Deliverable D2.2.4

Statistical properties of directed ad-hoc networks arising from random distributed transmitting powers of agents
Start date of the project: January 2004

Duration: 48 months

Project Coordinator: Prof. Dr. math. Friedhelm Meyer auf der Heide
Heinz Nixdorf Institute, University of Paderborn, Germany

Due date of deliverable: December 2006
Actual submission date: January 2007
Dissemination level: PU – public

Work Package 2.2: Dynamics and stability of networks

Participants: Università Degli Studi di Roma "La Sapienza" (UDRLS), Italy

Authors of deliverable: Guido Caldarelli (Guido.Caldarelli@roma1.infn.it)
Andrea Capocci (Andrea.Capocci@centrofermi.it)
Vito D.P. Servedio (Vito.Servedio@roma1.infn.it)
Abstract

In this deliverable we report the analysis made on the problem of networking of mobile agents with random emitting power. This report consists essentially of two contributions (DELIS-TR-0410 and DELIS-TR-0484). The former is a very simplified statistical model that wants to reproduce an elementary process of dynamics in a scale-free network. The main result is that scale invariance is a natural consequence of a suitably chosen extremal dynamics. In the second contribution we consider a more realistic multi-hop ad-hoc network (defined as an undirected network by considering two nodes as adjacent whenever their effective transmitting areas overlap). We show with simple probabilistic reasonings that a so constructed network is minimally connected if the effective transmitting area scales as $\frac{\log N - \log \log N}{N}$ with the number $N$ of nodes, irrespective of the area shape.

1 Introduction

The last two decades have witnessed a vertiginous increase of wireless devices in correspondence to an increased overall human mobility. An increasing activity has been devoted to the study of the theoretical foundation of these systems. This activity spans from hardware design to the topology of the structure and the software to run these systems. Here we present some contributions related to the effect of topologies in such systems. In our studies the effect of randomness is crucial and it is described by means of statistical models. One of the results we would like to obtain is a coherent description of the interplay between topology and dynamics in the network. This would allow to obtain a self-organised system able to cope with the various real situations. This is a very difficult problem and very little progress has been made on the self-organisation of these systems even in the very simplified version of toy statistical models.

In this area it is known that dynamical processes defined on graphs display a strong dependence on the topology [1, 2]. In the case of ad-hoc mobile network, there is another ingredient to take into account that is given by the difference of the various agents as for example in their emitting power. As for other fields where there is growing empirical evidence [3, 4] that many networks are in turn shaped by some variable associated to each vertex, we can capture this aspect through a ‘fitness’ or ‘hidden–variable’ model [5, 6]. Until now, these two facets of the same problem have been treated as separate, by considering on one hand dynamical processes on static networks [2], and on the other hand network formation mechanisms driven by quenched variables [5, 6, 7, 8, 9]. This may be perhaps justified for short time scales. However, in the long–term evolution it is crucial to understand the effects that these mechanisms have on each other, without ad hoc specifications of any fixed structure either in the topology or in the dynamical variables. Remarkably, the interplay of dynamics and topology can drive the network to a self–organized state that cannot be inferred by studying the two evolutionary processes as decoupled. Our main interest here is the complete description of a self–organized process where the dynamical variable acts also as the ‘hidden variable’ shaping network topology explicitly, as in the fitness model. In this model (see Fig.1) the probability $p$ to draw an edge, differently than the Erdős-Rényi model of Random Graphs [10] it depends upon the couple $i, j$ of vertices chosen and it is a function $f(x_i, x_j)$ of the fitnesses involved. Due to the

![Figure 1](image.png)

Figure 1: Every vertex $i$ is characterised by a fitness $x_i$. A graph is generated by assigning a function $f(x_i, x_j)$ with which we draw a link between a couple of vertices.

1
increased complexity of the problem, in this paper we choose the simplest possible dynamical rule for the hidden variable. We focus on the extremal dynamics defined in the Bak–Sneppen (BS) model [11], a traditional model of self–organized criticality (SOC) inspired by biological evolution.

2 Self–Organisation of Topology

Consider a graph with $N$ vertices, each regarded as a species having a fitness value $x_i$ initially drawn from a uniform distribution between 0 and 1. At each timestep the species with lowest fitness and all its neighbours undergo a mutation, and their fitness values are drawn anew from the same uniform distribution. We couple this rule with the fitness model assumption [5] that the network is formed by drawing a link between any two vertices $i$ and $j$ with fitness–dependent probability $f(x_i, x_j)$. When the fitness $x_i$ of a species $i$ is updated to $x_i'$, the links from $i$ to all the other vertices $j$ are drawn anew with probability $f(x_i', x_j)$. The ensemble spanned by this process is equivalent to the one obtained if the whole network (including the links between vertices whose fitness is unchanged) is updated at each timestep. Therefore the results that we obtain here hold also if heterogeneous updating timescales are introduced across the network, making the model very general in this respect. The master equation for the time–dependent fitness distribution $\rho(x, t)$ can be written as

$$\frac{\partial \rho(x, t)}{\partial t} = r^{in}(x, t) - r^{out}(x, t),$$

where $r^{in}(x, t)$ and $r^{out}(x, t)$ represent the average fractions of vertices with fitness $x$ which are added and removed at time $t$ respectively. If we look for the long–term stationary distribution $\rho(x)$ we must require that $\frac{\partial \rho(x, t)}{\partial t} = 0$, or $r^{in}(x) = r^{out}(x)$. These quantities are conditional on the value of the minimum fitness $m \equiv x_{\text{min}}$. When the minimum value is $m$ the number of updated values is

$$1 + k(m) = 1 + N \int_m^1 \rho(x) f(x, m) \, dx,$$

where $k(m)$ is the degree of the minimum–fitness vertex, and each of them is uniformly drawn between 0 and 1. Therefore $r^{in}(x|m) = [1 + k(m)]/N$. By contrast, $r^{out}(x|m)$ depends on $x$ according to the replacement rule. It is convenient to define $\tau$ as the value below which one fitness value (the minimum one) falls in the steady state, that is $N \int_0^\tau \rho(x) \, dx \equiv 1$. As happens for the random–neighbour variant of the BS model [12], the annealed nature of the links is such that, since $m$ is continuously replaced by uniformly distributed values, it is uniformly distributed

![Figure 2: A process of graph construction. Above: vertex $i$ has the minimum fitness and it is substituted by another with different fitness. Also its neighbours are substituted with others with newly extracted fitnesses. Below: with the new series of fitnesses a new graph is established](image)
between 0 and $\tau$. Thus the distribution for $m$ in the stationary state is $q(m) = \Theta(\tau - m)/\tau$. This also means that

$$\rho(x) = \begin{cases} 
  q(x)/N = (\tau N)^{-1} & x < \tau \\
  \rho_1(x) & x > \tau 
\end{cases}$$

(1)

with $\rho_1(x)$ to be determined. Now, $r^{\text{out}}(x|m) = 1/N$ when $x = m$ since the minimum is always replaced, while $r^{\text{out}}(x|m) = \rho(x)f(x,m)$ when $x > m$ since the probability that a vertex with fitness $x > m$ is updated is equal to the probability $f(x,m)$ that it is connected to the vertex with fitness $m$, no matter how old the link is. This is the key consideration that makes our model exactly solvable. Turning to the unconditional quantities $r^{\text{in}}(x) = \int q(m)r^{\text{in}}(x|m)dm$, $r^{\text{out}}(x) = \int q(m)r^{\text{out}}(x|m)dm$ and equating them yields

$$1 + \langle k_{\text{min}} \rangle = \begin{cases} 
  (\tau N)^{-1} & x < \tau \\
  \rho_1(x)\tau^{-1} \int_0^\tau f(x,m)dm & x > \tau 
\end{cases}$$

(2)

where $\langle k_{\text{min}} \rangle \equiv \int q(m)k(m)dm = \tau^{-1} \int_0^\tau k(m)dm$ is the average degree of the vertex with minimum fitness. Note that both terms appearing on the r.h.s. must be equal to the one on the l.h.s., which also means that they must be equal to each other. This allows to determine $\rho_1(x)$ and put it back into eq.(1), which finally yields

$$\rho(x) = \begin{cases} 
  (\tau N)^{-1} & x < \tau \\
  \frac{1}{N\int_0^\tau f(x,m)dm} & x > \tau 
\end{cases}$$

(3)

The value of $\tau$ can be determined through the normalization condition $\int_0^1 \rho(x)dx = 1$ which reads

$$\int_{\tau}^1 \frac{dx}{\int_0^\tau f(x,m)dm} = N - 1$$

(4)

This characterizes the stationary state completely. Once $\rho(x)$ is known, all the expected topological quantities can be determined as in the static fitness model [5, 8, 9]. Note that if $\tau$ is nonzero the model preserves the step–like behaviour displayed on static networks [11, 12, 13, 14], but here we find the novel feature that $\rho(x)$ is in general not uniform for $x > \tau$. Therefore the system spontaneously evolves from a random structure to a complex network with nontrivial dynamical and topological properties.

### 2.1 Analytical Results and Simulations

The above analytical solution holds for any form of $f(x, y)$. We now consider the choice of this function. First note that the null choice is $f(x, y) = p$, the network being a random graph. It is nonetheless an instructive simple case, and we briefly discuss it. This choice is asymptotically equivalent to the random neighbour variant [12] of the BS model, the average degree of each vertex being $d = pN$. Our analytical results read

$$\rho(x) = \begin{cases} 
  (\tau N)^{-1} & x < \tau \\
  (p\tau N)^{-1} & x > \tau 
\end{cases}$$

(5)

and, depending on how $p$ scales with $N$, eq.(4) implies

$$\tau = \frac{1}{1 + pN} \rightarrow \begin{cases} 
  1 & pN \rightarrow 0 \\
  (1 + d)^{-1} & pN = d \\
  0 & pN \rightarrow \infty 
\end{cases}$$

(6)
We note that these three dynamical regimes are tightly related to an underlying topological phase transition. As \( p \) decreases, the whole system splits up into a number of smaller subsets or clusters. Such process displays a critical behaviour near the threshold \( p_c \approx 1/N \) [1, 2]. Below \( p_c \), each node is isolated or linked to a small number of peers. Above \( p_c \), a large giant component emerges including a number of nodes of order \( O(N) \), whose fraction tends to 1 as \( p \to 1 \). This explains the dynamical regimes in eq.(6). If \( pN \to \infty \) (dense regime), then \( \tau \to 0 \) and \( \rho(x) \) is uniform between 0 and 1 as in the initial state, since an infinite number \( (k_{\text{min}}) = pN \) of fitnesses is continuously updated as on a complete graph. If \( pN = d \) with finite \( d > 1 \) (sparse regime), then \( \tau \) remains finite as \( N \to \infty \), and this is the case considered in ref.[12] that we recover correctly. Finally, if \( p \) falls faster than \( 1/N \) the graph is below the percolation threshold (subcritical regime); the updates cannot propagate and \( \tau \to 1 \), as for \( N \) isolated vertices. Therefore the dynamical transition is rooted in an underlying topological phase transition. This previously unrecognized property is fundamental and, as we show below, is also general.

A nontrivial form of \( f(x, y) \) must be chosen carefully. On static or fitness–independent networks \( x_i \) is usually interpreted as the fitness barrier against further mutation, and the links are interpreted as feeding relations [11]. However, once the topology depends on \( x \), the coupling we have introduced requires consistent interpretations of \( x \) and of the links. Also, the form of \( f(x, y) \) must be consistent with the feature that the updates of \( x \) propagate through the network determined by it. This very instructive aspect must characterize any model with coupled topology and dynamics, and reduces significantly the arbitrariness introduced in the static case. Here we suggest that the simplest self–consistent interpretation is the following. Since there is no external world in the model, the environment experienced by a species is simply the set of its ecological interactions. Now let \( x_i \) represent the fitness (rather than the barrier) of \( i \), and let a link between two species mean ‘being fit to coexist with each other’ (i.e. it represents an undirected, non–feeding interaction beneficial to both). The more a species is connected to other species, the more it is fit to the environment. This picture is self–consistent provided that the larger \( x \) and \( y \), the larger \( f(x, y) \). Following the results of refs.[15, 16], the simplest unbiased [16] choice for such a function is

\[
 f(x, y) = \frac{zxy}{1 + zxy} \tag{7}
\]

where \( z \) is a positive parameter controlling the number of links. With the above choice, \( \rho(x) \) can be computed analytically through eq.(3). However we write it in approximated form, exact when \( N \to \infty \), in order to solve also more complicated integrals involving it. We assume \( \langle f(x, m) \rangle \approx f(x, \langle m \rangle) \), that is \( \tau^{-1} \int_0^\tau f(x, m)dm \approx (zx\tau/2)/(1 + zx\tau/2) \). Then eq.(3) yields

\[
 \rho(x) = \begin{cases} (\tau N)^{-1} & x < \tau \\
 (\tau N)^{-1} + 2/(zN\tau^2x) & x > \tau 
\end{cases} \tag{8}
\]

and eq.(4) reads \( 1/\tau - 2z^{-1}\tau^{-2}\log \tau = N \), solved by

\[
 \tau = \sqrt{\frac{\phi(zN)}{zN}} \rightarrow \begin{cases} 1 & zN \to 0 \\
 \frac{\phi(d)/d}{\sqrt{d}} & zN = d \\
 0 & zN \to \infty \end{cases} \tag{9}
\]

for large \( N \), where \( \phi(x) \) is the ProductLog function defined as the solution of \( \phi e^\phi = x \). As for random graphs, we find a marked transition as the scaling of \( z \) changes from \( N^{-1} \) to more rapidly decaying. This suggests an analogous underlying percolation transition. As we show below, this is indeed the case. We can therefore still refer to the subcritical, sparse and dense regimes.

The main panel of Fig.3 shows the cumulative density function (CDF) of the fitness \( \rho_>(x) \equiv \int_x^\infty \rho(x')dx' \), while the inset shows a plot of \( \tau(zN) \). The theoretical results are in excellent agreement with large numerical simulations. As predicted by eq.(8), \( \rho(x) \) is the superposition of a uniform...
distribution and a power-law with exponent $-1$. For $z \ll 1$ we have $f(x, y) \approx zxy$ and $\rho(x) \propto x^{-1}$ for $x > \tau$. This purely power-law behaviour, that becomes exact in the sparse regime $z = d/N$ for $N \to \infty$, results in a logarithmic CDF looking like a straight line in log-linear axes. Note that, despite the value of the exponent, the presence of a nonzero lower threshold ensures that $\rho(x)$ is normalizable. This mechanism may provide a natural explanation for the onset of Pareto distributions with a finite minimum value in real systems. By contrast, for large $z$ the uniform part is nonvanishing and $\rho(x)$ deviates from the purely power-law behaviour. The decay of $\rho(x)$ for $x > \tau$ is a completely novel outcome of the extremal dynamics due to the feedback with the topology: now the fittest species at a given time is also the most likely to be connected to the least fit species and to mutate at the following timestep. Being more connected also means being more subject to changes. This enriches the coexistence patterns displayed on static networks.

### 2.2 Percolation properties

We now check the conjectured percolation transition. For different system sizes, we find that the cluster size distribution $P(s)$ displays power-law tails when the control parameter $d \equiv zN$ approaches a critical value $d_c = 1.32 \pm 0.05$ (corresponding to $z_c = d_c/N$), suggesting the onset of a second-order percolation-like phase transition. As shown in figure 4, $P(s) \propto s^{-\gamma}$ with $\gamma = 2.45 \pm 0.05$ at the phase transition. Fig. 5 shows that the average fraction of nodes in the largest component remains negligible for $d < d_c$, whereas it takes increasing finite values above $d_c$. As an additional check, following the method adopted in ref.[17], we have plotted the average size fraction of non-giant components, which diverges (in the infinite volume limit) when $P(s)$ decays algebraically as reported in the inset of fig. 5.

### 3 Ad-hoc networks

Connectivity of clients in an ad-hoc network is reached by means of signal transmitters scattered in the territory. Differently from base station antennas, transmitters of ad-hoc networks have a much lower transmitting power (e.g. up to 2W ERP), while the signal coverage of territory is ensured by their large number. Each transmitter emits with a given radiating power that in principle may differ from each other. Ad-hoc transmitters might be placed at fixed strategic points as for base stations antennas, for instance near public illumination lamps or onto traffic lights, but still the
Figure 4: Cluster size distribution $P(s)$. Far from the critical threshold ($d = 0.1$ and $d = 4$), $P(s)$ is well peaked. At $d_c = 1.32$, $P(s) \propto s^{-\gamma}$ with $\gamma = 2.45 \pm 0.05$. Here $N = 3200$.

Figure 5: The fraction of nodes in the giant component for different network sizes as a function of $d$. Inset: the non-giant component average size as a function of $d$ for $N = 6400$. 

high number of transmitting devices needed to cover the territory may result in a large maintenance effort. Maintenance difficulties can be overcome if transmitting devices are delivered to public users. A certain number of transmitters, the spatial density of which is to be determined later, should be assigned to human agents on voluntary basis. As incentive, the communication company providing the service would reduce that user’s fees or even grant her free communication channels. The choice of user’s acquired benefits is of course dictated by free market competition and straightforward calculation of costs. In this user delivered transmitting device scenario, maintenance costs are reduced by the implicit interaction with active users, who have high interest to maintain their device active in order not to lose their gained commercial benefits. Users have to carry their transmitting device with them all the time and provide for a possibly constant powering, by taking care of battery recharge. Transmitting devices might be formally indistinguishable from mere passive devices that are found nowadays. In fact, we think that the hardware technology necessary to build such transmitters is already available.

From a practical point of view there are some theoretical problems to be solved. Firstly, a suitable communication protocol should be implemented. With this respect it is worth to mention the theoretical works on the limits of wireless network capacity of Gupta and Kumar [18, 19]; the works of Jelasity et al. [20], who propose a gossip-based protocol for computing aggregate values over network components; the work of Ravelomanana [21], who faces the issue of address auto-configuration in the start up phase of network creation. Secondly, the conditions that ensure network connectivity must be studied. Here we shall examine which macroscopic properties should transmitter devices posses in order to generate a connected network with certain given probability. The macroscopic properties we refer to, are essentially the spatial density of transmitters and the probability distribution of their transmitting power.

Here and in the following, we shall assume that the signal is transmitted isotropically, i.e. the lines at constant radiating power are simply spheres. The spatial position of transmitters would then be bound to the spatial position of users. Although it is in principle possible to distinguish from different typologies of users (travel salesmen or housewives for instance, the former with high mobility, the latter more static), we shall consider their position at random in a given box. An ad-hoc network will be built associating a node with a single user and drawing links whenever transmitter centered circles with given effective radii overlap. From now on, we shall use indifferently the denomination of node, user, and transmitter. The time evolution of the network will be mimicked by different ensemble realizations of the system so that the time averaged properties of the system will be substituted by ensemble averages. Transmitters will be placed at random in a two-dimensional unit box with both horizontal and vertical coordinates extracted with uniform probability. Periodic boundary conditions will be used, since it was shown that, with the same finite number of nodes, they deliver more accurate results with respect to sharp boundaries [22]. Although the real situation is of course three-dimensional, we perform our analysis in two dimensions, restricting ourselves to the case of the average transmitting radius larger than usual building heights. A transmitting radius will be assigned to each of the transmitters. The undirected network arising when transmitting radii are all equal to a given value and two nodes are considered connected if their associated circle overlap, is often in the literature referred to as unit disk graph [23]. We shall go further and also analyze the properties of the directed network constructed by extracting the effective transmission radius of each node from a given probability distribution with finite support. Links in this directed network will be drawn from a transmitter towards all other transmitting devices lying inside its effective covered area. This kind of network would be very close to what we would get in reality.

3.1 Total transmitting power

We shall show here with the help of simple considerations that the total electromagnetic power necessary to cover with a signal a given area \( A \) embedded in a three-dimensional space, is independent
from the effective transmitting radius of transmitters. We shall see that this property stems from the fact that both the number of transmitters necessary to cover a certain area, and the radiated power per surface unit in three dimensions, scale with the inverse square of the effective transmitting radius. Let us indicate with \( P_i \) the effective radiating power of a transmitter necessary to maintain power \( W_i \) at a distance \( r_i \) from the source. At distance \( r_i \) from the transmitter the power \( W_i \) measured by a receiving device is simply proportional to \( P_i / r_i^2 \) [24]. Suppose a receiver is able to detect signals only above a certain threshold \( T \), which then defines the effective transmission radius of a transmitter. In case of different effective transmitting radii \( R_i \) we must have \( T = P_i / R_i^2 \) for each \( R_i \). If we have \( N_i \) transmitters with ERP \( P_i \) covering an area \( A \), we get for the total radiated power \( P \) the expression
\[
P = \sum N_i P_i = T \sum N_i R_i^2 \text{.}
\]
If the effective transmitting circles were not overlapping, then the area covered by the signal would be \( A = \sum N_i \pi R_i^2 \) and consequently \( P = TA/\pi \), expression that does depend neither on the \( N_i \) nor on the \( R_i \). We stress that this result is true if the transmitting circles do not overlap.

We shall see later in this paper that in the case transmission centers scattered with random uniform probability in the area \( A \), the number of transmitters with effective area \( \sigma = \pi R^2 \) needed to cover the space would be \( N = A^2 / \sigma \) rather than simply \( N = A / \sigma \). Intuitively we need more transmitters than in the overlap free case to cover a given area, since circles with random extracted centers are not optimized to cover holes. In this case, the total electric power needed to maintain the network would be \( P = TA / \pi \sum \log \frac{A}{\sigma} \). Consequently, the minimization of total power may be achieved by using transmitters with large effective transmitting radii, whose transmitting power should take into account both human health and the possibility to assure interference free communications.

4 Undirected ad-hoc networks

Given a certain number of nodes (transmitters) with an associated effective signal area in a unit square with periodic boundary conditions (PBC), the undirected ad-hoc network is constructed by joining two nodes whose signal areas overlap. In case of signal areas with a circular shape of given radius \( R \), a link is drawn whenever two nodes are less than the sum of their signal radii apart.

4.1 Fixed transmitting power

In this section we shall analyze the case of an undirected geometrical network, where nodes have the same transmitting radius \( R \). Two nodes are connected if their mutual distance is less than \( 2R \). Equivalently, two nodes are connected if one of them falls inside a circle of radius \( 2R \) centered on the other. In continuum percolation theory, the area of this circle with double radius is often to as excluded volume \( V_{ex} = 4\pi R^2 \). If we restrict our system in a unit square, the excluded volume represents directly the probability that two random nodes are connected. Since the \( N \) nodes are drawn independently and uniformly at random, the probability distribution of the degree (connectivity) \( k \) will be given by a binomial distribution
\[
\binom{N}{k} p^k (1 - p)^{N-k}
\]
with \( p = V_{ex} \), and the average degree will be simply given by \( \alpha = \langle k \rangle = Np = NV_{ex} \). At the thermodynamic limit, when \( N \) grows to infinity while \( \alpha \) remains constant, we can pose questions about the emergence of a macroscopic large cluster, i.e. a cluster with size growing as \( N \). Percolation theory predicts the existence of such cluster and that its appearance is ruled by a phase transition at some critical value of \( \alpha = \alpha_c \). Since there exists no analytical calculation able to devise this critical value yet, \( \alpha_c \) must be computed numerically and its value in 2-dimensions is found to lie around \( \alpha_c = 4.51 \) [22].
If we intend to realize an ad-hoc network with ideal transmitting devices with fixed transmitting radius and we need to ensure the presence of a macroscopic large cluster of nodes, then we have to choose a spatial node density greater than the critical value of \( N_c = \alpha c / V_{ex} \). What we really like to have is to set up a node spatial density in order to guarantee, with high probability, the emergence of a connected network, where all nodes are connected. The problem of connectivity has already been dealt with in literature and remains an hot topic today as well.

The basic question one poses is: given a certain number \( N \) of transmitters, how should one choose the excluded volume \( V_{ex} \) (or in our two-dimensional case the excluded area \( A \)) in order to have a totally connected network with high probability? An answer was already given by Gupta and Kumar [25], who showed that if \( N \) nodes are placed in a disk of unit area and each node transmits at a power level so as to cover a circular area

\[
A = \pi R^2 = \frac{\log N + c(N)}{N},
\]

then the resulting network is asymptotically connected with probability one if and only if \( c(N) \to +\infty \). The term \( c(N) \) is still allowed to grow slower than the logarithm, such that expressions like \( c(N) \approx \log \log N \) are feasible. Variants of this theorem can also be found in literature. We cite here the work of Xue and Kumar [26], who faced the problem from the point of view of the number of nearest neighbors needed to ensure total connectivity with high probability (they proved that the number of nearest neighbors should scale as \( O(\log N) \) and conjectured that the multiplicative constant should be strictly one).

**Minimally connected networks with arbitrary shaped agents**

In this section, we shall show with simple probabilistic reasonings, that in the case of a geometrical random network with \( N \) nodes placed uniformly at random in a unit square in two-dimensions and having excluded area \( A \), the network is asymptotically minimally connected if \( N = \log A / \log(1 - A) \) or equivalently \( A \approx (\log N - \log \log N) / N \). By “minimally connected” we denote a network with at least two of its nodes connected. The reasoning proceeds as follows.

As previously noticed, if we restrict to a unit square, the probability that a point in the square is covered by a particular node is simply \( A \). On the contrary, the complementary probability that a given point in the square is not covered by the area around a chosen node is \( 1 - A \). It follows immediately that the probability not to cover a given point in the square after \( N \) nodes have been placed randomly and independently is \( (1 - A)^N \). Since the area of the square is unity, the latter expression coincides with the average free space, i.e. uncovered space, in the square after \( N \) nodes have been placed. The network is minimally connected, as average, if the free space is less than the spanned area of the single transmitter, since in this case adding a new node would necessarily lead to an overlap of areas. Thus:

\[
(1 - A)^N < A
\]

or equivalently

\[
N > \frac{\log A}{\log(1 - A)}.
\]

Expression (12) may be well approximated as

\[
e^{-AN} < A
\]

since usually one has \( A \ll 1 \), i.e. the area spanned by the signal emitted by a transmitter is much less than the required covered territory area. The issue here is to solve the transcendent relation

\^[1\] We also carried on our analysis without this approximation and got same results. By using expression (14) instead of expression (12) we simplify the analytical treatment of the problem.
(14) with respect to \( A \). To achieve this, we rewrite relation (14) as
\[
e^{-AN} < 1
\]
and ask which expression for \( A(N) \) solves the associated equality. We substitute \( A \) with expression (11) and obtain
\[
e^{-c(N)} = 1.
\]
We observe that by imposing \( c(N) = -\log \log N \), we get for the left hand side
\[
\log N \quad \frac{\log N - \log \log N}{N}.
\]
expression that goes to one as \( N \) grows to infinity. For large \( N \) we find then for the area \( A(N) \) that ensures a minimally connected network, the approximate expression:
\[
A(N) \approx \log N - \log \log N.
\]
We notice that all the previous reasoning do not depend on the shape of the effective transmitting area.

**Connectivity with circular shaped agents**

We go back to the case of circular transmitting areas. The probability to have a circular area \( S = \pi r^2 \) still uncovered after \( N \) nodes with area \( A = \pi R^2 \) have been placed at random in a unit square, is equivalent to the probability not to have any node inside a circle of radius \( R + r \), that is \((1 - \pi (R + r)^2)^N \). If \( r = R \) we get
\[
(1 - 4\pi R^2)^N \approx e^{-4NA} = e^{-NA_{ex}}
\]
with \( A_{ex} \) denoting the excluded area as defined above in the text. By substituting expression (11) in the previous relation we get
\[
e^{-4NA} < \frac{1}{N^4}
\]
that thus gives an estimation of the probability that after \( N \) nodes carrying a transmission area \( A(N) \) have been placed in the box, a free uncovered circle of area \( A(N) \) emerges. In that case, if the next extracted node might fall inside this area and would be disconnected from the others. The complementary probability \((1 - \frac{1}{N^4})\), gives then a pessimistic estimation of the probability that a geometric network constructed as above would display disconnected sections. Finally, If \( N \to \infty \) and \( A(N) \) is chosen as in Eq. (11) the resulting network is connected with high probability.

Further, the average number of nearest neighbors of a node with transmission radius \( R \) in a unit box, i.e. its average connectivity \( \langle k \rangle \) was previously mentioned to be equal to the excluded volume or area times \( N \). In our case \( \langle k \rangle = 4NA(A(N) \), and by substituting again expression (11) we get
\[
\langle k \rangle = 4\log N + 4c(N) > 4\log N.
\]
Xue and Kumar demonstrated that if \( \langle k \rangle = O(\log N) \approx \gamma \log N \) then the corresponding random geometric network is connected and conjectured that \( \gamma \) should be unity [26]. Eq. (21), instead, suggests that the multiplicative constant \( \gamma \) should be larger than 4.
4.2 Random transmission power

As already discussed in the introduction, the real implementation of an ad-hoc multihop network would be characterized by the presence of nodes with transmitting power drawn from a random distribution density rather than presenting a fixed transmitting radius (Dirac’s delta peaked distribution density). In this section we shall generalize some network characteristic quantities already known for the fixed power case. In the following we shall speak indifferently about radius distribution density $p(R)$ and transmitting area distribution density $P(A)$. In the particular case of circular areas the respective distribution densities are connected by the equiprobability law

$$P(A) = p(R) \frac{dR}{dA} = \frac{p(\sqrt{A/\pi})}{2\sqrt{\pi A}}.$$  \hspace{1cm} (22)

The transmitting radius probability densities that we shall use for our analysis will be:

- **R** fixed radius: $p(R) = \delta(R)$ with $R = 0.01$;

- **2R** two radii: $p(R) = (\delta(R_1) + \delta(R_2))/2$ with $R_1 = 0.003$ and $R_2 = 0.02$;

- **unif** uniform distribution density: $p(R) = \Theta(R - R_1)\Theta(R_2 - R)/(R_2 - R_1)$ with $\Theta(x)$ representing the Heavyside theta function with value one if its argument is non negative. $R_1$ and $R_2$ are chosen as above.

- **bpl** bounded power-law: $p(R) = c\Theta(R - R_1)\Theta(R_2 - R)R^{-2}$, with $c$ normalization factor and $R_1$ and $R_2$ chosen as above.

We still shall consider our system inside a unit square with PBC. In Fig. 6 we show a random geometric network generated by the uniform radius distribution density above with 5000 nodes together with its topological representation.
Figure 7: Linear-Log plot. Numerically calculated free uncovered space as a function of the number of transmitters, for the radius probability distributions defined in the text at page 11. Slopes coincide with (minus) the average area for each distribution.

4.2.1 Free space

The probability to cover a point in the unit box by means of a randomly extracted node with given transmitting area $A$ is obtained by the product of the probability to extract area $A$ from its distribution density $P(A)$ times the probability $A$ that the point in the box falls in that area: $AP(A)$. Generally we get, after one node extraction, that the probability to cover a given point is $\sum_A AP(A) = \langle A \rangle$, where the discrete sum turns into an integral if $A$ is a continuous variable. We see that now the average transmitting area $\langle A \rangle$ takes the place of the area $A$ of the previous fixed transmitting power case of section 4.1. After $N$ independent node extractions the average free uncovered total space $F(N)$ will be

$$F(N) = (1 - \langle A \rangle)^N \approx e^{-N\langle A \rangle}$$

(23)

As for the case of fixed transmitting radius, it is possible to estimate the minimum number of nodes necessary to have a connected network once the distribution density of areas is fixed. In fact relation (12) now becomes

$$(1 - \langle A \rangle)^N < A_0$$

(24)

where $A_0$ is the minimum value of the transmitting areas. Thus, the network is connected when $N > \log A_0 / \log(1 - \langle A \rangle)$.

4.2.2 Degree distribution

The probability that a the circle of radius $r$ surrounding a given node, will intersect a circle of radius $R$ associated to another node is equal to the area of an effective circle of radius $r + R$ (sort of effective $r$-dependent excluded area) times the probability to get the node with radius $R$. The average value of the degree of the node with radius $r$ will be:

$$k(r) = N\pi \int (r + R)^2 p(R) dR.$$  

(25)

The previous equation defines also the degree distribution $P(k)$, by using the analogous of formula (22), i.e. $P(k) = p(r(k))r'(k)$. The average value of the degree in the network will be

$$\langle k \rangle = N\pi \int (r + R)^2 p(R)p(r) dR dr = 2\pi N[\langle r^2 \rangle + \langle r \rangle^2].$$

(26)
Figure 8: Numerically calculated relative size of the largest cluster in the network as a function of the average degree. In case of distribution $R$ (all nodes with same transmitting power), the percolative transition is known to appear at $\langle k \rangle \approx 4.51$. The radii distributions are defined in the text at page 11.

### 4.2.3 Size of largest cluster

The emergence of a macroscopic cluster of connected nodes is a crucial issue for a network that should be the support of communication transmissions. We know already that the case of a fixed transmission power yields a percolative phase transition at average connectivity $\langle k \rangle$ around the value of 4.51. In Fig. 8 we depict the order of the largest cluster in the network normalized to the total number of nodes in the graph. We indeed need better statistics than what is shown, in order to draw definitive conclusions. Nevertheless it seems that the choice of the transmitting power distribution might shift the position of the phase transition with respect to the average connectivity. Further analysis is required.

### 4.2.4 Clustering coefficient

In the theory of complex networks very often is analyzed a statistical quantity named clustering coefficient. Given a node in the network, its clustering coefficient is calculated as the fraction of all triangles formed by the node itself and its nearest neighbors, normalized to all possible triangle so defined. In a communication frame, the clustering coefficient might be of some importance, since if nodes a, b and c are connected, then the interruption of the direct communication between a and b would not preclude their mutual communication, as the information would pass through c. This idea should be developed in the near future. The clustering coefficient of a network is defined as the average clustering coefficient of its nodes. Fig. 9 depicts the total clustering coefficient of networks generated as above.

### 5 Directed ad-hoc networks

The undirected ad-hoc network as described in the previous sections are an ideal abstraction of what an effective working complex communicating network should be. It seems more reasonable to define a connection from node a to node b if the effective area of signal emitted by a embraces node b. In this way a more realistic network is defined directed. A next complication, which we shall not
deal with here, is to assign to each directed link a weight, possibly modeling the different bandwidth capabilities of nodes. For this latter case we already developed in the past year activity of WP2.2.1 the mathematical tools to analyze the relevant statistical properties of the network. The case of all node of the network with fixed transmitting power, can be easily generalized in the frame of directed networks, by simply replacing the excluded area with $\pi R^2$. In fact if two nodes are connected then the link must be reciprocal, while these two nodes must lie one in the transmitting effective area of the other. Differences from a trivial undirected network will arise when nodes may posses different transmitting radii.

5.1 Degree distributions

In a directed network there are two quantities related to the connectivity of a node: the in-degree (the number of incoming links) and the out-degree (number of outgoing links). Given a node $m$ with associated transmitting area $A_m$, the probability to draw a link towards another node is simply $A_m$ (we still restrict to a unit square). The out-degree distribution of node $m$ is then given by the binomial distribution of Eq. 10 with $k = k_{out,m}$ and $p = A_m$. The average $k_{out,m}$ is $\langle k_{out,m} \rangle = NA_m$ so that the the out-degree distribution $p(k_{out})$ will be essentially given by the distribution of the transmitting areas. The average out-degree is simply $\langle k_{out} \rangle = N \langle A \rangle$. The in-degree obeys the same relations.

5.2 Percolative phase transition

The questions about the existence of a percolative phase transition should be now addressed to the emergence of a macroscopic strongly connected component (SCC), defined to be a portion of graph whose nodes can be reached from all others. We shall carry this analysis in the nearest future.

6 Analysis software

We developed an open source software with which we are able to study many of the quantities dealt with in this report. It has the possibility to generates networks with given criteria ranging from the preferential attachment rule to the geometric random networks. It automatically performs ensemble
averages, by specifying the number of system copies to be handled. This software can be downloaded from the web page http://pil.phys.uniroma1.it/~servedio/Software.html.

References
