

## Research Statement

Hemanshu Kaul  
email: kaul@math.iit.edu

My research interests lie in Discrete Mathematics and Operations Research, including Graph Theory, Discrete Optimization, Discrete Geometry, Probabilistic Discrete Structures, and their interdisciplinary applications in Civil Engineering (Transportation Networks), Computer Science (Network Security and Resource allocation), ECE (Renewable Energy distribution), etc.

Discrete optimization models arise naturally in all branches of discrete mathematics and their applications. My motivation is to develop the understanding of hard optimization problems through a variety of approaches - using combinatorial and probabilistic tools to tackle problems, developing and analyzing the complexity of exact and approximation algorithms, and developing both practical and theoretical aspects of heuristics. I have worked with collaborators in Computer Science and in Engineering to apply these methodologies to interdisciplinary applications.

Graph Theory is the language used to describe and study pairwise relations, which explains its numerous and varied applications in engineering, social sciences, and natural sciences. For nontrivial applications of graph theory, it is essential to understand the substructures (and corresponding properties) in graphs and the parameters that imply their existence. My aims in graph theory are two-fold - to deepen the theoretical understanding of graphs as discrete structures, and to apply graph-theoretic techniques to problems in other disciplines.

**Research Activities:** I have, in the past and going forward, worked on both these aspects of research in Discrete Mathematics: theory and applications. I have actively broadened my research interests and found interdisciplinary connections by participating in the professional community and interacting with other researchers as well as mentoring future researchers. Over the past 5 years, I have been invited to give talks in 8 national and international conferences, and 7 colloquia and seminars around the country. I have recently finished editing, with a colleague from the Netherlands, a book featuring 11 survey articles written by experts in consensus and voting theory, clustering, location theory, mathematical biology, and optimization. I have organized special sessions in 3 AMS meetings, a regional conference in Graph Theory, a Chicago-area reading group in discrete mathematics, the department colloquium, and a research seminar. In addition to top research faculty from around the country, and from around Chicago, many of these events have involved both local and external undergraduate and graduate students.

**Research Mentorship:** I have worked with 10 undergraduate students (ranging from Freshmen to Seniors) and 4 graduate students (2 completed M.S. Theses and 2 current Ph.D. students) on research in both Graph Theory and Optimization. Two of my students have won the top two awards from the Illinois Mathematical Association of America for best research by an undergraduate student. I have also worked with a talented high school student on original research in graph theory. IIT has a small Ph.D. program in Applied Math, where many of the students are interested in Financial or Computational Math, so I, as well as my other current colleagues in the discrete math group, have not had an opportunity to work with Ph.D. students yet. However, I have an incoming Ph.D. student in Fall 2011 who will be working with me on his thesis research. In addition, we currently have two Ph.D. students in the first two years of their study who are interested in discrete math but haven't yet chosen their advisors. One of them has also

started his thesis research on combinatorial approximation algorithms with me in Fall 2011.

**Interdisciplinary Research Collaborations:** In addition to my collaborations outside IIT, I have also been building interdisciplinary research relations with members of the IIT faculty. I took the initiative to organize, with my colleagues in Applied Math, a Pan-IIT Interdisciplinary Seminar Series on “Networks and Optimization” which provided a forum for researchers in ECE, Civil Engineering, Computer Science, and Applied Math, to build joint research programs. This seminar directly led to many new interdisciplinary projects including some of my projects.

I have been part of research and grant proposals with faculty from Wireless Network and Communications Research Center (involving faculty from ECE and CS departments). We applied for a well-received (but unsuccessful) grant with NSF under the Network Ecosystems (NECO) section of the inter-disciplinary Networking Technology and Systems (NeTS) program in 2008. This project on “A Paradigm for a Healthy Nested Data Network in Medical Environments” utilized modifications of the art gallery problems to build secure and efficient data networks in a hospital. Although this grant proposal was not successful, it led to fruitful research by two of my students, an undergrad who proved new results on the problem of guarding orthogonal art galleries with holes (described later) and an M.S. student who implemented algorithms for solving the art gallery problem and also developed algorithms for solving a 2-stage stochastic art gallery problem with holes.

In 2007-2009, I worked with collaborators in the business school at IIT to develop an holistic optimization model for resource allocation in the Homeland security department. We wrote a white paper with our proposal for funding through the Department of Homeland Security.

In 2010, I was also invited to join a grant proposal for NSF’s Integrative Graduate Education and Research Training (IGERT) program on “Agent-Based Modeling, Optimization, and Control for Natural, Social and Engineered Systems”. I was involved in both teaching/course development and research (on Complex networks) aspects of this proposal.

Since 2010, I have been exploring Academic-Industry joint research with the founder-director of Salare Security, a software company that has done innovative work on cyber-network security. In 2010, with colleagues in ECE, we submitted a joint IIT-Salare proposal on “Automatic Identification and Mitigation of Unauthorized Information Leaking from Enterprise Networks” to Department of Homeland Security’s Long-Range BAA. Since then we have also looked into joint funding opportunities from NSF.

I am currently a lead member of a project on “Sustainable Transportation Decision-Making” with faculty members from Civil Engineering and Computer Science. We won the competitive ERIF grant for 2010 to build our research program. This work has already led to the submission of three papers, two algorithmic and one computational, on the issue of resource management in an environment with dependencies like transportation networks. We applied for an NSF grant in the highly competitive CDI (Cyber-enabled Discovery and Innovation) program that encourages large-scale interdisciplinary research in 2011. Although this application was unsuccessful, we are encouraged by the feedback to apply for a NSF grant under the Civil Infrastructures with a deadline of October 1st, 2011. We are also in contact with the transportation research center at Argonne National Lab for non-NSF funding opportunities.

**Research Projects:** In the following document, I will describe some of the main results of my research over the past five or so years since my graduation from UIUC and joining IIT and

a few of the problems that I am currently working on. Please refer to my ‘summary of papers’, attached as an appendix to this document, for an overview of all my papers.

### Discrete Optimization Models: Handling Dependencies

The question of how to optimize over discrete structures with underlying dependencies is an important question that has motivated much of my work in Discrete Optimization. In [27, 28], we worked on the Kauffman  $NK$  model, a stochastic combinatorial optimization model that has been used in mathematical biology, physics, and management science to model complex discrete systems with interacting components. It could be crudely described as an optimization problem to find a maximum weighted  $p$ -nary vector of length  $N$  with the weight of a vector defined in terms of its components’ weights and their ‘interaction’ with  $K$  ‘neighboring’ components.

We pioneered the use of tools such as stochastic networks, order statistics, graph coloring, concentration inequalities, high-dimensional geometry, that are distinct from the traditionally used tools like Markov chains applicable when  $K = 1$ ,  $N$  tends to infinity and the underlying distributions are exponential or negative exponential, to obtain new results that were independent of specific underlying distributions and especially applicable to  $K$  growing with  $N$ . We defined and studied a dependency graph to handle dependencies among underlying random variables in the  $NK$  model. Equitable coloring of the dependency graph, via Hajnal-Szemerédi (see [4]), was used to convert the problem of bounding order statistics of dependent random variables into that of independent random variables while incorporating qualitative information about the mutual dependencies between the underlying random variables: *Let  $X_1, \dots, X_n$  be identically distributed random variables with distribution  $F$ , and let the dependency graph of  $X_1, \dots, X_n$  have a  $t$ -equitable coloring. Then*

$$\mu_{\lfloor n/t \rfloor} \leq \mathbf{E}[X_{n:n}] \leq \mu_{\lceil n/t \rceil} + \sqrt{t-1} \sigma_{\lceil n/t \rceil}$$

where  $X_{n:n} = \max \{X_1, \dots, X_n\}$ , and  $\mu_k = \mathbf{E}[\max \{Y_1, \dots, Y_k\}]$ ,  $\sigma_k^2 = \text{Var}[\max \{Y_1, \dots, Y_k\}]$ , such that  $Y_1, \dots, Y_k$  i.i.d.  $F$ . This new method of bounding the order statistics of dependent random variables helped us to give sharp bounds when the underlying distribution is Normal using classical bounds on order statistics of independent random variables, and when the underlying distribution is standard uniform by bounding the volumes of certain high-dimensional bodies in the  $N$ -dimensional hypercube. For bounded underlying distributions, the global optimum value was shown to be concentrated around its mean.

My current major interdisciplinary project is also motivated by a similar question of how dependencies affect combinatorial optima. Here the dependencies between different transportation projects can affect their combined benefit. With collaborators in Civil Engineering and Computer Science, we are working on the problem of how to decide which projects to invest in and implement, so as to maximize the total utility of the highway transportation network after project implementation. In the past, decisions were based on calculating the benefit of each highway project independently, by measuring its local impacts (see [44, 33, 34]). However, this ignores two important factors. Firstly, local changes in a transportation network can lead to agglomerative changes in its global behavior. Secondly, multiple projects within a certain geographical area of the transportation network may be proposed for implementation simultaneously, which means that such projects cannot be considered independent of each other.

We have introduced new multi-commodity flow problems defined over the transportation network under study with different non-linear cost functions that are used to measure the global benefit of collections of projects. This model incorporates the global effect of a local project on the traffic flow in a transportation network. In the current methodology in transportation engineering, once the benefits were calculated based on only their local impact, the projects are chosen by solving a classical 0-1 knapsack problem which treats projects as being mutually independent. We introduce a generalization of the classical Knapsack problem where items under consideration are treated as vertices of an (hyper)graph and (hyper)edges indicate dependencies between them. Presence of edges between items picked for a solution changes the total benefit of those items together.

Graph Knapsack Problem (GKP) is defined as: Let  $G = (V, E)$  be an undirected graph, with a weight (budget) function on the vertices,  $w : V \rightarrow \mathbb{Z}^+$ , and a benefit function on vertices and edges  $b : E \cup V \rightarrow \mathbb{Z}$ . The vertices correspond to the items in the Knapsack problem. The benefit of a subgraph  $H = (V_H, E_H)$  is  $b(H) = \sum_{v \in V_H} b(v) + \sum_{e \in E_H} b(e)$  while its weight is  $w(H) = \sum_{v \in V_H} w(v)$ . Note that the benefits can be negative; negative weight edges model the case where two projects' benefits are less than the sum of their parts. The graphical-knapsack problem asks to find a subset of vertices  $S \subseteq V$  that maximizes the benefit of the induced subgraph,  $b(G[S])$  with the restriction that its weight  $w(G[S])$  is less than  $W$ .

GKP generalizes many graph optimization problems and other generalizations of the Knapsack problem (see [37, 22]), which are all NP-hard and most are even hard to approximate. In particular, it can be easily converted to the well-studied Quadratic Knapsack Problem (QKP) [38, 11] when the benefits are all non-negative. GKP also generalizes the Heaviest Subgraph problem (see [17]), finding the heaviest induced subgraph of an edge-weighted graph. This corresponds to GKP with edges having non-negative benefits while vertices have zero benefit, and, critically, the weight of each vertex is 1. We can also reduce the notoriously hard max-clique problem to GKP.

We have proposed and analyzed algorithms which give us a guaranteed approximation factor [25]. We have a FPTAS (fully polynomial time approximation scheme) when the underlying graph has bounded tree-width. We can also generalize this result to the hypergraph version of the GKP, where the dependencies between items are not just pairwise but  $k$ -wise for any  $k \geq 2$ . We also give a polynomial-time greedy algorithm which has a guaranteed approximation ratio of  $O(\min\{n, W\}/t)$  for any fixed  $t$ . These are the best known approximation ratios for the QKP. We also generalize these results to the hypergraph version of the problem. In our latest paper [24], we give a randomized algorithm based on non-linear Hyperbolic programs with non-linear rounding schemes and analyze it with the help of the Kim-Vu polynomial concentration bound [42] to improve the approximation ratio to  $O(w_{max}n^{1/2})$ , again the best approximation ratio known for QKP.

And in parallel to these theoretical developments, we have also built our collaboration with I-DOT (Illinois Department of Transportation) to study highway networks in Illinois using our research. Using these new mathematical models and algorithms, we have done a computational study [26] of the road network in the Chicago downtown loop area using latest real-life traffic data from I-DOT to analyze the decision-making for multiple projects under consideration there. The results clearly show that our models capture the effects of dependency between different projects. Ignoring these effects gives a false inflated benefit from a collection of projects leading to erroneous decision-making: the network-wide benefits with project interdependency consider-

ations tend to be lower than the corresponding benefits without interdependency considerations by 38 to 64 percent, and the network-wide benefits with project interdependency considerations begin to flatten out when the annualized budget reaches approximately \$7.5M with no additional benefits generated from travel time savings with higher levels of investment budgets.

## Graph Theory: Packing Problems

Graph packing is the edge-disjoint embedding of graphs. It is a fundamental concept in the theory of graphs that has been applied to the study of computational complexity, or evasiveness, of graph properties (see [35]) among others. The concept of graph packing generalizes various extremal graph problems, including problems on fixed subgraphs (such as the Hamiltonian Cycle problem), forbidden subgraphs (Turán-type problems), and equitable coloring (recall the application of equitable coloring to the  $NK$  model).

In [29] we extended the classical result of Sauer-Spencer(1976) on graph packing by characterizing its extremal graphs: If  $2\Delta(G)\Delta(H) \leq n$ , then  $G$  and  $H$  do not pack if and only if one of  $G$  or  $H$  is a perfect matching and the other either is  $K_{\frac{n}{2}, \frac{n}{2}}$  with  $\frac{n}{2}$  odd or contains  $K_{\frac{n}{2}+1}$ . This gives the Sauer-Spencer theorem as a corollary. In [30] we gave the current best result towards a long-standing conjecture of Bollobás and Eldridge [5] from 1978 (an asymptotic form of the conjecture was announced in a conference in 2008 but no manuscript has appeared yet):  $(\Delta(G) + 1)(\Delta(H) + 1) \leq (0.6n) + 1 \Rightarrow G$  and  $H$  pack. That is, we have, in a sense, improved the bound on the product of maximum degrees in the Sauer-Spencer theorem from  $(0.5)n$  to  $(0.6)n$  (although Sauer-Spencer bound is sharp in the form its written).

Two families of graphs are said to *pack* if there exists a pair of packable graphs, one from each family. Packing of families of graphs has also been applied to problems of evasiveness of graph properties, see [18]. In [3], we studied packing of families of graphs defined by fixed degree sequences which is equivalent to discrete tomography problem of reconstruction of the atomic structure of a crystalline lattice (see [15]). Let  $\pi_1$  and  $\pi_2$  be graphic  $n$ -tuples, with  $\pi_1 = (d_1^{(1)}, \dots, d_n^{(1)})$  and  $\pi_2 = (d_1^{(2)}, \dots, d_n^{(2)})$  (they need not be monotone). We say that  $\pi_1$  and  $\pi_2$  *pack* if there exist edge-disjoint graphs  $G_1$  and  $G_2$  with vertex set  $\{v_1, \dots, v_n\}$  such that the degrees of  $v_i$  in  $G_1$  and  $G_2$  are  $d_i^{(1)}$  and  $d_i^{(2)}$ , respectively. When packing graphs, we permit permuting the vertices to make  $G_1$  and  $G_2$  fit together, but when packing sequences, it is critical that the sequences not be permuted. This can be thought as packing families of labeled graphs.

We proved that two graphic  $n$ -tuples pack if  $\Delta \leq \sqrt{2\delta n} - (\delta - 1)$ , where  $\Delta$  and  $\delta$  denote the largest and smallest entries in  $\pi_1 + \pi_2$  (strict inequality when  $\delta = 1$ ); furthermore, the bound is sharp. If the two graphic sequences do not share any 0 terms then  $\Delta < \sqrt{2n}$  implies the two sequences pack. This can be thought as a Sauer-Spencer type of packing result for graph families.

Kundu's classical theorem characterizes when a graphic  $n$ -tuple has a realization containing a spanning subgraph that is  $k$ -regular. We consider extensions of Kundu's Theorem and conjecture the stronger statement that under the same obvious necessary conditions, when  $n$  is even there is a realization containing  $k$  edge-disjoint 1-factors (that is, a  $k$ -edge-colorable  $k$ -factor). (The case when  $n$  is odd can be handled easily.) This conjecture turns out to be identical to an old conjecture of Brualdi [7] from 1976 which remains unsolved and our results are the best progress towards it. We have proved the conjecture when the largest degree is at most  $n/2 + 1$ . We also prove the more difficult result that the conjecture holds when  $k \leq 3$ , by proving the

stronger statement that there is a realization containing a  $k$ -factor that has two 1-factors and a  $k - 2$ -factor, all edge disjoint.

It has been conjectured that any two trees can be packed in such a way that their union forms a planar graph ([19]), note that it is not possible to pack three trees in such a way. I have partial results towards this problem, including new proofs for planar packing of two copies of same tree, and planar packing of a spider and an arbitrary tree that I hope to extend to the full conjecture. I have also been exploring the topic of packing graphs of high girth in such a way that their union still has high girth. This natural problem has not been studied much in literature. Its easy to apply earlier packing results to get weak results of this nature but getting sharp results is proving challenging. There are several other open problems related to sufficient conditions for packing of graphs in terms of some fundamental graph parameters such as number of edges, average degree, etc. and particular graph classes such as trees (see [4]).

### **Algorithmic Graph Theory: Finding Large Subgraphs**

We have already discussed my work on the problem of finding the largest induced subgraph with fixed number of vertices as the Graph Knapsack Problem earlier. I also have preliminary results on the corresponding extremal question of largest induced subgraph on  $p$  vertices in a given graph on  $n$  vertices and  $m$  edges that is  $k$ -colorable, extending classical results of Chung, Erdős, and Spencer [13]. I am currently working on this problem.

The problem of finding a bipartite subgraph with maximum number of edges in a given graph is a classical problem in combinatorial optimization and extremal graph theory. There are numerous results and open-problems in the area, see [39] for an extensive survey. Since this is a NP-complete problem that is hard to approximate, heuristics are often used in practice to solve this problem. A common heuristic is to start with an arbitrary vertex bipartition  $V = X \cup Y$ , and then make local improvements if possible, by switching a vertex from one side to the other if that increases the number of edges in-between  $X$  and  $Y$ . Then, we are interested in questions like - What is the length of a longest such sequence of switches? What are the minimum number of such switches needed to reach a local optimum/ global optima? In [31], we essentially answer these questions as  $\frac{2}{25}n^2$  (note that  $\frac{1}{4}(n^2)$  is a trivial upper bound), and  $n/2$  (to reach a global optima!) respectively.

Consider the problem of finding in a given graph a  $K_r$ -minor-free subgraph with maximum number of edges. This problem was directly motivated by the classical problem of finding a maximum planar subgraph (see [8]). Since not much progress has been made on the maximum planar subgraph problem in the past ten years (see [9]), insights from maximum  $K_r$ -minor-free subgraph problem could lead to new progress. In [10], we give a  $\frac{7}{12}$ -approximation algorithm (current best) for the maximum  $K_4$ -minor-free subgraph (series-parallel subgraph) problem and, constructions and complexity results showing the limits of our approach. Unlike all previous algorithms for other such problems, the subgraph we generate is neither a tree nor an outer-planar graph. For the first time, we are able to analyze an algorithm that allows blocks of unbounded size in solution subgraph and is allowed to shrink or throw away previously selected blocks (non-greedy).

## Discrete Geometry and Graph coloring

The Art Gallery problem asks for the minimum number of guards sufficient to see every point of the interior of an  $n$ -vertex simple polygon (see [36]). A guard is assumed to be a point in the polygon that can see any point that can be connected with a line segment within the polygon. This is an active area of research in discrete and computational geometry from both theoretical and computational perspectives (like all problems under consideration here it is NP-hard, see [20], [16] for some approximation algorithms). I became interested in this problem as a result of an inter-disciplinary project with faculty in ECE on the design of wireless networks in a hospital building. From a geometric point-of-view, in [23], we, with an undergraduate student, gave an involved graph coloring argument over an auxiliary geometric graph to get the best current result:  $\lfloor (n + \frac{5}{3}h)/4 \rfloor$ , for the number of vertex guards needed for guarding an  $n$ -vertex orthogonal gallery (all boundary edges are vertical or horizontal) with  $h$  orthogonal holes (visual obstructions within the polygons). Shermer(1982) conjectured this to be  $\lfloor (n + h)/4 \rfloor$  and the best general bound so far ([36]) was  $\lfloor (n + 2h)/4 \rfloor$ .

In [12], we worked on the distinguishing chromatic number of a graph  $G$ ,  $\chi_D(G)$ , ([14]), the least number of colors needed for a proper coloring of  $G$  with the property that the only color-preserving automorphism of  $G$  is the identity. That is, we want to give a proper coloring of a graph that breaks all its symmetries, so that the coloring together with the structure of the graph uniquely determines the vertices. This can be thought of as an exact encoding of the vertices using only a proper coloring. We have results which illustrate that  $\chi_D(G)$  can be surprisingly at most 1 worse than  $\chi(G)$  under the cartesian products of graphs. The main theorem is : For every graph  $G$ , there exists a constant  $d_G$  (explicitly given) such that for all  $d \geq d_G$ ,  $\chi_D(G^d) \leq \chi(G) + 1$ , where  $G^d$  denotes the Cartesian product of  $d$  copies of  $G$ . Using our proof techniques, we also find the distinguishing chromatic number for Hypercubes, Hamming graphs (Cartesian products of complete graphs), and Cartesian products of complete multipartite graphs.

## References

- [1] Ahuja, R.K., Magnanti, T., Orlin, J.B. (1993): Network Flows, Prentice Hall.
- [2] Alon, N., Krivelevich, M., Sudakov, B. (2005), Max-Cut in H-free graphs, Comb. Prob. Comp., 14, 629–647.
- [3] Busch, A., Ferrara, M., Hartke, S., Jacobson, M.S., Kaul, H., West, D.B., (2011) Packing of Graphic Sequences, Journal of Graph Theory, to appear.
- [4] Bollobás, B. (1978): Extremal graph theory. LMS Monographs, 11, Academic Press, London-New York.
- [5] Bollobás, B., Eldridge, S.E. (1978): Packings of graphs and applications to computational complexity, J. Combin. Theory Ser. B 25, 105–124.
- [6] Bollobás, B., Scott, A.D. (2002): Problems and results on judicious partitions, Random Structures and Algorithms 21, 414-430.
- [7] Brualdi, R.A., (1976) Contribution to "Problèmes, in Problemes Combinatoires et Theorie des Graphes (Orsay, 1976), Colloq CNRS 260, 437.
- [8] Calinescu, G., Fernandes, C.G., (2007): Maximum Planar Subgraph (Chapter 56). In: Teofilo F. Gonzalez, Handbook of Approximation Algorithms And Metaheuristics, Chapman and Hall/CRC Press.

- [9] Calinescu, G., Fernandes, C.G., Finkler, U., Karloff, H. (1998): A Better Approximation Algorithm for Finding Planar Subgraphs, *Journal of Algorithms*, 269–302.
- [10] Calinescu, G., Fernandes, C.G., Kaul, H. (2011) Maximum Series-Parallel Subgraph: Approximation Algorithms, *Algorithmica*, to appear, 2011.  
A refereed conference version in: *Graph-Theoretic Concepts in Computer Science 2009*, *Lecture Notes in Computer Science*, Volume 5911, 2010, 54-65.
- [11] Caprara, A., Pisinger, D., Toth, P. (1999) Exact Solution of the Quadratic Knapsack Problem, *Inform. Journal on Computing*, 11, 125-137.
- [12] Choi, J., Hartke, S., Kaul, H. (2010) Distinguishing Chromatic Number of Cartesian Products of Graphs, *SIAM Journal on Discrete Mathematics*, Volume 24, 2010, 82-100.
- [13] Chung, F.R.K., Erdős, P., Spencer, J. (1985): Extremal Subgraphs for Two Graphs, *J. of Combinatorial Theory B*, 38, 248-260.
- [14] Collins, K., Trenk, A. (2006): The distinguishing chromatic number, *Electr. J. Comb.* 13.
- [15] Durr, C., Guinez, F., Matamala, M. (2009): Reconstructing 3-colored grids from horizontal and vertical projections is NP-hard, *Annual European Symposium (ESA 2009)*, LNCS 5757, 776-787.
- [16] Efrat, A., Har-Peled, S. (2002): Guarding galleries and terrains, *IFIP Conference Proceedings*, 23, 181–192.
- [17] Feige, U., Kortsarz, G., Peleg, D. (2001): The dense k-subgraph problem. *Algorithmica* 29, 410-421.
- [18] Friedgut, E., Kahn, J., Wigderson, A. (2002): Computing Graph Properties by Randomized Subcube Partitions, in *Randomization and Approximation Techniques in Computer Science*, 6th International Workshop, RANDOM, LNCS 2483, 105–113.
- [19] García, A., Hernando, C., Hurtado, F., Noy, M., Tejel, J. (2002): Packing trees into planar graphs, *J. Graph Theory*, 40, 172–181.
- [20] Ghosh, S. K. (1987): Approximation algorithms for art gallery problems, *Proc. Canadian Inform. Process. Soc. Congress*.
- [21] Ghosh, S.K., Burdick, J.W., Bhattacharya, A., Sarkar, S. (2008): Online Algorithms with discrete visibility: Exploring unknown polygonal environments, *IEEE Robotics and Automation Magazine*, 15, 67–76.
- [22] Hassin, R., Rubinstein, S., Tamir, A. (1997) Approximation algorithms for maximum dispersion. *Operations Research Letters* 21, 133-137.
- [23] Jo, Y., Kaul, H., Improved Bound for Guarding Orthogonal Art Galleries with Holes, submitted for publication.
- [24] Kapoor, S., Kaul, H., Graph Knapsack Problem: A Randomized Approximation Algorithm, submitted for publication.
- [25] Kapoor, S., Kaul, H., Pelsmajer, M., Graph Knapsack Problem, submitted for publication.
- [26] Kapoor, S., Kaul, H., Li, Z. and Veliou, E., Zhou, B., Lee, S., A New Methodology For Transportation Investment Decisions Considering Project Interdependencies, submitted for publication.
- [27] Kaul, H., Jacobson, S.H. (2006) Global Optima Results for the Kauffman  $NK$  Model, *Mathematical Programming*, Volume 106, 319-338.
- [28] Kaul, H., Jacobson, S.H. (2006) New Global Optima Results for the Kauffman  $NK$  Model: Handling Dependency, *Mathematical Programming*, Special issue on ‘Optimization under Uncertainty’, Volume 108, 475-494.
- [29] Kaul, H., Kostochka, A. (2007) Extremal Graphs for a Graph Packing Theorem of Sauer and Spencer, *Combinatorics, Probability and Computing*, Volume 16, 409-417.
- [30] Kaul, H., Kostochka, A., Yu, G. (2008) On a Graph Packing Conjecture of Bollobas, Eldridge, and Catlin, *Combinatorica*, Volume 28, 2008, 469-485.
- [31] Kaul, H., West, D.B. (2008) Long Local Searches for Large Bipartite Subgraph, *SIAM Journal on Discrete Mathematics*, Volume 22, 2008, 1138-1144.
- [32] Kaul, H., Mulder, H.M. (2010) Advances in Interdisciplinary Applied Discrete Mathematics, *Interdisciplinary Mathematical Sciences*, Volume 11, World Scientific Publishing, 275pp.

- [33] Li, Z., S. Madanu. 2009. Highway Project-Level Life-Cycle Benefit/Cost Analysis under Certainty, Risk, and Uncertainty: A Methodology with Case Study. *ASCE Journal of Transportation Engineering* 135(8), 516-526.
- [34] Li, Z., Sinha, K.C. (2004): A Methodology for Multicriteria Decision Making in Highway Asset Management. *Transportation Research Record* 1885, 79-87.
- [35] Lovasz, L., Young, N.E. (2002): Lecture Notes on Evasiveness of Graph Properties, Technical Report, <http://arxiv.org/abs/cs.CC/0205031> .
- [36] O'Rourke, J. (1987): *Art Gallery Theorems and Algorithms*, Oxford University Press.
- [37] Pferschy, U., Schauer, J. (2009): The Knapsack Problem with Conflict Graphs. *J. Graph Algorithms Appl.* 13, 233-249.
- [38] Pisinger, D. (2007) The quadratic knapsack problem - a survey, *Discrete Appl. Math.*, 155, 623-648.
- [39] Poljak, S., Tuza, Z. (1995): Maximum cuts and largest bipartite subgraphs, in W. Cook et al., editors, *Combinatorial Optimization*, DIMACS Series Vol. 20, 181-244.
- [40] Simonovits, M. (2002): Some of my favorite Erdos theorems and related results, theories, Paul Erdos and his mathematics, II (Budapest, 1999), 565–635, *Bolyai Soc. Math. Stud.*, 11.
- [41] Urrutia, J. (2000) Art Gallery and Illumination Problems, In: *Handbook of Computational Geometry*, Amsterdam, Elsevier 2000.
- [42] Vu, Van (2000) On the concentration of multivariate polynomials with small expectation. *Random Struct. Algorithms*, 16, 344-363.
- [43] Wahlström, M. (2004): Exact algorithms for finding minimum transversals in rank-3 hypergraphs, *J. Algorithms*, 51, 107–121.
- [44] Weissmann, J., R. Harrison, N.H. Burns, W.R. Hudson. (1990). Selecting Rehabilitation and Replacement Bridge Projects, *Extending the Life of Bridges*. ASTM STP 1100, 3-17.

## Appendix : Summary of Papers

- [1] **Global Optima Results for the Kauffman  $NK$  Model** (with S.H. Jacobson), *Mathematical Programming*, Volume 106, 2006, 319-338.

The Kauffman  $NK$  model is a stochastic combinatorial optimization model that has been used in theoretical biology, physics and management science to model complex systems with interacting components. It could be crudely described as an optimization problem to find a maximum valued (weighted)  $p$ -nary vector of length  $N$  with the value (weight) of a vector defined in terms of its components' weights and their 'interaction' with  $K$  'neighboring' components. This paper analyzes global optima of the  $NK$  model. Most previous papers focused on local optima and simulation based results.

We transform this NP-hard global optimization problem into a stochastic network model that is closely related to two well-studied problems in operations research - project duration in PERT networks and stochastic shortest path problem. This transformation leads to applicable strategies for explicit computation of bounds on the global optima (particularly with  $K$  either small or close to  $N$ ), such as a recursive scheme applicable as a dynamic program and simple stochastic networks that can be processed simultaneously. A general lower bound, which is sharp for  $K = 0$ , is obtained for the expected value of the global optimum of the  $NK$  model in terms of the order statistics of the underlying distribution. We also give a detailed analysis for the expectation and variance of the global optimum when  $K = N - 1$  and the underlying distribution is  $\mathbf{U}(0, 1)$ , by converting the analytic problem into a geometric one with estimation of volumes of certain bodies in the  $N$ -dimensional hypercube. The lower and upper bounds on the expectation obtained for this case show that there is a wide gap between the values of the local and the global optima. They also indicate that the complexity catastrophe, the tendency of the local

optima to collapse towards average behavior, does not arise for the global optima.

**[2] New Global Optima Results for the Kauffman  $NK$  Model: Handling Dependency**, (with S.H. Jacobson), *Mathematical Programming*, Special issue on 'Optimization under uncertainty', Volume 108, 2006, 475-494.

This paper generalizes and extends the work from the first paper [1] that focused on the analysis of the (independent) case  $K = N - 1$ . It presents new global optima results for the  $NK$  model by developing tools for handling the dependency between weight functions of different  $N$ -vectors due to overlapping weight contributions from their components, when  $K \leq N - 2$ . Previous papers used Markov chain theory to analyze the cases when  $K = 1$ ,  $N$  tends to infinity and the underlying distributions are exponential or negative exponential. The ideas developed here are more combinatorial in nature, independent of specific underlying distributions and especially applicable to  $K$  growing with  $N$ .

We define and study a dependency graph to handle dependencies among underlying random variables in the  $NK$  model. Equitable coloring of the dependency graph is used to bound general order statistics (with dependencies) and consequently, the expected value of the global optima,  $\mathbf{E}_{N,K}$ , for the  $NK$  model. These bounds convert the problem of bounding order statistics of dependent random variables into that of independent random variables while incorporating quantitative information about the mutual dependencies between the underlying random variables. An alternative upper bound on  $\mathbf{E}_{N,K}$  using direct arguments is also proposed. These ideas for handling dependence are applied to give a detailed analysis of  $\mathbf{E}_{N,K}$  for  $K$  close to  $N$  ( $K = N - \alpha$  and  $K = \beta N$ ,  $\alpha \in \mathbf{Z}^+$ ,  $\beta \in (0, 1)$ ) with sharp bounds when the underlying distribution is normal by using tools from order statistics theory, and when the underlying distribution is uniform by extending the geometric ideas from the first paper [1]. Finally, for bounded underlying distributions, the global optima is shown to be concentrated around its mean  $\mathbf{E}_{N,K}$ .

**[3] Extremal Graphs for a Graph Packing Theorem of Sauer and Spencer**, (with A. Kostochka), *Combinatorics, Probability and Computing*, Volume 16, 2007, 409-417.

Let  $G, H$  be graphs with maximum degrees  $\Delta(G) = \Delta_1$ ,  $\Delta(H) = \Delta_2$ , and orders  $n(G), n(H) \leq n$ .  $G$  and  $H$  are said to pack if there exist injective mappings of the vertex sets into  $[n]$ , such that the images of the edge sets do not intersect. In other words, either  $G$  or  $H$  is isomorphic to a subgraph of the complement of the other. The concept of graph packing generalizes various extremal graph problems, including problems on existence of fixed subgraphs (such as the Hamiltonian Cycle problem), forbidden subgraphs (Turán-type problems), and equitable coloring. One of the classical results in this area was by Sauer and Spencer (1978) : if  $2\Delta_1\Delta_2 < n$  then  $G$  and  $H$  pack.

We characterize the graphs that achieve equality in the condition on maximum degrees as given in the Sauer-Spencer result for packing of graphs. We show that: If  $2\Delta_1\Delta_2 \leq n$ , then  $G$  and  $H$  do not pack if and only if one of  $G$  or  $H$  is a perfect matching and the other either is  $K_{\frac{n}{2}, \frac{n}{2}}$  with  $\frac{n}{2}$  odd or contains  $K_{\frac{n}{2}+1}$ . This gives the Sauer-Spencer theorem as a corollary. This result can be thought of as small step towards the well-known Bollobás-Eldridge conjecture (described in [5] below).

**[4] Analyzing the Performance of Simultaneous Generalized Hill Climbing Algorithms**, (with D.E. Vaughan and S.H. Jacobson), *Computational Optimization and Applications*, Volume 37, 2007, 103-119.

Simultaneous generalized hill climbing (SGHC) algorithms provide a framework for using heuristics to simultaneously address sets of intractable discrete optimization problems where information is shared between the problems during the algorithm execution. A SGHC algorithm probabilistically moves between discrete optimization problems during its execution according to a (problem generation) probability function. Many well-known heuristics (Generalized Hill Climbing (GHC) algorithms), including simulated annealing, threshold accepting and pure local search, can be embedded within the SGHC algorithm

framework.

This paper shows that the solutions generated by an SGHC algorithm are a stochastic process that satisfies the Markov property. This allows the problem probability mass functions to be formulated for particular sets of problems based on the long-term behavior of the algorithm. Such results can be used to determine the proportion of iterations that an SGHC algorithm will spend optimizing over each discrete optimization problem. Sufficient conditions that guarantee that the algorithm spends an equal number of iterations in each discrete optimization problem are provided. SGHC algorithms can also be formulated such that the overall performance of the algorithm is independent of the initial discrete optimization problem chosen. Sufficient conditions are obtained guaranteeing that an SGHC algorithm will visit the globally optimal solution for each discrete optimization problem. Lastly, rates of convergence for SGHC algorithms are reported that show that given a rate of convergence for the embedded GHC algorithm, the SGHC algorithm can be designed to preserve this rate.

**[5] On a Graph Packing Conjecture of Bollobás, Eldridge, and Catlin**, (with A. Kostochka and G. Yu), *Combinatorica*, Volume 28, 2008, 469-485.

See the description in [4] above for definition of the graph packing problem. One of the earliest results in this area was by Sauer and Spencer(1978) : if  $\Delta_1\Delta_2 < \frac{1}{2}n$  then  $G$  and  $H$  pack. The main conjecture in the area was made by Bollobás and Eldridge in 1978 as an extension of the Sauer–Spencer result: if  $(\Delta_1 + 1)(\Delta_2 + 1) \leq n + 1$  then  $G$  and  $H$  pack. If true, this conjecture would be sharp, and would be a considerable extension of the Hajnal-Szemerédi theorem on equitable colorings. The conjecture has only been proved when  $\Delta_1 \leq 2$ , or  $\Delta_1 = 3$  and  $n$  is huge.

This paper focuses on proving a result of the form : for a fixed  $0 \leq \epsilon \leq 1$ ,  $(\Delta_1 + 1)(\Delta_2 + 1) \leq \frac{n}{2}(1 + \epsilon) + 1 \Rightarrow G$  and  $H$  pack. For  $\epsilon = 0$ , this is essentially the Sauer-Spencer result, while  $\epsilon = 1$  gives the Bollobás-Eldridge conjecture. Thus, for any  $\epsilon > 0$  this would improve the Sauer-Spencer result. We have proved this result for  $\epsilon = 0.2$ . That is, we have essentially improved the bound on the product of maximum degrees in the Sauer-Spencer theorem from  $(0.5)n$  to  $(0.6)n$ .

**[6] Long Local Searches for Large Bipartite Subgraphs**, (with D.B. West), *SIAM Journal on Discrete Mathematics*, Volume 22, 2008, 1138-1144.

We study a basic local-search algorithm for finding a bipartite subgraph with maximum number of edges in a given graph. Starting with an arbitrary vertex partition, move (‘flip’) a vertex from one partite set to the other if doing so increases the number of edges in the cut. We improve the previous best-known lower bound,  $\frac{1}{2}n^{3/2}$ , on the maximum number of flips possible in a graph with  $n$  vertices to  $\frac{2}{25}n^2$ . Note that  $\frac{1}{4}(n^2)$  is a trivial upper bound. We also prove better upper bounds like  $e(G) - \frac{1}{4}d^2$ , in terms of  $d$  the minimum degree of the graph. We also show that the minimum number of flips needed to reach a global optimum is at most  $n/2$ , answering a question of Cowen and West.

**[7] Distinguishing Chromatic Number of Cartesian Products of Graphs**, (with J. Choi and S. Hartke), *SIAM Journal on Discrete Mathematics*, Volume 24, 2010, 82-100.

The distinguishing chromatic number of a graph  $G$ ,  $\chi_D(G)$ , is the least number of colors needed for a proper coloring of  $G$  with the property that the only color-preserving automorphism of  $G$  is the identity. That is, we want to give a proper coloring of a graph that breaks all its symmetries, so that the coloring together with the structure of the graph uniquely determines the vertices. This can be thought of as an exact encoding of the vertices using only a proper coloring. It is a common extension of both the chromatic number and the distinguishing number of graphs.

The chromatic number,  $\chi(G)$ , is an immediate lower bound for  $\chi_D(G)$ . We show that  $\chi_D(G)$  can be surprisingly at most one worse than  $\chi(G)$  for  $G$  a Cartesian power of any graph. The main theorem is : For every graph  $G$ , there exists a constant  $d_G$  (explicitly given) such that for all  $d \geq d_G$ ,  $\chi_D(G^d) \leq \chi(G) + 1$ ,

where  $G^d$  denotes the Cartesian product of  $d$  copies of  $G$ . Using our proof techniques, we also find the distinguishing chromatic number for Hypercubes, Hamming graphs (Cartesian products of complete graphs), and Cartesian products of complete multipartite graphs.

**[8] Maximum Series-Parallel Subgraph: Approximation Algorithms**, (with G. Calinescu and C.G. Fernandes), *Algorithmica*, to appear, 2011.

A refereed conference version in: *Graph-Theoretic Concepts in Computer Science 2009*, Lecture Notes in Computer Science, Volume 5911, 2010, 54-65.

Consider the NP-hard problem of, given a simple graph  $G$ , to find a  $K_4$ -minor-free subgraph (series-parallel subgraph) of  $G$  with the maximum number of edges. The algorithm that, given a connected graph  $G$ , outputs a spanning tree of  $G$ , is a  $\frac{1}{2}$ -approximation. Indeed, if  $n$  is the number of vertices in  $G$ , any spanning tree in  $G$  has  $n - 1$  edges and any series-parallel graph on  $n$  vertices has at most  $2n - 3$  edges. We present a  $\frac{7}{12}$ -approximation algorithm (current best) for this problem and, constructions and computational complexity results showing the limits of our approach.

Unlike earlier algorithms for various planar subgraph problems, the subgraph we generate is not a tree or an outerplanar graph. For the first time, we are able to analyze an algorithm that allows blocks of unbounded size in solution subgraph and is allowed to shrink or throw away previously selected blocks.

**[9] Advances in Interdisciplinary Applied Discrete Mathematics**, (co-editor with H.M. Mulder), *Interdisciplinary Mathematical Sciences*, Volume 11, World Scientific Publishing, 2010, 275pp.

In the past 50 years, discrete mathematics has developed as a far-reaching and popular language for modeling fundamental problems in computer science, biology, sociology, operations research, economics, engineering, etc. This book focuses on fields such as consensus and voting theory, clustering, location theory, mathematical biology, and optimization that have seen an upsurge of new and exciting works over the past two decades using discrete models in modern applications. Featuring 11 survey articles written by experts in these fields, the articles emphasize the interconnectedness of the mathematical models and techniques used in various areas, and elucidate the possibilities for future interdisciplinary research.

**[10] Reductions for the Stable Set Problem**, (with E.C. Sewell and S.H. Jacobson), *Algorithmic Operations Research*, Vol 6(1), 2011, 40-55.

One approach to finding a maximum stable (independent) set (MSS) (or, equivalently a maximum clique or a minimum vertex cover) in a graph is to try to reduce the size of the problem by transforming the problem into an equivalent problem on a smaller graph. These reductions have been used to study properties of stability critical graphs and facets of the stable set polytope. They have also been used algorithmically in heuristics, polynomial-time algorithms for special classes of graphs, and exact algorithms.

This paper introduces several new reductions for the MSS problem, extends several well known reductions to the maximum weight stable set (MWSS) problem, demonstrates how reductions for the generalized stable set problem can be used in conjunction with probing to produce powerful new reductions for both the MSS and MWSS problems, and shows how hypergraphs can be used to expand the capabilities of clique projections. The effectiveness of these new reduction techniques are illustrated on a set of challenging MSS problems arising from Steiner Triple Systems.

**[11] Packing of Graphic  $n$ -tuples**, (with A. Busch, M. Ferrara, S. Hartke, M.S. Jacobson, and D.B. West), *Journal of Graph Theory*, to appear, 2011.

Let  $\pi_1$  and  $\pi_2$  be graphic  $n$ -tuples, with  $\pi_1 = (d_1^{(1)}, \dots, d_n^{(1)})$  and  $\pi_2 = (d_1^{(2)}, \dots, d_n^{(2)})$  (they need not be monotone). We say that  $\pi_1$  and  $\pi_2$  *pack* if there exist edge-disjoint graphs  $G_1$  and  $G_2$  with vertex set

$\{v_1, \dots, v_n\}$  such that the degrees of  $v_i$  in  $G_1$  and  $G_2$  are  $d_i^{(1)}$  and  $d_i^{(2)}$ , respectively. When packing graphs, we permit permuting the vertices to make  $G_1$  and  $G_2$  fit together, but when packing sequences, we do not permit the sequences to be permuted. From an optimization point-of-view, this can be thought of as an feasibility problem. This problem is related to certain multicommodity flow problems with applications in supply chain/logistics, and even x-ray tomography. However, packing even 2 bipartite sequences is NP-hard.

We prove that two graphic  $n$ -tuples pack if  $\Delta \leq \sqrt{2\delta n} - (\delta - 1)$ , where  $\Delta$  and  $\delta$  denote the largest and smallest entries in  $\pi_1 + \pi_2$  (strict inequality when  $\delta = 1$ ); furthermore, the bound is sharp. If the two graphic sequences do not share any 0 terms then  $\Delta < \sqrt{2n}$  implies the two sequences pack. This can be thought of as Sauer-Spencer type of packing result (see [4]).

Kundu's Theorem (1973) characterizes when a graphic  $n$ -tuple has a realization containing a spanning subgraph that is  $k$ -regular. We consider extensions of Kundu's Theorem and conjecture the stronger statement that in fact when  $n$  is even there is a realization containing  $k$  edge-disjoint 1-factors (that is, a  $k$ -edge-colorable  $k$ -factor). We prove the conjecture when the largest entry is at most  $n/2 + 1$ . We also prove the more difficult result that the conjecture holds when  $k \leq 3$ , by proving the stronger statement that there is a realization containing a  $k$ -factor that has two edge-disjoint 1-factors.

**[12] Forbidden Subgraphs of Unit Disk Graphs**, (with R. Martin), submitted for publication.

A graph  $G = (V, E)$  is called a unit disk graph (UDG) if there is a function  $f : V(G) \rightarrow \mathbb{R}^2$  such that for  $v, w \in V(G)$ ,  $\|f(v) - f(w)\| \leq 1$  if and only if  $\{v, w\} \in E(G)$ . UDGs are a natural model for wireless and radio networks. UDG is an example of a hereditary property. For every hereditary property  $\mathcal{H}$ , there exists a family of graphs  $\mathcal{F}(\mathcal{H})$  such that  $\mathcal{H} = \bigcap_{H \in \mathcal{F}(\mathcal{H})} \text{Forb}(H)$ , where  $\text{Forb}(H)$  is the family of graphs with no *induced* copy of  $H$ . This together with the fact that recognition of UDGs is NP-hard motivates the study of family of forbidden induced subgraphs of UDG. For example,  $K_{1,6}$  was proved to be a forbidden induced subgraph of UDG and this fact was used to give approximation algorithms for various NP-hard problems on UDGs.

We prove that  $K_{2,3}$  among other small graphs are forbidden in UDGs. We also describe a general construction for forbidden subgraphs that gives a total of  $2^{\binom{4}{2}} 2^{\binom{15}{2}}$  graphs on 19 vertices that are forbidden. All these results are generalized to Unit Ball graphs in  $d$ -dimensions. These results are then applied to an application of Szemerédi's regularity lemma to the edit-distance problem for large UDGs.

**[13] Guarding Orthogonal Art Galleries with Holes**, (with Y. Jo), submitted for publication.

The original art gallery problem (V.Klee, 1973) asked for the minimum number of guards sufficient to see every point of the interior of an  $n$ -vertex simple polygon (Art Gallery). Chvátal (1975) proved that  $\lfloor n/3 \rfloor$  guards are always sufficient. If all the edges of the given simple polygon are either horizontal or vertical, then such a polygon is called an orthogonal gallery. Kahn, Klawe and Kleitman (1983) proved that  $\lfloor n/4 \rfloor$  guards are sufficient for such a  $n$ -vertex gallery.

We study orthogonal gallery with holes, i.e., an orthogonal polygon enclosing some other orthogonal polygons called holes (interior of each hole is empty). In 1982, Shermer conjectured that any orthogonal polygon with  $n$  vertices and  $h$  holes can be guarded by  $\lfloor (n+h)/4 \rfloor$  vertex guards. This conjecture remains open. The best known result shows that  $\lfloor (n+2h)/4 \rfloor$  such guards suffice (O'Rourke 1987). We improve this bound to  $\lfloor (n + \frac{5}{3}h)/4 \rfloor$  using a graph coloring argument over a geometric graph.

**[14] Fall coloring of Graphs**, (with C. Mitillos), submitted for publication.

Fall coloring of graphs is a common extension of two fundamental NP-complete problems - Graph Coloring and Graph Domination. It asks for a partition of the vertices of a graph into sets which are both independent and dominating. The primary question in this topic is the feasibility of such a coloring existing for a given graph. Let  $Fall(G)$  be the set of values  $k$  such that  $G$  is  $k$ -Fall colorable ( $k$  is the

number of colors used). We construct a family of graphs with predetermined  $Fall(G)$  which shows that  $Fall(G)$  can be arbitrarily large with arbitrarily large gaps. We characterize Threshold graphs and Split graphs that are Fall colorable. We also show that the Cartesian product of a  $k$ -Fall-colorable graph with a  $k$ -colorable graph gives a  $k$ -Fall-colorable graph.

**[15] Approximating the Quadratic Knapsack Problem and its Generalization**, (with S. Kapoor and M. Pelsmajer), submitted for publication.

We study a generalization of the classical 0-1 knapsack problem. We generalize the benefit function over a subset of items by allowing for dependencies between items that could possibly modify their contribution to the benefit of the whole ensemble. In addition to the sum of the individual elements, the benefit function also includes terms involving the benefit contributions of various subsets of items. Our motivating example is a resource allocation problem, where the items are proposed projects, weight is the cost of implementation, and benefit measures the expected savings by implementing that project. Typically, it has been assumed that the projects are pairwise independent so their benefits are additive. However, in most applications the benefit of two projects can be less than or greater than the sum of the individual projects due to underlying dependencies among the projects.

To model this situation, we define the *Graph Knapsack Problem* (GKP). Let  $G = (V, E)$  be an undirected graph, with a weight function on the vertices,  $w : V \rightarrow \mathbb{Z}^+$ , and a benefit function on vertices and edges  $b : E \cup V \rightarrow \mathbb{Z}$ . The vertices correspond to the items in the Knapsack problem. The benefit of a subgraph  $H = (V_H, E_H)$  is  $b(H) = \sum_{v \in V_H} b(v) + \sum_{e \in E_H} b(e)$  while its weight is  $w(H) = \sum_{v \in V_H} w(v)$ . Note that the benefits can be negative; negative weight edges model the case where two projects' benefits are less than the sum of their parts. The graphical-knapsack problem asks to find a subset of vertices  $S \subseteq V$  that maximizes the benefit of the induced subgraph,  $b(G[S])$  with the restriction that its weight  $w(G[S])$  is less than  $W$ .

GKP generalizes many graph optimization problems and other generalizations of the Knapsack problem, which are all NP-hard and most are even hard to approximate. GKP can be expressed the well-studied Quadratic Knapsack Problem (QKP) when the benefits are non-negative. And, GKP generalizes the Heaviest Subgraph problem, finding the heaviest induced subgraph of an edge-weighted graph. This corresponds to GKP with edges having non-negative benefits while vertices have zero benefit, and the weight of each vertex is 1. We can also reduce the max-clique problem to GKP.

We give examples that show that algorithms proposed for Classical Knapsack problem and the Heaviest subgraph problem behave poorly when applied to GKP. We also give a polynomial-time greedy algorithm which has a guaranteed approximation ratio of  $O(\min\{n, W\}n/t)$  for any fixed  $t$ . These are the best known approximation ratios for the QKP. We give a FPTAS (Fully Polynomial time approximation scheme) when the underlying graph has bounded Tree-width. We can also generalize this result to the Hypergraph version of the GKP, where the dependencies between items are not just pairwise but  $k$ -wise for any  $k \geq 2$ .

**[16] A New Methodology For Transportation Investment Decisions Considering Project Interdependencies**, (with S. Kapoor, Z. Li, and E. Veliou, B. Zhou, C. Lee), under revision for publication.

We introduce a new model and algorithms for optimal transportation investment decision-making to support sustainable transportation. The objective is to decide which projects to invest in and implement, so as to maximize the total utility of the transportation network after project implementation viewed from economic (total cost to the transportation agency and the user), social (traffic mobility and safety), and environmental (energy consumption and vehicle emissions) dimensions. In the past, decisions were based on calculating the benefit of each project independently, by measuring its local impacts. However, this ignores two important factors. Firstly, local changes in a transportation network can lead to agglomerative changes in its global behavior. An addition or expansion of a single road segment can lead to better (or sometimes worse) traffic conditions elsewhere even far away from it. Secondly, multiple projects within

a certain geographical area or a major corridor of the transportation network may be proposed for implementation simultaneously, which means that such projects cannot be considered independent of each other. For the first time, both these factors are taken into account in our new models.

The benefit of a collection of projects is defined in terms of an appropriate multi-commodity flow problem defined over the transportation network under study with different non-linear cost functions. These benefits satisfy the global dependence properties described above. They are used to calculate the data needed for Graph Knapsack Problem, see [15], which is then solved to give an appropriate solution for the highway-investment decision-making. We have done a computational study (with help of graduate students in Transportation Engineering: E. Veliou, B. Zhou, C. Lee) of the road network in the Chicago downtown loop area using latest real-life traffic data from I-DOT (Illinois Department of Transportation) to analyze the decision-making for multiple projects under consideration there. Our results clearly show that our models capture the effects of dependency between different projects. Ignoring these effects gives a false inflated benefit from a collection of projects leading to erroneous decision-making: the network-wide benefits with project interdependency considerations tend to be lower than the corresponding benefits without interdependency considerations by 38 to 64 percent, and the network-wide benefits with project interdependency considerations begin to flatten out when the annualized budget reaches approximately \$7.5M with no additional benefits generated from travel time savings with higher levels of investment budgets.

**[17] A Randomized Approximation Algorithm for the Quadratic Knapsack Problem**, (with S. Kapoor), submitted for publication.

This paper continues the research from [15]. We give a randomized algorithm based on an innovative use of non-linear Hyperbolic programs with a non-linear rounding scheme, and analyze it with the help of the Kim-Vu polynomial concentration bounds to improve the approximation ratio to  $O(w_{max}n^{1/2})$ , where  $w_{max}$  is the maximum weight on the vertices. This is again the best known approximation bound for the QKP. To our knowledge, this is the first non-trivial application of Second-Order Cone Programming and nonlinear concentration bounds in approximation algorithms.

**[18] Lectures on Modern Probabilistic Methods for Discrete Mathematics**, (with D.B. West), unpublished manuscript.

These lecture notes based on courses offered by me at UIUC and IIT are being prepared for possible publication. These notes cover three areas and their applications that are not well covered in typical graduate courses in Discrete Mathematics: Concentration of Measure, Applications of Shannon Entropy, and Markov Chain Monte Carlo. The lectures present elementary proofs of basic results in these areas with the aim of making these topics accessible to uninitiated students and researchers in CS, ECE, OR, etc., in addition to those in mathematics. The focus is on developing the themes underlying the various methods and illustrating the final results through applications in graph theory, combinatorial optimization, and theoretical computer science. Most of the discussions and results would appear for the first time in a textbook.

**[19] Exact Algorithms for the Maximum Independent Set Problem**, (with E.C. Sewell and S.H. Jacobson), unpublished manuscript.

NP-hard optimization problems cannot be solved in polynomial time unless  $P = NP$ . To overcome this difficulty, a number of approaches including approximation algorithms and heuristics have been explored. But neither of these approaches guarantee exact solutions. In contrast to these approaches, super-polynomial time algorithms that solve NP-hard problems to optimality have been designed and analyzed. These exact algorithms lead to practical algorithms for moderate instance sizes, in addition to being used as sub-routines in hybrid algorithms with heuristics.

The maximum independent set problem (or, equivalently the maximum clique problem or the minimum vertex cover problem) is a fundamental NP-hard graph optimization problem. The running times of its exact algorithms has improved from  $O^*(1.2599^n)$  (Tarjan–Trojanowski, 1977) to  $O^*(1.1844^n)$  (Robson, 2001). In this paper, a new and faster algorithm, based partly on reductions from the previous paper [10] of the authors, is proposed.

**[20] A Threshold for Random Geometric Graphs with a Hamiltonian cycle**, (with J. Balogh), unpublished manuscript.

Naturally evolving massive networks like the social networks of acquaintances (used to model the spread of diseases, rumors, etc.), the internet, the power grid, airline traffic, etc., share common characteristics like short average distance and high clustering coefficient (average density of subgraphs induced by neighborhoods). *Random Geometric Graph* (RGG) has been studied as a model for such networks. It is constructed by placing  $n$  points randomly (according to uniform distribution or a poisson process) within a metric space, and putting an edge between any pair of points at most distance  $r = r(n)$  apart. As in classical random graphs, the fundamental questions concern the threshold functions for various graph structures. We study RGG on a bounded torus with  $l_\infty$ -metric and show that the threshold for 2-connectedness is about the same as the threshold for Hamiltonicity for these graphs, i.e., there exists a Hamiltonian cycle as soon as the graph becomes 2-connected.

**[21] Multi-Objective Optimization and Sensitivity Analysis For Discrete and Continuous Optimization Problems**, (with S.H. Jacobson, G.K. Kao, J.A. Stori and V. Venkat), Technical Report to Austral Engineering and Software, Inc., University of Illinois, Urbana, 2004.

This technical report focusses on providing tools for solving and analyzing various aspects of optimization problems with multiple objective functions. Multi-objective optimization problems frequently arise in real-life applications, but have not been studied in great detail in literature. My contributions to this report include surveying the previous research and proposing algorithms for sensitivity and post-optimality analysis of discrete multi-objective optimization problems as well as for solving mixed discrete-continuous multi-objective optimization problems.

**[22] On Queuenumber of Planar Graphs**, M.Sc. Thesis (under the guidance of S. Pemmaraju), Dept. of Mathematics, Indian Institute of Technology, Bombay, 1999.

The queuenumber of a graph is the minimum number of queues (FIFO) needed to process its edges (without nesting) over a total order of its vertices. Queue layouts of graphs are dual to stack (LIFO) layouts (better known as book embeddings and the corresponding pagenumber of a graph). Queue layouts have been applied to sorting permutations, parallel process scheduling, matrix computations, and graph drawing. This thesis studies a long standing open problem in queue layouts of graphs - whether queue number of planar graphs is bounded or not. The queue number of stellation of  $K_3$  is explored through a number of general labelings, using tools like Erdos-Szekeres theorem on long monotone subsequences in arbitrary sequences to analyze and show that these families of labelings give unbounded queue number. (The queuenumber of planar graphs remains unknown to date, 2011.)

**[23] An Elementary Construction of  $\mathbb{R}$  from  $\mathbb{Q}$** , *The Mathematica : Annual Journal of the Mathematics Society*, St. Stephen's College, Delhi, 1997.

This paper proposes an elementary construction of  $\mathbb{R}$  as a completion of  $\mathbb{Q}$ . The method taught in high school for approximating the value of an irrational number by upper and lower bounding sequences of rational numbers (essentially the decimal representation) is an illustration of the nested intervals property of reals. Any family of nested intervals of rationals with diameter tending to zero can be said

to define a real number. Just as rational Cauchy sequences need not converge in  $\mathbb{Q}$ , nested intervals in rationals need not have a non-empty intersection in  $\mathbb{Q}$ . This leads to  $\mathbb{R}$  being defined as the collection of families of nested intervals in  $\mathbb{Q}$  with diameter tending to zero. This construction is shown to follow the underlying theme of the Cantor's construction using Cauchy sequences and Dedekind's construction using Dedekind's cuts, and can be used as a motivation for these constructions.