

Table of Contents

Preface	1
Chapter 1: Fundamentals of Complex Networks and Gephi	7
Graph applications	8
Collaboration graphs	8
Who-talks-to-whom graphs	9
Information linkages	9
Technological networks	9
Natural-world networks	9
A network graph analysis primer	10
Paths and connectivity	11
Paths	11
Cycles	12
Connectivity	13
Network structure	13
Centrality	14
Components	17
Giant components and clustering	18
Homophily	19
Density	19
Network behaviors	20
Contagion and diffusion	20
Network growth	21
Overviewing Gephi	22
Primary windows	23
Data laboratory	23
Manual entry	24
CSV import	24
Excel import	25
MySQL import	25
Graph file import	25

Graph window	25
Preview window	26
Secondary windows – tabs	27
The filtering tab	27
The statistics tab	27
The layouts tab	28
Essential plugins	28
Clustering – Chinese Whispers	29
Data laboratory	29
Data laboratory helper	30
Exports	30
Sigma.js Exporter	30
Seadragon Web Export	30
Graph Streaming	30
ExportToEarth	31
Generator – the Complex Generators plugin	31
Layout	34
The Multipartite layout	34
The Hiveplot layout	34
The Concentric layout	35
The OpenOrd layout	35
The Circular layout	35
The Layered layout	35
The ARF layout	36
Additional plugins	36
Link Communities – metrics	36
Give color to nodes – tools	36
Summary	37
Chapter 2: A Network Graph Framework	39
A proposed process flow	40
Identifying an idea or topic	40
Determining the final output	42
Identifying the data sources	43
Formatting the data for Gephi	43
Importing data into Gephi	44
Viewing the initial graph layout	45
Selecting a layout	47
Analyzing the graph	48
Modifying the graph	49
Exporting the graph	50
Creating an example graph	51
Identifying the topic	51
Finding the data source	52

Formatting the data for Gephi	52
Importing the data	54
Viewing the initial network	55
Selecting an appropriate layout	56
The Force Atlas layout	56
The Fruchterman-Reingold layout	58
The Radial Axis layout	59
The Yifan hu layout	60
ARF	61
Analyzing the graph	62
Modifying the graph	65
Exporting the graph	68
Summary	68
Chapter 3: Selecting the Layout	69
<hr/>	
Overviewing the layout types	69
Force-based layouts	70
The ARF layout	71
Force Atlas	72
Force Atlas 2	73
Force Atlas 3D	74
The Fruchterman-Reingold algorithm	74
The OpenOrd algorithm	75
The Yifan Hu algorithm	76
The Yifan Hu Proportional layout	76
The Yifan Hu multilevel approach	77
Tree layouts	77
DAG layout	78
Circular layouts	78
The Circular layout	79
The Concentric layout	80
The Dual Circle layout	80
Radial layouts	81
The Hiveplot layout	82
The Radial Axis layout	82
Geographic layouts	83
The Geo layout	83
The Maps of Countries layout	84
Additional layouts	84
The Isometric layout	84
The Multipartite layout	85
The Layered layout	85
Network Splitter 3D	86
Additional layout tools	86

Assessing your graphing needs	87
Actual example – the Miles Davis network	89
Analysis goal	89
Dataset parameters	89
Network density	91
Network behaviors	91
Network display	91
Temporal elements	92
Interactivity	92
Layout strengths and weaknesses	93
Testing layouts	97
Testing the ARF layout	97
The Concentric layout	101
Testing the Radial Axis layout	103
Layout selection criteria	106
Graph aesthetics	107
Working example of graph aesthetics	108
Summary	111
Chapter 4: Network Patterns	113
Contagion and diffusion	114
Contagion	114
The SIR model	117
The SIS model	118
The SIRS model	118
Diffusion	119
Clustering and homophily	122
Clustering	123
Homophily	124
Network growth patterns	125
Using Gephi generators	126
Viewing a contagion network	131
Viewing network diffusion	136
Network clustering	140
Identifying homophily	144
Summary	148
Chapter 5: Working with Filters	149
The filtering theory	150
Primary filtering functions in Gephi	151
Attributes	152
Edges	153
Operator	154
Topology	154

Using simple filters	156
Using the Equal filter	157
Applying the regex function	159
Filtering edges	160
Using the Partition filter	162
Working with the Topology filters	164
Working with complex filters	168
Applying multiple filter conditions	168
Using subfilters	169
Working with Mask and Intersection conditions	175
Working with the UNION operator	178
Summary	180
Chapter 6: Graph Statistics	181
Overview of graph statistics	182
Network measures	182
Diameter	182
Eccentricity	183
Graph density	183
Average path length	183
Connected components	184
Erdos number	184
HITS	184
Edge betweenness	185
Centrality measures	185
Degree centrality (undirected graphs)	186
In-degree centrality (directed graphs)	187
Out-degree centrality (directed graphs)	187
Closeness centrality	187
Eigenvector centrality	187
Betweenness centrality	188
Clustering and neighborhood measures	188
Clustering coefficient	188
Number of triangles	189
Modularity	189
Link Communities	189
Neighborhood overlap and embeddedness	189
Interpreting graph statistics	190
Interpreting network measures	190
Interpreting centrality statistics	193
Degree centrality	193
In-degree centrality	194
Out-degree centrality	194
Closeness centrality	194
Eigenvector centrality	195
Betweenness centrality	195

Interpreting clustering statistics	196
Interpreting clustering coefficients	196
Number of triangles	197
Modularity	197
Link Communities	197
Embeddedness	197
Application of statistical measures	198
Basic statistical applications	198
Network statistics	199
Centrality statistics	201
Clustering statistics	205
Filtering using graph statistics	207
Summary	215
Chapter 7: Segmenting and Partitioning a Graph	217
Partitioning and clustering options	218
The Partition tab	219
The Ranking tab	220
Manual settings	220
Chinese Whispers	221
Markov clustering	221
Partitioning and clustering examples	222
Partitioning	222
Working with the Ranking tab	228
Using color and size options	228
Manual graph segmentation	231
Using the Chinese Whispers plugin	233
Using the Markov Clustering plugin	237
Summary	242
Chapter 8: Dynamic Networks	243
When to use DNA	244
Topology-based DNA	245
Generating a dynamic network	245
Understanding time intervals	246
Working with timelines	248
Preparing and importing data for DNA	249
Implementing and viewing a dynamic network	251
Creating time intervals in an existing project	251
Adding time intervals to a new project	254
Working with timelines	258
Applying the timeline	259
Timelines as filters	260

Attribute-based DNA	262
Preparing the data	262
Implementing and viewing dynamic attribute networks	264
Creating dynamic GEXF files	274
Summary	277
Chapter 9: Taking Your Graph Beyond Gephi	279
Overview of the available tools	280
Graph file exporters	280
CSV files	281
DL files	282
GDF files	282
GEXF files	283
GML files	284
GraphML files	284
NET files	285
VNA files	285
Image exporters	286
PNG export	286
SVG export	288
PDF export	292
Web exporters	294
Seadragon Web Export	294
SigmaExporter	296
Loxa Web Site Export	298
Exporting a web graph	300
Seadragon	300
Sigma.js Exporter	302
Loxa Web Site Export	308
Summary	314
Chapter 10: Putting It All Together	315
Using Gephi to understand existing networks	316
Creating new Gephi projects	324
Project 1 – Newman NetScience dataset	325
Exploring the network in Gephi	326
Deploying the project to the Web	331
Project 2 – high school network with dynamic edges	333
Using Gephi to explore the network	334
Creating the project as a PDF	338
Anticipating the future of network analysis	341
Summary	343

Appendix: Data Sources and Other Web Resources	345
Data sources	345
Web resources	346
Import processes	346
Bibliography	346
Index	349