

## Exploratory Social Network Analysis with Pajek

This is an extensively revised and expanded third edition of the successful textbook on analysis and visualization of social networks integrating theory, applications, and professional software for performing network analysis. The main structural concepts and their applications in social research are introduced with exercises. Pajek software and data sets are available, so readers can learn network analysis through application and case studies. In the end, readers will have the knowledge, skills, and tools to apply social network analysis across different disciplines. A fundamental redesign of the menu structure and the capability to analyze much larger networks required a new edition. This edition presents several new operations, e.g., community detection, generalized main paths searches, new network indices, advanced visualization approaches, and instructions for installing Pajek under MacOSX. This third edition is up-to-date with Pajek version 5 and it introduces PajekXXL for very large networks and Pajek3XL for huge networks.

Wouter de Nooy is Associate Professor in the Department of Communication Science at the University of Amsterdam, the Netherlands. He is a member of the Amsterdam School of Communication Research. His research interests include dynamic models for social networks and their application in culture and politics.

Andrej Mrvar is Professor of Social Science Informatics at the Faculty of Social Sciences, University of Ljubljana, Slovenia. He started developing the programs *Pajek*, *PajekXXL*, and *Pajek3XL* as part of his PhD dissertation at the Faculty of Computer and Information Science in 1996. Since then Pajek has been in constant development.

Vladimir Batagelj is Professor Emeritus of the University of Ljubljana, Slovenia. He is also a member of the Institute of Mathematics, Physics, and Mechanics (IMFM), Ljubljana, and of AMI, UP, Koper. His coauthored book *Generalized Blockmodeling* was awarded the 2007 Harrison White Outstanding Book Award by the Mathematical Sociology Section of the American Sociological Association. From the International Network for Social Network Analysis he was awarded the Georg Simmel Award (2007) and the Richards Software Award for the program *Pajek* (2013). In 2014 he coauthored a book entitled *Understanding Large Temporal Networks and Spatial Networks*.

Cambridge University Press  
978-1-108-47414-6 — Exploratory Social Network Analysis with Pajek  
3rd Edition  
Frontmatter  
[More Information](#)

---

## Structural Analysis in the Social Sciences

Series Editor: Mark Granovetter

The series *Structural Analysis in the Social Sciences* presents studies that analyze social behavior and institutions by reference to relations among such concrete social entities as persons, organizations, and nations. Relational analysis contrasts on the one hand with reductionist methodological individualism and on the other with macro-level determinism, whether based on technology, material conditions, economic conflict, adaptive evolution, or functional imperatives. In this more intellectually flexible structural middle ground, analysts situate actors and their relations in a variety of contexts. Since the series began in 1987, its authors have variously focused on small groups, history, culture, politics, kinship, aesthetics, economics, and complex organizations, creatively theorizing how these shape and in turn are shaped by social relations. Their style and methods have ranged widely, from intense, long-term ethnographic observation to highly abstract mathematical models. Their disciplinary affiliations have included history, anthropology, sociology, political science, business, economics, mathematics, and computer science. Some have made explicit use of social network analysis, including many of the cutting-edge and standard works of that approach, whereas others have kept formal analysis in the background and used “networks” as a fruitful orienting metaphor. All have in common a sophisticated and revealing approach that forcefully illuminates our complex social world.

### Recent Books in the Series

- Wouter De Nooy, Andrej Mrvar and Vladimir Batagelj, *Exploratory Social Network Analysis with Pajek: Revised and Expanded Edition for Updated Software. Third Edition*
- Darius Mehri, *Iran Auto*
- Navid Hassanpour, *Leading from the Periphery and Network Collective Action*
- Cheol-Sung Lee, *When Solidarity Works* (Second Edition)
- Benjamin Cornwell, *Social Sequence Analysis* (Second Edition)
- Mariela Szwarcberg, *Mobilizing Poor Voters* (Second Edition)
- Luke M. Gerdes, ed., *Illuminating Dark Networks* (Second Edition)
- Silvia Domínguez and Betina Hollstein, eds., *Mixed Methods in Studying Social Networks* (Second Edition)
- Dean Lusher, Johan Koskinen, and Garry Robins, eds., *Exponential Random Graph Models for Social Networks: Theory, Methods, and Applications* (Second Edition)
- Sean F. Everton, *Disrupting Dark Networks* (Second Edition)
- Wouter de Nooy, Andrej Mrvar, and Vladimir Batagelj, *Exploratory Social Network Analysis with Pajek* (First Edition)
- Noah E. Friedkin and Eugene C. Johnsen, *Social Influence Network Theory* (Second Edition)
- Zeev Maoz, *The Networks of Nations: The Evolution and Structure of International Networks, 1815–2002* (Second Edition)

(continued after the index)

Cambridge University Press  
978-1-108-47414-6 — Exploratory Social Network Analysis with Pajek  
3rd Edition  
Frontmatter  
[More Information](#)

---

# Exploratory Social Network Analysis with Pajek

*Revised and Expanded Edition for Updated  
Software. Third Edition*

WOUTER DE NOOY

*University of Amsterdam*

ANDREJ MRVAR

*University of Ljubljana*

VLADIMIR BATAGELJ

*University of Ljubljana*



**CAMBRIDGE**  
UNIVERSITY PRESS

**CAMBRIDGE**  
UNIVERSITY PRESS

University Printing House, Cambridge CB2 8BS, United Kingdom  
One Liberty Plaza, 20th Floor, New York, NY 10006, USA  
477 Williamstown Road, Port Melbourne, VIC 3207, Australia  
314–321, 3rd Floor, Plot 3, Splendor Forum, Jasola District Centre,  
New Delhi - 110025, India  
79 Anson Road, #06-04/06, Singapore 079906

Cambridge University Press is part of the University of Cambridge.

It furthers the University's mission by disseminating knowledge in the pursuit of education, learning, and research at the highest international levels of excellence.

[www.cambridge.org](http://www.cambridge.org)

Information on this title: [www.cambridge.org/9781108474146](http://www.cambridge.org/9781108474146)

DOI: 10.1017/9781108565691

© Wouter de Nooy, Andrej Mrvar, and Vladimir Batagelj 2018

This publication is in copyright. Subject to statutory exception and to the provisions of relevant collective licensing agreements, no reproduction of any part may take place without the written permission of Cambridge University Press.

First published 2018

Printed in the United States of America by Sheridan Books, Inc.

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

ISBN 978-1-108-47414-6 Hardback

ISBN 978-1-108-46227-3 Paperback

Cambridge University Press has no responsibility for the persistence or accuracy of URLs for external or third-party internet websites referred to in this publication and does not guarantee that any content on such websites is, or will remain, accurate or appropriate.

## Contents

<i>List of Figures</i>	page xiii
<i>List of Tables</i>	xix
<i>Preface to the Third Edition</i>	xxi
<i>Preface to the Second Edition</i>	xxiii
<i>Preface to the First Edition</i>	xxv
Overview	xxvi
Justification	xxviii
Acknowledgments	xxix

### PART I FUNDAMENTALS

1 Looking for Social Structure	3
1.1 Introduction	3
1.2 Sociometry and Sociogram	3
1.3 Exploratory Social Network Analysis	5
1.3.1 Network Definition	6
1.3.2 Manipulation	12
1.3.3 Calculation	15
1.3.4 Visualization	17
1.4 Assembling a Social Network	27
1.5 Summary	30
1.6 Questions	31
1.7 Assignment	32
1.8 Further Reading	32
1.9 Answers	33
2 Attributes and Relations	36
2.1 Introduction	36
2.2 Example: The World System	36
2.3 Partitions	38

viii	<i>Contents</i>	
	2.4	Reduction of a Network 45
	2.4.1	Local View 46
	2.4.2	Global View 48
	2.4.3	Contextual View 51
	2.5	Vectors and Coordinates 53
	2.6	Network Analysis and Statistics 61
	2.7	Summary 63
	2.8	Questions 64
	2.9	Assignment 65
	2.10	Further Reading 65
	2.11	Answers 66
	PART II COHESION	
	3	Cohesive Subgroups 73
	3.1	Introduction 73
	3.2	Example 73
	3.3	Density and Degree 75
	3.4	Components 79
	3.5	Cores 83
	3.6	Cliques and Complete Subnetworks 86
	3.7	Summary 92
	3.8	Questions 94
	3.9	Assignment 96
	3.10	Further Reading 96
	3.11	Answers 96
	4	Sentiments and Friendship 99
	4.1	Introduction 99
	4.2	Balance Theory 99
	4.3	Example 103
	4.4	Detecting Structural Balance and Clusterability 103
	4.5	Development in Time 109
	4.6	Summary 113
	4.7	Questions 113
	4.8	Assignment 115
	4.9	Further Reading 115
	4.10	Answers 116
	5	Affiliations 119
	5.1	Introduction 119
	5.2	Example 120
	5.3	Two-Mode and One-Mode Networks 121
	5.4	Islands 127



*Contents*

ix

5.5	Communities	132
5.6	The Third Dimension	135
5.7	Summary	139
5.8	Questions	140
5.9	Assignment	141
5.10	Further Reading	141
5.11	Answers	142
PART III BROKERAGE		
6	Center and Periphery	149
6.1	Introduction	149
6.2	Example	149
6.3	Distance	151
6.4	Betweenness	158
6.5	Eigenvector Centrality	160
6.6	Assortativity	162
6.7	Summary	164
6.8	Questions	165
6.9	Assignment	166
6.10	Further Reading	167
6.11	Answers	167
7	Brokers and Bridges	170
7.1	Introduction	170
7.2	Example	171
7.3	Bridges and Bi-Components	172
7.4	Ego-Networks and Constraint	177
7.5	Affiliations and Brokerage Roles	184
7.6	Summary	189
7.7	Questions	190
7.8	Assignment	191
7.9	Further Reading	193
7.10	Answers	194
8	Diffusion	197
8.1	Example	197
8.2	Contagion	200
8.3	Exposure and Thresholds	204
8.4	Critical Mass	211
8.5	Summary	216
8.6	Questions	217
8.7	Assignment	219
8.8	Further Reading	219
8.9	Answers	220

## PART IV RANKING

9	Prestige	225
9.1	Introduction	225
9.2	Example	226
9.3	Popularity and Indegree	227
9.4	Correlation	229
9.5	Domains	231
9.6	Proximity Prestige	235
9.7	Summary	238
9.8	Questions	238
9.9	Assignment	240
9.10	Further Reading	241
9.11	Answers	241
10	Ranking	244
10.1	Introduction	244
10.2	Example	245
10.3	Triadic Analysis	245
10.4	Acyclic Networks	253
10.5	Symmetric-Acyclic Decomposition	256
10.6	Summary	261
10.7	Questions	263
10.8	Assignment	265
10.9	Further Reading	265
10.10	Answers	266
11	Genealogies and Citations	269
11.1	Introduction	269
11.2	Example I: Genealogy of the Ragusan Nobility	269
11.3	Family Trees	270
11.4	Social Research on Genealogies	278
11.5	Example II: Citations among Papers on Network Centrality	289
11.6	Citations	291
11.7	Summary	304
11.8	Questions	304
11.9	Assignment 1	306
11.10	Assignment 2	306
11.11	Further Reading	306
11.12	Answers	307

## PART V MODELING

12	Blockmodels	315
12.1	Introduction	315

*Contents*

xi

12.2	Matrices and Permutation	316
12.3	Roles and Positions: Equivalence	322
12.4	Blockmodeling	331
	12.4.1 Blockmodel	332
	12.4.2 Blockmodeling	333
	12.4.3 Regular Equivalence	338
12.5	Summary	343
12.6	Questions	345
12.7	Assignment	347
12.8	Further Reading	348
12.9	Answers	348
13	Random Graph Models	353
13.1	Introduction	353
13.2	Example	355
13.3	Modeling Overall Network Structure	357
	13.3.1 Classic Uniform Models	358
	13.3.2 Small-World Models	362
	13.3.3 Preferential Attachment Models	366
13.4	Monte Carlo Simulation	373
13.5	Summary	377
13.6	Questions	379
13.7	Assignment	381
13.8	Further Reading	381
13.9	Answers	383
Appendix 1	Getting Started with Pajek	387
A1.1	Installation	387
A1.2	Network Data Formats	387
A1.3	Creating Network Files for Pajek	389
	A1.3.1 Within Pajek	389
	A1.3.2 Helper Software	391
	A1.3.3 Word Processor	392
	A1.3.4 Relational Database	394
A1.4	Limitations	400
A1.5	PajekXXL and Pajek3XL	400
A1.6	Updates of Pajek	402
Appendix 2	Exporting Visualizations	404
A2.1	Export Formats	404
	A2.1.1 Bitmap and JPEG	404
	A2.1.2 Encapsulated PostScript	405
	A2.1.3 Scalable Vector Graphics	406
	A2.1.4 VOSviewer	408
	A2.1.5 Virtual Reality Modeling Language and X3D	409

xii	<i>Contents</i>	
	A2.1.6 MDL MOL and Kinemages	410
A2.2	Layout Options	411
	A2.2.1 Top Frame on the Left: EPS/SVG Vertex Default	412
	A2.2.2 Bottom Frame on the Left: EPS/SVG Line Default	416
	A2.2.3 Top Frame on the Right	418
	A2.2.4 Middle Frame on the Right – Background Colors	419
	A2.2.5 Bottom Frame on the Right – EPS Border	420
Appendix 3	Installing Pajek on Mac OS X	421
Appendix 4	Shortcut Key Combinations	424
	A4.1 Main Screen	424
	A4.2 Hierarchy Edit Screen	425
	A4.3 Draw Screen	425
	<i>Glossary</i>	427
	<i>Index of Pajek and R Commands</i>	439
	<i>Subject Index</i>	445

## Figures

1	Dependencies between the chapters (for the second and third editions)	<i>page</i> xxvii
2	Sociogram of dining-table partners	4
3	Partial listing of a multiple relations network data file for Pajek	10
4	Pajek Main screen	11
5	Menu structure in Pajek	13
6	Dialog box in Pajek	14
7	Report screen in Pajek	16
8	Dialog box of <i>Network &gt; Info &gt; General</i> command	17
9	Draw screen in Pajek	19
10	Continue dialog box	21
11	A selected option in the Draw screen	23
12	<i>Options</i> menu of the Draw screen	24
13	Textual output from <i>[Draw]Info &gt; All Properties</i>	24
14	A 3-D rendering of the dining-table partners network	26
15	Empty network	29
16	Edit Network screen	29
17	World trade of manufactures of metal and world system position	40
18	Edit screen with partition according to world system position	41
19	Vertex colors according to a partition in Pajek	43
20	Trade ties within South America	46
21	The <i>Partitions</i> menu	47
22	World system positions in South America: (2) semiperiphery and (3) periphery	48
23	Trade in manufactures of metal among continents (imports in thousands of US dollars)	49

24	Trade among continents in the Draw screen	51
25	Contextual view of trade in South America	52
26	Geographical view of world trade in manufactures of metal, ca. 1994	54
27	<i>Vector&gt;Info</i> dialog box	55
28	Trade, position in the world system, and GDP per capita	58
29	Aggregate trade in manufactures of metal among world system positions	67
30	Contextual view of North American trade ties and (mean) GDP per capita	68
31	Visiting ties in Attiro	74
32	A simple unconnected directed network	79
33	Strong components (contours) and family–friendship groupings (vertex colors and numbers) in the network of Attiro	82
34	<i>k</i> -cores in the visiting network at Attiro	84
35	<i>k</i> -cores	85
36	Stacking or nesting of <i>k</i> -cores	85
37	The complete triad and an example	87
38	A hierarchy of cliques	89
39	Viewing a hierarchy in an Edit screen	90
40	Complete triads and family–friendship groupings (colors and numbers inside vertices)	91
41	Decision tree for the analysis of cohesive subgroups	93
42	A Person–Other–Object (X) triple	100
43	P–O–X triple as a signed digraph	100
44	A balanced network	102
45	First positive and negative choices between novices at <i>T4</i>	104
46	Output listing of a <i>Doreian–Mrvar Method</i> * command	108
47	Three solutions with one error	109
48	Partial listing of Sampson.net	110
49	Differences between two solutions with four classes	116
50	A fragment of the Scottish directorates network	122
51	One-mode network of firms created from the network in Figure 50	123
52	One-mode network of directors derived from Figure 50	124
53	Islands in the network of Scottish firms, 1904–1905 (contours added manually)	128
54	The islands in the network of Scottish firms (1904–1905) with industrial categories (class numbers) and capital (vertex size)	130

*List of Figures*

xv

55	Islands in three dimensions	136
56	Coordinate system of Pajek	136
57	A landscape of islands in the Scottish firms network	138
58	Communication ties within a sawmill	150
59	Star-network and line-network	152
60	Distances to or from Juan (vertex colors: Default GreyScale 1)	156
61	Geodesics between HP-1 and EM-4	157
62	Betweenness centrality in the sawmill	160
63	Communication network of striking employees	171
64	Cut-vertices (gray) and bi-components (manually circled) in the strike network	174
65	Hierarchy of bi-components and bridges in the strike network	177
66	Three connected triads	178
67	Alejandro's ego-network	179
68	Proportional strength of ties around Alejandro	181
69	Constraints on Alejandro	181
70	Energized constraint network	183
71	Five brokerage roles of actor $v$	185
72	Bob's ego-network	186
73	Constraint inside groups	188
74	Two overlapping cliques	192
75	Friendship ties among superintendents and year of adoption	199
76	Adoption of the modern math method: diffusion curve	201
77	Diffusion by contacts in a random network ( $N = 100$ ; vertex numbers indicate the distance from the source vertex)	201
78	Diffusion from a central and a marginal vertex	202
79	Adoption (vertex color) and exposure (in brackets) at the end of 1959	205
80	Modern math network with arcs pointing toward later adopters	209
81	Visiting ties and prestige leaders in San Juan Sur	226
82	<i>Partitions</i> menu in Pajek	231
83	Distances to family 47 (represented by the numbers within the vertices)	233
84	Proximity prestige in a small network	237
85	Student government discussion network	245
86	An example of a network with ranks	246
87	Triad types with their sequential numbers in Pajek	247
88	Strong components in the student government discussion network	255

xvi	<i>List of Figures</i>	
	89 Acyclic network with shrunk components	255
	90 Clusters of symmetric ties in the student government network	257
	91 Discussion network shrunk according to symmetric clusters	257
	92 Symmetric components in the (modified) student government discussion network	258
	93 The order of symmetric clusters according to the depth partition (acyclic)	260
	94 Ranks in the student government discussion network	261
	95 Three generations of descendants to Petrus Gondola (years of birth)	271
	96 Ore graph	272
	97 Descendants of Petrus Gondola and Ana Goce	274
	98 Shortest paths between Paucho and Margarita Gondola	275
	99 Structural relinking in an Ore graph	280
	100 P-graph	281
	101 Structural relinking in a P-graph	282
	102 Fragment of relinking grandchildren	285
	103 Centrality literature network in layers according to year of publication	290
	104 $k$ -cores in the centrality literature network (without isolates)	292
	105 Traversal weights in a citation network	293
	106 Forward local and key-route local (top), standard global and key-route global (middle), and backward local (bottom) main paths in the centrality literature network	299
	107 Main path component of the centrality literature network (not all names are shown here)	303
	108 Communication lines among striking employees	316
	109 The matrix of the strike network sorted by ethnic and age groups	318
	110 A network and a permutation	319
	111 Partial listing of the strike network as a binary matrix	320
	112 The strike network permuted according to ethnic and age groups	321
	113 Part of the permuted strike network displayed as a binary network	322
	114 Hypothetical ties among two instructors (i) and three students (s)	322



*List of Figures*

xvii

115	A dendrogram of similarities	324
116	Imports of miscellaneous manufactures of metal and world system position in 1980	325
117	Hierarchical clustering of the world trade network	328
118	Hierarchical clustering of countries in the Hierarchy Edit screen	329
119	An ideal core-periphery structure	331
120	Image matrix and shrunk network	332
121	Error in the imperfect core-periphery matrix	334
122	<i>Optimize Partition</i> dialog box	336
123	Output of the <i>Optimize Partition</i> procedure	336
124	<i>Random Start</i> dialog box	338
125	Matrix of the student government network	339
126	Image matrix and error matrix for the student government network	340
127	Assembling a blockmodel in Pajek	342
128	Random versions of a small friendship network	354
129	Political blogosphere, United States, February 8, 2005	356
130	Small-world random graph generation: Ring of local lines (left) and rewired lines (right)	363
131	Log-log degree distributions of the blogs network, absolute frequencies (left) and cumulative proportions (right)	369
132	<i>Read Network</i> dialog box	388
133	A network in Pajek matrix format	389
134	Editing vertex labels	390
135	Edit Network screen	390
136	An empty network in Pajek Arcs/Edges format	392
137	A network in the Pajek Arcs/Edges format	393
138	A network in the Pajek matrix format	393
139	A two-mode network in the Pajek Arcs/Edges format	394
140	Four tables in the world trade database (MS Access 2010)	395
141	Contents of the <i>Countries</i> table (partial)	395
142	A Lookup to the <i>Countries</i> table	396
143	Export a report to plain text	397
144	Tables and relations in the database of Scottish companies	399
145	The <i>Options</i> screen	411
146	Layout of a vertex and its label	412
147	Bezier curves and different vertex shapes (man, woman, and house)	414

xviii	<i>List of Figures</i>	
148	The $x/y$ ratio of a vertex	414
149	Visualization with Unicode symbols	416
150	The position and orientation of a line label	417
151	Gradients in SVG export: linear (left) and radial (right)	419
152	XQuartz webpage	422
153	WineHQ webpage	422
154	Pajek webpage	423
155	Pajek running on Mac OS X	423

## Tables

1	Tabular output of the command <i>Partition&gt;Info</i>	page 41
2	Distribution of GNP per capita in classes	56
3	Output of the <i>Info</i> command	62
4	Cross-tabulation of world system positions (rows) and GDP per capita (columns)	69
5	Frequency distribution of degree in the symmetrized network of visits	78
6	Error score with all choices at different moments ( $\alpha = .5$ )	112
7	Error score with first choices only ( $\alpha = .5$ )	117
8	Line multiplicity in the one-mode network of firms	126
9	Frequency tabulation of coordinator roles in the strike network	188
10	Adoption in the modern math network	203
11	Adoption rate and acceleration in the modern math diffusion curve	212
12	Fragment of Table 11	215
13	Indegree listing in Pajek	228
14	Input domain of f47	234
15	Size of input domains in the visiting relations network	235
16	Balance-theoretic models	250
17	Triad census of the example network	251
18	Triad census of the student government network	252
19	Number of children of Petrus Gondola and his male descendants	277
20	Size of sibling groups in 1200–1250 and 1300–1350	279
21	Birth cohorts among men and women	287
22	Traversal weights in the centrality literature network	297
23	Dissimilarity scores in the example network	324

xx	<i>List of Tables</i>	
24	Cross-tabulation of initial (rows) and optimal partition (columns)	337
25	Final image matrix of the world trade network	338
26	Monte Carlo simulation results: confidence intervals for the simple undirected blogs network	374
27	Names of colors in Pajek	413

## *Preface to the Third Edition*

Two major developments in program Pajek required a new edition of this book: a fundamental redesign of the menu structure and the capability to analyze much larger networks, containing billions of vertices. A proliferation of methods for analyzing a single network necessitated a reorganization of Pajek's menu structure, in particular the former *Net* menu, now called the *Network* menu. The new *Network* menu contains submenus for commands that apply to a particular type of network: a two-mode network, multiple relations network, acyclic network, temporal network, and signed network. It is much easier now to find analyses for special networks. Because the *Network* menu is used most intensively, we had to adjust most of the commands in the Application sections of this book.

Pajek's capability to analyze much larger networks is the second major development. As a result of changes in the Windows<sup>®</sup> operating system, Pajek can now handle networks with nearly one billion vertices. For even larger networks, PajekXXL and Pajek3XL have been developed, which can handle up to two and ten billion vertices respectively. PajekXXL and Pajek3XL have the same user interface as Pajek; if you can work with Pajek, you can also work with PajekXXL and Pajek3XL. What you need to know is that the latter two programs offer a very limited set of analyses to describe and partition huge networks. Subnetworks extracted from a huge network can be sent directly to (regular) Pajek for further analysis. Appendix A1.5 explains all of this.

Questions on how to install and use Pajek on Mac OS X have been asked repeatedly. This new edition brings an additional appendix (Appendix 3) containing detailed instructions on installing Pajek under Mac OS X. We hope that Mac users will find out that installing and running Pajek on Mac OS X is not a difficult technical problem: Installation takes only few additional minutes compared to installation under native Windows. When Pajek is installed, running Pajek under Mac OS X is the same as running it under Windows.

Finally, we took the opportunity to include some analyses requested by Pajek users. Chapter 1 now includes Pivot MDS and VOS mapping for network layout as well as the correlation between layout coordinates and network geodesics as a measure of layout performance. Chapter 2 introduces partitions on vertex labels using regular expressions and marking partition clusters with Unicode symbols in the Draw screen. It also discusses interactive FishEye magnification and the Adjusted Rand Index. Relaxed balance is explained in Chapter 4; community detection (Louvain Method and VOS Clustering) and the E-I Index appear in Chapter 5. Chapter 6 includes (degree) assortativity and the assortativity coefficient. A collection of main path methods (including key-route searches) and preprint transformation are explained and applied in Chapter 11. Appendices A1 and A2 have been updated, now containing goodies such as dragging and dropping data to a Pajek window, sending Pajek objects to Excel<sup>®</sup>, defining colors and transparency of vertices and lines, using Unicode symbols and additional vertex shapes (e.g., man, woman, and house), tooltips for vertex labels, drawing curved lines, and so on. We hope that you will continue enjoying social network analysis with Pajek.

The webpage to the third edition of this book (<http://mrvar.fdv.uni-lj.si/pajek/be3.htm>, mirror <http://mrvar2.fdv.uni-lj.si/pajek/be3.htm>) contains the example data sets, helper programs, and other online documents referenced in this book.

## *Preface to the Second Edition*

I go with him out in a shed in back and see he is selling a whole Harley machine in used parts, except for the frame, which the customer already has. He is selling them all for \$125. Not a bad price at all.

Coming back I comment, “He’ll know something about motorcycles before he gets *those* together.”

Bill laughs. “And that’s the best way to learn, too.”

Robert M. Pirsig, *Zen and the Art of Motorcycle Maintenance*

To some of its readers, this book is an introduction to social network analysis; to other readers, it is a manual to Pajek software (<http://mrvar.fdv.uni-lj.si/pajek>). To us, it is both. As Patrick Doreian argued in his review of our book (In: *Social Networks* 28 [2006] 269–274), an understanding of social network analysis is required for proper use of Pajek and, vice versa, understanding the concepts and logic of Pajek fosters comprehension of network concepts. In this second edition, we have aimed to strengthen both aspects, updating the discussion of the Pajek interface and commands to include several capabilities that have been implemented since we submitted the text of the first edition, such as multiplex networks (Section 1.3.1), eigenvector centrality (Section 6.5), matrix multiplication (Section 11.3), and using Pajek output in R (Chapters 5 and 13). The new capabilities cover some important advances in social network analysis, including random graph models to which we have dedicated a new chapter.

We expanded the Further Reading sections with references to seminal, much cited texts. This should allow the reader to trace the literature on the selected topic in bibliographic and citation databases. For more comprehensive lists of literature, we refer to two other volumes in this series: S. Wasserman and K. Faust, *Social Network Analysis: Methods and Applications* (Cambridge: Cambridge University Press, 1994) and P. J. Carrington, J. Scott, and S. Wasserman, *Models and Methods in Social Network*

xxiv      *Preface to the Second Edition*

*Analysis* (Cambridge: Cambridge University Press, 2005). A concise history of social network analysis is published by L. C. Freeman, *The Development of Social Network Analysis. A Study in the Sociology of Science* (Vancouver, Canada: Empirical Press, 2004).

We hope that this second edition will continue to stimulate analysts to sharpen their understanding of social networks and expand their command of network analytic tools.



## *Preface to the First Edition*

In the social sciences, social network analysis has become a powerful methodological tool alongside statistics. Network concepts have been defined, tested, and applied in research traditions throughout the social sciences, ranging from anthropology and sociology to business administration and history.

This book is the first textbook on social network analysis integrating theory, applications, and professional software for performing network analysis. It introduces structural concepts and their applications in social research with exercises to improve skills, questions to test the understanding, and case studies to practice network analysis. In the end, the reader has the knowledge, skills, and tools to apply social network analysis.

We stress learning by doing: Readers acquire a feel for network concepts by applying network analysis. To this end, we make ample use of professional computer software for network analysis and visualization: Pajek. This software, operating under Windows 95 and later, and all example data sets are provided on a Web site (<http://vlado.fmf.uni-lj.si/pub/networks/book/>) dedicated to this book. All the commands that are needed to produce the graphical and numerical results presented in this book are extensively discussed and illustrated. Step by step, the reader can perform the analyses presented in the book.

Note, however, that the graphical display on a computer screen will never exactly match the printed figures in this book. After all, a book is not a computer screen. Furthermore, newer versions of the software will appear, with features that may differ from the descriptions presented in this book. We strongly advise using the version of Pajek software supplied on the book's Web site (<http://vlado.fmf.uni-lj.si/pub/networks/book/>) while studying this book and then updating to a newer version of Pajek afterward, which can be downloaded from <http://vlado.fmf.uni-lj.si/pub/networks/pajek/default.htm>.

## Overview

This book contains five sections. The first section (Part I) presents the basic concepts of social network analysis. The next three sections present the three major research topics in social network analysis: cohesion (Part II), brokerage (Part III), and ranking (Part IV). We claim that all major applications of social network analysis in the social sciences relate to one or more of these three topics. The final section discusses an advanced technique (viz., blockmodeling), which integrates the three research topics (Part V).

The first section, titled Fundamentals, introduces the concept of a network, which is obviously the basic object of network analysis, and the concepts of a partition and a vector, which contain additional information on the network or store the results of analyses. In addition, this section helps the reader get started with Pajek software.

Part II on cohesion consists of three chapters, each of which presents measures of cohesion in a particular type of network: ordinary networks (Chapter 3), signed networks (Chapter 4), and valued networks (Chapter 5). Networks may contain different types of relations. The ordinary network just shows whether there is a tie between people, organizations, or countries. In contrast, signed networks are primarily used for storing relations that are either positive or negative such as affective relations: liking and disliking. Valued networks take into account the strength of ties, for example, the total value of the trade from one country to another or the number of directors shared by two companies.

Part III on brokerage focuses on social relations as channels of exchange. Certain positions within the network are heavily involved in the exchange and flow of information, goods, or services; whereas others are not. This is connected to the concepts of centrality and centralization (Chapter 6) or brokers and bridges (Chapter 7). Chapter 8 discusses an important application of these ideas, namely, the analysis of diffusion processes.

The direction of ties (e.g., who initiates the tie) is not very important in the section on brokerage, but it is central to ranking, presented in Part IV. Social ranking, it is assumed, is connected to asymmetric relations. In the case of positive relations, such as friendship nominations or advice seeking, people who receive many choices and reciprocate few choices are deemed as enjoying more prestige (Chapter 9). Patterns of asymmetric choices may reveal the stratification of a group or society into a hierarchy of layers (Chapter 10). Chapter 11 presents a particular type of asymmetry, namely, the asymmetry in social relations caused by time: genealogical descent and citation.

*Preface to the First Edition*

xxvii

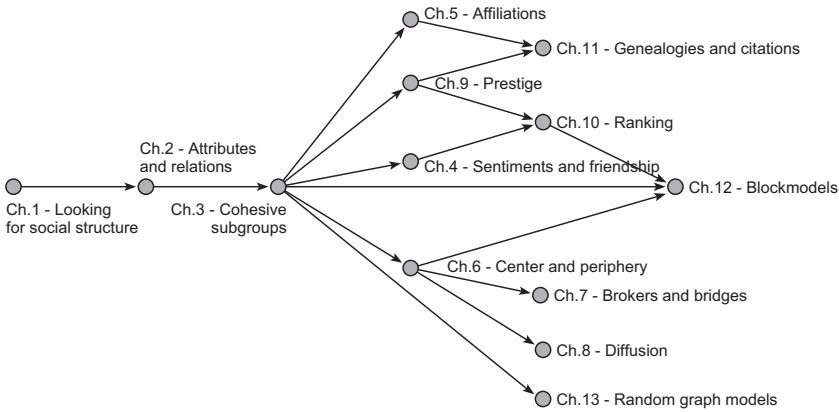


Figure 1. Dependencies between the chapters (for the second and third editions).

The final section, Part V, on roles concentrates on rather dense and small networks. This type of network can be visualized and stored efficiently by means of matrices. Blockmodeling is a suitable technique for analyzing cohesion, brokerage, and ranking in dense, small networks. It focuses on positions and social roles (Chapter 12).

The book is intended for researchers and managers who want to apply social network analysis and for courses on social network analysis in all social sciences as well as other disciplines using social methodology (e.g., history and business administration). Regardless of the context in which the book is used, Chapters 1, 2, and 3 must be studied to understand the topics of subsequent chapters and the logic of Pajek. Chapters 4 and 5 may be skipped if the researcher or student is not interested in networks with signed or valued relations, but we strongly advise including them to be familiar with these types of networks. In Parts III (Brokerage) and IV (Ranking), the first two chapters present basic concepts and the third chapter focuses on particular applications.

Figure 1 shows the dependencies among the chapters of this book. To study a particular chapter, all preceding chapters in this flowchart must have been studied before. Chapter 10, for instance, requires understanding of Chapters 1 through 4 and 9. Within the chapters, there are no sections that can be skipped.

In an undergraduate course, Part I and II should be included. A choice can be made between Part III and Part IV; or, alternatively, just the first chapter from each section may be selected. Part V on social roles and blockmodeling is quite advanced and more appropriate for a

postgraduate course. For managerial purposes, Part III is probably more interesting than Part IV.

### Justification

This book offers an introduction to social network analysis, which implies that it covers a limited set of topics and techniques, which we feel a beginner must master to be able to find his or her way in the field of social network analysis. We have made many decisions about what to include and what to exclude, and we want to justify our choices now.

As reflected in the title of this book, we restrict ourselves to exploratory social network analysis. The testing of hypotheses by means of statistical models falls outside the scope of this book. In social network analysis, hypothesis testing is important but complicated; it deserves a book on its own. Aiming our book at people who are new to social network analysis, our first priority is to have them explore the structure of social networks to give them a feel for the concepts and applications of network analysis. Exploration involves visualization and manipulation of concrete networks, whereas hypothesis testing boils down to numbers representing abstract parameters and probabilities. In our view, exploration yields the intuitive understanding of networks and basic network concepts that are a prerequisite for well-considered hypothesis testing.

From the vast array of network analytic techniques and indices we discuss only a few. We have no intention of presenting a survey of all structural techniques and indices because we fear that the readers will not be able to see the forest for the trees. We focus on as few techniques and indices as are needed to present and measure the underlying concept. With respect to the concept of cohesion, for instance, many structural indices have been proposed for identifying cohesive groups:  $n$ -cliques,  $n$ -clans,  $n$ -clubs,  $m$ -cores,  $k$ -cores,  $k$ -plexes, lambda sets, and so on. We discuss only components,  $k$ -cores, 3-cliques, and  $m$ -slices ( $m$ -cores) because they suffice to explain the basic parameters involved: density, connectivity, and strength of relations within cohesive subgroups.

Our choice is influenced by the software that we use because we have decided to restrict our discussion to indices and techniques that are incorporated in this software. Pajek software is designed to handle very large networks (up to millions of vertices). Therefore, this software package concentrates on efficient routines, which are capable of dealing with large networks. Some analytical techniques and structural indices are known to be inefficient (e.g., the detection of  $n$ -cliques), and for others no efficient algorithm has yet been found or implemented. This limits our options; we present only the detection of small cliques (of size 3), and we cannot extensively discuss an important concept such as  $k$ -connectivity. In

*Preface to the First Edition*

xxix

summary, this book is neither a complete catalogue of network analytic concepts and techniques nor an exhaustive manual to all commands of Pajek. It offers just enough concepts, techniques, and skills to understand and perform all major types of social network analysis.

In contrast to some other handbooks on social network analysis, we minimize mathematical notation and present all definitions verbatim. There are no mathematical formulae in the book. We assume that many students and researchers are interested in the application of social network analysis rather than in its mathematical properties. As a consequence, and this may be very surprising to seasoned network analysts, we do not introduce the matrix as a data format and display format for social networks until the end of the book.

Finally, there is a remark on the terminology used in the book. Social network analysis derives its basic concepts from mathematical graph theory. Unfortunately, different “vocabularies” exist within graph theory, using different concepts to refer to the same phenomena. Traditionally, social network analysts have used the terminology employed by Frank Harary, for example, in his book *Graph Theory* (Reading: Addison-Wesley, 1969). We choose, however, to follow the terminology that prevails in current textbooks on graph theory, for example, R. J. Wilson’s *Introduction to Graph Theory* (Edinburgh: Oliver and Boyd, 1972; published later by Wiley, New York). Thus, we hope to narrow the terminological gap between social network analysis and graph theory. As a result, we speak of a vertex instead of a node or a point, and some of our definitions and concepts differ from those proposed by Frank Harary.

## Acknowledgments

The text of this book has benefited from the comments and suggestions from our students at the University of Ljubljana and the Erasmus University Rotterdam, who were the first to use it. In addition, Michael Frishkopf and his students of musicology at the University of Alberta gave us helpful comments. Mark Granovetter, who welcomed this book to his series, and his colleague Sean Farley Everton have carefully read and commented on the chapters. In many ways, they have helped us make the book more coherent and understandable to the reader. We are also very grateful to an anonymous reviewer, who carefully scrutinized the book and made many valuable suggestions for improvements. Ed Parsons (Cambridge University Press) and Nancy Hulan (TechBooks) helped us through the production process. Finally, we thank the participants of the workshops we conducted at the Sunbelt International Conference on Social Network Analysis in New Orleans (XXII) and Cancun (XXIII) for their encouraging reactions to our manuscript.

Most data sets that are used in this book have been created from sociograms or listings printed in scientific articles and books. Notwithstanding our conviction that reported scientific results should be used and distributed freely, we have tried to trace the authors of these articles and books and ask for their approval. We are grateful to have obtained explicit permission for using and distributing the data sets from them. Authors or their representatives whom we have not reached are invited