Spatial structure of complex network and diffusion dynamics

Zi Hui

LUNAM Université, Université du Maine

Thèse de Doctorat

Spécialité: Physique

Rélisée au Laboratoire de Physique Statistique et Systèmes Complexes de l'ISMANS

par

Zi Hui

pour l'obtention du grade de Docteur en Sciences

Spatial structure of complex network and diffusion dynamics

Présentée le 8 avril 2013

Acceptée sur proposition du jury:

Prof. Q. Alexandre Wang (Directeur de thèse)

Prof. Jean-Marc Greneche (Co-Directeur)

Prof. Bénédicte Le Grand (Rapporteur)

Prof. Paul Lescot (Rapporteur)

Prof. Wei Li

Prof. Jean-Pierre Badiali

Prof. François Tsobnang

Acknowledgements

This booklet is the result of my PhD witness at the Université du Maine. Looking back at this important period of my life fills me with pleasure and satisfaction. But you would not be reading these words if I would not have benefited from the support and encouragement of a lot of people. Now is a perfect time to thank those who stand along the sidelines and who support me in various ways.

First and foremost, I would like to express my deepest gratitude to my thesis supervisor, Prof. Q. Alexandre Wang whose motivation and personal guidance was indispensable during my PhD study. I am pleased to continuously enjoy his knowledge spilling over to me, in the form of insightful comments, useful advices, innovative ideas and invaluable guidance. I also sincerely appreciated his continuous supports and encouragements, particularly his patience with me when things did not work out as foreseen. Additionally, I am very much indebted for the tremendous amount of trust and the constant source of inspiration he has provided during my journey in achieving the doctoral degree. He has provided a huge contribution toward the refinement of all the chapters.

Then my deep appreciation goes to Prof. Xu Cai who is my supervisor in my original university in China. He is the very person who takes me into the field of complexity science and gives me the chance to study in France. I was deeply influenced by his attitude and spirit on scientific research, as well as living. I also would like to gratefully thank Prof. Wei Li, my other supervisor in China, for his enormous assistance on my research work and living.

The members of ISMANS have contributed immensely to my personal and professional time at Le Mans. I would also like to thank the director of ISMANS, Prof. Mouad Lamrani for his support and encouragement. I would like to thank Dr. Cyril Pujos, Dr. Aziz El Kaabouchi, and Dr. Armen Allahverdyan for their friendly assistance and valuable discussions during my studies in ISMANS. I also thank to Prof. Alain Le Méhauté, Prof. François Tsobnang, Prof. Jean-Charles Craveur, Dr. Laurent Nivanen, Dr. Dominique Marceau, Dr. Gilles Brement, Mr. François Angoulvant, Mr. Perry Ngabo, Dr. Benoît Minisini, Mr. Luc Chanteloup, Mr. Jean-Philippe Bournot, Dr. Louis-Georges Tom, Dr. Benjamin Mercier, Dr. Pavlo Demianenko, Mrs. Sylvie Bacle, Mrs. Sandra Dugué, Miss Zélia Aveline, Miss Yukimi Duluard, Miss Charline Pasquier, Mrs. Leproux Audrey, and Miss Susanne Vilchez for creating very pleasant and stimulating ambiance in ISMANS. I would like to take this opportunity to thank all the other members in ISMANS for their

help and encouragement throughout these years.

I am also thankful to Prof. Jean-Marc Greneche, my associated supervisor, Prof. Youping Gao and Prof. Pierre Chauvet, my Comité de Suivi de Thèse. Their comments and reflections were indispensable to complete my PhD study and thesis.

I would also thank all the members in Institute of Particle Physics of Huazhong Normal University for their education and support during my undergraduate and postgraduate study.

My time at Le Mans is made enjoyable in large part due to the many friends and groups that becomes a part of my life. I am grateful for time spent with friends, for the memorable trips. I want to thank Ru Wang, Kai Zhao, Jian Jiang, Weibing Deng, Tonglin Ling and Wenxuan Zhang for their help and encouragement. Writing a PhD thesis is a long and lonely way which I would not have faced without their presence.

Last but not least, this thesis could not have been completed without the boundless support and sacrifices of my beloved family. No words can describe my gratitude to my parents, who always sacrifice and never hesitate to show how much they love and care for me. I would like also to thank all people, who due to lack of space, or by involuntary neglect, I do not mention.

Résumé

Dans le développement récent des sciences de réseau, réseaux contraints spatiales sont devenues un objet d'une enquête approfondie. Spatiales des réseaux de contraintes sont intégrées dans l'espace de configuration. Leurs structures et les dynamiques sont influencées par la distance spatiale. Ceci est prouvé par les données empiriques de plus en plus sur des systèmes réels montrant des lois exponentielles ou de distribution d'énergie distance spatiale de liens. Dans cette thèse, nous nous concentrons sur la structure de réseau spatial avec une distribution en loi de puissance spatiale. Plusieurs mécanismes de formation de la structure et de la dynamique de diffusion sur ces réseaux sont pris en considération.

D'abord, nous proposons un réseau évolutif construit en l'espace de configuration d'un mécanisme de concurrence entre le degré et les préférences de distance spatiale. Ce mécanisme est décrit par un $a\frac{k_i}{\sum_j k_j} + (1-a)\frac{r_{ni}^{-\alpha}}{\sum_j r_{nj}^{-\alpha}}$, où k_i est le degré du noeud i et r_{ni} est la distance spatiale entre les noeuds n et i. En réglant le paramètre a, le réseau peut être fait pour changer en continu à partir du réseau spatiale entraînée ($a = 0$) pour le réseau sans échelle ($a = 1$). La structure topologique de notre modèle est comparé aux données empiriques de réseau de courrier électronique avec un bon accord.

Sur cette base, nous nous concentrons sur la dynamique de diffusion sur le réseau axé sur spatiale ($a = 0$). Le premier modèle, nous avons utilisé est fréquemment employée dans l'étude de la propagation de l'épidémie: l'spatiale susceptible-infecté-susceptible (SIS) modèle. Ici, le taux de propagation entre deux noeuds connectés est inversement proportionnelle à leur distance spatiale. Le résultat montre que la diffusion efficace de temps augmente avec l'augmentation de α. L'existence d'seuil épidémique générique est observée, dont la valeur dépend du paramètre α Le seuil épidémique maximum et le ratio minimum fixe de noeuds infectés localiser simultanément dans le intervalle $1.5 < \alpha < 2$.

Puisque le réseau spatiale axée a bien défini la distance spatiale, ce modèle offre une occasion d'étudier la dynamique de diffusion en utilisant les techniques habituelles de la mécanique statistique. Tout d'abord, compte tenu du fait que la diffusion est anormale en général en raison de l'importante long plage de propagation, nous introduisons un coefficient de diffusion composite qui est la somme de la diffusion d'habitude constante D des lois de la Fick appliqué sur différentes distances de transfert possibles sur le réseau.

Comme prévu, ce coefficient composite diminue avec l'augmentation de α. et est une bonne mesure de l'efficacité de la diffusion. Notre seconde approche pour cette diffusion anormale est de calculer le déplacement quadratique moyen $\langle l^2 \rangle$ à identifier une constante de diffusion D' et le degré de la anomalousness γ avec l'aide de la loi de puissance $\langle l^2 \rangle = 4D't^\gamma$. D' comportements de la même manière que D, *i.e.*, elle diminue avec l'augmentation de α. γ est inférieur à l'unité (subdiffusion) et tend à un (diffusion normale) que α augmente.

Key Words préférence la distance spatiale, l'espace réseau axée sur, transition de phase, spatiale susceptible-infecté-susceptible modèle, soutenue rapport infecté noeuds, épidémie de seuil, la première loi de Fick, diffusion anormale, la diffusion spatiale, coefficient de diffusion

Abstract

In the recent development of network sciences, spatial constrained networks have become an object of extensive investigation. Spatial constrained networks are embedded in configuration space. Their structures and dynamics are influenced by spatial distance. This is proved by more and more empirical data on real systems showing exponential or power laws spatial distance distribution of links. In this dissertation, we focus on the structure of spatial network with power law spatial distribution. Several mechanisms of structure formation and diffusion dynamics on these networks are considered.

First we propose an evolutionary network constructed in the configuration space with a competing mechanism between the degree and the spatial distance preferences. This mechanism is described by $a\frac{k_i}{\sum_j k_j} + (1-a)\frac{r_{ni}^{-\alpha}}{\sum_j r_{nj}^{-\alpha}}$, where k_i is the degree of node i and r_{ni} is the spatial distance between nodes n and i. By adjusting parameter a, the network can be made to change continuously from the spatial driven network ($a = 0$) to the scale-free network ($a = 1$). The topological structure of our model is compared to the empirical data from email network with good agreement.

On this basis, we focus on the diffusion dynamics on spatial driven network ($a = 0$). The first model we used is frequently employed in the study of epidemic spreading: the spatial susceptible-infected-susceptible (SIS) model. Here the spreading rate between two connected nodes is inversely proportional to their spatial distance. The result shows that the effective spreading time increases with increasing α. The existence of generic epidemic threshold is observed, whose value depends on parameter α. The maximum epidemic threshold and the minimum stationary ratio of infected nodes simultaneously locate in the interval $1.5 < \alpha < 2$.

Since the spatial driven network has well defined spatial distance, this model offers an occasion to study the diffusion dynamics by using the usual techniques of statistical mechanics. First, considering the fact that the diffusion is anomalous in general due to the important long-range spreading, we introduce a composite diffusion coefficient which is the sum of the usual diffusion constant D of the Fick's laws applied over different possible transfer distances on the network. As expected, this composite coefficient decreases with increasing α and is a good measure of the efficiency of the diffusion. Our second approach to this anomalous diffusion is to calculate the mean square displacement $\langle l^2 \rangle$ to identify a

diffusion constant D' and the degree of the anomalousness γ with the help of the power law $\langle l^2 \rangle = 4D't^{\gamma}$. D' behaviors in the same way as D, *i.e.*, it decreases with increasing α. γ is smaller than unity (subdiffusion) and tends to one (normal diffusion) as α increases.

Key Words spatial distance preference, spatial driven network, phase transition, spatial susceptible-infected-susceptible model, steady infected nodes ratio, epidemic threshold, Fick's first law, anomalous diffusion, spatial diffusion, diffusion coefficient

Contents

Chapter 1

Introduction

Complex network structures describe a wide variety of systems in nature and society, and are represented by a collection of nodes and links. For example, the Internet is a complex network of routers and computers linked by various physical or wireless links; fads and ideas spread on the social network, whose nodes are human beings and whose links represent various social relationships; the cell is best described as a complex network of chemicals connected by chemical reactions; the World Wide Web (WWW) is an enormous virtual network of Web pages connected by hyperlinks, and many others (see Fig. 1.1 [1]). In order to determine its topology of the systems and to understand such interwoven systems, scientists investigate the mechanisms from these examples. And they encountered significant challenges as well. During the last decades, network became a subject of interest of scientists who wanted to discover the general laws governing systems formation and evolution. After that, such laws are applied to the majority of network models. The most important observation is that these systems can be represented by complex networks, which means that their properties cannot be simply reduced to a compound of individual components.

The origin of complex network can be traced back to the eighteenth century with the solution of the Königsberg bridge problem by the Swiss mathematician Leonhard Euler (see Fig. 1.2) [2]. This problem is referred as the first instance of a network theory application. In addition to the developments in mathematical graph theory, the study of networks has obtained many important achievements in some specialized contexts, as for instance in the social sciences. Social networks analysis started to develop in the early 1920s. More and more scientists focus on the relationships among social entities, such as communication between members of a group, trades among nations, or economic

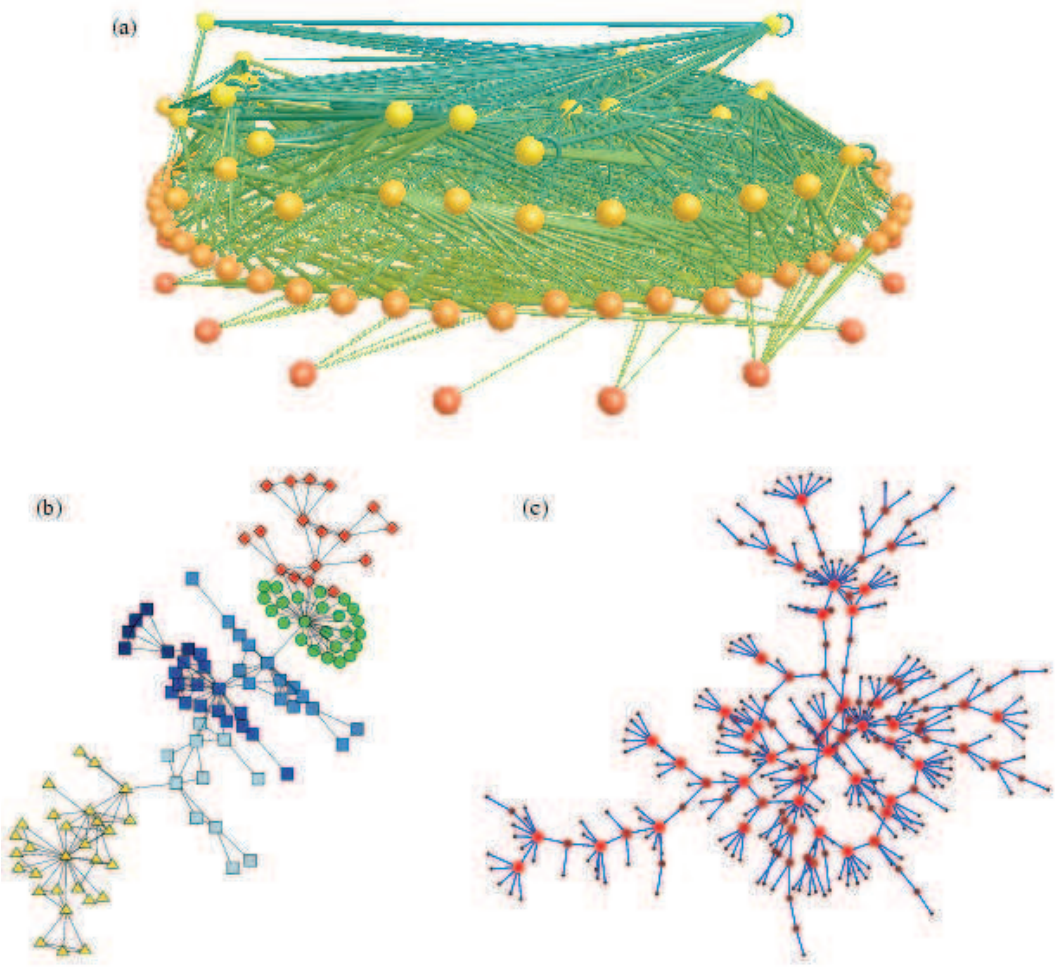

Fig. 1.1: Three examples of complex networks. (a) A food web of predator-prey interactions between species in a freshwater lake. (b) The network of collaborations between scientists at a private research institution. (c) A network of sexual contacts between individuals.

transactions between corporations. Since the last decade it has witnessed the birth of a new campaign of interests and researches in the study of complex network models. Random graphs are first studied by the Hungarian mathematicians Erdös and Rényi [3, 4, 5]. Thus this type of network is called ER random network. According to the ER model, we start with N isolated nodes, new links are connected with uniform probability p, creating a graph with approximately $pN(N-1)/2$ links distributed randomly. This model has guided our thinking about complex networks for decades since its application. But the growing interest in complex systems has prompted many scientists to reconsider this modeling paradigm and ask a simple question: are the real networks behind such diverse complex systems as the cell or the Internet fundamentally random? Our intuition clearly

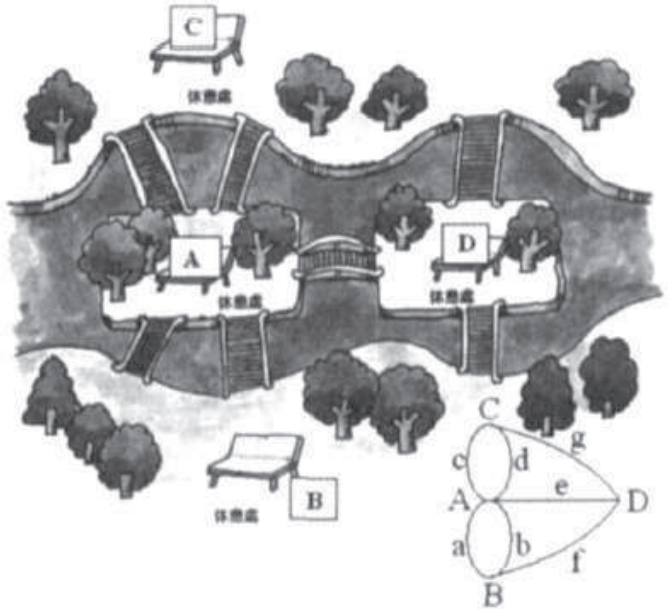

Fig. 1.2: Königsberg bridge problem: find a round trip that traversed each of the bridge of the prussian city of Königsberg exactly once. (A, B, C, D define four lands, a, b, c, d, e, f, g define seven bridges connecting four lands)

indicates that complex systems must display some organizing principles, which should be at some level encoded in their topology. But if the topology of these networks indeed deviates from a random graph, we need to develop tools and measurements to capture in quantitative terms the underlying organizing principles [6, 7]. The interest in networks was however renewed in 1998 by Watts and Strogatz [8] who extracted stylized facts from real world networks and proposed a simple, new model of small-world (SW) networks. The structure is characterized by the SW properties that the average path length over all nodes is as short as that in random graphs, and that the clustering coefficient, defined by the average ratio of the number of links connecting to its nearest neighbors of a node to the number of possible links between all these nearest neighbors, is as large as that in regular graphs. Large clustering coefficient means the high frequency of "the friend of a friend is also his/her friend." This renewal interest is reinforced after the publication, a year later, of an article about scale-free (SF) network by Albert and Barabási [9] on the existence of strong degree heterogeneities. Strong heterogeneities are in sharp contrast with the random graphs considered so far and the existence of strong fluctuations in real-world networks triggered a wealth of studies. Regarding the SF properties, the degree distribution follows a power-law, $P(k) \sim k^{-\gamma}$, $2 < \gamma < 3$; the fat-tail distribution consists of many nodes with low degrees and a few hubs with very high degrees. These properties

are widespread recently, including characterization of the Internet power law organization [10] and the identification of such a kind of connectivity in the WWW [11], resulting in the scale-free paradigm [12]. Subsequent discoveries suggest that many natural and artificial networks also exhibit scale-free organization, including metabolic networks [13], traffic networks [14, 15], protein networks [16], food webs [17], scientists collaborators [18], and many others. The visual examples of random, small-world, scale-free networks are shown in Fig. 1.3 from left to right.

Fig. 1.3: Examples of random, small-world, scale-free networks from left to right.

A decade later, we can find many books [20, 21, 22, 23, 24, 25] and reviews on this subject [1, 26, 27, 28]. These books and reviews discuss usually very quickly spatial aspects of networks. However, for many critical infrastructures, communication and biological networks, it's relevant to the spacial distance: most of the people have their friends and relatives in their neighborhood, power grids and transportation networks depend obviously on distance, many communication network devices have short wavelength radio frequency range, the length of axons in a brain has a cost, and the spread of contagious diseases is not uniform across territories [29]. An example of the spatial network of Facebook is shown in Fig. 1.4. In particular, in the important case of the brain, regions that are spatially close have a larger probability of being connected with than remote regions have as longer axons are more costly in terms of material and energy [30]. Wiring costs depending on distance is thus certainly an important aspect of brain networks and we can probably expect spatial networks to be very relevant in this rapidly evolving topic. Another particularly important example of such a spatial network is the Internet which is defined as the set of routers linked by physical cables with different lengths and latency times. More generally, the distance could be another parameter such as a social distance measured by salary, socioprofessional category differences, or any quantity which measures the cost associated with the formation of a link. All these examples show that these networks have nodes and links which are constrained by some spatial constrains and are

usually embedded in a space. The important effects on their topological properties and consequently on processes take place on such networks. The topological aspects of the network are then correlated to spatial aspects such as the location of the nodes and the length of links.

Spatial networks were actually the subject a long time ago of many studies in quantitative geography. Objects of studies in spatial are locations, activities, flows of individuals and goods. In the 1970s the scientists became to work in quantitative geography focused on networks evolving in time and space. Some books [31, 32] mentioned such modern questions in the complex system field morn than 40 years ago. In these books, the authors discuss the importance of space in the formation and the evolution of networks. They develop tools to characterize spatial networks and to discuss possible models. Maybe what were lacking at that time, were the datasets of large networks and the larger computer capabilities, fortunately a lot of interesting thoughts can be found in these early studies. Most of the important problems such as the location of nodes of a network, the evolution of transportation networks and their interaction with population and activity density are addressed in these earlier studies, but many important points still remain unclear and will certainly be benefit from the current knowledge on networks and complex systems. The advantages in complex networks help us to gain new insights in these difficult problems. In spatial and social relationship, it would be about the understanding of the evolution of social networks and the human mobility, the spatial structure of social relationship *et al.* and about how these different factors are entangled with each other, in order to propose an integrated approach of scale, mobility, and spatial distribution of activities at various scales.

In this dissertation, we have concerned the structure of evolutional spatial network and how does spatial structure works on different diffusion dynamics. The thesis is organized as following.

In the second chapter, we give an introduction to the basic concepts and notions in complex networks and present several topological measurements like degree, degree distribution, clustering coefficient, average path length, and so on. In the following, we present two important topological properties, small-world property and scale-free degree distribution, of the real networks in nature. After that, the illustrations of these models respectively to Erdös-Rényi, Watts-Strogatz, Barabási-Albert are included. The models are presented in sections organized according to their main types, including clustering coefficients, assortativity, entropies, centrality, subgraphs, spectral analysis, community-

Fig. 1.4: A visualization map about the Facebook users, developed by a Facebook intern. It took a sample of 10 million pairs of friends from the Facebook database and matched them with the corresponding coordinates of each city. The more friends and the larger distance between two countries the brighter the edges on a black-blue-white color scale.

based measurements and hierarchical measurements.

In the third chapter, we give a brief introduction of the spatial network. We first review the tools to characterize these networks and the empirical properties of some important spatial networks, and find that the spatial distance between two connected nodes follows power-law or exponential distribution. We then review five classes of models about spatial network formation which allows us to understand the main effects of the spatial constraints on the network properties. We present a spatial growth network model in the fourth class. The model grows following linear compound preferential attachment of degree and spatial distance. Several topological measurements are discussed. The topological structure of our model is compared to the empirical data from email network in this chapter as well.

In the fourth chapter, we focus on the epidemic spreading processes which take place on spatial networks. In the beginning, we introduce some epidemic spreading processes on classical network models. Such processes ignore the spatial effect. Then, we define a spatial epidemic spreading process on spatial driven model. The spreading probability is inversely proportional to the spatial distance. In addition, infected nodes ratio, the steady infected nodes ratio and the epidemic threshold are studied.

In the fifth chapter, we discuss the diffusive dynamics on spatial driven network based

on general diffusion equations. At first, we introduce some classical diffusion models, such as, Fick's first law, Fick's second law, reaction-diffusion system and anomalous diffusion. Diffusion coefficient is an important parameter indicative the diffusion mobility. We use two diffusion models, Fick's first law and anomalous diffusion, on the spatial network. In Fick's first law, the short-range diffusion and the long-range diffusion are both considered to calculate this coefficient. In anomalous diffusion, the diffusion process is sub-diffusion process. In this work, we want to reveal that the diffusion coefficient is affected by kinds of network properties.

Finally, we draw our conclusions and present the main perspectives of this study in the last chapter.

Chapter 2

Networks

2.1 Introduction

Networks present all around us in our life. Communication networks consisting of telephones and mobile phones, the electrical power grid, computing communication networks, airline networks are an important part of every day life. Society is also networked. The network of friendship between individuals, the collaboration network of scientific, and the networks of business relations between people and firms are examples of social and economic networks. In history, the study of networks can be traced back to Leonhard Euler's solution of the Königsberg bridge problem (consisting in searching a round trip that traversed each of the bridge of the prussian city of Königsberg exactly once), after which the theory of graphs has been useful for theoretical physics, biology, sociology and economy. Graphs are used for describing mathematical concepts in networks. Graphs represent the essential topological properties of a network by treating the network as a collection of nodes and links. For example, in computer networks, such as the Internet, computers can be represented by nodes, and the links are represented by the cables between them. In WWW networks, the nodes are the HTML pages, and the links represent the connected relationship between pages. This is a simple, yet powerful concept. Because of its simplicity, it considers different complex systems such as those described above, using the same mathematical tools and methods, the properties of the networks are similar. The last decade has witnessed the birth of a new movement of interest and research in the study of complex networks, which is triggered by two seminal papers. One is proposed by Watts and Strogatz on small-world networks, published in Nature in 1998, the other one is proposed by Barabási and Albert on scale-free networks, published in Science one year

later.

The success of complex networks is therefore to a large extent a consequence of their natural suitability to represent virtually any discrete system. Moreover, the organization and evolution of such networks, as well as dynamical processes on them, involve non-linear models and effects. The connectivity of networks is ultimately decisive in constraining and defining many aspects of systems dynamics. The key importance of this principle has been highlighted in many comprehensive surveys. For instance, the behavior of biological neuronal networks, one of the greatest remaining scientific challenges, is largely defined by connectivity. Due to its virtually unlimited generality for representing connectivity in the most diverse real systems in an integrative way, complex networks are promising for integration and unification of many aspects of modern science, including the inter-relationships between structure and dynamics. Such a potential has been confirmed with a diversity of applications for complex networks, encompassing areas such as epidemiology, genetics, ecology, physics, the Internet and WWW, computing, *et al.*. In fact, applications of complex networks are redefining the scientific method through incorporation of dynamic and multidisciplinary aspects of statistical physics and computer science.

The research on complex networks begin with the efforts of defining new concepts and measures to characterize the topology of real networks. The main result has been the identification of a plenty of statistical properties and unifying principles common to most of the real networks considered. These empirical results have initiated a revival of network modeling, since the models introduced in mathematical graph theory turned out to be far from the real needs. Scientists had to do with the development of new models to mimic the growth of a network and to recover the structural properties observed in real topologies. So this stage of the research was triggered by the expectancy that understanding and modeling the structure of a complex network would lead to a better knowledge of its evolutionary mechanisms, and to a better cottoning on its functional and dynamical behavior.

In the following section, we will introduce definitions and notations of the complex network, and discuss the basic quantities used to describe the topology of a network.

2.2 The structure of complex networks

2.2.1 Network definitions

Graph theory [34, 35, 36, 101] is the natural framework for the accurate mathematical handling of complex networks and formally, a complex network can be displayed as a graph. A graph is a pair $G = (V, E)$ of sets satisfying $E \subseteq [V]^2$; thus the elements of E are 2-element subsets of V. To avoid notational ambiguities, we shall always assume tacitly that $V \cap E = \varnothing$. The elements of V are the nodes (vertices) of the graph G, the elements of E are its edges (lines), where the edges connect paris of nodes $(p, q) \in E$ (with $p, q \in V$). If the starting node p and the ending node q of an edge is ordered, we see the graph as directed one, otherwise it is undirected. If $p = q$ we call it loops, which could be directed and undirected as well. In Fig. 2.1 [1], there are different ways in which networks may be more complex. In the following we are going to utilize simple undirected graphs which do not include any loops and multiple edges.

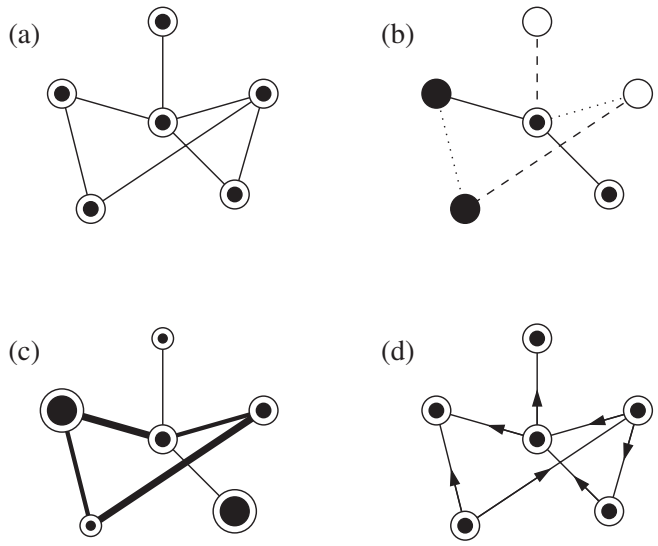

Fig. 2.1: Examples of different types of networks: (a) an undirected network with a single type of node and a single type of edge, (b) a network with a number of discrete node and edge types, (c) a network with varying the weight of node and edge, (d) a directed network in which each edge has a direction.

In computer science and mathematics, scientists always consider a matricial representation of a graph. If the number of nodes N denotes the size of the graph, the $N \times N$ adjacency matrix $\mathcal{A}$ may describe the graph with elements $a_{pq} = 1$ if an edge is presented between nodes p and q, otherwise $a_{pq} = 0$. Therefore, the adjacency matrix of a network

with undirected edges is symmetrical. Hence, it may be converted into an upper triangular matrix. For a network with directed edges, an element of the adjacency matrix, a_{pq}, equals 1 if there is an edge from vertex p to vertex q, and equals 0 otherwise. In weighted graphs, the a_{pq} elements of the adjacency matrix are equal to the weights of existing edges or equal to zero. The advantage of the adjacency matrix representation is the direct access of the edges in the matrix, but this description is robust. The disadvantage of the adjacency list description is that we cannot access the edges in one step, but it does not occupy much memory. The statistics of the adjacency matrix of a random network contains complete information about the structure of the net, and, in principle, one has to study just the adjacency matrix. Generally, this is not an easy task, so that, instead of this, only a very restricted set of structural characteristics is usually considered.

2.2.2 Network measurements

In order to characterize and represent complex networks, many measurements have been developed. Here we present some main features which are frequently used in the reference.

2.2.2.1 Degree

The degree (or connectivity) k_i of a node i is the number of edges connected with the node, and is defined in terms of the adjacency matrix as:

$$k_i = \sum_j a_{ij}. \tag{2.1}$$

The average degree $\langle k \rangle$ is simply

$$\langle k \rangle = \frac{2l}{N}, \tag{2.2}$$

where l is the total number of links, N is the size of the network.

If the graph is directed, the degree of the node i has two components: its out-degree $k_i^{\text{out}} = \sum_j a_{ij}$, and in-degree $k_i^{\text{in}} = \sum_j a_{ji}$. The total degree of the node i is then given by $k_i = k_i^{\text{out}} + k_i^{\text{in}}$. A list of the node degrees of a graph is called the degree sequence.

2.2.2.2 Degree distribution

An important network feature to be analyzed is the degree distribution $P(k)$, which gives the probability that a randomly chosen node has degree k. In the case of directed networks one needs to consider two distributions, out-degree distribution $P(k^{\text{out}})$ and in-degree distribution $P(k^{\text{in}})$.

Information on how the degree is distributed among the nodes of an undirected network can be obtained either by a plot of $P(k)$, or by the calculation of the moments of the distribution. The n-moment of $P(k)$ is defined as:

$$\langle k^n \rangle = \sum_k k^n P(k). \tag{2.3}$$

The first moment $\langle k \rangle$ is the mean degree of graph G. The second moment measures the fluctuations of the connectivity distribution.

2.2.2.3 Degree correlation

A large number of real networks are correlated in the sense that the probability that a node of degree k is connected to another node of degree, say k', depends on k. In these cases, it is necessary to introduce the conditional probability $P(k'|k)$, being defined as the probability that a link from a node of degree k points to a node of degree k' [33]. $P(k'|k)$ satisfies the normalization$\sum_{k'} P(k'|k) = 1$. While, thanks to the shortage of empirical data and consequent large fluctuations in the computed values, it is better to bring a more coarse, but less fluctuating measure, such as the average degree of the nearest neighbors of nodes with degree k [37, 62]:

$$k_{nn}(k) = \sum_{k'} k' P(k'|k). \tag{2.4}$$

The behavior of $k_{nn}(k)$ as a function of node degree k can be used to detect a property known as assortative in social networks, occurring when k_{nn} is an increasing function of k. Particularly, it has been indicated that the real world networks can be classified into two different classes: one is assortative correlation meaning that high degree nodes tend to be connected, such as social networks. Another is disassortative correlation meaning that high degree node tends to connect with low degree node, such as technological networks, biological networks, information networks.

2.2.2.4 Clustering coefficient

For the description of connections in the environment closest to a node, on introduces the so-called clustering coefficient. It is a local measure of the interconnection of the nodes, representing the probability that the neighbors of a given node are also connected (in social networks, roughly speaking, it is the probability that a friend of my friend is also my friend). Regarding a node i, the clustering coefficient c_i is defined as:

$$c_i = \frac{2e_i}{k_i(k_i - 1)} \tag{2.5}$$

where e_i is the number of links connecting neighbors of node i, $k_i(k_i - 1)/2$ the total number of possible connections (for peripheral nodes having $k_i = 1$, c_i is taken equal to zero). The number of edges e_i can be expressed in terms of the adjacency matrix:

$$e_i = \frac{1}{2} \sum_{p,q} a_{ip} a_{pq} a_{qi} \tag{2.6}$$

revealing that c_i is a measure of correlations in the adjacency matrix. Natural and artificial networks display very high clustering coefficients, an obvious deviation from random graph behavior. The clustering coefficient of the graph is then given by the average of c_i over all the nodes in graph G (see Fig. 2.2):

$$C = \langle c \rangle = \frac{1}{N} \sum_i c_i. \tag{2.7}$$

Fig. 2.2: Illustration of the definition of the clustering coefficient C. The individual vertices have local clustering coefficients, Eq. (2.5), of 1, 1, $\frac{1}{6}$, 0 and 0, for a mean value, Eq. (2.7), of $C = \frac{13}{30}$.

The behavior of $C(k)$ as a function of node degree, averaged over all nodes with degree k, has also been investigated, in order to characterize hierarchy and structural organization

of networks [37, 38, 39, 40]:

$$C(k) = \frac{1}{NP(k)} \sum_i c_i \delta_{k_i,k}. \tag{2.8}$$

A decreasing behavior of $C(k)$ with k has been empirically observed in some real world networks [40, 41, 42, 43].

2.2.2.5 Shortest path lengths

Shortest paths play an important role in the communication and transport within a network. It is useful to represent all the shortest path lengths of a graph G as a matrix D in which the element l_{pq} is the length of the geodesic from node p to node q. The maximum value of l_{pq} is called the diameter of the graph, and will be defined in the following as Diam(G). A measure of the typical separation between two nodes in the graph is given by the average path length, also known as characteristic path length, defined as:

$$L = \frac{1}{N(N-1)} \sum_{p,q \epsilon V, p \neq q} l_{pq}. \tag{2.9}$$

A problem with this definition is that L diverges if there are disconnected components in the graph. One possibility to avoid the divergence is to limit the summation in Eq. (2.9) only to couples of nodes belonging to the largest connected component.

2.2.2.6 Betweenness

The betweenness can be regarded as a measure of the importance of a node or an edge as a controller of the information which is flowing between the other nodes or edges in the network. The communication of two non-adjacent nodes, i and j, depending the nodes belonging to the shortest path connecting i and j. Consequently, a measure of the importance of a given node can be obtained by counting the number of shortest path going through it, and defining as node betweenness. Together with the degree and the closeness of a node (defined as the inverse of the mean distance from all other nodes), the betweenness is one of the standard measure of node centrality. It is originally introduced to quantify the importance of an individual in a social network [6, 7, 44]. More precisely, the betweenness b_o of a node o, is defined as [45, 46]:

$$b_o = \sum_{i,j \in N, i \neq j} \frac{n_{ij}(o)}{n_{ij}}, \tag{2.10}$$

where n_{ij} is the number of shortest paths connecting node i and j, and $n_{ij}(o)$ is the number of shortest paths connecting i and j and passing through o. There is an introduction of the standard algorithms to find shortest paths, such as the Dijkstra's algorithm and the breadth-first search method in Refs. [47, 48, 49]. In Ref. [50, 51], a fast algorithm recently is proposed to calculate the betweenness. In Refs. [52, 53, 54, 55, 56, 57], the betweenness distributions have been investigated. In Refs. [14, 58, 59, 60], the betweenness-betweenness correlations and betweenness-degree correlations have been explored respectively.

The edge betweenness of edge (i, j) is defined as the number of shortest paths between pairs of nodes which pass through the edge (i, j) [61]:

$$b_{ij} = \sum_{m \neq n} b_{ij}(m, n) = \sum_{m \neq n \neq i \neq j} \frac{n_{ij}(m, n)}{n_{mn}}, \tag{2.11}$$

where $n_{ij}(m, n)$ denotes all the numbers of shortest paths passing through edge (i, j). So, $b_{ij}(m, n)$ reflects the importance of the edge (i, j) connecting the node m and n.

2.2.2.7 Community structures

Given a graph G, a community (or cluster) is a subgraph G', whose nodes are tightly connected. Since the structural cohesion of the nodes in G' can be quantified in different ways, there are several definitions of community structures. Fig. 2.3 is a visualization of the friendship network of children in a US school studied by Moody [63], using individual's race as the principal divisions in the network.

The traditional method for extracting community structure from a network is cluster analysis [64], sometimes called hierarchical clustering. In this approach, one assigns a connection strength to paris of nodes in the network of interest. In general, each of the $n(n-1)/2$ possible pairs in a network of n nodes is assigned such a strength, not just those that are connected by an edges, although there are versions of the method where not all pairs are assigned a strength, in which case one can assume the remaining pairs to have a connection strength of zero. Then, starting with n isolated nodes, one adds edges in order to decrease node-node connection strength. One can pause at any point during the process and examine the component structure formed by the edges added so far. These components are taken to be the communities or clusters at that stage in the process. When all edges have been added into the system, all nodes are connected to all others, and there is only one community. The whole process can be represented by a tree or dendrogram of union operations between node sets where the communities at any level

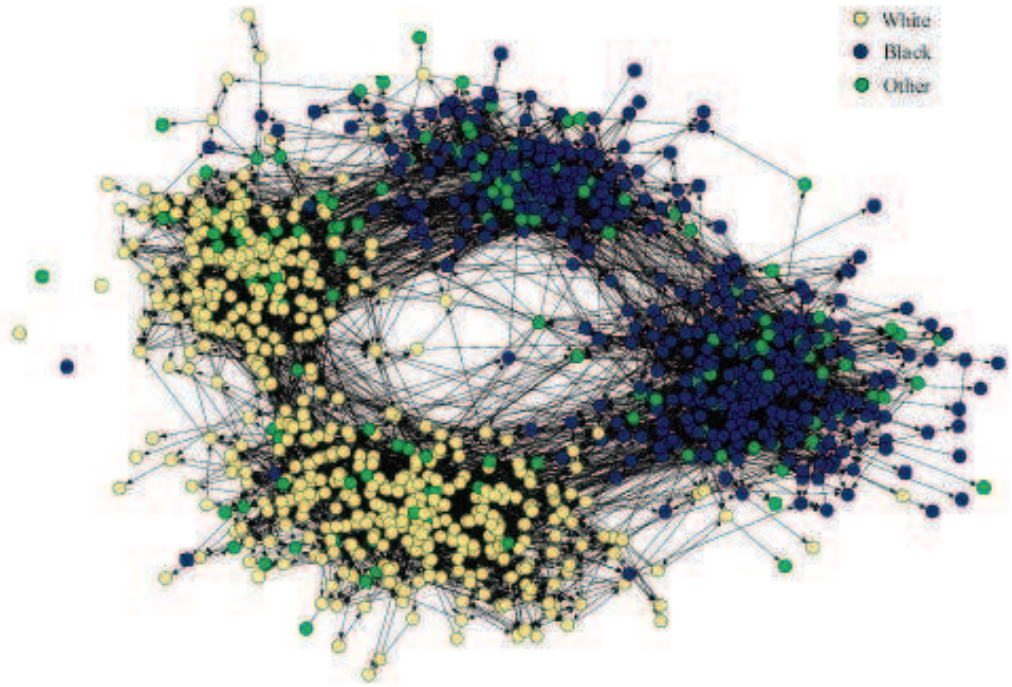

Fig. 2.3: Friendship network of children in a US school.

correspond to a horizontal cutting.

The strongest definition of community structure requires that all pairs of community members choose each other. Such a requirement leads to the definition of a clique. A clique is a maximal complete subgraph of three or more nodes, i.e. a subset of nodes all of which are adjacent to each other, and such that no other nodes exist adjacent to all of them. This definition can be extended by weakening the requirement of adjacency into a requirement of reachability: a n-clique is a maximal subgraph in which the largest geodesic distance between any two nodes is no larger than n. When $n = 1$, this definition represents a clique. 2-cliques are subgraphs where all nodes need not to be adjacent but are reachable through at most one intermediary. In 3-cliques all nodes are reachable through at most two intermediaries, and so on. Whereas the concept of n-clique includes increasing the permissible path lengths, an alternative possibility to weaken the strong assumption of cliques contains reducing the number of other nodes to which each node must be connected.

A various class of definitions is based on the relative frequency of links [27]. In this case, communities are seen as groups of nodes within which connections are dense, and between which connections are sparser [65, 66] (see Fig. 2.4). While the simplest formal

definitions in this case have been put forward in Refs. [67, 68], a less stringent definition is as follows: G' is a community if the sum of all degrees within G' is larger than that the sum of all degrees towards the rest of the graph [69]. The definitions proposed here are not the only possible choices. Several other definitions, perhaps more appropriate in some cases, can be found in Ref. [7].

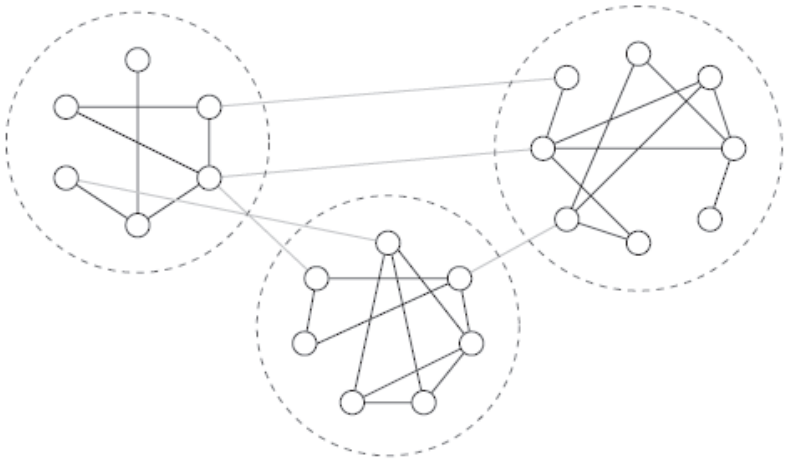

Fig. 2.4: Community can be defined as group of nodes such that there is a higher density of edges within group than between them. In this picture, there are three communities denoted by the dashed circles.

2.2.2.8 Motifs

A motif M is a pattern of interconnections occurring either in both of undirected and directed graph G at numerous significantly higher than in randomized graph. As a pattern of interconnections, the motif M is usually meant as a connected undirected or directed n-node graph which is a subgraph of G. An example of all the possible 3-nodes connected directed graphs is shown in Fig. 2.5 [27]. Alon and his coworkers originally introduced the concept of motifs in biological and other networks [70, 71, 72, 73, 74]. The study of the significant motifs in a graph G is based on matching algorithms counting the total number of occurrences of each n-node subgraph M in the original graph and in the randomized ones. Z-score describes the statistical significance of the motif, defined as

$$Z_M = \frac{n_M - \langle n_M^{rand} \rangle}{\sigma_{n_M}^{rand}}, \tag{2.12}$$

where n_M is the number of times the subgraph M appears in the graph G, $\langle n_M^{rand} \rangle$ and $\sigma_{n_M}^{rand}$ are the mean and the standard deviation of the number of appearances in the randomized network, respectively.

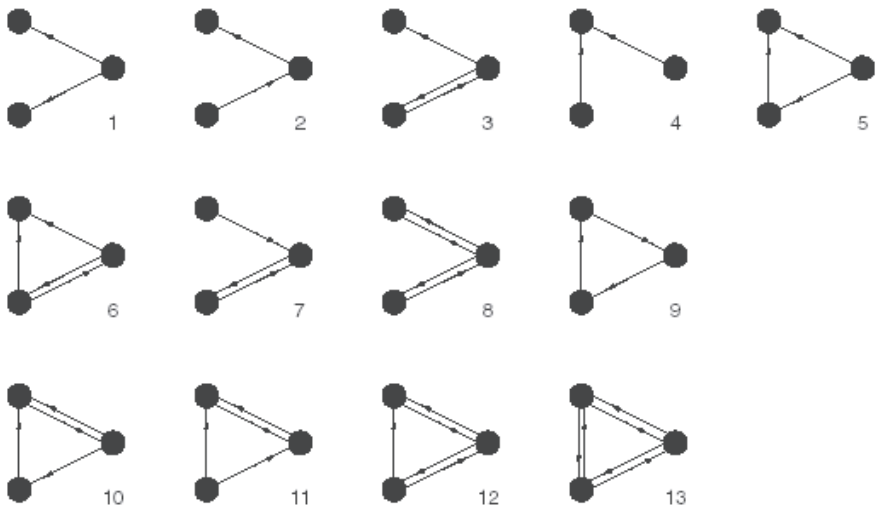

Fig. 2.5: All the 13 types of motifs involved in three-nodes directed connected subgraphs.

2.2.2.9 Graph spectra

The spectrum of a graph is the set of eigenvalues of its adjacency matrix $\mathcal{A}$. A graph $G_{N,K}$ has N eigenvalues λ_i $(i = 1, 2, \ldots, N)$, and N associated eigenvectors $\mathbf{V}_i$ $(i = 1, 2, \ldots, N)$. When G is undirected, without loops or multiple edges, $\mathcal{A}$ is real and symmetric, therefore the graph has real eigenvalues $\lambda_1 \leq \lambda_2 \leq \cdots \lambda_N$, and the eigenvectors corresponding to distinct eigenvalues are orthogonal. When G is directed, the eigenbalues can have imaginary part in the tournament graph with 3 nodes. It is graph invariant, and is a better way to compare graphs than $\mathcal{A}$ itself, which depends on the graph labeling. It also facilitates to define the graph spectral density [75], originating from the random matrix theory [97], as follows:

$$\rho(\lambda) = \frac{1}{N} \sum_{i=1}^{N} \delta(\lambda - \lambda_i), \tag{2.13}$$

which approaches a continuous function as $N \to \infty$ The eigenvalues and associated eigenvectors of a graph are intimately related to important topological features such as the diameter, the number of cycles, and the connectivity properties of the graph. Since its kth moment can be written as

$$\frac{1}{N} \sum_{j=1}^{N} (\lambda_j)^k = \frac{1}{N} \sum_{i_1, i_2, \ldots, i_k} \mathcal{A}_{i_1, i_2} \mathcal{A}_{i_2, i_3} \cdots \mathcal{A}_{i_k, i_1}, \tag{2.14}$$

i.e., the number of paths returning to the same node in the graph. It is worth to noticing that these paths can include nodes that were already visited.

From the Perron-Frobenius theorem, we know that a graph has a real eigenvalue λ_N

associated to a real non-negative eigenvector, and $|\lambda| \leq \lambda_N$ for each eigenvalue λ. For connected graph, λ_N has multiplicity 1 and $|\lambda| < \lambda_N$ for all eigenvalues λ different from λ_N. When nodes or edges are removed from the graph, the value of λ_N decreases. For a connected undirected graph, the largest eigenvalue λ_N is not degenerate, and every element of the corresponding eigenvector $\mathbf{V}_N$ is non-negative. All other eigenvectors have entries with mixed signs, as they are orthogonal to the eigenvector $\mathbf{V}_N$ corresponding to λ_N. The same theorem also states that either $k_{min} < \langle k \rangle < \lambda_N < k_{max}$, or $k_{min} = \langle k \rangle = \lambda_N = k_{max}$ in a connected graph.

For a random graph $G_{N,p}$ (p is the connection probability), it satisfies $p(N) = cN^{-z}$. When $z < 1$, there is an infinite cluster in the graph (cluster is a component of a graph and is a connected, isolated subgraph), and if $N \to \infty$, any node belongs to the infinite cluster. In this condition, the random graph's spectral density converges to a semicircular distribution (see Fig. 2.6 [76])

$$\rho(\lambda) = \begin{cases} \frac{\sqrt{4Np(1-p)-\lambda^2}}{2\pi Np(1-p)} & \text{if} \quad |\lambda| < 2\sqrt{Np(1-p)} \\ 0 & \text{otherwise} \end{cases} . \tag{2.15}$$

As Wigner's law [77, 78, 79] or the semicircle law, the above equation has well been applied to the quantum, solid state, and statistical physics [80, 81, 82]. In principle, the largest eigenvalue is isolated from the bulk of the spectrum, and it grows with the network size as pN. When $z > 1$, the spectral density is far away from the semicircle law. The important property of $\rho(\lambda)$ is that the odd moments are equal to zero, which means that if a path wants to come back to the original node, the only way is if it returns exactly following the same nodes.

2.3 Topology of real networks

Most of the interesting features of real-world networks that have attracted the attention of researchers in the last few years however concern the ways in which networks are not like random graphs. These networks in nature and in technology consist of a block of highly interconnected dynamical elements, such as social interacting species, coupled biological and chemical systems, the WWW. Real networks are nonrandom in some revealing ways that suggest some possible mechanisms. The first method to capture the global properties of such systems is to model them as graphs in which nodes mean the dynamical elements and the edges represent the interactions between the elements. During the last decade, the

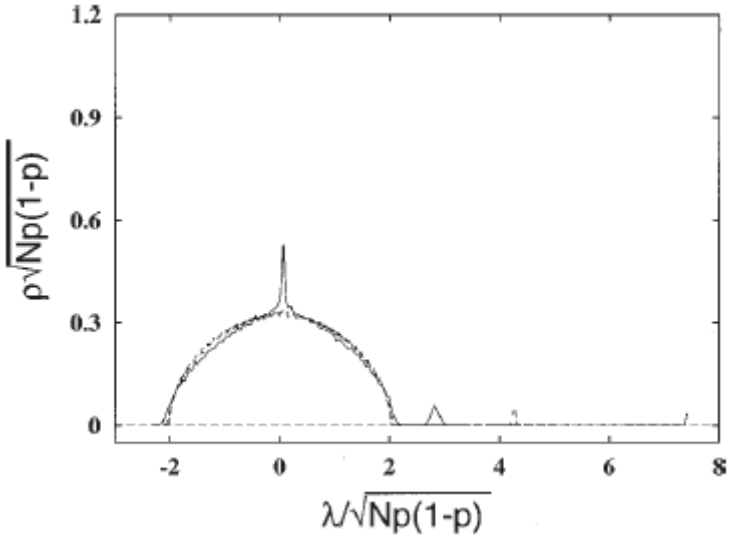

Fig. 2.6: Rescaled spectral density of three random graphs with $p = 0.05$, $N = 100$ (solid line), $N = 300$ (long-dashed line), and $N = 1000$ (shor-dasheline). The isolated peak corresponds to the principal eigenvalue.

increasing availability of the large databases, the optimized rating of computing facilities, as well as the achievements of reliable and powerful data analysis tools, have constituted a better machinery to investigate the topological features of many complex systems. In this section we discuss briefly the most significant topological properties, such as the small-world effect, and degree scale-free distributions.

2.3.1 The small-world property

In Ref. [1], they described the famous experiments carried out by Stanley Milgram in the 1960s, in which letters passed from person to person were able to reach a designated target individual in only a small number of steps around six in the published cases [7, 83, 84, 85]. This result is one of the first direct demonstrations of the small-world effect, the fact that most pairs of vertices in most networks seem to be connected by a short path through the network and is mathematically characterized by an average shortest path length L, defined as in Eq. (2.9), which depends at most logarithmically on the network size N. Nowadays, the small-world effect has been studied and verified directly in a large number of different networks [8, 51, 86, 87].

The small-world effect has obvious implications for the dynamics of processes taking place on networks. For example, if one considers the spread of information, or indeed

anything else, across a network, the small-world effect implies that this spread will be fast on most real-world networks. If it takes only six steps for a rumor to spread from any person to any other, for instance, then the rumor will spread much faster than if it takes a hundred steps, or a million. This affects the number of "hops" a packet must make to get from one computer to another on the Internet, the number of legs of a journey for an air or train traveler, the time it takes for a disease to spread throughout a population, and so forth.

On the other hand, at variance with random graphs, the small-world property in real network is often related to the presence of clustering, represented by high values of the clustering coefficient C, defined as in Eq. (2.7). For this reason, Watts and Strogatz have defined small-world networks as those networks having both a small value of average path length L, and a high clustering coefficient C. In the efficiency-based formalism, such a definition corresponds to networks with a high global efficiency $E_{\rm glob}$, and a high local efficiency $E_{\rm loc}$, i.e. to networks extremely efficient in exchanging information both at a global and at a local scale [88, 89].

2.3.2 Scale-free degree distributions

Exploring several large databases describing the topology of large networks, that span as diverse fields as the WWW or the citation patterns in science, recently Barabási and Albert have demonstrated [28], that independently of the nature of the system and the identity of its constituents, the probability $P(k)$ that a vertex in the network is connected to k other vertices decays as a power-law, following $P(k) \sim Ak^{-\gamma}$, with exponents varying in the range between 2 and 3. The average degree $\langle k \rangle$ in such networks is therefore well defined and bounded, however the variance $\sigma^2 = \langle k^2 \rangle - \langle k \rangle^2$ is dominated by the second moment of the distribution that diverges with the upper integration limit $k_{\rm max}$ as

$$\langle k^2 \rangle = \int_{k_{\rm min}}^{k_{\rm max}} k^2 P(k) \sim k_{\rm max}^{3-\gamma}. \tag{2.16}$$

Such networks have been called as scale-free networks [9, 28], since power-law is the same functional form for all scales. In fact, power-law is the only functional form $f(x)$ that remains invariable, apart from a multiplicative factor, under a rescaling of the independent variable x, being the only solution to the equation $f(ax) = bf(x)$. Power-law has a special role in statistical physics as their connections to the fractals [90] and phase transitions [91]. In scale-free networks, we represent a class of graphs with power-law in the degree

distribution. Of course, this does not necessarily denote that such graphs are scale-free with regard to other measurable structural properties [92]. These networks with highly inhomogeneous degree distribution, give rise to the simultaneous presence of a few hubs connected to many other nodes, and a large number of poorly connected nodes.

In finite-size networks, fat-tail degree distributions have natural cut-offs [93]. When we analyze the real networks, the data may have a very strong intrinsic noise due to the finiteness of the sample. Hence, when the size of the data is small, the degree distribution $P(k)$ will have big fat-tail. It is more reliable to measure the cumulative degree distribution $P_{\rm cum}(k)$, defined as $P_{\rm cum}(k) = \sum_{k'=k}^{\infty} P(k')$. Indeed, when summing up the original distribution $P(k)$, the statistical fluctuations generally present in the tails of the distribution are smoothed. Consequently, if $P(k) \sim k^{-\gamma}$, the relationship between the exponent γ and the slope of $P_{\rm cum}(k)$ in a log-log plot follows $\gamma = 1 + \gamma_{\rm cum}$. Another possible approach is that of performing an exponential binning of data [94].

2.4 Classical network models

2.4.1 Random graphs

2.4.1.1 Definition

The systematic study of random graphs was initiated by Erdős and Rényi in 1959 with the original purpose of studying [3, 95, 96] , by means of probabilistic methods, the properties of graphs as a function of the increasing number of random connections. The term random graph refers to the disordered nature of the arrangement of links between different nodes. Erdős and Rényi define a graph model in order to describe real world networks, which is called ER random graph model and denoted as $G_{N,M}$ [3]. Starting with N disconnected nodes, ER random graphs are generated by connecting couples of randomly selected nodes, prohibiting multiple connections, until the number of edges equals M, which are randomly chosen from the possible $\frac{N(N-1)}{2}$ edges. An alternative model for ER random graphs consists of N separated nodes, where each couple of nodes is linked with a probability p $(0 < p < 1)$ (see Fig. 2.7). Then the mean value of total number of edges in a $G_{N,p}$ graph is $M(p) = p\frac{N(N-1)}{2}$. In other words, $G_{N,p}$ is a group of all such graphs where a graph with e edges is realized with probability $p^M(1-p)^{\frac{N(N-1)}{2}-M}$. [3, 34, 95] The two models have a strong analogy, respectively, with the canonical and grand canonical ensembles

in statistical mechanics [98], and coincide in the limit of large N. Notice that the limit $N \to \infty$ is taken at fixed $\langle k \rangle$, which corresponds to fixing $2M/N$ in the first model and $p(N-1)$ in the second one. Although the first model seems to be more pertinent to applications, analytical calculations are easier and usually are performed in the second model.

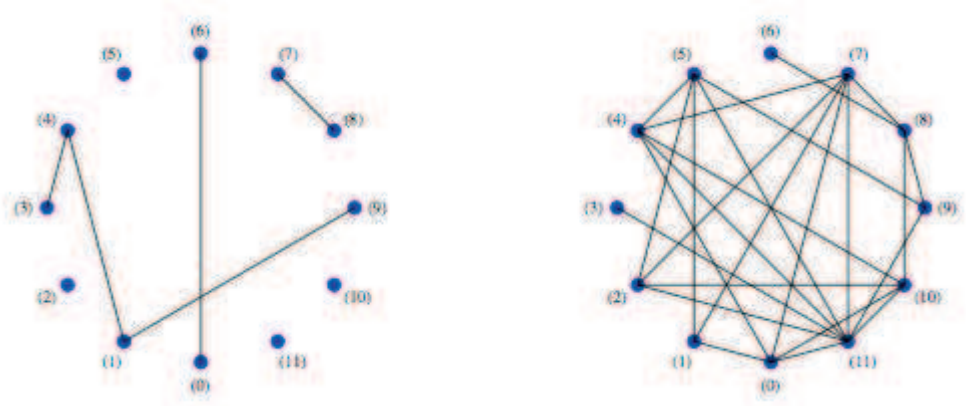

Fig. 2.7: Illustration of random graphs with $N = 12$ vertices and different connection probabilities $p = 0.0758$ (left) and $p = 0.3788$ (right).

2.4.1.2 Degree distribution

The average degree of the ER model is $\langle k \rangle = p(N-1) \sim pN$ for large N. The degree distribution $P(k)$ obeys a binomial distribution:

$$P(k) = \binom{N-1}{k} p^k (1-p)^{N-1-k}, \tag{2.17}$$

where p^k is the probability for the existence of k edges, $(1-p)^{N-1-k}$ is the probability for the absence of the remaining $N-1-K$ edges and $C_{N-1}^k = \binom{N-1}{k}$ is the number of different ways of choosing the end points of the k edges. If $N \sim \infty$, the degree distribution will be approximated by a Poisson distribution

$$P(k) = e^{-p(N-1)} \frac{p(N-1)^k}{k!} = e^{-\langle k \rangle} \frac{\langle k \rangle^k}{k!}. \tag{2.18}$$

The Poisson distribution indicates that a random graph can be characterized by an average degree without an obvious deviation of any other degree (see Fig. 2.8). The random graph model can be easily generalized to contain arbitrary degree distribution [99, 100], however, generalized random graphs can not be used to explore the origin of such distributions.

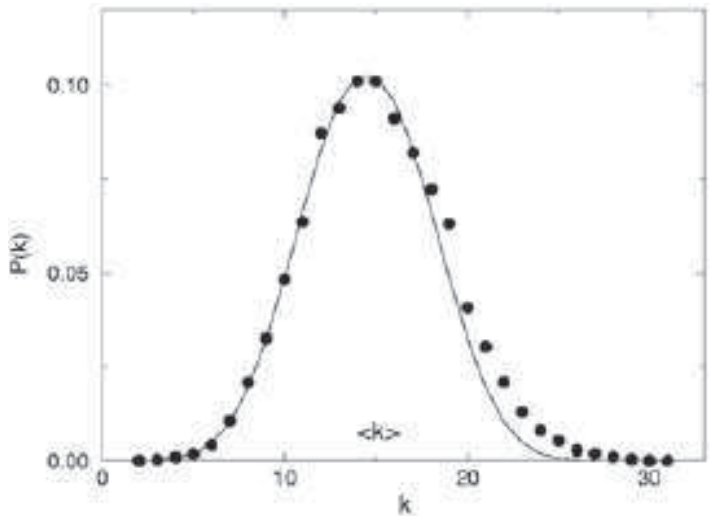

Fig. 2.8: Degree distribution resulting from numerical simulations of a random graph with $N = 10^5$ nodes and connection probability $p = 0.0015$. The picture compares $P(k)$ with the average value of the Poisson distribution, which suggests that the deviations are trivial.

2.4.1.3 Topological properties

The structural properties of ER random graphs vary as a function of p showing, in particular, a dramatic change at a critical probability $p_c = \frac{1}{N}$, corresponding to a critical average degree $\langle k \rangle_c = 1$. Erdős and Rényi proved that [34, 101]:

- If $\langle k \rangle \simeq pN < 1$ the edge density is low, isolated trees exist in the graph, and the cluster size distribution is exponential. The diameter (here the longest path length) of the graph is commensurate with the diameter of a tree.

- If $\langle k \rangle > 1$ a giant component appears which holds most of the nodes. The diameter of the graph is equal to the linear size of the giant cluster.

- If $\langle k \rangle > \ln N$ the graph is completely connected.

We can calculate the size S of the giant component from the following simple heuristic argument, which plays the order parameter role in the phase transition [1, 99, 102]. Let w be a randomly chosen node on the graph which does not belong to the giant cluster. In other words, $P_w(k) = w^k$ is the probability that the chosen node with degree k does not have any neighbors in the giant cluster. Averaging it over $P(k)$, we then find the following

self-consistency relation for w in the limit of large graph size:

$$w = \sum_{k=0}^{\infty} p_k w^k = e^{-\langle k \rangle} \sum_{k=0}^{\infty} \frac{(\langle k \rangle w)^k}{k!} = e^{\langle k \rangle (w-1)}. \tag{2.19}$$

The fraction S of the graph occupied by the giant component is $S = 1 - w$ and hence

$$S = 1 - e^{-\langle k \rangle S}. \tag{2.20}$$

By an argument only slightly more complex, which we give in the following section, we can show that the mean size $\langle s \rangle$ of the component to which a randomly chosen nodes belongs (for non-giant components) is

$$\langle s \rangle = \frac{1}{1 - \langle k \rangle - \langle k \rangle S} \tag{2.21}$$

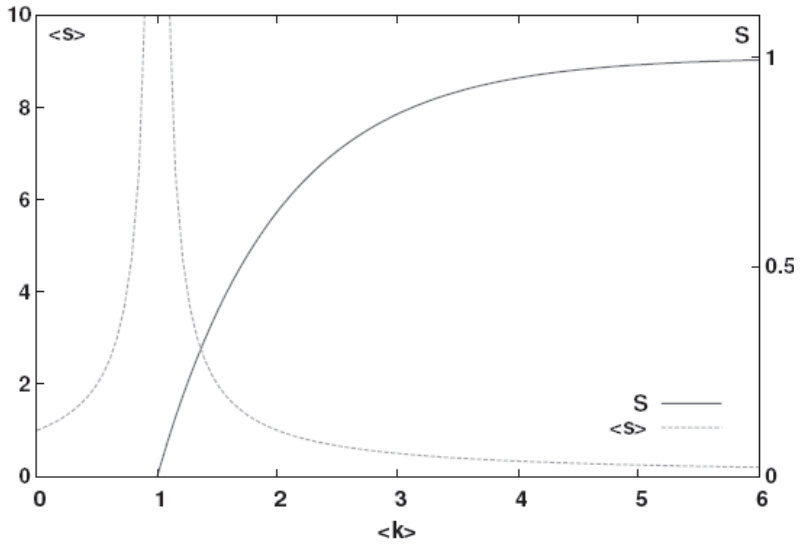

Fig. 2.9: Presentation of the topological phase transition in random graphs. Increasing the average degree per site, the S size of giant component grows at the critical point and the average component size $\langle s \rangle$ diverges.

The form of these two quantities is shown in Fig. 2.9 [1]. Eq. (2.21) is transcendental and has no closed-form solution. It is simple to see that if $\langle k \rangle < 1$, the only non-negative solution of S is $S = 0$ and if $\langle k \rangle > 1$ the finite solution is related to the size of the giant cluster. The phase transition point is in the position of $\langle k \rangle = 1$ where $\langle s \rangle$ diverges and the giant component appears in the graph. Fig. 2.9 shows such kind of transition.

ER random graph is completely non-correlated graph, so the connection between nodes in the graph is unrelated to their degree. When $p \geq lnN/N$, the random graph is almost

completely connected, the value of the graph diameter is around $D = lnN/ln(pN) = lnN/ln\langle k\rangle$. The average shortest path length of the graph is also the function of the size N:

$$L \sim \frac{lnN}{ln\langle k\rangle}, \tag{2.22}$$

which means that the average shortest path length of random graph is the increasing function of lnN, and is one of the feature of small-world effect. However, the probability of the connection between any two nodes in the ER random graph is identical and regardless of that whether they have a common neighbor, hence, the clustering coefficient is:

$$C = p = \frac{\langle k\rangle}{N}. \tag{2.23}$$

So, in the ER random graph, with the increasing of the network size N, the clustering coefficient decreases gradually. When $N \to \infty$, it will tend to zero, which is far smaller than those of networks in the real world.

2.4.2 Small-world model

Random graphs with arbitrary degree distribution show no clustering in the thermodynamic limit, in contrast to real world networks. It is therefore important to find methods to generate graphs which have a finite clustering coefficient and, at the same time, the small world property.

2.4.2.1 Definition

Wattz and Strogatz [8, 86, 103] have proposed a small-world model (WS model) which interpolates smoothly in between a regular lattice and an ER random graph, called small world model (see Fig. 2.10(a)).

Starting from a completely ordered graph (for simplify, a one-dimensional model, i.e. a ring lattice, is considered) consisted of n nodes each linked to K nearest neighbors, each link of the network is randomly rewired with probability p. We move one end of every link with the probability p to a new position chosen at random from the rest of the lattice. Obviously, when p is small, the situation has to be close to the original regular lattice. For large enough p, the network is similar to the classical random graph. Note that the periodical boundary conditions are not essential.

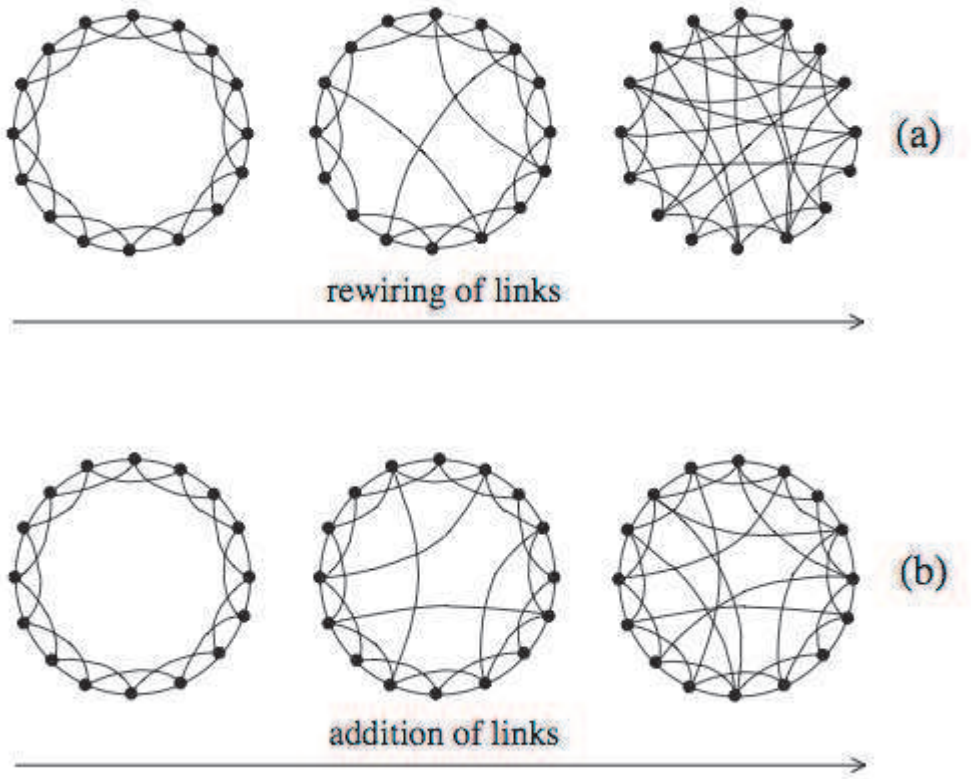

Fig. 2.10: Small-world networks in which the crossover from a regular lattice to a random network is realized. (a) The original Watts-Strogatz model with the rewiring of links. (b) The network with the addition of shortcuts is proposed by Newman and Watts.

A variation of the WS model has been suggested by Newman and Watts [104, 105]. Instead of rewiring links between sites as in Fig. 2.10(a), extra links, also called "shortcuts", are added between pairs of sites chosen at random, but no links are removed from the underlying lattice, see Fig. 2.10(b). This model is somewhat easier to analyze than the original WS model, because it is not possible for any region of the graph to become disconnected from the rest, whereas this can happen in the original model.

The small-world models illustrated in Fig. 2.10, have an intuitive justification for social networks. Most people are friends with their immediate neighbors. Neighbors on the same street, people that they work with or their relatives. But some people are also friends with a few far-away persons. Far away in a social sense, like people in other countries, people from other walks of life, acquaintances from previous eras of their lives, and so forth. These long-distance acquaintances are represented by the long-range links in the small-world models illustrated in Fig. 2.10.

2.4.2.2 Clustering coefficient

The WS model displays this duality for a wide range of the rewiring probability p. When $p = 0$, the clustering coefficient, which is large at the initial regular graph, does not depend on the size of the lattice but only on its topology. Since it depends only on the coordination number $z = 2K$ of the lattice. If the rewiring is introduced into the system by rewired edges, it is almost close to $C(p = 0)$, as long as a large fraction of original neighbors keep connected. The probability that three nodes which were connected at $p = 0$ still construct a triangle is $(1 - p)^3$ when $p > 0$, since there are three edges that need to keep intact. Barrat and Weigt [109] gave the formula as follows:

$$C(p) = C(0)(1 - p)^3 = \frac{3K - 3}{4K - 2}(1 - p)^3, \tag{2.24}$$

while for other definition, in which without rewiring, only shortcuts are added to the networks, Newman [106] gave that:

$$C(p) = \frac{3K - 3}{4K - 2 + 4Kp(p + 2)}. \tag{2.25}$$

2.4.2.3 Average path length

As we discussed above, in the WS model there is a change in the scaling of the characteristic path length L as the fraction p of the rewired edges is increased.In the limit $p \to 0$, the typical average path length L tends to $L = N/4K$, the model is "large world". In contrary, for large p, small-world behavior is typically characterized by logarithmic scaling $L \sim logN$, where the model becomes like a random graph. In between these two limits, there is a crossover regime. A widely accepted explanation that L satisfies a scaling relation of the form [107]:

$$L \sim \zeta g(N/\zeta), \tag{2.26}$$

where ζ is the correlation length that depends on p, and $g(x)$ is an unknown but universal scaling function that only depends on system dimension and lattice geometry, but not on N, ζ or p. The variation of ζ defines the crossover from large-world to small-world behavior. The behavior of L for small and large N, is reproduced by having ζ diverge as $p \to 0$ and

$$g(x) \sim \begin{cases} x & \text{for } x \gg 1 \\ \log x & \text{for } x \ll 1. \end{cases} \tag{2.27}$$

Eq.(2.26) has been corrected by a renormalization group treatments. From this treatment, we can derive the scaling form for L satisfying:

$$L \sim \frac{N}{K} f(NKp), \tag{2.28}$$

which is similar to Eq. (2.26), except for a factor of K, if $\zeta = 1/Kp$ and $g(x) = xf(x)$. According to the scaling form in above equation, the graph can get the transition from the large world regime to small world one by increasing p or by increasing N, since NKp is equal to the average number of shortcuts in the model. In the further progress, we would like to calculate the scaling function $f(x)$. For the normal small-world model, there is no exact solution, although some additional exact scaling forms have been found. A mean-field treatment of the model has been given by Newman [108], which approximately indicates that $f(x)$ is:

$$f(x) = \frac{1}{2\sqrt{x^2 + 2x}} \tanh^{-1} \frac{x}{\sqrt{x^2 + 2x}} \tag{2.29}$$

and then the average path length is

$$L = \frac{\zeta}{2K\sqrt{1 + 2\zeta/N}} \tanh^{-1} \frac{1}{1 + 2\zeta/N}. \tag{2.30}$$

The key result produced by small-world model is represented by the behavior of clustering coefficient and average path length, which is determined by the rewiring probability p and shown in Fig. 2.11. Starting with a K-regular network with $p = 0$, we study the behavior of both C and $\langle L \rangle$, when p is increasing. The key result is, when $p \approx 0.1 - 1$, the clustering coefficient remains nearly invariable, however, the average path length indicates an obvious decrease, reaching almost the value corresponding to a random graph. This transition arises from the emergence of short cuts, decreasing the average distance.

2.4.2.4 Degree distribution

Small-world model is still short of some of the peculiar features originated from empirical observations, leading to relatively homogeneous networks characterized by a Poissonian degree distribution. In the WS model for $p = 0$ each node has the same degree K. Therefor the degree distribution is a delta function centered at K. For $p > 0$ the degree distribution follows a Poisson distribution, since the network is homogeneous and relaxes into a random graph. During the rewiring process, only one end of a chosen link is reconnected, so each node has at least $K/2$ neighbors in the end. It suggests that for $k > 2$ there are no isolated nodes and the network is usually connected, unlike a random

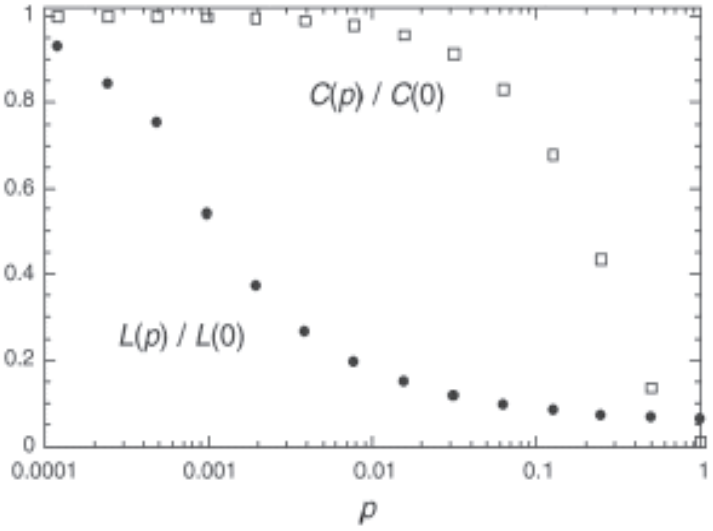

Fig. 2.11: Average clustering coefficient and average shortest path length, normalized to corresponding values obtained for regular lattices, as a function of the rewiring probability p. The high drop in $\langle L \rangle$ is due to the emergence of short-cuts in the networks, and occurs when the clustering coefficient C remains almost constant.

graph which consists of isolated clusters for a wide range of connection probabilities. The real degree of a node i can be rewritten as $k_i = K/2 + n_i$, with $n_i \geq 0$, where n_i can be divided into two kinds : it includes $n_i^1 \leqslant K/2$ edges which the rewiring process left unvaried with probability $(1-p)$ and $n_i^2 = n_i - n_i^1$ number of edges which have been rewired towards i (each with probability $p/(N-1)$) during the process. The probability distribution of these two kinds of edges can be written as:

$$P_1(n_i^1) = \binom{\frac{K}{2}}{n_i^1}(1-p)^{n_i^1} p^{\frac{K}{2}-n_i^1} \qquad \text{and} \qquad P_2(n_i^2) = \frac{(\frac{Kp}{2})^{n_i^2}}{n_i^2!} e^{-\frac{pK}{2}}. \tag{2.31}$$

if $N \gg 1$, the complete degree distribution becomes:

$$P_{WS}(k) = \sum_{n=0}^{min(k-\frac{K}{2},\frac{K}{2})} \binom{\frac{K}{2}}{n}(1-p)^n p^{\frac{K}{2}-n} \times \frac{(\frac{Kp}{2})^{k-\frac{K}{2}-n}}{(k-\frac{K}{2}-n)!} e^{-\frac{pK}{2}} \quad \text{if} \quad k \geq K/2. \tag{2.32}$$

The shape of the degree distribution is similar to that of a random graph, since in the latter condition the graph can contain isolated components, however a small-world network is always connected.

2.4.2.5 Spectral properties

The spectral density $\rho(\lambda)$ of a graph reveals important information about its topology. It comes as no surprise that the spectrum of the WS model depends on the rewiring

probability p. For $p = 0$ the network is regular and periodical; consequently $\rho(\lambda)$ contains numerous singularities as Fig. 2.12(a). For intermediate values of p, the singularities become blurred, but $\rho(\lambda)$ retains a strong skewness as Figs. 2.12(b) and 2.12(c). Finally, when $p \to 1$, $\rho(\lambda)$ approaches the semicircle law characterizing random graphs as Fig. 2.12(d). Thus the results in Fig. 2.12 allow us to conclude that a high number of triangles is a basic property of the WS model.

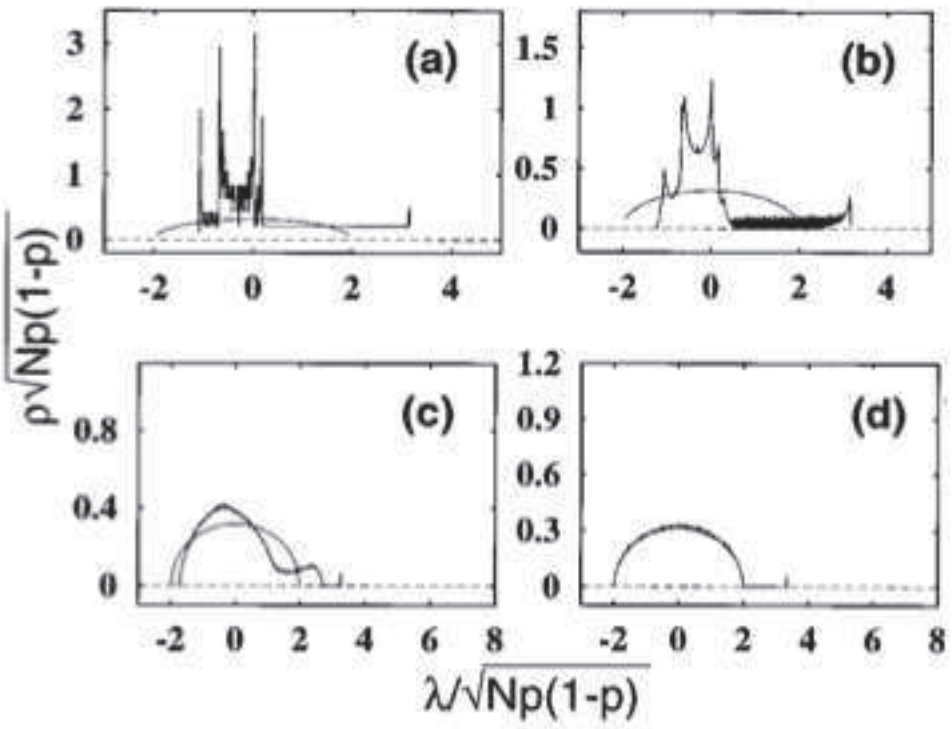

Fig. 2.12: Spectral density of small-world networks, compared to the semicircle law corresponding to random graphs (solid line). The rewiring probabilities are (a) $p = 0$; (b) $p = 0.01$; (c) $p = 0.3$; (d) $p = 1$.

2.4.3 Scale-free model

Most real world networks are formed by the continuous addition of new nodes to the system. The number of nodes, N, increases throughout the lifetime of the network, as it is the case for the WWW, which grows exponentially by the continuous addition of new web pages. The small world networks discussed in Section 2.4.2 are however constructed for a fixed number of nodes N, growth is not considered. In the real world, the size of the network evolutes with time.

2.4.3.1 Barabási-Albert model

In random network models, the probability that two vertices are connected is random and uniform. In contrast, most real networks exhibit the "rich-get-richer" phenomenon. Let us demonstrate the growth of a network with preferential linking using, as the simplest example, the Barabási-Albert model (the BA model) [9]. When the probability for a new node to connect to any of the existing nodes is not uniform, we speak of preferential connectivity. A newly created web page, to give an example, will make links to well known sites with a quite high probability. Popular web pages will therefore have both a high number of incoming links and a high growth rate for incoming links. Growth of nodes in terms of edges is therefore in general not uniform.

These two ingredients, growth and preferential attachment, inspired the introduction of the BA model, which led for the first time to a network with a power-law degree distribution. The algorithm of the BA model is the following:

(1) *Growth*: Starting with a small number (m_0) of nodes, at every time step, we add a new node with $m(\leqslant m_0)$ edges that link the new node to m different nodes already present in the system.

(2) *Preferential attachment*: We assume that the new nodes are selected with probability $\prod$. A new node will be connected to node i depends on the degree k_i of node i, such that

$$\prod(k_i) = \frac{k_i}{\sum_j k_j}. \tag{2.33}$$

After t time steps this procedure results in a network with $N = t + m_0$ nodes and mt edges. Numerical simulations indicated that this network evolves into a scale invariant state with the probability that a node has k edges following a power law with an exponent $\gamma_{BA} = 3$ (see Fig. 2.13). The scaling exponent is independent of m, the only parameter in the model.

2.4.3.2 Theoretical approaches

The dynamical properties of the scale-free model can be addressed using various analytic approaches. The BA model can also be solved exactly using theoretical approaches in the limit of large network size N [9, 111, 112, 113], thus confirming results are obtained numerically. These approaches include the continuum theory proposed by Barabási and

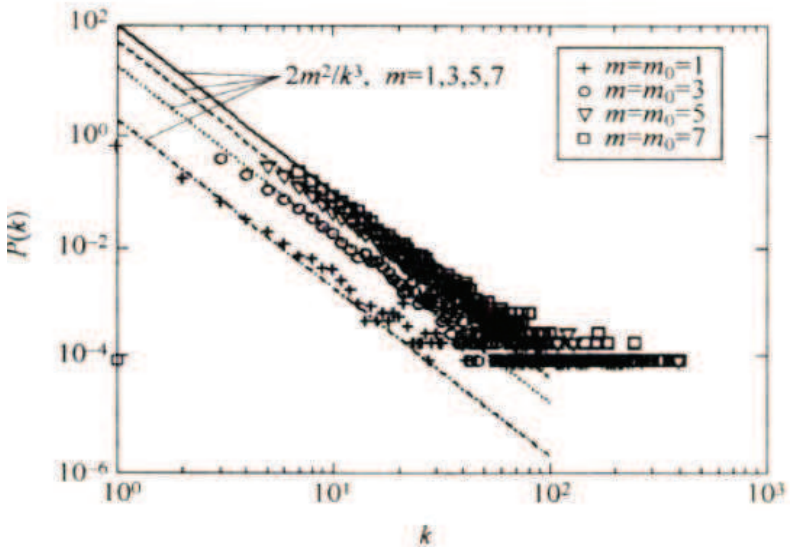

Fig. 2.13: Degree distribution of BA model for various m with system size $N = 3 \times 10^5$. The slope of the skew line is $\gamma = 3$.

Albert [9], the master equation approach proposed by Dorogovtsev, Mendes, and Samukhin [113], and the rate-equation approach introduced by Krapivsky, Redner, and Leyvraz [112]. We will give a brief introduction of them in this section.

The continuum approach focuses on the time dependence of the degree k_i of a given node i introduced by Barabási and Albert [9] and Barabási, Albert and Jeong [111]. This degree increases every time when a new node is added into the system and links to a given node i. The probability of this process is $\prod(k_i)$, following Eq. (2.33). Assuming that k_i is a continuous real variable, the rate at which k_i changes is expected to be proportional to $\prod(k_i)$. Consequently, k_i satisfies the dynamical equation:

$$\frac{\partial k_i}{\partial t} = m\Pi(k_i) = m\frac{k_i}{\sum_{j=1}^{N-1} k_j}. \tag{2.34}$$

The sum goes over all nodes except the newly added one in the system, thus the value is $\sum_j k_j = 2mt - m$, which results in

$$\frac{\partial k_i}{\partial t} = \frac{k_i}{2t}. \tag{2.35}$$

The solution of this equation is given with the initial condition that $k_i(t_i) = m$

$$k_i(t) = m\left(\frac{t}{t_i}\right)^{\beta}, \qquad \text{with} \quad \beta = \frac{1}{2}. \tag{2.36}$$

Eq. (2.36) suggests that the degree of all nodes evolves the same way as power law. The only difference is the intercept of the power law.

Based on Eq. (2.36), one can derive the probability that a node has a degree $k_i(t)$ smaller than k, $P(k_i(t) < k)$, as

$$P(k_i(t) < k) = P(t_i > \frac{m^{1/\beta}t}{k^{1/\beta}}). \tag{2.37}$$

Assuming that the nodes are added at equal time intervals to the network, the t_i values have a constant probability density:

$$P(t_i) = \frac{1}{m_0 + t}. \tag{2.38}$$

Substituting this into Eq. (2.37) we obtain:

$$P\left(t_i > \frac{m^{1/\beta}t}{k^{1/\beta}}\right) = 1 - \frac{m^{1/\beta}t}{k^{1/\beta}(t + m_0)}. \tag{2.39}$$

The degree distribution $P(k)$ can be obtained using:

$$P(k) = \frac{\partial P(k_i(t) < k)}{\partial k} = \frac{2m^{1/\beta}t}{m_0 + t}\frac{1}{k^{1/\beta+1}}, \tag{2.40}$$

When $t \to \infty$, we get:

$$P(k) \sim 2m^{1/\beta}k^{-\gamma}, \quad \text{with} \quad \gamma = \frac{1}{\beta} + 1 = 3, \tag{2.41}$$

which is independent of m, in agreement with the numerical results in Fig. 2.13.

As the power-law observed for real networks of different sizes, it is expected that a correct model should provide a time-independent degree distribution. From Eq. (2.40), it means that asymptotically. The degree distribution of the BA model is independent of time t and the system size $N = m_0 + t$. In other words, despite the continuous growth, the network reaches a stationary scale free state. In addition, Eq. (2.40) also suggests that the coefficient of the power law distribution is proportional to m^2.

The master equation approach calculates the probability $p(k, t_i, t)$ that at time t a node i introduced at time t_i has a degree k [113]. In the BA model, when a new node with m edges is added into the system, the degree of node i increases by 1 with a probability

$m\Pi(k) = k/2t$, otherwise, it keeps the same. Therefore, the master equation calculating $p(k, t_i, t)$ for the BA model has the form as follows:

$$p(k, t_i, t+1) = \frac{k-1}{2t} p(k-1, t_i, t) + \left(1 - \frac{k}{2t}\right) p(k, t_i, t). \tag{2.42}$$

The degree distribution can be written as

$$P(k) = \lim_{t\to\infty} \left(\sum_{t_i} p(k, t_i, t) \right) /t. \tag{2.43}$$

Eq. (2.43) means that $P(k)$ is the solution of the recursive equation

$$P(k) = \begin{cases} \frac{k-1}{k+2} P(k-1) & \text{for} \quad k \geq m+1 \\ 2/(m+2) & \text{for} \quad k = m \end{cases}, \tag{2.44}$$

giving

$$P(k) = \frac{2m(m+1)}{k(k+1)(k+2)}, \tag{2.45}$$

similar to Eq. (2.41) produced by the continuum theory.

For rate-equation approach, it focused on the average number $N_k(t)$ of nodes with k edges at time t [112]. When a new node is added into the system in the scale free model, the evolution of $N_k(t)$ satisfies

$$\frac{dN_k}{dt} = m \frac{(k-1)N_{k-1}(t) - kN_k(t)}{\sum_k kN_k(t)} + \delta_{k,m}, \tag{2.46}$$

where the first term on the right-hand side accounts for the new links that connect to nodes with $k-1$ links, thus increasing their degree to k. The second term on the right-hand side describes the new links connecting to nodes with k edges changing them into the nodes with $k+1$ edges, thus resulting in the decrease of the number of nodes with k links. The third term accounts for the new nodes with m links. According to the law of large numbers, the network has $N_k(t) = tP(k)$ and $\sum_k kN_k(t) = 2mt$. Substituting them into Eq. (2.46), the same recursive equation is obtained:

$$P(k) = \frac{k-1}{k+2} P(k-1). \tag{2.47}$$

Its solution is:

$$P(k) = \frac{2m(m+1)}{k(k+1)(k+2)} \sim 2m^2 k^{-3}, \tag{2.48}$$

which is the same result as predicted by the master equation approach.

These three approaches are completely equivalent and offer the same asymptotic results. They can be used interchangeably for calculating the scaling behavior of the degree distribution. Moreover, these approaches, not using a continuum assumption, seem more appropriate for obtaining exact results in more challenging network models.

2.4.3.3 Properties of the Barabási-Albert model

The average distance in the BA model is smaller than in a ER random graph with same N and $\langle k \rangle$, and increases logarithmically with N [28]. Analytical results predict a double logarithmic correction to the logarithmic dependence $L \sim logN/log(logN)$ [110]. Fig. 2.14 shows the clustering coefficient of the BA model with average degree $\langle k \rangle = 4$ and different sizes, compared with the clustering coefficient $C \simeq \langle k \rangle / N$ of a random graph. We find that the clustering coefficient of the BA network model is about five times higher than that of the random graph, and this factor slowly increases with the number of nodes. However, the clustering coefficient of the BA model decreases with the network size, following approximately a power law $C \sim N^{-0.75}$.

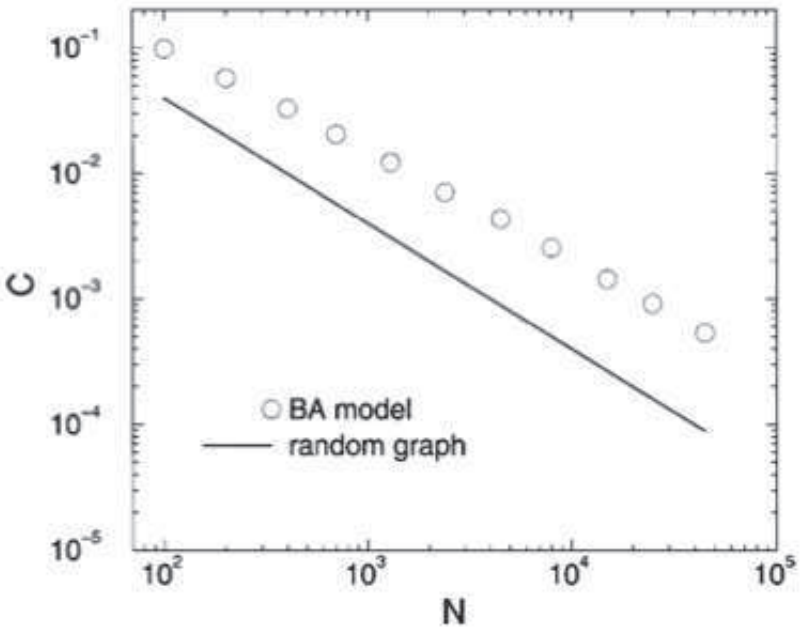

Fig. 2.14: Clustering coefficient versus size of the BA model with $\langle k \rangle$=4, compared with the clustering coefficient of a random graph, $C \simeq \langle k \rangle / N$.

The spectral density of the BA model is continuous, but it has a markedly different shape from the semicircular spectral density of random graphs [57, 76]. Numerical simulations indicate that the bulk of $\rho(\lambda)$ has a trianglelike shape with the top lying well above the semicircle and edges decaying as power-law (Fig. 2.15). This power-law decay is due to the eigenvectors localized on the nodes with the highest degree. As in the case of random graphs (and unlike small-world networks), the principal eigenvalue, λ_1, is clearly separated from the bulk of the spectrum. A lower bound for λ_1 can be given as the square root of the

network's largest degree k_1. The node degrees in the BA model increase as $N^{1/2}$; hence λ_1 increases approximately as $N^{1/4}$. Numerical results indicate that λ_1 deviates from the expected behavior for small network sizes, reaching it asymptotically for $N \to \infty$. This crossover indicates the presence of correlations between the longest row vectors, offering additional evidence for correlations in the BA model.

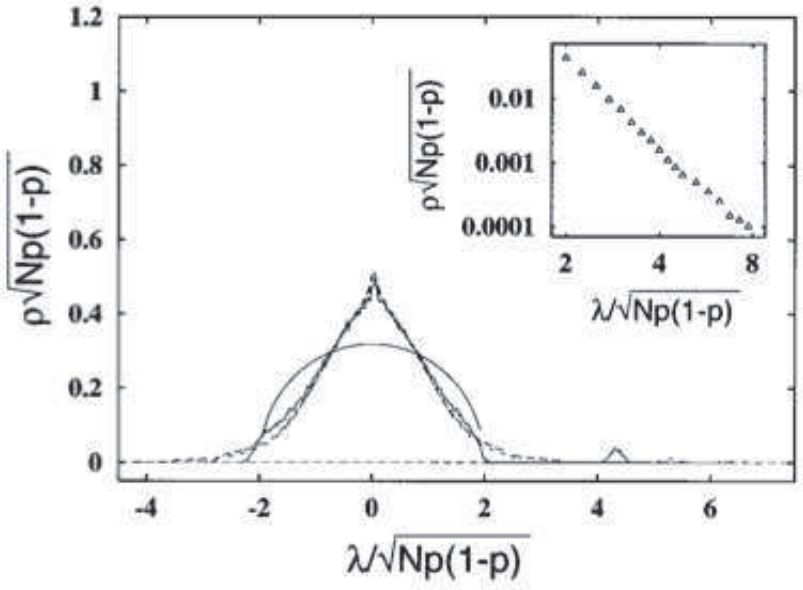

Fig. 2.15: Rescaled spectral density of three BA networks having $m = m_0 = 5$ and virous sizes N: solid line, $N = 100$; long-dashed line, $N = 300$; short-dashed line, $N = 1000$. The semicircle law corresponding to random graphs is drawn for comparison. The isolated peak corresponds to the largest eigenvalue, which increases as $N^{1/4}$. Inset: the edge of the spectral density decays as a power law.

The BA model has exceptional popularity, since it was first published, and many extensions appeared to make it more realistic e.g. introducing fitting parameter. Another modification when let the attachment probability $\Pi(k + k_0)$ to be dependent on an additional constant k_0, instead of k only. Modify k_0 between $-m < k_0 < \infty$, it scales the degree distribution [113, 114], which then goes as $P(k) \sim k^{-\epsilon}$, with $\epsilon = \gamma + k_0/m$, where $\gamma = 3$ is the original degree exponent. If one scales the attachment probability as a power-law $\Pi(k^{\varepsilon})$ and not linearly as before, when $\varepsilon = 1$ it is back to the original graph, however if $\varepsilon < 1$ or $\varepsilon > 1$, different network behaviors appear [112, 114]. Another extension of the model is given if we increase the average degree over time [115]. It seems to be realistic, since in the WWW the average degree increases by time as well. This modification changes the degree exponent too, and approximates the original BA network more close to real

systems.

Though the BA model gets great success, it is not able to capture all features, such as the behavior displayed by the clustering coefficient, or higher order correlations.

Chapter 3

Spatial network

3.1 Introduction

In the recent development of network sciences, spatial constraint networks have become an object of extensive investigation. These are the networks embedded in configuration space and influenced by spatial constraints. Recent findings have revealed that the spatial distance distribution follows power-law or exponential distribution [116, 118, 119, 133, 134]. These distributions are quite natural since, for instance, people tend to have their friends and relatives in their neighborhood, transportation networks often favor shorter distance trips, and many communication networks are mainly dominated by short radio ranges.

There are actually more and more observation of this kind of deviation. Barret *et al*. studied the North American airport network in Ref. [116]. Fig. 3.1 displays the distribution of the spatial distances of the direct flights, measured in kilometers. These distances correspond to Euclidean measures of the links and clearly show a fast decaying behavior reasonably fitted by an exponential. Masucci *et al*. [117] measured the street segment length distribution $P(l_1)$ in London, and found that the distribution decreases rapidly as $P(l_1) \sim {l_1}^{\gamma}$ with $\gamma \simeq 3.36$. In particular, as expected, space is also important in social networks. As a measure of the social tie, Lambiotte *et al*. [118] used mobile phone data for 3.3 millions customers in Belgium. The probability $P(d)$ that two connected individuals are separated by a spatial distance d follows $P(d) \sim d^{-2}$. Liben-Nowell *et al*. [119] studied one millions bloggers in the USA, which has the same phenomenon . The proportion of pairs of friends at distance d decays as $P(d) - \varepsilon \sim d^{-\alpha}$, where $\alpha \approx 1$ and $\varepsilon \approx 5 \times 10^{-6}$. In addition, this exponent α of order one is confirmed in another studies

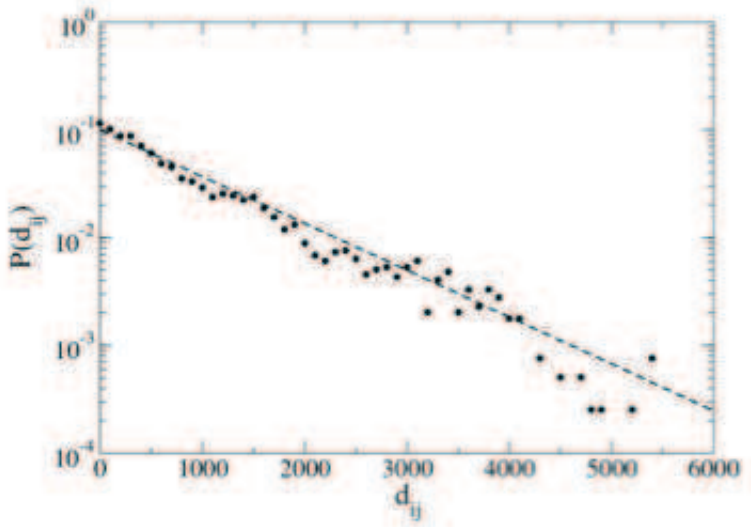

Fig. 3.1: Distribution of distances (in kms) between airports linked by a direct connection for the North American network. The straight line indicates an exponential decay with scale of order 1000 kms.

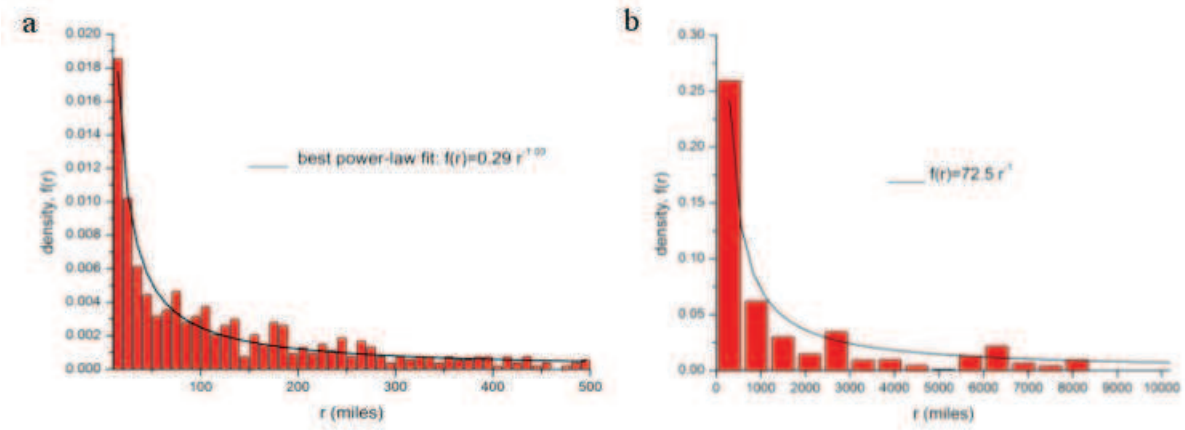

Fig. 3.2: Panel (a) describes the empirical density function of Facebook contacts as a histogram. The empirical distribution is in very good agreement with a scale-free power-law distribution. The solid line shows the best power-law fit to the empirical data, with an exponent of -1.03 and a standard error of 0.03. Panel (b) describes the empirical density function of Email contacts as a histogram. The solid line shows the power-law fit to the empirical data, with an exponent of -1.

[120] on Facebook users and on email communications (see Fig. 3.2).

In the following section, we will focus on the various models which attempt to reproduce these different effects. After that we will introduce a model with a kind of competition

between the degree and the spatial distance preferences. While the degree preferential attachment produces connections free from spatial constraints, the spatial distance preference favors closer connections. This competition is modulated by a parameter a.

3.2 Models of spatial network

To model these systems, scientists have proposed spatially constrained networks embedded in space. According to the generation rules, these networks can be categorized into five large classes [29]: The first class describes geometric graphs which are probably the simplest models of spatial networks. They are obtained for a set of vertices located in the plane and for a set of edges which are constructed according to some geometric condition (see for example [121, 122]). The second class concerns the ER model and its spatial generalization, including the spatial hidden variables models (see for example [123, 124]). These networks are obtained when the probability to connect two nodes depends on the distance between these nodes. The third class comprises spatial variants of the WS model and could be coined as spatial small-world (see for example [125, 126]). In these cases, the starting point is in d-dimensional lattice and random links are added according to a given probability distribution for their length. The fourth class concerns spatial growth models which can be considered as spatial extensions of the original growth model proposed by Barabási and Albert (see for example [127, 128, 129, 130, 131, 132, 133, 136, 137]). The last class concerns optimal networks obtained by the minimization of a 'cost' function.

In the first three classes, the network is stable; on the contrary, the fourth class is a growth network. We will focus on the fourth class in this chapter. This kind models have two base mechanisms, growth and preferential attachment. The process to generate such models always starts from a small 'seed' network. Then we introduce a new node n at each time step. This new node is allowed to make m connections towards nodes i with a probability $\Pi_{n\to i}$. In traditional preferential attachment, there is a propensity to connect a new node to an already well-connected one which is probably an important ingredient in the formation of various real-world networks. In spatial growth model, $\Pi_{n\to i}$ is a function of the spatial distance $r_{n,i}$ from node n to node i.

In one of the extended model, the network grows with addition of nodes randomly positioned in space. The new nodes connect to the existing nodes with probability $\Pi_{n\to i} \sim k_i r_{n,i}^{\alpha}$[131, 135]. Fig. 3.3 is the modulated example for this model. The case with $\alpha = 0$

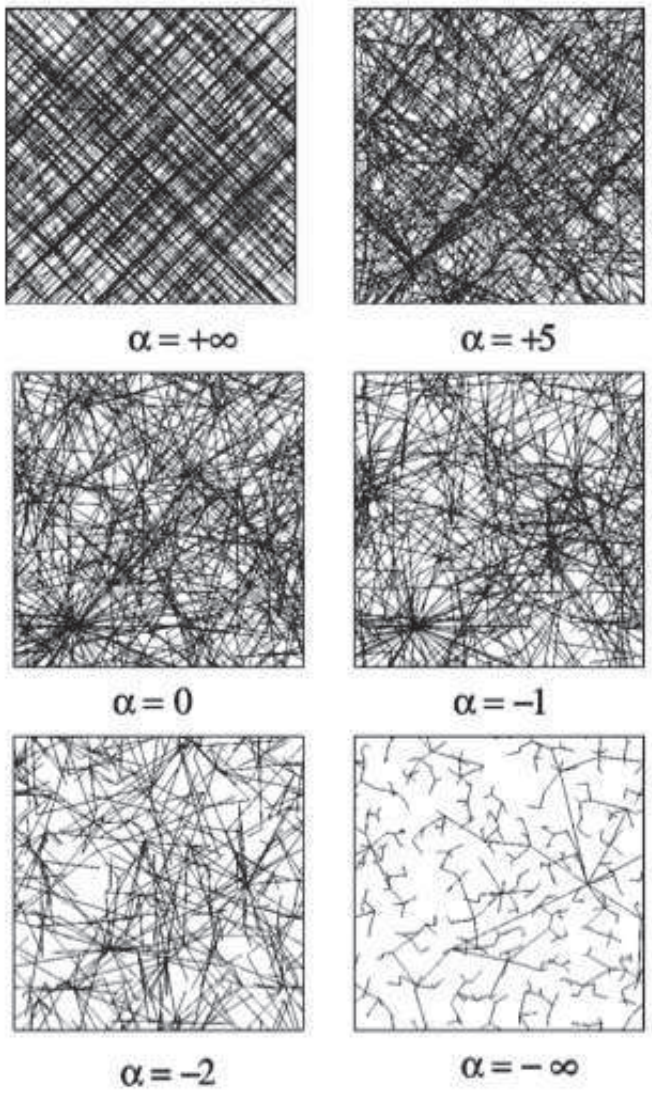

Fig. 3.3: Modulated scale-free networks within a unit square for different values of the modulation parameter α for the same distribution of 512 nodes. For $\alpha = +\infty$ and $-\infty$ a newly introduced node is linked only to its farthest and nearest predecessors, respectively, whereas for $\alpha = 0$ it is connected to one of the previous nodes according to the BA rule.

is the usual BA model. For the negative values of α, the largest value of the modulation factor $r_{n,i}^{\alpha}$ corresponds to the smallest value of r. Therefore, in the limit of $\alpha \to -\infty$, only the smallest value of r corresponding to the nearest node will contribute with probability 1. Similarly, for $\alpha > 0$ large r values will be more probable and the limit of $\alpha \to +\infty$ corresponds to only nonzero contribution from the furthest node. The distance distribution $p(r)$ is given by $p(r) \sim r^{-(\alpha-d+1)}$ as expected, where d is the dimension of the space. On the

other hand, the degree distribution is a power law for $\alpha > -1$ and a stretched exponential law for $\alpha < -1$. Another extension uses the connection probability $\Pi_{n\to i} \sim k_i^{\beta} r_{n,i}^{\alpha}$ and generates a power law degree distribution on a line in the $\alpha - \beta$ plane and in the zone limited by $\beta > 1$ and $\alpha < -0.5$ [127]. It is worth to notice that these results are got from the numerical simulation method only.

Since most of the real networks are growth and evolution with time, in this chapter, we are interested in the fourth class networks. We introduce a model with a kind of competition between the degree and the spatial distance preferences. While the degree preferential attachment produces connections free from spatial constraints, the spatial distance preference favors closer connections. This competition between short-range and long-range connections is modulated by a parameter.

3.3 The model

To construct the networks, the nodes are embedded on a one-dimensional ($d = 1$) ring of radius $R = 1/\pi$ or on a two-dimensional ($d = 2$) sphere of radius $R = 1/\pi$. The spatial distance r between a pair of nodes is defined as the shortest distance between them.

The model is constructed in the following way:

(1) *Initial condition*: We start with an initial state (t=m_0) of m_0+1 all-to-all connected nodes on the ring or the sphere.

(2) *Growth*: At every time step, a new node is added, which is randomly placed on the ring or the sphere.

(3) *Addition of edges*: The new node n connects with m ($m \leq m_0 + 1$) previous nodes, which are selected with the probability π_i

$$\pi_i = a\frac{k_i}{\sum_j k_j} + (1-a)\frac{r_{ni}^{-\alpha}}{\sum_j {r_{nj}}^{-\alpha}} \tag{3.1}$$

where k_i is the degree of node i, r_{ni} is the Euclidean distance between a new node n and a previous node i, $0 \leqslant a \leqslant 1$ and $0 \leqslant \alpha$. The growing process repeats steps (2) and (3) until the network reaches the desired size. Accordingly, at each step, the number of nodes increases by one, while the number of edges increases by m ($m = m_0 = 2$ in what follows if not mentioned).

This model has two limit cases : when $a = 1$, the network recovers the BA network

model, while the case of $a = 0$ and $\alpha = 0$ corresponds to the random growing process.

The numerical results described in this paper are the average of 20 simulations for different realization of networks under the same parameters with the network size of 10 000 nodes. We have also tried 50 000 nodes, but the result is almost the same.

3.3.1 Spatial driven model: $a = 0$

In this section, we focus on the behavior of pure spatial-driven model with $a = 0$ and $\alpha \neq 0$ in Eq. (3.1). The connection probability is

$$\pi_i = \frac{r_{ni}^{-\alpha}}{\sum_j {r_{nj}}^{-\alpha}}. \tag{3.2}$$

3.3.1.1 Degree distribution

The nodes are labeled by their birth times, $s = 0, 1, 2 \cdots t$. $p(k, s, t)$ is the probability that the node s has degree k at time t. The master equation of $p(k, s, t)$ is given by

$$p(k, s, t+1) = \frac{m}{t+1} p(k-1, s, t) + (1 - \frac{m}{t+1}) p(k, s, t). \tag{3.3}$$

The initial conditions are $p(k, s = 0, 1...m_0, t = m_0) = \delta_{k,m_0}$ and $p(k, t, t) = \delta_{k,m}$. $p(k, s, t+1)$ contains two parts. The first one comes from the nodes having degree $k-1$ at time t and selected to connect with the new node at time $t+1$. The second one comes from the nodes having degree k at time t and not selected at time $t+1$.

The degree distribution of the entire network can be written as

$$p(k, t) = \frac{1}{t+1} \sum_{s=0}^{t} p(k, s, t). \tag{3.4}$$

Combining Eqs. (3.3) and (3.4), we get the following equation for the degree distribution

$$(t+2)p(k, t+1) - (t+1)p(k, t) = mp(k-1, t) - mp(k, t) + \delta_{k,m} \tag{3.5}$$

let $t \to \infty$, $p(k, t)$ will approach a stationary distribution $p(k)$ [128]. Eq. (3.5) becomes

$$(m+1)p(k) - mp(k-1) = \delta_{k,m} \tag{3.6}$$

which means

$$p(k) = \begin{cases} \frac{m}{m+1} p(k-1) & \text{if } k > m \\ \frac{1}{m+1} & \text{if } k = m \end{cases} \tag{3.7}$$

The final degree distribution turns out to be

$$p(k) = \frac{1}{m+1} (\frac{m}{m+1})^{k-m}, (k \geq m) \tag{3.8}$$

which decays exponentially with k ($p(k) = 0$ for $k < m$), in agreement with the results of Refs. [128, 136]. Thus the spatial driven network is an exponential network like most small-world networks [128, 136, 137].

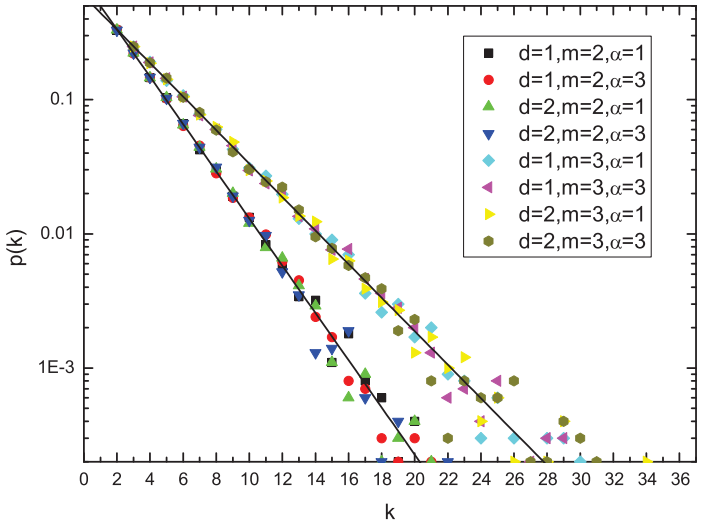

Fig. 3.4: The symbols represent the degree distribution $p(k)$ of a network of size $N = 10000$ grown according to our model for different m, α, $d = 1$ and $d = 2$. The solid lines are analytic results given by Eq. (3.8).

Fig. 3.4 shows the results of numerical simulation compared to the analytical results of Eq. (3.8) with a good agreement for different α and m. The degree distribution is only affected by the number of new edges m added at every time step.

3.3.1.2 Spatial distribution of link

When networks are embedded in space, the spatial distance r between the nodes is well-defined. The network evolution follows the purely spatial motivation as in Eq. (3.2).

When t is large enough, the nodes are homogeneously located on the one-dimension ring or two-dimension surface. The number of previous nodes at distance r from the new node is 2 when $d = 1$ and is proportional to $2\pi \sin \frac{r}{R}$ when $d = 2$. We define $\pi_t(r)$ as the probability of an added link between two specific nodes with distance r at time t, and $\triangle N(r,t)$ as the number of new links of length r that the network has at time t. $\triangle N(r,t)$ is given by the number of "neighbors" at distance r multiplied by the probability of link addition and by the number of new links, i.e., $\triangle N(r,t) = 2m\pi_t(r)$ for $d = 1$ and $\triangle N(r,t) \sim 2\pi m\pi_t(r) \sin \frac{r}{R}$ for $d = 2$.

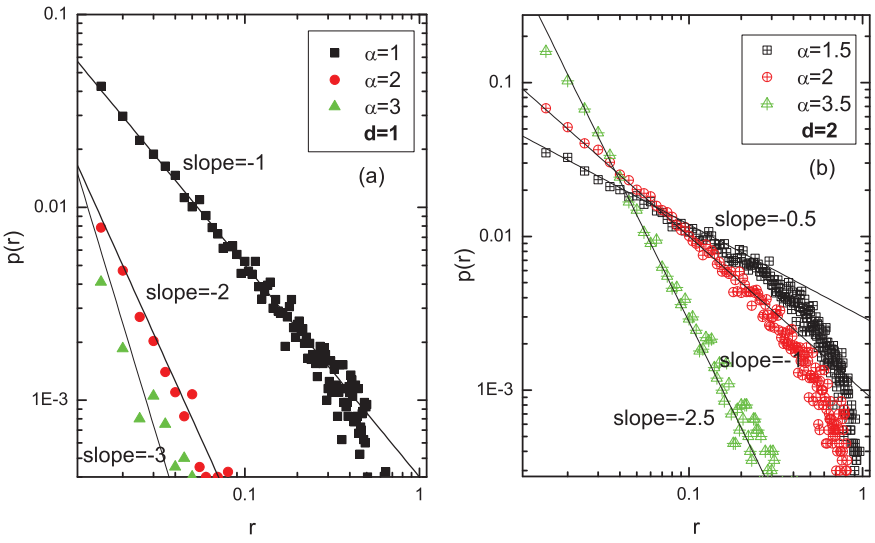

Fig. 3.5: (a) Probability $p(r)$ that a node has a connection at distance r for $\alpha = 1, 2, 3$ and d=1. (b) Probability $p(r)$ that a node has a connection at distance r for $\alpha = 1.5, 2, 3.5$ and $d = 2$. The straight lines are the results of analytic calculation with the slope $-(\alpha - d + 1)$.

In one-dimension, the number of links of length r at time t is

$$N(r,t) = \sum_{s=0}^{t} \triangle N(r,s) = 2m(\pi_0(r) + \pi_1(r) + \cdots + \pi_t(r)). \tag{3.9}$$

From Eq. (3.2), we get

$$\pi_s(r) \sim r^{-\alpha} \tag{3.10}$$

and

$$N(r,t) \sim r^{-\alpha}. \tag{3.11}$$

The spatial distribution of links is given by

$$p(r) = \frac{N(r,t)}{m(t+1)} \sim r^{-\alpha}. \tag{3.12}$$

For the case of two-dimension, we have

$$N(r,t) = \sum_{s=0}^{t} \triangle N(r,s) \sim 2\pi m(\pi_0(r) + \pi_1(r) + \cdots + \pi_t(r)) \sin\frac{r}{R}. \tag{3.13}$$

According to the Taylor series, the spatial distribution is given by

$$p(r) = \frac{N(r,t)}{m(t+1)} \sim r^{-\alpha+1}. \tag{3.14}$$

From the above results, we can derive $p(r) \sim r^{-(\alpha-d+1)}$. This result has also been found by Kosmidis and Manna using numerical simulation [129, 131]. Fig. 3.5 shows a good agreement between the results of numerical simulation and analytical calculation from Eqs. (3.12) and (3.14) for different values of α.

3.3.1.3 Clustering coefficient

The clustering coefficient of a single node i in network is defined as $c_i = \frac{2e_i}{k_i(k_i-1)}$ [8], where e_i is the total number of edges between all the k_i neighbors. The clustering coefficient C of the whole network is the average of c_i over all nodes.

Fig. 3.6 shows the behavior of the clustering coefficient C as a function of $1/N$ for different values of α. The variation of C follows $C \sim (1/N)^{\delta}$ where δ depends on α as shown in Fig. 3.7 for $d = 1$ and $d = 2$. The error bars are determined from the fitting with $C \sim (1/N)^{\delta}$. There are three regimes separated by $\alpha = d/2$ and $\alpha = 3d$ with different C behaviors. In the first regime $0 < \alpha \leqslant d/2$, the data follow $C \sim 1/N$, similar to the ER random graph [3]. In the second regime $d/2 < \alpha < 3d$, the exponent δ decreases from 1 to 0 (see Fig. 3.7). In the third regime $3d \leqslant \alpha$, C is independent from N ($\delta = 0$), similar to the OHO model presented by Ozik *et al.* [128, 136]. These regimes can be expressed as follows:

$$C \sim \begin{cases} 1/N & 0 < \alpha \leqslant d/2 \\ (1/N)^{\delta} & d/2 < \alpha < 3d. \\ constant & 3d \leqslant \alpha \end{cases} \tag{3.15}$$

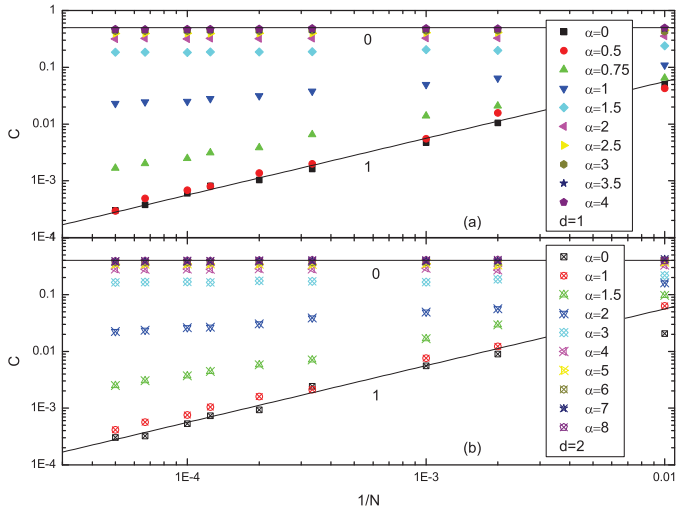

Fig. 3.6: (a) Clustering coefficient C as a function of $1/N$ for different α and $d = 1$. (b) Clustering coefficient C as a function of $1/N$ for different α and $d = 2$. Straight lines of slope 0 and 1 are best fits of the data.

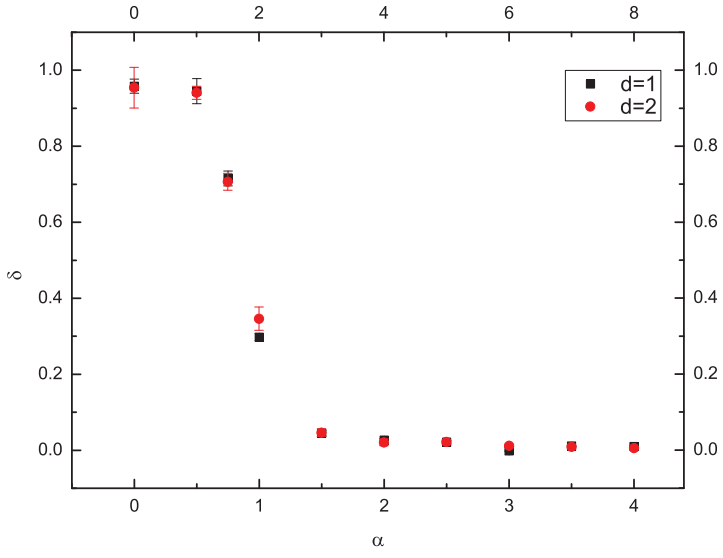

Fig. 3.7: Exponent δ as a function of α for d=1 and d=2. The left and the lower coordinates correspond to $d = 1$, the right and the upper coordinates correspond to $d = 2$. This figure shows that the α/d dependence of δ is free from dimension.

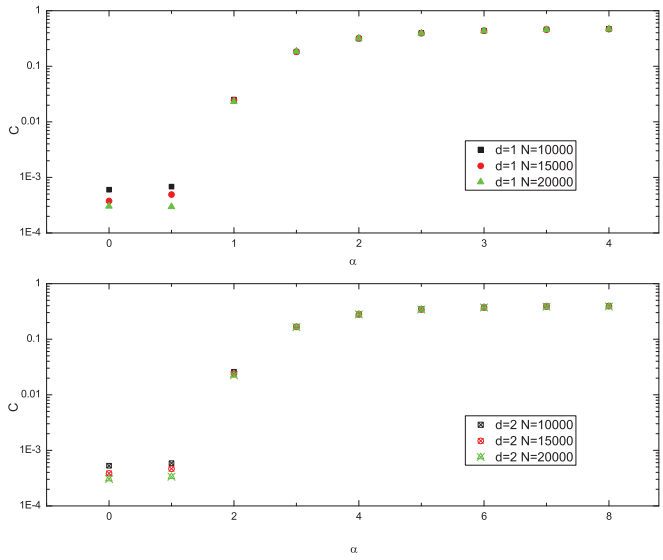

Fig. 3.8: Clustering coefficient C as a function of α with $N = 10000$, 15000 and 20000, for $d = 1$ and $d = 2$.

Now let us see how the clustering coefficient C depends on α when the network size is fixed (N=10 000, N=15 000, N=20 000). Fig. 3.8 shows that the critical point at $\alpha = d/2$ separates a phase of vanishing clustering ($0 < \alpha \leqslant d/2$) from a phase of increasing clustering ($d/2 < \alpha < 3d$), and that the point at $\alpha = 3d$ separates the phase of increasing clustering from a phase of constant clustering ($3d \leqslant \alpha$). The increasing clustering coefficient with increasing α is expected from the model because larger α favors smaller distance connection and higher clustering. This spatial effect on C is specifically notable in the second phase $d/2 < \alpha < 3d$.

3.3.1.4 Topological distance

We can see in Fig. 3.9 that the mean topological distance l of the network follows the function $l = \gamma logN$ for different α, which is a typical small world network behavior and different from what observed in Kosmidis's model embedded in regular lattices [129]. The α dependence of the slope γ is depicted in Fig. 3.10. The error bars are determined by the fitting with $l = \gamma logN$. We are interested in the behavior at the regime transition points $\alpha = d/2$ and $\alpha = 3d$. In the first regime $0 < \alpha \leqslant d/2$ and the third one $3d \leqslant \alpha$, γ is independent from α. In the second regime $d/2 < \alpha < 3d$, γ increases from the first one to

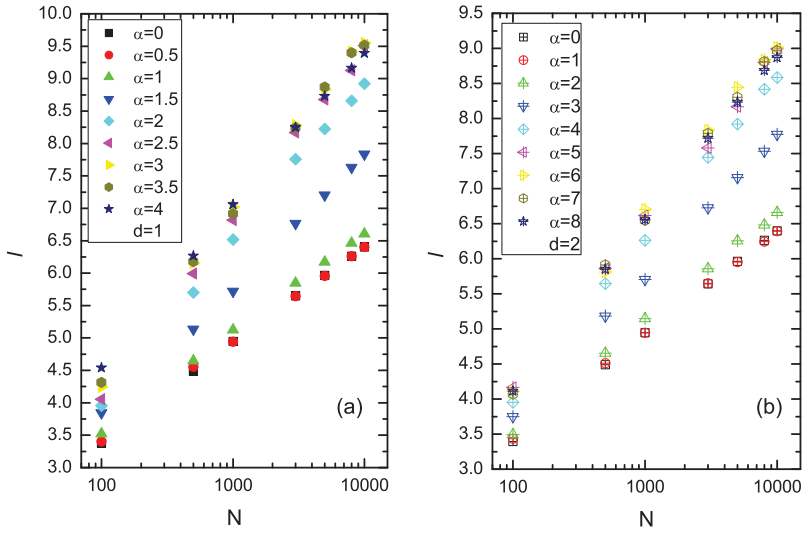

Fig. 3.9: (a) Topological distance l versus N for different α and $d = 1$. (b) Topological distance l versus N for different α and $d = 2$.

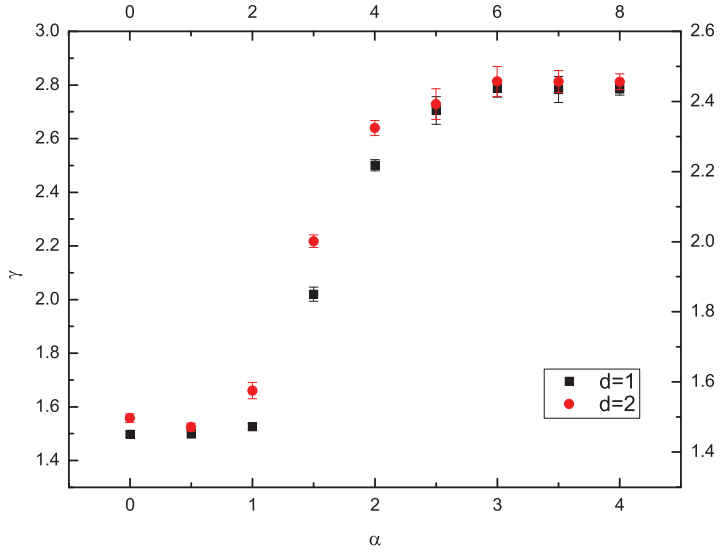

Fig. 3.10: α dependence of the slope γ. The left and the lower coordinates correspond to $d = 1$, the right and the upper coordinates correspond to $d = 2$.

the third one. It is worth mentioning that, in the ER random model, $l = logN/log\langle k\rangle$ or $\gamma = 1/log\langle k\rangle$. In the present case, $\langle k\rangle = 2m = 4$, leading to $\gamma \approx 1.66$, close to the value

of the first regime.

3.3.2 Network structure when $a \neq 0$

In many real space networks, the spatial distance between the nodes plays an important role in the formation of links, while the degree preferential attachment is also a natural property of linking. When $a \neq 0$ in our model, there is an interplay between the preference of larger degree and the preference of smaller distance in the evolution of the network. We note that in many models of network growth, the probability to get connected to a node at distance r is proportional to $r^{-\alpha}$. From the simulation results, when α increases, the degree distribution of the network gradually changes from power-law to stretched exponential [131, 135].

The degree increases each time when a new node n is added into the system and connects to a previous node i with probability π_i. Assume that k_i is a continuous variable, its increasing rate should be proportional to π_i and satisfy the following equation:

$$\frac{\partial k_i}{\partial t} = m\pi_i = ma\frac{k_i}{\sum_j k_j} + m(1-a)\frac{{r_{ni}}^{-\alpha}}{\sum_j {r_{nj}}^{-\alpha}}. \tag{3.16}$$

The sum in the first term on the right-hand side goes over all nodes except the new one, giving $\sum_j k_j = 2mt$. Since the new node is randomly located, and the existing nodes are uniformly distributed, when t is large, the change of the degree of node i must be independent of where node i is on the circle or the sphere, so that the second term reads $\frac{{r_{ni}}^{-\alpha}}{\sum_j {r_{nj}}^{-\alpha}} \approx \frac{1}{t}$. To prove this relationship we use the mean-field approximation in which the spatial distance preference probability $\pi_i^s = \frac{{r_{ni}}^{-\alpha}}{\sum_j {r_{nj}}^{-\alpha}}$ is represented by its average value. As we will show, the mean-field approximation turns out to be exact in the limit of large system size.

When t is large enough, the existing nodes are homogeneously located on the one-dimension ring or two-dimension surface. At time step t, we calculate π_i^s for T times, for each calculation the new node is placed randomly and labeled by $n_s = n_1, n_2, ...n_T$. The average value of π_i^s can be written as

$$\langle \pi_i^s \rangle = \frac{1}{T}\left(\frac{r_{n_1 i}^{-\alpha}}{\sum_{j=0}^{t-1} r_{n_1 j}^{-\alpha}} + \frac{r_{n_2 i}^{-\alpha}}{\sum_{j=0}^{t-1} r_{n_2 j}^{-\alpha}} + \cdots + \frac{r_{n_T i}^{-\alpha}}{\sum_{j=0}^{t-1} r_{n_T j}^{-\alpha}}\right). \tag{3.17}$$

Due to the uniform distribution of the existing nodes, we have $\sum_{j=0}^{t-1} r_{n_1 j}^{-\alpha} = \sum_{j=0}^{t-1} r_{n_2 j}^{-\alpha} = \cdots = \sum_{j=0}^{t-1} r_{n_T j}^{-\alpha}$. Hence

$$\langle \pi_i^s \rangle = \frac{1}{T} \frac{\sum_{l=1}^{T} r_{n_{li}}^{-\alpha}}{\sum_{j=0}^{t-1} r_{n_{sj}}^{-\alpha}} \tag{3.18}$$

In the denominator, $r_{n_{sj}}$ can be considered as a continuous variable between r_{min} and r_{max}, thus

$$\begin{aligned} \sum_{j=0}^{t-1} r_{n_{sj}}^{-\alpha} &= t\langle r_{n_s j}^{-\alpha} \rangle \\ &= t \int_{r_{min}}^{r_{max}} r_{n_{sj}}^{-\alpha} f(r_{n_{sj}}) dr_{n_{sj}} \end{aligned} \tag{3.19}$$

where $f(r_{n_{sj}})$ is the probability density of $r_{n_{sj}}$, satisfying $\int_{r_{min}}^{r_{max}} f(r_{n_{sj}}) dr_{n_{sj}} = 1$. For uniform distribution of nodes, $f(r_{n_{sj}})$ must be a constant, let it be c, and Eq. (3.19) can be written as

$$\sum_{j=0}^{t-1} r_{n_{sj}}^{-\alpha} = tc \int_{r_{min}}^{r_{max}} r_{n_{sj}}^{-\alpha} dr_{n_{sj}}. \tag{3.20}$$

This calculation also works for the sum over l from 1 to T, in the numerator. Since for large T, the new nodes are also uniformly distributed over $r_{n_{li}}$, so the probability density $f(r_{n_{li}})$ should be equal to c. Hence

$$\sum_{l=1}^{T} r_{n_{li}}^{-\alpha} = Tc \int_{r_{min}}^{r_{max}} r_{n_{li}}^{-\alpha} dr_{n_{li}}. \tag{3.21}$$

Take Eq. (3.20) and Eq. (3.21) to Eq. (3.18), Eq. (3.18) can be written as

$$\langle \pi_i^s \rangle = \frac{1}{T} \frac{Tc \int_{r_{min}}^{r_{max}} r_{n_{li}}^{-\alpha} dr_{n_{li}}}{tc \int_{r_{min}}^{r_{max}} r_{n_{sj}}^{-\alpha} dr_{n_{sj}}} = \frac{1}{t}. \tag{3.22}$$

From the above derivations and mean-field approximation, we can get

$$\pi_i^s \approx \langle \pi_i^s \rangle = \frac{1}{t}. \tag{3.23}$$

Substituting this result back into Eq. (3.16), the evolution of node $i's$ degree follows

$$\frac{\partial k_i}{\partial t} \approx a\frac{k_i}{2t} + (1-a)\frac{m}{t} = \frac{ak_i + 2(1-a)m}{2t}. \tag{3.24}$$

In the view of the initial condition $k_i(t_i) = m$ for the degree of a node added at time t_i, Eq. (3.24) has the solution

$$k_i = \frac{(2m - am)(\frac{t}{t_i})^{\frac{a}{2}} - 2(1-a)m}{a} \tag{3.25}$$

from which the probability $P(k_i(t) < k)$ that a node has degree $k_i(t)$ smaller than k is given by

$$P(k_i(t) < k) = P(t_i > \frac{t}{(\frac{ak+2(1-a)m}{2m-am})^{\frac{2}{a}}}). \tag{3.26}$$

Since the addition of nodes and links are carried out at equal time interval, the probability density at t_i is: $P_i(t_i) = 1/t_i$. Thus

$$P(k_i(t) < k) = 1 - (\frac{ak + 2(1-a)m}{2m - am})^{-\frac{2}{a}}. \tag{3.27}$$

Then the probability density $p(k)$ reads

$$p(k) = \frac{\partial P(k_i(t) < k)}{\partial k} = \frac{2}{2m - am}(\frac{a}{2m - am})^{-\frac{2+a}{a}}[k + \frac{2(1-a)m}{a}]^{-\frac{2+a}{a}}. \tag{3.28}$$

This is the "shifted power law" (SPL) function. When a changes from 0 to 1, the degree distribution gradually changes from an exponential law to a power law. In Fig. 3.11, the symbols represent the degree distribution $p(k)$ of our model for different values of a and α ($d = 1$ and 2). The solid lines are given by Eq. (3.28) with $m = 2$, $a = 0.1, 0.5$ and 0.9. The agreement between simulation results and analytical results means that the analysis from Eq. (3.16) to Eq. (3.28) is close to the numerical simulation with the model. SPL degree distribution is confirmed by empirical data as well [138, 139].

Fig. 3.12 shows the behavior of the clustering coefficient C for different a and α. In the first regime $0 < \alpha \leqslant d/2$ (see Fig. 3.12(a) for $d = 1$ and Fig. 3.12(d) for $d = 2$), $\ln C$ increases with a independently from α. Figs. 3.12(b) and 3.12(e) correspond to the the second regime $d/2 < \alpha < 3d$ in which $\ln C$ decreases linearly with a first and then

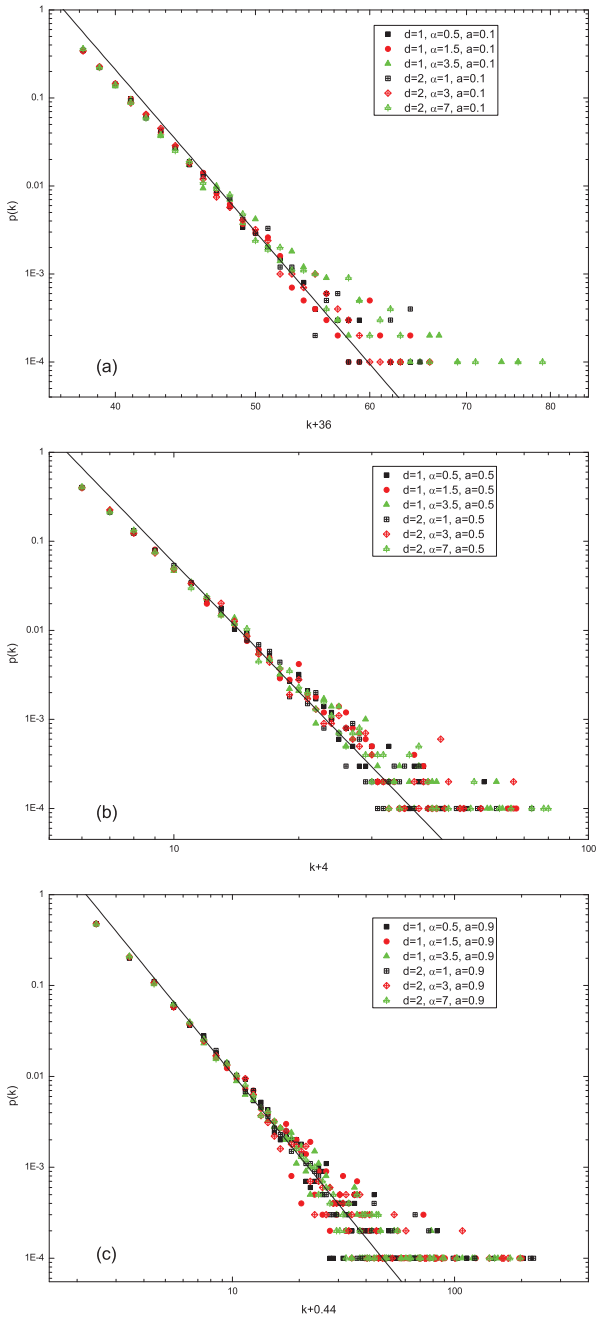

Fig. 3.11: (a) Degree distribution for $d = 1$, $a = 0.1$, $\alpha = 0.5, 1.5, 3.5$ and $d = 2$, $a = 0.1$, $\alpha = 1, 3, 7$. (b) Degree distribution for $d = 1$, $a = 0.5$, $\alpha = 0.5, 1.5, 3.5$ and $d = 2$, $a = 0.5$, $\alpha = 1, 3, 7$. (c) $d = 1$, $a = 0.9$, $\alpha = 0.5, 1.5, 3.5$ and $d = 2$, $a = 0.9$, $\alpha = 1, 3, 7$. The straight lines are analytic results given by Eq. (3.28) with the same a values as in the simulation, *i.e.* $a = 0.1, 0.5$ and 0.9, respectively.

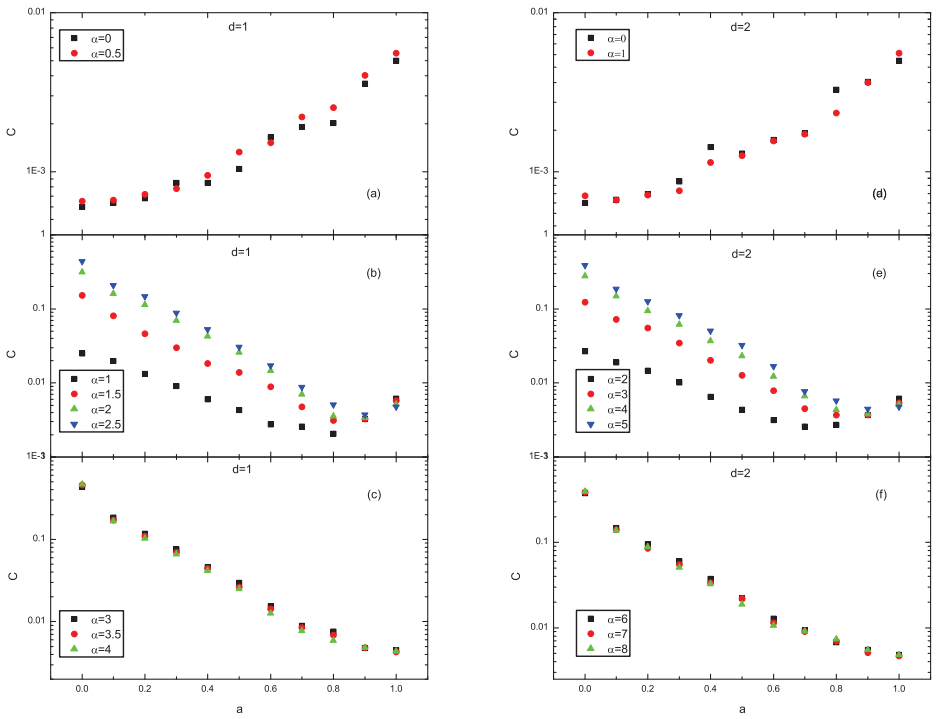

Fig. 3.12: (a) Clustering coefficient C as a function of a for $d = 1$ in the first regime for $\alpha = 0$ and 0.5. (b) C as a function of a for $d = 1$ in the second regime for $\alpha = 1$, 1.5, 2, 2.5. (c) C as a function of a for $d = 1$ in the third regime for $\alpha = 3$, 3.5, 4. (d) C as a function of a for $d = 2$ in the first regime for $\alpha = 0$, 1. (e) C as a function of a for $d = 2$ in the second regime for $\alpha = 2$, 3, 4, 5. (f) C as a function of a for $d = 2$ in the third regime for $\alpha = 6$, 7, 8.

increases slightly until $a = 1$. On the other hand, $\ln C$ increases with increasing α and its minimum seems to be dependent on α. In the third regime $3d \leqslant \alpha$ (see Fig. 3.12(c) for $d = 1$ and Fig. 3.12(f) for $d = 2$), $\ln C$ decreases with a independently from α.

3.4 Comparison with empirical data

Previous works have shown that the establishment of friendship and relationship is influenced by spatial constraints [119, 120, 134, 140]. In Ref. [120], Goldenberg and Levy collected data on the location of the receivers of more than 4400 email messages and found that the spatial distribution of the communication was a power law. This motivates us

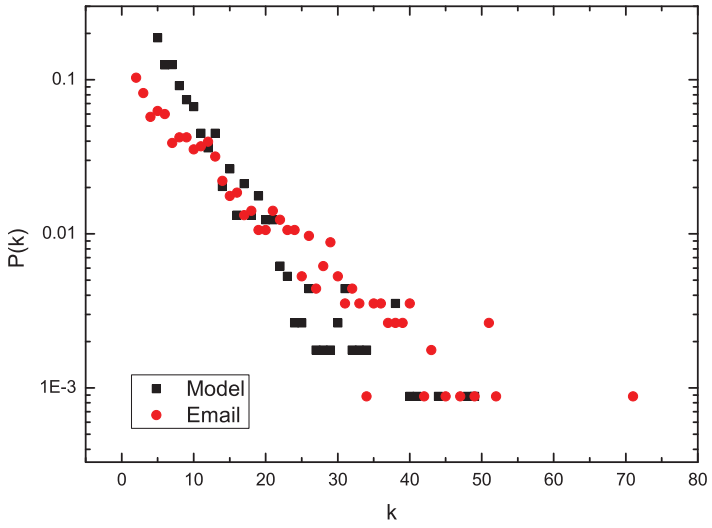

Fig. 3.13: Comparison of empirical data from email network and simulation results of the degree distribution. The simulation results come from our model with parameters $\alpha = 5$, $a = 0.02$, $N = 1133$, $m = m_0 = 5$.

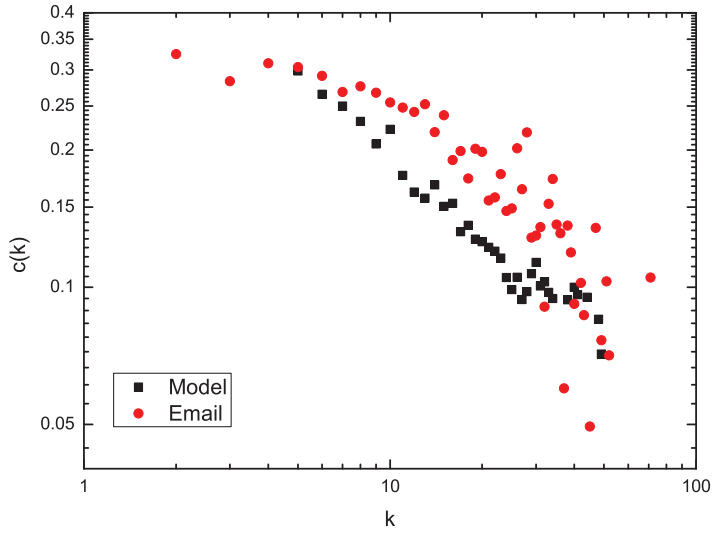

Fig. 3.14: Comparison of empirical data and simulation result of the k dependence of clustering coefficient. The simulation results come from our model with parameters $\alpha = 5$, $a = 0.02$, $N = 1133$, $m = m_0 = 5$.

to compare the topological structure of our model to the structure of the empirical data on email communication network [141], which consists of 1133 nodes and 5451 edges. In order to apply the two-dimensional model to the empirical data, we first set $N = 1133$ and $m = m_0 = 5$, then adjust $a = 0.02$ to match the degree distribution, finally set $\alpha = 5$ to match $\langle c \rangle$ and $\langle l \rangle$. The comparison of degree distribution is shown in Fig. 3.13. We employ the Kolmogorov-Smirnov test [142] in the software SPSS to compare the degree distribution of empirical data and that of our model. The result is $P - value = 0.826$ (when $P - value > 0.05$ means that the two comparative samples come from the same kind of distribution). Therefore we can conclude the email network [141] and our model follow the same kind of degree distribution. The topological parameters are listed in Table 3.1 with good agreement. The small value of a implies that the competition between degree and spatial distance preference for attachment *may* be present in the evolution of email systems and that the degree preference plays a much less important role than spatial distance. On the other hand, recent works [140], have shown that the degree dependence of clustering coefficient of several real networks follows a power law $C(k) \sim k^{-1}$. This tendency is confirmed by the email network data of [40, 141] and our simulation (see Fig. 3.14).

3.5 Conclusion

In this chapter, we have studied an evolutionary network in the configuration space with a model in which the probability of attachment is controlled by two competing factors: degree preference and spatial distance preference. These two factors are modulated by two parameters a and α.

When $a = 0$ and $\alpha \neq 0$, the model reduces to the spatial driven model, exhibiting phase transitions at some critical values of α. For the regime $0 < \alpha \leqslant d/2$, where d is the dimension of the embedding space, the network has short topological distance and vanishing clustering as in the ER random model. For $d/2 < \alpha < 3d$, the network has increasing clustering coefficient with increasing α. The topological distance is short as well. For $3d \leqslant \alpha$, the network has a constant clustering, similar to the OHO model. In all these regimes, the spatial distribution follows the power law $p(r) \sim r^{-(\alpha-d+1)}$, the degree distribution follows exponential law.

When $a \neq 0$, there will be more long-range links caused by degree preferential attach-

ment. The degree distribution follows shifted power law. When a changes from 0 to 1, the network property changes from the property of spatial driven network to the property of scale-free network.

The qualitatively consistent with empirical results reveals that the model has captured some basic mechanisms for the evolution of social communication networks. We hope that it will be helpful for further study and understanding of real networks whose evolution is influenced by the interplay of different even competing dynamics. The tunable parameter a can also become an object of investigation and of optimization when true physical processes such as epidemic diffusion, informational diffusion, opinion spreading, culture propagation, transport of matter and so forth, are considered in spatial network.

Table 3.1: Comparison of empirical data from email network and simulation results. N is the number of nodes, $\langle k \rangle$ is the average degree, $\langle c \rangle$ is the mean clustering coefficient, $\langle l \rangle$ is topological distance. The simulation results come from our model with parameters $\alpha = 5$, $a = 0.02$, $N = 1133$, $m = m_0 = 5$.

	N	$\langle k \rangle$	$\langle c \rangle$	$\langle l \rangle$
Email	1133	9.6222	0.2211	3.6060
Model	1133	10.0000	0.2214	3.5571

Chapter 4

Epidemic on spatial network

4.1 Introduction

The last two decades have seen several large-scale epidemics of international importance, including human, animal, and plant epidemics. Notable among these are SARS, foot-and-mouth disease, Dutch elm disease, citrus canker, sudden oak death, and rhizomania. Apart from wide spatial range, massive losses, and large costs for attempted containment, the epidemics have several factors in common. They all spread on complicated networks with a mixture of short-range and long-range links that are often difficult or impossible to identify, despite great effort invested in tracking contacts. There is also incomplete knowledge about the epidemiological status of individuals (humans, animals, herds or farms, fields or plants). An infected individual can initially go undetected or untreated, while spreading the disease to other individuals. The importance of space and mobility networks appears very clearly in the study of epidemic spread. Infectious diseases indeed spread because people interact and travel and the modeling of disease spread thus requires ideally the knowledge of the origin-destination matrix and of the social network.

Several spatial network models have been proposed in order to study the spatial characters and the spatial dependence of dynamical process. Guo *et al.* investigated local region immunization strategy on spatial distance preference model and found a critical immunization radius [143]. Sun *et al.* studied the susceptible-infected-susceptible (SIS) model on a growing network considering both topological and geographical structures and found that the epidemic threshold emerged in scale-free networks when the effect of spatial distance was taken into account [144]. Wang et al. compared two kinds of immunization

strategies, connection neighbor immunization and spatial neighbor immunization, on both scale-free network and small-world network [145], and found that spatial neighbor immunization always worked better than connection neighbor immunization on both of the networks. In real networks, the virus always spreads through human contact or air. Both of the factors are affected by spatial distance. On the one hand, scientists have revealed that on average the majority of our social relationships are in our neighborhood [29], on the other hand, the virus become weaker when it spreads by air, which has high probability to spread on short spatial distance than on long spatial distance. Thus, the spatial network is important for us to study the epidemic spreading.

In this chapter, we study the epidemic spreading process on a spatial driven network with Spatial SIS (S-SIS) model.

4.2 Epidemic models

Epidemic models are used to describe rapid outbreaks that occur in less than one year, while endemic models are used for studying diseases over longer periods, during which there is a renewal of susceptibles by births or recovery from temporary immunity. The two classic epidemic models, SIS model and SIR model, provide an intuitive basis for understanding more complex epidemiology modeling results. These divide the population into two or three classes: susceptible (S), meaning they don't have the disease of interest but can catch it if exposed to someone who does, infective (I) meaning they have the disease and can pass it on, and recovered (R), this class only exists in SIR model, meaning they have recovered from the disease and have permanent immunity, so that they can never get it again or pass it on. Some authors consider the R to stand for "removed", a general term that encompasses also the possibility that people may die of the disease and remove themselves from the infective pool in that fashion. Others consider the R to mean "refractory", which is the common term among those who study the closely related area of reaction diffusion processes

4.2.1 The SIR model

In traditional mathematical epidemiology [146, 147, 148], one then assumes that any susceptible individual has a uniform probability β per unit time of catching the disease from any infective one and that infective individuals recover and become immune at some

stochastically constant rate γ. The fractions s, i and r of individuals in the state S, I and R are then governed by the differential equations

$$\frac{\mathrm{d}s}{\mathrm{d}t} = -\beta is, \qquad \frac{\mathrm{d}i}{\mathrm{d}t} = \beta is - \gamma i, \qquad \frac{\mathrm{d}r}{\mathrm{d}t} = \gamma i. \tag{4.1}$$

One of the most important conclusions of this work is for the case of networks with power-law degree distributions, for which, as in the case of site percolation, there is no non-zero epidemic threshold so long as the exponent of the power-law is less than 3. Since most power-law networks satisfy this condition, we expect diseases always to propagate in these networks, regardless of transmission probability between individuals. This point was first made, in the context of models of computer virus epidemiology, by Pastor-Satorras and Vespignani [149, 150], although, as pointed out by Lloyd and May [151, 152], precursors of the same result can be seen in earlier work of May and Anderson [153]. May and Anderson studied traditional (fully mixed) differential equation models of epidemics, without network structure. They divided the population into activity classes with different values of the infection rate β. They showed that the variation of the number of infective individuals over time depends on the variance of this rate over the classes, and in particular that the disease always multiplies exponentially if the variance diverges precisely the situation in a network with a power-law degree distribution and exponent less than 3.

4.2.2 The SIS model

Not all diseases award immunity on their survivors. Diseases that, for instance, are not self-limiting but can be cured by medicine, can usually be caught again immediately by an unlucky patient. Tuberculosis and gonorrhea are two much-studied examples. Computer viruses also fall into this category; they can be "cured" by antivirus software, but without a permanent virus-checking program the computer has no way to fend off subsequent attacks by the same virus.

With diseases of this kind carriers which are cured move from the infective pool not to a recovered pool, but back into the susceptible one. A model with this type of dynamics is called an SIS model, for obvious reasons. In the simplest, fully mixed, single-population case, its dynamics are described by the differential equations

$$\frac{\mathrm{d}s}{\mathrm{d}t} = -\beta is + \gamma i, \qquad \frac{\mathrm{d}i}{\mathrm{d}t} = \beta is - \gamma i, \tag{4.2}$$

satisfying $s + i = 1$, where β and γ are, as before, the infection and recovery rates.

The SIS model is a model of endemic disease. Since carriers can be infected many times, it is possible, and does happen in some parameter regimes, that the disease will persist indefinitely, circulating around the population and never dying out. The equivalent of the SIR epidemic transition is the phase boundary between the parameter regimes in which the disease persists and those in which it does not. The SIS model cannot be solved exactly on a network as the SIR model can, but a detailed mean-field treatment has been given by Pastor-Satorras and Vespignani [149, 154] for SIS epidemics on the configuration model. Their approach is based on the differential equations, Eq. (4.2), but they allow the rate of infection β to vary between members of the population, rather than holding it constant. The quantity βi appearing in Eq. (4.2) represents the average rate at which susceptible individuals become infected by their neighbors.

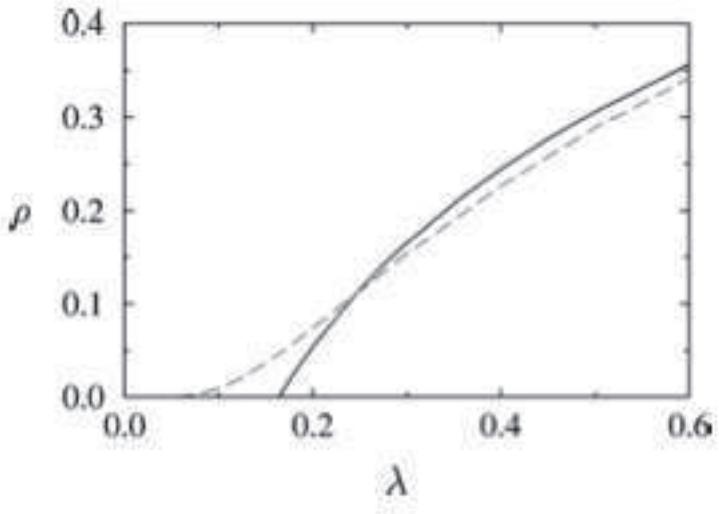

Fig. 4.1: Density of infected nodes ρ as a function of λ in the WS network (full line) and the BA network (dashed line).

When we focus on the WS model with $p = 1$, due to its exponential distribution, we make the replacement $\beta s \to \langle k \rangle \lambda (1 - i)$, where λ is the rate of infection via contact with a single infective individual. Eq. (4.2) approaches the analytical study of the SIS model by considering a single mean-field reaction equation for the density of infected nodes $i(t)$,

$$\partial_t i(t) = -i(t) + \lambda \langle k \rangle i(t)[1 - i(t)]. \tag{4.3}$$

The mean-field character of this equation stems from the fact that we have neglected the density correlations among the different nodes, independently of their respective connec-

tivities. After imposing the stationarity condition $\partial_t i(t) = 0$, we obtain the equation

$$i[-1 + \lambda\langle k\rangle(1 - i)] = 0 \tag{4.4}$$

for the steady state density i of infected nodes. This equation defines an epidemic threshold $\lambda_c = \langle k\rangle^{-1}$, as is shown in Fig. 4.1.

In BA networks, for a vertex of degree k, $\beta i \to k\lambda\Theta(\lambda)$, $\Theta(\lambda)$ is the probability that the neighbor at the other end of an edge will be infective. Note that Θ is a function of λ since presumably the probability of being infective will increase as the probability of passing on the disease increases. The remaining occurrences of the variables s and i are replaced by s_k and i_k, which are degree-dependent generalizations representing the fraction of nodes of degree k that are susceptible or infective, satisfying $s_k + i_k = 1$. We can write Eq. (4.2) as the single differential equation

$$\frac{\mathrm{d}i_k}{\mathrm{d}t} = k\lambda\Theta(\lambda)(1 - i_k) - i_k, \tag{4.5}$$

without loss of generality, set the recovery rate $\gamma = 1$. There is an approximation inherent in this formulation, since we have assumed that $\Theta(\lambda)$ is the same for all nodes, when in general it will be too dependent on degree. This is in the nature of a mean-field approximation, and can be expected to give a reasonable guide to the qualitative behavior of the system, although certain properties (particularly close to the phase transition) may be quantitatively mispredicted.

We can get the stationary solution of Eq. (4.5)

$$i_k = \frac{k\lambda\Theta(\lambda)}{1 + k\lambda\Theta(\lambda)}. \tag{4.6}$$

To calculate the value of $\Theta(\lambda)$, one averages the probability i_k of being infected over all nodes. Since $\Theta(\lambda)$ is defined as the probability that the node at the end of an edge is infective, i_k should be averaged over the distribution kp_k/z of the degrees of such nodes, where $z = \sum_k kp_k$ is, as usual, the mean degree. Thus

$$\Theta(\lambda) = \frac{1}{z}\sum_k kp_k i_k. \tag{4.7}$$

Taking Eq. (4.6) to Eqs. (4.5) and (4.7), we obtain

$$\frac{\lambda}{z}\sum_k \frac{k^2 p_k}{1 + k\lambda\Theta(\lambda)} = 1. \tag{4.8}$$

In order to perform an explicit calculation for the BA model, we use a continuous k approximation that allows the practical substitution of series with integrals. The full connectivity distribution is given by $p_k = 2m^2/k^{-3}$, where m is the minimum number of connection at each node. By noticing that the average connectivity is $\langle k \rangle = \int_m^\infty kP(k)dk = 2m$, Eq. (4.8) gives

$$\Theta(\lambda) = m\lambda\Theta(\lambda) \int_m^\infty \frac{1}{k^3} \frac{k^2}{1 + k\lambda\Theta(\lambda)}, \tag{4.9}$$

which has the solution

$$\Theta(\lambda) = \frac{e^{-1/m\lambda}}{\lambda m}(1 - e^{-1/m\lambda})^{-1}. \tag{4.10}$$

In order to find the behavior of the density of infected nodes we have to solve Eq. (4.6), which reads as

$$i = 2m^2\lambda\Theta(\lambda) \int_m^\infty \frac{1}{k^3} \frac{k}{1 + k\lambda\Theta(\lambda)}. \tag{4.11}$$

By substituting the obtained expression for $\Theta(\lambda)$ and solving the integral we find at the lowest order in λ

$$i \sim e^{-1/m\lambda}. \tag{4.12}$$

This result shows the surprising absence of any epidemic threshold or critical point in BA model (in Fig. 4.1). This can be intuitively understood by noticing that for usual lattices and mean-field models, the higher the node's connectivity, the smaller is the epidemic threshold.

4.3 Epidemic spreading on spatial network

In this section, we study epidemic spreading dynamic on spatial driven network, which we have mentioned in Section 3 with spatial SIS (S-SIS) model. Nodes in the network have two states: susceptible(S) and infected(I). For each infected node, it recovers and becomes susceptible with probability μ at each time step. In S-SIS model, initially, we choose an infected seed i in the network randomly. We assume that λ_0 denotes the infective rate of the virus itself. The cure rate μ is set to 1 without lack of generality, since it only affects the definition of the time scale of the infection propagation. In real world, the ability of transmission is closely related to the distance between the individuals, short distance relationship is easier to transmit virus than long distance relationship, which is an important factor for epidemic spreading. Inspire from previous works[143, 144], we define the transmission probability λ_{sj} from infective node j to susceptible node s is inversely proportional to their spatial distance as follows

$$\lambda_{sj} = \lambda_0 \frac{min_{k\in\Omega_j}(d_{kj})}{d_{sj}}, \tag{4.13}$$

where Ω_j is the set of all susceptible neighbors of node j. The distance between infected node and its closest neighbor is used to rescale λ_{sj} changing from 0 to λ_0. Therefore, for the closest susceptible neighbor, the infective rate is λ_0. For other susceptible neighbors, the infective rate is inversely proportional to their spatial distance. According to Eq. (4.13), we can obtain that the infective rate of susceptible node s depends on its spatial distance to all infective neighbors and can be written as $\lambda_s = 1 - \prod_{k\in\varphi_s}(1-\lambda_{sk})$, φ_s is the set of all infective neighbors of node s. Following the S-SIS model, infective node spreads the virus easily to its short distance neighbors. In real world, for the case of flu, if one is infected, the around people will have high probability to be infected.

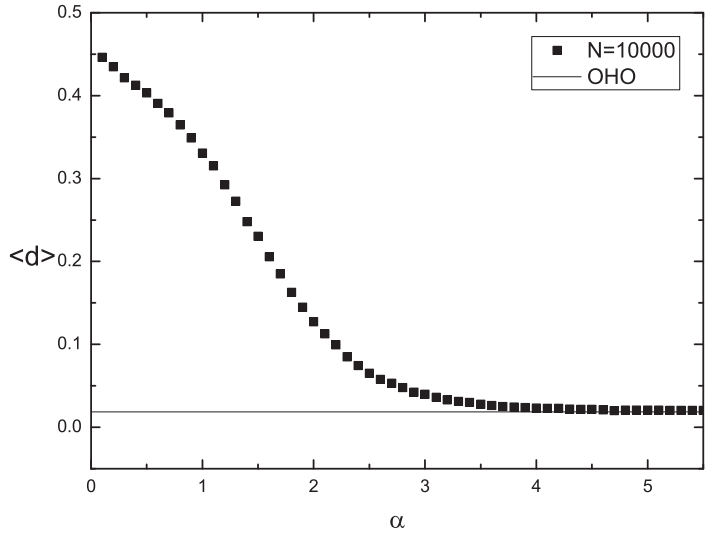

Fig. 4.2: The average spatial distance of all the links, $\langle d\rangle$, as a function of α for $N = 10000$. The symbols correspond to the spatial driven model, whereas the line stands for the OHO model [128].

4.4 Results and discussion

Fig. 4.3 shows the evolution of infected nodes ratio $\rho(t)$ as a function of time t for different spatial driven networks. Those networks grow with the same m (Fig. 4.3(a) with $m = 2$

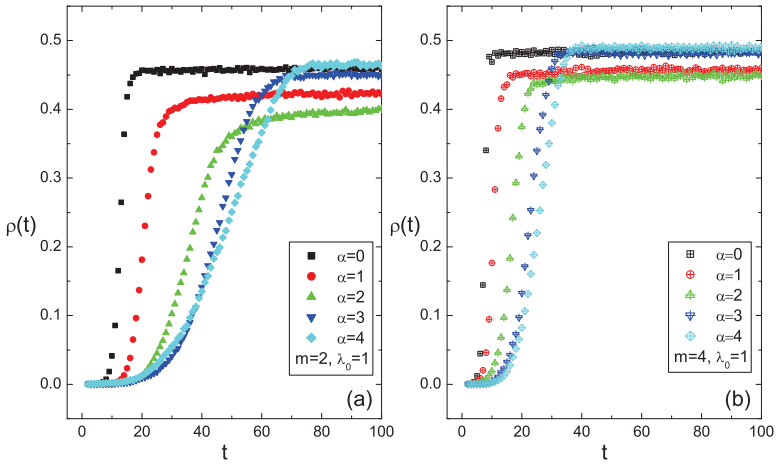

Fig. 4.3: The infected nodes ratio $\rho(t)$ as a function of time t in two special cases, (a) $m = 2$ or (b) $m = 4$, with $\alpha = 0, 1, 2, 3, 4$, respectively, the infective rate is $\lambda_0 = 1$. Simulations were performed on the networks with $N = 10000$.

and Fig. 4.3(b) with $m = 4$) and different α. We clearly notice a great influence of spatial structure on effective spreading time of virus. The effective spreading time is the number of steps the model needs to reach to the steady state. Short spatial distance network (larger α) needs longer effective spreading time to reach to the steady state than long spatial distance network (smaller α). For larger α, according to the spatial driven model, the nodes around the new one will be connected with high possibility, hence the local clustering and longer time. This property is independent on m and has been found by Xu *et al.* [155] as well. Compare Fig. 4.3(a) and Fig. 4.3(b), we reveal that the virus spreads quickly on the network with more neighbors. We can conclude that the virus needs shorter time to spread on the spatial driven network with smaller α and larger m.

During the spreading process, the node state changes between susceptible and infected. After T time steps, some nodes are infected more than one time. We define f as the infection frequency in a given period and discuss the infection frequency distribution $\rho(f)$ in Fig. 4.4. $\rho(f)$ gives the probability that a randomly chosen node is infected f times in a given period. According to the S-SIS model, since $\mu = 1$, for any infected nodes, it will be in susceptible state in next time step. Therefore, if the spreading period equals to T time steps, the maximum value of f will be $\frac{T}{2}$. In Fig. 4.4(a), the spreading process runs short time steps. The infection frequency distribution $\rho(f)$ decrease with the increasing of

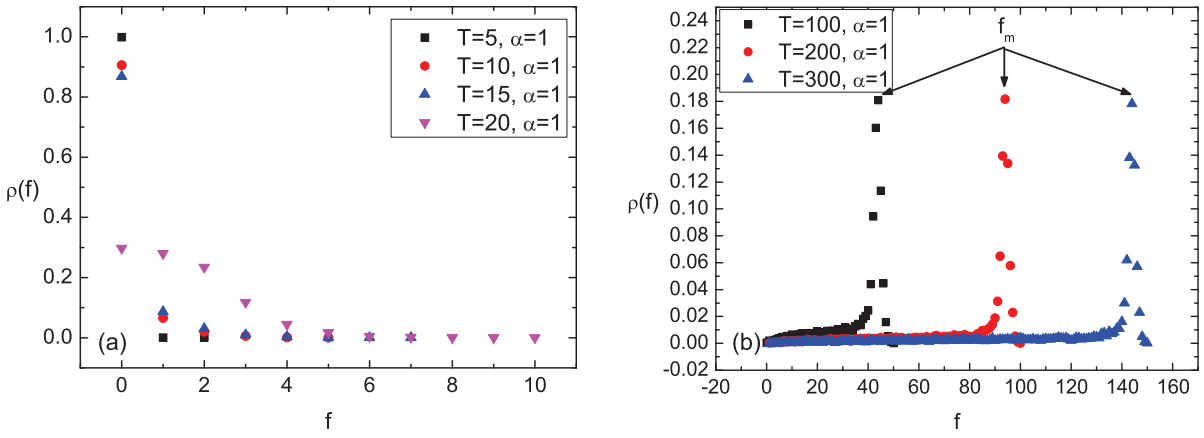

Fig. 4.4: The infection frequency distribution in different spreading period in the network with $\alpha = 1$ and $N = 10000$. (a) The spreading period is $T = 5, 10, 15, 20$ time steps. (b) The spreading period is $T = 100, 200, 300$ time steps.

infection frequency f. Most of nodes are never infected and only a few nodes are infected more than one times. In Fig. 4.4(b), the spreading process runs long enough time steps that the disease has spread out on the network. The infection frequency distribution always has a peak at the most probable frequency of infection f_m. The value of the peak $\rho(f_m)$ is independence with spreading period T. Fig. 4.5 shows that f_m grows lineally with the increasing of spreading period T and is close to $\frac{T}{2}$

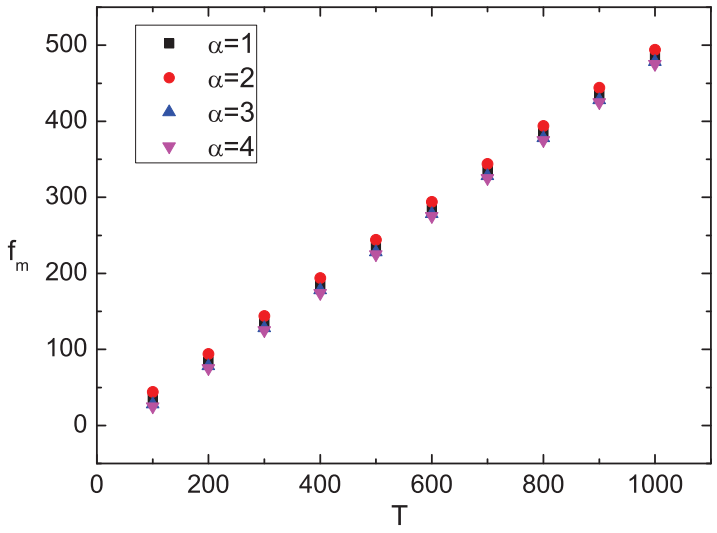

Fig. 4.5: The time dependence of the most probable frequency of infection f_m as a function of spreading period T.

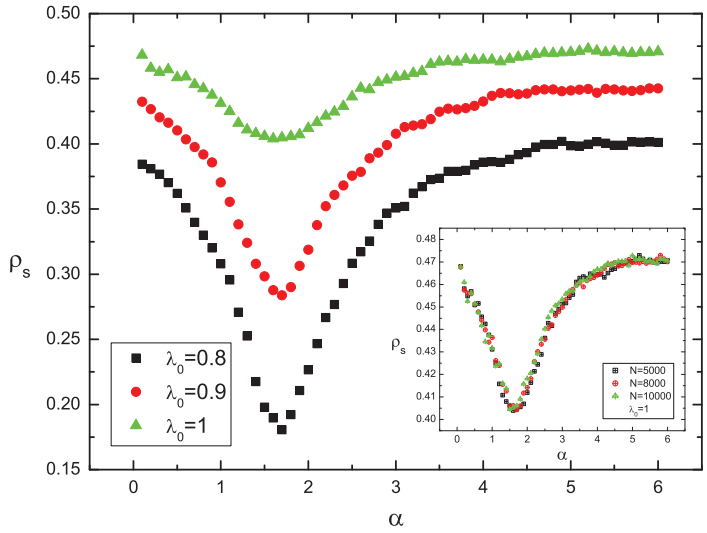

Fig. 4.6: The steady infected nodes ratio ρ_s as a function of α. The results are for three values of λ_0 and network size $N = 10000$. The inset corresponds to the results for different network sizes $N = 5000, 8000, 10000$ and $\lambda_0 = 1$.

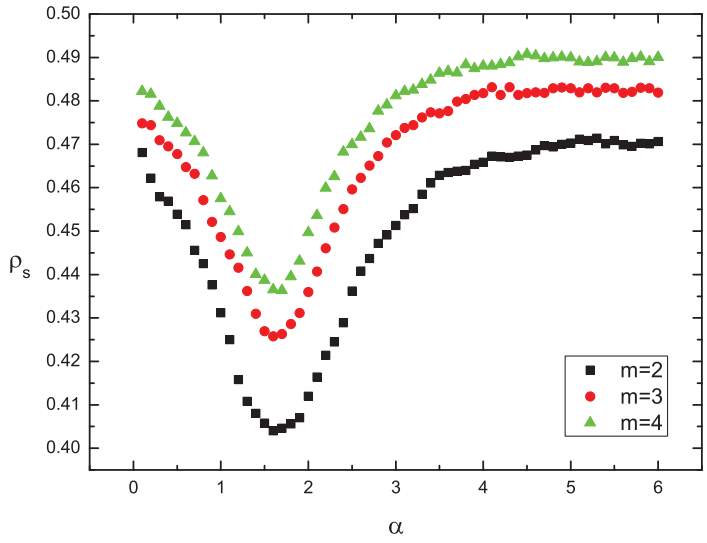

Fig. 4.7: The steady infected nodes ratio ρ_s as a function of α. Simulations were performed on different networks with $m = 2, 3$ and 4, respectively and $N = 10000$.

Furthermore, the steady infected nodes ratio ρ_s is various for different spatial structures. ρ_s is the average of $\rho(t)$ from $t = 9000$ to $t = 10000$. The results, presented in Fig. 4.6, show that the minimum of ρ_s for different λ_0 always exists when α is in the interval from 1.5 to 2. As shown in Fig. 4.6 inset, all curves play the same behavior and this property is independent on the size of the network. In Fig. 4.7, we apply the S-SIS model on three spatial driven networks with $m = 2, 3, 4$ and find that the minimum of ρ_s still exists when α is in the same interval. The relationship between spatial structure and ρ_s was disregarded by Guo *et al.* [143] and Sun *et al.* [144], who also studied the Spatial SIS model. From Fig. 4.2, we can find that when α is in the interval from 1.5 to 2 the average spatial distance $\langle d \rangle$ of the network is close to its middle value, long-range links and short-range links are homogeneous in the network. In this case, according to Eq. (4.13), $min_{k \in \Omega_j}(d_{kj})$ should get a smaller value thus the long-range links have smaller probability to transmit the virus, which causing the minimum of ρ_s. This interval separates the model to three parts. At the left side of the interval, for the smaller α, long-range links are general in the network, $min_{k \in \Omega_j}(d_{kj})$ has high probability to get a larger value, thus, long-range links have high probability to spread the virus. At the right side of the interval, for the larger α, short-range links are general in the network, $min_{k \in \Omega_j}(d_{kj})$ must be a smaller value and the virus can spread with high probability through these short-range links. Kleinberg reported the similar result on a transport network with long-range links following a power-law distribution, and found that the network has efficient navigation structure when the exponent is close to 2 [125]. This result comes from suitable spatial distance distribution as well [156].

In Fig. 4.8, we plot the distribution of ρ in four different spatial network planes at the same time step $t = 20$ and set that the infected seed is in the central of each plan. When $\alpha = 0.5$ and 1, the virus has spread from the central of the plane to the edge at $t = 20$. For smaller α, since the network has a lot of long-range links, the virus spreads to long distance easily. When $\alpha = 0.5$, the infected nodes are distributed homogenous in the plane. For $\alpha = 1$, the nodes close to the edge of the plane have smaller probability to be infected. While $\alpha = 2$ and 3, the virus only spreads surrounding the central at the same time step. In these cases, since network has less long-range links than the networks with $\alpha = 0.5$ and 1, the virus needs several steps to spread to the edge of the plane. Meanwhile, in the case of $\alpha = 2$ the infected nodes ratio ρ is smaller than the case of $\alpha = 3$ at the same position close to the central. Fig. 4.6 and Fig. 4.7 have uncovered the similar result that the S-SIS model is more difficult to spread on the spatial driven network with $\alpha = 2$

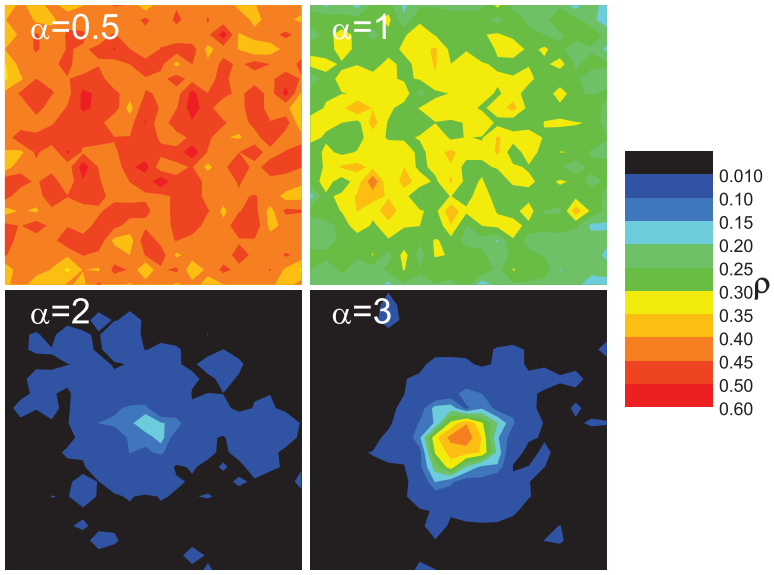

Fig. 4.8: The infected nodes ratio ρ in the plane for different values of α with $N = 10000$ and $t = 20$.

than on the network with $\alpha = 3$.

The epidemic threshold λ_c is a significant index to measure the dynamic processes of epidemic spreading on networks. Suppose λ_c is the epidemic threshold of a network which means that when $\lambda_0 < \lambda_c$ the infection dies out and when $\lambda_0 > \lambda_c$ the infection spreads and always exists in the network. In Fig. 4.9, we plot the steady infected nodes ratio ρ_s as a function of the original infective rate λ_0 for different α and try to find the simulation result of λ_c on different networks. Random growing process ($\alpha = 0$) and OHO model are two limiting cases for the spatial driven network. As shown in Fig. 4.9, the S-SIS model exhibits epidemic threshold on both random growing process ($\alpha = 0$) and OHO model. Meanwhile, in Fig. 4.9(a), it is observed that in the interval of $0 < \alpha < 1.5$ with the increasing of α, the value of epidemic threshold λ_c increases. On the contrary, in Fig. 4.9(b), in the interval of $2 < \alpha < 3.5$ with the increasing of α, λ_c decreases towards to the λ_c of OHO model. The maximum of λ_c exists in the network with parameter $1.5 < \alpha < 2$, where has the minimum of ρ_s. Therefore, S-SIS model exhibits epidemic threshold λ_c on all the spatial driven networks, whose value is determined by parameter α. But for the traditional SIS model, those networks have the same epidemic threshold because of the same degree distribution. This result means that spatial distance works in real world virus

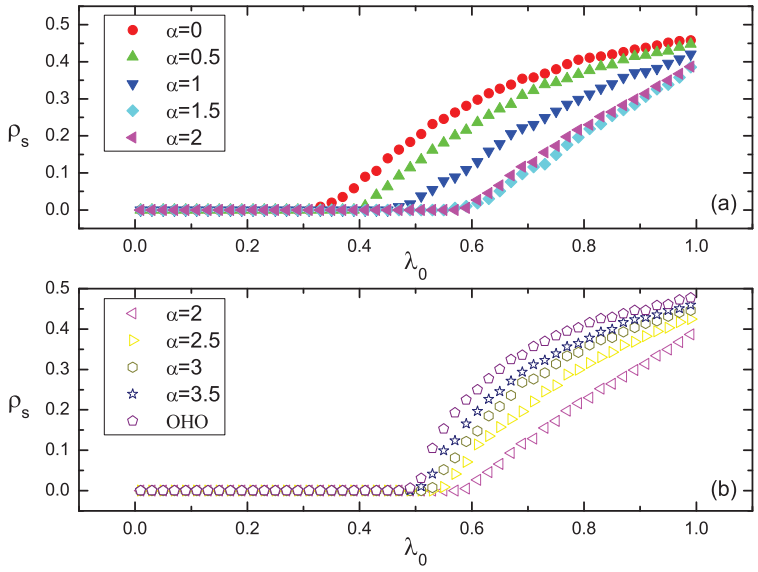

Fig. 4.9: The steady infected nodes ratio ρ_s as a function of λ_0 for different α and OHO model with $N = 10000$.

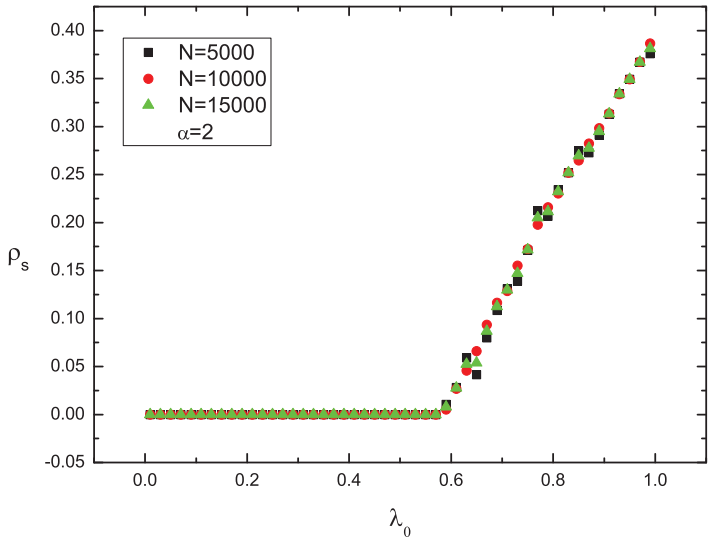

Fig. 4.10: The steady infected nodes ratio ρ_s as a function of λ_0 where $\alpha = 2$ and $N = 5000, 10000, 15000$.

spreading. In Fig. 4.10, we discuss λ_c on spatial driven network with $\alpha = 2$ for different network sizes and clearly observe that $\lambda_c \approx 0.58$, despite the variety of network size.

4.5 Conclusion

Many empirical works claim that topological structure and dynamical process are affected by the spatial structure. In theoretical modeling aspect, scientists begin to pay attention to these spatial networks. In this chapter, we study the epidemic spreading process on spatial driven network. Then we introduce the spatial constrain affect into traditional SIS epidemic spreading model. We present the S-SIS model, the spreading probability is not more a constant but rather inversely proportional to spatial distance for each links. Using the Monte Carlo method, we find that, when the network has smaller number of long distance connections, which needs longer time to reach to the steady state. The effective spreading time increases with increasing α. But the steady infected nodes ratio ρ_s has different behavior, which does not always decrease with increasing α and has the minimum value. The infection frequency distribution $\rho(f)$ is influenced by the spreading time T. When T is large enough, the distrbution $\rho(f)$ has a peak at the most probable frequency of infection f_m. In epidemic spreading work, epidemic threshold λ_c is an important problem continuously. In this work, we find that the epidemic threshold exists without exception, whose value is determined by parameter α. What's more, the maximum of λ_c and the minimum of ρ_s always exist in the network with α in the interval from 1.5 to 2. We hope that our work can help us to understand how virus spreads on different spatial networks. Moreover, taking the S-SIS model on the spatial network, proposed by Li *et al.* [157], in which, links follow a power-law distribution in the distance as well and are constrained by a maximum cost, will be an interesting topic in the future.

Chapter 5

Diffusion on spatial network

5.1 Introduction

Diffusive is a general dynamics in social networks such as spread of product innovations, rumors or fads propagation, " word-of-mouth" communications or some epidemic diseases among people. Introducing a new product, technology or idea in the market is an issue of major social economic relevance. Innovations do not necessarily spread at once, but often spread gradually through social and spatial networks. In fact, many products promote rather easily in a social system through a domino effect. At first stage a few innovators adopt the product, and this makes more likely that their neighbors do the same, then their neighbors' neighbors and so forth. One possible explanation for this phenomenon is that individuals' opinion heavily depends on the opinion of their interpersonal ties. Some of the opinion models are proposed under this principle. The dynamics of opinion among individuals is complex, because the individuals are. In mathematical model, opinions can be represented by numbers, the challenge is to find an adequate set of mathematical rules to describe the mechanisms responsible for the evolution and changes of them. The basic mechanism for opinion models is individual's opinion is effected by its neighbors. The "word-of-mouth" communication is a main dynamics in the opinion models. Each day, millions of conversations, e-mails, SMS, blog comments, instant messages or web pages containing various types of information are exchanged between people. Previous works have revealed that these actions are influenced by spatial constrains [119, 120]. Infectious diseases are spread because people interact and travel, which are constrained by spatial distance as well [158].

Various dynamics take place on spatial networks, and whose guiding idea is focus on the effects of space. V. Colizza *et al*. applied the spatio-temporal evolution model to the historical case of the Black Death, which occurred in the 14th century, only few traveling path were available and typical trips were limited to relatively short distances on the time scale of one day [159]. The main difficulty for their work was estimating the diffusion coefficient D. At that paper, they assumed that virus spreaded at a velocity of around 160 kms per year, and obtained $D \approx 10^4$ kms^2/year. But D was still an uncertainty parameter. A striking example can be seen in the SARS outbreak in 2003. From the evolution of the disease, V. Colizza *et al*. thought that pure spatial diffusion was not a good model anymore and that the global aspect of transportation network needed to be included in the modeling. They used metapopulation models to discuss the disease spreading in local but also with long-range jumps as well and found that reducing travel was not an efficient strategies for scale-free networks [159]. Hu *et al*. studied a malware propagation among WIFI routers and Wang *et al*. discussed the virus spread using Bluetooth and MMS in the spatially constrained networks [160, 161].

In the previous works, value of diffusion coefficient is a unresolved problem. But, for the traditional Fick's law, the diffusion coefficient depends on the temperature, viscosity of the fluid and the size of the particles [162]. Since the spatial driven network has well defined spatial distance, this model offers an occasion to use the usual diffusion equations to describe the diffusive dynamics. The purpose of our work is to calculate the diffusion coefficient in the spatial driven network and discuss the relationship between its value and the network spatial structure. Since our diffusion is not normal in general due to the long-range action, the diffusion coefficient must take into account the transfers of diffusive substances over all distances. And the coefficient is calculated from the Fick's first law for normal diffusion on all distances.

In the following, we will give an introduction of basic diffusion models, such as Fick's laws, anomalous diffusion and reaction-diffusion model. Then, based on the diffusion dynamics in spatial driven network, we use the Fick's first law and anomalous diffusion equation to get the diffusion coefficient. Some results and conclusions are presented finally.

5.2 Diffusion models

Over the twentieth century, a number of natural phenomena were modeled by diffusion and pattern formation processes. The former type of dynamics includes the established topics of atom and molecule diffusion [163] as well as heat diffusion through different materials [164]. In addition, econometricians have developed diffusion models to forecast the acceptance of new products and to understand their life-cycle [165]. Migration of animals, spreading of organisms and chemical substances are often investigated in terms of biological diffusion models [168]. More recently, complex biological and chemical patterns have been reproduced by systems of equations with diffusive and reactive terms [169]. These models range from simple diffusion equations (e.g. heat diffusion in a rod) to more sophisticated advectiondiffusion (e.g. chemical oceanography) and reactiondiffusion equations (e.g. chemical and biological patterns). Such two models are considered in the present paper in order to represent a reasonable range of natural and artificial phenomena: diffusion and the GrayScott reactiondiffusion models.

5.2.1 Fick's laws

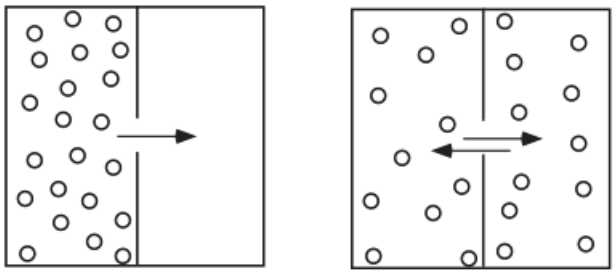

Fig. 5.1: Mass transport, diffusion as a consequence of existing spacial differences in concentration.

Diffusion usually occurs if there is a spatial difference in concentration of particles or heat, and it usually acts such as to reduce the spatial inhomogeneities in concentration. At any temperature different from absolute zero all atoms, irrespective of their state of aggregation (gaseous, liquid or solid), are constantly in motion. Since the movement of particles is associated with collisions, the path of a single particle is a zigzag one. However, an aggregation of "diffusing" particles has an observable drift from places of higher to places of lower concentration (Fig. 5.1). For this reason diffusion is known as a transport phenomenon.

5.2.1.1 Fick's first law

We first examine a simple one-dimensional model of diffusion. We consider diffusion of a trace amount of an impurity (or tracer) in a single-phase alloy. If the planer density of impurities at a position x is given by $\sigma(x)$ (measured in atoms/cm^2), and if the spacing between adjacent planes is $\triangle x$, then the volume concentration of impurities $c(x)$ is given by:

$$c(x) = \frac{\sigma(x)}{\triangle x}. \tag{5.1}$$

Further, we assume that we have one-dimensional random nearest neighbor jumps, and the diffusing atoms are chemically identical to, but distinguishable from the host atoms. This is the case for radioactive isotope tracers, for example. We define J_+ to be the flux of atoms to the right from the plane at x to the one at $x + \triangle x$. This is given by:

$$J_+ = \frac{\Phi\sigma(x)}{2}, \tag{5.2}$$

where Φ is the mean jump frequency, and the factor of $1/2$ accounts for the jumps being able to go in either the plus or minus x direction. We can also define J_- as the flux of atoms to the left from the plane at $x + \triangle x$ to the one at x, and we find:

$$J_- = \frac{\Phi\sigma(x + \triangle x)}{2}. \tag{5.3}$$

If we assume that Φ is not a function of concentration, then the net flux J is given by:

$$\begin{aligned} J &= J_+ + J_- \\ &= -\frac{1}{2}\Phi[\sigma(x + \triangle x) - \sigma(x)] \\ &= -\frac{1}{2}\Phi\triangle x[c(x + \triangle x) - c(x)] \\ &= -\frac{1}{2}\Phi(\triangle x)^2\frac{c(x + \triangle x) - c(x)}{\triangle x} \\ &\approx -\frac{1}{2}\Phi(\triangle x)^2\frac{\partial c}{\partial x} \\ &= -D\frac{\partial c}{\partial x}, \end{aligned} \tag{5.4}$$

where in the last step we have assigned, the quantity D is known as the diffusivity,

$$D = \frac{1}{2}\Phi(\triangle x)^2. \tag{5.5}$$

Eq. (5.4) which relates the concentration gradient to the flux is known as Fick's first law. As we go into this course, we will find that Fick's first law does not always hold, but it is

in fact a special case of the more general statement that a flux will be driven by a gradient in chemical potential. In many cases, the gradient in chemical potential is proportional to the gradient in concentration.

Actually, the flux is a vector quantity in that you might want to know the direction as well as magnitude of the atomic flow. Hence, Fick's first law can be written as a vector equation:

$$J = -D\nabla c, \tag{5.6}$$

where ∇ is the gradient operator.

5.2.1.2 Fick's second law

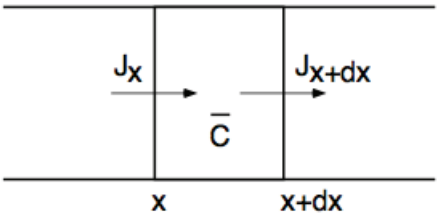

Fig. 5.2: Two cross-sections are separated by dx. The flux through the $1st$ section will not be the same as the flux through the second section.

Consider a volume element (between x and $x + dx$ of unit cross sectional area) of a membrane separating two finite volumes involved in a diffusion system (Fig. 5.2). The rate of change in concentration due to fluxes is:

$$\frac{\partial c}{\partial t} = \frac{J_x - J_{x+dx}}{dx}, \tag{5.7}$$

∂c is the average concentration in the volume element. Using a Taylor series we can expand J_{x+dx} about x and obtain:

$$J_{x+dx} = J_x + \frac{\partial J_x}{\partial x}dx + \frac{\partial^2 J_x}{\partial x^2}\frac{dx^2}{2} + \cdots . \tag{5.8}$$

Accordingly, as $dx \to 0$:

$$\frac{\partial}{\partial x}(D\frac{\partial c}{dx}) = \frac{\partial c}{\partial t} \tag{5.9}$$

and if D does not vary with x (which is normally the case) we have the formulation of Fick's second law:

$$\frac{\partial c}{\partial t} = D\frac{\partial^2 c}{\partial x^2}. \tag{5.10}$$

In physical terms this relationship states that the rate of compositional change is proportional to the "rate of change" of the concentration gradient rather than to the concentration gradient itself.

In infinite space, and if all particles start initially from $x = 0$, the solution of Eq. (5.10) is

$$c(x,t) = \frac{N_0}{sqrt4\pi Dt} e^{-x^2/4Dt}, \tag{5.11}$$

where N_0 is the total number of particles inside the volume under consideration. The solution obviously is identical to a Gaussian distribution with mean zero and variance $2Dt$. The variance is defined as

$$\langle x^2(t) \rangle = \int x^2 c(x,t) dx = 2Dt, \tag{5.12}$$

which is just identical to the mean square displacement, so that the results obtained earlier, using the simple version of the random walk or the stochastic differential equation of Langevin are again confirmed. Diffusion obeying Eq. (5.12) is called normal diffusion and is characteristic for the diffusion processes in systems that are equilibrium or very close to equilibrium.

5.2.2 Anomalous diffusion

Normal diffusion has as basic characteristic the linear scaling of the mean square displacement of the particles with time, $\langle r^2 \rangle \sim Dt$. Many different experiments though, including the one shown in the previous section, reveal deviations from normal diffusion, in that diffusion is either faster or slower, and which is termed anomalous diffusion [166, 167]. A useful characterization of the diffusion process is again through the scaling of the mean square displacement with time, where though now we are looking for a more general scaling of the form

$$\langle r^2(t) \rangle \sim t^{\delta}. \tag{5.13}$$

Diffusion is then classified through the scaling index δ. The case $\delta = 1$ is normal diffusion, all other cases are termed anomalous. The case $\delta > 1$ form the family of super-diffusive processes, including the particular case $\delta = 2$, which is called ballistic diffusion, and the cases $\delta < 1$ are sub-diffusion processes. If the trajectories of a sufficient number of particles inside a system are known, then plotting $\log\langle r^2 \rangle$ vs $\log t$ is an experimental way to determine the type of diffusion occurring in a given system.

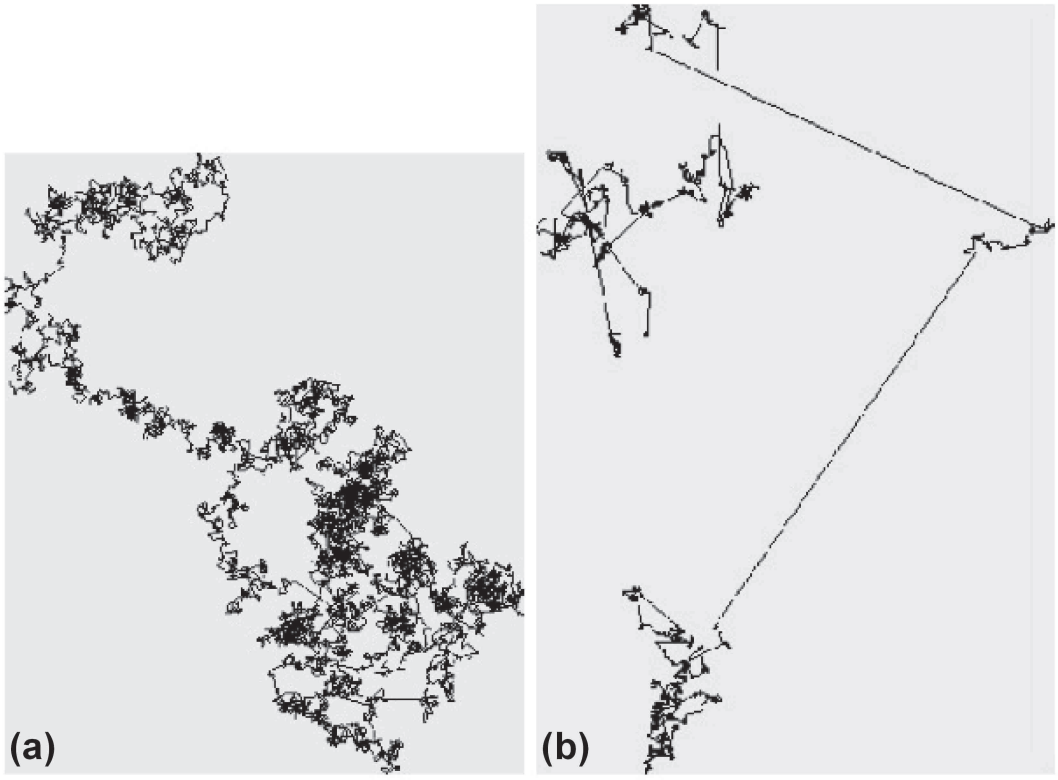

Fig. 5.3: (a) Random walk in dynamical systems close to normal diffusion (trajectory on the left), (b) random walk in dynamical systems close to anomalous diffusion (trajectory on the right).

As an illustration, let us consider a particle that is moving with constant velocity v and undergoes no collisions and experiences no friction forces. It then obviously holds that $r = vt$, so that $\langle r^2(t)\rangle \sim t^2$. Free particles are thus super-diffusive in the terminology used here, which is also the origin of the name ballistic for the case $\delta = 2$. Accelerated particles would even diffuse faster. The difference between normal and a anomalous diffusion is also illustrated in Fig. 5.4, where in the case of anomalous diffusion long "flights" are followed by efficient "trapping" of particles in localized spatial regions, in contrast to the more homogeneous picture of normal diffusion.

5.2.3 Reaction-diffusion system

Reaction-Diffusion systems are mathematical models which explain how the concentration of one or more substances distributed in space changes under the influence of two processes: local chemical reactions in which the substances are transformed into each other, and

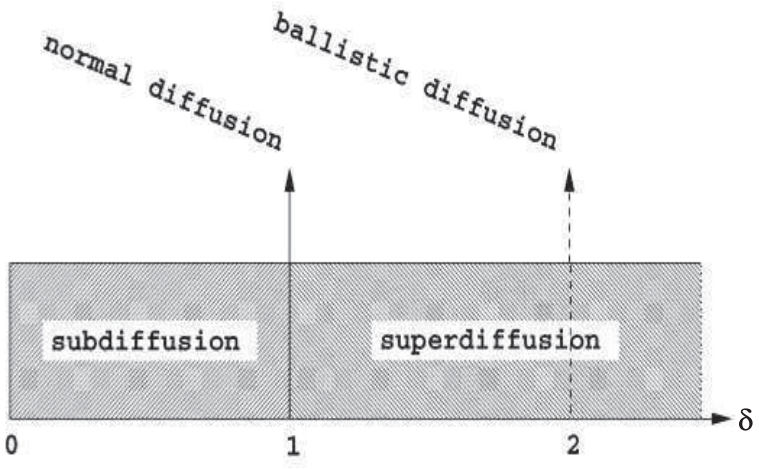

Fig. 5.4: Different domains of anomalous diffusion, defined through the mean squared displacement, Eq. (5.13), parametrised by the anomalous diffusion exponent δ: (a) subdiffusion for $0 < \delta < 1$, (b) superdiffusion for $\delta > 1$. On the threshold between sub- and superdiffusion is the normal Brownian diffusion located. Another special case is ballistic motion ($\gamma = 2$).

diffusion which causes the substances to spread out over a surface in space.

This description implies that reactiondiffusion systems are naturally applied in chemistry. However, the system can also describe dynamical processes of non-chemical nature. Examples are found in biology, geology and physics and ecology. Mathematically, reactiondiffusion systems take the form of semi-linear parabolic partial differential equations. They can be represented in the general form

$$\nabla_t c = D\triangle^2 c + R(c), \tag{5.14}$$

where each component of the vector $c(x,t)$ represents the concentration of one substance, D is a diagonal matrix of diffusion coefficients, and R accounts for all local reactions. The solutions of reactiondiffusion equations display a wide range of behaviors, including the formation of traveling waves and wave-like phenomena as well as other self-organized patterns like stripes, hexagons or more intricate structure like dissipative solitons.

5.3 Diffusion on spatial network

5.3.1 Spatial diffusion mechanism

The spatial network is a suitable frame to discuss the human spreading dynamics. The spatial network which we discusses in this chapter satisfies the spatial driven network model in Section 3. Fig. 5.5 shows the modulated example for spatial driven network model. The case with $\alpha = 1$, long-range connections are common in the network. With the increasing of α, the number of long-range connections become smaller, whereas for $\alpha = 4$ a newly introduced node is linked only to its nearer predecessors.

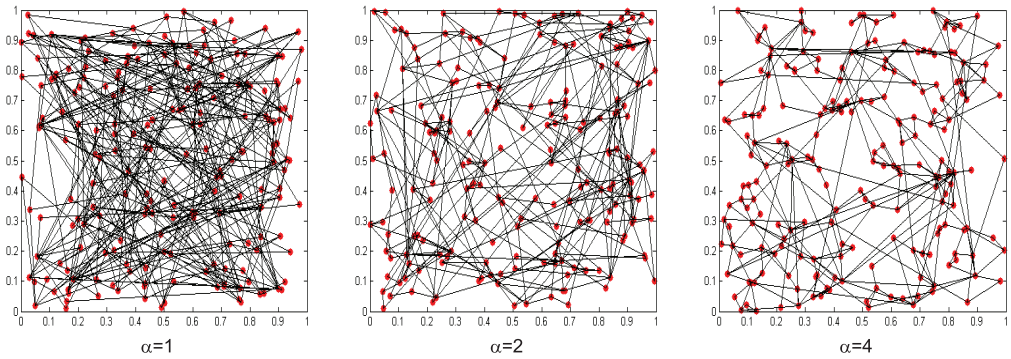

Fig. 5.5: Modulated spatial driven networks within a unit square for different values of the modulation parameter α for 200 nodes.

Based on some real social communication networks, we find that the communication probability is inversely proportional to the spatial distance [118, 120]. Arouse from the previous works [143, 144], we define the spatial diffusion mechanism: the diffusion probability λ_{sj} from the active agent s to its inactive neighbor j is inversely proportional to their spatial distance as follows

$$\lambda_{sj} = (\frac{l_{sj}}{min_{k\in\Omega_s} l_{sk}})^{-\beta}, \tag{5.15}$$

where Ω_s is the set of all inactive neighbors of node s. The distance between active agent and its closest neighbor is used to rescale λ_{sj} changing from 0 to 1. Therefore, for the closest inactive neighbor, the diffusion probability is 1; for other inactive neighbors, the diffusion probability is inversely proportional to the spatial distance from s to them with power β. We assume β is the parameter of diffusion product itself. For smaller β, the product is easy to diffusion; for larger β it is difficult.

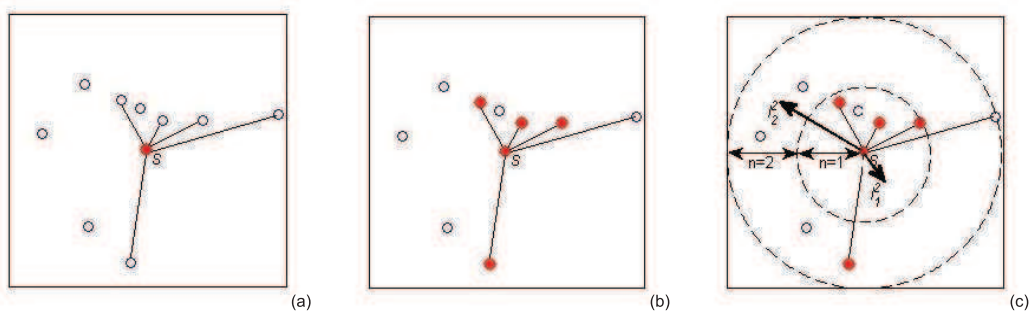

Fig. 5.6: (a) Example of a network and its original state. (b) One step diffusion on the network (a). (c) Divide the network into some rings to calculate the diffusion coefficient in simulation method.

5.3.2 Diffusion coefficient

Fick's first law relates the diffusive flux to the concentration under the assumption of steady state [162]. It postulates that the flux goes from regions of high concentration to regions of low concentration, with a magnitude that is proportional to the concentration gradient (spatial derivative), which can be written as:

$$J = -D\nabla c, \tag{5.16}$$

where D is the diffusion coefficient, J is the diffusion flux and c is the concentration.

In spreading network, the product always spreads from active agent to inactive one. When N agents are randomly distributed on 1×1 two-dimensional space, the density of nodes per unit area is N. For an active agent, its state density is $c = N$, otherwise $c = 0$. The product spreads from high state density agent to low state density agent. So we try to use Fick's first law to solve this problem. In order to get the diffusion coefficient, we make two hypothesis: the diffusion coefficient of the network is represented by its average values on different agents; it is unrelated with the number of active agents, thus, in the following, we only set one agent to active state and calculate the diffusion coefficient in one time step diffusion.

At first, we select an agent s randomly as an active seed, set s to the central of the planar and rescale all the nodes. As Fig. 5.6(a) shows, the red node s in the central represents the active agent, and the blue circles represent inactive agents. Using spatial

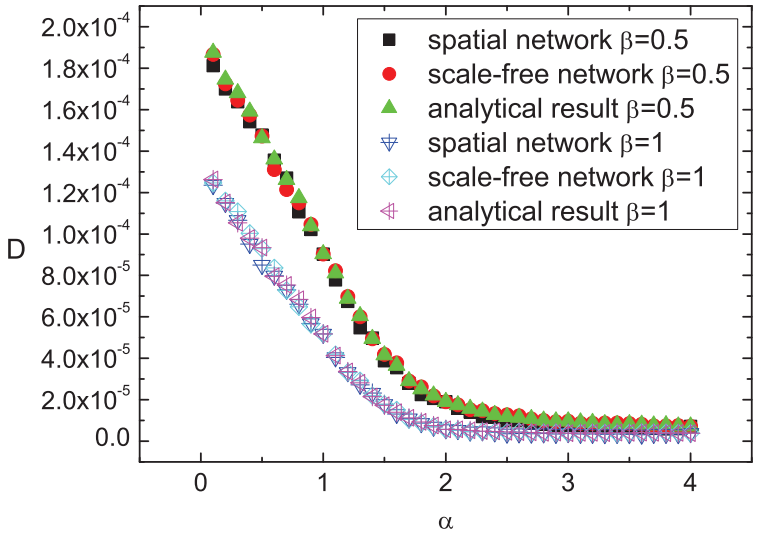

Fig. 5.7: Relation between the diffusion coefficient D and parameter α for $\beta = 0.5, 1$. D is calculated in three different ways.

diffusion mechanism of Eq. (5.15) and the Fick's first law of Eq. (5.16), the diffusion coefficient is given by

$$D = \sum_{j \in \Omega_s} \lambda_{sj} \frac{l_{sj}}{N} = \sum_{j \in \Omega_s} \frac{l_{sj}^{-\beta+1}}{(min_{k \in \Omega_s} l_{ks})^{-\beta} N}. \tag{5.17}$$

The diffusion flux in this model is the number of successful diffusion from seed s. Fig. 5.6(b) is the result of one step spatial diffusion on Fig. 5.6(a). Then we set agent s as the center and divide the planer to M rings, labeled as $n = 1, 2, \ldots, M$, to calculate D in simulation method (in Fig. 5.6(c)). The distance between the adjacent two rings is $\frac{1}{2M}$. The distance from the seed to the nth ring is $l_n^M = \frac{n-\frac{1}{2}}{2M}$, which is the average distance from the seed to the two adjacent circles. We assume that the distance between the seed and each node in the nth ring has the same value and equals to l_n^M. Since all the agents are randomly located on the planar, the number of agents in the nth ring is $N^M(n) = \frac{\pi(n^2-(n-1)^2)N}{4M^2}$. The approximation diffusion coefficient of the spatial network is the sum of sub-coefficient for all the parts

$$D^M = \sum_{n=1}^{M} d^M(n) + d_{out}^M = d^M(1) + d^M(2) + \cdots + d^M(M) + d_{out}^M, \tag{5.18}$$

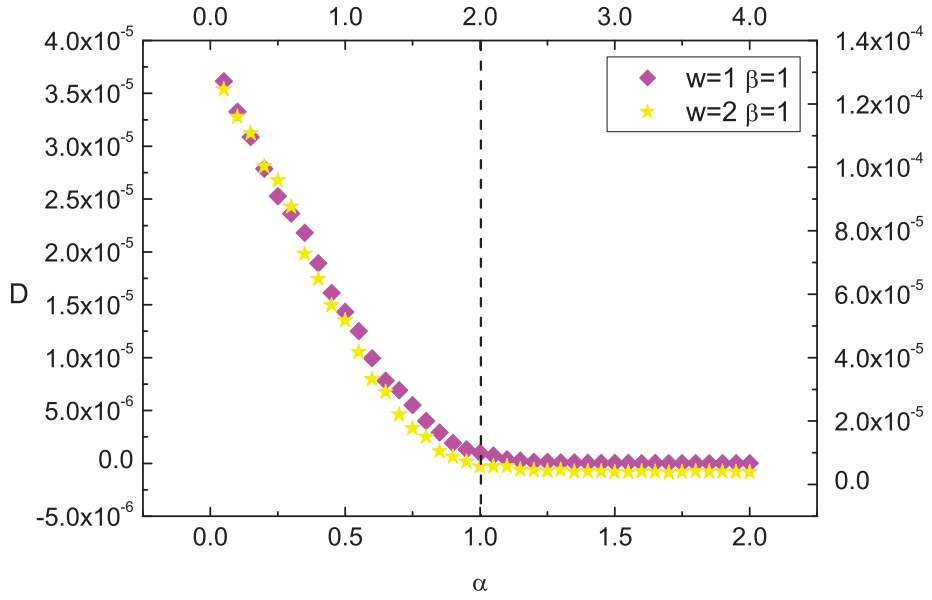

Fig. 5.8: Diffusion coefficient D as a function of parameter α for $w = 1$ and 2 with $\beta = 1$. The left and the lower coordinates correspond to the one-dimension ($w = 1$) line network space, and the right and the upper coordinates correspond to the two-dimension ($w = 2$) plane network space. The dash line follows $\alpha = w$.

d^M_{out} is the diffusion sub-coefficient of the nodes in the planar but out of the nth ring and $l_{out} = \frac{1}{2}$. $d^M(n)$ is the diffusion sub-coefficient contributed from the nodes in the nth ring, satisfying

$$d^M(n) = -\frac{J^M(n)}{\nabla c(n)}. \tag{5.19}$$

In Fig. 5.6(c), $M = 2$, $J^2(1) = 3$, $J^2(2) = 1$, $N = 10$, $l_1^2 = \frac{1}{8}$ and $l_2^2 = \frac{3}{8}$. According to Eq. (5.16), $d^2(1) = \frac{3}{80}$, $d^2(2) = \frac{3}{80}$, $D^2 = \frac{3}{40}$. When M$\to \infty$, $D^M \to D$.

According to the spatial network evolution mechanism, the number of agents, which connect to the seed, in the nth ring is

$$\begin{aligned} N_C^M(n) &\sim N^M(n)(l_n^M)^{-\alpha} \\ &\sim \frac{2\pi N(n-\frac{1}{2})^{-\alpha+1}}{(2M)^{-\alpha+2}}. \end{aligned} \tag{5.20}$$

Thus, the diffusion flux from the seed to the nth ring satisfies

$$J^M(n) \sim N_C^M(n)(l_n^M)^{-\beta}$$

$$\sim \frac{2\pi N(n-\frac{1}{2})^{-\alpha-\beta+1}}{(2M)^{-\alpha-\beta+2}}. \tag{5.21}$$

Applying Eq. (5.16), the diffusion sub-coefficient d of the nth ring is affected by n:

$$d^M(n) \sim (n-\frac{1}{2})^{-\alpha-\beta+2}. \tag{5.22}$$

For large M, each ring has less than two nodes. We use Eq. (5.16) to calculate d for each of the nodes, and D follows

$$D = -\sum_{p\in\Theta_s} \frac{1}{\frac{\Delta c}{l_{sp}}}, \tag{5.23}$$

where Θ_s is the set of active neighbors, which diffuse from seed s.

5.3.3 Numerical results

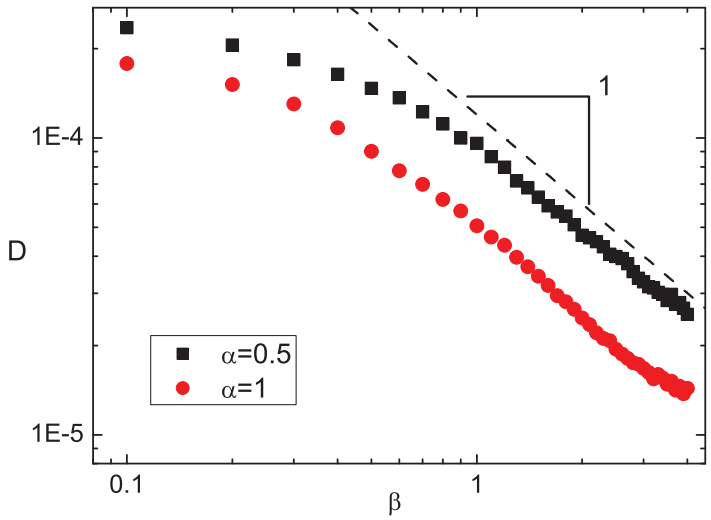

Fig. 5.9: Diffusion coefficient D as a function of parameter β for $\alpha = 0.5$ and 1. The dashed line has a slope -1.

For the traditional Fick's first law, the diffusion coefficient depends on the temperature, viscosity of the fluid and the size of the particles [162]. Now we are looking for which property affects the diffusion coefficient in spatial network. In Fig. 5.7, we discuss the relationship between D and parameter α. We compare three kinds of results for $\beta = 0.5$

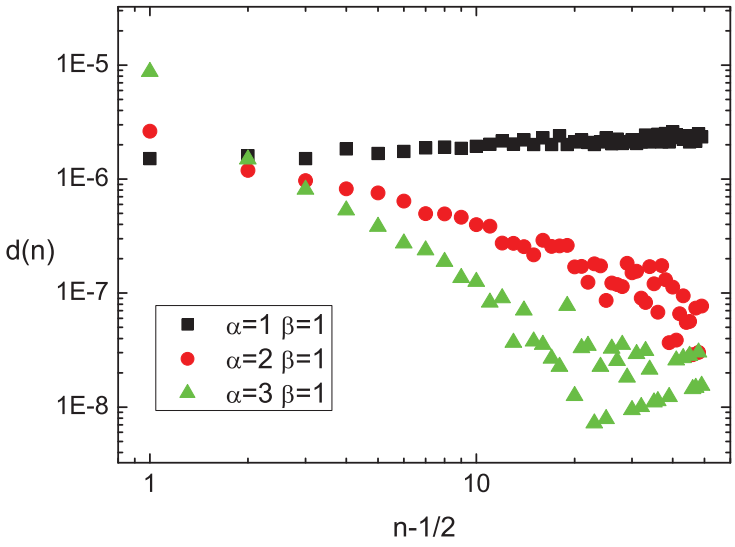

Fig. 5.10: Diffusion sub-coefficient d as a function of $n - \frac{1}{2}$ for $\alpha = 1, 2, 3$ and $\beta = 1$.

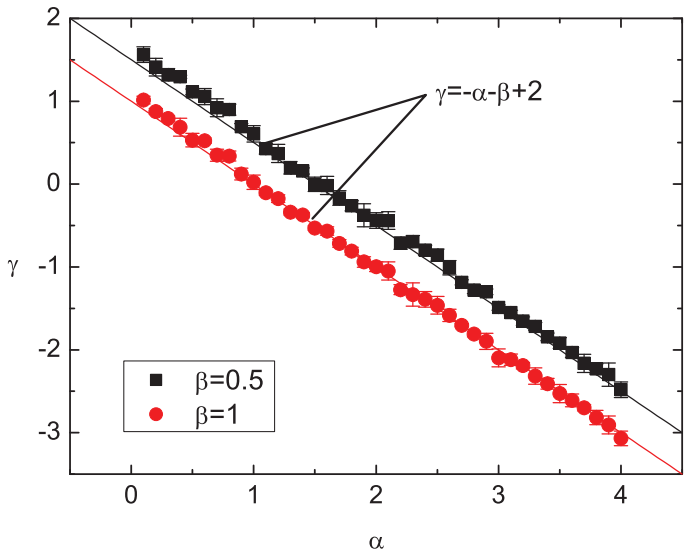

Fig. 5.11: α dependence of the slope γ for $\beta = 0.5, 1$. Points are best fit estimates to the simulation data. The straight lines are plots of Eq. (5.24).

and 1: the first one is the simulation result from Eq. (5.23) on spatial driven network; the second one is the simulation result from Eq. (5.23) on scale-free network [9] with the same

spatial distance distribution of links which we have used in the first one; the third one is the analytical result from Eq. (5.17) on the same spatial driven network. The simulation result matches the analytical result. In Fig. 5.7, we also find that the scale-free network and spatial driven network have the same diffusion coefficient. These two networks have the same spatial distance distribution of links but other various properties such as degree distribution, clustering coefficient, shortest path length. We can get that the diffusion coefficient is determined by the spatial distance distribution of links. Eq. (5.17) shows this property as well. D decreases with the increasing of α. We have done it as well in one-dimensional line ($w = 1$), which has the similar relationship. In Fig. 5.8, the left and the lower coordinates correspond to one-dimensional case ($w = 1$), and the right and the upper coordinates correspond to the two-dimensional case ($w = 2$), the dash line satisfies $\alpha = w$. When α smaller than the network dimension w, D decreases quickly; on the contrary, it decreases slowly.

β is the parameter of spatial diffusion mechanism corresponding to the property of diffusion particles in Fick's first law. In Fig. 5.9, we discuss how D is affected by parameter β on different spatial driven networks with $\alpha = 0.5$ and 1. For different networks, D always decreases with the increasing of β, which we have got from Eq. (5.17) as well. When β is larger than 1, D decreases following the power of β with exponent -1, as Fig. 5.9 shows.

In Fig. 5.10, the points suggest that for $\alpha = 1, 2$ and 3, d scales as $d \sim (n - \frac{1}{2})^{\gamma}$. The variation of the exponent γ with $n - \frac{1}{2}$ is shown in Fig. 5.11. The error bars are determined by the fitting with $d \sim (n - \frac{1}{2})^{\gamma}$. The results suggest that approximately

$$\gamma = -\alpha - \beta + 2 \tag{5.24}$$

Figs. 5.10 and 5.11 show a good agreement between the results of numerical simulation and analytical calculation from Eq. (5.22) for different values of α and β. When $\alpha < -\beta + 2$, long-range connection neighbors are in the main role of diffusion coefficient, on the contrary, short-range connection neighbors are in the main role.

5.4 Anomalous diffusion on spatial network

Then we introduce anomalous diffusion on two-dimension spatial driven network. For two-dimensional anomalous diffusion, the scaling of the mean square displacement $\langle l^2 \rangle$ with

time t follows

$$\langle l^2 \rangle = 4D't^{\delta}. \tag{5.25}$$

At first, we select an agent s randomly as an active seed, set s to the central of the planar and rescale all the nodes. The diffusion probability λ_{sj} from the active agent s to its inactive neighbor j is inversely proportional to their spatial distance and follows Eq. (5.15) as well. What's more, at each time step, only one inactive neighbor changes its state. For the active agent, it will change its state at the same time. Thus at every time step, there is only one active agent in the network. At time t, we mark the active agent as s_t. Thus $l(t)$ is the Euclidean distance between active seed s and active agent s_t. The variation of mean squared displacement for different spatial driven network with $\alpha = 1$, 2, 3, 4 and 5 is shown in Fig. 5.12 inset. The active agent diffuse from the active seed constraining by diffusion probability. Since the limited space effect, when diffusion time t is large enough, the active agent diffuse to any random place on the planar. The limited value of mean squared displacement can be calculated as

$$\frac{\int_{-0.5}^{0.5}\int_{-0.5}^{0.5}(x^2+y^2)dxdy}{\int_{-0.5}^{0.5}\int_{-0.5}^{0.5}dxdy} \approx 0.1667. \tag{5.26}$$

The diffusion process only works before the network approaches the limited space effect. When $\alpha = 1$, the active time is less than 20 time steps. The active time increase with the increasing of α. In order to reduce the limited space effect, we expand the network to 3×3 range with periodic boundary condition. The variation of mean squared displacement for different expand spatial driven network is shown in Fig. 5.12. The mean squared displacement of the active agent is a power law function of time as Eq. (5.25), whose slope δ is determined by the network parameter α. In a normal diffusion process, $\delta = 1$.

The relationship between slope δ and network parameter α is shown in Fig. 5.13. The error bars are determined by the fitting with Eq. (5.25) in diffusion process. When $\alpha < 3$, slope $\delta < 1$; otherwise $\delta \approx 1$. The diffusion process on spatial network with smaller α is sub-diffusion. δ increases with the increasing of α and tends to 1 (normal diffusion). Using Eq. (5.25) to fit the variation of the mean squared displacement, we can identify the diffusion constant D'. When α smaller than the network dimension 2, D' decreases quickly with the increasing of α; on the contrary, it decreases slowly (in Fig. 5.14). D' behaviors in the same way as D in Fig. 5.9. Network diffusion ability becomes weaker with increasing α.

In order to reveal the spatial network effect on spatial diffusion, we do realizations

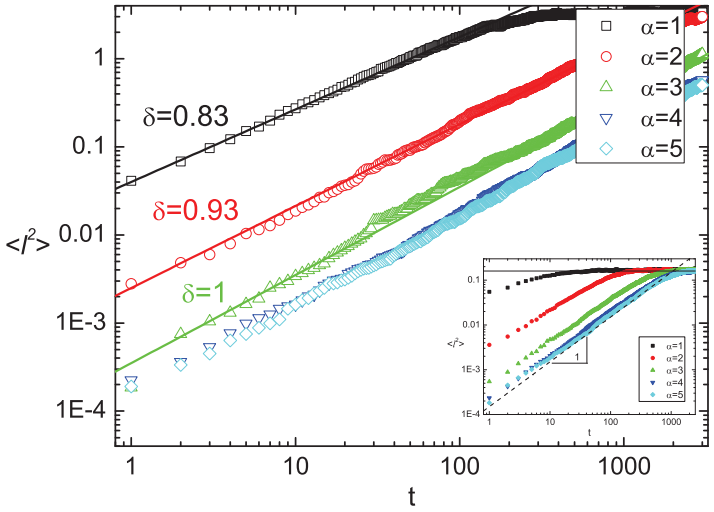

Fig. 5.12: The variation of mean squared displacement for 3×3 expand spatial driven network with different α. inset: The variation of mean squared displacement for 1×1 spatial driven network with different α. The full line satisfies $\langle l^2 \rangle = 0.1667$; the dash line has a slope 1. The nodes represent the ensemble average results for 1000 simulations on a network sample.

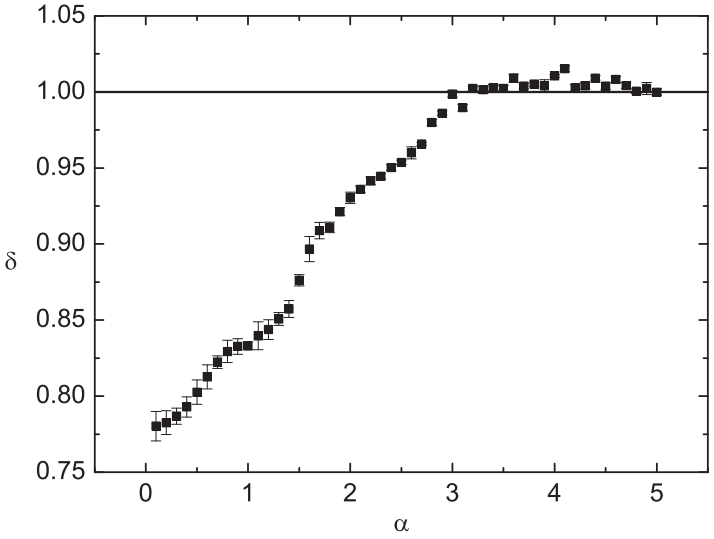

Fig. 5.13: The slope δ dependence of the parameter α. The straight line satisfies $\delta = 1$ for normal diffusion.

of the anomalous diffusion on different spatial networks with $\alpha = 1$ and 4. Fig. 5.15 shows the comparison of 100 steps simulation trajectories of anomalous diffusion on spatial network with $\alpha = 1$ and spatial network with $\alpha = 4$. When $\alpha = 1$, the network structure is beneficial for diffusion, existing a lot of long-range diffusion. In 100 time steps, the

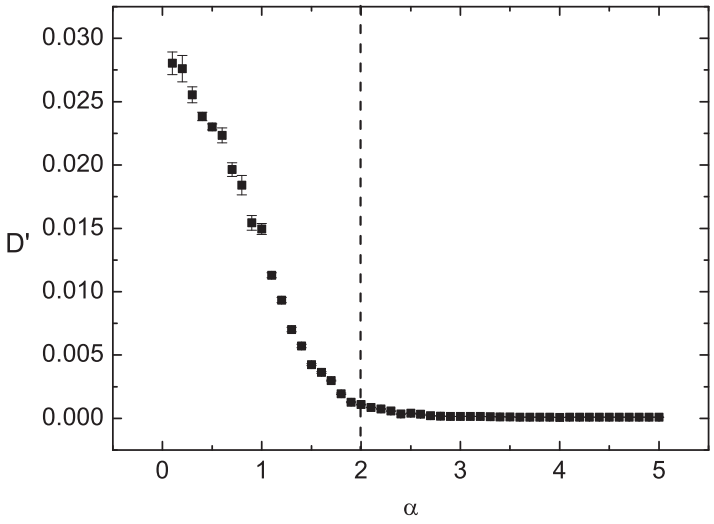

Fig. 5.14: Anomalous diffusion coefficient D' as a function of parameter α. The dash line follows $\alpha = 2$.

product has diffused on the whole planar. On the contrary, when $\alpha = 4$, the network structure is disadvantage for diffusion, only short-range diffusion are existed. The product diffuses very slowly.

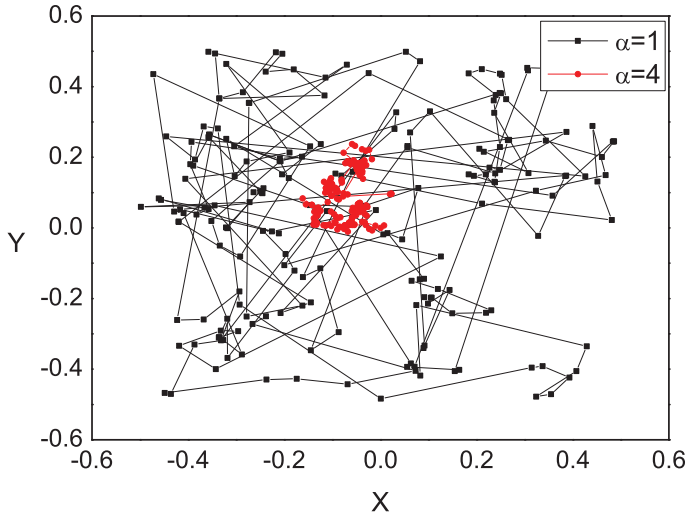

Fig. 5.15: Two trajectories of anomalous diffusion on spatial network (a) $\alpha = 1$ (b) $\alpha = 4$ with $N = 10000$. Both trajectories are simulation results for 100 time steps.

5.5 Conclusion

In this chapter, we use the Fick's first law and anomalous diffusion to solve the diffusion dynamics on spatial networks. In the method of Fick's first law, we introduce a composite diffusion coefficient D in order to measure the diffusion ability of the spatial driven networks. Since our diffusion is not normal in general due to the long-range spreading, this coefficient must take into account the transfers of diffusive substances over all distances. We get coefficient D using both of theoretical method and simulation method and find that D is determined by spatial distribution of links. In simulation method, we divide the planer to M rings, and apply the Fick's first law on each ring to get sub-coefficient d. For larger M, according to the simulation method, D is the sum of diffusion sub-coefficient for all the rings. When α is smaller than the topological dimension of the space, the diffusion coefficient decreases quickly with the increasing of α. On the contrary, it decreases slowly. For different diffusion mechanism, D always decreases with the increasing of β, where β is the parameter of diffusion mechanism. When $\beta > 1$, D decreases following the power of β with exponent -1. Finally, the diffusion sub-coefficient $d(n)$ is determined by the network structure and position of the ring, which follows $d(n) \sim (n-\frac{1}{2})^{-\alpha-\beta+2}$, where n is the rank of the ring. In the method of anomalous diffusion, the diffusion probability is inversely with the Euclidean distance between connected agents as well. From the variation of mean squared displacement, we get that the slope index δ in anomalous diffusion increases with the increasing of α and then approaches to 1. The diffusion process in our model is always sub-diffusion. The diffusion constant D' has the same behavior as D, which is decreases with the increasing of α.

In anomalous diffusion, the diffusion process is characterized through the scaling of the mean square displacement with time. This is a mature method to discuss the diffusion process. But it works only on the traditional anomalous diffusion process. Once the scaling property of the mean square displacement with time does not exist, the anomalous diffusion method does not work. In this case, the method of Fick's law is a good choice to discuss the diffusion process. In section 5.3, the diffusion dynamics works on more than one pair of agents. In this condition, the mean square displacement does not easy to get. Thus, using the Fick's law on all pairs of agents is a more suitable method.

Chapter 6

Summary and outlook

In this thesis, we have proposed a spatial network model in configuration space, and discussed diffusion dynamics on spatial driven model. Here, we shortly review our main results what we found in different subjects.

In the third chapter, we introduce an evolutional spatial network model in configuration space. The model grows following a competition between the degree and the spatial distance preferences. Network measures are discussed in this chapter. Controlling by parameters a and α, the model can be divided to three parts, spatial driven model, composed spatial model and Barabási-Albert network model. In spatial driven model, the spatial distribution follows the power law, the degree distribution follows exponential law. In composed spatial model, the degree distribution follows shifted power law. We also apply the spatial network model to fit an empirical data from email network. The qualitatively consistent reveals that the model has captured some basic mechanisms for the evolution of some real networks. In the future, we will apply this model to fit different kinds of empirical data and try to find the characteristics of the parameters in different real networks. We hope that it will be helpful for further study and understanding of real networks whose evolution is influenced by the interplay of different even competing dynamics. And introducing the competition into WS model will be another interesting spatial network model.

In the fourth chapter, we define spatial susceptible-infected-susceptible epidemic spreading process on spatial driven model. The spreading probability is inversely proportional to the spatial distance. Infected nodes ratio, the steady infected nodes ratio and the epidemic threshold are discussed. We find that, when the network has smaller number

of long distance connections, which needs longer time to reach to the steady state, and the effective spreading time increases with increasing α. But the steady infected nodes ratio ρ_s has different behavior, which has the minimum value. What's more, we find that the epidemic threshold exists without exception, whose value is determined by parameter α. The maximum of λ_c and the minimum of ρ_s always exist in the network with α in the interval from 1.5 to 2. In the following, we will introduce spatial property to other spreading processes, such as SIR model, opinion model et *al*.. We hope that our work can help us to understand how spreading process works on different spatial networks.

Since spatial network has well defined spatial distance, the spatial driven model offers an occasion to use the usual diffusion equations to describe the diffusive dynamics. In the fifth chapter, we introduce two diffusion methods, Fick's first law and anomalous diffusion, to study the diffusion process on spatial driven network. From these two methods, we can get two kinds of diffusion coefficient. Both methods reveal the same property, when α is smaller than the topological dimension of the space, the diffusion coefficient decreases quickly with the increasing of α. On the contrary, it decreases slowly. In the method of Fick's first law, for different diffusion mechanism, D always decreases with the increasing of β, where β is the parameter of diffusion mechanism. When $\beta > 1$, D decreases following the power of β with exponent -1. Finally, the diffusion sub-coefficient $d(n)$ is determined by the network structure and position of the ring, which follows $d(n) \sim (n - \frac{1}{2})^{-\alpha-\beta+2}$, where n is the rank of the ring. When we use the anomalous diffusion equation, we find that the diffusion process on spatial network is subdiffusion. The slope index δ increases with the increasing of α and then approaches to 1 (normal diffusion). For the future, we will try to apply the diffusion coefficient on different dynamic processes, such as opinion dynamic (voter model, Sznajd model), spreading dynamic (SIS model, SIR model) et *al*.. We hope that, via the future works, we can find the suitable diffusion coefficient for different dynamics processes.

Inspire from the empirical studies and models, we find that spatial networks attract more and more attention. Despite these various advances, there are still many open problems which could represent interesting research directions both at the theoretical and the applied levels.

There are a lot of interesting projects we can do on the spatial network model. We will try to define the efficiency index to value the spatial network, and try to find the relationship between the efficiency index and the dynamics behavior. We concern the fractal property of the model as well. Calculating the fractal dimension on spatial network

will be another problem. There have been several method to calculate the fractal dimension on complex network. But for spatial network, it will be quite different.

How transportation networks evolve is an old problem and was already the subject of many studies in the 1970s [31]. However, apart from some exceptions [170] this problem is still not very well understood. We are now in the position where data and tools are available and we can expect some interesting developments in this area. In parallel to empirical studies, we also need to develop theoretical ideas and models in order to describe the evolution of spatial networks. More generally, infrastructure and transportation networks are part of urban systems and we believe that the current understanding of spatial networks could help in understanding the structure and evolution of these systems. In particular, our knowledge of spatial networks could help in the understanding of important phenomenon such as urban sprawl and in the design of sustainable cities.

Data on spatial networks and in particular, road and other infrastructure networks are now available and these networks have been the subject of many studies. Also, with the emergence of geosocial applications on mobile phones for example we can expect interesting studies connecting spatial distributions and social behavior. This line of research already appeared in recent studies which tried to relate topological structures of networks with socio-economical indicators. In these studies, an important question concerns the correlations between topological quantities and social factors. For example, it would be interesting to know if we can understand some aspects of the spatial distribution of crime rates in terms of topological indicators of the road network.

Bibliography

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

[2] A. L. Barabási, Linked: how everything is connected to everything else and what it means for business, science, and everyday life, Plume, (2003).

[3] P. Erdős and A. Rényi, On random graphs, A. Publ. Math, Debrecen **6**, 290 (1959).

[4] B. Bollobás, Random graphs, Cambridge studies in advanced mathematics, 73 Cambridge University Press, Cambridge, New York, (2001).

[5] B. Bollobás, Graph theory: an introductory course, Graduate texts in mathematics, 63, Springer Verlag, New York, (1990).

[6] J. P. Scott, Social network analysis: A handbook, Sage Publications, London, 2nd edition, (2000).

[7] S. Wasserman, K. Faust, Social Networks Analysis, Cambridge University Press, Cambridge, (1994).

[8] D. J. Watts, S. H. Strogatz, Collective dynamics of 'small-world' networks, Nature, **393**, 440 (1998).

[9] A. L. Barabási, R. Albert, Emergence of scaling in random networks, Science, **286**, 509 (1999).

[10] M. Faloutsos, P. Faloutsos, and C. Faloutsos, On power law relationships of the Internet topology, Computer Communication Review, **29**, 251 (1999).

[11] R. Albert, H. Jeong, and A. L. Barabási, Diameter of the World Wide Web, Nature, **401** 130 (1999).

[12] R. Albert, H. Jeong, and A. L. Barabási, Error and attack tolerance of complex networks, Nature, **406**, 378 (2000).

[13] R. Guimerá and L. A. N. Amaral, Functional cartography of complex metabolic networks, Nature, **433**, 895 (2005).

[14] R. Guimera, S. Mossa, A. Turtschi, and L. A. N. Amaral, The worldwide air transporation network: anomalous centrality, community structure, and cities' global roles, Proc. Natl. Acad. Sci, **102**, 7794 (2005).

[15] M. T. Gastner and M. E. Newman, The spatial structure of networks, Eur. Phys. J. B, **49**, 247 (2006).

[16] H. Jeong, S. P. Mason, A. L. Barabási, and Z. N. Oltvai, Lethality and centrality in protein networks, Nature, **411**, 41 (2001).

[17] A. Krause, K. A. Frank, D. M. Mason, R. E. Ulanowicz, and W. M. Taylor, Compartments exposed in food-web structure, Nature, **426**, 282 (2003).

[18] M. E. J. Newman, The structure of scientific collaboration networks, Proceedings of the National Academy of Sciences, **98**, 404 (2001).

[19] G. Caldarelli, R. Marchetti, and L. Pietronero, The fractal properties of internet, Europhys. Lett, **52**, 386 (2000).

[20] A. Barrat, M. Barthélemy, and A. Vespignani, Dynamical processes in complex networks, Cambridge University Press, Cambridge, UK, (2008).

[21] G. Caldarelli, Scale-free networks, Oxford University Press, Oxford, (2007).

[22] R. Cohen and S. Havlin, Complex Networks: Structure, Robustness and Function, Cambridge University Press, Cambridge, (2010).

[23] S. N. Dorogovtsev and J. F. F. Mendes, Evolution of networks: From biological nets to the Internet and WWW, Oxford University Press, Oxford, (2003).

[24] M.E.J. Newman, Networks: an introduction, Oxford University Press, Oxford, UK, (2010).

[25] R. Pastor-Satorras and A. Vespignani, Evolution and structure of the Internet: A statistical physics approach, Cambridge University Press, Cambridge, (2003).

[26] S. N. Dorogovtsev and J. F. F. Mendes, Evolution of networks, Adv. Phys., **51**, 1079 (2002).

[27] S. Boccaletti, V. Latora, Y. Moreno, M. Chavez, D. U. Hwang, Complex networks: Structure and dynamics, Physics Reports, **424**, 175 (2006).

[28] R. Albert, A. L. Barabási, Statistical mechanics of complex networks, Rev. Mod. Phys, **74**, 47 (2002).

[29] M. Barthélemy, Spatial Networks, Physics Reports, **499**, 1 (2011).

[30] E. Bullmore and O. Sporns, Complex brain networks: graph theoretical analysis of structural and functional systems, Nature Reviews Neuroscience, **10**, 186 (2009).

[31] P. Haggett and R.J. Chorley, Network analysis in geography, Edward Arnold, London, (1969).

[32] R.J. Chorley and P. Haggett, Models in geography, Methuen and Co., London, (1967).

[33] S. Maslov and K. Sneppen, Specificity and stability in topology of protein networks, Science, **296**, 210 (2002).

[34] B. Bollobás, Random Graphs, Academic Press, London, (1985).

[35] B. Bollobás, Modern Graph Theory, Graducate Texts in Mathematics, Springer, New York, (1998).

[36] F. Harary, Graph Theory, Perseus, Cambridge, MA, (1995).

[37] A. Vazquez, R. Pastor-Satorras, and A. Vespignani. Large-scale topological and dynamical properties of internet. Phys. Rev. E, **65**, 066130 (2002).

[38] S. N. Dorogovtsev, A. V. Goltsev, and J. E. F. Mendes, Pseudofractal scale-free web, Phys. Rev. E, **65**, 066122 (2002).

[39] S. Jung, S. Kim, and B. Kahng, Geometric fractal growth model for scale free networks, Phys. Rev. E, **65**, 056101 (2002).

[40] E. Ravasz and A. L. Barabási, Hierarchical organization in complex networks, Phys. Rev. E, **67**, 026112 (2003).

[41] E. Ravasz, A. L. Somera, D. A. Mongru, Z. Oltvai, and A. L. Barabási, Hierarchical organization of modularity in metabolic networks, Science, **297**, 1551 (2002).

[42] A. Vazquez, Growing network with local rules: Preferential attachment, clustering hierarchy, and degree correlations, Phys. Rev. E, **67**, 056104 (2003).

[43] A. Capocci, G. Caldarelli, and P. De Los Rios, Quantitative description and modeling of real networks, Phys. Rev. E, **68**, 047101 (2003).

[44] V. Latora, M. Marchiori, A measure of centrality based on the network efficiency, preprint cond-mat/0402050 (2005).

[45] L. C. Freeman, Centrality in Social Networks: Conceptual clarification, Social Networks, **1**, 215 (1979).

[46] L. C. Freeman, A Set of Measures of Centrality Based on Betweenness, Sociometry **40**, 35 (1977).

[47] R. Sedgewick, Algorithms in C++, Part 5: Graph Algorithms, Addison Wesley, Boston MA, (1988).

[48] T. H. Cormen, C. E. Leiserson, R. L. Rivest, C. Stein, Introduction to Algorithms, MIT University Press, Cambridge, (2001).

[49] R. K. Ahuja, T. L. Magnati, J. B. Orlin, Network Flows: Theory, Algorithms, and Applications, Prentice-Hall, Englewood cliffs, NJ, (1993).

[50] U. Brandes, A faster algorithm for betweenness centrality, J. Math. Soc, **25**, 163 (2001).

[51] M. E. J. Newman, Scientific collaboration networks. II. Shortest paths, weighted networks, and centrality, Phys. Rev. E, **64**, 016132 (2001).

[52] K. I. Goh, E. S. Oh, H. Jeong, B. Kahng, D. Kim, Classification of scale-free networks, Proc. Natl. Acad. Sci, USA **99**, 125832 (2002).

[53] P. Crucitti, V. Latora, S. Porta, Centrality Measures in Spatial Networks of Urban Streets, preprint physics/0504163 (2005).

[54] M. Barthelemy, Betweenness centrality in large complex networks, Eur. Phys. J. B, **38**, 163 (2004).

[55] K. I. Goh, C. M. Ghim, B. Kahng, D. Kim, Comment on universal behavior of the load distribution in scale-free networks-Reply, Phys. Rev. Lett, **91**, 189804 (2003).

[56] M. Barthelemy, Comment on 'Universal behavior of load distribution in scale-free networks', Phys. Rev. Lett, **91**, 189803 (2003).

[57] K. I. Goh, B. Kahng, D. Kim, Universal behavior of load distribution in scale-free networks, Phys. Rev. Lett, **87**, 278701 (2001).

[58] P. Crucitti, V. Latora, M. Marchiori, A topological analysis of the Italian electric power grid, Physica A **338**, 92 (2004).

[59] R. Guimera and L. A. N. Amaral, Modeling the world-wide airport network, Eur. Phys. J. B, **38**, 381 (2004).

[60] K. I. Goh, E. Oh, B. Kahng, D. Kim, Betweenness centrality correlation in social networks, Phys. Rev. E, **67**, 017101 (2003).

[61] M. E. J. Newman, M. Girvan, Finding and evaluating community structure in networks, Phys. Rev. E, **69**, 026113 (2004).

[62] R. Pastor-Satorras, A. Vazquez, and A. Vespignani, Dynamical and correlation properties of the internet, Phys. Rev. Lett, **87**, 258701 (2001).

[63] J. Moody, Race, school integration, and friendship segregation in American, Am. J. Social, **107**, 679 (2001).

[64] B. Everitt, Cluter analysis, John Wiley, New York (1974).

[65] R. D. Alba, A graph-theoretic definition of a sociometric clique, J. Math. Sociol, **3**, 113 (1973).

[66] G. W. Flake, S. R. Lawrence, C. L. Giles, F. M. Coetzee, Self-organization and identification of web communities, IEEE Comput, **35**, 66 (2002).

[67] S. B. Seidman, Internal cohesion of LS sets in graphs, Social Networks, **5**, 97 (1983).

[68] S. P. Borgatti, M. G. Everett, P. R. Shirey, LS sets, lambda sets and other cohesive subsets, Social Networks, **12**, 337 (1990).

[69] F. Radicchi, C. Castellano, F. Cecconi, V. Loreto, D. Parisi, Defining and identifying communities in networks, Proc. Natl. Acad. Sci, USA **101**, 2658 (2004).

[70] S. Shen Orr, R. Milo, S. Mangan, U. Alon, Network motifs in the transcriptional regulation network of Escherichia coli, Nature Gen, **31**, 64 (2002).

[71] N. Kashtan, S. Itzkovitz, R. Milo, U. ALon, Efficient sampling algorithm for estimating subgraph concentrations and detecting network motifs, Bioinformatics, **20**, 1746 (2004).

[72] R. Milo, S. Shen-Orr, S. Itzkovitz, N. Kashtan, D. Chklovskii, U. Alon, Network Motifs: Simple Building Blocks of Complex Networks, Science, **298**, 824 (2002).

[73] R. Milo, S. Itzkovitz, N. Kashtan, R. Levitt, S. Shen-Orr, I. Ayzenshtat, M. Sheffer, U. Alon, Superfamilies of designed and evolved networks, Science, **303**, 1538 (2004).

[74] S. Mangan, U. Alon, Structure and function of the feed-forward loop network motif, PNAS, **100**, 11980 (2003).

[75] K. I. Goh, B. Kahng, D. Kim, Spectra and eigenvectors of scale-free networks, Phys. Rev. E, **64**, 051903 (2001).

[76] I. J. Farkas, I. Derenyi, A. L. Barabási, et. al., Spectra of "real-world" graphs: Beyond the semicircle law, Phys. Rev. E, **64**, 026704 (2001).

[77] E. P. Wigner, On the distribution of the roots of certain symmetric matrices, Ann. Math, **67**, 325 (1958).

[78] E. P. Wigner, Characteristic vectors of bordered matrices with infinite dimensions 2, Ann. Math, **65**, 203 (1957).

[79] E. P. Wigner, Characteristic vectors of bordered matrices with infinite dimensions 1, Ann. Math, **62**, 548 (1955).

[80] T. Guhr, A. M. Groeling, and H. A. Weidenmuller, Random matrix theories in quantum physics: common concepts, Physics. Reports, **299**, 189 (1998).

[81] A. Crisanti, G. Paladin, and A. Vulpiani, Products of Random Matrices in Statistical Physics (Springer, Berlin), (1993).

[82] M. L. Mehta, Random Matrices, 2nd ed. (Academic New York) (1991).

[83] S. Milgram, The Small World Problem, Psychol. Today, **1**, 60 (1967).

[84] J. Travers, S. Milgram, An Experimental Study of the Small World Problem, Sociometry, **32**, 425 (1969).

[85] C. Korte, S. Milgram, J. Personal, Acquaintance linking between white and negro populations: application of the small world problem, Social Psychol, **15**, 101 (1970).

[86] D. J. Watts, Small Worlds: The Dynamics of Networks between Order and Randomness, Princeton University Press, Princeton, NJ, (1999).

[87] M. E. J. Newman, Scientific collaboration networks. I. Network construction and fundamental results, Phys. Rev. E, **64**, 016131 (2001).

[88] V. Latora, M. Marchiori, Efficient Behavior of Small-World Networks, Phys. Rev. Lett, **87**, 198701 (2001).

[89] V. Latora, M. Marchiori, Economic Small-World Behavior in Weighted Networks, Eur. Phys. J. B, **32**, 249 (2003).

[90] H. E. Stanley, Introduction to Phase Transitions and Critical Phenomena, Oxford University Press, New York, (1971).

[91] K. J. Falconer, Fractal Geometry: Mathematical Foudations and Applications, Wiley, (1990).

[92] C. Song, S. Havlin, H. A. Makse, Self-Similarity of Complex Networks, Nature, **433**, 392 (2005).

[93] L. A. N. Amaral, A. Scala, M. Barthelemy, H. E. Stanley, Classes of small-world networks, Proc. Natl. Acad. Sci, **97**, 11149 (2000).

[94] R. Pastor-Satorras, A. Vespignani, Evolution and Structure of the Internet: A Statistical Physics Approach, Cambridge University Press, Cambridge, (2004).

[95] P. Erdős and A. Rényi, On the evolution of random graphs, Publ. Math. Inst. Hung. Acad. Sci, **5**, 17 (1960).

[96] P. Erdős and A. Rényi, On the strength of connectedness of a random graph, Acta Mathematica Scientia Hungary, **12**, 261 (1961).

[97] M. L. Mehta, Random matrices, Elsevier/Academic Press, Amsterdam, (2004).

[98] J. Park and M.E.J. Newman, The Statistical Mechanics of Networks, Phys. Rev. E, **70**, 066117 (2004).

[99] M. Molloy and B. Reed, A critical point for random graphs with a given degree sequence, Random Structures and Algorithms **6**, 161 (1995).

[100] M. E. J. Newman, S. H. Strogatz, and D. J. Watts, Random graphs with arbitrary degree distributions and their applications, Phys. Rev. E, **64**, 026118 (2001).

[101] D. B. West, Introduction to Graph Theory, Prentice Hall, Englewood Cliffs, NJ, (1995).

[102] M. Molloy and B. Reed, The Size of the Largest Component of a Random Graph on a Fixed Degree Sequence, Combinatories, Probability and Computing **7**, 295 (1998).

[103] D. J. Watts, Networks, dynamics and the small world phenomenon, Amer J. Sociol, **105**, 493 (1999).

[104] M. E. J. Newman, D. J. Watts, Renormalization group analysis of the small-world network model, Phys. Lett. A, **263**, 341 (1999).

[105] M. E. J. Newman, D. J. Watts, Scaling and percolation in the small-world network model, Phys. Rev. E, **60**, 7332 (1999).

[106] M. E. J. Newman, The structure and function of networks, Computer Physics Communications, **147**, 40 (2002).

[107] M. Barthelemy, and L. A. N. Amaral, Small-world networks: Evidence for a crossover picture, Phys. Rev. Lett, **82**, 3180 (1999).

[108] M. E. J. Newman, C. Moore, and D. J. Watts, Mean-field solution of the small-world network mdel, Phys. Rev. Lett, **84**, 3201 (2000).

[109] A. Barrat and M. Weigt, On the properties of small-world networks, Eur. Phys. J. B, **13**, 547 (2000).

(1976).

[110] B. Bollobas, O. Riordan, The diameter of a scalefree random graph, Preprint, Department of Mathematical Sciences, University of Memphis, (2002).

[111] A. L. Barabási, R. Albert, and H. Jeong, Mean-field theory for scale free random networks, Physica A, **272**, 173 (1999).

[112] P. L. Krapivsky, S. Redner, and F. Leyvraz, Connectivity of growing random networks, Phys. Rev. Lett, **85**, 4629 (2000).

[113] S. N. Dorogovtsev, J. F. F. Mendes, and A. N. Samukhin, Structure of growing networks with prederential linking, Phys. Rev. Lett, **85**, 4633 (2000).

[114] P. L. Krapivsky and S. Redner, Organization of growing random networks, Phys. Rev. E, **63**, 066123 (2001).

[115] S. N. Dorogovtsev and J. F. F. Mendes, Effect of the accelerating growth of communications networks on their structure, Phys. Rev. E, **63**, (2) 025101 (2001).

[116] A. Barrat, M. Barthélemy, and A. Vespignani, The effects of spatial constraints on the evolution of weighted complex networks, J. Stat. Mech., page P05003 (2005).

[117] A.P. Masucci, D. Smith, A. Crooks, and M. Batty, Random planar graphs and the london street network. Eur. Phys. J. B, **71**, 259 (2009).

[118] R. Lambiotte, V.D. Blondel, C. de Kerchove, E. Huens, C. Prieur, Z. Smoreda, and P. Van Dooren. Geographical dispersal of mobile communication networks, Physica A, **387**, 5317 (2008).

[119] D. Liben-Nowell, J. Nowak, R. Kumar, P. Raghavan, and A. Tomkins, Geographic routing in social networks. Proc. Natl Acad. Sci. (USA), **102**, 11623 (2005).

[120] J. Goldenberg and M. Levy, Distance is not dead: Social interaction and geographical distance in the internet era, arXiv, page 0906.3202 (2009).

[121] J. Dall and M. Christensen, Random geometric graphs, Phys. Rev. E, **66**, 016121 (2002).

[122] C. Herrmann, M. Barthélemy, and P. Provero, Connectivity distribution of spatial networks, Phys. Rev. E, **68**, 026128 (2003).

[123] B.M. Waxman, Routing of multipoint connections, IEEE J. Select. Areas. Commun., **6**,1617 (1988).

[124] C. McDiarmid, A. Steger, and D.J.A. Welsh, Random planar graphs, Journal of Combinatorial Theory, **93**, 187 (2005).

[125] J. M. Kleinberg, Navigation in a small world, Nature, **406**, 845 (2000).

[126] P. Sen and B.K. Chakrabarti, Small-world phenomena and the statistics of linear polymer, J. Phys. A, **34**, 7749 (2001).

[127] P. Sen and S. S. Manna, Clustering properties of a generalized critical Euclidean network, Phys. Rev. E, **68**, 026104 (2003).

[128] J. Ozik, B. R. Hunt and E. Ott, Growing networks with geographical attachment preference: Emergence of small worlds, Phys. Rev. E, **69**, 026108 (2004).

[129] K. Kosmidis, S. Havlin and A. Bunde, Structural properties of spatially embedded networks, Europhys. Lett., **82**, 48005 (2008).

[130] L. H. Wang, P. Pattison and G. Robins, A spatial model for social networks, Physica A, **360**, 99 (2006).

[131] S. S. Manna and P. Sen, Modulated scale-free network in Euclidean space, Phys. Rev. E, **66**, 066114 (2002).

[132] M. Barthélemy, Crossover from Scale-Free to Spatial Networks, Europhys. Lett., **63**, 915 (2003).

[133] S. -H. Yook, H. Jeong and A. -L. Barabási, Modeling the Internet's large-scale topology, Proc. Natl. Acad. Sci., **99**, 13382 (2002).

[134] Y. Hayashi, A Review of Recent Studies of Geographical Scale-Free Networks, IPSJ Digital Courier, **2**, 155 (2006).

[135] R. Xulvi-Brunet and I. M. Sokolov, Evolving networks with disadvantaged long-range connections, Phys. Rev. E, **66**, 026118 (2002).

[136] Z. Z. Zhang, S. G. Zhou, Z. Shen and J. H. Guan, From regular to growing small-world networks, Physica A, **385**, 765 (2007).

[137] Z. Z. Zhang, L. L. Rong and F. Comellas, Evolving small-world networks with geographical attachment preference, Physica A, **39**, 3253 (2006).

[138] P.-P. Zhang, K. Chen, Y. He, T. Zhou, B.-B. Su, Y.-D. Jin, H. Chang, Y.-P. Zhou, L.-C. Sun, B.-H. Wang and D.-R. He, Model and empirical study on some collaboration networks, Physica A **360**, 599 (2006).

[139] H. Chang, B.-B. Su, Y.-P. Zhou and D.-R. He, Assortativity and act degree distribution of some collaboration networks, Physica A **383**, 687 (2007).

[140] A. Vazquez, R. Pastor-Satorras and A. Vespignani, Internet topology at the router and autonomous system level, arXiv, page 0206.084 (2002).

[141] R. Guimerà, L. Danon, A. Díaz-Guilera, F. Giralt, and A. Arenas, Self-similar community structure in a network of human interactions, Phys. Rev. E **68**, 65103 (2003).

[142] M. L. Goldstein, S. A. Morris and G. G. Yen, Problems with fitting to the power-law distribution, Eur. Phys. J. B, **41**, 255 (2004).

[143] W. -P. Guo, X. Li and X. -F. Wang, Epidemics and immunization on Euclidean distance preferred small-world networks, Physica A, **380**, 684 (2007).

[144] P. Sun, X. -B. Cao, W. -B. Du and C. -L. Chen, The effect of geographical distance on epidemic spreading, Phys. Pro., **3**, 1811 (2010).

[145] B. Wang, K. Aihara and B. J. Kim, Comparison of immunization strategies in geographical networks, Phys. Lett. A, **373**, 3877 (2009).

[146] R. M. Anderson and R. M. May, Infectious Diseases of Humans, Oxford University Press, Oxford, (1991).

[147] N. T. J. Bailey, The Mathematical Theory of Infectious Diseases and Its Applications, Hafner Press, New York, (1975).

[148] H. W. Hethcote, The mathematics of infectious diseases, SIAM Review, **42**, 599 (2000).

[149] R. Pastor-Satorras and A. Vespignani, Epidemic spreading in scale-free networks, Phys. Rev. Lett., **86**, 3200 (2001).

[150] R. Pastor-Satorras and A. Vespignani, Epidemics and immunization in scale-free networks, in S. Bornholdt and H. G. Schuster (eds.), Handbook of Graphs and Networks, Wiley-VCH, Berlin, (2003).

[151] A. L. Lloyd and R. M. May, How viruses spread among computers and people, Science, **292**, 1316 (2001).

[152] R. M. May and A. L. Lloyd, Infection dynamics on scale-free networks, Phys. Rev. E, **64**, 066112 (2001).

[153] R. M. May and R. M. Anderson, The transmission dynamics of human immunodeficiency virus (HIV), Philos. Trans. R. Soc. London B, **321**, 565 (1988).

[154] R. Pastor-Satorras and A. Vespignani, Epidemic dynamics and endemic states in complex networks, Phys. Rev. E, **63**, 066117 (2001).

[155] X. -J. Xu, X. Zhang and J. F. F. Mendes, Impacts of preference and geography on epidemic spreading, Phys. Rev. E, **76**, 056109 (2007).

[156] M. Newman, The physics of networks, Phys. Today, **61**, 33 (2008).

[157] G. Li, S. D. S. Reis, A. A. Moreira, S. Havlin, H. E. Stanley and J. S. Andrade, Jr., Towards Design Principles for Optimal Transport Networks, Phys. Rev. Lett. **104**, 018701 (2010).

[158] A. Barrat, M. Barth*é*lemy, R. Pastor-Satorras and A. Vespignani, Proc. Natl. Acad. Sci., **101**, 3747 (2004).

[159] V. Colizza, M. Barth*é*lemy, A. Barrat and A. Vespignani, C. R. Biologies **330**, 364 (2007).

[160] H. Hu, S. Myers, V. Colizza and A. Vespignani, Proc. Natl. Acad. Sci. **106**, 1318 (2009).

[161] P. Wang, M. Gonzalez, C. A. Hidalgo and A. -L. Barab*á*si, Science, **324**, 1071 (2009).

[162] J. Crank, The mathematics of diffusion 2ed, Oxford University Press, Oxford, (1975).

[163] H. S. Carslaw and J. C. Jaeger, Conduction of Heat in Solids, Oxford University Press, Oxford, (1986).

[164] L. Kadanoff, Statistical Physics: Statics, Dynamics, and Renormalization, World Scientific, Singapore, (2000).

[165] P. Wilmott, S Howison and J. Dewynne, The Mathematics of Financial Derivatives: A Student Introduction, Cambridge University Press, Cambridge (1995).

[166] L. Vlahos, H. Isliker, Y. Kominis and K. Hizanidis, Normal and anomalous diffusion: A tutorial ,arXiv, page 0805.0419 (2008).

[167] R. Metzler and J. Klafter, The random walk's guide to anomalous diffusion: a fractional dynamics approach, Physics Reports. **339**, 1 (2000).

[168] A. Okubo and S. A. Levin, Diffusion and Ecological Problems: Mathematical Models, Springer, Berlin, (1980).

[169] P. Grindrod, The Theory and Applications of ReactionDiffusion Equations: Patterns and Waves, Oxford University Press, Oxford, (1996).

[170] D. Levinson and B. Yerra, Self-organization of surface transportation networks, Transportation Science, **40**, 179 (2006).

Publications

1. Z. Hui, X. Cai, J.-M. Greneche, Q. A. Wang, Diffusion on spatial network, Under review, Eur. Phys. J. B, 2012.

2. Z. Hui, W. Li, X. Cai, J.-M. Greneche, Q. A. Wang, Structure properties of evolutionary spatially embedded networks, Physica A, **392**, 1909 (2013).

3. Z. Hui, X. Cai, J.-M. Greneche, Q. A. Wang, Impacts of spatial structure on epidemic spreading, Int. J. Mod. Phys. C , **23**, 1250082 (2012).

4. Z. Hui, X. Cai, J.-M. Greneche, Q. A. Wang, Structure and collaboration relationship analysis in a scientific collaboration network, Chinese Sci Bull, **56**, 3702 (2011).

5. Hou Xiwen, Hui Zi, Ding Ruimin, Chen Xiaoyang and Gao Yu, Entanglement and decoherence in a quantum dimer, Chinese Phys., **15**, 2510 (2006).