1 Introduction

1.1 Course Topics

In the past 30 years, random graphs, and more generally, random discrete structures, have become the focus of research of large groups of mathematicians, computer scientists, physicists, and social scientists. All these groups contribute to this area differently: mathematicians try to develop models and study their properties in a formal way while physicists and computer scientists apply those models to study real-life networks and systems and, through simulations, to develop interesting and fruitful intuitions as to how to bring mathematical models closer to reality. The abrupt development of study in the theory and applications of random graphs and networks is in a large part due to the Internet and WWW revolution, which exploded at the end of the twentieth century and, in consequence, the worldwide popularity of different social media such as Facebook or Twitter, just to name the most influential ones.

Our textbook aims to give a gentle introduction to the mathematical foundations of random graphs and to build a platform to understand the nature of real-life networks.

Although the application of probabilistic methods to prove deterministic results in combinatorics, number theory, and in other areas of mathematics have a quite long history, dating back to results of Szele and Erdős in the 1940s, the crucial step was taken by Erdős and Rényi in their seminal paper titled "On the evolution of random graphs" in 1960 (see [43]). They studied the basic properties of a large uniformly chosen random graph and studied how it evolves through the process of adding random edges, one by one. At roughly the same time, we have the important contribution of Gilbert (see [54]) in which he studied binomial random graphs where edges are inserted independently with a fixed probability. The interest in random graphs grew significantly in the mid 1980s ignited by the publication of the book by Bollobás ([21]) and due to the tireless efforts of Paul Erdős, one of the titans of twentieth-century mathematics, who was promoting probabilistic combinatorics, cooperating with mathematicians all over the world and is recognized as a founding father of the whole area. Random graphs, at the beginning of the twenty-first century is recognized as a young but quickly maturing area of mathematics, with strong connections to computer science and physics. Computer science exploits various ways of applying probabilistic concepts in the analysis of algorithms and in the construction of randomized algorithms. A common ground of random graphs and physics is particularly visible in the analysis of phase transition phenomena and in percolation. In the past 20 years, one can observe a veritable tsunami of publications dealing with various models of random graphs introduced to analyze very large real-world networks: WWW linkage, social, neural, communication, information, and transportation networks, as well as a wide range of large-scale systems.

Nowadays, research in random graphs and networks is thriving, and the subject is included in the curriculum of many mathematics and computer science departments across the world. Our textbook should help readers not only gain mathematical knowledge about the basic results of the theory of random graphs but also allow them to better understand how to model and explore real-world networks.

1.2 Course Outline

The text is divided into three parts and presents the basic elements of the theory of random graphs and networks.

To help the reader navigate through the text and to be comfortable understanding proofs, we have decided to start with describing in the preliminary part (see Chapter 2) three of the main technical tools used throughout the text. Since, in general, we look at the typical properties of large, in terms of the number n of vertices (nodes) of random graphs, in the first section of Chapter 2, we show how to deal with often complicated expressions of their numerical characteristics (random variables), in terms of their rate of growth or decline as $n \to \infty$. We next turn our attention to bounds and asymptotic approximations for factorials and binomials, frequent ingredients of the mathematical expressions found in the book. Finally, we finish this introductory, purely technical, chapter with basic information about the probabilistic tools needed to study tail bounds, i.e., probabilities that a random variable exceeds (or is smaller than) some real value. In this context, we introduce and discuss the Markov, Chebyshev, and Chernoff–Hoeffding inequalities leaving the introduction of other, more advanced, probabilistic tools to the following chapters, where they are applied for the first time.

Part II of the text is devoted to the classic Erdős–Rényi–Gilbert uniform and binomial random graphs. In Chapter 3, we formally introduce these models and discuss their relationships. We also define and study the basic features of the asymptotic behavior of random graphs, i.e., the existence of thresholds for monotone properties.

In Chapter 4, we turn our attention to the process known as the evolution of a random graph, exploring how its typical component structure evolves as the number of the edges increases one by one. We describe this process in three phases: the subcritical phase where a random graph is sparse and is a collection of small tree components and components with exactly one cycle; the phase transition, where the giant component, of order comparable with the order of random graphs, emerges; the super-critical phase, where the giant component "absorbs" smaller ones, and a random graph becomes closer and closer to the moment when it gets fully connected.

Vertex degrees, one of the most important features of random graphs, are studied in Chapter 5 in two cases: when a random graph is sparse and when it is dense. We study not only the expected values of the number of vertices of a given degree but also their asymptotic distributions, as well as applications to the notoriously difficult problem of graph isomorphism.

Chapter 6 studies the connectivity and k-connectivity of a random graph, while Chapter 7 discusses the existence in a random graph of a fixed small subgraph, whose size (the number of vertices) does not depend on the size of the random graph itself, and studies the asymptotic distribution of the number of such subgraphs in a random graph.

Large subgraphs are considered in Chapter 8. Here, the thresholds for the existence of a perfect matching are established, first for a bipartite random graph, and next, for a general random graph. These results are proved using the well-known graph theory theorems of Hall and a weakening of the corresponding theorem of Tutte, respectively. After this, long paths and cycles in sparse random graphs are studied and the proof of the celebrated result discovering the threshold for the existence of the Hamilton cycle in a random graph is given. The chapter closes with a short section on the existence of isomorphic copies of certain spanning subgraphs of random graphs.

The last chapter, in Part II, Chapter 9, is devoted to the extremes of certain graph parameters. We look first at the diameter of random graphs, i.e., the extreme value of the shortest distance between a pair of vertices. Next, we look at the size of the largest independent set and the related value of the chromatic number of a random graph.

Part III concentrates on generalizations of the Erdős–Rényi–Gilbert models of random graphs whose features better reflect some characteristic properties of real-world networks such as edge dependence, global sparseness and local clustering, small diameter, and scale-free distribution of the number of vertices of a given degree. In the first section of Chapter 10, we consider a generalization of the binomial random graph where edge probabilities, although still independent, are different for each pair of endpoints, and study conditions for its connectedness. Next, a special case of a generalized binomial random graph is introduced, where the edge probability is a function of weights of the endpoints. This is known in the literature as the Chung–Lu model. Section 12.1 provides information about the volume and uniqueness of the giant component and the sizes of other components, with respect to the expected degree sequence. The final section of Chapter 10 introduces a tool, called the configuration model, to generate a close approximation to a random graph with a fixed degree sequence. Although promoted by Bollobás, this class of random graphs is often called the Molloy–Reed model.

In Chapter 11, the "small-world" phenomenon is discussed. This name bears the observation that large real-world networks are connected by relatively short paths although being globally sparse, in the sense that the number of edges is a bounded multiple of the number of vertices, their nodes/vertices. There are two random graph models presented in this chapter: the first due to Watts and Strogatz and the second due to Kleinberg illustrate this property. In particular, finding short paths in the Kleinberg model is amenable to a particularly simple algorithm.

In general, real-world networks have a dynamic character in terms of the continual addition/deletion of vertices/edges and so we are inclined to model them via random graph processes. This is the topic of Chapter 12. There we study the properties of a wide class of preferential attachment models, which share with real networks the property that their degree sequence exhibits a tail that decays polynomially (power law), as opposed to classical random graphs, whose tails decay exponentially. We give a detailed analysis and formal description of the so-called Barabási–Albert model, as well its generalization: spatial preferential attachment.

Chapter 13 introduces the reader to the binomial and geometric random intersection graphs. Those random graphs are very useful in modeling communities with similar preferences and communication systems.

Finally, Chapter 14 is devoted to a different aspect of graph randomness. Namely, we start with a graph and equip its edges with random weights. In this chapter, we consider three of the most basic combinatorial optimization problems, namely minimum-weight spanning trees, shortest paths, and minimum weight matchings in bipartite graphs.

Suggestions for Instructors and Self-Study

The textbook material is designed for a one-semester undergraduate/graduate course for mathematics and computer science students. The course might also be recommended for students of physics, interested in networks and the evolution of large systems as well as engineering students, specializing in telecommunication. The book is almost self-contained, there being few prerequisites, although a background in elementary graph theory and probability will be helpful.

We suggest that instructors start with Chapter 2 and spend the first week with students becoming familiar with the basic rules of asymptotic computation, finding leading terms in combinatorial expressions, choosing suitable bounds for the binomials, etc., as well as probabilistic tools for tail bounds.

The core of the course is Part II, which is devoted to studying the basic properties of the classical Erdős–Rényi–Gilbert uniform and binomial random graphs. We estimate that it will take between 8 and 10 weeks to cover the material from Part II. Our suggestion for the second part of the course is to start with inhomogeneous random graphs (Chapter 10), which covers the Chung–Lu and Molloy–Reed models, continue with the "small world" (Chapter 11), and conclude with Section 12.1, i.e., the basic preferential attachment model. Any remaining time may be spent either on one of the two sections on random intersection graphs (Chapter 13), especially for those interested in social or information networks, or selected sections of Chapter 14, especially for those interested in combinatorial optimization.

To help students develop their skills in asymptotic computations, as well as to give a better understanding of the covered topics, each section of the book is concluded with simple exercises mainly of a computational nature. We ask the reader to answer rather simple questions related to the material presented in a given section. Quite often however, in particular in the sections covering more advanced topics, we just ask the reader to verify equations, developed through complicated computations, where intermediate steps have been deliberately omitted. Finally, each chapter ends with an extensive set of problems of a different scale of complication, where more challenging problems are accompanied by hints or references to the literature.

Suggestions for Further Readings

A list of possible references for books in the classical theory of random graphs is rather short. There are two advanced books in this area, the first one by Béla Bollobás [21] and the second by Svante Janson, Tomasz Łuczak, and Andrzej Ruciński [66]. Both books give a panorama of the most important problems and methods of the theory of random graphs. Since the current book is, in large part, a slimmed-down version of our earlier book [52], we encourage the reader to consult it for natural extensions of several of the topics we discuss here. Someone taking the course based on our textbook may find it helpful to refer to a very nice and friendly introduction to the theoretical aspects of random networks, in the book by Fan Chung and Linyuan Lu [32]. One may also find interesting books by Remco van der Hofstad [60] and Rick Durrett [41], which give a deep probabilistic perspective on random graphs and networks.

We may also point to an extensive literature on random networks which studies their properties via simulations, simplified heuristic analysis, experiments, and testing. Although, in general, those studies lack mathematical accuracy, they can give good intuitions and insight that help understand the nature of real-life networks. From the publications in this lively area, we would like to recommend to our reader the very extensive coverage of its problems and results presented by Mark Newman [95].

Last but not least, we suggest, in particular to someone for whom random graphs will become a favorite area of further, deeper study, reading the original paper [43] by Paul Erdős and Alfred Rényi on the evolution of random graphs, the seed from which the whole area of random graphs grew to what it is today.