Matthias Dehmer
Editor

Structural Analysis of Complex Networks

Birkhäuser

Editor
Ao. Prof. Dr. habil. Matthias Dehmer
Institute for Bioinformatics and Translational Research
The Health and Life Sciences University
UMIT-Private Universität für Gesundheitswissenschaften
Eduard Wallnöfer-Zentrum 1
A-6060 Hall in Tirol, Austria

and

Institute of Discrete Mathematics and Geometry
Vienna University of Technology
Wiedner Hauptstrasse 8-10
A-1040 Vienna, Austria

ISBN 978-0-8176-4788-9 e-ISBN 978-0-8176-4789-6
DOI 10.1007/978-0-8176-4789-6
Springer Dordrecht Heidelberg London New York

Library of Congress Control Number: 2010938359

Mathematics Subject Classification: 05C05, 05C12, 05C75, 05C80, 05C85, 05D40, 68R10, 90B10, 92E10, 94C15

Printed on acid-free paper

www.birkhauser-science.com

Preface

Because of the increasing complexity and growth of real-world networks, their analysis by using classical graph-theoretic methods is oftentimes a difficult procedure. Thus, there is a strong need to combine graph-theoretic methods with mathematical techniques from other scientific disciplines, such as machine learning, statistics, and information theory, for analyzing complex networks more adequately.

The book *Structural Analysis of Complex Networks* presents theoretical as well as practice-oriented results for structurally exploring complex networks. Hence, the book does not only focus on classical graph-theoretical methods, it also shows the usefulness and potential of structural graph theory as a tool for solving interdisciplinary problems. Special emphasis is given to methods and areas which can be roughly summarized as follows:

- Graph-theoretical applications in, e.g., structural biology, computational biology, mathematical chemistry, and computational linguistics;
- Graph classes;
- General structural properties of networks;
- Graph colorings;
- Graph polynomials;
- Information measures for graphs, e.g., graph entropies;
- Metrical properties of graphs;
- Partitions and decompositions;
- Quantitative graph measures.

This book is intended for an interdisciplinary audience, covering topics from artificial intelligence, computer science, computational and systems biology, cognitive science, computational linguistics, discrete mathematics, machine learning, mathematical chemistry, and statistics, and it contains nineteen chapters that have been peer-reviewed according to the standards of international journals in applied mathematics. The chapters and some of their interrelations can be briefly described as follows.

Emmert-Streib starts the volume by surveying basic structural properties of complex networks, important graph classes, and graph measures used when performing network analysis quantitatively. The latter relates to determining the structural similarity between graphs and their structural complexity using entropic measures.

Further concepts used to explore networks structurally are provided by the next chapters authored by *Borowiecki*, *Goddard* et al., and *Ananchuen* et al. In particular, these chapters present techniques of graph partitioning, distances in graphs, and domination in graphs, respectively. The chapter written by *Fujii* also discusses entropy measures, but for infinite directed graphs. However, these measures are obtained by using operator theory and, hence, are differently defined than the ones presented in the chapter by Emmert-Streib; those are derived based on Shannon's entropy and can be interpreted as the structural information content of a graph. Then, the chapters authored by *Matsumoto*, *Kovář*, and *Brešar* et al. investigate multifaceted problems, like exploring infinite labeled graphs to study presentations of symbolic dynamical systems, special graph decompositions, and the examination of geodetic sets in graphs, which represents an important problem using metrical properties of graphs. *Ellis-Monaghan* et al. provide two chapters in this volume on graph polynomials: The first one emphasizes the Tutte polynomial and some closely related graph polynomials. The second chapter by *Ellis-Monaghan* et al. sheds light on interpretations of concrete polynomials and on interrelations between other graph polynomials and the Tutte polynomial. The problem of reconstructing graphs by examining specific properties of polynomials, here, the zeros of Krawtchouk polynomials, is tackled in the next chapter by *Stoll*. Quantitative methods to calculate the structural similarity or distance between two graphs have already been mentioned in Emmert-Streib's chapter. In particular, classical measures based on determining isomorphic subgraphs have already been mentioned there. The chapter written by *Lauri* treats the graph similarity problem in a similar manner, namely based on the number of common vertex-deleted subgraphs, and examines aspects of the computational complexity for calculating the mentioned graph similarity measure. The idea of defining structural distances between graphs is tackled in the next chapter, written by *Benadé*. More precisely, a chromatic metric is defined, and, remarkably, by applying this metric, the maximum distance between any two graphs is at most three.

The last six chapters of this volume use graph-theoretic techniques to solve challenging problems in, e.g., applied mathematics, computer science, quantum chemistry, electrical engineering, computational linguistics, structural biology and RNA structure analysis, computational biology, and mathematical chemistry. The chapter authored by *Cioabă* gives a broad overview on results for relating important structural properties of a graph to its eigenvalues. Also, *Cioabă* surveys important applications of graph spectra in subfields of the just-mentioned disciplines. To demonstrate the great potential of novel graph classes within computational linguistics, *Mehler* introduces a graph class consisting of hierarchical graphs called Minimum Spanning Markovian Trees and shows its usefulness by outlining concrete applications within semiotic network analysis. The chapter contributed by *Scripps* et al. starts by reviewing techniques to mine general complex networks, but mainly focuses on link-based classification, which often appears as an important problem in Web mining. The next two chapters, authored by *Washietl* et al. and *Mason* et al., explore graph-based problems in structural and computational biology, respectively. In particular, *Washietl* et al. investigate RNA structures represented by

graphs and review graph-theoretical methods for describing and comparing such structures. A problem that is currently of considerable interest in biological network analysis is addressed by *Mason* et al. and deals with surveying methods for predicting protein function based on complex interaction networks. An area in which graph-theoretical models and techniques have been intensely applied so far is mathematical chemistry. The volume concludes by presenting a chapter about a graph class that is meaningful in mathematical chemistry: *Vukičević* presents techniques for determining the existence and enumeration of what are called perfect matchings that correspond to Kekulé structures, which are well known in mathematical chemistry.

Many colleagues, whether consciously or unconsciously, provided input, help, and support before and during the formation of this book. In particular, I would like to thank Hamid Arabnia, Alireza Ashrafi, Alexandru T. Balaban, Subhash Basak, Igor Bass, Agnieszka Bergel, David Bialy, Danail Bonchev, Stefan Borgert, Mieczysław Borowiecki, Monique Borusiak, Ulrike Brandt, Mathieu Dutour, Michael Drmota, Abdol-Hossein Esfahanian, Maria Fonoberova, Bernhard Gittenberger, Arno Homburg, Jürgen Kilian, Elena Konstantinova, Reinhard Kutzelnigg, Dmitrii Lozovanu, Alexander Mehler, Tomás Madaras, Abbe Mowshowitz, Marina Popovscaia, Fred Sobik, Stefan Shetschew, Doru Stefanescu, Thomas Stoll, Kurt Varmuza, Ilona Wesarg, Bohdan Zelinka, Dongxiao Zhu, and all authors and co-authors of this book. I apologize to any who inadvertently have not been named.

I am deeply grateful to Armin Graber from UMIT for his strong support and for providing such a stimulating working atmosphere. Many thanks to Isabella Fritz, Bernd Haas, Gerd Lorünser, Brigitte Senn-Kircher, and Klaus Weinberger for their help and fruitful discussions. Moreover, I thank Frank Emmert-Streib for the extremely fruitful collaboration and many stimulating discussions we had over several years. Frank also provided the figures used to design the front cover of this book.

In addition, I would like to thank editors Tom Grasso, Rebecca Biega, and Regina Gorenshteyn from Birkhäuser Publishing (Boston), who have always been available and helpful. Last but not least, I would like to thank my wife Jana and my family — Marion Dehmer-Sehn and Werner Dehmer — for their unfailing support and encouragement.

Finally, I hope that this book will help to extend the enthusiasm and joy that I feel for this field to others, and that it will inspire people to apply graph theory to different scientific areas for the solution of challenging and interdisciplinary problems.

Hall in Tirol, April 2010 Matthias Dehmer

Contents

Preface v

Contributors xi

1 A Brief Introduction to Complex Networks and Their Analysis 1
Frank Emmert-Streib

2 Partitions of Graphs 27
Mieczysław Borowiecki

3 Distance in Graphs 49
Wayne Goddard and Ortrud R. Oellermann

4 Domination in Graphs 73
Nawarat Ananchuen, Watcharaphong Ananchuen, and Michael D. Plummer

5 Spectrum and Entropy for Infinite Directed Graphs 105
Jun Ichi Fujii

6 Application of Infinite Labeled Graphs to Symbolic Dynamical Systems 137
Kengo Matsumoto

7 Decompositions and Factorizations of Complete Graphs 169
Petr Kovář

8 Geodetic Sets in Graphs 197
Boštjan Brešar, Matjaž Kovše, and Aleksandra Tepeh

9 Graph Polynomials and Their Applications I: The Tutte Polynomial 219
Joanna A. Ellis-Monaghan and Criel Merino

10 Graph Polynomials and Their Applications II: Interrelations and Interpretations 257
Joanna A. Ellis-Monaghan and Criel Merino

11 Reconstruction Problems for Graphs, Krawtchouk Polynomials, and Diophantine Equations 293
Thomas Stoll

12 Subgraphs as a Measure of Similarity 319
Josef Lauri

13 A Chromatic Metric on Graphs 335
Gerhard Benadé

14 Some Applications of Eigenvalues of Graphs 357
Sebastian M. Cioabă

15 Minimum Spanning Markovian Trees: Introducing Context-Sensitivity into the Generation of Spanning Trees 381
Alexander Mehler

16 Link-Based Network Mining 403
Jerry Scripps, Ronald Nussbaum, Pang-Ning Tan, and Abdol-Hossein Esfahanian

17 Graph Representations and Algorithms in Computational Biology of RNA Secondary Structure 421
Stefan Washietl and Tanja Gesell

18 Inference of Protein Function from the Structure of Interaction Networks 439
Oliver Mason, Mark Verwoerd, and Peter Clifford

19 Applications of Perfect Matchings in Chemistry 463
Damir Vukičević

Index 483

Contributors

Nawarat Ananchuen Department of Mathematics, Faculty of Science, Silpakorn University, Nakorn Pathom 73000, Thailand, nawarat@su.ac.th

Watcharaphong Ananchuen School of Liberal Arts, Sukhothai Thammathirat Open University, Nonthaburi 11120, Thailand, laasawat@stou.ac.th

Gerhard Benadé School of Computer Science, Statistics and Mathematics, North-West University, Potchefstroom, South Africa, gerhard.benade@nwu.ac.za

Mieczysław Borowiecki Faculty of Mathematics, Computer Science and Econometrics, University of Zielona Góra, Podgórna 50, 65-246 Zielona Góra, Poland, M.Borowiecki@wmie.uz.zgora.pl

Boštjan Brešar Faculty of Natural Sciences and Mathematics, University of Maribor, Koroška 160, 2000 Maribor, Slovenia, bostjan.bresar@uni-mb.si

Sebastian M. Cioabă Department of Mathematical Sciences, University of Delaware, 501 Ewing Hall, Newark, DE 19716-2553, USA, cioaba@math.udel.edu

Peter Clifford Hamilton Institute, NUI Maynooth, Maynooth, Ireland, p@pclifford.net

Joanna A. Ellis-Monaghan Department of Mathematics, Saint Michael's College, One Winooski Park, Colchester, VT 05439, USA
and
Department of Mathematics and Statistics, University of Vermont, 16 Colchester Avenue, Burlington, VT 05405, USA, jellis-monaghan@smcvt.edu

Frank Emmert-Streib Computational Biology and Machine Learning, Center for Cancer Research and Cell Biology, School of Medicine, Dentistry and Biomedical Sciences, Queen's University Belfast, 97 Lisburn Road, Belfast BT9 7BL, UK, v@bio-complexity.com

Abdol-Hossein Esfahanian Computer Science and Engineering Department, 3115 Engineering Building, Michigan State University, East Lansing, MI 48824-1226, USA, esfahanian@cse.msu.edu

Jun Ichi Fujii Department of Arts and Sciences (Information Science), Osaka Kyoiku University, Asahigaoka, Kashiwara, Osaka 582-8582, Japan, fujii@cc.osaka-kyoiku.ac.jp

Tanja Gesell Center for Integrative Bioinformatics Vienna, Max F. Perutz Laboratories, Dr. Bohr-Gasse 9, 1030 Vienna, Austria
and
University of Vienna, Medical University of Vienna, and University of Veterinary Medicine, Vienna, Austria, tanja.gesell@univie.ac.at

Wayne Goddard School of Computing and Department of Mathematical Sciences, Clemson University, Clemson, SC 29634-1906, USA, goddard@clemson.edu

Petr Kovář Department of Applied Mathematics, Technical University Ostrava, 17. listopadu, 708 33 Ostrava–Poruba, Czech Republic, petr.kovar@vsb.cz

Matjaž Kovše Faculty of Natural Sciences and Mathematics, University of Maribor, Koroška 160 2000 Maribor, Slovenia matjaz.kovse@uni-mb.si

Josef Lauri Department of Mathematics, University of Malta, Tal-Qroqq, Malta, josef.lauri@um.edu.mt

Oliver Mason Hamilton Institute, NUI Maynooth, Maynooth, Ireland, oliver.mason@nuim.ie

Kengo Matsumoto Department of Mathematics, Joetsu University of Education, Joetsu 943-8512, Japan, kengo@juen.ac.jp

Alexander Mehler Goethe-University Frankfurt am Main, Senckenberganlage 31, 60325 Frankfurt am Main, Germany, Mehler@em.uni-frankfurt.de

Criel Merino Instituto de Matemáticas, Universidad Nacional Autónoma de México, Area de la Investigación Científica, Circuito Exterior, C.U., Coyoacán, 04510 México D.F., México, merino@matem.unam.mx

Ronald Nussbaum Computer Science and Engineering Department, 3115 Engineering Building, Michigan State University, East Lansing, MI 48824-1226, USA, ronald@cse.msu.edu

Ortrud R. Oellermann Department of Mathematics and Statistics, University of Winnipeg, Winnipeg, MB R3B 2E9, Canada o.oellermann@uwinnipeg.ca

Michael D. Plummer Department of Mathematics, Vanderbilt University, Nashville, TN 37240, USA, michael.d.plummer@vanderbilt.edu

Jerry Scripps School of Computing and Information Systems, 1 Campus Drive, Grand Valley State University, Allendale, MI 49401, USA, scrippsj@gvsu.edu

Thomas Stoll Faculty of Mathematics, School of Computer Science, University of Waterloo, Waterloo, ON, Canada, tstoll@cs.uwaterloo.ca

Pang-Ning Tan Computer Science and Engineering Department, 3115 Engineering Building, Michigan State University, East Lansing, MI 48824-1226, USA, ptan@cse.msu.edu

Aleksandra Tepeh University of Maribor, FEECS, Smetanova 17, 2000 Maribor, Slovenia, aleksandra.tepeh@uni-mb.si

Mark Verwoerd Hamilton Institute, NUI Maynooth, Maynooth, Ireland, mark.verwoerd@nuim.ie

Damir Vukičević Faculty of Mathematics and Natural Sciences, University of Split, Nikole Tesle 12, HR-21000 Split, Croatia, vukicevi@pmfst.hr

Stefan Washietl EMBL-European Bioinformatics Institute, Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SD, UK
and
Institute for Theoretical Chemistry, University of Vienna, Währingerstraße 17, A-1090 Vienna, Austria, washietl@ebi.ac.uk