Networks, Crowds, and Markets

Over the past decade there has been a growing public fascination with the complex connectedness of modern society. This connectedness is found in many incarnations: in the rapid growth of the Internet, in the ease with which global communication takes place, and in the ability of news and information as well as epidemics and financial crises to spread with surprising speed and intensity. These are phenomena that involve networks, incentives, and the aggregate behavior of groups of people; they are based on the links that connect us and the ways in which our decisions can have subtle consequences for others.

This introductory undergraduate textbook takes an interdisciplinary look at economics, sociology, computing and information science, and applied mathematics to understand networks and behavior. It describes the emerging field of study that is growing at the interface of these areas, addressing fundamental questions about how the social, economic, and technological worlds are connected.

David Easley is the Henry Scarborough Professor of Social Science and the Donald C. Opatrny '74 Chair of the Department of Economics at Cornell University. He was previously an Overseas Fellow of Churchill College, Cambridge. His research is in the fields of economics, finance, and decision theory. In economics, he focuses on learning, wealth dynamics, and natural selection in markets. In finance, his work focuses on market microstructure and asset pricing. In decision theory, he works on modeling decision making in complex environments. He is a Fellow of the Econometric Society and is Chair of the NASDAQ-OMX Economic Advisory Board.

Jon Kleinberg is the Tisch University Professor in the Computer Science Department at Cornell University. He is a member of the National Academy of Engineering and the American Academy of Arts and Sciences. His research focuses on issues at the interface of networks and information, with an emphasis on the social and information networks that underpin the Web and other online media. He is the recipient of MacArthur, Packard, and Sloan Foundation Fellowships; the Nevanlinna Prize; the ACM-Infosys Foundation Award; and the National Academy of Sciences Award for Initiatives in Research.

Networks, Crowds, and Markets

Reasoning about a Highly Connected World

David Easley

Cornell University

Jon Kleinberg

Cornell University

CAMBRIDGE
UNIVERSITY PRESS

CAMBRIDGE UNIVERSITY PRESS
Cambridge, New York, Melbourne, Madrid, Cape Town, Singapore,
São Paulo, Delhi, Dubai, Tokyo, Mexico City

Cambridge University Press
32 Avenue of the Americas, New York, NY 10013-2473, USA

www.cambridge.org
Information on this title: www.cambridge.org/9780521195331

First published 2010

Printed in the United States of America

A catalog record for this publication is available from the British Library.

Library of Congress Cataloging in Publication data

Easley, David.
Networks, crowds, and markets : reasoning about a highly connected world / David Easley, Jon Kleinberg.
p. cm.
Includes bibliographical references and index.
ISBN 978-0-521-19533-1 (hardback)
1. Telecommunication – Social aspects. 2. Information society. I. Kleinberg, Jon. II. Title.
HM851.E24 2010
303.48′33–dc22 2009050314

ISBN 978-0-521-19533-1 Hardback

Contents

Part III Markets and Strategic Interaction in Networks

Part IV Information Networks and the World Wide Web

Preface

Over the past decade, there has been a growing public fascination with the complex "connectedness" of modern society. This connectedness is found in many incarnations: in the rapid growth of the Internet and the Web, in the ease with which global communication now takes place, and in the ability of news and information as well as epidemics and financial crises to spread around the world with surprising speed and intensity. These are phenomena that involve networks, incentives, and the aggregate behavior of groups of people; they are based on the links that connect us and the ways in which each of our decisions can have subtle consequences for the outcomes of everyone else.

Motivated by these developments in the world, there has been a coming-together of multiple scientific disciplines in an effort to understand how highly connected systems operate. Each discipline has contributed techniques and perspectives that are characteristically its own, and the resulting research effort exhibits an intriguing blend of these different flavors. From computer science and applied mathematics has come a framework for reasoning about how complexity arises, often unexpectedly, in systems that we design; from economics has come a perspective on how people's behavior is affected by incentives and by their expectations about the behavior of others; and from sociology and the social sciences have come insights into the characteristic structures and interactions that arise within groups and populations. The resulting synthesis of ideas suggests the beginnings of a new area of study, focusing on the phenomena that take place within complex social, economic, and technological systems.

This book grew out of a course that we developed at Cornell, designed to introduce this topic and its underlying ideas to a broad student audience at an introductory level. The central concepts are fundamental and accessible ones, but they are dispersed across the research literatures of the many different fields contributing to the topic. The principal goal of this book is therefore to bring the essential ideas together in a single unified treatment and to present them in a way that requires as little background knowledge as possible.

Overview. The book is intended to be used at the introductory undergraduate level, and as such it has no formal prerequisites beyond a level of comfort with basic mathematical definitions at a precalculus level. In keeping with the introductory style, many of the ideas are developed in special cases and through illustrative examples; our goal is to take concepts and theories that are complex in their full generality and to provide simpler formulations where the essential ideas still come through.

In our use of the book, we find that many students are also interested in pursuing some of these topics more deeply, and so it is useful to provide pathways that lead from the introductory formulations into the more advanced literature on these topics. With this in mind, we provide optional sections labeled *Advanced Material* at the ends of most chapters. These advanced sections are qualitatively different from the other sections in the book; some draw on more advanced mathematics, and their presentation is at a more challenging level of conceptual complexity. Aside from the additional mathematical background required, however, even these advanced sections are self-contained; they are also strictly optional, in the sense that nothing elsewhere in the book depends on them.

Synopsis. The first chapter of the book provides a detailed description of the topics and issues that we cover. Here we give a briefer summary of the main focus areas.

The book is organized into seven parts of three to four chapters each. Parts I and II discuss the two main theories that underpin our investigations of networks and behavior: *graph theory*, which studies network structure, and *game theory*, which formulates models of behavior in environments where people's decisions affect each other's outcomes. Part III integrates these lines of thought into an analysis of the network structure of markets and the notion of power in such networks. Part IV pursues a different integration, discussing the World Wide Web as an information network, the problem of Web search, and the development of the markets that currently lie at the heart of the search industry. Parts V and VI study the dynamics of some of the fundamental processes that take place within networks and groups, including the ways in which people are influenced by the decisions of others. Part V pursues this topic at an aggregate scale, where we model interactions between an individual and the population as a whole. Part VI continues the analysis at the more fine-grained level of network structure, beginning with the question of influence and moving on to the dynamics of search processes and epidemics. Finally, Part VII considers how we can interpret fundamental social institutions – including markets, voting systems, and property rights – as mechanisms for productively shaping some of the phenomena we've been studying.

Use of the Book. The book is designed for teaching as well as for any reader who finds these topics interesting and wants to pursue them independently at a deeper level.

Several different types of courses can be taught from this book. When we teach from it at Cornell, the students in our class come from many different majors and have a wide variety of technical backgrounds; this diversity in the audience has served as our primary calibration in setting the introductory level of the book. Our course includes a portion of the material from each chapter; for the sake of concreteness, we provide the approximate weekly schedule we follow below. (There are three 50-minute lectures

each week, except that weeks 6 and 7 of our course contain only two lectures each. In each lecture, we don't necessarily include all the details from each indicated section.)

Week 1: Chapters 1; 2.1–2.3; 3.1–3.3, 3.5; 4.1
Week 2: Chapters 5.1–5.3; 6.1–6.4; 6.5–6.9
Week 3: Chapters 8.1–8.2; 9.1–9.6; 10.1–10.2
Week 4: Chapters 10.3; 10.4–10.5; 11.1–11.2
Week 5: Chapters 11.3–11.4; 12.1–12.3; 12.5–12.6
Week 6: Chapters 12.7–12.8; 13
Week 7: Chapter 14.1–14.2; 14.3–14.4
Week 8: Chapter 15.1–15.2; 15.3–15.4; 15.5–15.6, 15.8
Week 9: Chapter 16.1–16.2; 16.3–16.4; 16.5–16.7
Week 10: Chapters 17.1–17.2; 17.3–17.5; 18
Week 11: Chapters 19.1–19.2; 19.3; 19.4, 19.6
Week 12: Chapter 22.1–22.4; 22.5–22.9; 7.1–7.4
Week 13: Chapters 20.1–20.2; 20.3–20.6; 21.1–21.5
Week 14: Chapters 23.1–23.5; 23.6–23.9; 24

There are many other paths that a course could follow through the book. First, a number of new courses are being developed at the interface of computer science and economics, focusing particularly on the role of economic reasoning in the design and behavior of modern computing systems. The book can be used for such courses in several ways, building on four chapters as a foundation: Chapter 2 on graphs, Chapter 6 on games, Chapter 9 on auctions, and Chapter 10 on matching markets. From here, a more expansive version of such a course could cover the remainder of Parts II and III, all of Parts IV and V, Chapter 19, and portions of Part VII. A more focused and potentially shorter version of such a course concerned principally with auctions, markets, and the online applications of these ideas could be constructed from Chapters 2, 6, 9, 10, 13, 15, 17, 18, and 22, and drawing on parts of Chapters 11, 12, 14, 16, and 19. When these courses are taught at a more advanced level, the advanced sections at the ends of most of these chapters would be appropriate material; depending on the exact level of the course, the text of many of these chapters could be used to lead into the more advanced analysis in their respective final sections.

In a different but related direction, new courses are also being developed on the topic of social computing and information networks. The book can be used for courses of this type by emphasizing Chapters 2–6, 13, 14, 17–20, and 22; many such courses also include sponsored search markets as part of their coverage of the Web, which can be done by including Chapters 9, 10, and 15 as well. The advanced sections in the book can play a role here too, depending on the level of the course.

Finally, portions of the book can serve as self-contained modules in courses on broader topics. To pick just a few examples, one can assemble such modules on network algorithms (Sections 2.3, 3.6, 5.5, 8.3, 10.6, 14.2, 14.3, 14.6, 15.9, 20.3, 20.4, and 20.7); applications of game theory (Chapters 6–9 and 11; Sections 12.9, 15.3–15.6, 19.2, 19.3, 19.5–19.7, and 23.7–23.9); social network analysis (Chapters 2–5; Sections 12.1–12.3 and 12.5–12.8; and Chapters 18–20); the role of information in economic settings (Chapters 16 and 22, and Sections 23.6–23.10); and the analysis of large-scale network data sets (Sections 2.3, 3.2, 3.3, 3.6, 4.4, 5.3, 13.3, 13.4, 14.2–14.5, 18.2, 18.5, and 20.5). Most of these modules use graphs and/or games as fundamental

building blocks; for students not already familiar with these topics, Chapters 2 and 6, respectively, provide self-contained introductions.

Acknowledgments. Our work on this book took place in an environment at Cornell that was particularly conducive to interaction between computing and the social sciences. Our collaboration began as part of a project with Larry Blume, Eric Friedman, Joe Halpern, Dan Huttenlocher, and Éva Tardos funded by the National Science Foundation, followed by a campus-wide "theme project" on networks sponsored by Cornell's Institute for the Social Sciences, with a group that included Larry and Dan together with John Abowd, Geri Gay, Michael Macy, Kathleen O'Connor, Jeff Prince, and David Strang. Our approach to the material in the book draws on perspectives – ways of thinking about these topics and ways of talking about them – that we've learned and acquired from this interdisciplinary set of colleagues, a group that includes some of our closest professional collaborators.

The course on which the book is based grew out of discussions that were part of the Cornell theme project; the two of us had taught distinct portions of this material separately in graduate courses that we had developed, and Michael Kearns's *Networked Life* course at University of Pennsylvania demonstrated the vibrancy and relevance this material could have for an introductory undergraduate audience as well. We were intrigued by the prospect of combining different perspectives that hadn't previously appeared together – a process that would be educational not only to the students in the course but to us as well. Creating and teaching this new interdisciplinary course was made possible by the support of our departments, Computer Science and Economics, and by support from the Solomon Fund at Cornell University.

Once the book had begun to take shape, we benefited enormously from the feedback, suggestions, and experiences of colleagues who taught from early drafts of it. In particular, we thank Daron Acemoglu (MIT), Lada Adamic (Michigan), Allan Borodin (Toronto), Noshir Contractor (Northwestern), Jason Hartline (Northwestern), Nicole Immorlica (Northwestern), Ramesh Johari (Stanford), Samir Khuller (Maryland), Jure Leskovec (Stanford), David Liben-Nowell (Carleton), Peter Monge (USC), Asu Ozdaglar (MIT), Vijay Ramachandran (Colgate), R. Ravi (CMU), Chuck Severance (Michigan), Aravind Srinivasan (Maryland), and Luis von Ahn (CMU). The graduate and undergraduate teaching assistants in our own teaching of this subject have been very helpful as well; we thank Alex Ainslie, Lars Backstrom, Jacob Bank, Vlad Barash, Burak Bekdemir, Anand Bhaskar, Ben Cole, Bistra Dilkina, Eduard Dogaru, Ram Dubey, Ethan Feldman, Ken Ferguson, Narie Foster, Eric Frackleton, Christie Gibson, Vaibhav Goel, Scott Grabnic, Jon Guarino, Fahad Karim, Koralai Kirabaeva, Tian Liang, Austin Lin, Fang Liu, Max Mihm, Sameer Nurmohamed, Ben Pu, Tal Rusak, Mark Sandler, Stuart Tettemer, Ozgur Yonter, Chong-Suk Yoon, and Yisong Yue.

In addition to the instructors who used early drafts, a number of other people provided extensive comments on portions of the book, leading to many improvements in the text: Lada Adamic, Robert Kerr, Evie Kleinberg, Gueorgi Kossinets, Stephen Morris, David Parkes, Rahul Sami, Andrew Tomkins, and Johan Ugander. We also thank a further set of colleagues, in addition to those already listed, who have provided very useful advice

and suggestions on this project as it has proceeded: Bobby Kleinberg, Gene Kleinberg, Lillian Lee, Maureen O'Hara, Prabhakar Raghavan, and Steve Strogatz.

It has been a pleasure to be able to work with the editorial team at Cambridge University Press. Lauren Cowles, our main point of contact at Cambridge, has been an amazing source of advice and help, and we likewise very much appreciate the contributions of Scott Parris and David Tranah to this project, and Peggy Rote and her colleagues at Aptara for their work on the production of the book.

Finally, a profound thanks goes to our families, in continuing appreciation of their support and many other contributions.

David Easley
Jon Kleinberg
Ithaca, 2010