
Graph Theory and its Applications
Théorie des graphes et ses applications

(Org: **Sinan Aksoy** (Pacific Northwest National Laboratory), **Nicole Eikmeier** (Grinnell College) and/et **Stephen J. Young**
(Pacific Northwest National Laboratory))

ILYA AMBURG, PNNL

Localization in harmonic vectors of simplicial complexes

Simplicial complexes allow for modeling higher-order interactions beyond the standard pairwise graph paradigm, and naturally lend themselves to analysis of edge flows through the spectra of associated higher-order Laplacians. However, theoretical foundations and principled tools to justify and unify the various methods that have been introduced to study harmonic flows are sorely missing. In particular, at the heart of many edge signal analysis algorithms lie implicit assumptions regarding localization of harmonic vectors around “holes” in the complex. Here, we begin the principled study of localization properties of harmonic vectors. We provide theoretical insights for why localization of harmonics is expected given some assumptions about the underlying simplicial complex, and demonstrate that it occurs in empirical networks from a wide range of disciplines. We leverage these results to construct a hole localization algorithm that performs many orders of magnitude faster than state-of-the-art algorithms while yielding solutions with more desirable properties. The resulting localized harmonics also provide rich features, which we demonstrate increase performance in various machine learning tasks.

PHILIP CHODROW, Department of Mathematics, UCLA

Nonbacktracking Spectral Clustering of Nonuniform Hypergraphs

Spectral methods offer a tractable, global framework for clustering in graphs via eigenvector computations on graph matrices. Hypergraph data, in which entities interact on edges of arbitrary size, poses challenges for matrix representations and therefore for spectral clustering methods. Some spectral methods exist for hypergraphs in which all edges have the same size, but many data sets are nonuniform and violate this restriction. We study spectral clustering for nonuniform hypergraphs based on the hypergraph nonbacktracking operator. After reviewing the definition of this operator and its basic properties, we prove a theorem of Ihara-Bass type to enable faster computation of eigenpairs. We then propose an approximate coordinate ascent algorithm for inference in a hypergraph stochastic blockmodel via linearized belief-propagation. Based on this algorithm, we pose a conjecture about detectability thresholds in nonuniform hypergraph stochastic blockmodels. We perform experiments in real and synthetic data that underscore the benefits of hypergraph methods (over graph-based methods) when interactions of different sizes carry different information about cluster structure.

Joint work with Jamie Haddock (Harvey Mudd College) and Nicole Eikmeier (Grinnell College)

NIKITA DENISKIN, Scuola Normale Superiore

Vertex distinction and interlacing with subgraph centrality

Centrality measures are often used to determine the most important vertices of a graph. We focus on subgraph centrality, which is based on the matrix exponential of βA , where β is a parameter and A is the adjacency matrix of the graph.

Our objective is to study when two different vertices have the same subgraph centrality, and their interlacing values, i.e. how many times their relative ranking reverses as β varies from 0 to ∞ . We show a recent proof of Estrada’s conjecture, which is related to cospectral vertices and walk-regular graphs; we also show some bounds on the number of interlacing values for two vertices.

This is a joint work with Michele Benzi (Scuola Normale Superiore).

CHRISTIAN GAETZ, Harvard University

The hull metric, linear extensions, and Cayley graphs of Coxeter groups

A subset C of the vertices of a graph G is called *convex* if any vertex on a shortest path in G between two elements of C also lies in C . When G is the Cayley graph of the symmetric group for the generating set of simple transpositions, convex sets are in bijection with labeled posets, and the size of the convex set is the number of linear extensions of the poset, an important quantity for many applications. We reinterpret an inequality of Sidorenko for linear extensions as a triangle inequality for a “hull metric” on this graph and give new proofs. The existence of this hull metric seems to be quite rare for general graphs, but we conjecture it holds for the standard Cayley graphs of arbitrary Coxeter groups, and prove several special cases. Joint work with Yibo Gao.

REBECCA JONES, Brigham Young University
Predicting Roles in Dynamic Social Networks

In social networks, role detection is the process of identifying people who have similar types of connections or certain importance in the network. For example, a social media network might have influencers who can influence a large number of people, authorities who are experts on their subjects, hubs who point to authorities, and loners, people who have very few connections. Using Non-Negative Matrix Factorization (NMF), we can identify and describe these roles in a network based off of structural properties of the network as well as attributes of the nodes. While NMF has been applied in various forms to dynamic social networks, little work has been done in making predictions about roles. In this talk, I'll discuss how NMF and machine learning techniques can be applied to dynamic social networks to make predictions about how roles change.

MARI KAWAKATSU, Princeton University
Emergence of hierarchy in networked endorsement dynamics

Many social and biological systems are characterized by enduring hierarchies, including those organized around prestige in academia, dominance in animal groups, and desirability in online dating. Despite their ubiquity, the general mechanisms that explain the creation and endurance of such hierarchies are not well understood. We introduce a generative model for the dynamics of hierarchies using time-varying networks, in which new links are formed based on the preferences of nodes in the current network and old links are forgotten over time. The model produces a range of hierarchical structures, ranging from egalitarianism to bistable hierarchies, and we derive critical points that separate these regimes in the limit of long system memory. Importantly, our model supports statistical inference, allowing for a principled comparison of generative mechanisms using data. We apply the model to study hierarchical structures in empirical data on hiring patterns among mathematicians, dominance relations among parakeets, and friendships among members of a fraternity, observing several persistent patterns as well as interpretable differences in the generative mechanisms favored by each. Our work contributes to the growing literature on statistically grounded models of time-varying networks.

NICHOLAS LANDRY, University of Colorado Boulder
Hypergraph assortativity: A dynamical systems perspective

The largest eigenvalue of the matrix describing a network's contact structure is often important in predicting the behavior of dynamical processes. We extend this notion to hypergraphs and motivate the importance of an analogous eigenvalue, the *expansion eigenvalue*, for hypergraph dynamical processes. Using a mean-field approach, we derive an approximation to the expansion eigenvalue and its associated eigenvector in terms of the degree sequence for uncorrelated hypergraphs. We introduce a generative model for hypergraphs that includes degree assortativity, and use a perturbation approach to derive an approximation to the expansion eigenvalue and its corresponding eigenvector for assortative hypergraphs. We validate our results with both synthetic and empirical datasets. We define the *dynamical assortativity*, a dynamically sensible definition of assortativity for uniform hypergraphs, and describe how reducing the dynamical assortativity of hypergraphs through preferential rewiring can extinguish epidemics.

DINA MISTRY,
Modeling the spread and impact of COVID-19 in contact networks

In the early phase of the COVID-19 pandemic, efforts to mitigate and control the spread of the virus focused on non-pharmaceutical interventions (NPI) while research into vaccine development and anti-viral treatments were ongoing. The most common NPI were physical distancing and the closure of in-person schools and workplaces where possible. However, these interventions are not without their cost. Targeted interventions such as the closure of schools results in lost in-person school days for children. Mechanistic models in networks can be used to model the potential impact of targeted strategies while exploring different transmission scenarios. Here we performed analysis using Covasim (covasim.org), an open-source agent-based model of disease transmission in networks to investigate the dynamics of the COVID-19 pandemic in Seattle, Washington and explore the impact of diagnostic testing, scheduling, and structural interventions to mitigate transmission in schools and the greater community. Complex models aimed at evaluating targeted interventions also call for increased realism and accuracy of the descriptions of how people are in contact in their everyday lives. We use data-driven multi-layered networks generated by the SynthPops model (synthpops.org) to model the population of Seattle with realistic age distributions, age mixing patterns, and network structure. We found that no scenario of reopening schools is without risk of transmission or many days spent learning at home for children, however a focus on in-person learning for children in the youngest age groups presented the lowest risk scenario, provided sufficient countermeasures are implemented in schools such as diagnostic testing and screening.

ESMERALDA NĂSTASE, Xavier University
Graphs, Codes, and Compressed Sensing

The Interval-passing algorithm (IPA) used in compressed sensing is an iterative process that is used to recover a k -sparse signal $\mathbf{x} \in \mathbb{R}^n$ from a linear measurement vector \mathbf{y} where $\mathbf{y} = H\mathbf{x}$. The matrix H is called the measurement matrix. Similar to the iterative decoder used in decoding low-density parity-check (LDPC) codes, the IPA may be modeled by a bipartite graph called a Tanner graph. Yakimenka and Rosnes showed that graphical substructures called termatiko sets characterize when the IPA fails; the size of the smallest termatiko set is called the termatiko distance. In this talk, we present new results on the structure of different types of termatiko sets as well as bounds on the sizes of termatiko sets that exist in the graphs corresponding to certain classes of measurement matrices. This work gives new insight to designing good graphs (and corresponding measurement matrices) for compressed sensing.

GIOVANNI PETRI, ISI Foundation
Social contagion and norm emergence on simplicial complexes and hypergraphs

Complex networks have been successfully used to describe dynamical processes of social and biological importance. Two classic examples are the spread of diseases and the emergence of shared norms in populations of networked interacting individuals. However, pairwise interactions are often not enough to fully characterize contagion or coordination processes, where influence and reinforcement are at work. Here we present recent results on the higher-order generalization of the SIS process and of the naming game. First, we numerically show that a higher-order contagion model displays novel phenomena, such as a discontinuous transition induced by higher-order interactions. We show analytically that the transition is discontinuous and that a bistable region appears where healthy and endemic states co-exist. Our results help explain why critical masses are required to initiate social changes and contribute to the understanding of higher-order interactions in complex systems. We then turn to the naming game as a prototypical example of norm emergence and show that higher-order interactions can create interesting novel phenomenologies, for example, they can explain how –when communication among agents is inefficient– even very small committed minorities are able to bring the system to a tipping point and flip the majority in the system. We conclude with an outlook on higher-order models, posing new questions and paving the way for modelling dynamical processes on these networks.

JENNA POPE, Pacific Northwest National Laboratory
Geometric learning for sampling atomic conformations

A material's atomic conformation dictates the resulting electronic and mechanical properties; therefore, determining the experimentally realizable conformation is a key task in molecular modeling. Ab initio approaches allow one to compute the relaxed conformation with high accuracy, albeit at high computational cost. Alternatively, high-throughput sampling methods

allow one to quickly explore large areas of chemical space, but at reduced accuracy. Traditional Monte Carlo approaches to sampling atomic conformations rely on predefined structural perturbations of a structure's 3D point cloud, such as rotations around specific bonds or translations of specific atoms. However, in material systems under periodic boundary conditions, ideal sampling perturbations are not well defined as most material systems do not have covalent bonds around which to rotate and atom translations lead to cascading atomic motions across the entire material. Therefore, we redefine the problem to sample graphs, rather than 3D point clouds. We designed a graph neural network (GNN) to sample atomic conformations learned from a training set of relevant structures in lieu of predefined perturbations. This talk will discuss the definition of material graphs and sampling of new material graphs via GNN.

VALERIE POULIN, Tutte Institute for Mathematics and Computing
Evaluating Graph Clustering Algorithms: Beware!

An impressive number of graph clustering algorithms have been proposed, studied and compared over the past decades. To identify better graph clustering techniques, one needs a way to score the techniques against one another. A typical method is to compare values of some similarity measure between ground truth partitions of given graphs and the partitions produced by the different algorithms on those graphs. However, the choice of the similarity measure used is crucial and has a huge impact on the conclusions made.

In graph clustering comparison studies, set partition similarities are used as accuracy measures. Typically, a member of the pair-counting family such as Adjusted Rand Index, or of the Shannon information-based family such as Adjusted Mutual Information is used to assess the superiority of a graph clustering algorithm over another. These measures are designed for comparing set partitions and not graph partitions specifically. We call them *graph-agnostic* as they ignore the graph structure.

In this work, we introduce a family of *graph-aware* similarity measures with their adjusted forms for graph partitions and prove that the combination of both graph-agnostic and graph-aware measures is critical for effectively comparing graph partitions.

SANDIP ROY, Washington State University
Algebraic Graph Theory Techniques for Resilience in Critical Infrastructures

As critical infrastructures become pervasively cyber-enabled, they are increasingly being subjected to heterogeneous threats arising from both the cyber world and physical world. At the same time, individual threats are having growing impact on infrastructure performance, as the infrastructures become increasingly complex and stressed. The growing susceptibility of critical infrastructures to threats has fostered research on the design of resilient decision algorithms and controls for infrastructures. In this talk, I will argue that algebraic graph theory concepts can support the design of resilient decision-making capabilities for infrastructures. In particular, using case studies from several domains (air traffic management, microgrid control), I will show that algebraic graph theory metrics give insight into the vulnerability of infrastructures to threats, and also enable positioning of control resources to mitigate these vulnerabilities.

SOPHIE SPIRKL, University of Waterloo
New results on polynomial χ -boundedness

The number of colours required to colour a graph G (the chromatic number $\chi(G)$) is at least its clique number, that is, the maximum size of a set of pairwise adjacent vertices. A class of graphs is χ -bounded if the converse is approximately true, that is, the chromatic number is at most some function of the clique number. In this talk, we are interested in when this function can be chosen as a polynomial. I will discuss some recent results in the case of forbidding a single graph as an induced subgraph.

Joint work with Alex Scott and Paul Seymour.