Analytically solvable model of probabilistic network dynamics

M. A. M. de Aguiar,[1,2] Irving R. Epstein,[1,3] and Yaneer Bar-Yam[1]
[1]*New England Complex Systems Institute, Cambridge, Massachusetts 02138, USA*
[2]*Instituto de Física Gleb Wataghin, Universidade Estadual de Campinas, 13083-970 Campinas, São Paulo, Brazil*
[3]*Department of Chemistry, MS015, Brandeis University, Waltham, Massachusetts 02454, USA*
(Received 16 March 2005; revised manuscript received 29 August 2005; published 7 December 2005)

We present a simple model of network dynamics that can be solved analytically for fully connected networks. We obtain the dynamics of response of the system to perturbations. The analytical solution is an excellent approximation for random networks. A comparison with the scale-free network, though qualitatively similar, shows the effect of distinct topology.

DOI: 10.1103/PhysRevE.72.067102 PACS number(s): 89.75.Hc, 02.50.Ey, 84.35.+i

Recent advances in the understanding of complex social [1], biological [2], and technological [3] systems have revealed widespread if not universal properties of the topology of networks of association, interaction, and communication. These properties include small-world global connectivity [4], scale-free local connectivity distribution [5], and characteristic local motif structures [6]. Central to our understanding of complex systems [7] is characterizing their response to environmental stimuli. While much of the focus has been on robustness to random perturbation or directed attack [8], the effectiveness of response requires satisfying a wider range of conditions including, for example, sensitivity to particular stimuli [9]. Indeed, one of the main functions of biological and social systems is the detection of specific stimuli that require collective (large scale) response in seeking desirable resources (foraging) or responding to dangers ("fight or flight"). Understanding system function from network topology requires mapping the topology onto system response [10].

Insofar as the network characterizes the internal interactions, it must provide key information about response to external perturbations. In this paper we propose to use a model of influence similar to that studied in physics and the social sciences as the "voter model" and in biology as a model for propagation of mutations through a populations. We study the probabilistic discrete dynamics of this model and solve it analytically for fully connected networks, and for uniform node degree networks after aggregating states into state classes. We show that the model can reveal the differences in dynamic response of distinct topologies. The exact solution we obtain can serve as a basis for perturbative studies of diverse topologies.

We consider a general network with N nodes. To each node i is assigned an internal state σ_i that can take the values 0 or 1. At each time step the state of a node is updated according to the following rule: either the state does not change, which happens with probability p, or, with probability $(1-p)$, it copies the state of one of its neighbors. This process describes many systems in which mutual influence occurs, for example, high school students changing their dress style based upon that of others, or the propagation of a mutation through a species [11]. For a fully connected network and $p=1/N$ the model is equivalent to the Moran process [12], and for $p=0$ to the voter model [13]. Since the states of a node are abstract labels, the change of one node to adopt the state of another can be considered a general model of influence propagation, with each node state a label for its own relevant physical, biological, or social property. We will study the impact of a perturbation that changes the state of a subset of nodes as it propagates through the system.

The 2^N states of the network can be labeled by a string of zeroes and ones describing the internal state of each node in sequence $(\sigma_{N-1},\sigma_{N-2}\ldots,\sigma_1,\sigma_0)$. Alternatively, the states can be labeled by integers via $x=\sum_{j=0}^{N-1}\sigma_j 2^j$, with x varying between 0 and 2^N-1.

Let $P_t(x)$ be the probability of finding the network in the state x at time t and let the network evolve through asynchronous updates, where a single node is allowed to change at each time step. To find how this probability changes with time we define the auxiliary state $\tilde{x}_k$ which is equal to x at all nodes except at node k, which has the opposite internal state. The probability of finding the network in the state x at time $t+1$ can now be written as a sum of three terms: (a) the probability that the network was in state x at time t and that the selected node did not change plus (b) the probability that it was in the state x and the selected node copied the state of an identical neighbor plus (c) the probability that the network was in the state $\tilde{x}_k$ at time t and that the node k was selected and its state $\tilde{\sigma}_k=1-\sigma_k$ changed to σ_k:

$$P_{t+1}(x)=pP_t(x)+\frac{(1-p)}{N}\sum_k\{P_t(x)\mathrm{Prob}[\sigma_\mathrm{k}\to\sigma_\mathrm{k}]+P_t(\tilde{x}_k)\mathrm{Prob}[\tilde{\sigma}_\mathrm{k}\to\sigma_\mathrm{k}]\}.$$

The probability $\mathrm{Prob}[\sigma_\mathrm{k}\to\sigma_\mathrm{k}]$ is just the number of neighbors of node k in the state σ_k divided by the total number of neighbors (the degree) $d_k=\sum_{i=0}^{N-1}C_{ik}$, where C_{ik} is the connectivity (or adjacency) matrix. This can be written as

$$\frac{1}{d_k}\sum_{i=0}^{N-1}C_{ik}|1-\sigma_i-\sigma_k|.$$

The probability $\mathrm{Prob}[\tilde{\sigma}_\mathrm{k}\to\sigma_\mathrm{k}]$ is also given by this formula, since $\tilde{\sigma}_k=1-\sigma_k$ and $\tilde{\sigma}_i=\sigma_i$ for $i\neq k$. Using these relations, we obtain the following master equation for the network dynamics:

$$P_{t+1}(x) = pP_t(x) + \frac{(1-p)}{N}\sum_{k=0}^{N-1}\frac{1}{d_k}\sum_{i=0}^{N-1} C_{ik} \times |1-\sigma_i-\sigma_k|[P_t(x)+P_t(\tilde{x}_k)]. \quad (1)$$

Finding $P_t(x)$ for networks with arbitrary topologies can be very difficult. However, the problem can be completely solved for fully connected networks, where $d_k=N-1$. In this case, by symmetry, the states of the network can be labeled simply by counting the number of nodes in the internal state 1, given by $n(x)=\Sigma_i\sigma_i$. The probability of finding the network in the state labeled by n is related to $P(x)$ by

$$P(n(x)) = P(x)B(N,n), \quad (2)$$

where $B(N,n)=N!/[n!(N-n)!]$ is a binomial coefficient. The last two terms on the right-hand side of Eq. (1) can be simplified if we separate the sum over k into the cases $\sigma_k=1$ and $\sigma_k=0$ and multiply both sides of this equation by $B(N,n)$. After some algebra we obtain

$$P_{t+1}(n) = pP_t(n) + \frac{1-p}{N(N-1)}\{[n(n-1) + (N-n)(N-n-1)]P_t(n) + (N-n+1)(n-1)P_t(n-1) + (N-n-1)(n+1)P_t(n+1)\}. \quad (3)$$

For *uniform networks* where $d_k=d_0<N-1$ is the same for all nodes, we can use this equation if we classify together all the network states with the same $n(x)$. In this case the factor d_k in the denominator is replaced by d_0 in Eq. (1). Averaging with respect to the set of states labeled by n, the counting of the number of neighbors must be multiplied by $d_0/(N-1)$, so that d_0 cancels and $(N-1)$ remains in the denominator. Therefore Eq. (3) holds in this case as well. For random networks the degree of each node is nearly constant; tests reported below show that Eq. (3) is indeed a reasonable approximation for the dynamics in this case.

We now calculate the transition probabilities $P_t(n)$, defining a vector P_t of $N+1$ components. The master equation (3) can be written in matrix form as $P_{t+1}=UP_t$ where the *evolution matrix* U is tridiagonal. The propagation of an initial probability vector requires the calculation of powers of U. Alternatively, we can diagonalize U and use its eigenvectors as a basis. In what follows we shall calculate explicitly all the eigenvalues and eigenvectors of U, obtaining the complete solution of the dynamical problem.

The eigenvalues of U can be calculated for small matrices and extrapolated to matrices of arbitrary size. They are given by

$$\lambda_r = 1 - \frac{1-p}{N(N-1)}r(r-1)$$

with $r=0,1,\ldots,N$. The only degeneracy occurs for $\lambda_0=\lambda_1=1$. The other eigenvalues are all smaller than 1 and decrease towards $\lambda_N=p$.

Since U is not symmetric, its eigenvectors do not form an orthogonal set. Let $|a_r\rangle$ and $\langle b_r|$ be the right and left eigenvectors of U, with components a_{rm} and b_{rm}. Then

$$\sum_{r=0}^{N}\frac{1}{\Gamma_r}|a_r\rangle\langle b_r| = 1,$$

where $=\langle b_{r'}|a_r\rangle=\Gamma_r\delta_{rr'}$ and $\Gamma_r=\Sigma_m a_{rm}b_{rm}$.

An initial vector $|v(0)\rangle$ containing the information about the probability of the different states at time zero can be projected using this resolution of unity and easily evolved:

$$|v(t)\rangle = U^t|v(0)\rangle = \sum_{r=0}^{N}\frac{1}{\Gamma_r}\langle b_r|v(0)\rangle\lambda_r^t|a_r\rangle.$$

The transition probability between two network states with $n=M$ and $n=L$ after a time t can be calculated by taking the components of the initial vector as $v_m(0)=\delta_{M,m}$ and projecting the evolved state onto the state with components $\delta_{L,m}$:

$$P(L,t;M,0) = \sum_{r=0}^{N}\frac{1}{\Gamma_r}b_{rM}a_{rL}\lambda_r^t.$$

The coefficients a_{rm} follow a recursion relation that can be derived directly from the eigenvalue equation for U. For $r=0$ and $r=1$ the eigenvectors can be found immediately:

$$|a_0\rangle = \begin{pmatrix}1\\0\\0\\\vdots\\0\\1\end{pmatrix}, \quad |a_1\rangle = \begin{pmatrix}1\\0\\0\\\vdots\\0\\-1\end{pmatrix} \quad (4)$$

and

$$|b_0\rangle = (1 \quad 1 \quad 1 \quad \cdots \quad 1 \quad 1),$$

$$|b_1\rangle = (N \quad N-2 \quad N-4 \quad \cdots \quad -N+2 \quad -N). \quad (5)$$

In order to calculate the remaining eigenvectors, we define the auxiliary eigenvalues μ_r by

$$\mu_r = (1-\lambda_r)\frac{N(N-1)}{1-p} = r(r-1) \quad (6)$$

and the auxiliary coefficients

$$A_{rm} = m(N-m)a_{rm}, \quad m=1,2,\ldots,N-1. \quad (7)$$

The recursion relation for the A_{rm} can be written explicitly as

$$A_{rm+1} - 2A_{rm} + A_{rm-1} = -\frac{\mu_r}{N}\left(\frac{A_{rm}}{m} + \frac{A_{rm}}{N-m}\right). \quad (8)$$

A generating function is now defined as

$$f_r(x) = \sum_{m=1}^{N-1} A_{rm}x^m$$

(note that $A_{r0}=A_{rN}=0$). Multiplying Eq. (8) by x^m and summing over m we get, on the left-hand side,

$$\frac{f_r}{x}(1-x)^2 - A_{r1} + A_{rN-1}x^N.$$

In order to write down the right-hand side of Eq. (8) we define the auxiliary functions

$$g_r(x) = \sum_{m=1}^{N-1} \frac{A_{rm}}{m} x^m, \quad h_r(x) = \sum_{m=1}^{N-1} \frac{A_{rm}}{N-m} x^m. \tag{9}$$

It is easy to check that $dg_r/dx = f_r/x$ and $dh_r/dx = Nh_r/x - f_r/x$. After multiplying Eq. (8) by x^m and summing over m, we differentiate both sides with respect to x to obtain

$$\frac{d}{dx}\left[\frac{(1-x)^2}{x} f_r(x) - A_{rN-1}x^N\right] = -\frac{\mu_r}{x} h_r(x). \tag{10}$$

The solution of the differential equation for h_r can be obtained in terms of its Green function, satisfying $dG/dx - NG/x = \delta(x-y)$. In this case G is given by $(x/y)^N$ if $x > y$ and zero otherwise. Therefore

$$h_r(x) = x^N\left(\alpha - \int_{-\infty}^{x} \frac{f_r(y)}{y^{N+1}} dy\right). \tag{11}$$

Substituting Eq. (11) into Eq. (10), re-arranging the terms and differentiating once again with respect to x, we obtain

$$\frac{d}{dx}\left[\frac{1}{x^{N-1}}\frac{d}{dx}\left(\frac{(1-x)^2}{x} f_r(x)\right)\right] = \mu_r \frac{f_r(x)}{x^{N+1}}.$$

Finally, defining

$$F_r(x) = \frac{(1-x)^2}{x} f_r(x) \tag{12}$$

we obtain the differential equation

$$F_r'' - \frac{N-1}{x} F_r' - \frac{\mu_r}{x}\frac{F_r}{(1-x)^2} = 0. \tag{13}$$

Now let $\phi_r(x) = \Sigma a_{rm} x^m = (g_r + h_r)/N$. Differentiating with respect to x, using Eq. (11), dividing by x^{N-1} and differentiating again, we find

$$\phi_r'' - \frac{N-1}{x}\phi_r' - \frac{1}{x}\frac{F_r}{(1-x)^2} = 0.$$

Comparing with Eq. (13) we see that $\phi_r = F_r/\mu_r$. Therefore, except for normalization, *the generating function for the coefficients* a_{rm}, $\phi_r(x)$, *is equal to* $F_r(x)$.

For $r=0$ or $r=1$, $\mu_r = 0$ and the two independent solutions of Eq. (13) are $F_0(x) = 1 + x^N$ and $F_1(x) = 1 - x^N$, which correspond to the two degenerate eigenvectors $|a_0\rangle$ and $|a_1\rangle$. For $r=2$ and $r=3$ the solution can also be found explicitly; the general formula can then be extrapolated from these simple cases. We find

$$F_r(x) = (1-x)^{1-r}\left[1 + \sum_{p=1}^{r-1} d_{rp} x^p\right]. \tag{14}$$

with

$$d_{rp} = (-1)^p \frac{B(r-1,p)B(N+r-1,p)}{B(N-1,p)}.$$

Finally, the coefficients of the rth eigenvector are given by

$$a_{rm} = \frac{1}{m!}\left.\frac{d^m F_r(x)}{dx^m}\right|_{x=0}.$$

Working out the derivatives, we find the explicit formula valid for $r \geq 2$:

$$a_{rm} = \sum_{p=0}^{r-1} B(m-p+r-2, r-2) d_{rp} \tag{15}$$

for $m = 1, 2, \ldots, N-1$, with $a_{r0} = 1$ and $a_{rN} = (-1)^r$.

From the recursion relations we find that the coefficients of the left eigenvectors are given by

$$b_{rm} = a_{rm}[m(N-m)/N] \tag{16}$$

for $m = 1, 2, \ldots, N-1$, with $b_{r0} = b_{rN} = 0$. Finally, the normalization factors Γ_r can also be obtained explicitly:

$$\Gamma_r = \frac{r!\, B(N+r-1, r)}{(2r-1)B(N,r)} \tag{17}$$

for $r = 2, \ldots, N$ and $\Gamma_0 = 2$, $\Gamma_1 = 2N$.

We apply our calculations to consider a perturbation propagating through the system. Specifically, we seek the probability that a perturbation that initially affects M nodes will lead to a response by L nodes at a time t later. Starting with $n = M$ sites with the value 1 and all other sites 0 we look at the transition probability $P(L,t;M,0)$ of finding the system at a later time t with $n = L$. Except for the three special cases $L = 0$, M, N, this probability must start from zero, go through a maximum and then asymptotically go to zero. The case $L = M$ starts from one and decreases monotonically to zero, and the cases $L = 0$, N start from zero and increase

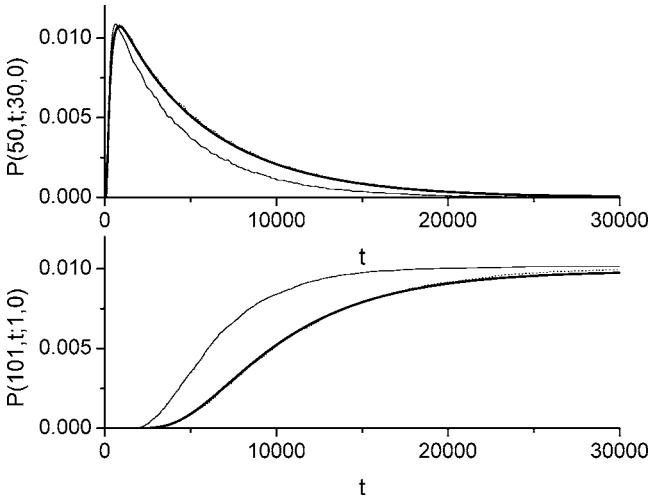

FIG. 1. Transition probabilities for a network with $N = 101$ and $p = 0.1$. The lines correspond to our theoretical calculation (thick) and to simulations for random (dotted) and scale-free (thin) networks, both with an average of six connections per node. The numerical probabilities were computed running the simulations 2×10^5 times. The dotted line is nearly indistinguishable from the thick line.

monotonically to their nonzero equilibrium values, as they are the absorbing steady states of the system. The initial and final state conditions are independent of topology; however, the intermediate dynamics can be affected by the topology and reveal key differences in system response to perturbations. $P(L,t;M,0)$ can be computed using Eqs. (4), (5), (15), (16), and (17):

$$P(L,t;M,0) = \left[\frac{N-M}{N} - \frac{3M(N-M)}{(N+1)N}\lambda_2^t\right]\delta_{L0} + \left[\frac{M}{N} - \frac{3M(N-M)}{(N+1)N}\lambda_2^t\right]\delta_{LN} + \frac{6M(N-M)}{N(N^2-1)}(1-\delta_{L0})(1-\delta_{LN})\lambda_2^t + \Sigma_{r=3}^{N}\frac{1}{\Gamma_r}b_{rM}a_{rL}\lambda_r^t.$$

All eigenvalues (except for λ_0 and λ_1) are smaller than 1, so the transition probability at long times is dominated by $\lambda_0=\lambda_1=1$ and by the second largest eigenvalue λ_2, whose contribution we obtain explicitly. For large networks $\lambda_2^t \approx \exp-[2(1-p)t/N^2]$, so that the characteristic duration of the transition process is $\tau=N^2/2(1-p)$, which increases with the square of the network size. The time to the maximum can be estimated as $t_{\max}=N|L-M|/2(1-p)$. The asymptotic probability of finding the network in states $n=0$ or $n=N$, having started from $n=M$, is $(N-M)/N$ and M/N, respectively.

Figure 1 shows a comparison between our analytical result and numerical calculations for a random and a scale-free network with $N=101$. The estimates $\tau \approx 6000$ and $t_{\max} \approx 1000$ [for Fig. 1(a)] are in agreement with the numerical calculation. The fluctuations (not shown) of the transition probabilities for random and scale-free networks are very similar, which is presumably due to self-averaging. The probability of initially selecting a highly connected node in a scale-free network is low, but as the perturbation spreads, such nodes will be encountered with increasing frequency, because their influence scales with their connectivity.

Summarizing, we have completely solved a simple dynamical process on fully connected networks. Our analytical result is an excellent approximation for the average behavior of sparsely connected random networks. For scale-free networks the qualitative behavior is the same, but there are deviations reflecting the significant topological differences, showing the utility of the dynamic model in describing the distinct responses of different topologies.

M.A.M.A. acknowledges financial support from CNPq and FAPESP. I.R.E. was supported in part by the David and Lucile Packard Foundation.

[1] S. Wasserman and K. Faust, *Social Network Analysis* (Cambridge University Press, Cambridge, UK, 1994); D. J. de S. Price, Science **149**, 510 (1965); M. E. J. Newman, Proc. Natl. Acad. Sci. U.S.A. **98**, 404 (2001); Phys. Rev. E **64**, 016131 (2001); **64**, 016132 (2001); D. Braha and Y. Bar-Yam, *ibid.* **69**, 016113 (2004).

[2] J. Cohen, F. Briand, and C. Newman, *Community Food Webs: Data and Theory* (Springer, Berlin, 1990); C. Koch and G. Laurent, Science **284**, 96 (1999); H. Jeong, B. Tombor, R. Albert, Z. N. Oltvai, and A.-L. Barabási, Nature (London) **407**, 651 (2000).

[3] M. Faloutsos, P. Faloutsos, and C. Faloutsos, Comput. Commun. Rev. **29**, 251 (1999); R. Albert, H. Jeong, and A.-L. Barabási, Nature (London) **401**, 130 (1999).

[4] D. J. Watts and S. H. Strogatz, Nature (London) **393**, 440 (1998).

[5] R. Albert and A.-L. Barabási, Rev. Mod. Phys. **74**, 47 (2002).

[6] R. Milo *et al.*, Science **298**, 824 (2002).

[7] Y. Bar-Yam, *Dynamics of Complex Systems* (Perseus, Cambridge, MA, 1997).

[8] R. Albert, H. Jeong, and A.-L. Barabási, Nature (London) **406**, 378 (2000); B. Shargel, H. Sayama, I. R. Epstein, and Y. Bar-Yam, Phys. Rev. Lett. **90**, 068701 (2003).

[9] Y. Bar-Yam and I. R. Epstein, Proc. Natl. Acad. Sci. U.S.A. **101**, 4341 (2004).

[10] D. H. Zanette, Phys. Rev. E **64**, 050901(R) (2001); M. Kuperman and G. Abramson, Phys. Rev. Lett. **86**, 2909 (2001); M. A. de Menezes and A.-L. Barabási, *ibid.* **92**, 028701 (2004).

[11] E. M. Rauch and Y. Bar-Yam, Nature (London) **431**, 449 (2004).

[12] P. A. P. Moran, Proc. Cambridge Philos. Soc. **54**, 60 (1958); G. A. Watterson, Ann. Math. Stat. **32**, 716 (1961); C. Cannings, Adv. Appl. Probab. **6**, 260 (1974); K. Gladstien, SIAM J. Appl. Math. **34**, 630 (1978).

[13] T. M. Liggett, *Interacting Particle Systems* (Springer, New York, 1985); D. Vilone and C. Castellano, Phys. Rev. E **69**, 016109 (2004); V. Sood and S. Redner, Phys. Rev. Lett. **94**, 178701 (2005).