

Research Statement (brief summary of research activity in past 20 years)

My research career started around 1990 in India when I worked on a few problems in computational geometry. After arriving at Pennsylvania State University as a graduate student, I continued working further in *geometric problems* (with P. Berman), on *graph theory* (M. Fürer and others) and on *lower bounds for continuous function circuits* (with G. Schnitger and others). Two notable problems that I worked on during this time are the *rectilinear polygon cover* problems and on *computational separation between sigmoidal vs. threshold circuits*. In 1992 I moved to University of Minnesota to finish my doctoral studies under D.-Z. Du. During this time, I worked on topics such as *learning-theoretic aspects of continuous valued circuits* (with E. Sontag from Rutgers and H. Siegelmann) and on *routing problems in hypercubes and multi-stage networks* (with S. Bhattacharya, F. Hwang and Y.-C. Yao). After finishing my PhD, I had a brief post-doctoral position in DIMACS with E. Sontag where we worked out optimal *sample complexity bounds for learning recurrent perceptron models*. My second post-doctoral position with M. Li (University of Waterloo) and T. Jiang (then at McMaster University) shifted my focus to computational biology. Our work on *nearest neighbor interchange distance for phylogenetic trees* settled a computational complexity issue for this problem that was open for a very long time, and our work on the *syntenic distance between multi-chromosome genomes* (also with S. Kannan and E. Sweedyk) provided the first approximation for this problem.

From 1996 until 2001 I was at Rutgers University Camden Campus working first as a visiting assistant professor and later as a tenure-track assistant professor. In 2001 I joined UIC as an assistant professor, where currently I am a professor of computer science with courtesy appointment in bioengineering. During these years I have worked on problems related to applications in several areas such as *computational biology*, *computational economics and finance*, *social networks* and *hybrid systems*, as well as *purely theoretical computational complexity questions*. I have also been instrumental in developing several softwares that are heavily used by the relevant biologists. I was awarded the NSF CAREER award in 2004 in recognition of my research activities.

Here are some highlights of my research activities in the past 20 years or so:

- I am in the “proof-correction” stage of publishing a *graduate level textbook* with J. Liang entitled “*Models and Algorithms for Biomolecules and Molecular Networks*” to be published by John Wiley & Sons very soon. The *first part* of the book will cover surface and volume representation of the structures of biomolecules based on the idealized ball model. The underlying geometric constructs as well as their computation will then be discussed. This will be followed by the chapter on constructing effective scoring functions in different functional forms using either the statistical approach or the optimization approach, with the goal of identifying native-like protein structures or protein-protein interfaces, as well as constructing a general fitness landscape for protein design. The topic of sampling and estimation that can be used to generate biomolecular structures and to estimate their evolutionary patterns are then discussed, with equal emphasis on the Metropolis Monte Carlo (or Markov Chain Monte Carlo) approach and the chain growth (or sequential Monte Carlo) approach. This is followed by a chapter covering the topic of stochastic networks formed by interacting biomolecules and the framework of discrete chemical master equations, as well as computational methods for direct numerical computation and for sampling reaction trajectories of the probabilistic landscape of these networks. The *second part* of the book will cover interaction networks of biomolecules. We will discuss stochastic models for networks with small copy numbers of molecular species, as those arising in genetic circuits,

protein synthesis, and transcription binding, and algorithms of computing the properties of stochastic molecular networks. We will then cover signal transduction networks that arise, for example, in complex interactions between the numerous constituents such as DNAs, RNAs, proteins and small molecules in a complex biochemical system such as a cell. We will also discuss the experimental protocols and algorithmic methodologies necessary to synthesize these networks. Of special interest will be the synthesis these networks from double-causal experimental evidences and methods for reverse engineering of such networks based on suitable experimental protocols.

- My research (with P. Berman) on *rectilinear polygon cover* problems provided the first known APX-hardness result for a geometric problem (as stated by D. Hochbaum).
- My research work (with G. Schnitger) on the *separation of computational power between discrete threshold circuits and continuous-valued analog circuits* improved upon the previous separation bounds and is the best known separation bound as of yet.
- My research work (with X. He, T. Jiang, M. Li, J. Tromp and L. Zhang) on the *nearest-neighbor-interchange distance* answered a question on the computational complexity of this problem that was open for about 25 years.
- My research work (with P. Berman) on off-line *throughput maximization for job scheduling* on one or more machines improved the approximation bounds for these problems previously known. The 2-approximation for the single machine case is to the best of our knowledge still the best possible approximation ratio known.
- My research work (with E. Sontag) on checking *equivalence under certain semiring congruence for the state-space isomorphism problem for hybrid systems* solved an open problem that was raised about 13 years back.
- My series of research works (with P. Berman, S. Muthukrishnan and S. Ramaswami) on *several tiling and packing problems involving rectangles and hyper-rectangles* improved the computational complexity bounds of previously best known results for these problems. Some of the approximation bounds are still the best possible. In a separate work (with P. Berman, P. Bertone, M. Gerstein, M.-Y. Kao and M. Snyder) we worked on the *genome tiling* problem and provided an efficient algorithm.
- My research work (with P. Berman and S. Muthukrishnan) on *binary space partitioning bounds for rectangles* provided an optimal upper bounds (the corresponding matching lower bounds were obtained by A. Dumitrescu and J. Mitchell).
- My research work (with K. Akcoglu, J. Aspnes and M.-Y. Kao) on *opportunity cost algorithms for combinatorial auctions* provided an efficient algorithm for approximating a revenue-maximizing set of bids an auctioneer should accept when bids are restricted to be connected node subsets of an underlying object graph. The graph-theoretic components in the analysis were further developed about 10 years later by A. Borodin and others in studying specific classes of graphs called elimination graphs.
- My research work (with P. Berman and M.-Y. Kao) on *test set and string barcoding* problems provided provably optimal approximability results for these problems.
- My research work (with P. Berman, D. Mubayi, R. Sloan, G. Turàn and Y. Zhang) on *inverse protein folding problem* provided optimal algorithms for this problem under the canonical model in 2D and 3D.
- My research work (with P. Berman and E. Sontag) on *set multicover* problems provided improved randomized approximation and showed how this problem can be used in modular response analysis in reverse engineering of biological regulatory networks.
- My research work (with G. A. Enciso, E. Sontag and Y. Zhang) provided an efficient algorithm via semi-definite programming methods to *decompose a complex biological system into monotone sub-systems*.

- My series of research works (with M. Ashley, T. Y. Berger-Wolf, P. Berman, I. Caballero, W. Chaovalitwongse, C.-A. Chou, P. Govindan, A. A. Khokhar, Z. Liang and S. Sheikh) provided a comprehensive in-depth investigation (both theoretical and experimental) of the *kinship reconstruction problem from microsatellite data in population genetics*. Our research also led to investigation of computational complexity aspects of triangle packing and an “implicit” set cover problem, and culminated in the development of the KINALYZER software that has been subsequently used by other population biologists.
- My research work (with A. Binkowski, J. Dundas and J. Liang) provided an efficient computational method of *topology independent protein structural alignment* in structural bioinformatics.
- My series of research works (with R. Albert, R. Dondi, S. Kachalo, E. Sontag, A. Zelikovsky, K. Westbrook and R. Zhang) developed and thoroughly (both theoretically and experimentally) investigated a novel method of *synthesizing and analyzing biological signal transduction networks from double-causal experimental data*. This research also culminated in the development of the NET-SYNTHESIS software that has been subsequently used by other biologists.
- My research work (with P. Berman and M. Karpinski) on the *transitive reduction problems for directed graphs* provided 1.5-approximation algorithms for the minimization version and 2-approximation algorithms for the maximization versions of these problems using appropriate primal-dual polytopes.
- My research work (with J. Jun and I. Mandoiu) developed efficient algorithms for *primer selection methods for detection of genomic inversions and deletions* in biological sequences.
- My research work (with A. Bhattacharya, D. Mubayi and G. Turán) provided the best-known (at that time) inapproximability results for *minimizing a special class of Boolean formula called Horn formula*.
- My research work (with M. Comi, M. Schapira and V. Srinivasan) investigated *two- and multi-party approximately privacy preserving communication protocols* (under the geometric privacy model) for tiling functions.
- My series of research works (with D. Ayala, J. Lin, O. Wolfson and B. Xu) comprehensively investigate *game-theoretic frameworks for parking slot assignments and navigational guidance for parking*. The research also culminated in submission of a US patent.
- My research work (with R. Albert, A. Gitter, G. Gürsoy, R. Hegde, P. Pal, G. S. Sivanathan and E. Sontag) provided a new computationally efficient measure of *topological redundancy for biological and social networks*.
- My research work (with D. Desai) provided *approximation algorithms and hardness results for modularity clustering* which are currently the best known results. We also investigated a connection between modularity clustering and the small set expansion problem.
- My research work (with S. Muthukrishnan) on *stochastic budget optimization in internet advertising* provided the first known non-trivial poly-logarithmic approximation for these problems as well as the first known hardness results of getting better than logarithmic approximation ratios in the various parameters involved.
- My series of research works (with P. Berman, L. Kaligounder and M. Karpinski) on the *vulnerability of financial networks for scenarios such as the over-the-counter derivatives market* provided the first known computational complexity results for computing the stability. We also performed a comprehensive empirical evaluation of this stability measure to derive topological properties and parameter combinations that may be used to flag the network as a possible fragile network.
- My research work (with R. Albert and N. Mobasher) provided interesting *topological implications of Gromov-hyperbolicity for biological and social networks*.

- My research work (with B. Atassi, B. P. Beirne, E. C. Dragut, W. Meng, A. Neyestani and C. Yu) provided efficient algorithms for *merging query results from local search engines for geo-referenced objects*.
- My not-yet-published research work (with A. D. M. Gunawan and L. Zhang) answered an open problem in the study of phylogenetic networks by providing a polynomial-time solution for the so-called *tree containment problem for stable phylogenetic networks*.
- My not-yet-published research work (with M. Karpinski, N. Mobasher and F. Yahyanejad) provided constructive non-trivial bounds on *node expansions and cut-sizes for hyperbolic graphs*, and provided algorithmic consequences of these bounds and their related proof techniques for a few problems related to cuts and paths for hyperbolic graphs.