GraphXD Workshop 2018

https://graphxd.github.io/workshop/2018.html

Berkeley Institute for Data Science (BIDS)
Doe Memorial Library, UC Berkeley

March 27–29, 2018

Background

Graphs arise in many fields. A sociologist interested in social interactions in a community may record their observations in a graph of individuals connected by friendship. A geneticist interested in how genes are co-expressed may represent their data as a graph of genes with edges connecting genes that are expressed together. An online airline purchasing system may use a graph of airports connected by edges with direct flights between them.

Progress in understanding graphs and developing new graph algorithms in a number of diverse fields is hindered by the fact that researchers who use them typically don't have the opportunity to communicate with others who work on similar problems in different domains.

What is GraphXD?

Graphs and Complex Networks Across Domains (GraphXD) connects scientists, researchers, and theorists interested in graphs from a variety of fields through:

1. **Seminars**

   (09/28/2017) "Graph Clustering Algorithms" Tselil Schramm (Simons Institute, UC Berkeley)

   (10/19/2017) "Spectral Sparsification of Graphs" Nikhil Srivastava (Mathematics, UC Berkeley)

   (11/30/2017) "Data-Driven Methods for Learning Sparse Graphical Models" Somayeh Sojoudi (EECS and Mechanical Engineering, UC Berkeley)

   (02/26/2018) "Vector Representations of Graphs and the Maximum Cut Problem" David P. Williamson (Operations Research & Information Engineering, Cornell University)

2. **Workshops**

   This is the inaugural workshop.

Participants

- Katelyn Arnemann (Neuroscience, UC Berkeley)
- Carl Boettiger (Environmental Science, Policy, and Management, UC Berkeley)
- Seth Bromberger (Critical Infrastructure Security, Lawrence Livermore National Laboratory)
- Aydın Buluç (Computational Research Division, Lawrence Berkeley National Laboratory)
- Thomas Caswell (National Synchrotron Light Source II, Brookhaven National Laboratory)
- Rob Fatland (Research Computing, University of Washington)
- Jessica Forde (Jupyter)
- Kimon Fountoulakis (Statistics and ICSI, UC Berkeley)
- Marília Palumbo Gaiarsa (Entomology, UC Riverside)
- Jason Grout (Scientific Software, Bloomberg)
- Aric Hagberg (Theoretical Division, Los Alamos National Laboratory)
- Chris Holdgraf (Berkeley Institute of Data Science, UC Berkeley)
- Paul Ivanov (Scientific Software, Bloomberg)
- Caigui Jiang (ICSI Vision Group, UC Berkeley)
- Rasmus Kyng (Theory of Computation, Harvard University)
- Rachel Lawrence (Computer Science, UC Berkeley)
- Mike MacNeil (Data Analysis and Visualization, Lawrence Berkeley National Laboratory)
- Charlotte Mazel-Cabasse (Berkeley Institute of Data Science, UC Berkeley)
- Jarrod Millman (Biostatistics, UC Berkeley)
- Dmitriy Morozov (Data Analysis and Visualization, Lawrence Berkeley National Laboratory)
- Rajarshi Mukherjee (Biostatistics, UC Berkeley)
- Soumendu Mukherjee (Statistics, UC Berkeley)
- Russell Neches (Biochemistry & Biophysics, UC San Francisco)
- M Pacer (Berkeley Institute of Data Science, UC Berkeley)
- Lauren Ponisio (Entomology, UC Riverside)
- Nick Ryder (Mathematics, UC Berkeley)
- Aaron Schild (Computer Science, UC Berkeley)
- Ludwig Schmidt (Computer Science, UC Berkeley)
- Dan Schult (Mathematics, Colgate University)
- Camille Scott (Computer Science, UC Davis)
- Deborah Sunter (Energy and Resources, UC Berkeley)
- Daniela Ushizima (Computational Research Division, Lawrence Berkeley National Laboratory)
- Stéfan van der Walt (Berkeley Institute of Data Science, UC Berkeley)
- Maryam Vareth (Radiology and Biomedical Imaging, UC San Francisco)
- Nelle Varoquaux (Statistics, UC Berkeley)
- Chris Warner (Biophysics, UC Berkeley)
- Brian Xia (Artificial and Human Intelligence, Siemens)
- Stella Yu (Computer Science, Vision Science, and ICSI Vision Group, UC Berkeley)

Agenda

Tuesday, March 27

Time	Activity
9:00–9:30	**Breakfast**
9:30–10:00	K. Jarrod Millman, Graphs and complex networks across domains
10:00–10:30	Lauren Ponisio, Understanding the ecology and evolution of communities through networks, part I
10:30–11:00	Marília Gaiarsa, Understanding the ecology and evolution of communities through networks, part II
11:00–11:30	**Tea**
11:30–12:00	Nick Ryder, A history of spectral graph theory and its applications, part I
12:00–12:30	Aaron Schild, A history of spectral graph theory and its applications, part II
12:30–2:00	**Lunch**
2:00–3:00	Aric Hagberg, Exploring network structure, dynamics, and function using NetworkX
3:00–3:30	Aydın Buluç, Graph abstractions in computational genomics
3:30–4:00	**Tea**
4:00–4:30	Katelyn Arnemann, Challenges for graph theory in human neuroscience
4:30–5:00	Kimon Fountoulakis, Variational perspective on local graph clustering
5:00–7:00	**BBQ** (near Brown's California Cafe)

Wednesday, March 28

Time	Activity
9:00–9:30	**Breakfast**
9:30–10:00	Camille Scott, Sequence assembly graphs and their construction
10:00–11:00	Rasmus Kyng, How to solve problems on graphs using linear equations, and how to solve linear equations using graphs
11:00–11:30	**Tea**
11:30–12:00	Ludwig Schmidt, Linear regression with graph constraints
12:00–12:30	Discussion
12:30–2:00	**Lunch**
2:00–3:00	Planning
3:00–5:00	Self-organized activities

Thursday, March 29

Time	Activity
9:00–9:30	**Breakfast**
9:30–12:30	Self-organized activities
12:30–1:30	**Lunch**
1:30–3:00	Self-organized activities
3:00–4:00	Discussion
4:00–5:00	Reflection

Questions

Reflection

GraphXD aims to:

- foster a community of shared interest from disparate fields including mathematics, computer science, statistics, physics, biology, sociology, and any other field with an interest in graph data;

- develop a shared vocabulary and identify common principles, algorithms, and tools for understanding graphs; and

- learn from one another while strengthening ties across disciplinary boundaries.

Currently, the workshop is our main avenue for pursuing these aims. Since this is our inaugural workshop, we would like you to help us figure out whether this is the right approach (or even if these aims are worth pursuing).

During the workshop, we would like you to think about whether the lack of communication among researchers from diverse fields interested in graphs hinders progress. If so, in what ways. Do you have specific examples? What would it look like if there was better communication? Could a workshop like this help? Would other approaches work better? How can we use our time at the workshop to acheive a large impact? Did we invite the right people? Who should we invite (e.g., specific individuals, individuals from specific domains)? Should events like this be stand alone or should they be organized in conjunction with existing conferences / events? If so, which? Is there a role for education (e.g., would new classes help)? Should we try to organize a group to meet regularly? If so, how should we do this?

At the end of the workshop, you will be asked to submit four to five paragraphs reflecting on your thoughts and experience at the workshop.

Specific questions

Here are some more specific questions you should reflect on during the workshop.

- With an eye towards improving future GraphXD workshops, what changes would you recommend?

- Imagine a version of GraphXD designed from scratch and executed perfectly; what goals should we hope to achieve?

- What outputs (e.g., whitepapers, journal articles, books) should we consider targeting in higher-level areas such as education, reproducibility, data sharing, and software?

- What did you hope to get out of the workshop? Were those wishes realized? If not, why not?

- Describe one thing you enjoyed learning at the workshop.

- How could we improve interactions at the workshop? E.g., is the talk/discussions/self-organized activity balance effective? where there too few or too many breaks?

- What did you do during the self-organized activity portion of the workshop?

- Did you start any new projects or meet any new collaborators at the event? If not, how could we make it easier for you to do so.