Spreading and shortest paths in systems with sparse long-range connections

Cristian F. Moukarzel
Instituto de Física, Universidade Federal Fluminense, CEP 24210-340, Niterói, RJ, Brazil
(Received 6 August 1999)

Spreading according to simple rules (e.g., of fire or diseases) and shortest-path distances are studied on d-dimensional systems with a small density p per site of *long-range connections* (''small-world'' lattices). The volume $V(t)$ covered by the spreading quantity on an infinite system is exactly calculated in all dimensions as a function of time t. From this, the average shortest-path distance $\ell(r)$ can be calculated as a function of Euclidean distance r. It is found that $\ell(r)\sim r$ for $r<r_c=[2p\Gamma_d(d-1)!]^{-1/d}\log(2p\Gamma_d L^d)$ and $\ell(r)\sim r_c$ for $r>r_c$. The characteristic length r_c, which governs the behavior of shortest-path lengths, *diverges* logarithmically with L for all $p>0$. [S1063-651X(99)50312-7]

PACS number(s): 05.10.−a, 05.40.−a, 05.50.+q

Regular d-dimensional lattices with a small density p per site of long-ranged bonds (or ''small-world'' networks) [1] model the effect of weak unstructured (mean-field) interactions in a system where the dominant interactions have a regular d-dimensional structure, and have many applications in physics as well as in other sciences [1–13]. The recent observation [1] that a very small density of long-range connections has a drastic influence on shortest-path (or *chemical*) distance properties has triggered a lot of interest in these systems.

Most of the recent work in this field has been essentially numerical [1–13]. Here we report on an analytical calculation of the average shortest-path distance $\ell(r)$ between two points separated by an Euclidean distance r, as a function of p and d. We first study for this purpose the closely related problem of *spreading*, which is defined as follows. Consider (see, e.g., [1,11] and references therein) some influence (e.g., a forest fire, or an infectious disease) that spreads according to the following simple law: at each time step, the fire or disease propagates from a burnt (or infected) site to all unburnt (uninfected) sites connected to it by a link. Long-range connections, or *shortcuts* represent sparks that start new fires far away from the original front, or individuals who, when first infected, move to a random location amongst the uninfected population. For the dynamics of this simple problem, an important network property is the set of shortest-path distances $\{\ell_{ij}\}$, where ℓ_{ij} is defined as the minimum number of links one has to traverse between i and j. On isotropic d-dimensional lattices ℓ_{ij} is proportional to d_{ij}^E, the Euclidean distance between i and j. On regular lattices, both the number of sites within an Euclidean distance r from i, and the number of sites within r nearest-neighbor steps from i behave as r^d.

Consider now a randomly connected network, made up of L^d sites sitting on a regular d-dimensional lattice, but connected at random with an average coordination number C (i.e., a total of $L^dC/2$ bonds). The number of sites in a volume of radius r is still r^d, but we can visit $\sim C^k$ sites in k steps. Thus, all L^d sites can be visited in $\mathcal{O}(\log L^d$ steps, and therefore the typical shortest-path distance $\overline{\ell}$ is of order $\log L$, much *shorter* than the typical Euclidean distance $\overline{r}\sim L$.

''Small-world'' networks [1] are intermediate between the regular lattice, where $\overline{\ell}\sim L$, and the totally random network, where $\overline{\ell}\sim\log L$. They consist of a regular d-dimensional lattice with $N=L^d$ sites, on which pL^d *additional* long-range bonds have been connected between randomly chosen pairs of sites. The key finding of Watts and Strogatz [1] is that a vanishingly small density p of long-range bonds is enough to make shortest-path distances proportional to $\log L$ instead of L. If $L^dp\ll 1$, the system typically contains no shortcuts, and the average shortest-path distance $\overline{\ell}=1/N^2\Sigma_{\langle ij\rangle}[\ell_{ij}]_p$ scales as L. If, on the other hand, $L^dp\gg 1$, one finds $\overline{\ell}\sim\log L$ [1,4,7]. For any fixed density p of long-range bonds, a *persistence-size* $L^*(p)$ exists [4,5], above which shortest-path distances are only logarithmically increasing with L. This persistence size diverges as $p^{-1/d}$ [4,7,9] when $p\to 0$. The precise nature of the small-world transition at $p=0$ is still controversial [7,9,11,13].

In this work we calculate the volume $V(t)$ that is covered, on a small-world network, by a spreading quantity as a function of time [11] when the spreading law is the simple rule above. Using our knowledge of $V(t)$ an exact expression for the average shortest-path $\ell(r)$ is derived. Our result for $\ell(r)$ has been verified numerically [13].

Assume a disease spreads with constant radial velocity $v=1$ from an original infection site A, as shown in Fig. 1. Let $\rho=2p$ be the density of *shortcut-ends* on the system. We work on the continuum for simplicity, so that the infected volume $V(t)$ will initially grow as a sphere of radius t and surface $\Gamma_d t^{d-1}$. We call the sphere stemming from A ''primary sphere.''

Each time the primary sphere hits a shortcut end, which happens with probability $\rho\Gamma_d t^{d-1}$ per unit time, a new sphere (''secondary'') starts to grow from a random point in uninfected space (the other end of the shortcut). These in turn later give rise to further secondary spheres in the same fashion.

Following Newman and Watts (NW) [11], we notice that the total infected volume is the sum of the primary volume $\Gamma_d\int_0^t\tau^{d-1}d\tau$ plus a contribution $V(t-\tau)$ for each new sphere born at time τ. Thus, in the continuum the average total infected volume satisfies

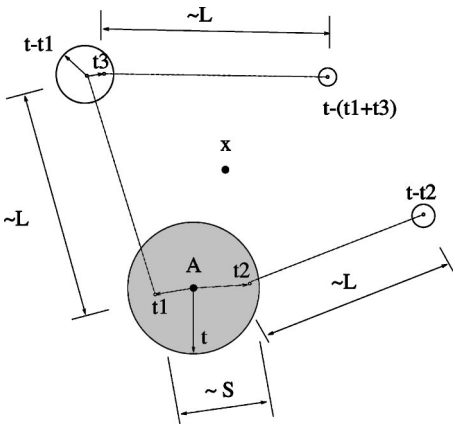

FIG. 1. Consider the spreading of fire or diseases on small-world systems. Assume that ρL^d points are drawn at random in d-dimensional space (open dots). The mean distance S between neighboring points is proportional to $\rho^{-1/d}$. Now connect pairs of points at random. The mean separation between ''mates'' will be of the order of L, the system size. Paired points represent long-range bonds (shaded lines), across which fire or diseases travel instantaneously. Now a disease starts to spread from A. Each time that the resulting sphere hits a shortcut-end, a secondary sphere will be born somewhere else. These in turn later give rise to other secondary spheres. The proliferation of secondary spheres produces an exponentially fast growth of the infected volume for times $t>S$.

$$V(t)=\Gamma_d\int_0^t \tau^{d-1}\{1+\rho V(t-\tau)\}d\tau, \tag{1}$$

This equation has been solved by NW in one dimension [11]. For general d we rewrite (1) in terms of rescaled variables $\tilde{V}=\rho V$ and $\tilde{t}=(\rho\Gamma_d(d-1)!)^{1/d}t$ as

$$\tilde{V}(\tilde{t})=\frac{1}{(d-1)!}\int_0^{\tilde{t}}(\tilde{t}-\tilde{\tau})^{d-1}\{1+\tilde{V}(\tilde{\tau})\}d\tilde{\tau}. \tag{2}$$

It is interesting to notice that $\tilde{V}$ is the total number of infected shortcut-ends, while $\tilde{t}^d/d!$ is the total number of shortcut-ends infected by the primary sphere. On an infinite system, the functional relation (2) that links these two variables has no parameters except for the space dimensionality d.

On a system of finite volume $\Gamma_d L^d/d$, an important parameter is the rescaled linear size $\tilde{L}=[\rho\Gamma_d(d-1)!]^{1/d}L$, or equivalently the total number N_s of shortcut-ends in the system: $N_s=\tilde{L}^d/d!$

Deriving Eq. (2) d times with respect to $\tilde{t}$ we obtain

$$\frac{\partial^d}{\partial\tilde{t}^d}V(\tilde{t})=1+V(\tilde{t}), \tag{3}$$

whose solution is

$$\tilde{V}(\tilde{t})=\sum_{k=1}^{\infty}\frac{\tilde{t}^{dk}}{(dk)!}. \tag{4}$$

Notice that Eq. (4) is a series expansion of $(e^{\tilde{t}}-1)$ with all powers not multiples of d removed. Thus, Eq. (4) can be written as a sum of d exponentials, each with a different d-root of 1 in its argument. In this way, powers which are not multiples of d cancel out,

$$\tilde{V}(\tilde{t})=\frac{1}{d}\sum_{n=0}^{d-1}\exp\{\mu_d^n\tilde{t}\}-1, \tag{5}$$

where $\mu_d=e^{i2\pi/d}$. Some specific examples are

$$\tilde{V}(\tilde{t})=e^{\tilde{t}}-1 \quad \text{in 1D},$$
$$\tilde{V}(\tilde{t})=\cosh\tilde{t}-1 \quad \text{in 2D},$$
$$\tilde{V}(\tilde{t})=\frac{e^{\tilde{t}}+e^{\mu_3\tilde{t}}+e^{\mu_3{}^2\tilde{t}}}{3}-1 \quad \text{in 3D}.$$

The one-dimensional solution is the same as that derived by NW [11].

A general property of Eq. (4) is that $\tilde{V}$ grows as $\tilde{t}^d/d!$ for $\tilde{t}<1$, and later exponentially as $e^{\tilde{t}}/d$. Thus, the characteristic timescale [11] for the spreading process is $t^*=(\rho\Gamma_d(d-1)!)^{-1/d}$.

Notice that Eq. (1), and thus also Eq. (4), only hold in an infinite system. On a finite system with $\tilde{L}^d\gg 1$, $\tilde{V}$ will saturate after a time $\tilde{t}_{sat}$ that can be estimated by equating $\tilde{V}\sim e^{\tilde{t}_{sat}}/d\sim\tilde{L}^d/d!$ and therefore

$$\tilde{t}_{sat}\sim\log[\tilde{L}^d/(d-1)!], \tag{6}$$

which can be rewritten as

$$t_{sat}\sim[\rho\Gamma_d(d-1)!]^{-1/d}\log(\rho\Gamma_d L). \tag{7}$$

If on the other hand $\tilde{L}^d\ll 1$, the spreading stops at $\tilde{t}_{sat}=\tilde{L}$, before reaching the exponential growth regime. Thus, for a finite system with $\tilde{L}^d\gg 1$ one has

$$\tilde{V}(\tilde{t})d\sim\begin{cases}\tilde{t}^d/d! & \text{for } \tilde{t}\ll 1\\ e^{\tilde{t}}/d & \text{for } 1\ll\tilde{t}<\tilde{t}_{sat}\sim\log\left(\dfrac{\tilde{L}^d}{(d-1)!}\right)\\ \tilde{L}^d/d! & \text{for } \tilde{t}>\tilde{t}_{sat}.\end{cases} \tag{8}$$

Assume now that $\tilde{L}^d\gg 1$. Because of the exponentially fast spreading process, the fraction of the total volume covered by the disease is negligible for $\tilde{t}<\tilde{t}_{sat}$ and saturates to 1 abruptly at $t=t_{sat}$. Therefore, on a large system most of the points become infected essentially at the same time $\tilde{t}_{sat}$.

Now let us see how to calculate the average shortest-path distance $\ell(r)$ as a function of the Euclidean separation r between two points. Since we assumed that the disease spreads with unit velocity, it is clear that the time t at which a point x becomes first infected is exactly the shortest-path distance $\ell(A,x)$ from A to x. By definition, no part of the finite system remains uninfected after $t=t_{sat}$, so we conclude that no shortest-path distance can be larger than t_{sat} on a finite system. Assuming that $\ell(r)$ cannot decrease with increasing r, we conclude that

$$\ell(r) = t_{sat} \quad \text{for} \quad r \geq t_{sat}. \tag{9}$$

In order to calculate $\ell(r)$ for $r < t_{sat}$, let us write $V(t) = V_1(t) + V_2(t)$, where V_1 is the primary volume and V_2 the volume infected by secondary spheres. Let $p_2(t)$ be the probability to become infected by the secondary infection exactly at time $t \leq t_{sat}$. Consequently $I_2(t) = \int_0^t p_2(\tau) d\tau$ is the probability for a point to become infected at time t or earlier. Assuming that $p_2(t)$ is known, it is easy to calculate the average shortest-path distance $\ell(r)$ as a function of Euclidean distance r, according to the following. If an individual at x becomes infected by a secondary sphere at time $\tau < d^E(A,x)$, its shortest-path distance $\ell(A,x)$ to A is τ. Otherwise if x is still uninfected at time $t = d^E(A,x)$ {which happens with probability $1 - I_2[d^E(A,x)]$}, then $\ell(A,x) = d^E(A,x)$, since at that time the primary sphere hits x with probability 1. Therefore, the average shortest-path satisfies

$$\ell(r) = \int_0^r t p_2(t) dt + r\{1 - I_2(r)\} \tag{10}$$

$$= r - \int_0^r I_2(t) dt. \tag{11}$$

The fact that the secondary volume V_2 is randomly distributed in space makes this problem relatively simple. The probability $I_2(t)$ for a point to be infected by the secondary version of the disease at time t or earlier is simply $I_2 = V_2(t)/(1/d\Gamma_d L^d)$, i.e., the fraction of the total volume which is covered by the secondary infection. Thus,

$$\ell(r) = r - \frac{d}{\Gamma_d L^d} \int_0^r V_2(t) dt. \tag{12}$$

If there are no shortcuts on the system, V_2 is zero at all times and thus $\ell(r) = r$ as expected. But it is also clear from this expression that $\ell(r) = r$ when $L \to \infty$, for all *finite* r, i.e., in the thermodynamic limit the shortest-pat distance $\ell(r)$ coincides with the Euclidean distance r for all *finite* r, no matter what ρ is.

On a finite system with $\tilde{L}^d \gg 1$, $V_2(t)/L^d$ is negligible for all $t < t_{sat}$, as we have already noticed. Therefore, $\ell(r) = r$ if $r < t_{sat}$. Combining this with Eq. (9) we have

$$\ell(r) \approx \begin{cases} r & \text{for } r < r_c = [\rho\Gamma_d(d-1)!]^{-1/d} \log(\rho\Gamma_d L^d) \\ r_c & \text{for } r \geq r_c. \end{cases} \tag{13}$$

Detailed knowledge of $\ell(r)$ for $r \approx r_c$ would only be possible if the finite-size effects that we ignored in Eq. (2) were exactly known, but the interesting remark is that the lack of this knowledge has little or no importance for $r \neq r_c$.

We thus see that on a *finite* system, a characteristic length $r_c = [\rho\Gamma_d(d-1)!]^{-1/d} \log(\rho\Gamma_d L^d)$ exists, that governs the behavior of average shortest-path distances as a function of Euclidean separation. This characteristic length diverges when $L \to \infty$, for any $\rho > 0$. The typical separation $s = \rho^{-1/d}$ between shortcut ends [11], which is size-independent, is *not relevant* for $\ell(r)$. The validity of Eq. (13) has been verified numerically in one dimension recently [13].

It is interesting to notice that the rescaled shortest-path distance $\tilde{\ell} = \ell/r_c$ is a simple function of the rescaled Euclidean distance $\tilde{r} = r/r_c$,

$$\tilde{\ell}(\tilde{r}) = \begin{cases} \tilde{r} & \text{for } \tilde{r} < 1 \\ 1 & \text{for } \tilde{r} \geq 1. \end{cases} \tag{14}$$

Using Eq. (13) we can calculate $\overline{\ell}(\rho, L)$, the (global) average shortest-path length [1,4,7], when $\tilde{L}^d \gg 1$. One has

$$\overline{\ell}(\rho, L) = \frac{d}{L^d} \int_0^L \ell(r) r^{d-1} dr$$

$$= \frac{d}{L^d} \int_0^{r_c} r^d dr + \frac{r_c d}{L^d} \int_{r_c}^L r^{d-1} dr \tag{15}$$

$$= r_c \left[1 - \frac{1}{d+1} \left(\frac{r_c}{L} \right)^d \right]. \tag{16}$$

So that the ''order parameter'' $\mathcal{L} = \overline{\ell}/L$ [9,13] reads

$$\mathcal{L} = z \left(1 - \frac{z^d}{d+1} \right), \tag{17}$$

where $z = r_c/L$.

When $\rho \to 0$ faster than $L^{1/d}$ (so that $\tilde{L}^d \ll 1$), formula (15) holds with $r_c \to L$, and thus $\mathcal{L} \to d/(d+1)$ as expected. On the other hand if $\rho > 0$ one has that $r_c \ll L$ when $L \to \infty$, and thus $\mathcal{L} \to 0$ in this limit. Therefore, $\mathcal{L}$ undergoes a discontinuity at $\rho = 0$ in the $L \to \infty$ limit [9,13].

Notice that $r_c/L = \log(\rho\Gamma_d L^d)/L[\rho\Gamma_d(d-1)!]^{1/d} \sim \log(L/s)/(L/s)$, where $s \sim \rho^{-1/d}$ [11] is the mean separation between shortcut-ends. Thus, $\mathcal{L}$ can be written as a function of L/s only. Therefore, if we measure $\mathcal{L}$ on systems with several values of L and ρ and plot the data versus L/s, we would find that they *collapse* [4,5,7,11]. Because of this behavior it has been suggested [7,11] that the transition at $\rho = 0$ is a *critical point* with a size-independent characteristic length $\xi \sim s \sim \rho^{-1/d}$. Our results here and in previous work [9,13] suggest that this is not the case. According to our calculation, the only characteristic length in regard to shortest-paths is r_c, and it diverges with system size L. On the other hand, the scaling behavior of $\mathcal{L}(L,p)$ is entirely compatible with a normal *first-order* transition [9,13].

We have thus shown that, on a finite system with $L^d p \gg 1$, two widely separated timescales for spreading can be identified. The first one $t^* = [2p\Gamma_d(d-1)!]^{-1/d}$ determines the crossover from normal (i.e., proportional to t^d) to exponential spreading. A much larger timescale t_{sat} given by Eq. (7) determines the saturation of the spreading process. This second timescale coincides with the lengthscale r_c at which the behavior of shortest path lengths $\ell(r)$ saturates, as given by Eq. (13).

It is clear from our calculation that r_c diverges with L because the locations of the secondary spheres are uncorrelated with the location of the primary infection. In other words, because on a system of size L, the typical separation between both ends of a shortcut scales as L. A different situation would certainly arise if shortcuts had a length-

dependent distribution. For example one can connect each site i, with probability p, to a single other site j, chosen with probability $r_{ij}^{-\alpha}$, where α is a free parameter. For $\alpha \to 0$, this model is the same as discussed here, while for α large one would only have short-range connections and thus there would be no short-distance regime, even for $p=1$. We are presently studying this more general model [14].

Note added. For recent related analytical work, see [15].

I wish to acknowledge useful discussions with M. Argollo de Menezes. This work was supported by FAPERJ.

[1] D. J. Watts and S. H. Strogatz, Nature (London) **393**, 440 (1998).
[2] H. Herzel, Fractals **6**, 4 (1998).
[3] S. A. Pandit and R. E. Amritkar, Phys. Rev. E **60**, 1119 (1999).
[4] M. Barthélémy and L. A. N. Amaral, Phys. Rev. Lett. **82**, 3180 (1999).
[5] A. Barrat, e-print cond-mat/9903323.
[6] R. Monasson, e-print cond-mat/9903347.
[7] M. E. J. Newman and D. J. Watts, Phys. Lett. A (to be published).
[8] A. Barrat and M. Weigt, Eur. Phys. J. B (to be published).
[9] M. Argollo de Menezes, C. Moukarzel, and T. J. P. Penna, e-print cond-mat/9903426.
[10] R. Kasturirangan, e-print cond-mat/9904055.
[11] M. E. J. Newman and D. J. Watts, Phys. Rev. E (to be published).
[12] R. V. Kulkarni, E. Almaas, and D. Stroud, e-print cond-mat/9905066.
[13] C. Moukarzel and M. Argollo de Menezes, e-print cond-mat/9905131.
[14] M. Argollo de Menezes, C. Moukarzel, and T. J. P. Penna (unpublished).
[15] M. E. J. Newman, C. Moore, and D. J. Watts, e-print cond-mat/9909165; R. V. Kulkarni, E. Almaas, and D. Stroud, e-print cond-mat/9908216; S. N. Dorogovtsev and J. F. F. Mendes, e-print cond-mat/9907445.