Dynamics of Epidemic Spreading in Complex Networks

Jorge Olivares Rodriguez

Abstract

Several mathematical models, mostly based on graph theory, have been developed over the past decades to predict and explain the behavior within complex interconnected dynamical systems. Here we present an introduction to the current theoretical models on epidemic spreading mechanisms in complex networks. We show using several mathematical models how the topology of the network determines the behavior of the epidemic. An analysis of the strengths and weaknesses of each model presented is also included. The study of the patterns emerging from epidemic spreading in complex networks can not only help prevent and contain the spread of computer viruses or biological epidemics, but may also help understand similar phenomena in other areas such as in social media.

1 Introduction

Some of the most challenging yet rich problems in physics are those concerned with understanding and modeling the behavior of many body systems. However, these systems are not confined to physical objects such as atoms, electrons or other particles and their constituents. Biological, social and technological systems are formed of a large number of particles (often quite different from each other) that interact between each other in complex ways. From these interactions along with the large populations forming these systems, emerge rich behaviors often with complicated underlying dynamics. An example of this and the main focus of this paper is the dynamics of epidemics and the emergence of outbreaks in complex networks. This topic has been widely studied for a long time particularly for its applicability on biological systems. However, epidemic phenomena is not limited to biological diseases, any networked system can experience it. For example social networks experience the spreading of rumors, while computer viruses or an internet meme can spread through technological networks.

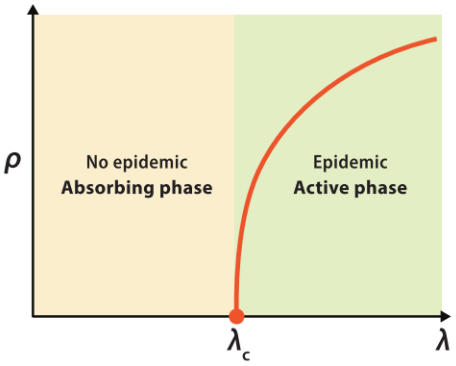


Figure 1: Order parameter and critical (transition) point indicating the onset of a sustained outbreak within a population [1]

Models for epidemic spreading have caught the attention of the statistical physics community due to the similarity with the models for non equilibrium problems in statistical physics [1]. The concept of *epidemic threashold*, or the point at which an outbreak occurs within a population resembles that of a phase transition. A control parameter λ is varied and the transition is characterized by an order parameter ρ, which has a non zero value in one phase (or state) and zero in another. Just as in statistical physics, the transition takes place at a particular value of the control parameter, the *transition point* λ_c.

2 Introduction to Networks and Graph Theory

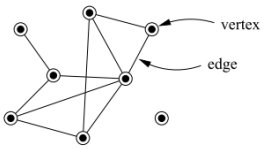


Figure 2: An example of a graph [2]

All of the systems mentioned in the introduction form networks. A network is a collection of items that interact with each other. Mathematically, networks are known as graphs and they are the subject of study of graph theory.[1]

2.1 Terminology in Graph Theory

Each element or item in the graph or network is called a *vertex* or *node* and the connections representing the interactions or relationships between them are known as edges, links or *ties*.[2] These are the most basic elements of a graph[2], but layers of complexity can be added by introducing rules and constraints to the interactions or by making the nodes distinguishable. By adding layers of complexity

to a graph one introduces new topological metrics, as in the case of weighed graphs where one can measure the strength of the node, $s_i = \sum_j w_{i,j}$. The *degree* of a node is the number of edges connected to it.[1, 2] In a undirected network, it can be represented by $k_i = \sum_j a_{i,j}$. In undirected networks, the degree distribution $P(k)$ is the probability that a randomly selected node has a degree k or for networks of finite size, this represents the fraction of nodes with degree k. Using this degree distribution one can then define the moments of the degree distribution as $\langle k^n \rangle = \sum_k P(k)k^n$. It is possible to have correlation between the different degrees of the nodes in the network. In this case, the correlation is denoted by $P(k'|k)$ which is the probability that an edge going from a node with degree k is connected to one with degree k'.[1] If the network is uncorrelated then $P(k'|k)$ is independent of the original vertex with degree k.[1]

2.2 Real Networks and Models

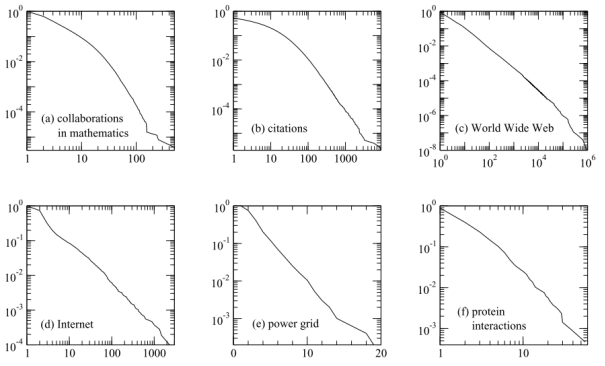


Figure 3: Degree (or connectivity) distribution function for various real networks [5]

The large size and complexity between the interactions taking place within real networks have made their study extremely challenging. This has made the topology of most networks for the most part not known. It was not until recently that scientists in a large variety of fields could began to test the mathematical and computational models they had created and developed over more than half a century based on short term observations on network dynamics. The recent availability and collection of data in large scales pertaining to the behavior of different physical, biological and social systems in recent years has allowed for the testing and refinement of these models.[2, 1, 3] Data collection on real networks has revealed non trivial topological information arising from their complex internal structure. [3] This is important since the topology of a graph directly influences its connectivity.[3] Here we present a brief description of three models that have been traditionally used to describe complex networks: Classical Random Graph theory(or Erdös-Rényi Model), the Small World Network theory (through the Watt-Strogartz model) and Scale Free Network theory (via the Barabási -Albert Model). The first two are shortly described but we will be focusing closely on the BA model as most real networks such as power grids and the internet have been shown to be best described by this scale free network theory as shown in figure 3.

2.2.1 The Random-Graph Theory (Erdös-Rényi Model)

Random graph theory, originally developed by Erdös and Rényi, is the oldest and the most exhaustively researched network models used to describe complex networks.[3]. Known also as *classical random graph* theory, in its simplest form, one builds a graph $G_p(N)$ from a set of N unlinked nodes, where each one of the total $M = N(N-1)/2$ possible edges can be

drawn with a probability p.[1] Note that by doing this the total number of edges, m, is now a random variable with an expectation value $\langle m \rangle = pM$.[5]Using this model we can then define the *ensemble* $G_p(N)$ of all graphs where a graph having m edges will appear with a probability $p^m(1-p)^{M-m}$.[2] In classic random graph theory, the number of ways in which k edges can be drawn from any given node follows a binomial distribution,

$$P(k = k_i) = C_{N-1}^{k} p^k (1-p)^{N-1-k} \tag{1}$$

where p^k and $(1-p)^{N-1-k}$ represent the probability of k edges and the absence of additional edges. The binomial coefficient represents all equivalent ways of selecting the k end points for these edges. Any two distinct nodes, i and j, $P(k_i = k)$ and $P(k_j = k)$ can be regarded as independent random variables[5]. This implies that the degree distribution of this network must be then related to the probability that these random variables take on a given value and therefore to the number of nodes with degree k. The average number of nodes with degree k is then

$$\lambda_k = NP(k_i = k) = N\binom{N-1}{k} p^k (1-p)^{N-1-k} \tag{2}$$

in the limit of large N[2]. In this same limit then the distribution of the number of nodes with degree k approaches a Poisson distribution $P(k) \approx e^{-\lambda_k} \frac{\lambda_k^k}{k!}$. The degree distribution then follows a Poisson distribution with mean value λ_k where this function is not constant[5]. However since the standard deviation of the distribution is $\sigma_k = \sqrt{\lambda_k}$, one can claim that the number of nodes with degree k does not diverge much from $\lambda_k = \langle k \rangle$ and then to a good approximation, in the large N limit $P(k) \approx e^{-\langle k \rangle} \frac{\langle k \rangle^k}{k!}$. The structure of the random graph varies with the value of p[2]. The clustering coefficient for this model is $\langle c \rangle = p$. This model is only useful for describing networks that are purely stochastic. Even though for $\langle k \rangle > 1$, the networks generated by classical random graphs are small in diameter much like real world networks. However this is where the similarities end as other important topological properties such as high clustering cannot be explained thorough this model [1].

2.2.2 Small World Networks and the Watts-Strogatz Model

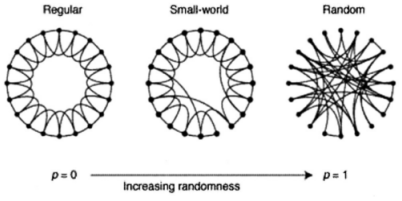


Figure 4: Increasing randomness as a function of p_{WS} in the WS model. [5]

The main goal of this model was to create a system that would transition from a highly ordered graph into a random network. Commonly known as the *small world network*[3], it is also the first attempt to obtain a network with a small diameter and large clustering coefficient[1], an improvement over the ER model. It consists of a one dimensional lattice of N nodes with periodic boundary conditions and where each node is bonded symmetrically to its neighbors(the starting highly ordered configuration)up to an order n[1] (nearest neighbor for $n = 1$, next nearest for $n = 2$

and so on and where the number $2n$ is known as the coordination number of a node [3]).Then, each bond is rewired with a probability p_{WS} to a randomly chosen node, with the only constraint that no two nodes can have more than one bond and that no node can bond to itself [3]. Effectively the rewiring adds long range shortcuts to the the network which reduces the average shortest path length preserving clustering [1]. For the extreme limits where $p_W S = 0$ the network is higly clustered and the average distance between any two nodes grows linearly with N[3], while for $p_W S = 1$, the system becomes a random graph, clustering is mostly gone and the average distance grows logarithmically with network size(see figure 4).[3] It was noted by WS that within the interval $0 < p_W S < 0.01$ the model exhibits the small world properties of high clustering and small diameter. Even though, one of the common short comings of the ER and WS models is that the probability of finding a highly connected node (large k) decreases exponentially with k, and hubs with large connectivity are practically non existent in these networks, a common characteristic in real networks.[4, 5] Nonetheless, this particular model is well suited for the study of social networks.[1]

2.2.3 Scale Free Networks and the Barabasi-Albert model

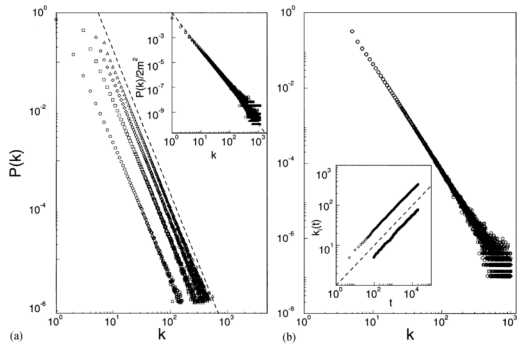


Figure 5: (a)Degree distribution $P(k)$ of the BA modle with $N = m_0 + t = 300000$ and $m = m_0$ for the values of 1(circles), 3 (squares), 5 (diamonds) and 7 (triangles). Dashed line represents $\gamma = 2.9$, and the inset shows the rescaled distribution $P(k)/2m^2$ for the same values of m with $\gamma = 3$. (b) Degree distribution for $m = m_0 = 5$ with system sizes N of 100000(circles), 150000 (squares) and 200000 (diamonds). Inset shows time evolution of two nodes for $t_1 = 5$ and $t_2 = 95$ where the dashed line has a slope of 0.5 [3]

One of the common topological properties in real networks is a degree distribution that is free of scale, i.e. a power law distribution over several orders of magnitude as it can be seen from figure 3.[1, 2, 5, 3, **?**]The last models have taken the approach of explaining the final form that networks take, trying to construct graphs with the appropriate topology . However, Barabasi and Albert realized that in order to develop a model that would be scale free one would have to focus on the *dynamics* of the network.[5, 4] They formulated two rules governing the evolution of the network from which the final topology of the network could be calculated.[5][4] They realized that both models considered a fixed number of nodes N that are either randomly connected or reconnected but without N changing, which is unrealistic, since most real networks are open and tend to grow in size.[4, 2, 1, 5] Secondly, the two previous network models assume that all connections (or rewiring) between the nodes are random and *independent* of the degree of the node.[3, 4, 5] Real networks, on the other hand, seem to exhibit *preferential attachment*[5], where new nodes have a higher probability of forming a link with existing nodes of higher degree. [3, 5, 4] Surprisinly, the introduction of only these two elements into the network

creates a power law degree distribution.[3, 5]. The Barabási and Albert (BA) model consists of the implementation of these two ingredients via the following algorithm:

1. Growth is incorporated by assuming that the network begins with a small number of nodes, m_0 and an additional node is added every time interval a new node is added with $m \leq m_0$ edges.

2. Preferential attachment is introduced by assuming that every time a new node is introduced it will attach to an existing node with a probability(Π) proportional to its degree k_i, that is[3, 5]

$$\Pi(k_i) = \frac{k_i}{\sum_j k_j} \tag{3}$$

Then the total number of nodes in the network after t steps have passed is $N = t + m_0$ with mt edges. Computer simulations have shown that this algorithm produces networks that are indeed scale invariant with a scaling exponent $\gamma = 2.9 \pm 0.1$, as shown in figure 5[3, 5].

In order to calculate analytically the time dependence of the connectivity, a mean field approach must be used. Assuming k is a continuous variable, the expression for preferential attachment can be seen as a rate of change for k_i, as the degree of any node will increase every time a new node enters the system and attaches to i, with a probability $\Pi(k_i)$. Therefore,

$$\frac{\partial k_i}{\partial t} = m\Pi(k_i) = m\frac{k_i}{\sum_j^{N-1} k_j} = m\frac{k_i}{2mt - m} = \frac{k_i}{2t} \tag{4}$$

In this case, $N = m_0 + t - 1$ and the change in connectivity at each time step is m due to the introduction of each node. Since every edge increases the connectivity of two nodes, after t time steps, $\sum_j^{N-1} k_j = 2mt$. Which can be solved using the additional constraint that every node k_i added at any time t_i begins with a connectivity $k_i(t_i) = m$, so that $k_i = m\left(\frac{t}{t_i}\right)^{1/2}$. Indicating that the degree of all the nodes evolve following the same power law. This also suggests that older nodes, with smaller t_i, in the network have larger connectivities than those introduced later on as we would expect, a phenomenon that many real networks experience [3, 4]. This allows us to relate the probabilities between finding a node with degree k in the network and the time steps taken by [3]

$$P[k_i(t) < k] = P\left[t_i > \frac{m^2}{k^2}t\right] \tag{5}$$

and from the assumption that nodes are added to the network at equal time intervals, the probability density of t_i is $P(t_i) = 1/(m_0 + t)$, and using the definition,

$$P\left[t_i > \frac{m^2}{k^2}t\right] = 1 - P\left[t_i \leq \frac{m^2}{k^2}t\right] = 1 - P(t_i)\frac{m^2}{k^2}t = 1 - \frac{m^2 t}{k^2(t + m_0)} \tag{6}$$

The probability density for the degree distribution is

$$P(k) = \frac{\partial P[k_i(t) < k]}{\partial k} = \frac{2m^2 t}{k^3(m_0 + t)} \tag{7}$$

where in the limit as $t \to \infty$,

$$P(k) \to \frac{2m^2}{k^3} \tag{8}$$

indicating that $\gamma = 3$ consistent with the numerical simulations.

3 The Epidemiological Process

Epidemiological models aim to predict and describe the dynamics of the contagion process within a population. Several have been developed and extensively studied to describe this process, each with added refinements to better reflect or constraint the spreading behavior. The most simple models usually assume that the total population of the network under study is held constant, ignoring all properties associated with the individuals composing the network (i.e. demographics).[1] Refinement in these models usually consists of the addition of phases or stages in which a particular member of the network (a node) can be in at any given time. The simplest models assume only two stages, healthy and infected. Amongst these are the Suceptible-Infected-Susceptible (SIS) model and the Suceptible-Infected-Removed (SIR) model.[2, 1] SIS for example has only two states, where an individual can either be infected with the disease or healthy and susceptible to become infected. The SIR model introduces an additional stage where the individual can be removed from the infection process once the node either recovers from being infected, developing an immunity to the disease or dying.[1]

3.1 Classical Epidemiology

As mentioned above, SIS and SIR are the most common models used to study the evolution of epidemics. In the SIR model, the amount of infected individuals, in the long time limit always tends to zero. However for SIS, depending on the initial conditions of the epidemic, an *endemic* or stationary state, characterized by a constant fraction of individuals may exist in this same limit.[1] Classically and in both models, the evolution of the epidemic is dictated by the infection and recovery processes[1]. In these models the transitions $I \to R$ and $I \to S$ are mostly time dependent and independent of the interactions with other individuals in the population. On the contrary, the transition $S \to I$ only occurs when a susceptible individual

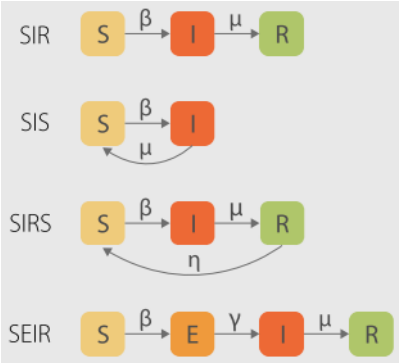


Figure 6: Examples of the different epidemiological models in use along with the transition rates between stages. [1]

comes into contact with an infectious one. Classical epidemiology takes the differences between these mechanisms into account along with simplifications that make epidemic modeling solvable.

Epidemiologist define the average time spend by an individual in the infected stage as $1/\mu$ where μ can also be interpreted as the recovery probability within some time interval. In the continuous approximation, and assuming a Poisson process, the probability that an individual remains in the infectious stage is $P_I(\tau) = \mu e^{-\mu\tau}$ with $\tau = 1/\mu$. Meanwhile, the transition $S \to I$ requires a more careful description. It is assumed that individuals within a particular population interact mostly randomly (*homogeneous mixing* approximation). In addition, the larger the number of infected individuals in contact with a susceptible one, the higher the probability of infection. This allows for the definition of the force of infection, or the probability (risk) that a susceptible individual will become infected at any time interval, $\alpha = \bar{\beta}\frac{N^I}{N}$, where $\bar{\beta}$ is specific to a disease and a populations contact pattern and $N^I/N = \rho^I$ is the fraction of infected individuals. Occasionally $\bar{\beta}$ can be written as $\bar{\beta} = \beta k$ where k is the average number of contacts with other individuals. In a mean field description, this is analogous to a mass action law, commonly used in epidemiology.[1] The force of infection is only dependent on the density of infected individuals, where

$$\left\langle \frac{d\rho}{dt} \right\rangle = f\rho^{\alpha} \tag{9}$$

that is, the average change of the density of individuals in a particular stage is equal to the force times the density of individuals in that stage. From the above, we could approach epidemic modeling as a *reaction diffusion process* where individuals are placed in different stages of the model at any given time and they can transition between them according to a specific set of rules. This is shown in figure 6,where in the continuous limit, each transition between stage is represented by a *rate*.[1] The allowed transitions in the SIS can be summarized by

$$S + I \longrightarrow 2I \tag{10}$$
$$I \longrightarrow S \tag{11}$$

where the first equation has an infection rate β and the second a recovery rate μ. Meanwhile, for the SIR model,

$$S + I \longrightarrow 2I \tag{12}$$
$$I \longrightarrow R \tag{13}$$

Often stages are added to these simple models in order to reflect real world behaviors. In the case of biological diseases, for example, one can add an additional susceptible stage to SIR as a way of incorporating temporal immunity to a particular pathogen, where a new

variable η is introduced, indicating the rate at which immunity is lost. Other examples include the SEIR model where the additional stage includes exposed individuals E which may be infected but not yet able to infect others.[1]

Classic understanding of epidemics is based on the use of difference equations that describe the movement of individuals between stages in the continuous time limit.The *homogeneous mixing approximation* (HMA) is employed, where the assumptions that individuals interact randomly between each other and are well mixed are made. Well mixing of individuals allows us to average over their particular demographics and treat everyone indistinguishably from each other as long as they belong within a particular stage of the model. Under this approximation, the epidemic state is encoded in the number or density of individuals at a particular stage α, $\rho^\alpha = N^\alpha/N$.For both SIS and SIR[1],

$$\frac{d\rho^I}{dt} = (\beta\rho^S - \mu)\rho^I \qquad\qquad \frac{d\rho^S}{dt} = (-\beta\rho^S + \chi)\rho^I \tag{14}$$

where χ is either μ or zero for SIS and SIR respectively. These systems of equations are subject to the normalization condition $\sum_\alpha \rho^\alpha = 1$. These equations can all be solved if they are linearized in the limit when $\rho^I \approx 0$, valid at the early stages of an epidemic, where

$$\rho^S = 1 - \rho^I \approx 1 \qquad\qquad \frac{d\rho^I}{dt} = (\beta\rho^S - \mu)\rho^I \approx (\beta - \mu)\rho^I \tag{15}$$

where the solution $\rho^I(t) = \rho(0)e^{(\beta-\mu)t}$ represents the behavior at the beginning of contagion. The number of infected individuals will only grow exponentially if $\beta - \mu > 0$ or

$$R_0 = \frac{\beta}{\mu} > 1 \tag{16}$$

where R_0 is known as the *basic reproduction number*[1], describing the average number of secondary infections caused by a primary case. Only if $R_0 > 1$ outbreaks occur. For $R_0 < 1$, in the thermodynamic limit of infinite population, the infection vanishes in the case of SIR or all individuals become healthy again in SIS. So far the models have been deterministic and the stochastic nature of disease spreading has not been taken into account. Here, $R_0 > 1$ is necessary and sufficient for an epidemic outbreak to occur, but in stochastic systems it is only necessary.[1]Models that take this stoichastic nature into account are considered next,but the focus will be models with Poisson like processes. It is worth mentioning that more complex behaviors have been explored by using other non-poissonian distributions[1]. Furthermore, the classical deterministic model assumes random and homogeneous mixing, but in real networks, the interactions between individuals vary. Every node in the network has a different number of edges. For example, more social individuals will have greater contact with others and have a higher chance of becoming infected. R_0 in classical epidemiological theory does not seem to be illustrate this dependence on the network heterogeneity. These heterogeneous connectivity patters are a consequence of the network topology and in turn affect epidemic behavior.

4 The SIR Model

The SIR model is well suited to describe a large number of epidemic phenomena, from the spread of computer viruses, disease spreding with adquired immunity after infection, rumors in social circles etc.[1] This model will be then one of the main focus of the following sections. In addition, we will define the densities with the following variables

$$S(t) = \rho^S(t) \qquad I(t) = \rho^I(t) \qquad R(t) = \rho^R(t) \tag{17}$$

The order parameter (or prevalence) in the SIR model is defined as the number of removed individuals at the end of the epidemic, that is $R_\infty = \lim_{t\to\infty} R(t)$[1]. The differential equations that describe this model are

$$\frac{dS}{dt} = -\lambda\bar{k}I(t)S(t) \qquad \frac{dI}{dt} = -\mu I(t) + \lambda\bar{k}I(t)S(t) \qquad \frac{dR}{dt} = \mu I(t) \tag{18}$$

where λ is the microscopic infection rate $S \to I$ and $\bar{k}$ is the rate of contacts for all individuals assumed to be constant. Using the initial conditions $R(0) = 0, I(0) \approx 0$ and $S(0) \approx 1$, and combining the differential equations for the time derivative of R and S,

$$\dot{S} = -\lambda\bar{k}S\dot{I} \qquad \Rightarrow S(t) = e^{-\lambda\bar{k}R(t)}$$

Using this result and the normalization condition,

$$\dot{R}(t) = I(t) = (1 - R(t) - S(t)) = 1 - R(t) - e^{-\lambda\bar{k}R(t)} \tag{19}$$

$$\lim_{t\to\infty} \dot{R(t)} = 0 = \lim_{t\to\infty}(1 - R(t) - e^{-\lambda\bar{k}R(t)}) \Rightarrow R_\infty = 1 - e^{-\lambda\bar{k}R_\infty} \tag{20}$$

where we let the order parameter converge to a constant in the long time limit. This last expression is denoted as the self consistent equation for the order parameter [6]. $R_\infty = 0$ is always a solution (which we already knew), but we are more interested in non zero solutions. Since $0 \leq R_\infty \leq 1$ and noting that for large values of R_∞ the right hand side of the self consistent equation tends to 1, a non zero crossing point will exists is the slope of this function exceeds the slope of the left hand side at zero, as this is just a linear function that always increases. This constraint can be expressed as

$$\frac{d}{dR_\infty}\left[1 - e^{-\lambda\bar{k}R(t)}\right]_{R_\infty=0} > 1 \Rightarrow \lambda > \frac{1}{\bar{k}} \tag{21}$$

This is equivalent to the condition $\lambda > \lambda_c$ for $\lambda_c = 1/\bar{k}$

4.1 The SIR Model in Complex Networks

To introduce contact heterogeneity in the network, while using the SIR model, we will assume that the network has a general degree distribution $P(k)$, that is the network is uncorrelated.[6] A more complex but accurate description would use the degree correlation $P(k'|k)$. In this picture, the densities are now also functions of the degree of the node and described by the variables $S_k(t), I_k(t)$ and $R_k(t)$, with normalization $S_k(t) + I_k(t) + R_k(t) = 1$[1]. And since the probability that a node is connected to another node of degree l has a probability $lP(l)$, we introduce the factor

$$\Theta(t) = \frac{\sum_k kP(k)I_k(t)}{\sum_j jP(j)} = \frac{\sum_k kP(k)I_k(t)}{\langle k \rangle} \tag{22}$$

which takes into account the probability that an edge in the network is connected to an infected site.[1, 6] Then the densities must obey the differential equations

$$\dot{S_k} = -\lambda k S_k(t)\Theta(t) \qquad \dot{I_k} = -I_k(t) + \lambda k I_k(t) S_k(t)\Theta(t) \qquad \frac{dR}{dt} = I_k(t) \tag{23}$$

Using the initial conditions $R_k(0) = 0, I_k(0) = I^0$ and $S_k(0) = 1 - I_k(0)$ and in the limit $I^0 << 1$, then $I_k(0) \approx 0$ and $S_k(0) \approx 1$, and by integrating the equation for the susceptible density,

$$S_k(t) = e^{-\lambda k \phi(t)} \tag{24}$$

$$\phi(t) = \int_0^t dt' \Theta(t') = \int_0^t dt' \frac{\sum_k kP(k)I_k(t)}{\langle k \rangle} = \frac{1}{\langle k \rangle} \sum_k kP(k) \int_0^t dt' \frac{dR_k(t')}{dt'} \tag{25}$$

$$= \frac{1}{\langle k \rangle} \sum_k kP(k)R_k(t) \tag{26}$$

The function ϕ cannot be solved analitically in general[1]. However we can obtain useful information by examining its time evolution and behavior in the long time limit[1][6]. The order parameter in the long time limit is

$$R_\infty = \sum_k P(k)(1 - e^{-\lambda k \phi_\infty}) \tag{27}$$

derived just as equation 19. Meanwhile for the time derivative of ϕ is

$$\dot{\phi(t)} = \frac{1}{\langle k \rangle} \sum_k kP(k)\dot{R}_k(t) = \frac{1}{\langle k \rangle} \sum_k kP(k)I_k(t) = \frac{1}{\langle k \rangle} \sum_k kP(k)(1 - R(t) - e^{-\lambda k \phi(t)}) \tag{28}$$

$$= \frac{1}{\langle k \rangle} (\sum_k kP(k) - \sum_k kP(k)R(t) - \sum_k kP(k)e^{-\lambda k \phi(t)}) \tag{29}$$

$$= 1 - \phi(t) - \frac{1}{\langle k \rangle} \sum_k kP(k)e^{-\lambda k \phi(t)} \tag{30}$$

Under the assumption that in the long time limit $\phi(t)$ converges to a constant, and $\lim_{t\to\infty}\phi(t) = \phi_\infty$, then

$$\phi_\infty = 1 - \frac{1}{\langle k\rangle}\sum_k kP(k)e^{-\lambda k\phi_\infty} \tag{31}$$

This is the self consistency equation for ϕ_∞ where the value of zero is always a solution. In analogy with equation 21. A non zero solution is met by [6]

$$\frac{d}{d\phi_\infty}\left[1 - \frac{1}{\langle k\rangle}\sum_k kP(k)e^{-\lambda k\phi_\infty}\right]_{R_\infty=0} > 1 \qquad \Rightarrow \frac{1}{\langle k\rangle}\sum_k kP(k)(\lambda k) = \lambda\frac{\langle k^2\rangle}{\langle k\rangle} \tag{32}$$

which allows us to define the epidemic threshold at $\lambda_c = \frac{\langle k\rangle}{\langle k^2\rangle}$.

4.1.1 SIR model in Watts-Strogatz Model

For a WS network with $p = 1$, each node has at least m neighbors, the connectivity distribution, for $k \geq m$ is given by $P(k) = \frac{m^{k-m}}{(k-m)!}e^{-m}$.[6] For networks whose moments are all convergent, equations 27 and 28 can both be solved in the limit where $\phi(t)$ is small. In the case of equation 27,

$$R_\infty = -\sum_{k,n=1} P(k)\frac{(-\lambda k\phi_\infty)^n}{n!} = -\sum_{k,n=0} P(k)\frac{(-\lambda k\phi_\infty)^{(n+1)}}{(n+1)!} \approx \lambda\phi_\infty\sum_k kP(k) \tag{33}$$

$$\approx \lambda\phi_\infty\langle k\rangle \tag{34}$$

while for equation 28,

$$\dot{\phi(t)} = -\phi(t) - \frac{1}{\langle k\rangle}\sum_{k,n=1} kP(k)\frac{(-\lambda k\phi_\infty)^n}{n!} = -\phi(t) - \frac{1}{\langle k\rangle}\sum_{k,n=0} k^{n+2}P(k)\frac{(-\lambda\phi_\infty)^{n+1}}{(n+1)!} \tag{35}$$

$$= \phi(t)\left(-1 - \frac{\lambda\langle k^{n+2}\rangle}{\langle k\rangle}\sum_{n=0}\frac{(-\lambda\phi_\infty)^n}{(n+1)!}\right) \approx \phi(t)\left(-1 + \frac{\lambda\langle k^2\rangle}{\langle k\rangle} - \frac{\lambda^2\langle k^3\rangle}{\langle k\rangle}\right) \tag{36}$$

This can be integrated to give [6]

$$\phi(t) = \frac{2(\lambda-\lambda_c)\langle k\rangle}{\langle k^3\rangle\lambda^2 + Ae^{-(\lambda-\lambda_c)t/\lambda_c}} \tag{37}$$

for A being an integration constant and $\lambda_c = \langle k\rangle/\langle k^2\rangle$. And using the degree distribution for WS networks [6],

$$\lambda_c = \frac{2}{1+4m} \tag{38}$$

4.1.2 SIR in Power Law Distributed Networks

From the degree distribution for BA networks, equation 8, the first and second moments are

$$\langle k \rangle = 2m^2 \int_m^\infty k k^{-3} = 2m \quad \langle k^2 \rangle = 2m^2 \int_m^\infty k^2 k^{-3} = 2m^2 \lim_{k_{max} \to \infty} \ln(k_{max}/m) \to \infty \quad (39)$$

Then the epidemic threshold tends to zero always. In the long time limit,

$$R_\infty = 1 - 2m^2 \int_m^\infty k^{-3} e^{-\lambda k \phi_\infty} \approx 2\lambda m \phi_\infty \quad (40)$$

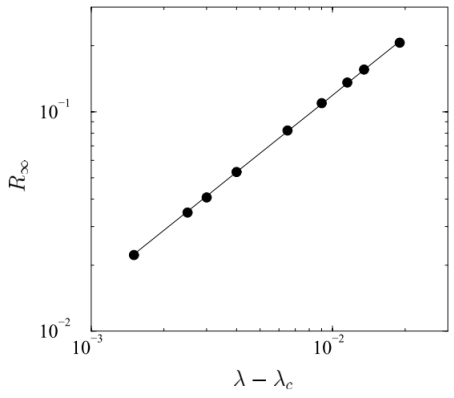
where the integral can be expressed in terms of the incomplete gamma function and expanded for small ϕ_∞.[6] In addition, one can solve the expression for the time derivative of ϕ for small $\phi(t)$ by using incomplete gamma functions to find that in the long time limit and where γ is the Euler Mascheroni constant, $\phi_\infty \approx \frac{e^{1-\gamma}}{\lambda m} e^{-\frac{1}{\lambda m}}$. Using this result, the epidemic prevalence

$$R_\infty \propto e^{1/\lambda m} \quad (41)$$

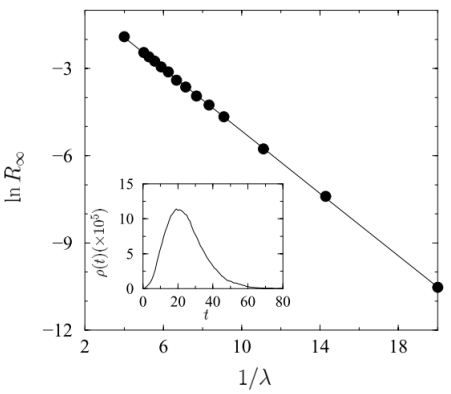
where we see that R_∞ is never zero for $\lambda \neq 0$.

5 Results and Final Remarks

The last two sections lead us to conclude that, under the same model, the topology of the network has a strong influence on the behavior of the spread of contagion. Particularly in the determination of the the critical value at which an outbreak will occur. The analytic results were tested through numerical simulations by Moreno *et al.*[6]. The results of their simulations are shown in figure 7, where the theoretical predictions seem to agree with the simulations. In the case of small world networks, where connectivity fluctuations are small and finite, the epidemic threshold is also finite. However in the case of scale free networks where the fluctuations in connectivity tend to diverge, the epidemic threshold vanishes, making them prone to constant outbreaks, at least in the thermodynamic limit. Large fluctuations on the connectivity of a network can enhance epidemic spreading on a network. This is reasonable as we would expect that for a scale free highly heterogeneous network, an infection beginning at any node will tend to cause an outbreak since low connectivity points are likely connected to high connectivity ones providing a pathway for it to occur. These results were the product of the addition of layers of complexity to the classical theory of epidemiology. If other refinements are included, such as the use of the correlation function $P(k'|k)$ or by studying systems of finite size, we may find a richer topological dependence of the order parameter and critical point. These considerations, if solvable, may lead to a more accurate description of epidemic dynamics, and ultimately to a practical strategy to contain them or even control them.



(a) Order parameter R_∞ as a function of $(\lambda - \lambda_c)$ in WS network with $N = 10^6$. The value $\lambda_c = 0.184(5)$ and $\beta = 0.9(1)$ where obtained from the numerical calculations for the line fit of the form $R_\infty \propto (\lambda - \lambda_c)^\beta$



(b) Order parameter R_∞ as a function of $1/\lambda$ in BA network with $N = 10^6$. The semilogarithmic plot of the numerical results vs the line fit validates the analytic result.

Figure 7: Results of the comparison between the numerical simulations and the analytic solution performed on the SIR model[6]

References

[1] Romualdo Pastor-Satorras, Claudio Castellano, Piet Van Mieghem, and Alessandro Vespignani. Epidemic processes in complex networks. *Rev. Mod. Phys.*, 87:925–979, Aug 2015.

[2] M. E. J. Newman. The structure and function of complex networks. *SIAM Review*, 45(2):167–256, 2003.

[3] Albert-Lszl Barabsi, Rka Albert, and Hawoong Jeong. Mean-field theory for scale-free random networks. *Physica A: Statistical Mechanics and its Applications*, 272(1):173 – 187, 1999.

[4] Albert-László Barabási and Réka Albert. Emergence of scaling in random networks. *Science*, 286(5439):509–512, 1999.

[5] Réka Albert and Albert-László Barabási. Statistical mechanics of complex networks. *Rev. Mod. Phys.*, 74:47–97, Jan 2002.

[6] Y. Moreno, R. Pastor-Satorras, and A. Vespignani. Epidemic outbreaks in complex heterogeneous networks. *The European Physical Journal B - Condensed Matter and Complex Systems*, 26(4):521–529, Apr 2002.