Contents

List of Contributors  XV

1  Using the DiffCorr Package to Analyze and Visualize Differential Correlations in Biological Networks  1
Atsushi Fukushima and Kozo Nishida

1.1 Introduction  1
1.1.1 An Introduction to Omics and Systems Biology  1
1.1.2 Correlation Networks in Omics and Systems Biology  1
1.1.3 Network Modules and Differential Network Approaches  2
1.1.4 Aims of this Chapter  4
1.2 What is DiffCorr?  4
1.2.1 Background  4
1.2.2 Methods  5
1.2.3 Main Functions in DiffCorr  5
1.2.4 Installing the DiffCorr Package  6
1.3 Constructing Co-Expression (Correlation) Networks from Omics Data – Transcriptome Data set  8
1.3.1 Downloading the Transcriptome Data set  8
1.3.2 Data Filtering  9
1.3.3 Calculation of the Correlation and Visualization of Correlation Networks  11
1.3.4 Graph Clustering  15
1.3.5 Gene Ontology Enrichment Analysis  17
1.4 Differential Correlation Enrichment Analysis by DiffCorr Package  21
1.4.1 Calculation of Differential Co-Expression between Organs in Arabidopsis  21
1.4.2 Exploring the Metabolome Data of Flavonoid-Deficient Arabidopsis  26
1.4.3 Avoiding Pitfalls in (Differential) Correlation Analysis  29
1.5 Conclusion  30
Acknowledgments  30
Conflicts of Interest  30
References  30
## Contents

### 2 Analytical Models and Methods for Anomaly Detection in Dynamic, Attributed Graphs 35

**Benjamin A. Miller, Nicholas Arcolano, Stephen Kelley, and Nadya T. Bliss**

- 2.1 Introduction 35
- 2.2 Chapter Definitions and Notation 36
- 2.3 Anomaly Detection in Graph Data 37
  - 2.3.1 Neighborhood-Based Techniques 37
  - 2.3.2 Frequent Subgraph Techniques 38
  - 2.3.3 Anomalies in Random Graphs 39
- 2.4 Random Graph Models 41
  - 2.4.1 Models with Attributes 41
  - 2.4.2 Dynamic Graph Models 43
- 2.5 Spectral Subgraph Detection in Dynamic, Attributed Graphs 44
  - 2.5.1 Problem Model 44
  - 2.5.2 Filter Optimization 46
  - 2.5.3 Residuals Analysis in Attributed Graphs 47
- 2.6 Implementation in R 50
- 2.7 Demonstration in Random Synthetic Backgrounds 51
- 2.8 Data Analysis Example 55
- 2.9 Summary 58

### Acknowledgments 58

### References 59

### 3 Bayesian Computational Algorithms for Social Network Analysis 63

**Alberto Caimo and Isabella Gollini**

- 3.1 Introduction 63
- 3.2 Social Networks as Random Graphs 64
- 3.3 Statistical Modeling Approaches to Social Network Analysis 64
  - 3.3.1 Exponential Random Graph Models (ERGMs) 65
  - 3.3.2 Latent Space Models (LSMs) 65
- 3.4 Bayesian Inference for Social Network Models 66
  - 3.4.1 R-Based Software Tools 67
  - 3.5 Data 67
  - 3.5.1 Bayesian Inference for Exponential Random Graph Models 68
  - 3.5.2 Bayesian Inference for Latent Space Models 71
  - 3.5.3 Predictive Goodness-of-Fit (GoF) Diagnostics 76
- 3.6 Conclusions 80

### References 81

### 4 Threshold Degradation in R Using iDEMO 83

**Chien-Yu Peng and Ya-Shan Cheng**

- 4.1 Introduction 83
- 4.2 Statistical Overview: Degradation Models 85
  - 4.2.1 Wiener Degradation-Based Process 85
  - 4.2.1.1 Lifetime Information 86

### References 81
4.2.1.2 Log-Likelihood Function 87
4.2.2 Gamma Degradation-Based Process 88
4.2.2.1 Lifetime Information 88
4.2.2.2 Log-Likelihood Function 89
4.2.3 Inverse Gaussian Degradation-Based Process 89
4.2.3.1 Lifetime Distribution 90
4.2.3.2 Log-Likelihood Function 91
4.2.4 Model Selection Criteria 91
4.2.5 Choice of $\Lambda(t)$ 91
4.2.6 Threshold Degradation 92
4.3 iDEMO Interface and Functions 92
4.3.1 Overview of the Package iDEMO Functionality 93
4.3.2 Data Input Format 93
4.3.3 Starting the iDEMO 93
4.3.3.1 Import Data 94
4.3.3.2 Basic Information 95
4.3.3.3 Degradation Model Selection 96
4.3.4 Single Degradation Model Analysis 96
4.3.4.1 Parameter Estimation 97
4.3.4.2 Lifetime Information 98
4.3.5 Odds and Ends 101
4.3.6 Computational Details 101
4.4 Case Applications 101
4.4.1 Laser Example 102
4.4.2 Fatigue Example 106
4.4.3 ADT Example 112
4.5 Concluding Remarks 122
References 122

5 Optimization of Stratified Sampling with the R Package SamplingStrata: Applications to Network Data 125
Marco Ballin and Giulio Barcaroli
5.1 Networks and Stratified Sampling 125
5.2 The R Package SamplingStrata 126
5.2.1 General Setting 126
5.2.2 A General Procedure for the Optimization of Strata in a Frame 130
5.2.3 An Example 132
5.3 Application to Networks 139
5.3.1 Use of Networks as Frames 139
5.3.2 Sampling Massive Networks 145
5.4 Conclusions 149
References 149
6 Exploring the Role of Small Molecules in Biological Systems Using Network Approaches 151
Rajarshi Guha and Sourav Das
6.1 The Role of Networks in Drug Discovery 152
6.2 R for Network Analyses 153
6.3 Linking Small Molecules to Targets, Pathways, and Diseases 154
6.3.1 Drug–Target Networks 154
6.3.2 Disease Networks 155
6.3.3 SAR Networks 156
6.3.4 Assay Networks 157
6.3.5 Scaffold Networks 158
6.3.6 Scaffold-Document Networks 159
6.4 R as a Platform for Network Analyses in Drug Discovery 162
6.5 Discussion 165
Acknowledgments 165
References 166

7 Performing Network Alignments with R 173
Qiang Huang and Ling-Yun Wu
7.1 Introduction 173
7.2 Problems, Models, and Algorithms 175
7.2.1 Problems 176
7.2.1.1 Pairwise Network Alignment 176
7.2.1.2 Network Querying 178
7.2.1.3 Multiple Network Alignment 179
7.2.2 Models and Algorithms 180
7.2.3 Comparison and Challenges 180
7.2.3.1 NQ Versus PNA 180
7.2.3.2 PNA Versus MNA 182
7.2.3.3 Challenges 182
7.3 Algorithms Based on Conditional Random Fields 183
7.3.1 CNetQ for Network Querying 183
7.3.1.1 General Framework 183
7.3.1.2 Feature Function 185
7.3.1.3 Gap Penalty 185
7.3.1.4 Network Simplification 186
7.3.1.5 Real Examples 186
7.3.2 CNetA for Pairwise Network Alignment 186
7.3.2.1 Iterative Bidirectional Mapping Strategy 187
7.3.2.2 Simulated Data 188
7.3.2.3 Comparison 188
7.3.2.4 Evaluation Measures 189
7.3.3 CNetMA for Multiple Network Alignment 189
7.3.3.1 Græmlin 189
7.3.3.2 IsoRank 190
<table>
<thead>
<tr>
<th>Sections</th>
<th>Pages</th>
</tr>
</thead>
<tbody>
<tr>
<td>7.3.3.3 MNA Examples</td>
<td>190</td>
</tr>
<tr>
<td>7.3.3.4 CNetMA</td>
<td>191</td>
</tr>
<tr>
<td>7.4 Performing Network Alignments with R</td>
<td>193</td>
</tr>
<tr>
<td>7.4.1 Installation</td>
<td>193</td>
</tr>
<tr>
<td>7.4.1.1 CRF Package</td>
<td>193</td>
</tr>
<tr>
<td>7.4.1.2 Corbi Package</td>
<td>193</td>
</tr>
<tr>
<td>7.4.2 Usage</td>
<td>193</td>
</tr>
<tr>
<td>7.4.2.1 Input File Format</td>
<td>194</td>
</tr>
<tr>
<td>7.4.2.2 Output File Format</td>
<td>194</td>
</tr>
<tr>
<td>7.4.2.3 Arguments</td>
<td>194</td>
</tr>
<tr>
<td>7.4.3 Examples</td>
<td>195</td>
</tr>
<tr>
<td>7.4.3.1 Network Querying</td>
<td>195</td>
</tr>
<tr>
<td>7.4.3.2 Pairwise Network Alignment</td>
<td>195</td>
</tr>
<tr>
<td>7.4.4 Web Services and Tool Functions</td>
<td>196</td>
</tr>
<tr>
<td>7.5 Discussion</td>
<td>196</td>
</tr>
<tr>
<td>References</td>
<td>197</td>
</tr>
</tbody>
</table>

8 \( \ell_p \)-Penalized Methods in High-Dimensional Gaussian Markov Random Fields 201

Luigi Augugliaro, Angelo M. Mineo, and Ernst C. Wit

8.1 Introduction 201
8.2 Graph Theory: Terminology and Basic Topological Notions 202
8.3 Probabilistic Graphical Models 203
8.4 Markov Random Field 204
8.4.1 Ising Model and Extensions 205
8.4.2 Gaussian Markov Random Fields 206
8.5 Sparse Inference in High-dimensional GMRFs 207
8.5.1 Neighborhood Selection 207
8.5.2 The R Package simone 209
8.5.3 Osteolytic Lesions Data Set: An Analysis by Neighborhood Selection Method 210
8.5.4 Graphical Lasso Estimator 215
8.5.5 The R Package glasso: Computing the Gradient and Coefficient Solution Path on a Simulated Data Set 217
8.5.6 Computational Aspects of the glasso Estimator: the Block-Coordinate Descent Algorithm 223
8.5.7 Faster Computation via Exact Covariance Thresholding 225
8.5.8 Lung Cancer Microarray Data: An Analysis by glasso Estimator 227
8.5.9 The Joint Graphical Lasso 233
8.5.10 Computational Aspects of the jglasso Estimator: ADMM Algorithm 235
8.5.11 The R Package JGL 239
8.5.12 Lung Cancer Microarray Data: An Analysis by jglasso Estimator 241
8.5.13 Structured Graphical Lasso 243
9 Cluster Analysis of Social Networks Using R 267

Malika Charrad

9.1 Introduction 267
9.2 Cluster Analysis in Social Networks 268
9.2.1 Social Network Data 268
9.2.1.1 The Data as a Graph 268
9.2.1.2 The Data as a Matrix 269
9.2.2 Clustering in Social Networks 269
9.3 Cluster Analysis in Social Networks Using R 270
9.3.1 R Packages for Cluster Analysis 270
9.3.2 Data Loading and Formatting 270
9.3.2.1 Removing Zero Edges 271
9.3.2.2 Coercing the Data into a Graph Object 271
9.3.2.3 Creating Social and Task Subgraphs 272
9.3.3 Agglomerative Hierarchical Clustering 274
9.3.3.1 Measuring Similarity/Dissimilarity 274
9.3.3.2 Clustering 275
9.3.3.3 Cluster Validity 276
9.3.4 Edge Betweenness Clustering Algorithm 279
9.3.5 Fast Greedy Modularity Optimization Algorithm 281
9.3.6 Walktrap Algorithm 283
9.4 Discussion and Further Readings 285
References 286

10 Inference and Analysis of Gene Regulatory Networks in R 289

Ricardo de M. Simoes, Matthias Dehmer, Constantine Mitsiades, and Frank Emmert-Streib

10.1 Introduction 289
10.2 Multiple Myeloma 290
10.3 Installation of Required R Packages from CRAN and Bioconductor 291
10.4 Data Preprocessing 292
10.5 Bc3net Gene Regulatory Network Inference 294
10.6 Retrieving and Generating Gene Sets for a Functional Analysis 297
10.7 Pathway and Other Gene Set Collections 298
10.7.1 Functional Enrichment Analysis of Gene Regulatory Networks 300
11 Visualization of Biological Networks Using NetBioV

Shailesh Tripathi, Salissou Moutari, Matthias Dehmer, and Frank Emmert-Streib

11.1 Introduction

11.2 Network Visualization

11.3 NetBioV

11.3.1 Global Network Layouts

11.3.2 Modular Network Layout

11.3.3 Layered Network (Multiroot) Layout

11.3.4 Other Features

11.3.4.1 Information Flow

11.3.4.2 Spiral View

11.3.4.3 Color Schemes, Node Labeling

11.3.4.4 Interface to R and Customization

11.4 Example: Visualization of Networks Using NetBioV

11.4.1 Loading Library and Data

11.4.2 Global Layout Style

11.4.2.1 R Code in Figure 11.4

11.4.3 Modular Layout Style

11.4.3.1 R Code in Figure 11.5

11.4.4 Layered Layout Style

11.4.4.1 R Code in Figure 11.6

11.5 Conclusion

11.6 Appendix

11.6.1 R Code for the Visualization in Figures 11.2 and 11.3

11.7 Spiral View

11.7.1 Spiral Layout Style in Figure 11.7

References

Index