

Commenced Publication in 1973

Founding and Former Series Editors:

Gerhard Goos, Juris Hartmanis, and Jan van Leeuwen

Editorial Board

David Hutchison

Lancaster University, UK

Takeo Kanade

Carnegie Mellon University, Pittsburgh, PA, USA

Josef Kittler

University of Surrey, Guildford, UK

Jon M. Kleinberg

Cornell University, Ithaca, NY, USA

Alfred Kobsa

University of California, Irvine, CA, USA

Friedemann Mattern

ETH Zurich, Switzerland

John C. Mitchell

Stanford University, CA, USA

Moni Naor

Weizmann Institute of Science, Rehovot, Israel

Oscar Nierstrasz

University of Bern, Switzerland

C. Pandu Rangan

Indian Institute of Technology, Madras, India

Bernhard Steffen

TU Dortmund University, Germany

Demetri Terzopoulos

University of California, Los Angeles, CA, USA

Doug Tygar

University of California, Berkeley, CA, USA

Gerhard Weikum

Max Planck Institute for Informatics, Saarbruecken, Germany

Andreas Kerren Helen C. Purchase
Matthew O. Ward (Eds.)

Multivariate Network Visualization

Dagstuhl Seminar #13201
Dagstuhl Castle, Germany, May 12-17, 2013
Revised Discussions

Volume Editors

Andreas Kerren

Linnaeus University, Faculty of Technology (FTK)

Department of Computer Science

Vejdes Plats 7, 351 95 Växjö, Sweden

E-mail: kerren@acm.org

Helen C. Purchase

University of Glasgow, School of Computing Science

18 Lilybank Gardens, Glasgow G12 1RZ, UK

E-mail: hcp@dcs.gla.ac.uk

Matthew O. Ward

Worcester Polytechnic Institute, Department of Computer Science

100 Institute Road, Worcester, MA 01609-2280, USA

E-mail: matt@cs.wpi.edu

ISSN 0302-9743

e-ISSN 1611-3349

ISBN 978-3-319-06792-6

e-ISBN 978-3-319-06793-3

DOI 10.1007/978-3-319-06793-3

Springer Cham Heidelberg New York Dordrecht London

Library of Congress Control Number: 2014937318

LNCS Sublibrary: SL 3 – Information Systems and Application, incl. Internet/Web and HCI

© Springer International Publishing Switzerland 2014

This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed. Exempted from this legal reservation are brief excerpts in connection with reviews or scholarly analysis or material supplied specifically for the purpose of being entered and executed on a computer system, for exclusive use by the purchaser of the work. Duplication of this publication or parts thereof is permitted only under the provisions of the Copyright Law of the Publisher's location, in its current version, and permission for use must always be obtained from Springer. Permissions for use may be obtained through RightsLink at the Copyright Clearance Center. Violations are liable to prosecution under the respective Copyright Law.

The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

While the advice and information in this book are believed to be true and accurate at the date of publication, neither the authors nor the editors nor the publisher can accept any legal responsibility for any errors or omissions that may be made. The publisher makes no warranty, express or implied, with respect to the material contained herein.

Typesetting: Camera-ready by author, data conversion by Scientific Publishing Services, Chennai, India

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

Preface

During May 12–17, 2013, a seminar on “Information Visualization – Towards Multivariate Network Visualization” (no. 13201) took place at the International Conference and Research Center for Computer Science, Dagstuhl Castle, Germany. The center was initiated by the German government to promote computer science research at the international level. It seeks to foster dialog among the research community, advance academic education and professional development, and transfer knowledge between academia and industry.

Information visualization (InfoVis) is a research area that focuses on the use of visualization techniques to help people understand and analyze data as well as relations between data. While related fields such as scientific visualization involve the presentation of data that have some physical or geometric correspondence (for example, climate patterns, molecular formations, transport networks), InfoVis centers on abstract information without such correspondences, i.e., it is not possible to map this information into the physical world in most cases. Examples of such abstract data are symbolic, tabular, networked, hierarchical, or textual information sources—for example, genealogies, demographic data of a population, or financial trends.

The goal of this third Dagstuhl Seminar on Information Visualization was to bring together theoreticians and practitioners from InfoVis, HCI, and graph drawing with a special *focus on multivariate network visualization*, i.e., on graphs where the nodes and/or edges have additional (multidimensional) attributes. The integration of multivariate data into complex networks and their visual analysis is one of the big challenges not only in visualization, but also in many application areas. Thus, in order to support discussions related to the visualization of real-world data, we also invited researchers from selected application areas, especially bioinformatics, social sciences, and software engineering. The unique “Dagstuhl climate” ensured an open and undisturbed atmosphere to discuss the state-of-the-art, new directions, and open challenges of multivariate network visualization.

This book is the outcome of Dagstuhl Seminar no. 13201. It documents and extends the findings and discussions of the various sessions in detail. During

the last day of the seminar, the most important topics for publication were identified and assigned to interested participants. The resulting author groups worked together to write book chapters on the chosen topics.

We would like to thank all participants of the seminar for the lively discussions and contributions during the seminar as well as the scientific directorate of Dagstuhl Castle for giving us the possibility to organize this event. The abstracts and presentation slides can be found on the Dagstuhl website for this seminar¹. There is an online document that reports on all activities during the seminar². We are also grateful to all the authors for their valuable time and contributions to the book. Last but not least, the seminar and thereby this book would not have been possible without the great help of the staff of Dagstuhl Castle. We would like to acknowledge all of them for their assistance.

January 2014

Andreas Kerren
Helen C. Purchase
Matthew O. Ward

¹ <http://www.dagstuhl.de/13201>

² <http://dx.doi.org/10.4230/DagRep.3.5.19>

List of Contributors

James Abello

Rutgers University, DIMACS Center for Discrete Mathematics
and Theoretical Computer Science, USA
abello@dimacs.rutgers.edu

Daniel Archambault

Swansea University, Department of Computer Science, UK
d.w.archambault@swansea.ac.uk

Katy Börner

Indiana University Bloomington, Department of Library
and Information Science, USA
katy@indiana.edu

Stephan Diehl

University of Trier, Department of Computer Science, Germany
diehl@uni-trier.de

Tim Dwyer

Monash University Clayton, Caulfield School of Information Technology,
Australia
tim.dwyer@monash.edu

Niklas Elmqvist

Purdue University, School of Electrical & Computer Engineering, USA
elm@purdue.edu

Jean-Daniel Fekete

INRIA, France
Jean-Daniel.Fekete@inria.fr

Liang Gou

IBM Research Almaden, USA
lgou@us.ibm.com

Hans Hagen

University of Kaiserslautern, Department of Computer Science, Germany
hagen@cs.uni-kl.de

Danny Holten

SynerScope, Eindhoven, The Netherlands
danny.holten@synerscope.com

Christophe Hurter

École Nationale de l'Aviation Civile, Interactive Computing Laboratory,
France
christophe.hurter@enac.fr

T.J. Jankun-Kelly

Mississippi State University, Department of Computer Science
and Engineering, USA
tjk@acm.org

Jessie Kennedy

Edinburgh Napier University, School of Computing, UK
j.kennedy@napier.ac.uk

Andreas Kerren

Linnaeus University Växjö, Department of Computer Science, Sweden
kerren@acm.org

Stephen Kobourov

University of Arizona, Department of Computer Science, USA
kobourov@cs.arizona.edu

Oliver Kohlbacher

University of Tübingen, Center for Bioinformatics,
Quantitative Biology Center, Department of Computer Science,
and Faculty of Medicine, Germany
oliver.kohlbacher@uni-tuebingen.de

Kwan-Liu Ma

University of California at Davis, Department of Computer Science, USA
ma@cs.ucdavis.edu

Silvia Miksch

Vienna University of Technology, Institute of Software Technology
and Interactive Systems, Austria
miksch@ifs.tuwien.ac.at

Chris Muelder

University of California at Davis, Department of Computer Science, USA
cwmuelder@ucdavis.edu

Martin Nöllenburg

Karlsruhe Institute of Technology, Institute of Theoretical Informatics,
Germany
noellenburg@kit.edu

A. Johannes Pretorius

University of Leeds, School of Computing, UK
a.j.pretorius@leeds.ac.uk

Helen C. Purchase

University of Glasgow, School of Computing Science, UK
hcp@dcs.gla.ac.uk

Jonathan C. Roberts

Bangor University, School of Computer Science, UK
j.c.roberts@bangor.ac.uk

Falk Schreiber

Martin Luther University Halle-Wittenberg, Institute of Computer
Science, Germany; Leibniz Institute of Plant Genetics and Crop Plant
Research (IPK), Germany; and Monash University Melbourne,
Clayton School of Information Technology, Australia
schreibe@ipk-gatersleben.de

John T. Stasko

Georgia Institute of Technology, School of Interactive Computing, USA
stasko@cc.gatech.edu

Alexandru C. Telea

University of Groningen, Institute Johann Bernoulli, The Netherlands
a.c.telea@rug.nl

Jarke J. van Wijk

Eindhoven University of Technology, Department of Mathematics
and Computer Science, The Netherlands
vanwijk@win.tue.nl

Tatiana von Landesberger

Technische Universität Darmstadt, Department of Computer Science,
Germany
tatiana.von-landesberger@gris.tu-darmstadt.de

Matthew O. Ward

Worcester Polytechnic Institute, Department of Computer Science, USA
matt@cs.wpi.edu

Chris Weaver

University of Oklahoma, School of Computer Science, USA
weaver@cs.ou.edu

Michael Wybrow

Caulfield School of Information Technology, Monash University Caulfield,
Australia
Michael.Wybrow@monash.edu

Kai Xu

Middlesex University London, School of Science and Technology, UK
k.xu@mdx.ac.uk

Jing Yang

University of North Carolina at Charlotte, Computer Science
Department, USA
Jing.Yang@uncc.edu

Dirk Zeckzer

Leipzig University, Institute of Computer Science, Germany
zeckzer@informatik.uni-leipzig.de

Michelle X. Zhou

IBM Research Almaden, USA
mzhou@us.ibm.com

Björn Zimmer

Linnaeus University Växjö, Department of Computer Science, Sweden
bjorn.zimmer@lnu.se

Contents

Preface	V
List of Contributors	VII
1 Introduction to Multivariate Network Visualization	1
<i>Andreas Kerren, Helen C. Purchase, Matthew O. Ward</i>	
1.1 Multivariate Networks: Definitions and Terminology	2
1.2 Existing Visualizations	3
1.3 Outline of This Book	6
References	7

Part I: Application Domains – Characteristics and Challenges

2 Multivariate Networks in Software Engineering	13
<i>Stephan Diehl, Alexandru C. Telea</i>	
2.1 Aims and Scope	13
2.1.1 History and Definitions	14
2.1.2 Importance	14
2.2 Data Characteristics	15
2.2.1 Entities	15
2.2.2 Relations	15
2.2.3 Attributes	16
2.2.4 Software as Multivariate Time-Dependent Graphs ...	17
2.2.5 Reference Implementation	17
2.2.6 Software Data <i>vs.</i> other InfoVis Domains	19
2.3 Applications	21
2.3.1 Structure Visualization	21
2.3.2 Behavior Visualization	26
2.3.3 Evolution Visualization	28
2.4 Challenges and Future Directions	32
References	34

3	Multivariate Social Network Visual Analytics.....	37
	<i>Chris Muelder, Liang Gou, Kwan-Liu Ma, Michelle X. Zhou</i>	
3.1	Data Characteristics	37
3.1.1	Traditional Data Collection	38
3.1.2	Data Collection from Online Social Media Networks...	38
3.2	Task Characteristics	39
3.2.1	Understanding Social Network Nodes	39
3.2.2	Understanding Social Network Links	39
3.2.3	Understanding Social Networks	39
3.3	Examples of Technologies.....	40
3.3.1	Clustering	40
3.3.2	Network Centralities	41
3.3.3	Centrality Derivatives	41
3.3.4	Traditional Network Layouts.....	42
3.3.5	Improved Network Layouts	43
3.3.6	Multivariate Social Networks	44
3.3.7	Dynamic Social Networks	46
3.3.8	Egocentric Approaches	47
3.3.9	TreeNetViz: Revealing Patterns of Networks with Hierarchical Attributes	47
3.3.10	SocialNetSense: Making Sense of Multivariate Social Networks	50
3.3.11	Summary.....	54
3.4	Challenges and Future Directions.....	55
	References	56
4	Multivariate Networks in the Life Sciences	61
	<i>Oliver Kohlbacher, Falk Schreiber, Matthew O. Ward</i>	
4.1	Characteristics of Data and Tasks	62
4.1.1	Types of Biological Networks	62
4.1.2	Data Mapping and Multivariate Networks	63
4.2	Use Cases	66
4.2.1	Signaling	66
4.2.2	Genetic Linkage	66
4.2.3	Relationship Discovery Based on Document Analysis.....	67
4.2.4	Gene Regulation and Transcriptome Data	68
4.3	Challenges	69
4.3.1	Scale	70
4.3.2	Uncertainty/Ambiguity	70
4.3.3	Heterogeneity	71
4.3.4	Interactivity	71
4.3.5	Standardization	71
4.4	Summary and Conclusions.....	72
	References	72

Part II: Topics in Multivariate Network Research

5	Tasks for Multivariate Network Analysis	77
	<i>A. Johannes Pretorius, Helen C. Purchase, John T. Stasko</i>	
5.1	Entities and Properties	78
5.2	Tasks	79
	5.2.1 General Task Taxonomy	80
	5.2.2 Tasks for Multivariate Network Analysis	83
5.3	Discussion	92
5.4	Conclusion	94
	References	95
6	Interaction in the Visualization of Multivariate Networks	97
	<i>Michael Wybrow, Niklas Elmqvist, Jean-Daniel Fekete, Tatiana von Landesberger, Jarke J. van Wijk, Björn Zimmer</i>	
6.1	Background	98
6.2	Classification of Interactions	101
	6.2.1 View-Level Interactions	102
	6.2.2 Visual Structure-Level Interactions	104
	6.2.3 Data-Level Interactions	106
6.3	Exemplars	109
	6.3.1 GraphDice	110
	6.3.2 GraphTrail	111
	6.3.3 State Transition Networks	112
	6.3.4 Parallel Node-Link Bands	113
6.4	Recommendations and Guidelines	115
	6.4.1 Learnability	115
	6.4.2 Flexibility	116
	6.4.3 Robustness	117
6.5	Challenges and Vision	118
	References	121
7	Novel Visual Metaphors for Multivariate Networks	127
	<i>Jonathan C. Roberts, Jing Yang, Oliver Kohlbacher, Matthew O. Ward, Michelle X. Zhou</i>	
7.1	Background	128
	7.1.1 Semantically Rich Data	128
	7.1.2 Where Ideas Come From?	129
	7.1.3 The Ideation Process	131
	7.1.4 The Visual Mapping Process	132
7.2	Classes of Metaphors	135
	7.2.1 Nature-Inspired	135
	7.2.2 Non-physical	140

7.3	Man-Made	141
7.4	Visualization-Inspired	142
7.5	Proposed New Ideas	144
7.5.1	Graphs/Networks	145
7.5.2	Hierarchies	145
7.6	Summary and Conclusions	147
	References	147
8	Temporal Multivariate Networks	151
	<i>Daniel Archambault, James Abello, Jessie Kennedy,</i>	
	<i>Stephen Kobourov, Kwan-Liu Ma, Silvia Miksch,</i>	
	<i>Chris Muelder, Alexandru C. Telea</i>	
8.1	Definitions	151
8.1.1	Structure, Behavior, and Evolution	152
8.1.2	Formal Definitions of Temporal Multivariate Networks	153
8.2	Refining Our Models and Definitions for Time	154
8.3	Survey of Representations and Algorithms	156
8.3.1	Static Graph Layouts	157
8.3.2	Dynamic Graph Layouts	158
8.3.3	Animation versus Small Multiples	161
8.3.4	Mental Map Preservation	162
8.3.5	Alternative Representations	162
8.3.6	Static Temporal Plots	163
8.3.7	Dynamic Graph Analytics	164
8.4	Applications to Software Engineering	165
8.5	Open Problems	167
8.5.1	Attribute Dimensionality	167
8.5.2	Capturing Patterns	168
8.5.3	Data Size	168
8.6	Summary and Conclusions	169
	References	169
9	Heterogeneous Networks on Multiple Levels	175
	<i>Falk Schreiber, Andreas Kerren, Katy Börner,</i>	
	<i>Hans Hagen, Dirk Zeckzer</i>	
9.1	Formal Description of Used Data Structures	177
9.2	Application Domains	179
9.2.1	Life Sciences / Biology	179
9.2.2	Social Science	185
9.2.3	Software Engineering	190
9.3	Visualization	196
9.3.1	Approaches for Networks at Multiple Levels	197
9.3.2	Challenges and Future Directions	200
	References	201

10 Scalability Considerations for Multivariate Graph Visualization 207
T.J. Jankun-Kelly, Tim Dwyer, Danny Holten, Christophe Hurter, Martin Nöllenburg, Chris Weaver, Kai Xu

10.1 Limits of Visualization 208
 10.1.1 Limits of Visual Acuity 208
 10.1.2 Cognitive Limits 210
 10.1.3 Leveraging the Graphics Card (GPU) 211

10.2 Design Strategies for Scalable Multivariate Graph Visualization 215
 10.2.1 Data Transformation and Reduction 217
 10.2.2 Visual Mapping 222
 10.2.3 View Transformation 223

10.3 Studies on Scalability in Graph Visualization 224
 10.3.1 Data Transformation and Reduction 224
 10.3.2 Visual Mapping 225
 10.3.3 Navigation and Interaction 227

10.4 Challenges and Future Directions 229
 References 229

Author Index 237