PROGRAM

Communities in Networks

Comnets @ Netsci 2022

https://hocinecherifi.wixsite.com/comnets2022

JULY 11, 2022

Time (EDT)	Authors	Title
8:15 - 8:30		***Opening***
8:30 - 9:15	*Santo Fortunato*	*Communities in networks: modularity and embeddings*
9:15 - 10:00	*Vincent Traag*	*Mapping the landscape of science using community detection*
10:00- 10:20		***Cofffee break***
10:20 - 10:40	*Sang Hoon Lee*	*Refinement for community structures of bipartite networks*
10:40 - 11:00	*Bojan Evkoski, Igor Mozetic and Petra Kralj Novak*	*Extending LFR - The Unconstrained Friedman-Nemenyi Benchmark*
11:00 - 11:20	*Konstantin Avrachenkov, Maximilien Dreveton and Lasse Leskelä*	*Recovering communities in temporal networks using persistent edges*
11:20 - 11:40	*Zhen Su, Henning Meyerhenke and Jürgen Kurths*	*The climatic interdependence of extreme-rainfall events around the globe: combining consensus clustering and mutual correspondences*
11:40 - 12:00	*Stephany Rajeh and Hocine Cherifi*	*Extracting the Modular Backbone of Weighted Complex Networks*

KEYNOTE: Santo Fortunato Indiana University Bloomington USA
Communities in networks: modularity and embeddings

A central question in network community detection is to establish how modular a network is. The modularity by Newman and Girvan addresses that but it has well-known shortcomings. We propose a measure based on the concept of robustness, that avoids such problems by design: modularity is the probability to find trivial partitions when the structure of the network is randomly perturbed. Also, I will discuss the effectiveness of graph embeddings at discovering community structure. I especially focus on techniques based on random walks, like the popular node2vec, and show their similarities with spectral clustering techniques. This allows to derive results on the detectability limits of such techniques, and suggests ways to improve their performance.

KEYNOTE: Vincent Traag Leiden University The Netherlands
Mapping the landscape of science using community detection

The scientific literature continues to grow at astounding rates, amounting to millions of publications each year. Making sense of this large corpus is challenging. Although existing classifications and ontologies of science help to navigate the literature, they show various problems. One problem is that such classifications are often based on journals, which is increasingly problematic with mega-journals such as PLOS ONE and Scientific Reports. Another problem is that such classifications are often not sufficiently granular. Applying community detection in citation networks helps to build more fine-grained classifications of science that transcend traditional disciplinary and journal boundaries. I will present an algorithm we use to classify up to 60 million publications. The resulting classifications are useful for better understanding the landscape of science, and I will showcase various applications. At the same time, such algorithmically derived classifications come with certain challenges, and I will discuss some of them.

JULY 12, 2022

Time (EDT)	Authors	Title
8:15 - 8:30		***Opening***
08:30 - 9:15	*Peter J. Mucha*	*Finite-state parameter space maps for pruning partitions in modularity-based community detection*
09:15 - 10:00	*Aaron Clauset*	*Evaluating overfit and underfit in models of network community structure*
10:00- 10:10		***Coffee break***
10:10 - 10:30	*Gáspár Sámuel Balogh, Bianka Kovács and Gergely Palla*	*Extremely modular structure of growing hyperbolic networks*
10:30 - 10:50	*Stephen Eubank, Ritwick Mishra, Madhurima Nath and Abhijin Adiga*	*Communities in Directed Weighted Food Networks using Moore-Shannon Network Reliability*
10:50 - 11:10	*Antonio Lopolito, Andrea Nigri, Rocco Caferra and Piergiuseppe Morone*	*Rethinking interactions: managing social distances mitigating health and economic consequences during COVID-19*
11:30 - 11:50	*Francesco De Nicolò, Alfonso Monaco, Giuseppe Ambrosio,*	*Territorial Developent as an innovation driver: a complex network approach*
11:10 - 11:30	*Issa Moussa Diop, Cherif Diallo, Chantal Cherifi and Hocine Cherifi*	*The component structure of complex networks: a case study*
11:50 - 12:00		***Closing***

KEYNOTE: Peter J. Mucha Department of Mathematics, Dartmouth College, USA
Finite-state parameter space maps for pruning partitions in modularity-based community detection

Numerous software packages are available and widely used for community detection, but many of these require parameters to be selected (or assume default values) that are not always obvious to application domain experts. For example, the best use of modularity-based methods includes setting a parameter to control the resolution. Moreover, most of the algorithms are pseudo-random heuristic approximations. As such, one frequently needs to reconcile numerous different partitions of nodes into communities while simultaneously exploring the parameter space. These problems are exacerbated when community detection is extended to multilayer networks, because of the addition of at least one parameter to specify the coupling between layers. To address these difficulties, we combine recent developments into a simple framework for pruning a set of partitions to a subset that are self-consistent by an equivalence with stochastic block model (SBM) inference.Specifically, by first restricting attention to a subset of partitions obtained by intersecting half-spaces, iterative procedures proposed by Newman (2016) and by Pamfil et al. (2019) yield a finite-state map on this restricted subset. Implementing these pruning steps highlights only a small number of "stable" (fixed point) partitions in our examples. Our framework works for single networks and multilayer networks, as well as for restricting to a fixed number of communities when desired. We also derive resolution parameter upper bounds for fitting a constrained SBM of K blocks and demonstrate that these bounds hold in practice, further guiding parameter space regions to consider. Our code for implementing these procedures is available at http://github.com/ragibson/ModularityPruning.

KEYNOTE: Aaron Clauset The University of Colorado at Boulder, USA
Evaluating overfit and underfit in models of network community structure

Community detection is a network-analog of clustering in vector data: it seeks a typically unsupervised decomposition of a network into groups based on statistical regularities in network connectivity. Although we now have a dizzying number of algorithms for extracting such decompositions, the 'No Free Lunch' theorem for community detection implies that no algorithm can be optimal across all inputs, and therefore all algorithms make tradeoffs in how they decompose a network into communities. Despite 20 years of work, we still know little in practice about how different algorithms over- or under-fit when applied to real networks, and we lack good tools for reliably assessing and comparing such behavior across algorithms. In this talk, I'll describe a broad assessment of 16 state-of-the-art community detection algorithms applied to a novel benchmark corpus of 572 structurally diverse real-world networks, showing that (i) algorithms vary widely in the number

and composition of communities they find when given the same input; (ii) algorithms themselves can be clustered into distinct groups based on the similarities of their outputs when applied to real-world networks; and (iii) algorithmic differences induce wide variation in accuracy on link-based learning tasks. We argue that two such tasks, link prediction and a new task we call 'link description', can be used to assess a bias-variance-like tradeoff to assess an algorithms general behavior, which we use to quantify each algorithm's overall tendency to over- or under-fit to network data. We close with a brief discussion of future directions in community detection.