

You are here: Program » Seminar Calendar » Seminar Homepage

Seminar Calendar[All Events](#)[Dagstuhl Seminars](#)[Dagstuhl Perspectives](#)[GI-Dagstuhl Seminars](#)[Summer Schools](#)[Events](#)[Research Guests](#)[Expenses](#)[Planning your visit](#)<https://www.dagstuhl.de/21152>**April 11 – 16 , 2021, Dagstuhl Seminar 21152**

Multi-Level Graph Representation for Big Data Arising in Science Mapping

Organizers

Katy Börner (Indiana University – Bloomington, US)

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



1 / 4

For support, please contact

[✉ Dagstuhl Service Team](#)

Documents

Dagstuhl Report, Volume 11, Issue 3 [Aims & Scope](#)[List of Participants](#)[Shared Documents](#)[Dagstuhl Seminar Schedule \[pdf\]](#)

Places & Spaces: Mapping Science

Drawing from across cultures and across scholarly disciplines, the Places & Spaces: Mapping Science traveling exhibit demonstrates the power of maps (visualizations) to address vital questions about the contours and content of scientific knowledge.

[Read more and view samples...](#)

Summary

Networks are all around us. At any moment in time we are driven by, and are an integral part of, many interconnected and dynamically changing networks. Our species has evolved in the context of diverse ecological, biological, social, and other types of networks over thousands of years. We created telephone and power networks, road and airline networks, the Internet and the World Wide Web. We model biological processes with metabolic and protein networks, fake news and rumors with epidemic networks, and even the brain with neural networks. The study of such large-scale networks is one of the prominent generators of big data. Analyzing, exploring, and understanding these complex, interdependent, multi-level networks requires new, more efficient,

Documentation

In the series Dagstuhl Reports each Dagstuhl Seminar and Dagstuhl Perspectives Workshop is documented. The seminar organizers, in cooperation with the collector, prepare a report that includes contributions from the participants' talks together with a summary of the seminar.

Download [📄 overview leaflet \(PDF\)](#).

Publications

Furthermore, a comprehensive peer-reviewed collection of research papers can be published in the series Dagstuhl Follow-Ups.

Dagstuhl's Impact

Please inform us when a publication was published as a result from your seminar. These publications are listed in the category Dagstuhl's Impact and are presented on a special shelf on the ground floor of the library.

and more intuitive graph analysis and visualization approaches. This is confirmed in a 2017 VLDB study by Sahu et al. about how graphs are used in practice, where analysts identified scalability and visualization as the most important issues to address.

Recent advances in data, algorithms, and computing infrastructures make it possible to map humankind's collective scholarly knowledge and technology expertise by using topic maps on which "continents" represent major areas of science (e.g., mathematics, physics, or medicine) and zooming reveals successively more detailed subareas. Basemaps of science and technology (S&T) are generated by analyzing citations links between millions of publications and/or patents. "Data overlays" (e.g., showing all publications by one scholar, institution, or country or the career trajectory of a scholar as a pathway) are generated by science-locating relevant publication records based on topical similarity. Science maps are widely used to compare expertise profiles, to understand career trajectories, and to communicate emerging areas. The recent National Academy of Science Colloquium on *Modeling and Visualizing Science Developments* co-organized by Börner showcased the utility of predictive modeling and large-scale mapping efforts, e.g., in support of ranking institutions, analyzing job market developments, and innovation diffusion and technology adoption.

Despite the demonstrated utility of large-scale S&T maps, current approaches do not scale to the hundreds of millions of data records now available. Most users have a hard time reading large-scale networks and few can traverse or derive knowledge from multi-level presentations of networks. Most maps of science, technology, or jobs data support exactly one level of detail. Very few even support two levels such as the UCSD map of science. A key challenge is designing efficient and effective methods to visualize and interact with more than 100 million scholarly publications at multiple levels of resolution.

Given results from our prior studies on the effectiveness and memorability of map-like visualization of large graphs, we are interested to bring together leading experts to design a multi-level, large-scale map of science that can be used by experts and the general public alike.

The notion of multiple-levels-of-detail graph representation can be captured with Multi-Level Graph Sketches (MLGS) that take a static map-like representation to a multi-level setting needed for exploring and interacting with large, real-world networks. Using the familiar map metaphor, multi-level graph algorithms can make it possible to identify important nodes, major pathways, and clusters across multiple levels. Specifically, we aim to develop efficient algorithms with theoretical and practical guarantees for creating Multi-Level Graph Sketches (MLGS) in support of visual analytics tasks for large network exploration, navigation, and communication. Unlike existing methods for visualizing multi-level networks based on meta-nodes and meta-edges, the MLGS approach can provide real nodes (prototypes) and real paths (backbones) for each level, similar to geographic maps that show real cities and real roads at every level of detail.

Research questions included: (1) designing efficient algorithms for MLGS: advance the state-of-the-art in graph algorithms by generalizing the notion of graph spanners to multiple levels. (2) utilizing MLGS algorithms in visualization: applying the MLGS representation in the context of network analysis and visualization for interacting with large networks, which combines the MLGS approach with clustering, layout, and map-like visualization. (3) developing a new approach for science classification, lookup, and topical mapping service in support of data-driven decision making by students, teachers, and

administrators. (4) validating the new algorithms and visualizations: evaluation based on quantitative metrics such as efficiency/scalability, stress/distortion and precision/recall, as well as qualitative metrics based on human subject studies of the utility of visualizations along with readability, engagement, and memorability.

The main goal of this seminar was to bring together researchers coming from information visualization, psychology, cognitive science, human-computer interaction, graph drawing, computational geometry, cartography, and GIS with interests in "science of science" to discuss novel graph mining and layout algorithms and their application to the development of science mapping standards and services.

Due to the pandemic, we had a "hybrid" seminar with only 5 in-person participants and 25 by-zoom participants; see Fig. 1. With participants from more than 10 different countries and at least 5 different time-zones (some as much as 9 hours behind Central Dagstuhl time), this was a new experience for most of us and definitely different from previous Dagstuhl seminars. Nevertheless, we attempted to adapt by having additional evening events and moving the traditional Wednesday excursion to the morning.

As this was a highly interdisciplinary seminar, we started the event with talks that introduced the state of the art in the participating fields on a high level. After the introductory presentations, we presented our initial research problems and discussed further questions in an open-problem session.

A set of 4-6 research problems was then finalized and the formation of working groups for these problems was completed by the end of the second day. The remaining three days were dedicated to working-group meetings, progress reports, and initial write-ups.

The main expected outcome of this seminar will be a special issue of the journal IEEE Computer Graphics and Applications on the main topics of the seminar. Specifically, we expect 4-6 research papers on the problems discussed by the working groups. Longer term goals include: improved collaborations and communications between the different communities brought together for this seminar, improved maps of science for SciMap2020, and new multi-level graph algorithms and approaches.

Summary text license

© Creative Commons BY 4.0

© Katy Börner and Stephen G. Kobourov

Classification

- Data Structures And Algorithms
- Human-Computer Interaction
- Social And Information Networks

Keywords

- Science of science
- Multi-level graph algorithms
- Network visualization