The final report of the project NKFI-SSN-117879

October 20, 2020

In the frame of this project both basic mathematical research and algorithmic research were pursued on the Hungarian side. We used every opportunity to foster collaboration between these groups: summer schools, conference visits and study trips made it possible to interchange ideas and learn from the other fields. There was an intensive collaboration with the Slovenian partners which was realized by several scientific meetings and common workshops both in Budapest an in Ljubljana. Furthermore, a conference entitled "Middle-European Conference on Applied Theoretical Computer Science" was organized in which the Slovenian-Hungarian cooperation was crucial.

During the project, Gábor Tardos's research was recognized by electing him a member of Academia Europea and a corresponding member of the Hungarian Academy of Sciences and by asking him to give an invited lecture at the International Congress of Mathematicians (Rio de Janeiro, 2018). Ervin Győri's research was recognized by the Tibor Szele prize (issued by the Bolyai Mathematical Society) in 2018.

The project was very fruitful in terms of scientific research and publications. See the 85 items in the attached publication list. Here we give a sample of the results only.

László Tóth, László Hajdu, András London and others worked on the community structure and diffusion models in graphs. The community structure of a graph is a classification of its points. The task is to find such classification that gives meaningful information of the structure of the graph due to some natural requirements. A possible approach is using statistically validated networks, which reduce the size of the original system by considering only the links which are statistically significant. Key new concepts were developed and possible applications were shown on various real data sets including the actor-movie network, co-authorship network and word co-occurrence networks. A method was developed which is able to detect patterns in urban public transport.

Diffusion models in graphs is a general approach to model information spreading through network structures. One of the most widely used models is the so called Independent Cascade Model. A generalization of this model has been introduced recently for which optimization problems can also be considered. For large graphs mainly greedy heuristics are effective. Different approaches of the greedy type methodology were studied and analyzed . The first results have been presented at the StuCosRec Conference and the Hungarian Operations Research Conference.

András London and coauthors analyzed the performance of the Markowitz portfolio optimization method on the Budapest Stock Exchange data set using two different filtering techniques defined for correlation matrices. The results show that the estimated risk is much closer to the realized risk using filtering methods. Bootstrap analysis shows that ratio between the realized return and the estimated risk (Sharpe ratio) is also improved by filtering.

A larger group lead by Miklós Krész worked on scheduling problems. Evaluating heuristic algorithms is hard in the absence of proven performance guarantees. The most popular method is to use benchmark problems. The group aimed to develop method to generate these benchmark problems that characterize the problem properly. PCB assembly optimization is applied for this problem.

A new method was developed for crew scheduling using the graph coloring approach. Another classical graph theoretical problem used for scheduling is matching in graphs. Inverse optimization problems are about searching parameters with minimum cost such that a prespecified solution becomes optimal. They studied the inverse problems for vertex-weighted matchings by which the structure of the optimal solutions can be understood. It can help to develop new methods in scheduling applications.

The algorithmic methodology of scheduling connected to public transportation is a well-studied field, but the specific practical requirements are rarely considered in basic research. The group studied a set of practical questions in the above fashion summarized below. Disruptions happen on a daily basis to the schedules of a transportation company. Because of this, quick and efficient solution methods and frameworks are needed to restore the order of transportation as soon as possible. This was considered for a daily horizon, and models and methods are presented that give good solutions in acceptable running time. Long term planning is important is public transportation, and the assignment of buses over such a planning period is not really considered in the literature. When creating vehicle schedules that have similar underlying structure, it should be important to consider this during the scheduling process, and aim for solutions as close to each other as possible. The group gave heuristic approaches to consider similarity while creating vehicle schedules with similar input.

The paper of Bekesi, Dávid and Krész deals with schedule assignment over a planning period for inter-city bus transportation. Here, buses are assigned daily schedules over a longer planning period, while their parking and maintenance tasks also have to be taken into account. In order to better visualize the structure of the solution, linearly priced times automata were applied during the solution method. Another paper of Dávid, Krész and Hegyháti deals with integrating the phases of vehicle scheduling and assignment for a single day. Instead of creating a simple vehicle schedule, exact buses are also assigned to this, and all vehicle-specific tasks (like parking, refueling) belonging to this bus are inserted into this schedule as well. The paper gave a set-partitioning model for the problem, and solved it using column generation.

The community structure of a graph is a classification of its vertices. The task is to find such a classification providing meaningful information based on the graph structure. The classical approach is considering a partition of the vertices, however the fine structure of the graph with overlapping community detection gives the framework more sophisticated information extraction. In the newly published paper of Bóta, Krész and Hajdu have developed an efficient overlapping community detection algorithm in directed graphs and have demonstrated its use in the maximization of influence spread. Diffusion models in graphs is a general approach to model information spreading through network structures. One of the most widely used models is the so called Independent Cascade Model. Since the maximization problem is NP-hard, numerous approximation algorithms have been developed during the last years. Their new approach is based on the community structure by which a community value can be calculated for each vertex. This community value is used after that for ordering the vertices with giving the chance to use those vertices in the approximation which has high community values. It turned out that results achieved by the the new approach are very promising int he Independent Cascade Model.

A London, M. Krész and J. Németh developed a new model for probabilistic forecasting using graph-based rating method. They provided a forward-looking type graph-based approach and applied it to predict football game outcomes by simply using the historical game results data of the investigated competition. Although the application they presented in the published paper is special, they note that our method can be applied to forecast general graph processes. A talk was given in the 21st IS - 2018 Slovenian conference and a short paper published in the corresponding proceedings. The research about the community detection in public transportation networks have been summarized in a paper published by L. Hajdu, A. Bóta, M. Krész, A. Khani and L. M. Gardner. They utilized known contact patterns from a public transit assignment model in a major metropolitan city, and proposed the development of two novel network structures, each of which elucidate certain aspects of passenger travel behavior.

A. Tóth and M. Krész presented their research about universal tule definition for scheduling problems in the VOCAL 2018 conference and published an article in the conference proceedings. B. Dvid and M. Krsz used a robust flow network to create schedules for multiple days at the same time. In this approach all input trips of every considered day is used to build the flow network, then through the failure of certain arcs, solutions can be produced for every single day. This result was presented in International Conference of Practice and Theory of Automatic Timetabling. Another problem in public transit application considered by the same authors is a planning horizon of several days, where buses have to be assigned to existing schedules, and their vehicle-specific tasks (parking, maintenance) are also considered. A mathematical model is developed for the problem, and solutions for periods of up to 3 weeks are presented.

Leila Badakhshian and Gyula Katona found the domination number of the following bipartite graph: The vertices in the first class are the k-element subsets of an n-element set. The second class consists of all elements. Two vertices are adjacent iff the set contains the element. Consider now the analogous problem where one side consist of the k-element subsets, the other one of the two-element subsets. Two vertices are adjacent iff the k-element subsets is a superset of the two-element one. Toward their general conjecture for the asymptotic behavior of this domination number (constant times n^2) they proved lower and upper estimates on the constants. Together with Zs. Tuza, they later managed to prove the full conjecture.

An old theorem of Erős and Gallai determined the Turán nmber of a path of k vertices. G.O.H. Katona, C. Xiao, J. Xiao and O. Zamora posed a conjecture for the Turán number of the square of a path and proved it for $k = 5$ and 6 (3 and 4 were known).

Péter Erdős and István Miklós worked on degree sequence problems with several co-authors. They defined stable degree sequences and showed that the swap Markov chain is always rapidly mixing on the realizations of stable degree sequences. They also showed that the swap Markov chain is rapidly mixing if there is a short alternating cycle in any realization on each pair of vertices. They also showed the following. If there is no short alternating cycle on one of the pair of points, then the spanned subgraph on the shortest alternating cycle has a prescribed degree sequence. This degree sequence is not stable, however, the swap Markov chain is rapidly mixing on its realizations.

I. Miklós, P. Erdős, T. Mezei and D. Soltész worked on degree sequence problems. The key definition here has been the stable degree sequences. A class of degree sequences is stable if the number of realizations of any member of them are within a polynomial factor from the number of realizations of any degree sequence within an L_1 distance 2. They proved the following theorems: a) The swap Markov chain mixes rapidly on the realizations of any stable degree sequences. b) A degree sequence is stable if any edge in any of its realizations is on a short (constant length) alternating cycle of edges and non-edges. They also gave the degree sequence of the spanned subgraph on the shortest alternating cycle. They showed that this class of degree sequences is not stable, however, the swap Markov chain is rapidly mixing on them.

István Miklós also worked with several supervised BSc students. They give necessary and sufficient conditions when two tree degree sequences have edge disjoint caterpillar realizations. They also conjectured that arbitrary number of tree degree sequences always have edge disjoint caterpillar realizations if each vertex is a leaf in at most one of the vertices. They proved that this conjecture is true if the number of tree degree sequences is at most five. They gave necessary and sufficient conditions when a degree sequence has half regular factorizations.

István Miklós together with its student, Letong Hong, gave necessary and sufficient perturbations to transform any r-edge coloring of an r-regular bipartite graph into any another r-edge coloring via r-edge colorings.

Ervin Győri and several coauthors worked on various extremal graph problems.

Ervin Győri and coauthors studied terminal pairability problems when the base graph (where the edge disjoint paths are wanted) is a complete bipartite graph. It is interesting that the cases when the demand graph is bipartite or not are very different. In both cases, they essentially improved the best known sufficient upper bounds on the maximum degree. They determined the maximum number of edges in the demand graph when the problem is always resolvable.

Beka Ergemlidze, Ervin Győri, Abhishek Methuku and Nika Salia significantly improved the previous best upper bound on the maximum number of triangles in a pentagon-free graph. The problem was initiated by Bollobás and Győri more than 10 years ago and the best known bounds were improved several times. The present upper bound is about 70% above the lower bound, about one fifth of the first bound and about one third of the last improvement.

Ervin Győri, Michael D. Plummer, Dong Ye and Xiaoya Zha studied cycles in claw-free graphs. Their main result determines the maximum of k such that for any choice of $k + 1$ vertices, a 3-connected claw-free graph contains a cycle through the first k vertices but avoids the last one. It turned out that this maximum is $k = 5$.

Ervin Győri and Tamás Róbert Mezei studied a variant of the art galery problem where the guards are mobile. Let us denote by $p(P)$, $mh(P)$, $mv(P)$ the minimum number of point guards, vertical mobile guards and horizontal mobile guards for the arrangement P. They prove sharp bounds on the ratio $(mh(P) + mv(P))/p(P)$.

L. Colucci, E. Győri and A. Methuku improved results of Hoppen, Kohayakawa and Lefmann about the maximum number of edge colorings without monochromatic copies of a star of a fixed size that a graph on n vertices may admit.

E. Győri, N. Salia, C. Tompkins and O. Zamora consider hypergraph and multi-hypergraph variants.of a classical conjecture of Erdős and Sós on the Turán number of a tree.

Lucas Colucci and Ervin Győri extend a result of Griggs and Yeh about the maximum possible value of the $L(2, 1)$-labeling number of a graph in terms of its maximum degree to oriented graphs. They consider the problem both in the usual definition of the oriented $L(2,1)$-labeling number and in some variants.

E. Győri, N. Lemons, N. Salia and O. Zamora study the structure of r-uniform hypergraphs containing no Berge cycles of length at least k for $k \leq r$, and determine that such hypergraphs have some special substructure. In particular they determine the extremal number of such hypergraphs, giving an affirmative answer to the conjectured value when $k = r$ and giving a a simple proof to a recent result of Kostochka-Luo when $k < r$.

Ervin Győri, Nika Salia, Casey Tompkins and Oscar Zamora, asymptotically determined the inverse Turán number of paths of length 4 and 5. They also improved the bound for even cycles.

E. Győri, M. Plummer, D. Ye and X. Zha show that a 3-connected claw-free graph always has a cycle passing through any given five vertices but avoiding any other one specified vertex. They also show that this result is sharp by exhibiting an infinite family of 3-connected claw-free graphs in which there is no cycle containing a certain set of six vertices but avoiding a seventh specified vertex.

Andrei Kupavskii, János Pach and Gábor Tardos studied tilings of the plane with a pairwise noncongruent triangles. Answering a question of R. Nandakumar they proved that no such tiling is possible with triangles of equal area and perimeter. They also showed that such a tiling is possible with equal area and bounded perimeter triangles as well as with equal perimeter triangles with area bounded away from zero. Regarding tilings with equilateral triangles with sizes bounded away from zero they show that that it either contains two triangles sharing an edge or it is periodic.

D. Korándi, G. Tardos, I. Tomon and C. Weidert worked on the extremal theory of ordered graph. They proved for a large class of ordered bipartite forests that any ordered graph on n vertices avoiding them must have at most $n^{1+o(1)}$ edges. Tardos also published a survey on the extremal theory of ordered graphs in the proceedings of the International Congress of Mathematitians.

D. Gerbner, A. Methuku, D. Nagy, D. Pálvölgyi, G. Tardos and M. Vizer established the extremal theory of edge-ordered graphs. A conference version of this paper appeared in the proceedings of Eurocomb. Tardos also published a survey paper on the topic in the proceedings of the British Combinatorics Conference.

Sándor Szabó and Bogdán Zaválnij worked on heuristic algorithms for estimating the maximal clique size. Note that finding even reasonable estimates is NP-hard. Some new combinatorial heuristics were found. Generating good benchmark problems for these heuristic algorithms (see above) were also considered and they found constructions of graphs that prove to be extremely hard for current solvers, and also showed the possibility of constructing such graphs for future solvers as well. On the other hand the same problems can be tackled with continuous approach as well. They used LP reformulation to show this possibility, but as the problems tend to be huge they also recommended a new way of solving these kinds of LP problems and published a conference paper of alternative LP solver approach using solution of initial value problem of systems of ODEs. This work is also shown to be a base of efficient sequential algorithms for exact maximum clique search. This method seems to be well parallelizable and suitable for upper and lower bound calculation of clique search problems. They presented their new algorithm on the SYNASC2018 conference. The paper appeared in the IEEE conference proceedings, showing that the proposed algorithms is among the three best worldwide. One important side of the proposed approach is that they used a sequence of k-clique searches to prove a maximum clique optimality. That is, they proposed to use a series of NP-complete decision problems for solving the NP-hard optimization problem. Bogdan Zavalnij was invited to present this idea on the UG Workshop at the Zuze Institute, Berlin in January. Connected to this research a journal paper about constructing infinitely many hard clique search instances were published in the Slovenian Informatica journal. This research, apart from theoretical point of view, has many applications as well. S. Szabó and B. Zaválnij applied the k-clique algorithm, to hypergraph coloring, and proposed a partial answer to the open problem of C-D coloring of hypergraphs by Voloshin. The paper was published in the journal Discrete and Applied Mathematics. They also participated on the ECCO2019 conference, where they applied the k-clique search to hard scheduling problems, namely the Job Shop scheduling problem. Bogdan Zavalnij was also invited to host a lecture at the University of Szeged on the “Data science, Networks and Modeling” seminar. The lecture detailed several possibilities of using graphs and clique search for modeling real life and theoretical applications.

Szabó and Zaválnij also made advances on the reformulation of job scheduling problems into graphs and clique search. This was presented on an international conference, and on the workshop ECMI 2018.

Matjaz Depolli, Sándor Szabó and Bogdán Zaválnij worked on efficient heuristics aiding solution of subgraph isomorphism. The problem is well known to be solvable by graph modeling and clique search. The researchers showed several coloring schemes that can greatly shorten the usual branch-and-bound algorithms. The results were presented on the workshop on ECMI 2018.

Hierarchical clustering approaches are especially useful to understand large-scale structural features of large graphs and appears in various domains. A variant, namely the single-linkage clustering is closely related to the minimal spanning tree of the graph. Andras London and Andras Pluhar investigated the evolution of spanning trees of correlation networks defined through time- series correlation data, and also in a (positional) game-theoretic scenario. Concretely, they investigated special types of Maker-Breaker games defined on graphs. They restrict Makers possible moves that resembles the way that was introduced by Espig, Frieze, Krivelevich and Pedgen. They required that the subgraph induced by Makers edges must be connected throughout the game (see Prim's minimum spanning tree algorithm).

S. Szab, B. Zavalnij and Pablo San Segundo continued their work on different approximation and parallelization methods for maximum clique and k-clique problem. They published a chapter in the book "Ultrascale Computing Systems"', edited by Carretero, J., Jeannot, E. and Zomaya, A., which was published by IET.

Béla Csaba and Judit Nagy-György worked on graph embeddings. A nice result obtained claimed that if the Ore-degree of an vertex graph is at most 5, then it can be embedded in any n vertex graph of minimal degree at least $2n/3$. Another of their results is about which pair of degree-sequences can be simultaneously realized by a simple graph and one of its subgraphs.

Béla Csaba and Bálint Vásárhelyi constructed a class of bounded degree bipartite graphs with a small separator and large bandwidth. Furthermore, they also prove that graphs from this class are spanning subgraphs of graphs with minimum degree just slightly larger than $n/2$.

The notion of discrepancy is used for several contexts, even in the theory of graphs. In a paper accepted for the Electronic Journal of Combinatorics J. Balogh, B. Csaba, Y. Jing and A. Pluhar consider a family S_G of (spanning) subgraphs of a graph G of certain types (e.g., spanning trees, Hamiltonian cycles) and the -1, 1 labels assigned to the edges. They prove bounds on the sum of the labels that hold for all elements of S_G.

Béla Csaba introduced a new graph decomposition method, which works for relatively small or sparse graphs, and can be used to substitute the regularity lemma of Szemerédi in some graph embedding problems.

R. Cymer and M. Krész considered a graph decomposition problem motivated by constraint programming. They gave a complexity analysis based on the structure of perfect matchings is given for the most efficient basic filtering algorithms in constraint programming with respect to the role of edges in matchings.