

Social networks and the spread of epidemics

Carlo Piccardi

Published online: 21 August 2013
© Centro P.RI.ST.EM, Università Commerciale Luigi Bocconi 2013

Abstract In addition to the modern web-based “social networks”, where contacts are only virtual, there is a more traditional network of physical contacts among individuals, which is ultimately responsible for the transmission of all types of diseases. In this paper, we introduce the basic models of networks used to describe the social contacts within a population. Then we study how to transfer the traditional epidemic models, based on the “homogeneous mixing” assumption, to the new framework. We will see that, moving from the classical to the network modelling approach, results become more complex and somehow unexpected, as the structure of the social network plays a fundamental role.

Keywords Social networks · Epidemics · Epidemiology · Network theory · Homogeneous network · Heterogeneous network · Scale-free network · Small-world network · SIS model · Erdős number

The phrase “social network” has been used in recent years to denote the virtual communities of users of websites such as Facebook, Twitter and the like. Thus, the connections between individuals (“friends” in Facebook or “followers” in Twitter) is virtual, mediated by the World Wide Web and hence transmitted by the technological structure of the Internet.

However, well before the advent of the Internet, quantitative sociology used this term to designate sets of

individuals connected by more traditional links, mostly personal contacts, entailing physical proximity [16]. For instance, an adult living and working in a city has social contacts, though with different frequencies, with his relatives, coworkers, friends, public transportation users, and so on. Clearly, this is the kind of social network we are interested in, to understand and model the processes involved in epidemics, most of which spread by contact between individuals: some by simple proximity (for instance, sneezing or shaking hands can transmit flu or measles), other ones by less casual contacts (as in sexually transmitted diseases, or the spread of hepatitis or HIV through infected material).

The connection between epidemiology and mathematics goes back a long way. The first significant use of a mathematical model dates back to 1927, when Kermack and McKendrick suggested a system of differential equations in order to explain the evolution data of a plague epidemic in India [7]. Their crucial assumption lay in describing the population as a kind of perfectly mixed gas, in which every individual is equivalent to every other and has, in a fixed time interval, the same probability of meeting any other individual. It is through the meeting of a healthy and not immune individual (called *susceptible*, S) and of an infected one (called *infective*, I) that the infection can be transmitted, with a probability depending on the epidemic’s virulence: the susceptible person becomes infected.

This modelling approach was the main one in mathematical epidemiology for about 70 years, with countless variations and extensions that made it possible, for instance, to adapt the model to different pathogens or to refine it by distinguishing individuals according to age or sex [1]. Let us see an example, one of the simplest: in the SIS model it is assumed that each individual may get infected and recover,

C. Piccardi (✉)
DEIB, Dipartimento di Elettronica, Informazione e
Bioingegneria, Politecnico di Milano, Piazza Leonardo da Vinci,
32, 20133 Milan, Italy
e-mail: carlo.piccardi@polimi.it

passing through the transitions $S \rightarrow I \rightarrow S \dots$ an unlimited number of times. If we denote by $y(t)$ the proportion of infectives in the population at time t , so that $s(t) = 1 - y(t)$ is the proportion of susceptibles, we have:

$$\dot{y}(t) = -\gamma y(t) + \beta y(t)s(t) = -\gamma y(t) + \beta y(t)(1 - y(t))$$

The first summand on the right-hand side represents the flow $I \rightarrow S$ of recovered people ($1/\gamma$ is the average time after which an individual recovers), while the second summand is the crucial one to describe the $S \rightarrow I$ process of infection: the probability of each S meeting an I is proportional to the density $y(t)$ of the latter, so the number of $S \rightarrow I$ transitions is proportional to the product $y(t)s(t)$. The transmission coefficient β depends on both the social structure, which dictates the frequency of meetings, and the virulence of the pathogen.

What does the SIS model predict? It is easy to verify that there exist two equilibrium solutions (Fig. 1): $y = 0$, which corresponds trivially to the absence of infectives, and $y = 1 - \gamma/\beta$, which is positive and asymptotically stable if $\beta > \beta_{cr} = \gamma$, while for $\beta < \beta_{cr}$ the solution $y = 0$ is the asymptotically stable one. We may deduce that the number of infected people grows with β , as we could have guessed, but, more importantly, that there exists a threshold value β_{cr} under which the epidemic is bound to die out. Indeed, if β is too small, the few infected people in the population are not able to generate a sufficient number of other infected individuals, due to the scarcity of contacts or the low virulence of the pathogen. The existence of this threshold has always been one of the cornerstones of epidemiological modelling [1].

1 From traditional models to network models

The complexity of present day societies, especially in the most advanced countries, puts a strain on the kind of modelling described above. While the hypothesis of a

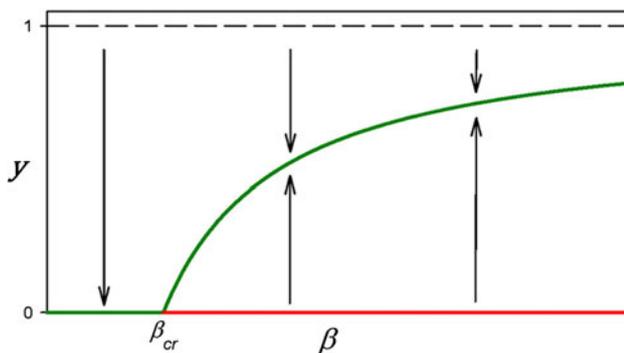


Fig. 1 In the classical SIS model, the proportion of infectives at the equilibrium grows with the transmission coefficient β , but the epidemic dies out (that is, $y \rightarrow 0$) if $\beta < \beta_{cr}$

“perfect mixing” in the population has always been debatable (since everybody has a different network of social interactions), the assumption that all individuals are about equivalent in terms of number of contacts is even less solid: several studies have shown that people are highly diverse, in the sense that a population typically include a small set of individuals (the so-called “hubs”) having a high number of social contacts, within the whole of a population having generally a middle-low connectivity (see examples in Fig. 2). We might rightly guess that the hubs play a crucial role in the spread of epidemics.

Another point of view involves the actual details of spatial spread. The SIS model sketched above does not take spatial factors into account, since it imagines the population to be confined in a “container”, where it is perfectly mixed. If we want to describe the spread of the epidemic in a region, we have to derive a partial differential equation model in the variable $y(t, x)$, where x is the spatial variable. The simplest and most intuitive generalisations of the SIS model assume a diffusive term, which means that an epidemic originating at a point expands with a wavefront covering ever larger areas. This is a phenomenon that has been observed very often in the past. Figure 3 shows as an example the Black Death epidemic that had decimated Europe by the middle of the fourteenth century: having appeared in Sicily in 1347, the epidemic gradually spread northwards, covering the whole of Europe within 3 years. The spreading was conveyed by the slow and not so numerous journeys by travellers and merchants between adjacent villages and cities.

The present situation is far different, as is clear from the second case shown in Fig. 3, that of the SARS epidemic in summer 2003. The epidemic originated in the Far East (Hong Kong), but within a few weeks important centres of infections appeared in several cities around the world, especially those harbouring significant Chinese communities. The process of spreading was twofold. On one hand, just like the case discussed above, it was of a “local” kind: the network of social relations mainly includes people who are nearby (relatives, coworkers, friends and so on) with whom contacts are frequent. This was the earlier mechanism of spreading, and indeed, it caused a very high incidence of the disease in crowded Hong Kong and its environs. But this was followed by a long-range spreading carried by air transport, in this case toward the main emigration areas. In addition to local contacts, which are often abundant and frequent, many individual also have “remote” contacts, typically few and uncommon, but on the whole sufficient to spread the epidemic far away in a very short time. This is a process which was clearly impossible not just at the time of the Black Death, but even a few decades ago, and which cannot be possibly described by the conventional diffusive models, but requires network models instead.

Fig. 2 The number of proximity contacts in the residents of a big city (left, [6]) or the number of sexual partners in the previous 12 months declared by the individuals of a community (right, [8]) vary greatly, suggesting a strong heterogeneity in the connectivity of the individuals themselves

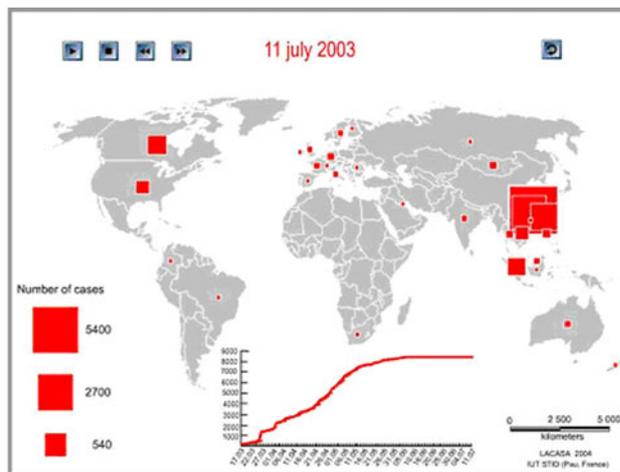
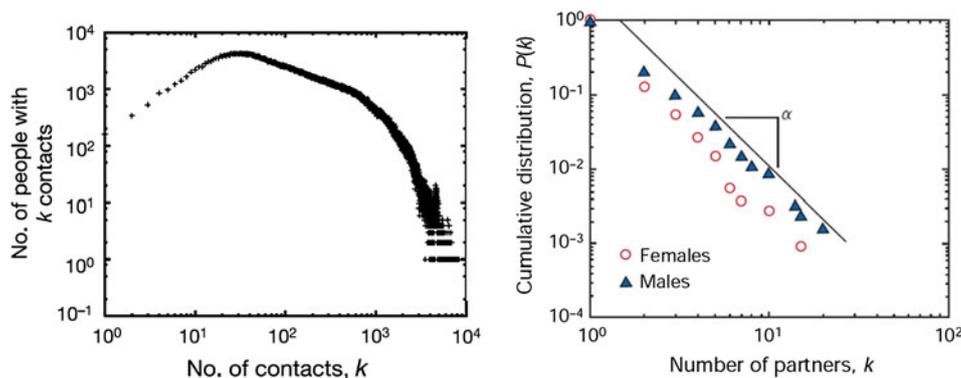


Fig. 3 Left the geographical spreading of the Black Death in the fourteenth century is described by a classical diffusive model [13, p. 105]; right in the case of the 2003 SARS epidemic, the epidemic spreads over large distances rapidly, due to air transport [3]

2 Elements of network theory

A network is defined by a *graph* consisting of N nodes and L edges (or links) [2, 5, 11, 12]. In the case of social networks, the nodes represent individuals, while the links describe the social relationship that exists between pairs of individuals. As already mentioned, this may be a virtual relationship, if we consider recent, web-based “social networks”, or an actual physical connection, if the network describes associations between relatives, friends, colleagues, etc. Since we are dealing with the spread of epidemics, we may assume the network to be *undirected*, as the pathogen can transmit both in the direction $i \rightarrow j$ and in the direction $j \rightarrow i$. Moreover, we shall confine ourselves to the simplest case, and omit assigning a *weight* $w_{ij} > 0$ to each $i \leftrightarrow j$ link, which might for instance give the frequency of social contacts between different pairs of individuals. Under these hypotheses, the network is completely described by its *adjacency matrix* $A = [a_{ij}]$, an $N \times N$ matrix in which $a_{ij} = a_{ji} = 1$ if the $i \leftrightarrow j$ link exists, while $a_{ij} = a_{ji} = 0$ otherwise: A is therefore a symmetrical matrix (Fig. 4).

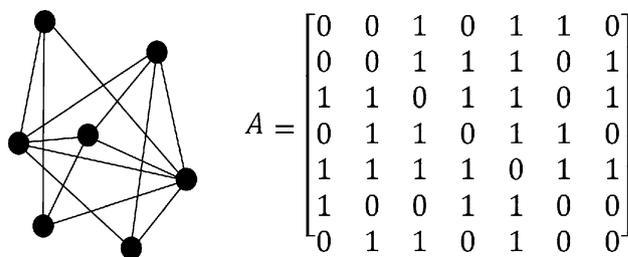


Fig. 4 A network with $N = 7$ nodes and $L = 14$ links, and its adjacency matrix A

The network is said to be *connected* if for every pair (i, j) of nodes there is a path joining i and j in the graph. The *distance* d_{ij} is then defined as the length (measured by the number of links) of the shortest path (i, j) . Hence, we can introduce the notions of *diameter* D of the network, that is, the maximum of distances d_{ij} , and of *average distance* d , that is, the average of all d_{ij} values of the $N(N - 1)/2$ distinct pairs of nodes. As we shall see later, a crucial feature of real-life networks, and in particular of social ones, is a surprisingly small average distance, even when their size N is very large.

The *degree* k_i is the number of contacts of the node i , that is, the number of links incident to i : so we have $k_i = \sum_j a_{ij}$. The single most important description of a network is given by its *degree distribution* $P(k)$ which, for each $k = 1, 2, \dots$, gives the relative abundance of k -degree nodes in the network:

$$P(k) = \frac{\text{number of nodes of degree } k}{N}.$$

Since $\sum_k P(k) = 1$ and $0 \leq P(k) \leq 1$ for each k , the degree distribution is formally analogous to a discrete probability distribution. In this sense, the value $P(k)$ can be taken to be the probability that a randomly chosen node has degree k . The *average degree* m of the network is given by

$$m = \frac{1}{N} \sum_{i=1}^N k_i = \sum_k kP(k) = \frac{2L}{N}$$

where the last expression is a consequence of the fact that each of the L links contributes 1 to the degree of both the nodes it connects.

A network is said to be (strictly) *homogeneous* when all of its vertices have the same degree $k_i = m$, hence $P(m) = 1$ and $P(k) = 0$ for $k \neq m$. This is a situation that never occurs in real-life networks, but it is however possible to find examples of *almost homogeneous* networks, that is, in which the node degrees never lie too far from the average value m . An important example is given by the so-called *Erdős–Rényi networks*: having fixed N and L (and hence the average degree $m = 2L/N$), the network is constructed by assigning links to L pairs of nodes selected randomly among the $N(N-1)/2$ possible ones. For sufficiently large N and L (formally: for $N \rightarrow \infty$ with m fixed), we can prove that $P(k)$ converges to a Poisson distribution with mean m : what this means is that most nodes have a degree not far from m which can so be taken as a representative value (or “typical scale”) of the degree of the nodes. So, in terms of social networks, an Erdős–Rényi network represents a community of individuals in which everybody has approximately the same number of acquaintances (Fig. 5).

In the last 15 years, extensive analysis of representative data for real-life networks has clearly shown that (almost) homogeneous degree distributions are the exception rather than the rule. Instead, most networks have strongly heterogeneous degree distributions, where few nodes with a very high degree (the so-called *hubs*) can be found together with a majority of medium-low degree nodes. Such examples have been found almost everywhere, from biology to the WWW, from transport to trade networks, to mention just a few instances. In a social network, this means that in the population there are individuals who, due to their kind of activities, are in touch with a very large number of people. Alongside these individuals, there are others who, at the opposite end of the distribution, have just a few contacts. For such a network, the average degree m , while mathematically well-defined, has no informative value, since the deviation from it may be huge: the network has no “characteristic scale” and is indeed called *scale-free*.

In a celebrated paper, Barabási and Albert [4] proposed a simple and elegant algorithm to create a scale-free network artificially. Having fixed the average degree m (integer and even, for the sake of simplicity), the network is iteratively constructed by adding at each step a new node with $m/2$ links, to be connected to $m/2$ existing nodes. The crucial ingredient of the algorithm lies in selecting these nodes, which are picked up randomly by assigning to each node a probability of being selected proportional to its current degree k_i . In other words, the new node will tend to connect to nodes that are already well connected, with an effect of positive feedback that, in the end, is what is actually responsible for the creation of hubs. Adding more and more nodes (and hence for $N \rightarrow \infty$), the degree distribution converges toward a *power law* of the form $P(k) \sim k^{-3}$. It is important to remark that such a distribution has a divergent variance, that is, $\sigma^2 = \sum_k (k-m)^2 P(k) \rightarrow \infty$ when $N \rightarrow \infty$. This property does not hold for real-life networks with a finite N . However, for such networks, if N is very large, σ^2 will tend to be very large

Fig. 5 *Left* an almost homogeneous (Erdős–Rényi) network with $N = 100$ and $L = 300$. *Right* a Poisson degree distribution with average degree m [15]

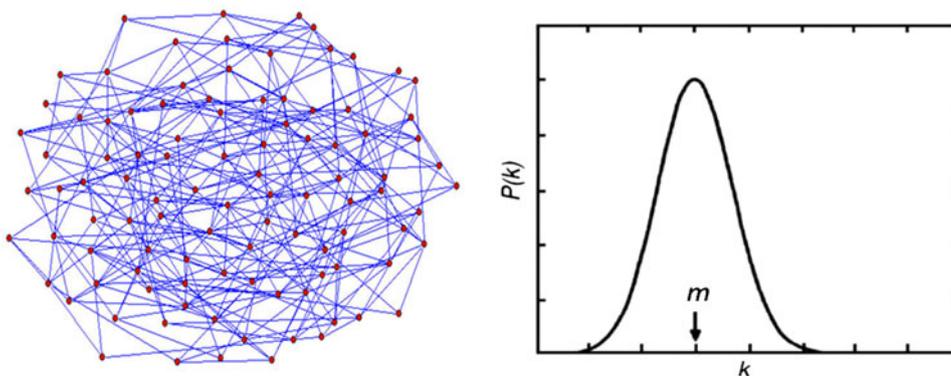


Fig. 6 *Left* a simplified example of a non-homogeneous network. *Right* a power law degree distribution (truncated to a finite k) in a logarithmic scale; the node degrees vary from 1 to 100 (from [15])

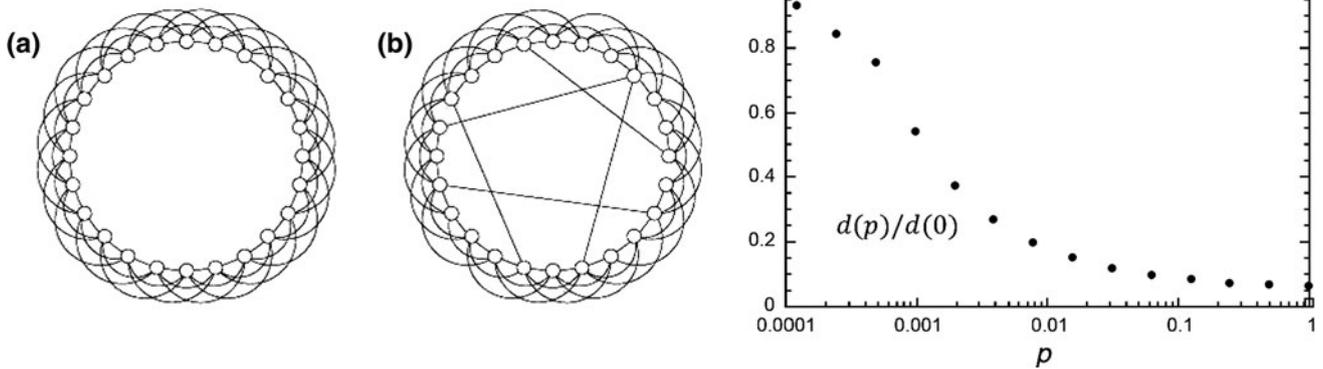
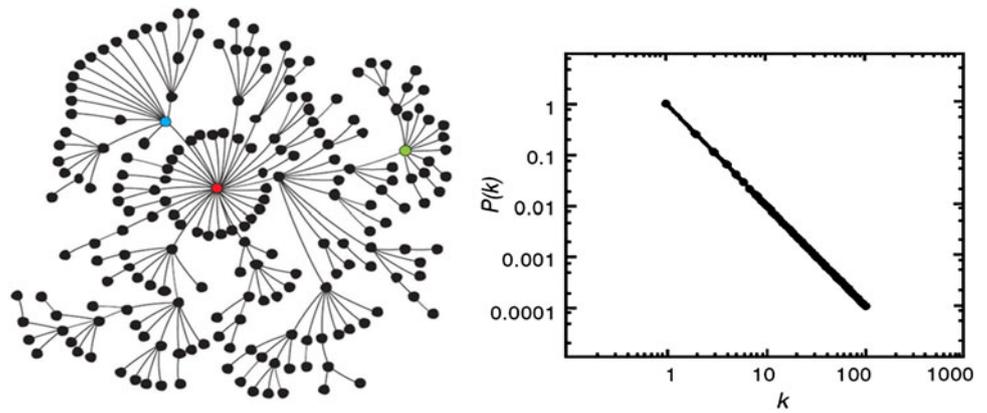


Fig. 7 With just a small proportion p of “long-distance” links, a regular network (**a left**) acquires the small-world property (**b right**); the average distance $d(p)$ decreases rapidly with p (modified from [17])

too, with interesting consequences on the mechanism of spreading of epidemics, as we shall see (Fig. 6).

Another feature clearly emerging from the study of real-life data involves the average distance d , which is often surprisingly small with respect to the size N of the network. In the context of social networks, this property was first highlighted in the 1960s by the now famous experiment by Milgram [9], which gave rise to the “theory” of the “six degrees of separation”, now fully a part of popular culture. According to this theory, in the social network consisting of all the inhabitants of the Earth (in which N equals some billion), the distance between two arbitrary individuals is on the average six. In other words, assuming two individuals to be linked if they are related by kinship, friendship etc., it would be possible to connect any two human beings by a chain consisting on the average of six links. While this belief is based on extrapolating from experiments performed on a far smaller scale, there is reason to believe that the actual average distance d is not much larger, and in any case several orders of magnitude smaller than N . A network with this property is called a *small-world* network. More precisely, if we consider a set of networks parametrised by

N , we say that the set has the small-world property if d grows as $\log N$, and hence far slower than N .

In 1998, Watts and Strogatz [17] provided a simple experimental explanation of the small-world property, showing that d is small when there is even a slight proportion of “long-distance” connections. The experiment works as follows: we start with a network with a regular geometry, that is, in which all connections are of “local” type; then we randomly select a (small) proportion $0 < p < 1$ of the links, for each of these we disconnect one randomly chosen endpoint, and reconnect it to a node again chosen at random (“rewiring”). As the diagram in Fig. 7 shows, a value as small as $p = 0.001$ (that is, just one thousandth of the links is rewired) is sufficient to halve the average distance d , while $p = 0.01$ reduces it to 20 % of the original value. It can be shown that, while $d \sim N$ in the regular network, we have $d \sim \log N$ in the rewired one. So, long-distance connections turn a regular network into a small-world one and, in the case of epidemic spreading, they allow a rapid transmission of the contagion in all areas of the network: this is what happened for the SARS epidemic mentioned earlier.

3 Epidemic spreading in a network

Having described a population with a network model, the spreading of an epidemic can be seen as a dynamic process that uses the network as its support. The most direct and effective class of models exploitable to this end is that of *probabilistic cellular automata*. In their simplest version, these are discrete-time models, that is, models in which the events happen at times $t = 0, \Delta, 2\Delta, \dots$, having chosen Δ as our discretisation interval. At time t , each node i is in a state x_t^i belonging to a finite set $\Sigma = \{0, 1, \dots, q - 1\}$, typically of a small cardinality q . At time $t + \Delta$, the node i will change to a state $x_{t+\Delta}^i$ depending, according to probabilistic rules to be specified, on x_t^i but also on the state of the nodes directly linked to i (the so-called “neighbours” of i). The cellular automaton is completely defined when we have specified the network, giving its adjacency matrix A , and the interaction rules among adjacent nodes. Having fixed the initial state x_0^i for all the nodes and allowing the system to evolve, what we obtain is a realization of a stochastic process, since the interaction rules are probabilistic. The first question we might ask is rather obvious: what properties are common to all possible instances? This means we are asking whether the epidemic is able to spread or not, and in the former case, which global value it may reach, which nodes are most frequently hit, and so on. The second question is subtler: does the answer to the first question depend on the topological properties of the network? We are asking, in other words, whether the structural features of the network (the degree distribution, for instance) play a role in the result of the spreading process.

As we have done above, for the sake of simplicity let us fix our attention on the SIS process. There are two possible states: susceptible and infected. So we take $\Sigma = \{0, 1\}$, defining $x^i = 0$ if i is susceptible, while $x^i = 1$ if i is infected. For the recovery process $I \rightarrow S$, assume that if the node i is infected at time t , then with probability $\gamma\Delta$ it will be healthy (so susceptible again) at time $t + \Delta$, while it will stay infected with probability $(1 - \gamma\Delta)$. As for the infection process $S \rightarrow I$, assume that the probability for the susceptible node i to become infected is $\rho\Delta \sum_{j=1}^N a_{ij}x_t^j$, that is, proportional to the number of infected neighbours (recall that $a_{ij} = 1$ if j is a neighbour of i , otherwise $a_{ij} = 0$).

This rule set can be easily transferred to instructions in a programming environment (Matlab, for instance): having defined the network by its adjacency matrix, it is possible to run several simulations and to analyse the results in several ways. However, we want to know if it is possible to get an aggregate description, possibly up to some approximation, of the results we might expect from those simulations: this means to write a *mean field model* for a process

intrinsically distributed. Let us start by writing a balance of the number of infected individuals $Y \in [0, N]$ between two consecutive (discrete) times:

$$Y(t + \Delta) = Y(t) - \gamma\Delta Y(t) + \rho\Delta\Phi(t)(N - Y(t))$$

where $N - Y(t)$ is the number of susceptible individuals, while $\Phi(t)$ is an estimate of the mean value of infected neighbours for every susceptible individual. Notice that we have gone from a *microscopic* description, typical of a cellular automaton, where the number of infected neighbours may well be different for each susceptible individual, to a *macroscopic* one, where we have to use averaged values. Now, if the degree distribution is (almost) *homogeneous*, we can assume that every node has m neighbours (the average degree); moreover, by ignoring possible spatial inhomogeneities, we may assume that a proportion of these neighbours equal to $y(t) = Y(t)/N$ is infected, that is, a proportion equal to that for the whole network. Incorporating these hypotheses in the balance equation, dividing by N , and taking the limit as $\Delta \rightarrow 0$, we find

$$\dot{y}(t) = -\gamma y(t) + \rho m y(t)(1 - y(t))$$

We immediately notice that this equation is the same as for the “classical” SIS model discussed above; that model is therefore valid not only under the hypothesis of a perfectly mixed population, but also in a network where the contact structure is fixed, on the condition that all individuals have (approximately) the same number of contacts. By comparing the two formulas, we notice that $\beta = \rho m$: the derivation through the network model makes it clear that the transmission coefficient β consists of two factors: the first one (ρ) purely biological and representative of the pathogen’s virulence, the second one (m) related to the structure of the social network. Just as the “classical” SIS model predicted, in a homogeneous network as well the epidemic can spread and survive only if $\beta > \beta_{cr} = \gamma$, hence only if $\rho > \rho_{cr} = \gamma/m$, otherwise it will die out.

The situation is quite different if the social network has a strongly *heterogeneous* degree distribution (scale-free network), as shown by Pastor-Satorras and Vespignani in a celebrated paper [14]. The hypothesis that the number of neighbours can be described by $\Phi(t) = m y(t)$ is not acceptable anymore, since the number of neighbours can stray far, node by node, from the average value m . Hence, it is necessary to substitute for the single equation in $y(t)$ a system of differential equations in the variables $y_k(t)$, $k = 1, 2, \dots$, one for each degree appearing in the network, where y_k describes the proportion of infected individuals *within the set of nodes of degree k* :

$$\dot{y}_k(t) = -\gamma y_k(t) + \rho \Phi_k(t)(1 - y_k(t)), \quad k = 1, 2, \dots$$

The term $\Phi_k(t)$ now represents the expected number of infected neighbours for a degree k node; as it can be shown,

it is proportional to k and is of the form $\Phi_k(t) = kF(y_1(t), y_1(t), \dots)$. Finally, the global proportion of infected can be written as $y(t) = \sum_k P(k) y_k(t)$.

The properties of the model just described can be studied partly in an analytic way and partly by means of numerical analysis: the results parallel the simulations run directly via probabilistic cellular automata. Whichever approach we prefer, the results agree in pointing out the disappearance of the threshold value ρ_{cr} for networks of arbitrarily large size N : the epidemic is able to survive and spread in the network even for an arbitrarily low virulence. More precisely, it can be shown (by stability analysis) that the threshold lies, for any degree distribution, at the value

$$\rho_{cr} = \frac{\gamma m}{\sigma^2 + m^2}$$

where m and σ^2 are, respectively, the mean and the variance of the degree distribution. In a strictly homogeneous network, $\sigma^2 = 0$, so $\rho_{cr} = \gamma/m$, as we have already seen. In a scale-free network as discussed above, on the other hand, when $N \rightarrow \infty$ we have $\sigma^2 \rightarrow \infty$, and so $\rho_{cr} \rightarrow 0$. In concrete terms, in large networks the threshold value ρ_{cr} is exceedingly small, even if not equal to zero.

As it is apparent from Fig. 8, for a very small ρ , the value y where the epidemic reaches a steady state in a strongly heterogeneous network is different from zero, but very low. Moreover, with a few algebraic steps, it can be shown that, again in a state of equilibrium, the proportion y_k of infected nodes of degree k is

$$y_k = \frac{1}{1 + \alpha/k},$$

where α is a function of the problem data but does not depend on k . So the proportion of infected nodes increases monotonically with k , and $y_k \rightarrow 1$ for an arbitrarily large k . In other words, the higher the values we take in account for the degree k , the fewer the nodes we will find ($P(k) \rightarrow 0$), but the higher will be the probability that they be infected ($y_k \rightarrow 1$). This is the process underlying a contagion within a heterogeneous network: once the epidemic is in the network, the hubs, due to their extremely high number of connections, have a high probability of getting infected. In a way, they are the “watchmen” of the epidemic, which is thus able to avoid extinction and survive, even if at low levels. Indeed, this is the scenario observed in some kinds of disease [10] as well as, turning to technology, the typical pattern in the spreading and persistence of computer viruses in the Internet [14].

The example we have studied here, while relatively simple, is sufficient to stress how, in going from “classical” to “network” models, we can find unexpected and, in any case, more complex and elaborate results. In the last

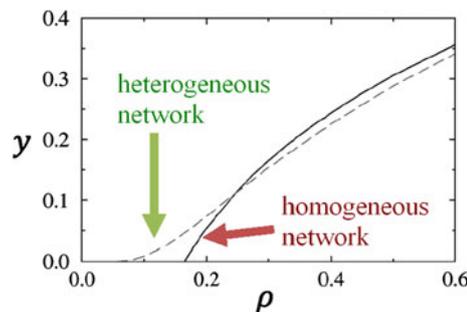


Fig. 8 The threshold value ρ_{cr} , under which the epidemic dies out, decreases when the variance of the degree distribution $P(k)$ increases. So, in heterogeneous networks we may have $\rho_{cr} \rightarrow 0$ when $N \rightarrow \infty$: the epidemic can survive and spread even with an arbitrarily low virulence (modified from [14])

10 years, research has extended in countless directions, and in the near future will continue to extend even farther. As for epidemics, models more complex than a simple SIS process have been used, as we might easily imagine, in order to describe different features of the many kinds of existing diseases. Regarding social networks, it is apparent that a “static” description such as the one given here, that is, one in which the individuals are fixed and each of them maintains the same contacts forever, is overly simplified. More realistic descriptions are bound to include the processes of birth and death of individuals (with the associated linking/unlinking of connections), their mobility (a network variable in time) or, in the most complex case, the reciprocal interaction between epidemics and topology: links vary in time as a function of the epidemiological state of the involved individuals (adaptive networks). All these extensions, as it is easy to understand, make studying networks more and more difficult and complex. However, it is an effort well worth making: on the whole, infectious diseases cause more than 15 million deaths worldwide every year (25 % of all deaths [10]). Let us hope that the joint study of epidemics and networks through mathematical models will help to lower this tragic toll.

Translated from the Italian by Daniele A. Gewurz

Appendix: Erdős number

An unusual social network is the one defined by *collaborations between mathematicians*: each node represents a scholar, while two nodes are linked if the two scholars have written and published at least one joint paper. Although the central connected core of the networks (that is, the main component that leaves out isolated nodes and small components unconnected with the rest) consists of no less than 268,000 nodes, the average distance between any two scholars is definitely low: about 7. This is a clear evidence

of the small-world property. A particularly special node is that corresponding to Paul Erdős (1913–1996), a very prolific Hungarian mathematician (more than 1,400 published papers), but chiefly a very collaborative one (509 different co-authors). In network jargon, this means that the Erdős node has 509 neighbours or, equivalently, that it has degree equal to 509. For any given mathematician, his or her *Erdős number* E is the distance between the corresponding node and the Erdős node. By the above, there are 509 scholars with $E = 1$. Due to the already mentioned small-world property, E may turn out to be surprisingly low: the present author, even though he does not frequently contribute to mathematical journals, has $E = 4$, slightly less than the average value, which is about 4.65. For further details, as well as to compute one's own Erdős number, see the *Erdős Number Project* web page (<http://www.oakland.edu/enp/>).

References

1. Anderson, R., May, R.: *Infectious Diseases of Humans: Dynamics and Control*. Oxford University Press, Oxford; 1991
2. Barrat, A., Barthélemy, M., Vespignani, A.: *Dynamical Processes on Complex Networks*. Cambridge University Press, Cambridge; 2008
3. Banos, A., Lacasa, J.: Spatio-temporal exploration of SARS epidemic. *Cybergeo Eur. J. Geogr.* [online]; 2007. <http://cybergeo.revues.org/12803>
4. Barabási, A.L., Albert, R.: Emergence of scaling in random networks. *Science* **286**, 509–512 (1999)
5. Boccaletti, S., Latora, V., Moreno, Y., Chavez, M., Hwang, D.H.: Complex networks: structure and dynamics. *Phys. Rep.* **424**, 175–308 (2006)
6. Eubank, S., Guclu, H., Kumar, V.S.A., Marathe, M.V., Srinivasan, A., Toroczkai, Z., Wang, N.: Modelling disease outbreaks in realistic urban social networks. *Nature* **429**, 180–184 (2004)
7. Kermack, W.O., McKendrick, A.G.: A contribution to the mathematical theory of epidemics. *Proc. R. Soc. A* **115**, 700–721 (1927)
8. Liljeros, F., Edling, C.R., Amaral, L.A.N., Stanley, H.E., Åberg, Y.: The web of human sexual contacts. *Nature* **411**, 907–908 (2001)
9. Milgram, S.: The small world problem. *Psychol. Today* **2**, 60–67 (1967)
10. Morens, D.M., Folkers, G.K., Fauci, A.S.: The challenge of emerging and re-emerging infectious diseases. *Nature* **430**, 242–249 (2004)
11. Newman, M.E.J.: The structure and function of complex networks. *SIAM Rev.* **45**, 167–256 (2003)
12. Newman, M.E.J.: *Networks: An Introduction*. Oxford University Press, Oxford; 2010
13. *Oxford Atlas of World History*, Oxford University Press, Oxford, p. 105; 1999
14. Pastor-Satorras, R., Vespignani, A.: Epidemic spreading in scale-free networks. *Phys. Rev. Lett.* **86**, 3200–3203 (2001)
15. Wang, X.F., Chen, G.: Complex networks: small-world, scale-free and beyond. *IEEE Circuits Syst. Mag.* **3**, 6–20 (2003)
16. Wasserman, S., Faust, K.: *Social Network Analysis*. Cambridge University Press, Cambridge; 1994
17. Watts, D.J., Strogatz, S.H.: Collective dynamics of 'small-world' networks. *Nature* **393**, 440–442 (1998)

Author Biography



Carlo Piccardi is full professor of Systems and Control at Politecnico di Milano, where he teaches dynamics of complex systems and networks. His research activity is focused on modeling, control, and identification of dynamical systems, with applications in very diverse fields including vehicle dynamics, water resources, epidemic models, and social sciences and economics.