Chapter 1
Epidemic Models: Their Spread, Analysis and Invasions in Scale-Free Networks

Somnath Tagore

Abstract The mission of this chapter is to introduce the concept of epidemic outbursts in network structures, especially in case of scale-free networks. The invasion phenomena of epidemics have been of tremendous interest among the scientific community over many years, due to its large scale implementation in real world networks. This chapter seeks to make readers understand the critical issues involved in epidemics such as propagation, spread and their combat which can be further used to design synthetic and robust network architectures. The primary concern in this chapter focuses on the concept of Susceptible-Infectious-Recovered (SIR) and Susceptible-Infectious-Susceptible (SIS) models with their implementation in scale-free networks, followed by developing strategies for identifying the damage caused in the network. The relevance of this chapter can be understood when methods discussed in this chapter could be related to contemporary networks for improving their performance in terms of robustness. The patterns by which epidemics spread through groups are determined by the properties of the pathogen carrying it, length of its infectious period, its severity as well as by network structures within the population. Thus, accurately modeling the underlying network is crucial to understand the spread as well as prevention of an epidemic. Moreover, implementing immunization strategies helps control and terminate theses epidemics.

1.1 Scale-Free Networks

The degree distribution of individuals is one of the most standard and efficient network measures that is existent today. In most of the synthetic as well as practical networks, many individuals have lesser number of connected neighbours than others.

S. Tagore (✉)
Department of Biotechnology and Bioinformatics,
Padmashree Dr. D.Y. Patil University, Navi Mumbai 400614, India
e-mail: somnathtagore@yahoo.co.in

D. Król et al. (eds.), *Propagation Phenomena in Real World Networks*,
Intelligent Systems Reference Library 85, DOI 10.1007/978-3-319-15916-4_1

For instance, random networks, small worlds display lesser variation in terms of neighbourhood sizes, whereas spatial networks have Poisson-like degree distributions. Moreover, as highly connected individuals are of more importance considering disease transmission, incorporating them into the current network is of outmost importance [4]. This is essential in case of capturing the complexities of disease spread. Architecturally, scale-free networks are heterogenous in nature and can be dynamically constructed by adding new individuals to the current network structure one at a time. This strategy is similar to naturally forming links, especially in case of social networks. Moreover, the newly connected nodes or individuals link to the already existent ones (with larger connections) in a manner that is preferential in nature. This connectivity can be understood by a power-law plot with the number of contacts per individual, a property which is regularly observed in case of several other networks like that of power grids, world-wide-web, to name a few [14].

Epidemiologists have worked hard on understanding the heterogeneity of scale-free networks for populations for a long time. Highly connected individuals as well as hub participants have played essential roles in the spread and maintenance of infections and diseases. Figure 1.1 illustrates the architecture of a system consisting of a population of individuals. It has several essential components, namely, nodes, links, newly connected nodes, hubs and sub-groups respectively. Here, nodes correspond to individuals and their relations are shown as links. Similarly, newly connected nodes correspond to those which are recently added to the network, such as initiation of new relations between already existing and unknown individuals [24]. Hubs are

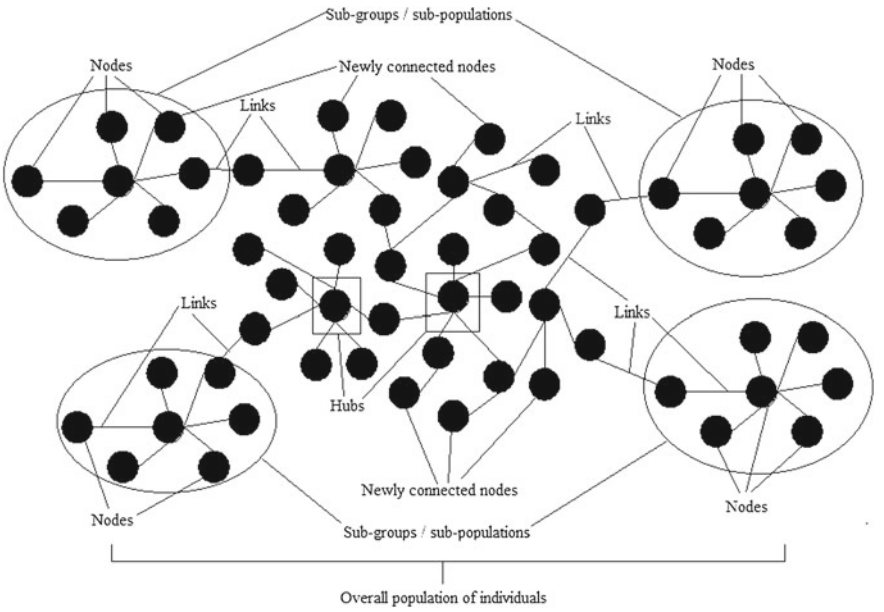

Fig. 1.1 A synthetic scale-free network and its characteristics

those nodes which are highly connected, such as individuals who are very popular among others and have many relations and/or friends. Lastly, sub-groups correspond to certain sections of the population which have individuals with closely associated relationships, such as group of nodes which are highly dense in nature, or having high clustering coefficient. Furthermore, it is important in having large number of contacts as the individuals are at greater risk of infection and, once infected, can transmit it to others. For instance, hub individuals of such high-risk individuals help in maintaining sexually transmitted diseases (STDs) in different populations where majority belong to long-term monogamous relationships, whereas in case of SARS epidemic, a significant proportion of all infections are due to high risk connected individuals. Furthermore, the preferential attachment model proposed by Barabási and Albert [4] defined the existence of individuals of having large connectivity does not require random vaccination for preventing epidemics. Moreover, if there is an upper limit on the connectivity of individuals, random immunization can be performed to control infection.

Likewise, the dynamics of infectious diseases has been extensively studied in case of scale-free as well as small-world and random networks. In small-world networks, most of the nodes may not be direct neighbors, but can be reached from all other nodes via less number of hops, that are number of nodes between start and terminating nodes. Also, in these networks distance, *dist*, between two random nodes increases proportionally to the logarithm of the number of nodes, *tot*, in the network [15], i.e.,

$$dist \propto \log tot \tag{1.1}$$

Watts and Strogatz [24] identified a class of small-world networks and categorized them as random graphs. These were classified on the basis of two independent features, namely, average shortest path length and clustering coefficient. As per Erdős-Rényi model, random graphs have a smaller average shortest path length and small clustering coefficient. Watts and Strogatz on the other hand demonstrated that various real-world networks have a smaller average shortest path length along with high clustering coefficient greater than expected randomly. It has been observed that it is difficult to block and/or terminate an epidemic in scale-free networks with slow tails. It has especially been seen in case the network correlations among infections and individuals are absent. Another reason for this effect is the presence of hubs, where infections could be sustained and reduced by target-specific selections [17].

1.1.1 Power-Law

It has been well known that real-world networks ranging from social to computers are scale-free in nature, whose degree distribution follows an asymptotic power-law. These are characterized by degree distribution following a power law,

$$P(conn) \approx conn^{-\eta} \tag{1.2}$$

for the number of connections, *conn* for individuals and η is an exponent. Barabási and Albert [4] analyzed the topology of a portion of the world-wide-web and identified 'hubs'. The terminals had larger number of connections than others and the whole network followed a power-law distribution. They also found that these networks have heavy-tailed degree distributions and thus termed them as 'scale-free'. Likewise, models for epidemic spread in static heavy-tailed networks have illustrated that with a degree distribution having moments resulted in lesser prevalence and/or termination for smaller rates of infection [14]. Moreover, beyond a particular threshold, this prevalence turns to non-zero. Similarly, it has been seen that for networks following power-law,

$$moment > \eta - 1 \tag{1.3}$$

does not exist and the prevalence is non-zero for any infection rates. Due to this reason, epidemics are difficult to handle and terminate in static networks having power-law degree distributions. Figure 1.2 illustrates a power-law plot between $P(conn)$

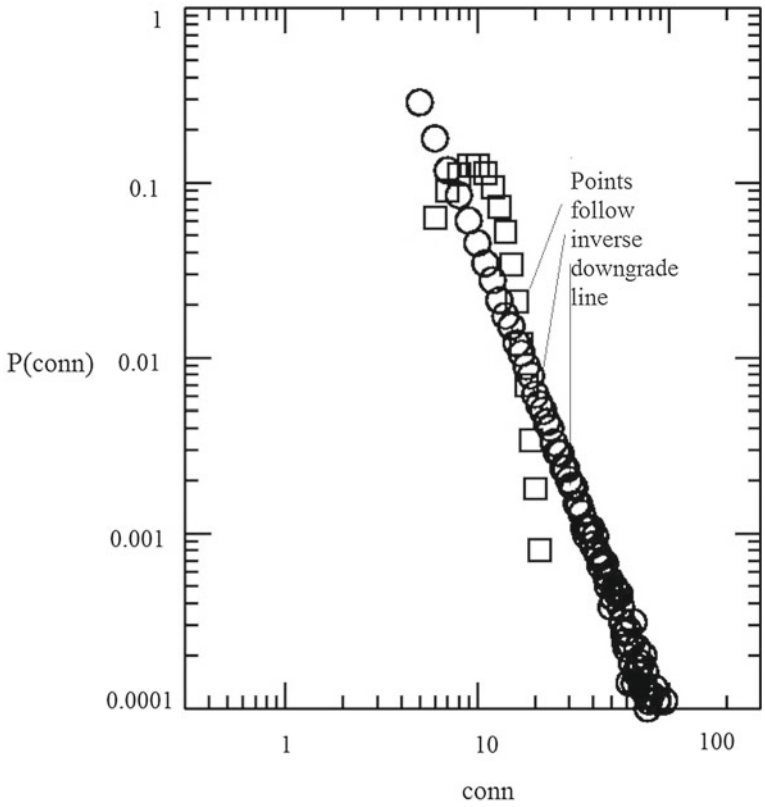

Fig. 1.2 Power-law curve illustrating $P(conn)$ versus *conn* in log-log scale

versus *conn* in log-log scale. It shows that points in this figure follow a inverse downgrade line in log-log scale, satisfying 'scale-free' behavior.

Likewise, in various instances, networks are not static but dynamic (i.e., they evolve in time) via some rewiring processes, in which edges are detached and reattached according to some dynamic rule. Steady states of rewiring networks have been studied in the past. More often, it has been observed that depending on the average connectivity and rewiring rates, networks reach a scale-free steady state, with an exponent, η, represented using dynamical rates [17].

1.2 Epidemics

The study of epidemics has always been of interest in areas where biological applications coincide with social issues. For instance, epidemics like influenza, measles, and STDs, can pass through large group of individuals, populations, and/or persist over longer timescales at low levels. These might even experience sudden changes of increasing and decreasing prevalence. Furthermore, in some cases, single infection outbreaks may have significant effects on a complete population group [1].

Epidemic spreading can also occur on complex networks with vertices representing individuals and the links representing interactions among individuals. Thus, spreading of diseases can occur over the network of individuals as spreading of computer viruses occur over the world-wide-web. The underlying network in epidemic models is considered to be static while the individual states vary from infected to non-infected individuals according to certain probabilistic rules. Furthermore, the evolution of an infected group of individuals in time can be studied by focusing on the average density of infected individuals in steady state. Lastly, the spread as well as growth of epidemics can also be monitored by studying the architecture of the network of individuals as well as its statistical properties [2].

1.2.1 Branching

One of the essential properties of epidemic spread is its branching pattern, thereby infecting healthy individuals over a time period. This branching pattern of epidemic progression can be classified on the basis of their infection initiation, spread and further spread (Fig. 1.3) [5].

1. Infection initiation: If an infected individual comes in contact with a group of individuals, the infection is transmitted to each with a probability p, independent of one another. Furthermore, if the same individual meets k others while being infected, these k individuals form the infected set. Due to this random disease transmission from the initially infected individual, those directly connected to it get infected.

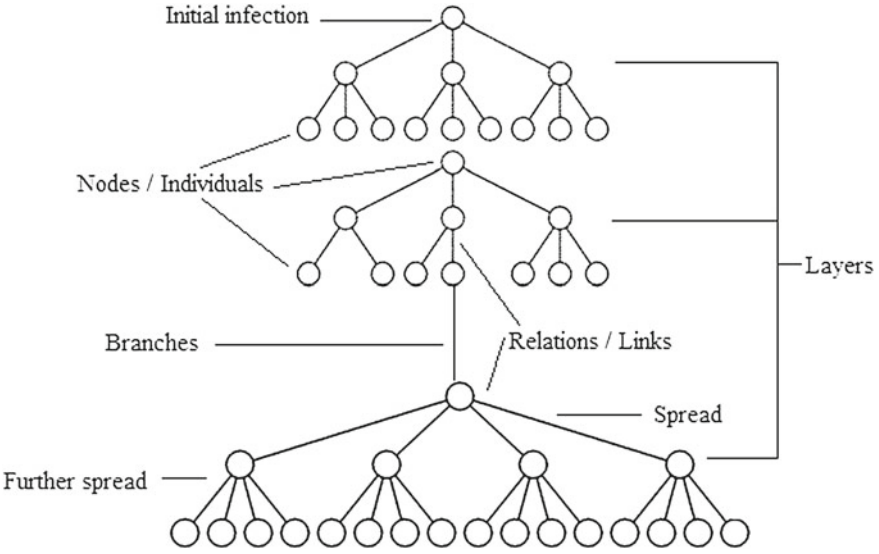

Fig. 1.3 Branching modes and patterns in epidemic progression

2. Spread: Every individual in the original infected set meets k other individuals, which results in k^2 individuals.
3. Further spread: The infection spreads further with each individual in the present infected set connecting to k healthy individuals with a probability p independent of individual infection.

1.2.1.1 Reproductive Number

If infection in a branching process reaches an individual set and fails to infect healthy individuals, then termination of the infection occurs, which leads to no further progression and infection of other healthy individuals. Thus, there may be two possibilities for an infection in a branching process model. Either it reaches a site infecting no further and terminating out, or it continues to infect healthy individuals through contact processes. The quantity which can be used to identify whether an infection persist or fades out is defined as *basic reproductive number* [6].

This basic reproductive number, τ, is the expected number of newly infected individuals caused by a single already infected individual. In case where every individual meets k new people and infects each with probability p, the basic reproductive number is represented as

$$\tau = pk \tag{1.4}$$

It is quite essential as it helps in identifying whether or not an infection can spread through a population of healthy individuals. The concept of τ was first proposed by Alfred Lotka, and applied in the area of epidemiology by MacDonald [13].

For non-complex population models, τ can be identified if information for 'death rate' is present. Thus, considering death rate, d, and birth rate, b, at the same time,

$$\tau = \frac{b}{d} \tag{1.5}$$

Moreover, τ can also be used to determine whether an infection will terminate, i.e., $\tau < 1$ or it becomes an epidemic, i.e., $\tau > 1$. But, it cannot be used for comparing different infections at the same time on the basis of multiple parameters. Several methods, such as identifying eigenvalues, Jacobian matrix, birth rate, equilibrium states, population statistics can well be used to analyze and handle τ [18].

1.2.1.2 Branching Models

There are some standard branching models that are existent for analyzing the progress of infection in a healthy population or network. The first one, *Reed-Frost model*, considers a homogeneous close set consisting of total number of individuals, *tot*. Let *num* designate the number of individuals susceptible to infection at time $t = 0$ and m_{num} the number of individuals infected by the infection at any time t [19]. Here,

$$num + m_{num} = tot \tag{1.6}$$

$$m_{num} = num \tag{1.7}$$

Here, Eq. 1.7 is in case of a smaller population. It is assumed that an individual x is infected at time t, whereas any individual y comes in contact with x with a probability $\frac{a}{num}$, where $a > 0$. Likewise, if y is susceptible to infection then it becomes infected at time $t + 1$ and x is removed from the population (Fig. 1.4a). In this figure, x or $v_1(*)$ represents the infection start site, $y(v_3)$, v_2 are individuals that are susceptible to infection, $num = 0$, $tot = 11$, and $m_{num} = 1$.

The second one, *3-clique model* constructs a 3-clique sub-network randomly by assigning a set of *tot* individuals. Here, for individual/vertex pair (v_i, v_j) with probability p_1, the pair is included along with vertices triples (v_i, v_j, v_k) with probability p_2. Thus, the corresponding pairs (v_i, v_j), (v_j, v_k) and (v_k, v_i) are also included. This creates a network

$$G = g_1 \bigcup g_2 \tag{1.8}$$

Here, g_1, g_2 are two independent graphs, where g_1 is a Bernoulli graph with edge probability p_1 and g_2 with all possible triangles existing independently with a probability p_2 (Fig. 1.4b). In this figure, $g_1 = (v_1, v_2, v_3)$, $g_2 = (v_4, v_5, v_6)$, $g_3 = (v_7, v_8, v_9)$ are the three 3-clique sub-networks with $tot = 9$, and $G = g_1 \bigcup g_2 \bigcup g_3$ respectively [21].

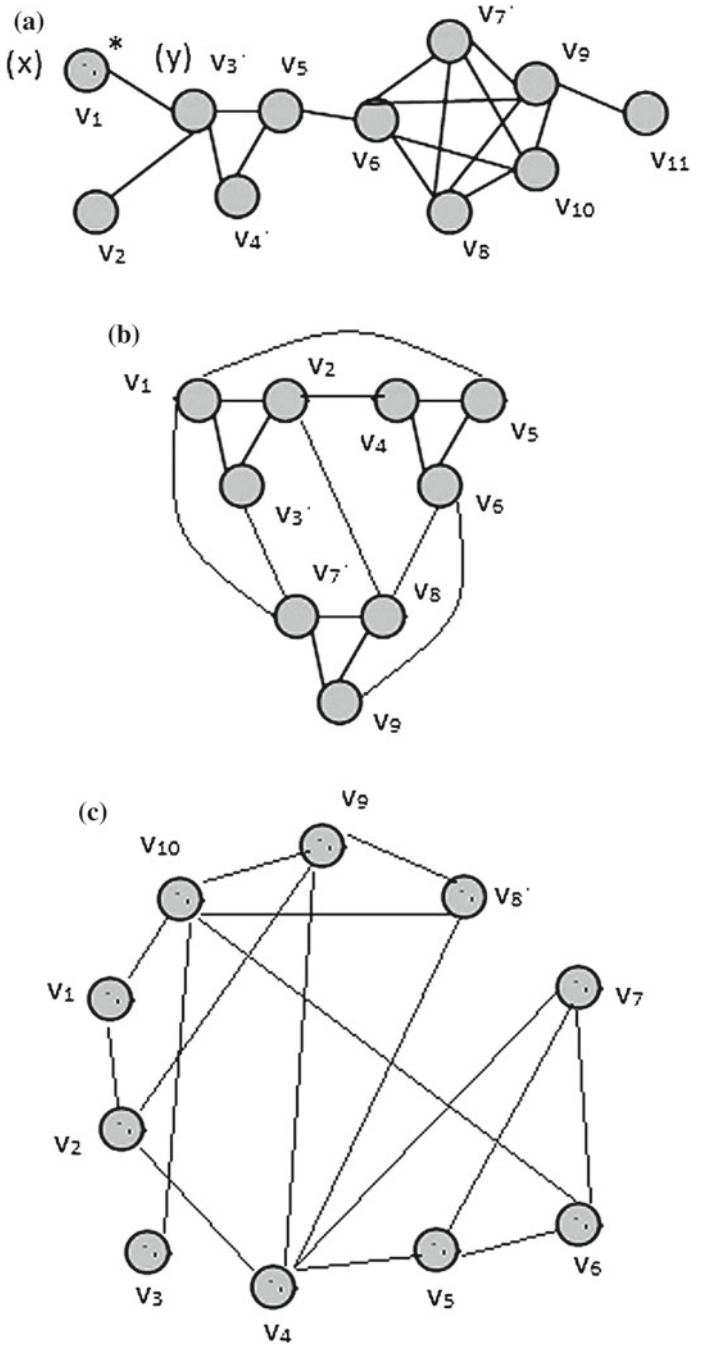

Fig. 1.4 Types of branching models illustrated in synthetic networks: **a** *Reed-Frost*, **b** *3-clique*, **c** *Household*

The third one, *Household model* assumes that for a given a set of *tot* individuals or vertices, g_1 is a Bernoulli graph consisting of $\frac{tot}{b}$ disjoint $b-$cliques, where $b \ll tot$ with edge probability p_2. Thus, the network G is formed as the superposition of the graphs g_1 and g_2, i.e., $G = g_1 \bigcup g_2$. Moreover, g_1 fragments the population into mutually exclusive groups whereas g_2 describes the relations among individuals in the population. Thus, g_1 does not allow any infection spread, as there are no connections between the groups. But, when the relationship structure g_2 is added, the groups are linked together and the infection can now spread using relationship connections (Fig. 1.4c). In this figure, $tot = 10$ where the individuals (v_1 to v_{10}) are linked on the basis of randomly assigned p_2 and $b = 4 \ll tot = 10$.

1.3 Network Architectures

The interconnected architecture of various networks have been of primary interest to researchers in various scientific areas. In interconnected networks, failure in vertex links in one network can cause failure of dependent vertices in other networks. This results in cascading failures. Similarly, in case of networks without dependencies among vertices, the level of information flow between the interconnected vertices affects the epidemic transition on subset levels. Furthermore, percolation threshold in interacting networks are lower than in single networks, with the appearance of a giant component in certain cases (Fig. 1.5). A giant component is a connected sub-graph of a random graph containing a constant fraction of total vertices of the entire graph. These are extremely prominent in Erdős-Rényi graphs, where each edge connecting vertex pairs for a set of n vertices remains independently of one another with a probability p. Here, if $p \leq \frac{1-\varepsilon}{n}$ for any constant $\varepsilon > 0$, then all the connected components have size $O(\log n)$, and giant component is absent. But, for $p \geq \frac{1+\varepsilon}{n}$ a single giant component may reside. Figure 1.5a–d illustrate the formation of a giant component in a random graph with $p = 0.002, 0.006, 0.009$ in Fig. 1.5b–d respectively [23].

Thus, it is essential to identify the conditions which results in an epidemic spread in one network, with the presence of minimal isolated infections on other network components. Moreover, depending on the parameters of individual sub-networks and their internal connectivities, connecting them to one another creates marginal effect on the spread of epidemic. Thus, identifying these conditions resulting in analyzing spread of epidemic process is very essential. In this case, two different interconnected network modules can be determined, namely, strongly and weakly coupled. In the strongly coupled one, all modules are simultaneously either infection free or part of an epidemic, whereas in the weakly coupled one a new mixed phase exists, where the infection is epidemic on only one module, and not in others [25].

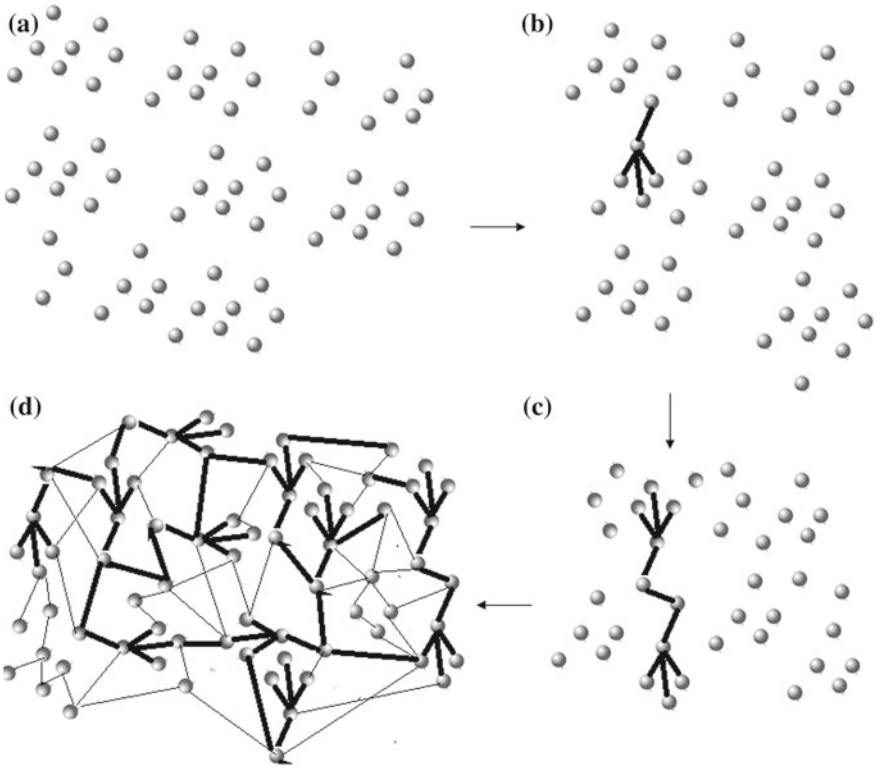

Fig. 1.5 Emergence of giant component in an interconnected network. **a** Original network, **b** emergence of giant component, **c** further emergence, and **d** final architecture

1.3.1 Concurrency

Generally, epidemic models consider contact networks to be static in nature, where all links are existent throughout the infection course. Moreover, a property of infection is that these are contagious and spread at a rate faster than the initially infected contact. But, in cases like HIV, which spreads through a population over longer time scales, the course of infection spread is heavily dependent on the properties of the contact individuals. The reason for this being, certain individuals may have lesser contacts at any single point in time and their identities can shift significantly with the infection progress [25].

Thus, for modeling the contact network in such infections, transient contacts are considered which may not last through the whole epidemic course, but only for particular amount of time. In such cases, it is assumed that the contact links are undirected. Furthermore, different individual timings do not affect those having potential to spread an infection but the timing pattern also influences the severity of the overall epidemic spread. Similarly, individuals may also be involved in

concurrent partnerships having two or more actively involved ones that overlap in time. Thus, the concurrent pattern causes the infection to circulate vigorously through the network [22].

1.4 Propagation Phenomena in Real World Networks

In the last decade, considerable amount of work has been done in characterizing as well as analyzing and understanding the topological properties of networks. It has been established that scale-free behavior is one of the most fundamental concepts for understanding the organization various real-world networks. This scale-free property has a resounding effect on all aspect of dynamic processes in the network, which includes percolation. Likewise, for a wide range of scale-free networks, epidemic threshold is not existent, and infections with low spreading rate prevail over the entire population [10]. Furthermore, properties of networks such as topological fractality etc. correlate to many aspects of the network structure and function. Also, some of the recent developments have shown that the correlation between degree and betweenness centrality of individuals is extremely weak in fractal network models in comparison with non-fractal models [20].

Likewise, it is seen that fractal scale-free networks are dis-assortative, making such scale-free networks more robust against targeted perturbations on hubs nodes. Moreover, one can also relate fractality to infection dynamics in case of specifically designed deterministic networks. Deterministic networks allow computing functional, structural as well as topological properties. Similarly, in case of complex networks, determination of topological characteristics has shown that these are scale-free as well as highly clustered, but do not display small-world features. Also, by mapping a standard Susceptible, Infected, Recovered (SIR) model to a percolation problem, one can also find that there exists certain finite epidemic threshold. In certain cases, the transmission rate needs to exceed a critical value for the infection to spread and prevail. This also specifies that the fractal networks are robust to infections [11]. Meanwhile, scale-free networks exhibit various essential characteristics such as power-law degree distribution, large clustering coefficient, large-world phenomenon, to name a few [16].

1.5 Network Definition and Measurement

Network analysis can be used to describe the evolution and spread of information in the populations along with understanding their internal dynamics and architecture. Specifically, importance should be given to the nature of connections, and whether a relationship between x and y individuals provide a relationship between y and x as well. Likewise, this information could be further utilized for identifying transitivity-based measures of cohesion (Fig. 1.6).

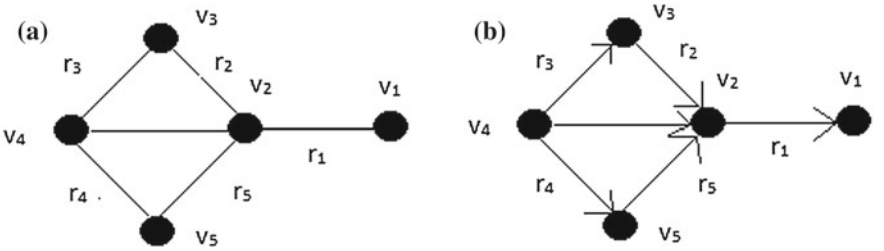

Fig. 1.6 Architectural properties in a hypothetical. **a** Undirected, **b** directed network

Meanwhile, research in networks also provide some quantitative tools for describing and characterizing networks. *Degree* of a vertex is the number of connectivities for each vertex in the form of links. For instance, $degree(v_4) = 3$, $degree(v_2) = 4$ (for undirected graph (Fig. 1.6a)). Similarly for Fig. 1.6b, $degree_{in}(v_2) = 3$ (number of incoming links), $degree_{out}(v_2) = 1$ (number of outgoing links). *Clustering coefficient (CC)* of a vertex is the compactness of the network, i.e., $CC(v_i) = \frac{2*link}{degree(degree-1)}$, where *degree* = degree of vertex v_i, *link* = number of links among neighbors of v_i. For instance, in Fig. 1.6a, $CC(v_2) = 0.33$, $CC(v_4) = 0.6$, etc. Likewise, *Shortest path* is the minimum number of links that needs to be parsed for traveling between two vertices. For instance, in Fig. 1.6a, shortest path between v_4 and $v_1 = (v_4, v_2, v_1)$. *Diameter* of network is the maximum distance between any two vertices or the longest of the shortest walks. Thus, in Fig. 1.6b, from v_4, one has (v_4, v_3, v_2, v_1), (v_4, v_2, v_1), (v_4, v_5, v_2, v_1), from v_3, we have $(v_3, v_4, v_5, v_2, v_1)$, (v_3, v_4, v_2, v_1), (v_3, v_2, v_1), from v_5, we have $(v_5, v_4, v_3, v_2, v_1)$, (v_5, v_4, v_2, v_1), (v_5, v_2, v_1). Out of these the longest of the shortest walks = $(v_3, v_4, v_5, v_2, v_1)$, $(v_5, v_4, v_3, v_2, v_1) = 4$. Thus, diameter = 4 [15].

Radius of a network is the minimum eccentricity (eccentricity of a vertex v_i is the greatest geodesic distance), i.e., distance between two vertices in a network is the number of edges in a shortest path connecting them between v_i and any other vertex of any vertex. For instance, in Fig. 1.6b, radius of network = 2. *Betweenness centrality* ($g(v_i)$) is equal to the number of shortest paths from all vertices to all others that pass through vertex v_i, i.e.,

$$g(v_i) = \frac{v_x v_y(v_i)}{v_x v_y} \tag{1.9}$$

where $v_x v_y$ is total number of shortest paths from vertex v_x to vertex v_y and $v_x v_y(v_i)$ is the number of those paths that pass through v_i. Thus, in Fig. 1.6b, $g(v_4) = 0.77$. Similarly, *Closeness centrality* ($c(v_i)$) of a vertex v_i describes the total distance of v_i to all other vertices in the network, i.e., sum the shortest paths of v_i to all other vertices in the network. For instance, in Fig. 1.6b, $c(v_4) = (v_4, v_3, v_2, v_1) + (v_4, v_2, v_1) + (v_4, v_5, v_2, v_1) = 8$. Lastly, *Stress centrality* ($s(v_i)$) is the simple accumulation of

the number of shortest paths between all vertex pairs, sometimes interchangeable with betweenness centrality [14].

Use of 'adjacency matrix', $A_{v_i v_j}$, describing the connections within a population is also persistent. Likewise, various network quantities can be ascertained from the adjacency matrix. For instance, size of a population is defined as the average number of contacts per individual, i.e.,

$$\overline{num} = \frac{1}{tot} \sum_{v_i v_j} A_{v_i v_j} \tag{1.10}$$

The powers of adjacency matrix can be used to calculate measures of transitivity [14].

1.5.1 Data Collection Process

One of the key pre-requisites of network analysis is initial data collection. For performing a complete mixing network analysis for individuals residing in a population, every relationship information is essential. This data provides great difficulty in handling the entire population, as well as handling complicated network evaluation issues. The reason being, individuals have contacts, and recall problems are quite probable. Moreover, evaluation of contacts requires certain information which may not always be readily present. Likewise, in case of epidemiological networks, connections are included if they explain relationships capable of permitting the transfer of infection. But, in most of the cases, clarity of defining such relations is absent. Thus, various types of relationships bestow risks and judgments that needs to be sorted for understanding likely transmission routes. One can also consider weighted networks in which links are not merely present or absent but are given scores or weights according to their strength [9].

Furthermore, different infections are passed by different routes, and a mixing network is infection specific. For instance, a network used in HIV transmission is different from the one used to examine influenza. Similarly, in case of airborne infections like influenza and measles, various networks need to be considered because differing levels of interaction are required to constitute a contact. The problems with network definition and measurement imply that any mixing networks that are obtained will depend on the assumptions and protocols of the data collection process.

Three main standard techniques can be employed to gather such information, namely, *infection searching*, *complete contact searching* and *diary-based studies* [9].

1.5.1.1 Infection Searching

After an epidemic spread, major emphasis is laid on determining the source and spread of infection. Thus, each infected individual is linked to one other from

whom infection is spread as well as from whom the infection is transmitted. As all connections represent actual transmission events, infection searching methods do not suffer from problems with the link definition, but interactions not responsible for this infection transmission are removed. Thus, the networks observed are of closed architecture, without any loops, walks, cliques and complete sub-graphs [15].

Infection searching is a preliminary method for infectious diseases with low prevalence. These can also be simulated using several mathematical techniques based on differential equations, control theories etc., assuming a homogeneous mixing of population. It can also be simulated in a manner so that infected individuals are identified and cured at a rate proportional to the number of neighbors it has, analogous to the infection process. But, it does not allow to compare various infection searching budgets and thus a discrete-event simulation need to be undertaken. Moreover, a number of studies have shown that analyses based on realistic models of disease transmission in healthy networks yields significant projections of infection spread than projections created using compartmental models [8]. Furthermore, depending on the number of contacts for any infected individuals, their susceptible neighbors are traced and removed. This is followed by identifying infection searching techniques that yields different numbers of newly infected individuals on the spread of the disease.

1.5.1.2 Complete Contact Searching

Contact searching identifies potential transmission contacts from an initially infected individual by revealing some new individual set who are prone to infection and can be subject of further searching effort. Nevertheless, it suffers from network definition issues; is time consuming and depends on complete information about individuals and their relationships. It has been used as a control strategy, in case of STDs. Its main objective of contact searching is identifying asymptomatically infected individuals who are either treated or quarantined.

Complete contact searching deals with identifying the susceptible and/or infected individuals of already infected ones and conducting simulations and/or testing them for degree of infection spread, treating them as well as searching their neighbors for immunization. For instance, STDs have been found to be difficult for immunization. The reason being, these have specifically long asymptomatic periods, during which the virus can replicate and the infection is transmitted to healthy, closely related neighbors. This is rapidly followed by severe effects, ultimately leading to the termination of the affected individual. Likewise, recognizing these infections as global epidemic has led to the development of treatments that allow them to be managed by suppressing the replication of the infection for as long as possible. Thus, complete contact searching act as an essential strategy even in case when the infection seems incurable [7].

1.5.1.3 Diary-Based Studies

Diary-based studies consider individuals recording contacts as they occur and allow a larger number of individuals to be sampled in detail. Thus, this variation from the population approach of other tracing methods to the individual-level scale is possible. But, this approach suffers from several disadvantages. For instance, the data collection is at the discretion of the subjects and is difficult for researchers to link this information into a comprehensive network, as the individual identifies contacts that are not uniquely recorded [3].

Diary-based studies require the individuals to be part of some coherent group, residing in small communities. Also, it is quite probable that this kind of a study may result in a large number of disconnected sub-groups, with each of them representing some locally connected set of individuals. Diary-based studies can be beneficial in case of identifying infected and susceptible individuals as well as the degree of infectivity. These also provide a comprehensive network for diseases that spread by point-to-point contact and can be used to investigate the patterns infection spread.

1.6 Robustness

Robustness is an essential connectivity property of power-law graph. It defines that power-law graphs are robust under random attack, but vulnerable under targeted attack. Recent studies have shown that the robustness of power-law graph under random and targeted attacks are simulated displaying that power-law graphs are very robust under random errors but vulnerable when a small fraction of high degree vertices or links are removed. Furthermore, some studies have also shown that if vertices are deleted at random, then as long as any positive proportion remains, the graph induced on the remaining vertices has a component of order of the total number of vertices [15].

Many a times it can be observed that a network of individuals may be subject to sudden change in the internal and/or external environment, due to some perturbation events. For this reason, a balance needs to be maintained against perturbations while being adaptable in the presence of changes, a property known as robustness. Studies on the topological and functional properties of such networks have achieved some progress, but still have limited understanding of their robustness. Furthermore, more important a path is, higher is the chance to have a backup path. Thus, removing a link or an individual from any sub-network may also lead to blocking the information flow within that sub-network. The robustness of a model can also be assessed by means of altering the various parameters and components associated with forming a particular link. Robustness of a network can also be studied with respect to ‘resilience’, a method of analyzing the sensitivities of internal constituents under external perturbation, that may be random or targeted in nature [18].

1.7 Models of Infections

Basic disease models discuss the number of individuals in a population that are susceptible, infected and/or recovered from a particular infection. For this purpose, various differential equation based models have been used to simulate the events of action during the infection spread. In this scenario, various details of the infection progression are neglected, along with the difference in response between individuals. Models of infections can be categorized as SIR and Susceptible, Infected, Susceptible (SIS) [9].

1.7.1 Susceptible-Infected-Recovered (SIR)

The SIR model considers individuals to have long-lasting immunity, and divides the population into those susceptible to the disease (S), infected (I) and recovered (R). Thus, the total number of individuals (T) considered in the population is

$$T = S + I + R \tag{1.11}$$

the transition rate from S to I is κ and the recovery rate from I to R is ρ. Thus, the SIR model can be represented as

$$\frac{dS}{dT} = \gamma(T - S) - \kappa \frac{I}{T} S \tag{1.12}$$

$$\frac{dI}{dT} = \kappa \frac{I}{T} S - (\gamma + \rho) I \tag{1.13}$$

$$\frac{dR}{dT} = \rho I - \lambda R \tag{1.14}$$

Likewise, the reproductivity (θ) of an infection can be identified as the average number of secondary instances a typical single infected instance will cause in a population with no immunity. It determines whether infections spreads through a population; if $\theta < 1$, the infection terminates in the long run; $\theta > 1$, the infection spreads in a population. Larger the value of θ, more difficult is to control the epidemic [12].

Furthermore, the proportion of the population that needs to be immunized can be calculated by

$$\theta = \frac{\kappa}{\gamma + \rho} \tag{1.15}$$

Similarly, for $S(0)$, $I(0)$, $R(0)$, and $\theta <= 1$,

$$\lim_{t \to \infty} (S(t), I(t), R(t)) \to (T, 0, 0) \tag{1.16}$$

known as disease free stability, whereas if $\theta > 1$ and $I(0) > 0$, then

$$\lim_{t\to\infty}(S(t), I(t), R(t)) \to (\frac{T}{\theta}, \frac{\gamma T}{\kappa}(\theta - 1), \frac{\rho T}{\kappa}(\theta - 1)) \tag{1.17}$$

known as endemic stability can be identified. Depending upon these instances, immunization strategies can be initiated [6].

1.7.1.1 Extensions to SIR Model

Although the contact network in a general SIR model can be arbitrarily complex, the infection dynamics can still being studied as well as modeled in a simple fashion. Contagion probabilities are set to a uniform value, i.e., p, and contagiousness has a kind of 'on-off' property, i.e., an individual is equally contagious for each of the t_I steps while it has the infection, where 1 is present state of the system. One can extend the idea that contagion is more likely between certain pairs of individuals or vertices by assigning a separate probability p_{v_i,v_j} to each pair of individuals or vertices v_i and v_j, for which v_i is linked to v_j in a directed contact network.

Likewise, other extensions of the contact model involves separating the I state into a sequence of early, middle, and late periods of the infection. For instance, it could be used to model an infection with a high contagious incubation period, followed by a less contagious period while symptoms are being expressed [16].

1.7.1.2 Percolations of SIR Model

In most of the cases, SIR epidemics are thought of dynamic processes, in which the network state evolves step-by-step over time. It captures the temporal dynamics of the infection as it spreads through a population. The SIR model has been found to be suitable for infections, which provides lifelong immunity, like measles. In this case, a property termed as the force of infection is existent, a function of the number of infectious individuals is. It also contains information about the interactions between individuals that lead to the transmission of infection.

One can also have a static view of the epidemics where SIR model for $t_I = 1$. This means that considering a point in an SIR epidemic when a vertex v_i has just become infectious, has one chance to infect v_j (since $t_I = 1$), with probability p. One can visualize the outcome of this probabilistic process and also assume that for each edge in the contact network, a probability signifying the relationship is identified. Furthermore, one can also use the open and blocked healthy edges to represent the course of the infection spread. A vertex v_i will become infected during the epidemic if and only if there is a path to v_i from one of the initially infected nodes that consists entirely of open edges [3].

1.7.2 Susceptible-Infected-Susceptible (SIS)

The SIS model can be represented as

$$\frac{dS}{dT} = \rho I - \kappa S \tag{1.18}$$

$$\frac{dI}{dT} = \kappa S - \rho I \tag{1.19}$$

Removed state is absent in this case. Moreover, after a vertex is over with the Infectious state, it reverts back to the Susceptible state and is ready to initiate the infection again. Due to this alternation between the S and I states, the model is referred to as SIS model. The mechanics of SIS model can be discussed as follows [2].

1. At the initial stage, some vertices remain in I state and all others are in S state.
2. Each vertex v_i that enters the I state and remains infected for a certain number of steps t_I.
3. During each of these t_I steps, v_i has a probability p of passing the infection to each of its susceptible directly linked neighbors.
4. After t_I steps, v_i no longer remains infected, and returns back to the S state.

The SIS model is predominantly used for simulating and understanding the progress of STDs, where repeat infections are existent, like gonorrhoea. Moreover, certain assumptions with regard to random mixing between individuals within each pair of sub-networks are present. In this scenario, the number of neighbors for each individual is considerably smaller than the total population size. Such models generally avoid random-mixing assumptions thereby assigning each individual to a specific set of contacts that they can infect.

1.7.2.1 Life Cycle of SIS

An SIS epidemic, can run for long time duration as it can cycle through the vertices multiple number of times. If at any time during the SIS epidemic all vertices are simultaneously free of the infection, then the epidemic terminates forever. The reason being, no infected individuals exist that can pass the infection to others. In case if the network is finite in nature, a stage would arise when all attempts for further infection of healthy individuals would simultaneously fail for t_I steps in a row.

Likewise, for contact networks where the structure is mathematically tractable, a particular critical value of the contagion probability p is existent, an SIS epidemic undergoes a rapid shift from one that terminates out quickly to one that persists for a long time. In this case, the critical value of the contagion probability depends on the structure of the problem set [1].

1.8 Epidemic Invasions, Propagations and Outbursts

The patterns by which epidemics spread through vertex groups is determined by the properties of the pathogen, length of its infectious period, severity and the network structures. The path for an infection spread are given by a population state, with existence of direct contacts between the individuals or vertices. The functioning of network system depends on the nature of interaction between their individuals. This is essentially because of the effect of infection-causing individuals and topology of networks. To analyze the complexity of epidemics, it is important to understand the underlying principles of its distribution in the history of its existence. In recent years it has been seen that the study of disease dynamics in social networks is relevant with the spread of viruses and the nature of diseases [9].

Moreover, the pathogen and the network are closely intertwined with even within the same group of individuals, the contact networks for two different infections are different structures. This depends on respective modes of transmission of infections. For instance, a highly contagious infection, involving airborne transmission, the contact network includes a huge number of links, including any pair of individuals that are in contact with one another. Likewise, for an infection requiring close contact, the contact network is much sparser, with fewer pairs of individuals connected by links [7].

1.9 Combat and Immunization

Immunization is a site percolation problem where each immunized individual is considered to be a site which is removed from the infected network. Its aim is to transfer the percolation threshold that leads to minimization of the number of infected individuals. The model of SIR and immunization is regarded as a site-bond percolation model, and immunization is considered successful if the infected a network is below a predefined percolation threshold. Furthermore, immunizing randomly selected individuals requires targeting a large fraction, *frac*, of the entire population. For instance, some infections require 80–100 % immunization. Meanwhile, target-based immunization of the hubs requires global information about the network in question, rendering it impractical in many cases, which is very difficult in certain cases [6].

Likewise, social networks possess a broad distribution of the number of links, *conn*, connecting individuals and analyzing them illustrate that that a large fraction, *frac*, of the individuals need to be immunized before the integrity of the infected network is compromised. This is essentially true for scale-free networks, where $P(conn) \approx conn^{-\eta}$, $2 < \eta < 3$, where the network remains connected even after removal of most of its individuals or vertices. In this scenario, a random immunization strategy requires that most of the individuals need to be immunized before an epidemic is terminated [8].

For various infections, it may be difficult to reach a critical level of immunization for terminating the infection. In this case, each individual that is immunized is given immunity against the infection, but also provides protection to other healthy individuals within the population. Based on the SIR model, one can only achieve half of the critical immunization level which reduces the level of infection in the population by half. A crucial property of immunization is that these strategies are not perfect and being immunized does not always confer immunity. In this case, the critical threshold applies to a portion of the total population that needs to be immunized. For instance, if the immunization fails to generate immunity in a portion, *por*, of those immunized, then to achieve immunity one needs to immunize a portion

$$Im = \frac{\tau - 1}{\tau(1 - por)} \tag{1.20}$$

Here, *Im* denotes immunity strength. Thus, in case if *por* is huge it is difficult to remove infection using this strategy or provides partial immunity. It may also invoke in various manners: the immunization reduces the susceptibility of an individual to a particular infection, may reduce subsequent transmission if the individual becomes infected, or it may increase recovery.

Such immunization strategies require the immunized individuals to become infected and shift into a separate infected group, after which the critical immunization threshold (S_I) needs to be established. Thus, if *CIL* is the number of secondary infected individuals affected by an initial infectious individual, then

$$CIL = \frac{\tau - 1}{\tau - S_I} \tag{1.21}$$

Thus, S_I needs to be less than one, else it is not possible to remove the infection. But, one also needs to note that an immunization works equally efficiently if it reduces the transmission or susceptibility and increases the recovery rate. Moreover, when the immunization strategy fails to generate any protection in a proportion *por* of those immunized, the rest $1 - por$ are fully protected. In this scenario, it can be not possible to remove the infection using random immunization. Thus, targeted immunization provides better protection than random-based [13].

1.9.1 Complex Topologies and Heterogeneous Structures

In case of homogenous networks, the average degree, $\overline{conn}$, fluctuates less and can assume $conn \simeq \overline{conn}$, i.e., the number of links are approximately equal to average degree. However, networks can also be heterogeneous. Likewise, in a homogeneous network such as a random graph, $P(conn)$ decays faster exponentially whereas for heterogenous networks it decays as a power law for large *conn*.

The effect of heterogeneity on epidemic behavior studied in details for many years for scale-free networks. These studies are mainly concerned with the stationary

limit and existence of an endemic phase. An essential result of this analysis is the expression of basic reproductive number which in this case is $\tau \infty \frac{\overline{conn^2}}{\overline{conn}}$. Here, τ is proportional to the second moment of degree, which finally diverges for increasing network sizes [15].

1.9.2 Damage Patterns

It has been noticed that the degree of interconnection in between individuals for all form of networks is quite unprecedented. Whereas, interconnection increases the spread of information in social networks, another exhaustively studied area contributes to the spread of infection throughout the healthy network. This rapid spreading is done due to less stringency of its passage through the network. Moreover, initial sickness nature and time of infection are unavailable most of the time, and the only available information is related to the evolution of the sick-reporting process. Thus, given complete knowledge of the network topology, the objective is to determine if the infection is an epidemic, or if individuals have become infected via an independent infection mechanism that is external to the network, and not propagated through the connected links.

If one considers a computer network undergoing cascading failures due to worm propagation whereas random failures due to misconfiguration independent of infected nodes, there are two possible causes of the sickness, namely, *random* and *infectious spread*. In case of *random* sickness, infection spreads randomly and uniformly over the network where the network plays no role in spreading the infection; and *infectious spread*, where the infection is caused through a contagion that spreads through the network, with individual nodes being infected by direct neighbors with a certain probability [6].

1.9.2.1 Random Sickness

In random damage, each individual becomes infected with an independent probability ψ_1. At time t, each infected individual reports damage with an independent probability ψ_2. Thus, on an average, a fraction ψ of the network reports being infected, where

$$\psi = \psi_1.\psi_2 \tag{1.22}$$

It is already known that social networks possess a broad distribution of the number of links, k, originating from an individual. Computer networks, both physical and logical are also known to possess wide, scale-free, distributions. Studies of percolation on broad-scale networks display that a large fraction, fc, of the individuals need to be immunized before the integrity of the network is compromised. This is particularly true for scale-free networks, where the percolation threshold tends to 1, and the network remains contagious even after removal of most of its infected individuals [9].

1.9.2.2 Infection Spread

When the hub individuals are targeted first, removal of just a fraction of these results in the breakdown of the network. This has led to the suggestion of targeted immunization of hubs. To implement this approach, the number for connections of each individual needs to be known. During infection spread, at time 0, a randomly selected individual in the network becomes infected. When a healthy individual becomes infected, a time is set for each outgoing link to an adjacent individual that is not infected, with expiration time exponentially distributed with unit average. Upon expiration of a link's time, the corresponding individual becomes infected, and in-turn begins infecting its neighbors [7].

1.9.3 Immunity

In general, for an epidemic to occur in a susceptible population the basic reproductive rate must be greater than 1. In many circumstances not all contacts will be susceptible to infection. In this case, some contacts remain immune, due to prior infection which may have conferred life-long immunity, or due to some previous immunization. Therefore, not all individuals are infected and the average number of secondary infections decrease. Similarly, the epidemic threshold in this case is the number of susceptible individuals within a population that is required for an epidemic to occur. Similarly, the herd immunity is the proportion of population immune to a particular infection. If this is achieved due to immunization, then each case leads to a new case and the infection becomes more stable within the population [6].

One of the simplest immunization procedure consists of random introduction of immune individuals in the population for achieving uniform immunization density. In this case, for a fixed spreading rate, ξ, the relevant control parameter in the density of immune individuals present in the network, the immunity, *imm*. At the mean-field level, the presence of a uniform immunity reduces ξ by a factor $1 - imm$, i.e., the probability of identifying and infecting a susceptible and non-immune individual becomes $\xi(1-imm)$. For homogeneous networks, one observes that, for aconstant ξ, the stationary prevalence is given by

$$\rho_{imm} = 0 \tag{1.23}$$

for $imm > imm_c$ and

$$\rho_{imm} = (imm_c - imm)/(1 - imm) \tag{1.24}$$

for $imm \leq imm_c$ Here imm_c is the critical immunization value above which the density of infected individuals in the stationary state is null and depends on ξ as $imm_c = 1 - \frac{\xi_c}{\xi}$.

Thus, for a uniform immunization level larger than imm_c, the network is completely protected and no large epidemic outbreaks are possible. On the contrary, uniform immunization strategies on scale-free heterogenous networks are totally ineffective. The presence of uniform immunization elocally depresses the infections prevalence for any value of ξ, and it is difficult to identify any critical fraction of immunized individuals that ensures the eradication of infection [2].

1.10 Understanding Cascading Failures, Natural Disturbances

Cascading, or epidemic processes are those where the actions, infections or failure of certain individuals increase the susceptibility of others. This results in the successive spread of infections from a small set of initially infected individuals to a larger set. Initially developed as a way to study human disease propagation, cascades ares useful models in a wide range of application. The vast majority of work on cascading processes focused on understanding how the graph structure of the network affects the spread of cascades. One can also focus on several critical issues for understanding the cascading features in network for which studying the architecture of the network is crucial [5].

The standard independent cascade epidemic model assumes that the network is directed graph $G = (V, E)$, for every directed edge between v_i, v_j, we say v_i is a parent and v_j is a child of the corresponding other vertex. Parent may infect child along an edge, but the reverse cannot happen. Let V denote the set of parents of each vertex v_i, and for convenience $v_i \in V$ is included. Epidemics proceed in discrete time where all vertices are initially in the susceptible state. At time 0, each vertex independently becomes active, with probability p_{init}. This set of initially active vertices are called 'seeds'. In each time step, the active vertices probabilistically infects its susceptible children; if vertex v_i is active at time t, it infects each susceptible child v_j with probability $p_{v_i v_j}$, independently. Correspondingly, a vertex v_j susceptible at time t becomes active in the next time step, i.e., at time $t + 1$, if any one of its parents infects it. Finally, a vertex remains active for only one time slot, after which it becomes inactive and does not spread the infection further as well as cannot be infected again either [5]. Thus, in this kind of an SIR epidemic, where some vertices remain forever susceptible because the epidemic never reaches them, while others transition, susceptible $\rightarrow$ active for one time step $\rightarrow$ inactive.

1.11 Conclusions

In this chapter, we discussed some critical issues regarding epidemics and their outbursts in static as well as dynamic network structures. We mainly focused on SIR and SIS models as well as identifying key strategies for identifying the

damage caused in networks. We also discussed the various modeling techniques for studying cascading failures. Epidemics pass through populations and persists over long time periods. Thus, efficient modeling of the underlying network plays a crucial role in understanding the spread and prevention of an epidemic. Social, biological, and communication systems can be explained as complex networks with their degree distribution follows a power law, $P(conn) \approx conn^{-\eta}$, for the number of connections, *conn* for individuals, representing scale-free (SF) networks. We also discussed certain issues on epidemic spreading in SF networks characterized by complex topologies with basic epidemic models describing the proportion of individuals susceptible, infected and recovered from a particular disease. Likewise, we also explained the significance of the basic reproduction rate of an infection, that can be identified as the average number of secondary instances a typical single infected instance will cause in a population with no immunity. Also, we explained how determining the complete nature of a network required knowledge of every individual in a population and their relationships as, the problems with network definition and measurement depend on the assumptions of data collection processes. Nevertheless, we also illustrated the importance of invasion resistance methods, with temporary immunity generating oscillations in localized parts of the network, with certain patches following large numbers of infections in concentrated areas. Similarly, we also explained the significance of damages, namely, random, where the damage spreads randomly and uniformly over the network and in particular the network plays no role in spreading the damage; and infectious spread, where the damage spreads through the network, with one node infecting others with some probability.

References

1. Anderson, R., May, R.: Infectious Diseases of Humans: Dynamics and Control. Oxford University Press, Oxford (1991)
2. Bailey, N.: The Mathematical Theory of Infectious Diseases and its Applications. Hafner Press, New York (1975)
3. Bak, P., Chen, K., Tang, C.: A forest-fire model and some thoughts on turbulence. Phys. Lett. A. **147**, 297–300 (1990)
4. Barabási, A., Albert, R.: Emergence of scaling in random networks. Science **286**(5439), 509–512 (1999)
5. Casals, M., Guzman, K., Cayla, J.: Mathematical models used in the study of infectious diseases. Rev. Esp. Salud. Publica. **83**(5), 689–95 (2009)
6. Ellis, L.: Spread of Epidemic Disease on Networks. Wiley, New York (2001)
7. Keeling, M., Eames, K.: Networks and epidemic models. J. R. Soc. Interf. **2**(4), 295–307 (2005)
8. Kenah, E., Robins, J.: Network-based analysis of stochastic sir epidemic models with random and proportionate mixing. J. Theor. Biol. **249**(4), 706–722 (2007)
9. Kot, M.: Elements of Mathematical Ecology. Cambridge University Press, Cambridge (2001)
10. Król, D.: On modelling social propagation phenomenon. In: Nguyen, N., Attachoo, B., Trawiński, B., Somboonviwat, K. (eds.) Intelligent Information and Database Systems. Lecture Notes in Computer Science, pp. 227–236 (2014)
11. Król, D.: Propagation phenomenon in complex networks: theory and practice. New Gener. Comput. **32**(3–4), 187–192 (2014)

12. Lotka, A.: Relation between birth rates and death rates. Science **26**, 121–130 (1907)
13. MacDonald, G.: The analysis of malaria epidemics. Trop. Dis. Bull. **50**(10), 871–889 (1953)
14. Mason, O., Verwoerd, M.: Graph theory and networks in biology. IET Syst. Biol. **1**(2), 89–119 (2007)
15. Murray, J.: Mathematical Biology. Springer, New York (2002)
16. Newman, M.: Spread of epidemic disease on networks. Phys. Rev. E. Stat. Nonlin. Soft Matter Phys. **66**(1), 0168,128 (2002)
17. Nokes, D., Anderson, R.: The use of mathematical models in the epidemiology study of infectious diseases and in the design of mass vaccination programmes. Epidemiol. Infect **101**, 1–20 (1988)
18. Rhodes, C., Anderson, R.: Forest-fire as a model for the dynamics of disease epidemics. J. Frankl. Inst. **335**(2), 199–211 (1998)
19. Rhodes, C., Jensen, H., Anderson, R.: On the critical behaviour of simple epidemics. Proc. Biol. Sci. **264**(1388), 1639–46 (1997)
20. Sobol, I.: Sensitivity estimates for nonliner mathematical models. Math. Model. Comput. Exp. **1**(12), 577–596 (1993)
21. Tran, L., Rizk, M., Liao, J.: Ensemble modeling of metabolic networks. Biophys. J. **95**(12), 5606–5617 (2008)
22. Trapman, P.: On analytical approaches to epidemics on networks. Theor. Popul. Biol. **71**(2), 160–73 (2007)
23. Vallabhajosyula, R., Raval, A.: Computational modeling in systems biology. Methods Mol. Biol. **662**(5), 97–120 (2010)
24. Watts, D., Strogatz, S.: Collective dynamics of 'small-world' networks. Nature **393**(6684), 440–442 (1998)
25. Zinck, R., Grimm, V.: Unifying wildfire models from ecology and statistical physics. Am. Nat. **174**, E170–E185 (2009)