Young European Probabilists Workshop 2024:
Interplay between local and global structures of graphs

11th March 2024 - 15th March 2024

The YEP Workshop is a yearly event organized by EURANDOM at Eindhoven University of Technology and this is the 19th workshop in this successful series. It is geared towards learning, development and networking opportunities for young researchers.

The topic of the 2024 workshop is "Interplay between local and global graph structures". An exciting area of research that covers a wide range of sub-topics within probability theory, random graph theory, and combinatorics, among others. A lot of cutting-edge research is currently undertaken in this area, so that perfectly lends itself as the subject of the 2024 workshop.

Contents

1 Schedule 2

2 Mini-courses 3

3 Invited Talks 4

4 Contributed Talks 6

5 Posters 8

6 3 Minute Thesis Participants 9

7 Directions 10

1 Schedule

	Monday 11th	Tuesday 12th	Wednesday 13th	Thursday 14th	Friday 15th
9:30 - 10:00		**Souvik Dhara**	**Souvik Dhara**	**Ross Kang**	**Cécile Mailler**
10:00 - 10:30	**Souvik Dhara**				
10:30 - 11:00		*Break*	*Break*	*Break*	*Break*
11:00 - 11:30	*Break*	**Ross Kang**	**Ross Kang**	**Cécile Mailler**	**Cécile Mailler**
11:30 - 12:00	**Souvik Dhara**				
12:00 - 12:30					
12:30 - 13:00					
13:00 - 13:30	*Lunch*	*Lunch*	*Lunch*	*Lunch*	*Lunch*
13:30 - 14:00					
14:00 - 14:30	**Michael Schaub**	**Johannes Lengler**	**Gábor Lugosi**	**Nicolas Broutin**	
14:30 - 15:00					
15:00 - 15:30	*Break*	*Break*	*Break*	*Break*	
15:30 - 16:00	**Jasper van der Kolk**	**Felix Stamm**	**Ollie Baker**	**Serte Donderwinkel**	
16:00 - 16:30	**Riccardo Michiela**	**Nick Schleicher**	**David Geldbach**		
16:30 - 17:00	*Break*	*Break*	*Break*	*Break*	
17:00 - 17:30	**Poster session & Reception**	**How to apply for funding**	**3MT**	**Carmen van-de-l'Isle**	
17:30 - 18:00				**Matthias Lienau**	
18:00 - 18:30					
18:30 - 19:00					
19:00 - 19:30		***Dinner***			
19:30 - 20:00					

2 Mini-courses

Souvik Dhara

The community detection problem

Recovering latent communities is a key unsupervised learning task in network data. This topic has received significant attention due to its far-reaching applications, ranging from identifying echo chambers in social networks to discerning brain regions predictive of psychosis onset. We will explore key theoretical aspects in community detection from the lens of random graphs, and discuss recent advances in this direction focused on community detection with missing data.

Ross Kang

The local occupancy method in combinatorics

An old and central problem in combinatorics is to find maximum independent sets –vertex subsets inducing no edges– in triangle-free graphs –graphs whose neighbourhood subsets induce no edges. One might cast this as transferring local structure (triangle-freeness) to global structure (large independent set). I will survey some recent attempts to capture state-of-the-art progress on this and a large family of related problems from extremal combinatorics, especially in the bounded maximum degree case, within a unified framework. The framework is built around the probabilistic method, and a view of independent sets from the perspective of statistical physics –the hard-core model. It reduces bounds on the extremal (deterministic) problems to verification of a condition in the hard-core model dubbed local occupancy. I hope to highlight how the random (regular) graph –as well as related discrete random processes– offers intuition in such problems. I plan to mention a number of easily-stated related open problems, where all known techniques fall short.

Cécile Mailler

The random recursive tree and other "small world" random trees.

In this mini-course, we focus on "small-world" trees, i.e. trees whose height (distance between the root and the leaf farthest to the root) is logarithmic in the number of nodes. These include the random recursive tree, the preferential attachment tree, the random binary search tree, and, more generally, most of the "split trees" defined by Devroye in the late 80's. In this course, we look at several properties of some of these trees: eg., their degree distribution, their profile, and their height.

3 Invited Talks

Nicolas Broutin

The scaling limit of critical hypercube percolation

I will talk about the critical phase of bond percolation on the Hamming hypercube $\{0,1\}^m$, and more precisely of the scaling limit of the large connected components. We will see that the limit is the same for critical Erdős-Rényi random graphs, which is much easier to study. This is based on a joint work with Arthur Blanc-Renaudie and Asaf Nachmias.

Serte Donderwinkel

Random trees are short (but not too short)

I will discuss some new upper and lower bounds on the height of random trees. The first result is that, under very general assumptions, trees with a given degree sequence, simply generated trees and Bienaymé-Galton-Watson trees of size n have height $O(\sqrt{n})$ with Gaussian tails (and height $o(\sqrt{n})$ in the high variance regime). Moreover, we show that all critical Bienaymé-Galton-Watson trees of size n have height $\omega(\log n)$. The proofs are mostly combinatorial and are based on the Foata-Fuchs bijection between trees and sequences. The results resolve various conjectures from the literature and are based on a work with Louigi Addario-Berry and a work with Louigi Addario-Berry and Igor Kortchemski.

Johannes Lengler

Geometric inhomogeneous random graphs in non-metric spaces

Geometric Inhomogeneous Random Graphs (GIRGs) are a random graph model where each vertex draws i.i.d. two random variables: 1) a weight, which corresponds to the expected degree, and which we assume to be power-law distributed, and 2) a position in some geometric space. Then each pair of vertices connects with some probability that depends on their distance and their weights. This model has many properties that are also found in real-world social networks, like the power-law degree distribution, ultra-small distances, a high clustering coefficient, a very rich community structure, small entropy (high compressibility), and more.

Traditionally, the model is studied on a cube, ball or torus with the usual Euclidean geometry, which is of course a metric space. But the model actually works for a large class of geometries with very different distance functions. Many of the desired properties hold for any distance function which satisfies just some very weak regularity conditions. In particular, some of the most interesting distance functions do not satisfy the triangle inequality. While most properties remain unchanged, some (local and global) properties do change, and I will argue that the resulting models are a better fit for real social networks than the ones that come from traditional metric spaces.

Gábor Lugosi

Increasing paths in random temporal graphs

We consider a version of the classical Erdős–Rényi random graph G(n,p) where each edge has a distinct random time stamp, and connectivity is constrained to sequences of edges with increasing time stamps. We study the lengths of increasing paths: the lengths of the shortest and longest paths between typical vertices, the maxima of these lengths from a given vertex, as well as the maxima between any two vertices; this covers the (temporal) diameter.

This talk is based on joint work with Nicolas Broutin and Nina Kamčev.

Michael Schaub

Learning from networks with unobserved edges

In many applications, we are confronted with the following system identification scenario: a dynamical process is observed that describes the state of a system at particular times. Based on these observations, dynamical interactions are inferred between the entities observed. In the context of a distributed system, this typically corresponds to a "network identification" task: find the weighted edges of the graph of interconnections. However, often, the number of samples obtained from such a process is far too few to identify the edges of the network exactly. Can one still reliably infer some aspects of the underlying system? Motivated by this question, the following identification problem is considered: instead of trying to infer the exact network, the aim is to recover a low-dimensional statistical model of the network based on the observed signals on the nodes. More concretely, the focus is on observations that consist of snapshots of a diffusive process that evolves over the unknown network. The model of the unobserved network is generated from an independent draw from a latent stochastic block model (SBM), and the goal is to infer both the partition of the nodes into blocks, as well as the parameters of this SBM. Simple spectral algorithms are presented that provably solve the partition and parameter inference problems with high accuracy. Some possible variations and extensions of this problem setup are further discussed.

4 Contributed Talks

Ollie Baker

Entropy of random geometric graphs

The entropy of a graph allows us to quantify the complexity of a network structure. Specifically, a random geometric graph has a wide range of applications, from modelling communication networks to molecular structures. In this talk, I will outline some background results in the entropy of graphs embedded in space, and detail some new insights into recent research questions, such as how entropy varies with density, and how we can bound the entropy of a random geometric graph using probability and information theory.

David Geldbach

Continuum asymptotics for tree growth models

We classify the forward dynamics of all (plane) tree-valued Markov chains $(T_n, n \geq 1)$ with uniform backward dynamics. That is, to obtain the distribution of T_{n-1} we remove a uniform leaf from T_n. Every such Markov chain is classified by a decorated planar real tree. We also show that under an inhomogeneous rescaling after trimming leaves $(T_n, n \geq 1)$ converges to a random real tree in the Gromov–Prokhorov metric.

Carmen van-de-l'Isle

The symbiotic contact process on random trees

The symbiotic contact process can be thought of as a two type generalisation of the contact process which can be used to model the spread of two symbiotic diseases. Each site can either be infected with type A, type B, both, or neither. Infections of either type at a given site occur at a rate of lambda multiplied by the number of neighbours infected by that type. Recoveries of either type at a given site occur at rate 1 if only one type is present, or at a lower rate mu if both types are present, hence the symbiotic name. Both the contact process and the symbiotic contact process have two critical infection rates on a Galton-Watson tree, one determining weak survival, and the other strong survival. Here, weak survival refers to the event where at least one A infection and at least one B infection is present at all times. Strong survival is the event that the root of the tree is infected with both A and B infections at the same time infinitely often. In this talk, I will prove that for small values of mu the weak critical infection rate for the symbiotic model is strictly smaller than the critical rate for the contact process. I will also discuss the more complicated case of strong survival for both processes.

Jasper van der Kolk

Emergence of geometric turing patterns in complex networks

Turing patterns, arising from the interplay between competing species of diffusive particles, have long been an important concept for describing nonequilibrium self-organization in nature and have been extensively investigated in many chemical and biological systems. Historically, these patterns have been studied in extended systems and lattices. Recently, the Turing instability was found to produce topological patterns in networks with scale-free degree distributions and the small-world property, although with an apparent absence of geometric organization. While hints of explicitly geometric patterns in simple network models (e.g., Watts-Strogatz) have been found, the question of the exact nature and morphology of geometric Turing patterns in heterogeneous complex networks remained unresolved. In this work, we study the Turing instability in the framework of geometric random graph models, where the network topology is explained using an underlying geometric space. We demonstrate that not only can geometric patterns be observed, their wavelength can also be estimated by studying the eigenvectors of the annealed graph Laplacian. Finally, we show that Turing patterns can be found in geometric embeddings of real networks. These results indicate that there is a profound connection between the function of a network and its hidden geometry, even when the associated dynamical processes are exclusively determined by the network topology."

Matthias Lienau

Large components in the subcritical Norros-Reittu model

The Norros-Reittu model is a random graph with n vertices and i.i.d. weights assigned to them. The number of edges between any two vertices follows an independent Poisson distribution whose parameter is increasing in the weights of the two vertices. Choosing a suitable weight distribution leads to a power-law behaviour of the degree distribution as observed in many real-world complex networks. We study this model in the subcritical regime, i.e. in the absence of a giant component. For each component, we count the vertices and show convergence of the corresponding point process to a Poisson process. More generally, one can also count only specific vertices per component, like leaves. From this one can deduce asymptotic results on the size of the largest component or the maximal number of leaves in a single component. The results also apply to the Chung-Lu model and the generalised random graph.

Riccardo Michielan

Detecting network geometry using global and local triangle-based statistics

In this talk we will discuss why encoding geometry in random graphs, embedding vertices in a proper metric space, helps revealing key properties of complex networks, including small-worldness, clustering, and community structure. Moreover, to address the challenge of distinguishing network geometry, framed as an hypothesis testing problem, we will introduce a novel triangle-based statistic which discounts triangles formed on high-degree vertices.

Nick Schleicher

Component size of early vertices in preferential attachment

We study the size of the connected component of early typical vertices in an inhomogeneous random graph of preferential attachment type. The main tools are a coupling of the graph to a killed branching random walk and a result for the number of particles absorbed in the killing barrier.

Felix Stamm

Generating random graphs that (approximately) or exactly preserve network centrality measures

"Neighborhood Structure Configuration Models" are a way to generate synthetic networks that are probably indistinguishable from an original network in terms of the Weisfeiler-Leman graph isomorphism test. By doing so we can ensure some meso scale structure of the original network is still present in synthesized networks and thus network centralities like Katz, Eigenvector, and PageRank of the synthetic network may be preserved. The MCMC sampling scheme used to generate samples is easy and computationally inexpensive allowing for Neighborhood Structure Configuration Models to be applicable to large networks without problems.

5 Posters

1. **Ollie Baker**

   Entropy of soft random geometric graphs

2. **Vinay Kumar Bindiganavile Ramadas**

   Community recovery on block models with geometric kernels

3. **Stijn Cambie**

   Packing of list colourings

4. **Martijn Gösgens**

   The projection method: a geometric framework for community detection

5. **Mahamat Abdelkerim Issa**

   Flexible approach to modeling the climate-health warning system

6. **Yongjin Lee**

   Estimating number of ways to giving acyclic orientations on random graphs

7. **Marta Milewska**

   Dynamic local convergence

8. **Niek Mooij**

   Finding large independent sets in networks using competitive dynamics

9. **Bharath Roy Choudhury**

   Genealogy of records of stochastic processes with stationary increments as unimodular trees

10. **Christel Sirocchi**

    Leveraging local and global graph properties in distributed averaging on networked systems

11. **Felix Stamm**

    Neighborhood structure configuration models

12. **Rebecca Steiner**

    Broadcasting on random recursive trees

13. **Mike de Vries**

    Graphs with uniform independent sets

6 3 Minute Thesis Participants

1. Stijn Cambie
2. Alberto Fiorentino
3. Martijn Gösgens
4. Carmen van-de-l'Isle
5. Marilyn Korfhage
6. Yongjin Lee
7. Susana Morala
8. Manish Pandey
9. Lucas Schätze

7 Directions

All presentations and activities (except for the dinner on Tuesday evening) will be held on the fourth floor of the MetaForum building. This building can be found after the water fountain, when entering the campus from the train station and is highlighted by a red circle in the map below. The presentations will be held in the room MF11/12 and most activities will be either in this same room, or in the Eurandom lounge. Both MF11/12 and the lounge are located on the 4th floor, on the opposite side of the elevators.