Report from Dagstuhl Seminar 13201

Information Visualization – Towards Multivariate Network Visualization

Edited by

Andreas Kerren[1], Helen C. Purchase[2], and Matthew O. Ward[3]

1 **Linnaeus University – Växjö, SE,** kerren@acm.org
2 **University of Glasgow, GB,** helen.purchase@glasgow.ac.uk
3 **Worcester Polytechnic Institute, US,** matt@cs.wpi.edu

Abstract

Information Visualization (InfoVis) focuses on the use of visualization techniques to help people understand and analyze large and complex data sets. The aim of this third Dagstuhl Seminar on Information Visualization was to bring together theoreticians and practitioners from Information Visualization, 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. To support discussions related to the visualization of real world data, researchers from selected application areas, especially bioinformatics, social sciences, and software engineering, were also invited. During the seminar, working groups on six different topics were formed and enabled a critical reflection on ongoing research efforts, the state of the field in multivariate network visualization, and key research challenges today. This report documents the program and the outcomes of Dagstuhl Seminar 13201 "Information Visualization – Towards Multivariate Network Visualization".

Seminar 12.–17. May, 2013 – www.dagstuhl.de/13201

1998 ACM Subject Classification E.1 [Data Structures]: Graphs and networks, H.5.2 [Information Interfaces and Presentation]: User Interfaces—Graphical user interfaces (GUI), I.3.6 [Computer Graphics]: Methodology and Techniques—Interaction techniques

Keywords and phrases Information visualization, visualization, network visualization, graph drawing, visual analytics, network analysis, interaction

Digital Object Identifier 10.4230/DagRep.3.5.19

Edited in cooperation with Björn Zimmer (Linnaeus University – Växjö, SE)

1 Executive Summary

Andreas Kerren
Helen C. Purchase
Matthew O. Ward

Introduction

Information Visualization (InfoVis) is a research area that focuses on the use of visualization techniques to help people understand and analyze data. While related fields such as Scientific Visualization involve the presentation of data that has some physical or geometric correspondence, Information Visualization 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.

The first two Dagstuhl Seminars on Information Visualization aimed to cover more general aspects of our field, such as interaction, evaluation, data wrangling, and collaboration, or focused on higher level topics, for instance, the value of InfoVis or the importance of aesthetics. Besides the Dagstuhl reports that are typically published directly after a seminar [1, 2, 4, 5], there were also follow-up publications for both seminars. The participants of Seminar #07221 wrote book chapters which have been consolidated into a Springer book [7]; the organizers of the same seminar published a workshop report in the Information Visualization journal [6]. For the second Seminar #10241, a special issue in the same journal was published [3].

The goal of this third Dagstuhl Seminar on Information Visualization was to bring together theoreticians and practitioners from Information Visualization, 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.

Seminar Topics

The following themes were discussed during the seminar. The seminar allowed attendees to critically reflect on current research efforts, the state of field, and key research challenges today. Participants also were encouraged to demonstrate their system prototypes and tools relevant to the seminar topics. In consequence, topics emerged in the seminar week and were the focus of deeper discussions too.

- **Focus on biochemistry/bioinformatics:** In the life sciences, huge data sets are generated by high-throughput experimental techniques. Consequently, biologists use computational methods to support data analysis. The information in many experimental data sets can be either represented as networks or interpreted in the context of various networks. How can our current techniques help to analyze primary and secondary data in the context of such networks, and how can different network types be combined?
- **Focus on social science:** Graph drawing techniques have been used for several years for the visualization and analysis of social networks, but other social science fields (e.g., geography, politics, cartography, and economics) also make use of data visualization. How can (or do) our network visualizations support these domains?
- **Focus on software engineering:** In the application domain of software engineering, various graphs and data attached to graphs (e.g., software metrics) play a dominant role in the static and dynamic analysis of programs. Which of these problems are conceptually similar to graph-related problems in biology or social sciences and how can multivariate network visualization support specific tasks, such as software architecture recovery?
- **Approaches and methods:** There already exist a number of technical approaches, algorithms, and methods to interactively visualize multivariate networks. Which ones are suitable for solving specific tasks in our applications areas? What is their potential? What are their limitations? By identifying the range of approaches that do exist, can we see the potential for new, innovative visualization ideas?

- **Challenges in visualizing multivariate networks:** Multivariate networks are large and complex and their complexity will increase in the future. Thus, not all problems can be solved in the short term. What are the current challenges?
- **Time-dependent/dynamic networks:** Many networks that are considered in practice change over time with respect to their topology and/or their attributes. How can we best visualize networks and attributes that change over time?
- **Interaction:** How can we best support the navigation, exploration and modification of multivariate networks?
- **Multiple networks at different scales:** How can we integrate, combine, compare more than one multivariate network at different scales? In this context, the term of so-called multi-modal networks is often used in literature. What does this term mean exactly? Can we visualize a range of different information types concurrently?
- **Tasks:** What range of tasks can multivariate network visualization support? Are there general tasks for all application domains?
- **Novel metaphors:** What type of visualization metaphors should we use beyond node-link diagrams? What would be the benefit in doing so?

Outcomes

The organizers and participants decided to write a book on multivariate network visualization to be published as LNCS issue by Springer. The possibility of publishing this Springer book was confirmed by the Editor-in-Chief of LNCS already before the start of the seminar. Working groups have been invited to submit a book chapter building on their discussions and findings, and writing is underway. The final chapters are to be submitted by November 3, 2013, with a planned publication date of Spring 2014. A preliminary book structure was presented at the end of the seminar:

1. Introduction
   a. Definition of multivariate networks, typical representations
2. Domain Application Data Characteristics in Context of Multivariate Networks
   a. Biology
   b. Social Sciences
   c. Software Engineering
3. Tasks
4. Interaction
5. Metaphors (Visual Mappings beyond Node-Link)
6. Multiple and Multi-Domain Networks
7. Temporal Networks
8. Scalability
9. Summary/Conclusion

The Dagstuhl team performed an evaluation at the end of the seminar week. The results of this survey (scientific quality, inspiration to new ideas/projects/research/papers, insights from neighboring fields, ...) were throughout very good to excellent. Only a few single improvements were proposed by participants, for example, more junior researchers should be invited to come into contact with world-class researchers. And more domain experts should be invited to be spread out across the breakout groups. Another issue was that the time available for group work should be extended in future seminars.

References

1 Andreas Kerren, Catherine Plaisant, and John T. Stasko. 10241 Abstracts Collection: Information Visualization. In Andreas Kerren, Catherine Plaisant, and John T. Stasko, editors, *Information Visualization*, number 10241 in Dagstuhl Seminar Proceedings, Dagstuhl, Germany, 2010. Schloss Dagstuhl – Leibniz-Zentrum fuer Informatik, Germany.

2 Andreas Kerren, Catherine Plaisant, and John T. Stasko. 10241 Executive Summary: Information Visualization. In Andreas Kerren, Catherine Plaisant, and John T. Stasko, editors, *Information Visualization*, number 10241 in Dagstuhl Seminar Proceedings, Dagstuhl, Germany, 2010. Schloss Dagstuhl – Leibniz-Zentrum fuer Informatik, Germany.

3 Andreas Kerren, Catherine Plaisant, and John T. Stasko. Information Visualization: State of the Field and New Research Directions. *Information Visualization*, 10(4):269–270, 2011.

4 Andreas Kerren, John T. Stasko, Jean-Daniel Fekete, and Chris North. 07221 Abstracts Collection: Information Visualization – Human-Centered Issues in Visual Representation, Interaction, and Evaluation. In Jean-Daniel Fekete, Andreas Kerren, Chris North, and John T. Stasko, editors, *Information Visualization – Human-Centered Issues in Visual Representation, Interaction, and Evaluation*, number 07221 in Dagstuhl Seminar Proceedings, Dagstuhl, Germany, 2007. Internationales Begegnungs- und Forschungszentrum für Informatik (IBFI), Schloss Dagstuhl, Germany.

5 Andreas Kerren, John T. Stasko, Jean-Daniel Fekete, and Chris North. 07221 Executive Summary: Information Visualization – Human-Centered Issues in Visual Representation, Interaction, and Evaluation. In Jean-Daniel Fekete, Andreas Kerren, Chris North, and John T. Stasko, editors, *Information Visualization – Human-Centered Issues in Visual Representation, Interaction, and Evaluation*, number 07221 in Dagstuhl Seminar Proceedings, Dagstuhl, Germany, 2007. Internationales Begegnungs- und Forschungszentrum für Informatik (IBFI), Schloss Dagstuhl, Germany.

6 Andreas Kerren, John T. Stasko, Jean-Daniel Fekete, and Chris North. Workshop Report: Information Visualization – Human-Centered Issues in Visual Representation, Interaction, and Evaluation. *Information Visualization*, 6(3):189–196, 2007.

7 Andreas Kerren, John T. Stasko, Jean-Daniel Fekete, and Chris North, editors. *Information Visualization: Human-Centered Issues and Perspectives*, volume 4950 of *Lecture Notes in Computer Science*. Springer, Berlin, Heidelberg, 2008.

2 Table of Contents

3 Seminar Program and Activities

Participation and Program

41 people from 9 countries participated in this seminar. Most attendees came from the US, United Kingdom and Germany; others came from France, The Netherlands, and other European countries as shown in Figure 1. At least two domain experts from the application domains (bioinformatics, social sciences, software engineering) participated in the seminar.

The program aimed to generate lively discussions. Before the seminar started, the organizers asked the domain experts to prepare talks (45 minutes) that highlight the characteristics of their domain-specific data sets and the tasks that analysts perform on the data. At the same time, all participants were called to volunteer for three survey talks (45 minutes) which show the state-of-the-art of visualization and interaction techniques applied to the data considered in each domain; in addition three talks of the same length were invited to present some of the technical challenges of multivariate network visualization within each domain. Individual presentations on new techniques were also welcome.

Thus, the first half of the seminar week was mainly dedicated to the presentations of the invited talks followed by discussions. In the second half, primarily breakout groups were formed to deal with specific topics (briefly illuminated in the next subsection) including group reports to the whole audience. On Thursday, the organizers alloted two slots for scientific presentations on techniques and tool demos. Table 1 provides an overview of the final seminar schedule.

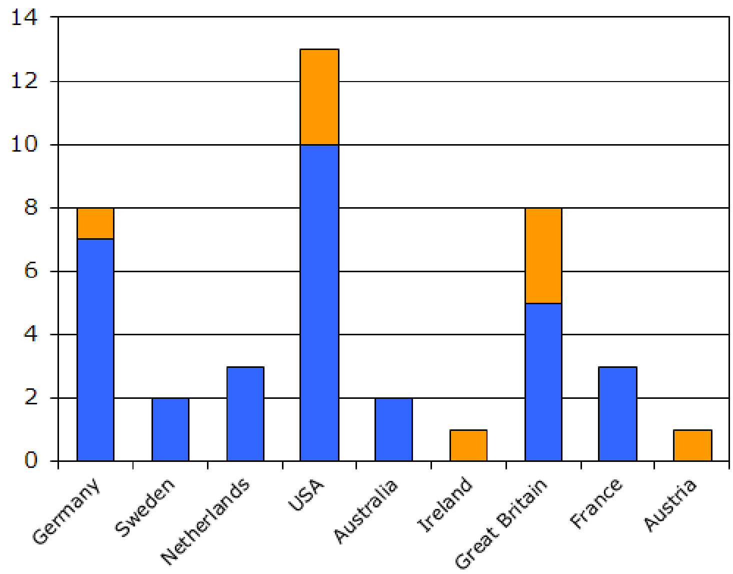

Figure 1 Attendee Statistics of Seminar #13201. Blue colored bars represent male and orange colored bars female participants.

Table 1 Final structure of the seminar. The main discussion topic of Monday was network visualization with a focus on bioinformatics. Tuesday was mainly focused on social network visualization, whereas Wednesday morning was used to discuss visualizations in context of networks that occur in software engineering. Thursday and Friday were mainly focused on general techniques for the visualization of multivariate networks as well as on group work and group reporting.

Monday	Tuesday	Wednesday	Thursday	Friday
Welcome Self Presentations	Characteristics-Talks (social) Presentations (social, 1 talk)	Characteristics/ Survey-Talk (SE) Challenges-Talk (SE)	Presentations (general, 3 talks)	Synergy Session & Sum Up
Self Presentations Characteristics-Talks (bio)	Challenges-Talk (social) Presentations (social, 2 talks)	Breakout Groups Group Reporting	Breakout Groups	Book Planning & Closing Remarks
Survey-Talk (bio) Challenges-Talk (bio)	Discussion: Topics for Breakout Groups	Social Event (Völklinger Hütte)	Presentations/ Demos (general, 4 talks)	
Presentations (general, 3 talks)	Breakout Groups		Group Reporting	

Activities

Invited Talks

The titles and presenters of invited talks for each application domain are listed in the following. Abstracts for the individual talks can be found in Sect. 4.

- Multivariate Networks in Biology
  - *Matthew O. Ward and Carsten Görg:* Biological Multivariate Network Visualizations – A Partial Survey [survey]
  - *Oliver Kohlbacher:* Characteristics of Biological Data in the Age of Omics – Part 1 [characteristics]
  - *Falk Schreiber:* Characteristics of Biological Data – Part 2 [characteristics]
  - *Jessie Kennedy, Oliver Kohlbacher, and Falk Schreiber:* Challenges in Visualising Biological Data [challenges]
- Multivariate Networks in Social Sciences
  - *Lothar Krempel:* Social Science Applications of Multivariate Network Visualization [characteristics]
  - *Maura Conway:* Visualising Terrorism and Violent (Online) Political Extremism Data [characteristics]
  - *Michelle X. Zhou:* Top Challenges in Visualizing Multivariate Personal Networks [challenges]
- Multivariate Networks in Software Engineering
  - *Stephan Diehl and Alexandru C. Telea:* Multivariate Graphs in Software Engineering [characteristics, survey, challenges]

Breakout Groups

As already mentioned above, the program included breakout sessions on six specific topics, i.e, six working groups discussed one topic at a time in parallel sessions. The themes were based on topics discussed in the original seminar proposal as well as topics that emerged in the first session on Tuesday afternoon. The detailed working group reports are presented in Sect. 5. In the following, we list the different groups:

1. Temporal Multivariate Networks
2. Interaction for Multivariate Networks
3. Multiple and Multi-domain Networks
4. Scalability of Multivariate Graph Visualization
5. Tasks for Multivariate Network Analysis
6. Novel Visual Metaphors for Multivariate Networks

Scientific Talks and Demos

In addition, a number of speakers gave a scientific talk and/or a tool demo on a theme related to the research questions of the seminar. In sum, 14 talks/demos were given during the seminar (cf. Sect. 4 for details):

- *James Abello:* Multivariate Time-Variant Graphs and Simplicial Complexes: Work in Progress
- *Daniel Archambault:* Effective Visualization of Information Cascades
- *Katy Börner:* The Information Visualization MOOC
- *Tim Dwyer:* CodeMap: Visualizing Software Dependencies in Microsoft Visual Studio
- *Peter Eades:* On the Faithfulness of Graph Visualisations
- *Niklas Elmqvist:* Multi-Phase/Variate/Modal Graphs: Visualization, Interaction, and Evaluation
- *Benjamin David Hennig:* The Complexity of Space
- *Christophe Hurter:* Scalable Multivariate Data Exploration Tools
- *Stephen G. Kobourov:* Maps of Computer Science
- *Kwan-Liu Ma:* Visualization for Studying Social Networks
- *Silvia Miksch:* A Matter of Time: Visual Analytics of Dynamic Social Networks
- *Martin Nöllenburg:* Column-based Graph Layout for Argument Maps
- *John T. Stasko:* Multivariate Network Data & Attribute-based Layout: Two Examples
- *Kai Xu:* GraphScape: Integrated Multivariate Network Visualization

The content of these talks, given for all seminar attendees, raised further key issues and helped the groups to discuss their individual theme from various perspectives.

4 Overview of Talks

4.1 Multivariate Time-Variant Graphs and Simplicial Complexes: Work in Progress

James Abello (Rutgers University – Piscataway, US)

One of the ultimate goals of time-variant data analysis is to synthesize the structure, behavior, and evolution, of the encoded information. Folklore evolution and information flow in communication networks constitute useful data sources to identify some fundamental time-variant data facets. On the other hand, simplicial complexes are useful mathematical tools that aid in the formulation and identification of some of the essential computational questions associated with time-variant graph data. In this case, contractions to planar graphs offer visual appealing representations of the data space that can be used as the main mechanisms for data interaction. These slides offer quick entry points to Computational Folkloristics, Discrepancy in Communication Networks, and Planar Graph Contractability.

4.2 Effective Visualization of Information Cascades

Daniel Archambault (Swansea University, GB)

Dynamic graph attributes appear in many applications, including social media analysis. An attribute is a value associated with a node or edge of the graph and a dynamic attribute is one that changes its value over time. In graphs used in social media analysis, nodes may receive and transmit dynamic attributes to their neighbors, and long chains of these transmissions are known as cascades. Typical methods for visualizing cascades include animations and small multiples representations where nodes change color with the changing attribute value. We present the results of a formal user study that tests the effectiveness of dynamic attribute visualization on graphs. We test the task of locating nodes which amplify the propagation of a cascade and factors such as animation and small multiples, and force-directed and hierarchical layouts. Overall, we found that small multiples was significantly faster than animation with no significant difference in terms of error rate. Participants generally preferred animation to small multiples and a hierarchical layout to a force-directed layouts. Considering each presentation method separately, when comparing force-directed layouts to hierarchical layouts, hierarchical layouts were found to be significantly faster for both presentation methods and significantly more accurate for animation. Thus, for our task, this experiment supports the use of a small multiples interfaces with hierarchically drawn graphs for the visualization of dynamic attributes.

4.3 The Information Visualization MOOC

Katy Börner (Indiana University – Bloomington, US)

The talk discusses the structure and design of the Information Visualization MOOC taught in Spring 2013 and students from 93 different countries attended. Among other topics, the course covers:

- Data analysis algorithms that enable extraction of patterns and trends in data.
- Major temporal, geospatial, topical, and network visualization techniques.
- Discussions of systems that drive research and development.

Everybody who registers at http://ivmooc.cns.iu.edu gains free access to the Scholarly Database (26 million paper, patent, and grant records) and the Sci2 Tool (100+ algorithms and tools). First 'learning analytics' visualizations of student activity and collaboration networks are presented as well.

4.4 Visualizing Terrorism and Violent (Online) Political Extremism Data

Maura Conway (Dublin City University, IE)

This presentation explored the ways in which data on terrorism and violent (online) political extremism has been visualized to-date and ways in which the InfoVis community could contribute in this research area going forward. There are three major reasons for visualizing terrorism and violent political extremism data:

1. To answer specific questions.
2. To explore the data and thereby potentially identify new avenues for research and analysis.
3. To communicate research and results to other researchers, policy makers, security professionals, and others.

Some well-known visualizations in this area include those of specific terrorist networks, including the 9/11 attackers social network(s) (Krebs 2002); time-lapse visualizations of the event data contained in the Global Terrorism Database (GTD); and mapping of various aspects of the Northern Ireland conflict by the CAIN project. The presentation focused, in particular, on the 'new' data available to researchers into terrorism and violent political extremism resulting from the latter migrating some of their activity to the Internet, including social networking sites, which has not been extensively visualized to-date. Suggestions for ways in which the InfoVis community could productively collaborate with scholars in the field of terrorism and violent (online) political extremism were then identified, including mapping terrorist attacks and incidences of violent political extremism; visualizing online VPE content distribution networks; visualizing VPE data collected from across online social networking sites, and similar.

4.5 Multivariate Graphs in Software Engineering

Stephan Diehl (Universität Trier, DE) and Alexandru C. Telea (University of Groningen, NL)

In this presentation we first talk about the importance of visualization in software engineering in general. Next, we give examples of the typical kinds of data used in software engineering in terms of items, attributes, and relations. Since all these can change over time, we conclude that dynamic multivariate compound graphs provide a good common model. Next, we present many examples of visualization tools that show multivariate data and dynamic graphs related to the structure, behavior and evolution of software. For each tool, we give a typical task to be solved with the tool. We conclude by suggesting some challenges for future research in the area.

4.6 CodeMap: Visualizing Software Dependencies in Microsoft Visual Studio

Tim Dwyer (Monash University Melbourne, AU)

I worked for four years with the Visual Studio product group at Microsoft to develop a software dependency visualization tool. In that time we conducted many user studies to try to better understand how graph visualization can help developers with their most difficult tasks. Based on the feedback from these studies we iterated a tool that attempts to support a developer's working memory with a visual scratchpad that co-exists with their established, code-centric workflow. In this talk I will demonstrate the tool we have developed: CodeMap. I hope to share some of the insights we gained from our studies, which sometimes ran counter to own expectations and some of the tacit assumptions of the visualization fraternity.

4.7 On the Faithfulness of Graph Visualisations

Peter Eades (The University of Sydney, AU)

We argue that the classical readability criteria for visualizing graphs, though necessary, are not sufficient for effective graph visualization. We introduce another kind of criterion, generically called "faithfulness", that we believe is necessary in addition to readability. Intuitively, a graph drawing algorithm is "faithful" if it maps different graphs to distinct drawings. In other words, a faithful graph drawing algorithm never maps distinct graphs to the same drawing. This concept is extended to a task-oriented model of visualization.

4.8 Multi-Phase/Variate/Modal Graphs: Visualization, Interaction, and Evaluation

Niklas Elmqvist (Purdue University, US)

In this talk, I will review our recent work on visualizing, navigating in, and interacting with various types of complex graphs: those that (1) evolve over time, (2) contain multivariate data, and (3) involve different types of nodes and links. Specific projects include TimeMatrix, GraphDice, Dynamic Insets, the COE Explorer, and Parallel Node-Link Bands. I will also discuss several human subjects evaluations we performed on various graph-related tasks. Taken together, these projects begin to form an outline for the design space of complex (multi-*) graph visualization.

4.9 The Complexity of Space

Benjamin David Hennig (University of Sheffield, GB)

How can we reduce the complexity of flows to a significant visual representation relating to our basic geographic understanding of the world? From a geographic perspective, this relates strongly to theoretical and conceptual questions of space, which can be understood in manifold ways. Transferring such concepts into geographic visualizations, as demonstrated with a novel gridded approach to applying density-equalizing map transformations, may provide a useful new idea to visualizing flows in other ways. This talk outlines the thinking behind these ideas and presents some examples of how this might contribute to new ways of showing geographic flows.

4.10 Scalable Multivariate Data Exploration Tools

Christophe Hurter (ENAC – Toulouse, FR)

Interactive data exploration and manipulation are often hindered by the size of the datasets. This is particularly true for 3D datasets where the problem is exacerbated by occlusion, important adjacencies, and entangled patterns. These complexities make visual interaction via common filtering techniques difficult. In this presentation, I described a set of techniques aimed at performing real-time multi-dimensional data deformation with the intention of helping people to easily select, analyze, and eliminate specific spatial-and data patterns. Our interaction techniques allow animation between view configurations, semantic filtering and view deformation. Any subset of the data can be selected at will at any step along the animation. Selected data can be filtered and deformed in order to remove occlusion and ease complex data selections. I applied our techniques to the following domain areas: 3D medical imaging, and multivariate network. The technique is simple, flexible and interactive with large datasets (up to 50 of millions displayed data points).

4.11 Challenges in Visualising Biological Data

Jessie Kennedy (Edinburgh Napier University, GB), Oliver Kohlbacher (Universität Tübingen, DE), and Falk Schreiber (IPK Gatersleben & MLU Halle, DE)

Past Dagstuhl Seminars already identified some challenges in visualising biological networks, as described in the paper "A graph-drawing perspective to some open problems in molecular biology" by Albrecht et al. (GD '09). In this presentation, we extend this work by other technical challenges (such as the visual analysis of ontologies or the important aspect of uncertainty in context of multivariate networks) as well as personal challenges of visualization researchers (like the need to understand biological problems or communicating with biologists). Open visualization problems are highlighted with the help of many practical tool examples.

4.12 Maps of Computer Science

Stephen G. Kobourov (University of Arizona – Tucson, US)

We describe a practical approach for visual exploration of research papers. Specifically, we use the titles of papers from the DBLP database to create maps of computer science (MOCS). Words and phrases from the paper titles are the cities in the map, and countries are created based on word and phrase similarity, calculated using co-occurrence. With the help of heatmaps, we can visualize the profile of a particular conference or journal over the base map. Similarly, heatmap profiles can be made of individual researchers or groups such as a department. The visualization system also makes it possible to change the data used to generate the base map. For example, a specific journal or conference can be used to generate the base map and then the heatmap overlays can be used to show the evolution of research topics in the field over the years. As before, individual researchers or research groups profiles can be visualized using heatmap overlays but this time over the journal or conference base map. Finally, research papers or abstracts easily generate visual abstracts giving a visual representation of the distribution of topics in the paper. We outline a modular and extensible system for term extraction using natural language processing techniques, and show the applicability of methods of information retrieval to calculation of term similarity and creation of a topic map. The system is available at mocs.cs.arizona.edu.

4.13 Characteristics of Biological Data in the Age of Omics – Part 1

Oliver Kohlbacher (Universität Tübingen, DE)

This talk gives a brief overview of omics data and its relationship to networks. First, I introduce the usual terminology and show the differences between classical data and omics data. Omics data in systems biology either represent a network or is interpreted in the context of a network. Based on these fundamental characteristics, a hierarchy of biological networks can be identified. The various network elements at the different levels in the hierarchy differ not only in their meaning, but also in their scale.

4.14 Social Science Applications of Multivariate Network Visualization

Lothar Krempel (MPI für Gesellschaftsforschung – Köln, DE)

The social network perspective seeks to understand how the embeddedness of actors provides specific opportunities for action (social capital). Network data allow to characterize actors according to their local or global centrality, their status or equivalent roles, to identify clusters, cliques or communities. These metrics can help to understand how social systems work. Additional attributes of the actors or their relations can greatly enhance the analysis.

Visualizations are an indispensable tool to inspect such complex datasets. I present results from scientific cooperations, where we able to gain new insights into social science questions. Data on capital ties among the largest companies—as collected by governmental agencies—can be used to trace a historic process, i.e., how national economies transform under the regime of internationalization and financial liberalization. Letters among scientists in the 18th century can be used to generate historical science maps. Growth rates in car trade and the composition of trade in parts and components vs. assembled cars give detailed insight into the state of an international division of labor. Data on awards to universities per funding area allow for insight into the specialization of German universities.

4.15 Visualization for Studying Social Networks

Kwan-Liu Ma (University of California – Davis, US)

The network datasets produced by studies in social science have several characteristics. The datasets usually contain collections of small networks. They are mostly multivariate and categorical. There are often missing, incomplete data due to the nature of the data collection methods. Studying the resulting data suggest the need of egocentric visualization techniques. In my presentation, I share with you my experience in working with sociologists to develop egocentric visualization techniques for studying their data. I describe the particular datasets I have worked on and the visualization designs we have derived to meet their needs. As the data collection methods are changing with the widespread use of mobile and web applications, we anticipate a rapid growth in data size in the near future. I also discuss a few techniques for simplifying and visualizing large networks.

4.16 A Matter of Time: Visual Analytics of Dynamic Social Networks

Silvia Miksch (TU Wien, AT)

Several techniques have been proposed to examine static network data, but the visual analysis of dynamic network data is an emerging research field with several open questions. The dynamic and multi-relational nature of this data poses the challenge of understanding both its topological structure and how it changes over time. In this talk, I will present the applied research project ViENA (Visual Enterprise Network Analytics) and how a visual analytics approach supports the examination of dynamic networks according to specific user tasks. Finally, I show possible future directions and challenges.

4.17 Column-based Graph Layout for Argument Maps

Martin Nöllenburg (KIT – Karlsruhe Institute of Technology, DE)

An argument map is a diagram that captures the logical structure of the various arguments and statements in a debate. It can be modeled as a directed hierarchical graph, where vertices are represented as variable-size boxes containing components of an argument and edges indicating relationships as support or attack. We present a new algorithm of Betz et al. (GD'12) for column-based argument map layout following the topology-shape-metrics framework. It uses rectangular fixed-width boxes and orthogonal polyline edges with at most 4 bends. The algorithm is implemented in the Argunet tool (www.argunet.org), of which we show a short demo.

4.18 Characteristics of Biological Data – Part 2

Falk Schreiber (IPK Gatersleben & MLU Halle, DE)

Today, it is of great importance to combine omics, image, volume, and network data with mapping, analysis and visualization possibilities. In this talk, I show the problem of multi-domain biological data and the need of an integrated analysis and visualization of such multi-domain biological data.

4.19 Multivariate Network Data & Attribute-based Layout: Two Examples

John T. Stasko (Georgia Institute of Technology, US)

In this talk I present short demos of two systems. The first, dotlink360, visualizes financial and descriptive information about companies in the mobile computing ecosystem. More importantly, it visualizes alliances and partnerships between these companies. Among others, the system presents four network visualizations using attribute-based node positioning. The second demo presents IEEE InfoVis Conference papers. It shows the total number of Google Scholar citations for each, along with internal citations between the papers. Our technique does not draw edges between the papers but instead uses interaction to highlight the references.

4.20 Biological Multivariate Network Visualizations – A Partial Survey

Matthew O. Ward (Worcester Polytechnic Institute, US) and Carsten Görg (University of Colorado, US)

In this talk, we present a variety of examples for multivariate network visualizations of biological data. We first review a number of different encodings used to map multivariate data to network nodes. We then introduce plugins for the Cytoscape network visualization framework, including the Hanalyzer plugin for exploring knowledge networks in the context of experimental data, the MetaMapp plugin for identifying differentially regulated metabolites, and the Cerebral plugin for analyzing multiple experimental conditions on a graph with biological context. We also review stand-alone tools, including Chilibot for identifying relationships between genes and proteins in the biomedical literature, enRoute for context preserving pathway visualization and analysis, and a system for uncertainty-aware visual analysis of biochemical reaction networks. We describe the networks according to their node mappings, edge mappings, and analysis tasks they support.

4.21 GraphScape: Integrated Multivariate Network Visualization

Kai Xu (Middlesex University, GB)

We introduce a method, GraphScape, to visualize multivariate networks, i.e., graphs with multivariate data associated with their nodes. GraphScape adopts a landscape metaphor with network structure displayed on a 2D plane, and the surface height in the third dimension represents node attribute. More than one attribute can be visualized simultaneously by using multiple surfaces. In addition, GraphScape can be easily combined with existing methods to further increase the total number of attributes visualized. One of the major goals of GraphScape is to reveal multivariate graph clustering, which is based on both network structure and node attributes. This is achieved by a new layout algorithm and an innovative way of constructing attribute surfaces, which also allows visual clustering at different scales through interaction.

4.22 Top Challenges in Visualizing Multivariate Personal Networks

Michelle X. Zhou (IBM Almaden Center – San José, US)

Given the abundant digital footprints left by people, we are developing people analytics and engagement technologies to facilitate deeper understanding of people and to optimize individualized engagements (e.g., doctors-patients and marketers-consumers). To help users either gain a better understanding of themselves or others, we use a personal network to represent each individual. In a personal network, each individual is a multivariate node, of which traits are automatically derived from one's linguistic footprints, while each link represents the individual's social contact also automatically extracted and characterized from one's interaction and communication activities with others. Based on the characteristics of data, value, variety, veracity, velocity, volume, and user capabilities, I list a set of challenges in creating, exploring, and improving personal networks.

5 Working Groups

This section describes results from each of the six working groups and identifies the attendees contributing to each group. The names of those people who reported for the working groups (group leaders) are underlined.

5.1 WG: Temporal Multivariate Networks

James Abello, Daniel Archambault, Maura Conway, Stephan Diehl, Carsten Görg, Benjamin David Hennig, Jessie Kennedy, Stephen G. Kobourov, Lothar Krempel, Kwan-Liu Ma, Silvia Miksch, and Alexandru C. Telea

The breakout group aimed to characterize the temporal dimension in visualizing multivariate networks exemplified through the study of three application domains that were discussed in the seminar (biology, software engineering and social sciences). As a first step, we classified characteristic domain examples of networks into three different categories (cf. Figure 2): the *structure* describes arrangement and relationships between parts or elements (i.e., the network itself). *Behavior* means the action or reaction of something under given circumstances, and *evolution* is the gradual development of something over time. Based on this underlying model, the group members plan to structure their book chapter as follows: after a presentation of the various types of graphs for temporal networks, we try to specify typical tasks that usually have to be performed with temporal graphs. Then, a domain characterization will be discussed as already mentioned above. Finally, the group will do a survey/classification of existing visualization techniques for temporal multivariate networks followed by highlighting visualization challenges in this context.

	Biology	Software Engineering	Social Networks
Structure	Biological entity, e.g. gene and gene interactions	Modules and couplings	A twitter community network
Behaviour	Expression level	Program trace	Tweet, retweet, mention and follow activity
Evolution	Organism development / experimental conditions	Different versions of code	Changes in a twitter community structure

Figure 2 Domain examples classified according to structure, behavior and evolution.

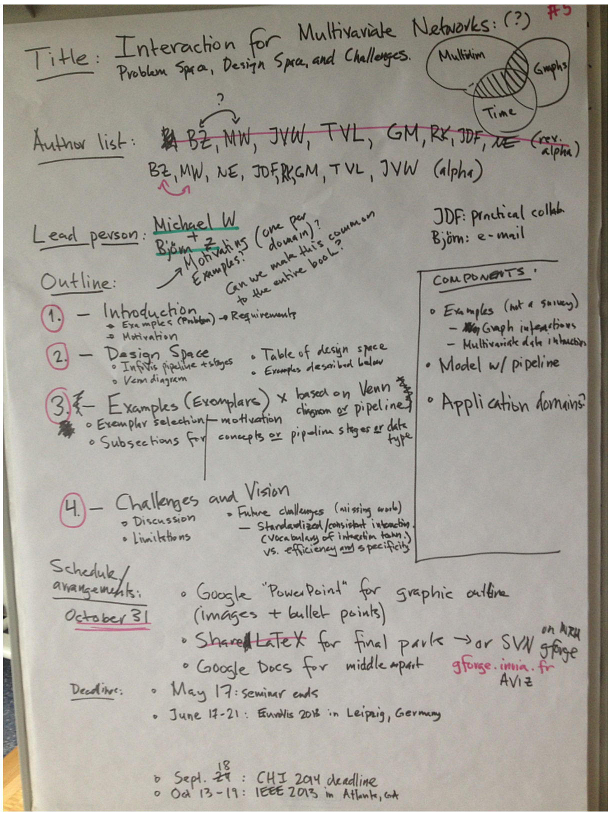

Figure 3 Possible structure of the corresponding "interaction" chapter.

5.2 WG: Interaction for Multivariate Networks

Niklas Elmqvist, Jean-Daniel Fekete, Robert Kosara, Guy Melançon, Jarke J. van Wijk, Tatiana von Landesberger, Michael Wybrow, and Björn Zimmer

The interaction group started out discussing existing interaction techniques for networks and multivariate data sets to order to get an overview of the field. The group examined in greater detail the problems and requirements that arise during the analysis process of graphs with multivariate attributes. It was decided to categorise interaction techniques by their place in the stages of the standard InfoVis pipeline (cf. Figure 3), i.e., whether they operate at data, visual encoding or view level. The group planned and discussed the overall structure and content for the interaction book chapter and decided how they would continue to work on it after the seminar. It was decided that the chapter would not just take the form of a survey of existing techniques, but rather be a guide to the use of interaction in the visualisation of multivariate networks; describing various popular interaction techniques, discussing application of novel approaches, and exploring the remaining challenges in this space.

5.3 WG: Multiple and Multi-domain Networks

Katy Börner, Hans Hagen, Andreas Kerren, and Falk Schreiber

Multiple networks and multi-domain networks occur in several application fields, where their integration, combination, comparison, and visualization is one of the big challenges. The working group started to analyze the general characteristics of this type of data and identified examples in the three application domains of the seminar: biology, social sciences, and software engineering. Figure 4 shows the outcome of this discussion. Conceptually, there are sets of multivariate networks at two or more levels. Each level may describe a specific scale, and within each level several related multimodal or heterogeneous networks are represented. To take an example from biology: within a level different molecular-biological networks such as metabolism, protein interaction and gene regulation networks may be integrated, across levels the scale could be from a molecular biological network to an evolutionary network. We allow $n : m$ mappings within the same level, but only $1 : n$ mappings across levels that must be consecutive. This leads to a structured data set that is the basis for further visual analyses. At the end of the seminar, the group proposed a set of example visual representations, discussed application domain examples, and presented the structure of the planned book chapter.

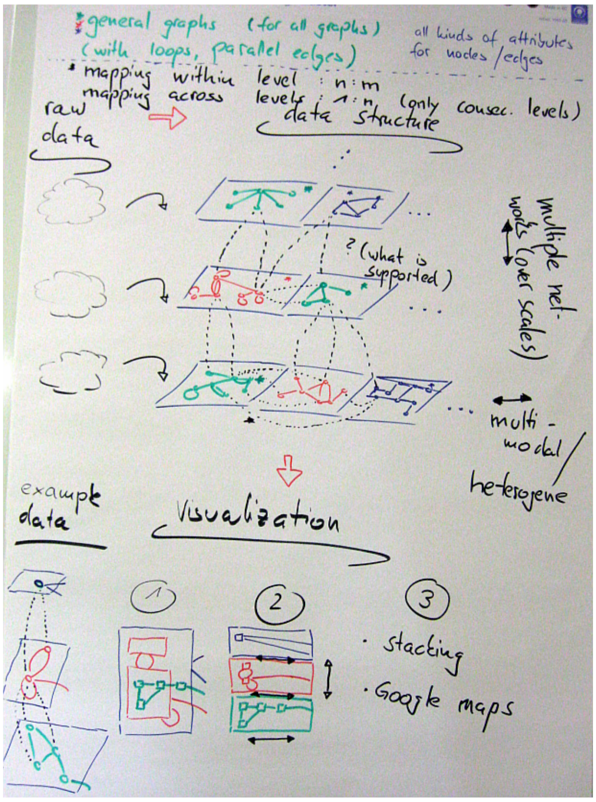

Figure 4 Sketch to understand the fundamental problem of analyzing multiple networks at different scale.

5.4 WG: Scalability of Multivariate Graph Visualization

Tim Dwyer, Danny Holten, Christophe Hurter, T. J. Jankun-Kelly, Martin Nöllenburg, and Kai Xu

The working group for Scalability worked under the mantra "Draw what you need when you need it". As a starting off point, we assumed that Shneiderman's visual seeking mantra needs a caveat: choose an overview that is appropriate to the task—i.e. rarely do users want to see the whole universe. Large (especially scale-free graphs) inevitably turn into hairballs—layout can help: but it happens sooner or later even to the best layout methods. Matrix views do not have the problem of edge crossings, but they have a limited ability to convey transitive structure, also the amount of display area they require grows in the square of the number of nodes. Thus, we end up with displays of hairballs or white noise. Multivariate structures just add further complication. Thus, to understand the issues of complexity in scalable multivariate graphs, we focused our discussion on five dimensions: the display resolution (2- vs. 3-dimensions), the limitations of visual acuity on dense graph perception, the related limitations of cognitive resources in understanding complex graphs, software architectural limitations on the processing very large graphs, and how other visual dimensions (such as color or over plotting) can affect our understanding. These dimensions will be used to build up suggestions for grounded approaches for effectively using large, multivariate graphs.

5.5 WG: Tasks for Multivariate Network Analysis

Peter Eades, Helen Gibson, Daniel A. Keim, A. Johannes Pretorius, Helen C. Purchase, and John T. Stasko

This workgroup set out to characterize the tasks associated with multivariate network analysis. A *task* was considered to be an activity that a user wishes to accomplish by means of interaction with a visual representation of a multivariate network.

All workgroup participants agreed that a distinction could be made between *low-level* tasks and *high-level* tasks. The group felt that this characterization is independent of the application domain. Low-level tasks are atomic interactive analysis activities. Examples include retrieving an attribute value of a node, or determining if nodes are related to each other. High-level tasks cater more specifically for the relational and multivariate nature of multivariate networks and are composed of low-level tasks. Examples include analyzing the overall structure of the network, or finding clusters of nodes that have specific attribute values.

Other issues discussed included interaction as a means of performing tasks (a topic we felt might best be left to the interaction group) and network comparisons (a topic that might best be included in the work of the temporal or multiple network groups).

Having thus defined the scope of our "task" chapter, we can draw on the existing literature that defines tasks for (non-multivariate) graphs and other visualization methods. This literature also often distinguishes between low- and high-level tasks. From our discussions it emerged that there may be another level of tasks somewhere between low- and high-level. However, this issue was not resolved as it proved to be very difficult to map to meaningful tasks of intermediary complexity, both from low-level and from high-level tasks.

5.6 WG: Novel Visual Metaphors for Multivariate Networks

Oliver Kohlbacher, Jonathan C. Roberts, Matthew O. Ward, Jing Yang, and Michelle X. Zhou

There are opportunities to display multivariate networks in different ways. While node-link representations are very common, there are a lot of other possibilities. We look for novel visual metaphors and visual interaction techniques for multivariate networks and discuss the data visual mapping process involving them. The metaphors are organized into nature inspired, geographical, non-geographical, man-made, non-physical, and traditional-visualization inspired classes. Besides exemplar work from each class, several case studies of prior work on novel metaphors are presented with in-depth discussions on advantages and limitations. Finally, a gallery of new metaphors is presented in the book chapter.

6 Open Problems in Multivariate Network Visualization

Not all the topics identified during the seminar could be addressed in the working groups (and will not be addressed in the planned book) and might be considered for a future Dagstuhl seminar. They include the following:

- Data quality: How can we represent uncertainty, noise, errors, missing data?
- Dimensions: How can we best characterize/reduce the high-dimensional variable space in multivariate networks?
- Evaluation and Experimental Design: What tasks do scientists perform with visualizations? How do we measure understanding of such visualizations? What is the role of aesthetics?
- Existing Practices: How can we merge the best practices of graph drawing and information visualization?
- Scalability: How can we make visualizations of large data sets meaningful?
- Selective Visualization: What should be shown and when?
- Faithfulness in Visualization: How can we ensure the validity of visualizations?
- Toolkits and Standards: What progress has been made to date on achieving toolkits and standard data/information representations?
- Challenges for the Next Ten Years: What challenges can be solved and what challenges do we expect to come up during the next ten years?

7 Acknowledgments

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 of organizing this event. Björn Zimmer gathered the abstracts for the overview of talks in Sect. 4 of this document and the slides of all presenters. These slides can be found on the materials website of the seminar. In addition, many attendees agreed to take notes during the breakout sessions. These notes were the basis for writing this executive summary (incl. Sect. 5) and are also available for download on the Dagstuhl web pages of the seminar. Last but not least, the seminar would not have been possible without the great help of the staff at Dagstuhl Castle. We acknowledge all of them and their assistance.

Participants

- James Abello
Rutgers Univ. – Piscataway, US
- Daniel Archambault
Swansea University, GB
- Katy Börner
Indiana University – Bloomington, US
- Maura Conway
Dublin City University, IE
- Stephan Diehl
Universität Trier, DE
- Tim Dwyer
Monash Univ. Melbourne, AU
- Peter Eades
The University of Sydney, AU
- Niklas Elmqvist
Purdue University, US
- Jean-Daniel Fekete
INRIA Saclay – Île-de-France – Orsay, FR
- Helen Gibson
University of Northumbria, GB
- Carsten Görg
University of Colorado, US
- Hans Hagen
TU Kaiserslautern, DE
- Benjamin David Hennig
University of Sheffield, GB
- Danny Holten
SynerScope BV, NL
- Christophe Hurter
ENAC – Toulouse, FR
- T. J. Jankun-Kelly
Mississippi State University, US
- Daniel A. Keim
Universität Konstanz, DE
- Jessie Kennedy
Edinburgh Napier University, GB
- Andreas Kerren
Linnaeus University – Växjö, SE
- Stephen G. Kobourov
Univ. of Arizona – Tucson, US
- Oliver Kohlbacher
Universität Tübingen, DE
- Robert Kosara
Tableau Software – Seattle, US
- Lothar Krempel
MPI für Gesellschaftsforschung – Köln, DE
- Kwan-Liu Ma
Univ. of California – Davis, US
- Guy Melançon
University of Bordeaux, FR
- Silvia Miksch
TU Wien, AT
- Martin Nöllenburg
KIT – Karlsruhe Institute of Technology, DE
- A. Johannes Pretorius
University of Leeds, GB
- Helen C. Purchase
University of Glasgow, GB
- Jonathan C. Roberts
Bangor University, GB
- Falk Schreiber
IPK Gatersleben & MLU Halle, DE
- John T. Stasko
Georgia Inst. of Technology, US
- Alexandru C. Telea
University of Groningen, NL
- Jarke J. van Wijk
TU Eindhoven, NL
- Tatiana von Landesberger
TU Darmstadt, DE
- Matthew O. Ward
Worcester Polytechnic Inst., US
- Michael Wybrow
Monash Univ. Melbourne, AU
- Kai Xu
Middlesex University, GB
- Jing Yang
University of North Carolina at Charlotte, US
- Michelle X. Zhou
IBM Almaden Center – San José, US
- Björn Zimmer
Linnaeus University – Växjö, SE