10th International Conference on Network Analysis NET 2020

Nizhny Novgorod, June 15-16, 2020

Laboratory of Algorithms and Technologies for Network Analysis
National Research University
Higher School of Economics, Nizhny Novgorod
Laboratory of Advanced Combinatorics and Network Applications,
National Research University
Moscow Institute of Physics and Technology

Location: Zoom

Conference chair

Panos M. Pardalos, University of Florida, USA and LATNA, NRU HSE

Program Committee

Fuad Aleskerov, NRU HSE

Mikhail Batsyn, NRU HSE

Fedor Fomin, Bergen University, Norway and St Petersburg department of Steklov Mathematical Institute

Valery Kalyagin, NRU HSE

Yury Kochetov, Russian Academy of Sciences, Novosibirsk

Alexander Koldanov, NRU HSE

Valentina Kuskova, NRU HSE, Laboratory of Applied Network Research

Dmitriy Malyshev, NRU HSE

Oleg Prokopyev, University of Pittsburgh, USA

Andrey Raigorodskii, Moscow Institute of Physics and Technology, Moscow State University, Yandex

Nikolay Zolotykh, Lobachevsky State University, Nizhny Novgorod

Andrey Savchenko, NRU HSE

Conference Organizers

Panos M. Pardalos, University of Florida, USA

Andrey Raigorodski, MIPT – Yandex, Russia

Valery Kalyagin, National Research University Higher School of Economics, Russia

Monday, June 15

15:00 – 15:30 Panos Pardalos, University of Florida and LATNA HSE
Inverse Combinatorial Optimization Problems
15:30 – 16:00 Fuad Aleskerov, HSE and Institute of Control Sciences RAS
New Centrality Measures in the Networks and their Applications
16:00 – 16:30 Tiago Schieber, Federal University of Minas Gerais, Brazil
Diffusion Capacity in Multilayer Networks: the Role of the Interlayer Paths
16:30 – 17:00 Mario Guarracino, University of Cassino and Southern Lazi, Italy, *A Short Journey through Whole Graph Embedding Techniques*

17:00 – 17:20 Coffee break

17:20 – 17:40 Evgueny Burashnikov, LATNA HSE
Branch-and-bound and Dynamic Programming Approaches for the Knapsack Problem
17:40 – 18:00 Dmitry Sirotkin, LATNA HSE
NP-hardness of Problems, Connected with an Auto-battler Video-game Genre
18:00 – 18:30 My Thai, University of Florida, USA
Fast Maximization of Non-submodular Functions on Integer Lattice
18:30 – 19:00 Oleg Prokopyev, University of Pittsburgh, USA
Fortification Against Cascade Propagation Under Uncertainty

Tuesday, June 16

15:00 – 15:30 Nenad Mladenovic, Khalifa University, Abu Dhabi, United Arab Emirates, *Less is More Approach in Optimization*

15:30 – 16:00 Pierre Miasnikof, University of Toronto, Canada

Graphs in Metric Space

16:00 – 16:30 Gregory Khvatsky, HSE Moscow

Named-Entities Attributes Prediction based on their Position in Semantic Graph: Application for Policy Actors Identification

16:30 – 17:00 Theodore Trafalis, University of Oklahoma, USA

Latency-Based Evacuation Modeling Using Dynamic Traffic Assignment

17:00 – 17:20 Coffee break

17:20 – 17:40 Anastasia Sokolova. LATNA HSE

Open-set Face Identification Algorithms with Automatic Detection of Rare Data

17:40 – 18:00 Dmitry Semenov, LATNA HSE

Building a Confidence Set of Connected Stocks

18:00 – 18:30 Andrei Raigorodsky, MIPT and Yandex

Modularity in Different Graph Models

18:30 – 19:00 Roman Belavkin, Middlesex University London, UK

Distance-degree Duality and Phase Transitions in Large Graphs

Opening talk

Inverse Combinatorial Optimization Problems

Panos M.Pardalos

www.ise.ufl.edu/pardalos

Given an optimization problem and a feasible solution to it, the corresponding inverse optimization problem is to find a minimal adjustment of the cost vector under some norm such that the given solution becomes optimum. Inverse optimization problems have been applied in diverse areas, ranging from geophysical sciences, traffic networks, communication networks, facility location problems, finance, electricity markets, and medical decision-making. It has been studied in various optimization frameworks including linear programming, combinatorial optimization, conic, integer and mixed-integer programming, variational inequalities, and countably infinite linear problems and robust optimization. In this talk, we mainly concentrate on inverse combinatorial optimization problems (ICOP). We will introduce some classes of ICOP as well as general methods to solve them. Some open problems are proposed. We also discuss some generalized inverse optimization problems. We introduce inverse optimization problems on spanning trees and mainly concentrate on the inverse max+sum spanning tree problems (IMMST) in which the original problem aims to minimize the sum of a maximum weight and a sum cost of a s-panning tree (see references).

References

[1] X.C. Guan, X.Y. He, P.M. Pardalos and B.W. Zhang (2017), Inverse max+sum spanning tree problem under hamming distance by modifying the sum-cost vector, Journal of Global Optimization, 69(4), 911-925.

[2] X.C. Guan, P.M. Pardalos and B.W. Zhang (2018), Inverse max+sum spanning tree problem under weighted l1 norm by modifying the sum-cost vector, Optimization Letters, 12(5), 1065-1077.

[3] X.C. Guan, P.M. Pardalos and X. Zuo (2015), Inverse max+sum spanning tree problem by modifying the sum-cost vector under weighted l∞ norm. Journal of Global Optimization, 61(1), 165-182.

Invited talks

New centrality measures in the networks and their applications

Fuad Aleskerov

National Research University Higher School of Economics and
Institute of Control Sciences of Russian Academy of Sciences

We propose a new class of centrality measures in networks which are defined as a linear combination of in-degree centrality index, bundle index taking into account parameters of the vertices and group influence of the vertices to a vertex, and a pivotal index taking into account most influential vertex in the group. The approach is illustrated via two examples - migration and foreign trade.

Distance-degree duality and phase transitions in large graphs

Roman Belavkin

Department of Computer Science, Middlesex University London

We consider two variational problems for random graphs: 1) Maximize entropy of a degree distribution subject to a constraint on the average of log-degree; 2) Minimize the average distance (path length) in a graph subject to a constraint on mutual information between vertexes. Solution to the first optimization problem is the power-law degree sequence, which is known to characterize many large networks. Solution to the second optimization problem leads to the preferential attachment procedure, which is known to generate power-law graphs. We show that for large networks these optimization problems are dual of each other. In particular, the average distance from a vertex to all other vertexes in a very large graph can be related negative logarithm of its degree. The logarithm of a degree turns out to be useful in the analysis of a phase transition in power-law graphs at exponent parameter is equal to one, above which such graphs are known to be almost surely disconnected.

Branch-and-bound and dynamic programming approaches for the knapsack problem

Evgueny Burashnikov
Lab LATNA
National Research University Higher School of Economics, Nizhny Novgorod

The knapsack problem is one of the most popular NP-hard problems in combinatorial optimization. For 0-1 knapsack problem there are two famous approaches which guarantee the optimality of the solutions: Branch-and-Bound (BnB) and Dynamic Programming (DP) algorithms. For DP we suggest some preprocessing and rules which help us to avoid unneeded calculations. For BnB we provide an algorithm which uses the same cache of solutions as DP does, but fills only the necessary part of this cache and does it gradually during the search.

A short journey through whole graph embedding techniques

Ilaria Granata, Lucia Maddalena, Ichcha Manipur
National Research Council of Italy
Mario Manzo
University of Naples "L'Orientale"
Mario R Guarracino
University of Cassino and Southern Lazi, mario.guarracino@gmail.com

Networks provide suitable models in many applications, ranging from social to life sciences. Such representations are able to capture interactions and dependencies among variables or observations, thus providing simple and powerful modeling of phenomena. Whole graph embedding involves the projection of graphs into a vector space, while retaining their structural properties. In recent years, several embedding techniques using graph kernels, matrix factorization, and deep learning architectures have been developed to learn low dimensional graph representations. These embeddings can then be used for feature extraction, graph clustering or for building classification models. In this paper, we survey embedding techniques which jointly embed whole graphs for classification tasks. We compare them and

evaluate their performance on undirected synthetic and real world network datasets.

Named-Entities Attributes Prediction based on their Position in Semantic Graph: Application for Policy Actors Identification

Valentina Kuskova, Dmitry Zaytsev, **Gregory Khvatsky**
National Research University Higher School of Economics, Moscow, Russia
vkuskova@hse.ru, dzaytsev@hse.ru, grigoriy.khvatskiy@outlook.com

In this paper we propose an algorithm for named-entity recognition without dictionary using only machine learning procedures. This algorithm based on syntactic analysis trying to find patterns in syntax to predict existing dictionary which was made by human. There are plenty of areas were such ma-chine learning procedures can be helpful: from policy-making to business ad-ministration. We apply the proposed model to the real world example of the Russian education policy.

Graphs in Metric Space

Pierre Miasnikof (1), Alexander Y. Shestopaloff (2,3), Leonidas Pitsoulis (4), and Yuri Lawryshyn (1)
1 University of Toronto, Toronto, ON, Canada
2 Queen Mary University of London, London, United Kingdom
3 The Alan Turing Institute, London, United Kingdom
4 Aristotle University of Thessaloniki, Thessaloniki, Greece

With a view on graph clustering, we present a definition of vertex-to-vertex distance which is based on shared connectivity. Our definition differs from the usual shortest-path geodesic distance. Using this definition, we represent graphs as vertices in a metric space, where all vertex-pairs are separated by a known distance. One advantage of using such a graph representation is that it allows a formulation of the graph clustering problem as a distance minimization problem. We argue that vertices sharing more connections are closer to each other than vertices sharing fewer connections. Our thesis is centered on the widely accepted notion that strong clusters are formed by high levels of induced subgraph density,

where subgraphs represent clusters. These clusters are formed by grouping vertices deemed to be similar in their connectivity. At the cluster level (induced subgraph level), our thesis translates into low mean intra-cluster distances. In this article, we compare three distance measures from the literature. Our benchmark is the accuracy of each measure's reflection of intra-cluster density, when aggregated (averaged) at the cluster level. We conduct our tests on synthetic graphs generated using the planted partition model, where clusters and intra-cluster density are known in advance. We examine correlations between mean intra-cluster distances and intracluster densities. Our numerical experiments show that Jaccard and Otsuka-Ochiai offer very accurate measures of density, when averaged over vertex pairs within clusters.

Less is more approach in Optimization

Nenad Mladenovic

Faculty of Engineering, Khalifa University, Abu Dhabi, United Arab Emirates.

In all kinds of human creative work comes the moment when adding more ingredients does not provide a better outcome. The reaction to the inclusion of redundant ingredients is represented by the approach most commonly called "Less is more", in almost all scientific and artistic disciplines. It is the case in Architecture, Music, Physics, Medicine, Neurosciences, Teaching, Cuisine, etc. My collaborators and I have recently proposed 'Less is more approach' (LIMA) in Optimization. Its main idea is to find the minimum number of search ingredients in solving some optimization problems, which makes some optimization methods more efficient than the currently best. LIMA appeared as a reaction to more and more complex hybrid heuristic methods that combine many different ideas, without a real explanation of why those ideas are combined. Combining several heuristics to get a new hybrid method has a price of losing efficiency and user-friendliness, the two very important desired properties of any heuristic. Indeed, despite the simplicity of LIMA, we got significantly better results than more complex heuristics in solving several classical optimization problems. Thus,

including many ideas in the search does not necessarily lead to better computational results; on the contrary, sometimes less can yield more.

Fortification Against Cascade Propagation Under Uncertainty

Oleg Prokopyev

University of Pittsburgh, USA

Network cascades represent a number of real-life applications: social influence, electrical grid failures, and so on. The commonality between these phenomena is that they begin from a set of seed nodes and spread to other regions of the network. We consider a variant of a critical node detection problem dubbed the robust critical node fortification problem, wherein the decision-maker wishes to fortify nodes (within a budget) to limit the spread of cascading behavior under uncertain conditions. In particular, the arc weights - how much influence one node has on another in the cascade process - are uncertain but are known to lie in some range bounded by a worst-case budget uncertainty. The cascading behavior is modeled using the Linear Threshold model. This problem is shown to be NP-hard, even in the deterministic case. We formulate a mixed-integer program (MIP) to solve the deterministic problem, and improve its continuous relaxation via nonlinear constraints and convexification. The robust problem is computationally more difficult, and we present an MIP-based Expand-and-cut exact solution algorithm. Insights from these exact solutions motivate two novel (inter-related) centrality measures, and a centrality-based heuristic that obtains high-quality solutions within a few seconds. Finally, extensive computational results are given to validate our theoretical developments, as well as provide insights into structural properties of the robust problem and its solution.

Modularity in different graph models

Andrey Raigorodski
MIPT and Yandex

In the talk, I'll overview some recent progress in bounding modularity for various graph models.

NP-hardness of problems, connected with an auto-battler video-game genre

Dmitry Sirotkin
Lab LATNA
National Research University Higher School of Economics, Nizhny Novgorod

Auto-battler is a relatively new video-game genre, where several players compete against each other. Each player recruits a team of fantasy creatures and puts them to fight with other players' teams. In a team-building not only raw strength of creatures is important, but also their affiliation: each creature can be a part of several specific alliances - like mages, warriors, druids, etc. Gathering in a team several members of one alliance confers different benefits for said team. Moreover, since one creature can be a part of several alliances, they overlap and a team can easily benefit from several active alliances, making the process of team building quite complicated. We formalize this problem in terms of linear programming and prove, that it is NP-hard. Also, we solve this problem for one representative of the genre with the simplex method.

Diffusion capacity in multilayer networks: the role of the interlayer paths

Tiago Schieber
Federal University of Minas Gerais, Brazil

The past decade has witnessed an important advance in the representation and analysis of complex systems. The understanding that a myriad of phenomena

consists of interdependent processes that run at many levels has successfully spurred research efforts into methods to express their interrelations. The emergence of multilayer representations has improved the interpretation of a diverse set of critical phenomena, ranging from super-diffusion to multi-state synchronization. This work introduces the concept of “Diffusion Capacity”, a measure that quantifies the potential of a network element to propagate information, conditioned by its structural connectivity. For that, a novel method to efficiently measure the network's diffusion capacity is proposed. The method considers the node distance distribution from weighted networks and uses the cumulative Jensen-Shannon divergence. Focusing on inter-layer connections, we investigate when it is advantageous, in terms of diffusion, for a network element to be part of a coupled structure. We present some examples with synthetic and real-world data to show the wide potential of our measure.

Building a confidence set of connected stocks.

Dmitry Semenov, Petr Koldanov
Lab LATNA
National Research University Higher School of Economics, Nizhny Novgorod

The research considers the task of analyzing the relationships between stocks based on the results of observations of their returns. The aim is to build and study the properties of statistical procedures for constructing a set of stock pairs, which with a given probability contains all pairs of stocks, links (edge weights) between which are more than a given threshold. Such set will be called a confidence set for the edges of the threshold graph. As the measure of similarity between stock returns, the popular Pearson measure, the sign measure of similarity or the Fechner measure, and the Kendall rank measure are used. To construct such sets standard tests are used to test individual hypotheses about the values of the considered similarity measures. While testing a lot of hypothesis simultaneously in order to control the probability of a least one 1st type error Bonferroni, Holm and Hochberg multiple statistical procedures are used and investigated. The dependence of the

confidence probability, the size of the constructed confidence set and the average number of links that did not fall into the confidence set, on the model of joint behavior of stock returns, such as individual tests and the volume of observations, are studied. The results of statistical modeling show the stability of procedures based on a combination of individual Kendall and Fechner tests, and the instability of procedures based on a combination of individual Pearson tests. At the same time, the percentage of strong links that were erroneously not included in the constructed set by the Pearson procedure is quite small. The research is supported by RFFI grant 18-07-00524.

Open-set Face Identification Algorithms with Automatic Detection of Rare Data

Anastasia Sokolova

Lab LATNA

National Research University Higher School of Economics, Nizhny Novgorod

Nowadays, the main issue for facial classification is bad quality of images, different scale or illumination, changes in the appearance and etc. Also, well-trained classifiers make a mistake on such images and often define them as one class objects. Therefore, in this talk, I focus on the problem of detecting "out-of-distribution" facial images. A special dataset was collected by different transformations and anomalies detection. This dataset was used to train a neural network that can find "out-of-distribution" facial images. I present the k-NN classifier that takes into account information about the distribution of the images. The problem of insufficient computational efficiency of such approach is highlighted. A hierarchical classifier of image descriptors is introduced. Finally, I present experimental comparison of the proposed approach with the well-known classification methods for feature vectors extracted by various convolutional networks.

Fast Maximization of Non-submodular Functions on Integer Lattice

My Thai
University of Florida, USA

The optimization of submodular functions on the integer lattice has received much attention recently, but the objective functions of many real-world applications are non-submodular. This talk discusses two approximation algorithms for maximizing a non-submodular function on the integer lattice subject to a cardinality constraint; these are the first algorithms for this purpose that have polynomial query complexity. We also introduce a general framework for influence maximization on the integer lattice that generalizes prior works on this topic, and we demonstrate the efficiency of our algorithms in this context.

Latency-Based Evacuation Modeling Using Dynamic Traffic Assignment

Hamoud Bin Obaid
Department of Industrial Engineering, King Saud University, Saudi Arabia,
and
Theodore B.Trafalis
Department of Industrial and Systems Engineering, University of Oklahoma
ttrafalis@ou.edu

In this talk, a latency-based dynamic traffic assignment (LB DTA) modeling approach is introduced to optimize the dynamic evacuation process in emergencies. The model is developed as a mixed integer linear programming (MILP) model using a piecewise linear approximation of the load-dependent travel time function on each link. The objective is to minimize the network clearance time E, where E is the total travel time of the last group of evacuees arriving at the safe destination since the beginning of the evacuation process. The model is tested on small networks and the computational results are compared with the Cell Transmission Model (CTM) Other objectives are also discussed such as the average travel time (ATT) and the average evacuation time (AET).