



Matthias Dehmer
Frank Emmert-Streib
Alexander Mehler
Editors

Towards an Information Theory of Complex Networks

Statistical Methods and Applications



Editors

Matthias Dehmer
UMIT
Institute of Bioinformatics
and Translational Research
Eduard-Wallnöfer-Zentrum I
A-6060 Hall in Tirol
Austria
mathias.dehmer@umit.at

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

Alexander Mehler
Faculty of Computer Science
and Mathematics
Goethe-University Frankfurt
am Main Robert-Mayer-Straße 10
P.O. Box: 154
D-60325 Frankfurt am Main
Germany
mehler@em.uni-frankfurt.de

ISBN 978-0-8176-4903-6 e-ISBN 978-0-8176-4904-3
DOI 10.1007/978-0-8176-4904-3
Springer New York Dordrecht Heidelberg London

Library of Congress Control Number: 2011932673

Mathematics Subject Classification (2010): 68R10, 68P30, 94C15

© Springer Science+Business Media, LLC 2011

All rights reserved. This work may not be translated or copied in whole or in part without the written permission of the publisher (Springer Science+Business Media, LLC, 233 Spring Street, New York, NY 10013, USA), except for brief excerpts in connection with reviews or scholarly analysis. Use in connection with any form of information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed is forbidden.

The use in this publication of trade names, trademarks, service marks, and similar terms, even if they are not identified as such, is not to be taken as an expression of opinion as to whether or not they are subject to proprietary rights.

Printed on acid-free paper

www.birkhauser-science.com

Preface

For more than a decade, complex network analysis has evolved as a methodological paradigm for a multitude of disciplines, including physics, chemistry, biology, geography, sociology, computer science, statistics, media science, and linguistics. Researchers in these fields share an interest in information processing subject to the networking of their corresponding research object, for instance, genes, molecules, individuals, semes, memes, etc. They start with the insight that any of these research objects is *extrinsically* characterized, if not constituted, by its networking with objects of the same provenance. In this way, networks, for example, gene networks, food networks, city networks, networks of words, sentences, texts, or web documents become important research objects in more and more disciplines.

This book, in line with these research developments, presents theoretical and practical results of statistical models of complex networks in the formal sciences, the natural sciences, and the humanities. One of its goals is to advocate and promote combinations of graph-theoretic, information-theoretic, and statistical methods as a way to better understand and characterize real-world networks.

On the one hand, networks appear as paradigmatic objects of approaches throughout the natural and social sciences and the humanities. On the other hand, networks are—irrespective of their disciplinary provenance—known for characteristic distributions of graph-theoretic invariants which affect their robustness and efficiency in information processing. The main goal of this book is to further develop information-theoretic notions and to elaborate statistical models of information processing in such complex networks. In this way, the book includes first steps toward establishing a statistical information theory as a unified basis for complex network analysis across a multitude of scientific disciplines.

The book presents work on the statistics of complex networks together with applications of information theory in a range of disciplines such as quantitative biology, quantitative chemistry, quantitative sociology, and quantitative linguistics. It aims to integrate models of invariants of network topologies and dynamic aspects of information processing in these networks or by means of these networks.

Thus, the book is in support of sharing and elaborating models and methods that may help researchers get insights into complex problems emerging from interdisciplinary reasoning.

The book is divided into two parts: Chaps. 1–4 deal with formal-theoretical issues of network modeling, while Chaps. 5–13 further develop and apply these methods to empirical networks from a wide range of areas. The book starts with a theoretical contribution by *Abbe Mowshowitz* on the entropy of digraphs and infinite graphs. The aim is to provide insights into more complex graph models that go beyond the majority of network models based on finite undirected graphs. The chapter by *Nicolas Bonichon*, *Cyril Gavoille*, and *Nicolas Hanusse* presents an information-theoretic upper bound of planar graphs by means of the newly introduced notion of well-orderly maps. Such a technique might be useful when studying properties of the very important notion of planar graphs. *Terence Chan* and *Raymond W. Yeung* study a statistical inference problem using network models. *Richard Berkovits*, *Lukas Jahnke*, and *Jan W. Kantelhardt* examine phase transitions within complex networks that help to examine their structural properties.

The remainder of the book combines the theoretical stance of the first section with an empirical analysis of real networks. *Elena Konstantinova* provides a survey on information-theoretic measures used in chemical graph theory. *Prabhat K. Sahu* and *Shyi-Long Lee* develop a model of chemical graphs by example of molecular networks. Exploring the spectral characteristics of these graphs, they provide a successful classification of chemical graphs.

Biological or, more specifically, ecological networks are dealt with by *Robert E. Ulanowicz* who describes a framework of quantifying patterns of the interaction of networked trophic processes from the point of view of information theory. Ecological networks are also the focus of the chapter of *Linda J. Moniz*, *James D. Nichols*, *Jonathan M. Nichols*, *Evan G. Cooch*, and *Louis M. Pecora*, who provide an approach to modeling the interaction dynamics of ecosystems and their change. A comprehensive view of ontologically disparate networks is given by *Cristian R. Munteanu*, *J. Dorado*, *A. Pazos Sierra*, *F. Prado-Prado*, *L.G. Pérez-Montoto*, *S. Vilar*, *F.M. Ubeira*, *A. Sanchez-González*, *M. Cruz-Monteagudo*, *S. Arrasate*, *N. Sotomayor*, *E. Lete*, *A. Duardo-Sánchez*, *A. Díaz-López*, *G. Patlewicz*, and *H. González-Díaz* who use the notion of entropy centrality to compare various systems such as chemical, biological, crime, and legislative networks, thereby showing the interdisciplinary expressiveness of complex network theory.

The book continues with two contributions to linguistic networks: *Alexander Mehler* develops a framework for analyzing the topology of social ontologies as they evolve within Wikipedia and contrasts them with nonsocial, formal ontologies. *Olga Abramov* and *Tatjana Lokot* present a comparative, classificatory study of morphological networks by means of several measures of graph entropy.

Edward B. Allen discusses the measurement of the complexity and error probability of software systems represented as hypergraphs. Finally, in the chapter by *Philippe Blanchard* and *Dimitri Volchenkov*, random walks are studied as a kind of Markov process on graphs that allow insights into the dynamics of networks as diverse as city and trade and exchange networks.

With such a broad field, it is clear that the present book addresses an interdisciplinary readership. It does not simply promote transdisciplinary research. Rather, it is about interdisciplinary research that may be the starting point of developing an overarching network science.

Matthias Dehmer
Frank Emmert-Streib
Alexander Mehler

Acknowledgments

Many colleagues have provided us with input, help, and support (consciously or unconsciously) before and during the preparation of this book. In particular, we would like to thank Andreas Albrecht, Gökmen Altay, Gabriel Altmann, Alain Barrat, Igor Bass, David Bialy, Philippe Blanchard, Danail Bonchev, Stefan Borgert, Mieczysław Borowiecki, Andrey A. Dobrynin, Michael Drmota, Ramon Ferrer i Cancho, Maria and Gheorghe Duca, Maria Fonoberova, Armin Graber, Martin Grabner, Peter Gritzmann, Ivan Gutman, Peter Hamilton, Wilfried Imrich, Patrick Johnston, Elena Konstantinova, D. D. Lozovanu, Dennis McCance, Abbe Mowshowitz, Arcady Mushegian, Andrei Perjan, Armindo Salvador, Maximilian Schich, Heinz Georg Schuster, Helmut Schwegler, Andre Ribeiro, Burghard Rieger, Brigitte Senn-Kircher, Fred Sobik, Doru Stefanescu, John Storey, Shailesh Tripathi, Kurt Varmuza, Bohdan Zelinka, and Shu-Dong Zhang. Additionally, Matthias Dehmer thanks Armin Graber for strong support and providing a fruitful atmosphere at UMIT. Finally, we would like to thank our editor Tom Grasso who has been always available and helpful.

The work on the chapters of Philippe Blanchard and Dimitri Volchenkov, Olga Abramov, and Alexander Mehler have been supported by the German Federal Ministry of Education and Research (BMBF) through the project *Linguistic Networks*.¹ We gratefully acknowledge this financial support.

¹www.linguistic-networks.net.

Contents

1	Entropy of Digraphs and Infinite Networks	1
	A. Mowshowitz	
2	An Information-Theoretic Upper Bound on Planar Graphs Using Well-Orderly Maps.....	17
	Nicolas Bonichon, Cyril Gavoille, and Nicolas Hanusse	
3	Probabilistic Inference Using Function Factorization and Divergence Minimization	47
	Terence H. Chan and Raymond W. Yeung	
4	Wave Localization on Complex Networks	75
	Richard Berkovits, Lukas Jahnke, and Jan W. Kantelhardt	
5	Information-Theoretic Methods in Chemical Graph Theory.....	97
	Elena Konstantinova	
6	On the Development and Application of Net-Sign Graph Theory	127
	Prabhat K. Sahu and Shyi-Long Lee	
7	The Central Role of Information Theory in Ecology	153
	Robert E. Ulanowicz	
8	Inferences About Coupling from Ecological Surveillance Monitoring: Approaches Based on Nonlinear Dynamics and Information Theory	169
	L.J. Moniz, J.D. Nichols, J.M. Nichols, E.G. Cooch, and L.M. Pecora	

9	Markov Entropy Centrality: Chemical, Biological, Crime, and Legislative Networks	199
	C.R. Munteanu, J. Dorado, Alejandro Pazos-Sierra, F. Prado-Prado, L.G. Pérez-Montoto, S. Vilar, F.M. Ubeira, A. Sanchez-Gonzaléz, M. Cruz-Monteagudo, S. Arrasate, N. Sotomayor, E. Lete, A. Duardo-Sánchez, A. Díaz-López, G. Patlewicz, and H. González-Díaz	
10	Social Ontologies as Generalized Nearly Acyclic Directed Graphs: A Quantitative Graph Model of Social Tagging	259
	Alexander Mehler	
11	Typology by Means of Language Networks: Applying Information Theoretic Measures to Morphological Derivation Networks	321
	Olga Abramov and Tatiana Lokot	
12	Information Theory-Based Measurement of Software	347
	Edward B. Allen	
13	Fair and Biased Random Walks on Undirected Graphs and Related Entropies.....	365
	Philippe Blanchard and Dimitri Volchenkov	

Contributors

Olga Abramov University of Bielefeld, Universitätsstraße 25, 33615 Bielefeld, Germany, olga.abramov@uni-bielefeld.de

Edward B. Allen Department of Computer Science and Engineering, Mississippi State University, Box 9637, Mississippi State, MS 39762, USA, edward.allen@computer.org

S. Arrasate Department of Organic Chemistry II, Faculty of Science and Technology, University of the Basque Country/Euskal Herriko Unibertsitatea, Apt. 644, 48080 Bilbao, Spain, sonia.arrasate@ehu.es

Richard Berkovits Minerva Center and Department of Physics, Bar-Ilan University, Ramat-Gan 52900, Israel, berkov@mail.biu.ac.il

Philippe Blanchard Bielefeld – Bonn Stochastic Research Center (BiBoS), University of Bielefeld, Postfach 100131, 33501 Bielefeld, Germany

Nicolas Bonichon LaBRI, University of Bordeaux, 351 Cours de la libération, 33405 Bordeaux, France, bonichon@labri.fr

Terence H. Chan Institute for Telecommunications Research, University of South Australia, Adelaide, SA 5095, Australia, hlchan6@gmail.com; terence.chan@unisa.edu.au

E.G. Cooch Department of Natural Resources, Cornell University, Ithaca, NY 14853, USA, evan.cooch@cornell.edu

M. Cruz-Monteagudo CEQA, Faculty of Chemistry and Pharmacy, UCLV, Santa Clara 54830, Cuba, gmailkelem@yahoo.es

A. Díaz-López Department of Special Public Law, Faculty of Law, University of Santiago de Compostela, 15782 Santiago de Compostela, Spain, antonio.lopez.diaz@usc.es

J. Dorado Department of Information and Communication Technologies, Computer Science Faculty, University of A Coruña, 15071 A Coruña, Spain, julian@udc.es

A. Duardo-Sánchez Department of Special Public Law, Faculty of Law, University of Santiago de Compostela, 15782 Santiago de Compostela, Spain, aliuskaduardo@yahoo.es

Cyril Gavoille LaBRI, University of Bordeaux, 351 Cours de la libération, 33405 Bordeaux, France, gavoille@labri.fr

H. González-Díaz Department of Microbiology and Parasitology, Faculty of Pharmacy, University of Santiago de Compostela, 15782 Santiago de Compostela, Spain, humberto.gonzalez@usc.es

Nicolas Hanusse LaBRI, CNRS – University of Bordeaux, 351 Cours de la libération, 33405 Bordeaux, France, hanusse@labri.fr

Lukas Jahnke Martin-Luther-Universität Halle-Wittenberg, 06099 Halle, Germany

Jan W. Kantelhardt Martin-Luther-Universität Halle-Wittenberg, 06099 Halle, Germany

Elena Konstantinova Sabolev Institute of Mathematics, Siberian Branch of Russian Academy of Sciences, 630090 Novosibirsk, Russia, e.konsta@math.nsc.ru

Shyi-Long Lee Department of Chemistry and Biochemistry, National Chung Cheng University, Chia-Yi, 621 Taiwan, chesll@ccu.edu.tw

E. Lete Department of Organic Chemistry II, Faculty of Science and Technology, University of the Basque Country/Euskal Herriko Unibertsitatea, Apt. 644, 48080 Bilbao, Spain, esther.lete@ehu.es

Tatiana Lokot Faculty of Technology, University of Bielefeld, Universitätsstr. 25, 33615 Bielefeld, Germany, tlokot@math.uni-bielefeld.de

Alexander Mehler Faculty of Computer Science and Mathematics, Goethe University Frankfurt am Main, D-60325 Frankfurt am Main, Germany, mehler@em.uni-frankfurt.de

L.J. Moniz Johns Hopkins University, Applied Physics Laboratory, 11100 Johns Hopkins Road, Laurel, MD 20723, USA, lindano@comcast.net

Abbe Mowshowitz Department of Computer Science, The City College of New York (CUNY), 138th Street at Convent Avenue, New York, NY 10031, USA, abbe@cs.ccny.cuny.edu

C.R. Munteanu Department of Information and Communication Technologies, Computer Science Faculty, University of A Coruña, 15071 A Coruña, Spain, cmunteaunu@udc.es

J.D. Nichols U.S. Geological Survey, Patuxent Wildlife Research Center, Laurel, MD 20708, USA, jnichols@usgs.gov

J.M. Nichols Naval Research Laboratory, Optical Sciences Division, Code 5673, Washington, DC 20375, USA, jonathan.nichols@nrl.navy.mil

G. Patlewicz Institute for Health and Consumer Protection (IHPC), Joint Research Centre (JRC), European Commission, via E. Fermi 2749–21027 Ispra (Varese), Italy
DuPont Haskell Global Centers for Health and Environmental Sciences, Newark, DE 19711, USA, Grace.Y.Tier@usa.dupont.com

Alejandro Pazos-Sierra Department of Information and Communication Technologies, Computer Science Faculty, University of A Coruña, 15071 A Coruña, Spain, apazos@udc.es

L.M. Pecora Naval Research Laboratory, Code 6362, Washington, DC 20375, USA, pecora@anvil.nrl.navy.mil

L.G. Pérez-Montoto Faculty of Pharmacy, University of Santiago de Compostela, 15782 Santiago de Compostela, Spain, lgmp2002@yahoo.es

F. Prado-Prado Faculty of Pharmacy, University of Santiago de Compostela, 15782 Santiago de Compostela, Spain, feno11@hotmail.com

Prabhat K. Sahu Institut für Physikalische und Theoretische Chemie, Universität Würzburg, Am Hubland, 97074 Würzburg, Germany

Department of Chemistry and Biochemistry, National Chung Cheng University, Chia-Yi, 621 Taiwan, sahu@chemie.uni-wuerzburg.de

A. Sanchez-Gonzaléz Department of Inorganic Chemistry, Faculty of Pharmacy, University of Santiago de Compostela, 15782 Santiago de Compostela, Spain, angeles.sanchez@usc.es

N. Sotomayor Department of Organic Chemistry II, Faculty of Science and Technology, University of the Basque Country/Euskal Herriko Unibertsitatea, Apt. 644, 48080 Bilbao, Spain, nuria.sotomayor@ehu.es

F.M. Ubeira Department of Microbiology and Parasitology, Faculty of Pharmacy, University of Santiago de Compostela, 15782 Santiago de Compostela, Spain, fm.ubeira@usc.es

Robert E. Ulanowicz Department of Biology, University of Florida, Gainesville, FL 32611-8525, USA

University of Maryland Center for Environmental Science, Solomons, MD 20688-0038, USA, ulan@umces.edu

S. Vilar Faculty of Pharmacy, University of Santiago de Compostela, 15782 Santiago de Compostela, Spain, qosanti@yahoo.es

Dimitri Volchenkov The Center of Excellence Cognitive Interaction Technology (CITEC), University of Bielefeld, Postfach 100131, 33501 Bielefeld, Germany, volchenk@physik.uni-bielefeld.de

Raymond W. Yeung Department of Information Engineering, The Chinese University of Hong Kong, whyeung@ie.cuhk.edu.hk