Preface

The internet affects many aspects of our lives, such as how we store and retrieve information, conduct business, and communicate. For example, information is no longer only stored in printed form, but is represented on-line via a complex set of interconnected web pages. The *web graph* has vertices representing web pages, with edges corresponding to the links between pages. The web graph is a real-world network which has undergone intensive study in the last decade by theoreticians and experimentalists. Does this graph have interesting properties? Are there good, rigorous mathematical models for these properties? Can we exploit the graph structure of the web to help search it for information? The answer to all three questions is, of course, yes!

The study of the web graph, or *internet mathematics* as it is now often called, is an active field of study. As the subject is new, there is often a lack of consensus on the central topics, models, even notation, with key questions not always evident. As the subject is fast-breaking, a large arsenal of techniques are required to model and analyze properties of the web. However, possessing the right mathematical tools and a familiarity with current research developments is an important first step. This book should supply a solid mathematical introduction to internet mathematics, and will encourage interest in an emerging and fascinating area of graph theory and theoretical computer science.

The book resulted from lecture notes for an Atlantic Association for Research in the Mathematical Sciences (AARMS) Summer School graduate course *Massive Networks and Internet Mathematics* taught in July 2006 at Dalhousie University in Halifax. A version of the course was taught twice before at Wilfrid Laurier University in Waterloo. As such, the book is appropriate for graduate students or keen undergraduate students in mathematics, computer science, engineering, or physics, whose background includes elementary graph theory, linear algebra, and probability theory. The text is also useful to professional mathematicians, scientists, or engineers interested in learning more about the web graph and graph theory in general. We emphasize that our view is clearly on the mathematics surrounding the web graph. Further, the topics covered are by no means exhaustive.

The book is largely self-contained, and references are given where proofs are omitted. There are over 100 exercises at the end of the chapters and many worked examples, all making the book suitable for either a course or for self-study. Open problems are stated in the exercises and elsewhere.

The book consists of seven chapters. Chapter 1 supplies the requisite background and notation in graph theory and discrete probability used throughout the remaining chapters. We describe the graph and probability theory as well as notation that acts as the foundation for the remaining chapters. The web graph and its key properties are introduced in Chapter 2. Here the reader will learn, among other things, about power law degree distributions and the small world property. Various real-world, self-organizing networks, ranging from technological, biological, to social, are discussed in this chapter. In Chapter 3, an introduction is given to techniques and properties of the classical $G(n, p)$ random graph. Random graph theory supplies the backbone for much of internet mathematics; the techniques used here will be used in later chapters. Chapter 4 surveys the mathematics of stochastic web graph models. Several models are reviewed and analyzed for their degree sequence and other parameters. The topic of searching the web is presented in Chapter 5, where the key web ranking algorithm PageRank—used by the search engine Google—is described. The chapter includes a discussion of the linear algebra and Markov chains used in modern web ranking algorithms. In Chapter 6, we describe the interaction between infinite graph theory and web graph models. The view of massive real-world networks as infinite graphs is relatively new, and it ties in well with the existing theory on the infinite random graph. There are myriad facets to research on the web graph, so as a result we finish in Chapter 7 with three distinct topics on web graph research: spectra of power law graphs, modelling viruses on the web, and domination in web graph models.

How to read this book? The key chapters for a course in internet mathematics are Chapters 1 to 5, inclusive. Chapters 6 and 7 may be completely or partially omitted in a one-semester course. The topics in those chapters are well suited for reading projects. All chapters contain exercises (some with references), and so the book is well suited for assignments and self-study. A web page will be maintained for the book at

`http://info.wlu.ca/~wwwmath/faculty/bonato/webgraph.html`

which will contain useful links and additional information such as corrections or addenda.

As with any undertaking of this nature, there are many people to acknowledge and thank. A sincere thank you goes out to Jon Borwein, David Langstroth, Ron Fitzgerald, Ina Mette, Natalya Pluzhnikov, and everyone at the AMS and AARMS for their generous support of this project. I would like to thank all the mathematicians I have discussed internet mathematics and graph theory with over the years; in particular, I would like to thank Cathy Baker, Kathie Cameron, Peter Cameron, Colin Cooper, Dejan Delić, Fan Chung Graham, Geňa Hahn, Jon Kleinberg, Jarik Nešetřil, Richard Nowakowski, and Joel Spencer. I especially wish to acknowledge Jeannette Janssen, whose infectious enthusiasm and brilliance attracted me to and kept me fascinated with the subject. Thanks to Dejan Delić, Douglas Hamlyn, Jeannette Janssen, Pawel Prałat, Laleh Samarbakhsh, Changping Wang, and the anonymous referees for their careful reading of early drafts of the book. Without the constant love and support of my family—Doug, Anna Maria, Paul, Lisa—this book would not have been written. I dedicate the book to the memory of my sister Paula.