Preface

In the five years since the first edition of *Data Networks* appeared, the networking field has changed in a number of important ways. Perhaps the most fundamental has been the rapid development of optical fiber technology. This has created almost limitless opportunities for new digital networks of greatly enhanced capabilities. In the near term, the link capacities available for data networks, both wide area and local, are increasing by many orders of magnitude. In the longer term, public broadband integrated service networks that provide integrated data, voice, and video on a universal scale are now technologically feasible. These networks of the future appear at first to have almost nothing in common with the data networks of the last 20 years, but in fact, many of the underlying principles are the same. This edition is designed both to provide a fundamental understanding of these common principles and to provide insight into some of the new principles that are evolving for future networks.

Our approach to helping the reader understand the basic principles of networking is to provide a balance between the description of existing networks and the development of analytical tools. The descriptive material is used to illustrate the underlying concepts, and the analytical material is used to generate a deeper and more precise understanding of the concepts. Although the analytical material can be used to analyze the performance of various networks, we believe that its more important use is in sharpening one's conceptual and intuitive understanding of the field; that is, analysis should precede design rather than follow it.

The book is designed to be used at a number of levels, varying from a senior undergraduate elective, to a first year graduate course, to a more advanced graduate course, to a reference work for designers and researchers in the field. The material has been tested in a number of graduate courses at M.I.T. and in a number of short courses at

varying levels. The book assumes some background in elementary probability and some background in either electrical engineering or computer science, but aside from this, the material is self-contained.

Throughout the book, major concepts and principles are first explained in a simple non-mathematical way. This is followed by careful descriptions of modelling issues and then by mathematical analysis. Finally, the insights to be gained from the analysis are explained and examples are given to clarify the more subtle issues. Figures are liberally used throughout to illustrate the ideas. For lower-level courses, the analysis can be glossed over; this allows the beginning and intermediate-level student to grasp the basic ideas, while enabling the more advanced student to acquire deeper understanding and the ability to do research in the field.

Chapter 1 provides a broad introduction to the subject and also develops the layering concept. This layering allows the various issues of data networks to be developed in a largely independent fashion, thus making it possible to read the subsequent chapters in any desired depth (including omission) without seriously hindering the ability to understand other chapters.

Chapter 2 covers the two lowest layers of the above layering and also discusses a number of closely related aspects of the higher layers. The treatment of the lowest, or physical, layer provides a brief overview of how binary digits are transmitted over physical communication media. The effort here is to provide just enough material so that the student can relate the abstraction of digital transmission to physical phenomena. The next layer, data link control, allows packets to be transmitted reliably over communication links. This provides an introduction into the distributed algorithms, or protocols, that must be used at the ends of the link to provide the desired reliability. These protocols are less important in modern high speed networks than in older networks, but the concepts are used repeatedly at many layers in all kinds of networks. The remainder of the chapter focuses on other point to point protocols that allow the end points of a link, or the end points of a network session, to cooperate in providing some required service.

Chapter 3 develops the queueing theory used for performance analysis of multiaccess schemes (Chapter 4) and, to a lesser extent, routing algorithms (Chapter 5). Less analytical courses will probably omit most of this chapter, simply adopting the results on faith. Little's theorem and the Poisson process should be covered however, since they are simple and greatly enhance understanding of the subsequent chapters. This chapter is rich in results, often developed in a far simpler way than found in the queueing literature. This simplicity is achieved by considering only steady-state behavior and by sometimes sacrificing rigor for clarity and insight. Mathematically sophisticated readers will be able to supply the extra details for rigor by themselves, while for most readers the extra details would obscure the line of argument.

Chapter 4 develops the topic of multiaccess communication, including local area networks, metropolitan area networks, satellite networks, and radio networks. Less theoretical courses will probably skip the last half of section 4.2, all of section 4.3, and most of section 4.4, getting quickly to local area networks in section 4.5. Conceptually, one gains a great deal of insight into the nature of distributed algorithms in this chapter.

Preface

Chapter 5 develops the subject of routing. The material is graduated in order of increasing difficulty and depth, so readers can go as far as they are comfortable. Along with routing itself, which is treated in greater depth than elsewhere in the literature, further insights are gained into distributed algorithms. There is also a treatment of topological design and a section on recovery from link failures.

Chapter 6 deals with flow control (or congestion control as it is sometimes called). The first three sections are primarily descriptive, describing first the objectives and the problems in achieving these objectives, and then two general approaches, window flow control, and rate control. The fourth section describes the ways that flow control is handled in several existing networks. The last section is more advanced and analytical, describing various algorithms to select session rates in rate control schemes.

A topic that is not treated in any depth in the book is that of higher-layer protocols, namely the various processes required in the computers and devices using the network to communicate meaningfully with each other given the capability of reliable transport of packets through the network provided by the lower layers. This topic is different in nature than the other topics covered and would have doubled the size of the book if treated in depth.

We apologize in advance for the amount of jargon and acronyms in the book. We felt it was necessary to include at least the most commonly used acronyms in the field, both to allow readers to converse with other workers in the field and also for the reference value of being able to find out what these acronyms mean.

An extensive set of problems are given at the end of each chapter except the first. They range from simple exercises to gain familiarity with the basic concepts and techniques to advanced problems extending the results in the text. Solutions of the problems are given in a manual available to instructors from Prentice-Hall.

Each chapter also contains a brief section of sources and suggestions for further reading. Again, we apologize in advance to the many authors whose contributions have not been mentioned. The literature in the data network field is vast, and we limited ourselves to references that we found most useful, or that contain material supplementing the text.

The stimulating teaching and research environment at M.I.T. has been an ideal setting for the development of this book. In particular we are indebted to the many students who have used this material in courses. Their comments have helped greatly in clarifying the topics. We are equally indebted to the many colleagues and advanced graduate students who have provided detailed critiques of the various chapters. Special thanks go to our colleague Pierre Humblet whose advice, knowledge, and deep insight have been invaluable. In addition, Erdal Arikan, David Castanon, Robert Cooper, Tony Ephremides, Eli Gafni, Marianne Gardner, Inder Gopal, Paul Green, Ellen Hahne, Bruce Hajek, Michael Hluchyi, Robert Kennedy, John Spinelli, and John Tsitsiklis have all been very helpful. We are also grateful to Nancy Young for typing the many revisions. Our editors at Prentice-Hall have also been very helpful and cooperative in producing the final text under a very tight schedule. Finally we wish to acknowledge the research sup-

port of DARPA under grant ONR-N00014-84-K-0357, NSF under grants ECS-8310698, ECS-8217668, 8802991-NCR, and DDM-8903385, and ARO under grants DAAG 29-84-K-000 and DAAL03-86-K-0171.

Dimitri Bertsekas

Robert Gallager