the new task is selected given the old task has priority x, and

$$q_1(x, t) = p[1 - R_1(x, t)] + (1 - p)\frac{1}{2}$$ (3)

is the probability that the old task is selected given the new task has priority x. In the stationary state, $R_1(x, t + 1) = R_1(x, t)$, from (1) we obtain

$$R_1(x) = \frac{1 + p}{2p}\left[1 - \frac{1}{1 + \frac{2p}{1-p}R(x)}\right].$$ (4)

To analyze this result let us consider its limiting cases. When $p \to 0$ (4) results in

$$\lim_{p\to 0} R_1(x) = R(x).$$ (5)

Indeed, this limit corresponds to the random selection protocol and, therefore, the priority distribution of old tasks is equal to that of new tasks. On the other hand, when $p \to 1$ from (4) we obtain

$$\lim_{p\to 1} R_1(x) = \begin{cases} 0, & x = 0 \\ 1, & x > 0 \end{cases}.$$ (6)

That is, $\rho_1(x)$ is concentrated around $x = 0$. This result implies that in the limit $p \to 1$ the new task will always be selected for execution, resulting in a waiting time $\tau = 1$. We are going to return to this result after computing the waiting time distribution. The progression between these two limiting cases is illustrated in Fig. 1, where we plot $R_1(x)$ (4) as a function of x for a uniform distribution in

$0 \le x \le 1$. As p increases from zero, $R_1(x)$ deviates more from $R(x)$, resulting in a higher probability to obtain smaller values of x. When $p = 0.999$, $R_1(x)$ grows to a value of almost one in a very short x range close to $x = 0$, approaching the limit distribution (6).

Next we turn our attention to the waiting time distribution. Consider a task with priority x that has just been added to the queue. The selection of this task is independent from one step to the other. Therefore, the probability that it waits τ steps is given by the product of the probability that it is not selected in the first $\tau - 1$ steps and that it is selected in the τth step. The probability that it is not selected in the first step is $q_1(x)$, while the probability that it is not selected in the subsequent steps is $q(x)$. The integration over the new task's priorities results in

$$P_\tau = \begin{cases} \int_0^\infty dR(x)[1 - q_1(x)], & \tau = 1 \\ \int_0^\infty dR(x)q_1(x)[1 - q(x)]q(x)^{\tau-2}, & \tau > 1 \end{cases}.$$ (7)

Using (2)–(4) and integrating (7) we finally obtain

$$P_\tau = \begin{cases} 1 - \frac{1-p^2}{4p}\ln\frac{1+p}{1-p}, & \tau = 1 \\ \frac{1-p^2}{4p}\left[(\frac{1+p}{2})^{\tau-1} - (\frac{1-p}{2})^{\tau-1}\right]\frac{1}{\tau-1}, & \tau > 1 \end{cases}.$$ (8)

Note that P_τ is independent of the $\rho(x)$. Indeed, what matters for a task selection is its relative priority with respect to other tasks, resulting that all dependences with x in (2)–(4) and (7) appears via $R(x)$.

As before, let us consider the limiting cases. In the limit $p \to 0$ from (8) it follows that

$$\lim_{p\to 0} P_\tau = \left(\frac{1}{2}\right)^\tau$$ (9)

for $\tau \ge 1$. This limit corresponds with the random selection protocol, where a task is selected with probability $1/2$
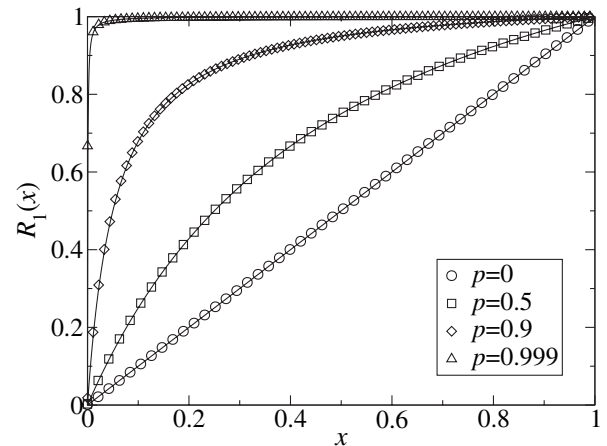

FIG. 1. Old task priority distribution for the case of a uniform new task priority distribution function, $\rho(x) = 1$ and $R(x) = x$ in $0 \le x \le 1$, as obtained from (4) (lines) and numerical simulations (points). The case $p = 0$ corresponds with the random selection protocol with $R_1(x) = R(x) = x$.

on each step. In the other limit, $p \to 1$, we obtain

$$\lim_{p \to 1} P_\tau = \begin{cases} 1 + \mathcal{O}(\frac{1-p}{2} \ln(1-p)), & \tau = 1 \\ \mathcal{O}(\frac{1-p}{2}) \frac{1}{\tau - 1}, & \tau > 1 \end{cases}. \quad (10)$$

In this case, almost all tasks have a waiting time $\tau = 1$, while the waiting time of tasks that are not selected in the first step follows a power law probability distribution. This picture is illustrated by a direct plot of P_τ in (8). In Fig. 2 we plot P_τ vs τ for a uniform distribution in $0 \le x \le 1$. For $p = 0.9$ the probability distribution P_τ is dominated by an exponential cutoff. This exponential cutoff can be derived from (8) by taking the limit $\tau \to \infty$ with p fixed, resulting in

$$P_\tau \sim \frac{1 - p^2}{4} \frac{1}{\tau} \exp\left(-\frac{\tau}{\tau_0}\right), \quad (11)$$

where

$$\tau_0 = \left(\ln\frac{2}{1+p}\right)^{-1}. \quad (12)$$

When $p \to 1$ we obtain that $\tau_0 \to \infty$ and, therefore, the exponential cutoff is shifted to higher τ values, while the power law behavior $P_\tau \sim 1/\tau$ becomes more evident. The P_τ vs τ curve systematically shifts, however, to lower values for $\tau > 1$, indicating that this power law applies to a vanishing task fraction [see Fig. 2 and (11)]. In turn, $P_1 \to 1$ when $p \to 1$, as it is corroborated by the direct plot of P_1 as a function of p (see inset of Fig. 2).

Another characteristic magnitude of a queueing system is the average waiting time of an active task before its execution. For $L = 2$ we can calculate the average waiting time directly from (8), obtaining
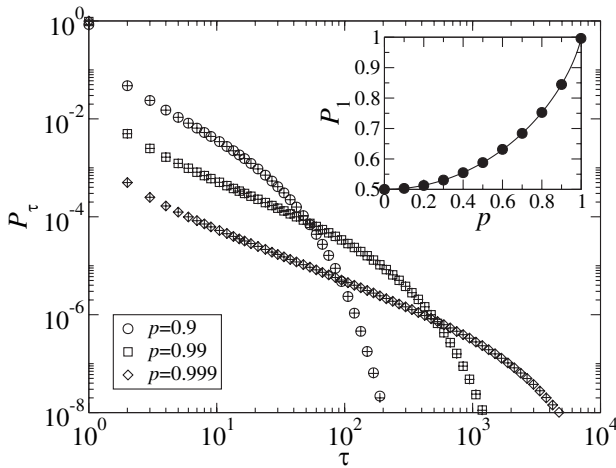

FIG. 2. Waiting time probability distribution for the case of a uniform new task priority distribution, $\rho(x) = 1$ and $R(x) = x$ in $0 \le x \le 1$, as obtained from (8) (pluses) and numerical simulations (open symbols). The inset shows the fraction of tasks with waiting time $\tau = 1$ as a function of p, as obtained from (8) (line) and numerical simulations (points).

$$\langle\tau\rangle = \begin{cases} 2, & 0 \le p < 1 \\ 1, & p = 1 \end{cases}. \quad (13)$$

This average is restricted to those tasks that are executed and, therefore, for $p = 1$ we are excluding the task that remains indefinitely in the queue. As we show next, we can extend this result for $L > 2$ using simple arguments.

On each step there are L task in the queue and one of them is executed. Therefore,

$$\sum_{i=1}^{t} \tau_i + \sum_{i=1}^{L-1} \tau_i' = Lt, \quad (14)$$

where τ_i is the waiting time of the task executed at the ith step and τ_i', $i = 1, \ldots, L - 1$, are the resident times of the $L - 1$ tasks that are still active at the t step. From (14) it follows that

$$\langle\tau\rangle = \lim_{t \to \infty} \frac{1}{t} \sum_{i=1}^{t} \tau_i = L - \lim_{t \to \infty} \frac{1}{t} \sum_{i=1}^{L-1} \tau_i'. \quad (15)$$

For $0 \le p < 1$ the numerical simulations indicate that all active tasks are expected to be executed [5]. Therefore $\langle\tau'\rangle \le \langle\tau\rangle$ and the last term in (15) vanishes when $t \to \infty$. In contrast, for $p = 1$ the numerical simulations [5] indicate that after some transient time the most recently added task is always executed, while $L - 1$ tasks remain indefinitely in the queue. In this case $\tau_i' \sim t$ when $t \to \infty$ and the last term in (15) is of the order of $L - 1$ when $t \to \infty$. Based on these arguments we conjecture that the average waiting time of executed tasks is given by

$$\langle\tau\rangle = \begin{cases} L, & 0 \le p < 1 \\ 1, & p = 1 \end{cases}. \quad (16)$$

This result was proved for $L = 2$ (13), and for $L > 2$ it is corroborated by the numerical simulations (see Fig. 3). It is
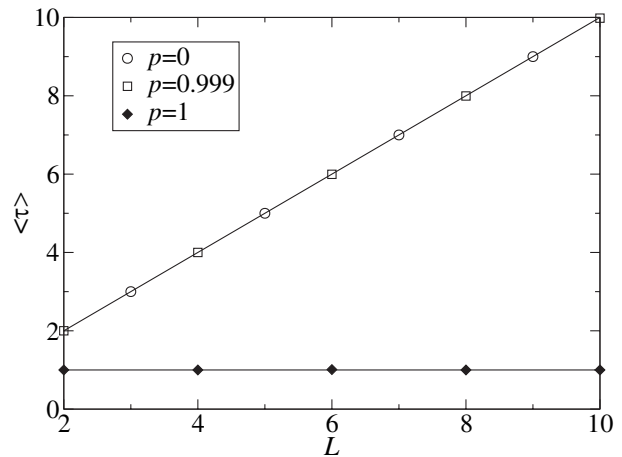

FIG. 3. Average waiting time of executed tasks vs the list size for the case of a uniform new task priority distribution function, $\rho(x) = 1$ and $R(x) = x$ in $0 \le x \le 1$, as obtained from (16) (lines) and numerical simulations (points).

important to note that the equality in (15) is independent of the selection protocol, allowing us to reach more general conclusions beyond the Barabási model. Using (15) we obtain

$$\langle \tau \rangle \le L. \tag{17}$$

From this constraint it follows that P_τ must decay faster than τ^{-2} when $\tau \to \infty$. Thus, when $\tau \to \infty$ either

$$P_\tau \sim a\tau^{-\alpha}, \tag{18}$$

where a is a constant and $\alpha > 2$, or

$$P_\tau = \tau^{-\alpha} f\left(\frac{\tau}{\tau_0}\right), \tag{19}$$

where $\tau_0 > 0$ and $f(x) = \mathcal{O}(bx^{\alpha-2})$ when $x \to \infty$, where b is a constant. For instance, for the Barabási model with $L = 2$ and $0 \le p < 1$, $\alpha = 1$ and $f(x)$ decays exponentially (11), in agreement with (19).

The empirical evidence is in favor of the second scenario with $\alpha = 1$ [5,10,11] or $\alpha = 3/2$ [11,17]. This observation is in agreement with our expectation of a natural cutoff. For instance, we might go on a trip and not check Email for several days, but sooner or later we are going to do it because we receive and transmit important information using this communication media. This cutoff time is expected to be more related to the necessity of performing a given task rather than to the interaction with other tasks. In this sense, the random selection of a task in the Barabási model could be interpreted as a task whose priority suddenly increases to the maximum value because the need to execute it after being on the queue for a long time. This indicates future directions of research, considering the case when the priority of old tasks may also change with time [18].

The singular behavior of the Barabási model is a consequence of the extremal dynamics rule: the task with the highest priority is selected first. Therefore, the conclusions obtained here are also relevant to other complex system evolving with extremal dynamics [19–21]. In this more general context the waiting time represents the lifetime before selection, an important quantity in evolution models [19,20] and optimization methods based in extremal dynamics [21]. Further research is required, however, to determine the influence of other factors such as correlations among neighbors, which are absent in the Barabási queueing model.

[1] H. R. Anderson, *Fixed Broadband Wireless System Design* (Wiley, New York, 2003).
[2] J. H. Greene, *Production and Inventory Control Handbook* (McGraw-Hill, New York, 1997).
[3] P. Reynolds, *Call Center Staffing* (The Call Center School Press, Lebanon, Tennessee, 2003).
[4] W. Feller, *An Introduction to Probability Theory and its Applications* (Wiley, New York, 1966), Vol. II.
[5] A.-L. Barabási, Nature (London) **435**, 207 (2005).
[6] J. Masoliver, M. Montero, and G. H. Weiss, Phys. Rev. E **67**, 021112 (2003).
[7] C. Dewes, A. Wichmann, and A. Feldman, in *Proceedings of the 2003 ACM SIGCOMM Conference on Internet Measurement (IMC-03)* (ACM, New York, 2003).
[8] S. D. Kleban and S. H. Clearwater, in *Proceedings of SC'03, November 15-21, Phoenix, Arizona, USA* (ACM, Phoenix, 2003).
[9] V. Paxson and S. Floyd, IEEE/ACM Transactions in Networking **3**, 226 (1995).
[10] A. Vázquez, Z. Dezső, J. Oliveira, K.-I. Goh, I. Kondor, and A.-L. Barabási, physics/0510117.
[11] Z. Dezső, E. Almaas, A. Lukács, and A.-L. B. B. Rácz, I. Szakadát, physics/0505087.
[12] D. Gross and C. M. Harris, *Queueing Theory* (John Wiley & Sons, New York, 1998).
[13] Y. Sugiyama and H. Yamada, Phys. Rev. E **55**, 7749 (1997).
[14] T. Ohira and R. Sawatari, Phys. Rev. E **58**, 193 (1998).
[15] A. Arenas, A. Diaz-Guilera, and R. Guimera, Phys. Rev. Lett. **86**, 3196 (2001).
[16] R. Solé and S. Valverde, Physica A (Amsterdam) **289**, 595 (2001).
[17] J. G. Oliveira and A.-L. Barabási, Nature (London) **437**, 1251 (2005).
[18] P. Blanchard and M.-O. Hongler, (to be published).
[19] P. Bak and K. Sneppen, Phys. Rev. Lett. **71**, 4083 (1993).
[20] F. Slanina and M. Kotrla, Phys. Rev. Lett. **83**, 5587 (1999).
[21] S. Boettcher and A. G. Percus, Phys. Rev. Lett. **86**, 5211 (2001).