

# Contents

## Part I Introduction

<b>1</b>	<b>A First Encounter</b> . . . . .	3
1.1	Introduction to Network Analysis . . . . .	3
1.2	Data . . . . .	5
1.2.1	From Relationship to Graph . . . . .	6
1.2.2	First Probes into the Data . . . . .	8
1.2.3	Measuring Indirect Effects . . . . .	12
1.2.4	Distributions . . . . .	13
1.3	Network Analysis Literacy: A Primer . . . . .	15
1.3.1	Visualizations . . . . .	15
1.4	Approaches to Network Analysis . . . . .	18
1.5	Outlook . . . . .	19
1.6	Recommended Reading . . . . .	20
	References . . . . .	21
<b>2</b>	<b>Graph Theory, Social Network Analysis, and Network Science</b> . . . . .	23
2.1	Introduction . . . . .	23
2.2	The Basis . . . . .	24
2.2.1	Graph Theory . . . . .	24
2.2.2	The Origins of Social Network Analysis in Sociology . . . . .	27
2.2.3	Typical Viewpoints of Social Network Analysis . . . . .	30
2.2.4	Network Science . . . . .	31
2.3	Universal Structures versus Individual Features . . . . .	35
2.3.1	Statistical Physics and Early Complex Network Analysis . . . . .	37
2.3.2	Statistical Physics and Complex Network Analysis . . . . .	38
2.3.3	Complex Network Analysis in Other Disciplines . . . . .	40

- 2.4 Network Analysis Literacy: General Requirements . . . . . 42
  - 2.4.1 Implementations and Verbal Descriptions of  
Network Analytic Measures: A Primer . . . . . 42
  - 2.4.2 Interpreting a Measure’s Value: A Primer . . . . . 43
  - 2.4.3 Interpretation by Trained Domain Experts . . . . . 45
  - 2.4.4 Interpretation by Academic Experts . . . . . 48
  - 2.4.5 The Widespread Use of Scientific Rituals . . . . . 49
  - 2.4.6 The Interpretation of Network Analytic Measures . . . . . 49
- 2.5 Recommended Reading . . . . . 53
- 2.6 Exercise . . . . . 53
- References. . . . . 53
- 3 Definitions . . . . . 57**
  - 3.1 Introduction . . . . . 57
  - 3.2 Mathematical Abbreviations . . . . . 58
  - 3.3 Set Theoretic Terms . . . . . 58
    - 3.3.1 Function . . . . . 60
    - 3.3.2 Partitions and Hierarchical Clustering . . . . . 60
  - 3.4 Mathematical Operators . . . . . 61
  - 3.5 Graph Theoretic Definitions . . . . . 61
    - 3.5.1 Distances in Graphs . . . . . 63
    - 3.5.2 Degrees and Walks in Graphs . . . . . 63
    - 3.5.3 Graph Families . . . . . 65
  - 3.6 Data Structures for Graphs . . . . . 66
    - 3.6.1 Basic Data Structures . . . . . 67
    - 3.6.2 Basic Data Structures for Simple Graphs . . . . . 68
    - 3.6.3 Data Structures and Definitions  
for Directed Graphs . . . . . 71
    - 3.6.4 Weighted Graphs . . . . . 72
    - 3.6.5 Bipartite and Affiliation Networks . . . . . 73
    - 3.6.6 Multiplex Networks . . . . . 74
  - 3.7 Graph File Formats . . . . . 74
    - 3.7.1 Graph Formats for Visualization . . . . . 77
  - 3.8 A Little Bit of Linear Algebra . . . . . 77
    - 3.8.1 Scalar Product . . . . . 77
  - 3.9 Normalization . . . . . 78
    - 3.9.1 Covariance . . . . . 78
    - 3.9.2 Correlation Coefficient . . . . . 79
  - 3.10 Algorithms and Runtime Complexity . . . . . 80
  - 3.11 Plots and Diagrams . . . . . 81
  - 3.12 Distributions . . . . . 82
  - 3.13 A Bit of Statistics . . . . . 82
  - 3.14 Markov Chains . . . . . 83
    - 3.14.1 Properties of Markov Chains . . . . . 85

3.15	Further Reading . . . . .	86
3.16	Exercises. . . . .	86
	References. . . . .	88

**Part II Methods**

<b>4</b>	<b>Classic Network Analytic Measures . . . . .</b>	<b>91</b>
4.1	Introduction . . . . .	91
4.2	Direct Statistics. . . . .	92
4.3	Distance Based Measures . . . . .	93
4.4	Degree Based Measures . . . . .	95
	4.4.1 Degree Distribution . . . . .	95
	4.4.2 Assortativity . . . . .	95
4.5	Mutuality, Transitivity, and the Clustering Coefficient . . . . .	99
	4.5.1 Mutuality or Reciprocity . . . . .	99
	4.5.2 Transitivity . . . . .	100
4.6	Density . . . . .	102
4.7	Summary . . . . .	104
4.8	Further Reading . . . . .	105
4.9	Exercises. . . . .	105
	References. . . . .	107
<b>5</b>	<b>Network Representations of Complex Systems . . . . .</b>	<b>109</b>
5.1	Introduction . . . . .	109
5.2	Why Networks are only Models of Complex Systems . . . . .	109
	5.2.1 Edges as Abstract Representations of Real-World Relationships . . . . .	111
	5.2.2 Types of Network Representations . . . . .	113
5.3	Phases of a Network Analytic Project. . . . .	117
	5.3.1 Trilemma of Complex Network Analysis . . . . .	119
5.4	Defining the Entity of Interest. . . . .	121
	5.4.1 Network Boundary . . . . .	122
	5.4.2 Observing Entities . . . . .	123
	5.4.3 Entity Resolution. . . . .	126
5.5	Relationships and Mathematical Relations . . . . .	127
	5.5.1 Classic Relationships Analyzed in Complex Networks . . . . .	130
5.6	Weighted and Dynamic Graphs . . . . .	131
	5.6.1 Observing and Representing Weighted Relationships . . . . .	131
	5.6.2 Dynamic Networks . . . . .	132
	5.6.3 Transformation into Undirected, Unweighted Networks . . . . .	133
5.7	One-Mode Projections of Bipartite Graphs . . . . .	137
	5.7.1 Classic One-Mode Projections. . . . .	137

- 5.7.2 Show Case: Co-authorship Networks. . . . . 139
- 5.8 An Example: Metabolic Networks . . . . . 141
- 5.9 Summary . . . . . 145
- 5.10 Further Reading . . . . . 145
- 5.11 Exercises. . . . . 146
- References. . . . . 147
- 6 Random Graphs and Network Models . . . . . 149**
  - 6.1 Introduction . . . . . 149
  - 6.2 The Set of All Graphs with the Same Number of Nodes . . . . . 150
    - 6.2.1 The  $G(n,m)$  Random Graph Model . . . . . 152
  - 6.3 The Classic Random Graph Model. . . . . 154
  - 6.4 The Small-World Model: Explaining the Small-World Phenomenon . . . . . 158
    - 6.4.1 The Small-World Model (WS-Model) . . . . . 162
  - 6.5 The Preferential Attachment Model (BA-Model) . . . . . 165
    - 6.5.1 Scale-Freeness . . . . . 166
  - 6.6 When is a Random Graph Model Explanatory? . . . . . 170
  - 6.7 Summary . . . . . 174
  - 6.8 Further Reading . . . . . 175
  - 6.9 Exercises. . . . . 177
  - References. . . . . 179
- 7 Random Graphs as Null Models . . . . . 183**
  - 7.1 Introduction . . . . . 183
  - 7.2 Assessing the Significance of a Structural Feature . . . . . 183
    - 7.2.1 Reciprocity Revisited I . . . . . 184
    - 7.2.2 What is the Best Null Model for Assessing Reciprocity in General? . . . . . 186
    - 7.2.3 Node Similarity and Co-occurrence. . . . . 187
  - 7.3 Fixed and Expected Degree Sequence Models . . . . . 191
    - 7.3.1 Stub or Configuration Method. . . . . 193
    - 7.3.2 Simple Independence Model (SIM)—Approximating the Configuration Model . . . . . 194
    - 7.3.3 Chung-Lu-Model: Expected Degree Sequences . . . . . 196
    - 7.3.4 Fixed Degree Sequence Model . . . . . 197
  - 7.4 The Philosophy behind Identifying Statistically Significant Structural Features. . . . . 198
  - 7.5 History of Assessing the Significance of Real-World Network Structures . . . . . 201
    - 7.5.1 Network Motifs . . . . . 202
    - 7.5.2 The Algorithm. . . . . 203
    - 7.5.3 Biologically Meaningful Motifs. . . . . 206
    - 7.5.4 Choosing the Best Null Model . . . . . 207

7.6	Summary . . . . .	208
7.7	Further Reading . . . . .	209
7.8	Exercises. . . . .	210
	References. . . . .	213
<b>8</b>	<b>Understanding and Designing Network Measures. . . . .</b>	<b>215</b>
8.1	Introduction . . . . .	215
8.2	Beware of verbal Descriptions—Why Mathematical Equations are Necessary . . . . .	216
8.2.1	Reciprocity . . . . .	218
8.3	Profile of a Measure’s Behavior . . . . .	221
8.3.1	Applicability . . . . .	222
8.3.2	Range of the Measure and Extremal Graphs . . . . .	225
8.3.3	Scalability . . . . .	226
8.3.4	Size Independence/Comparability . . . . .	227
8.3.5	Robustness. . . . .	228
8.3.6	Assumptions . . . . .	228
8.4	How to Design a Network Analytic Measure . . . . .	230
8.4.1	Generalizing a Method . . . . .	231
8.4.2	Another Interpretation of the Degree in Weighted Graphs. . . . .	235
8.4.3	Clustering Coefficient for Bipartite Graphs . . . . .	235
8.5	Summary . . . . .	238
8.6	Recommended Reading . . . . .	238
8.7	Exercises. . . . .	239
	References. . . . .	241
<b>9</b>	<b>Centrality Indices . . . . .</b>	<b>243</b>
9.1	Introduction . . . . .	243
9.2	What is a Centrality Index? . . . . .	244
9.3	Classic Centrality Indices . . . . .	246
9.3.1	Degree-Like Centralities . . . . .	246
9.3.2	Closeness-Like Centralities . . . . .	249
9.3.3	Stress and betweenness-Like Centralities. . . . .	250
9.3.4	Correlation between Different Centrality Indices . . . . .	255
9.3.5	Comparing Centrality Values in Different Networks. . . . .	256
9.3.6	The Centralization of a Graph . . . . .	258
9.4	Generalizing Centrality Indices. . . . .	259
9.4.1	Centrality Indices for Networks between Different Groups of Nodes. . . . .	259
9.4.2	Centrality Indices for Directed Networks. . . . .	260
9.4.3	Centrality Indices for Weighted Networks. . . . .	260
9.5	Characterizations of Centrality Indices . . . . .	261
9.5.1	The Graph-Theoretic Perspective. . . . .	261
9.5.2	Network Flow Processes and Centrality Indices . . . . .	264

9.6	Centrality-Based Visualization of Graphs . . . . .	264
9.7	Applications of Centrality Indices. . . . .	265
9.7.1	Centrality Distributions as General Structural Descriptors. . . . .	266
9.7.2	Correlation between Centrality Indices and External Properties . . . . .	268
9.7.3	Centrality Indices as Process-Based Predictors . . . . .	270
9.8	Summary . . . . .	271
9.9	Further Reading . . . . .	271
9.10	Exercises. . . . .	272
	References. . . . .	274

### Part III Literacy

<b>10</b>	<b>Literacy: Data Quality, Entities, and Nodes . . . . .</b>	<b>279</b>
10.1	Introduction . . . . .	279
10.2	Describing a Network Representation Transparently. . . . .	280
10.3	Bad Data. . . . .	282
10.3.1	Bad Data: Protein-Protein Interaction Networks. . . . .	282
10.3.2	Bad Data: BGP Routing Data . . . . .	286
10.3.3	Inferred Transcription Network Data . . . . .	287
10.4	Network Boundary . . . . .	289
10.4.1	When is a Node a Node . . . . .	289
10.5	Sampling Effects. . . . .	293
10.5.1	Dynamic and Time-Thresholded Data . . . . .	296
10.6	Evaluating Sampling Strategies. . . . .	297
10.6.1	Evaluating BGP/Traceroute Data. . . . .	298
10.7	Data Biases. . . . .	299
10.7.1	Data Biases in Protein-Protein Interaction Data. . . . .	299
10.7.2	Data Biases in Surveys . . . . .	300
10.7.3	Estimating the Degree of a Node in a Network. . . . .	302
10.8	Curating Complex Networks. . . . .	304
10.9	Summary . . . . .	306
10.10	Further Reading . . . . .	306
10.11	Exercises. . . . .	306
	References. . . . .	309
<b>11</b>	<b>Literacy: Relationships and Relations . . . . .</b>	<b>313</b>
11.1	Introduction . . . . .	313
11.2	When is an Edge an Edge?. . . . .	314
11.3	Aggregations in Time and Space . . . . .	318
11.3.1	Aggregation in Time . . . . .	318
11.3.2	Aggregation in Space . . . . .	319
11.3.3	Choosing an Appropriate Observation Period . . . . .	320

- 11.4 Weighted Relationships. . . . . 322
  - 11.4.1 Interrelationship with Chosen Method . . . . . 322
  - 11.4.2 Dynamic Weights . . . . . 325
  - 11.4.3 Thresholding . . . . . 326
- 11.5 Proxy Relationships . . . . . 327
  - 11.5.1 Proxies for Sexual Relationship Networks. . . . . 327
  - 11.5.2 Online Social Network Data as Proxies. . . . . 329
  - 11.5.3 With Whom do We Discuss Important Matter. . . . . 330
  - 11.5.4 Co-authorship versus Collaboration. . . . . 332
  - 11.5.5 Interchangeability of Social Relations . . . . . 332
  - 11.5.6 Observational versus Recalled Interactions . . . . . 334
  - 11.5.7 Email Interaction versus Communication  
Networks . . . . . 334
  - 11.5.8 Internet Network Data and Their Proxies. . . . . 336
- 11.6 Relations that don't Lend Themselves to a Network  
Representation . . . . . 338
  - 11.6.1 Information Contained in Relations. . . . . 338
  - 11.6.2 Mathematical Relations without Network  
Processes . . . . . 340
  - 11.6.3 Aggregating Paths into Complex Networks. . . . . 341
  - 11.6.4 Relationships, Network Processes, and Complex  
Networks . . . . . 344
- 11.7 Horizons of Network Processes . . . . . 348
- 11.8 Data Responsibility. . . . . 350
  - 11.8.1 Evaluating Existing Network Data for Re-use. . . . . 351
  - 11.8.2 Data Hygiene, Producer and Consumer Rules. . . . . 353
  - 11.8.3 Producer Rules: Making Data Reusable. . . . . 354
  - 11.8.4 Consumer Rules: Validating Data . . . . . 356
- 11.9 Aim of Analysis (A-Rules). . . . . 357
  - 11.9.1 Publishers' Responsibility . . . . . 357
- 11.10 Summary . . . . . 358
- 11.11 Further Reading . . . . . 359
- References. . . . . 359
- 12 Literacy: When Is a Network Model Explanatory? . . . . . 363**
  - 12.1 Introduction . . . . . 363
  - 12.2 Models of Networks and Processes. . . . . 365
    - 12.2.1 What is a Scientific Model?. . . . . 366
    - 12.2.2 Modelling Processes on Complex Networks . . . . . 370
    - 12.2.3 Evolution of Models . . . . . 371
  - 12.3 Structure, Function, and Behavior of Network Models. . . . . 372
    - 12.3.1 Interpretation of 'Smallness' as a Function . . . . . 373
    - 12.3.2 Properties and Behavior of "Scale-Free"  
Networks . . . . . 377

- 12.4 Explanatory Models . . . . . 381
  - 12.4.1 When Preferential Attachment is not Enough . . . . . 382
  - 12.4.2 Networks with a “Scale-Free” Degree Distribution  
Which are not “Scale-Free” . . . . . 383
  - 12.4.3 The Internet—A “Scale-Free” Network without a  
Hub-Dominated Architecture . . . . . 384
  - 12.4.4 Shrinking Diameters in the Evolution of Complex  
Networks . . . . . 385
  - 12.4.5 Measuring Preferential Attachment . . . . . 385
- 12.5 Summary . . . . . 387
- 12.6 Further Reading . . . . . 389
- References. . . . . 392
- 13 Literacy: Choosing the Best Null Model . . . . . 395**
  - 13.1 Introduction . . . . . 395
  - 13.2 Assessing the Small-World Phenomenon . . . . . 398
    - 13.2.1 Clustering Coefficient in One-Mode Projections  
of Bipartite Graphs . . . . . 399
  - 13.3 The Rich-Club Coefficient . . . . . 401
  - 13.4 Reciprocity Revisited II . . . . . 405
  - 13.5 A New Perspective on One-Mode Projections . . . . . 407
    - 13.5.1 The Simple Independence Model SIM. . . . . 408
    - 13.5.2 An Example: MovieLens. . . . . 410
    - 13.5.3 Discussion of the SIM. . . . . 415
    - 13.5.4 The Fixed Degree Sequence Model FDSM for  
Bipartite Graphs. . . . . 418
  - 13.6 Evaluating Expectation Models by a Gold Standard or  
Ground Truth . . . . . 419
    - 13.6.1 Building the OMP . . . . . 420
    - 13.6.2 Is There a Weighted FDSM?. . . . . 421
  - 13.7 Can the Configuration Model Replace the FDSM?. . . . . 422
  - 13.8 Summary . . . . . 425
  - 13.9 Further Reading . . . . . 426
  - 13.10 Exercises. . . . . 427
  - References. . . . . 428
- 14 Literacy Interpretation . . . . . 431**
  - 14.1 Introduction . . . . . 431
  - 14.2 The Interpretation of Measures in the Context  
of a Complex System . . . . . 432
  - 14.3 Interpretation of Distance-Based Measures . . . . . 435
    - 14.3.1 Robustness Measures Based on Distance. . . . . 435
    - 14.3.2 Comparing Average Distances of Different  
Networks . . . . . 440
    - 14.3.3 Interpretation of Low Average Distances  
in Metabolic Networks . . . . . 441



- 14.4 Centrality Index Literacy . . . . . 443
  - 14.4.1 Borgatti’s Flow Concept . . . . . 444
  - 14.4.2 Interpretation of Classic Centrality Indices . . . . . 445
  - 14.4.3 Air Transportation Networks . . . . . 447
  - 14.4.4 Multiplex Air-Transportation Networks . . . . . 450
  - 14.4.5 Designing Interpretable Centrality Indices . . . . . 454
- 14.5 Explorative Applications of Distance Based Measures . . . . . 455
- 14.6 The Centrality of Moscow in the 12th and 13th Century . . . . . 457
- 14.7 Sexual Contact Networks . . . . . 462
  - 14.7.1 From Data to Network. . . . . 463
  - 14.7.2 The Human Web of Sexual Contacts. . . . . 463
  - 14.7.3 An Assessment of Preferential Attachment  
as a Mechanism for Human Sexual Network  
Formation . . . . . 466
- 14.8 Post-Hoc Analysis. . . . . 467
- 14.9 Verbal Description of Findings. . . . . 469
- 14.10 Summary . . . . . 470
- 14.11 Exercises. . . . . 471
- References. . . . . 472
- 15 Ethics in Network Analysis. . . . . 475**
  - 15.1 Why Ethical Network Analysis Needs Network Analysis  
Literacy. . . . . 475
  - 15.2 The Wegman Report. . . . . 476
    - 15.2.1 Discrediting a Scientist by Co-authorship-Network  
Analysis. . . . . 476
  - 15.3 Who Owns a Relationship?. . . . . 479
  - 15.4 Prediction Based on Network Analysis. . . . . 482
  - 15.5 Summary . . . . . 483
  - References. . . . . 484
- Appendix A: The Structure and Typical Outlets of Network  
Analytic Papers. . . . . 487**
- Appendix B: Glossary. . . . . 493**
- Appendix C: Solutions . . . . . 499**
- Author Index. . . . . 529**
- Subject Index. . . . . 531**



<http://www.springer.com/978-3-7091-0740-9>

Network Analysis Literacy

A Practical Approach to the Analysis of Networks

Zweig, K.A.

2016, XXIII, 535 p. 126 illus., 14 illus. in color.,

Hardcover

ISBN: 978-3-7091-0740-9