Discrete Algorithms Seminar

List of Presentations

Fall 2007

August 29, 2007

Title: "Approximating the Spanning Star Forest Problem and Its Applications to Genomic Sequence Alignment."

Authors: C. T. Nguyen, Jian Shen, Minmei Hou, Li Sheng, and Webb Miller.

Journal or Conference: SODA 2007.

Abstract: This paper studies the algorithmic issues of the spanning star forest problem. We prove the following results: (1) there is a polynomial-time approximation scheme for planar unweighted graphs; (2) there is a polynomial-time algorithm with approximation ratio 3/5 for unweighted graphs; (3) it is NP-hard to approximate the problem within ratio $545/546 + e$ for unweighted graphs; (4) there is a linear-time algorithm to compute the maximum star forest of a weighted tree; (5) there is a polynomial-time algorithm with approximation ratio 1/2 for weighted graphs. We also show how to apply this spanning star forest model to aligning multiple genomic sequences over a tandem duplication region.

Speaker: Andres Varon.

September 05, 2007

Title: "Clearing Algorithms for Barter Exchange Markets: Enabling Nationwide Kidney Exchanges."

Authors: David Abraham, Avrim Blum, and Tuomas Sandholm.

Journal or Conference: EC 2007.

Abstract: In barter-exchange markets, agents seek to swap their items with one another, in order to improve their own utilities. These swaps consist of cycles of agents, with each agent receiving the item of the next agent in the cycle. We focus mainly on the upcoming national kidney-exchange market, where patients with kidney disease can obtain compatible donors by swapping their own willing but incompatible donors. With over 70,000 patients already waiting for a cadaver kidney in the US, this market is seen as the only ethical way to significantly reduce the 4,000 deaths per year attributed to kidney disease. The clearing problem involves finding a social welfare maximizing exchange when the maximum length of a cycle is fixed. Long cycles are forbidden, since, for incentive reasons, all transplants in a cycle must be performed simultaneously. Also, in barter-exchanges generally, more agents are affected if one drops out of a longer cycle. We prove that the clearing problem with this cycle-length constraint is NP-hard. Solving it exactly is one of the main challenges in establishing a national kidney exchange. We present the first algorithm capable of clearing these markets on a nationwide scale. The key is incremental problem formulation. We adapt two paradigms for the task: constraint generation and column generation. For each, we develop techniques that dramatically improve both runtime and memory usage. We conclude that column generation scales drastically better than constraint generation. Our algorithm also supports several generalizations, as demanded by real-world kidney exchanges. Our algorithm replaced CPLEX as the clearing algorithm of the Alliance for Paired Donation, one of the leading kidney exchanges. The match runs are conducted every two weeks and transplants based on our optimizations have already been conducted.

Description: From a blog that comments on this paper:
http://geomblog.blogspot.com/2007/08/saving-lives-with-exact-algorithms.html

Saving lives with exact algorithms:

It's not often you get to say this in a paper:

> We aim for exact algorithms [because] ... any loss of optimality could lead to unnecessary patient deaths.

Anyone who's gone through an algorithms class will at some point hear about stable marriage algorithms, and how the method is used to match hospitals and newly minted MDs looking for residencies.

Consider now the far more serious problem of matching kidney transplant candidates to potential donors. Because transplant lists are long, and cadaver donors are few,

marketplaces matching healthy donors to recipients have sprung up in the US. For complicated ethical reasons (which are not without controversy), such exchanges are not made for money, and are viewed as gifts.

So what happens if a donor kidney doesn't match the potential recipient? Ordinarily, nothing. Suppose however that there was another donor-recipient pair with a similar mismatch, and if only the donors were swapped, both transplants could go through? What about if a 3-way cycle of matchings could be found? This is called a 'market clearing', and is the subject of a paper by Abraham, Blum and Sandholm to appear in EC.

Link Paper:

`http://www.cs.cmu.edu/~sandholm/kidneyExchange.EC07.withGrantInfo.pdf`

Speaker: Matthew P. Johnson.

September 19, 2007

Title: "Cheap Labor Can Be Expensive."

Authors: Ning Chen and Anna R. Karlin.

Journal or Conference: SODA 2007.

Abstract: We study markets in which consumers are trying to hire a team of agents to perform a complex task. Each agent in the market prices their labor, and based on these prices, consumers hire the cheapest available team capable of doing the job they need done. We define the cheap labor cost in such a market as the ratio of the best Nash equilibrium of the original market and the best possible Nash equilibrium of any of its submarkets, where "bes" is defined with respect to consumers, i.e., we are looking at Nash equilibria in which the consumer pays the least. This definition is motivated by a "Braess-style" paradox: in certain kinds of marketplaces, competition, in the form of the availability of "cheap labor", can actually cause the prices paid by consumers to go up. We present tight bounds on the cheap labor cost for a variety of markets including s-t path markets, matroid markets and perfect bipartite matching markets. The differences in cheap labor cost across markets demonstrate the complex relationship between the combinatorial structure of the marketplace and the advantages or more precisely, disadvantages to consumers due to competition.

Link Paper: `http://www.cs.washington.edu/homes/ning/paper/cheap-labor.pdf`

Speaker: Simon Shamoun.

September 26, 2007

Title: "The Ferry Cover Problem."

Authors: Michael Lampis and Valia Mitsou.

Journal or Conference: FUN 2007: Fourth International Conference on FUN WITH ALGORITHMS; (http://www.di.unipi.it/fun07/).

Abstract: In the classical wolf-goat-cabbage puzzle, a ferry boat man must ferry three items across a river using a boat that has room for only one, without leaving two incompatible items on the same bank alone. In this paper we define and study a family of optimization problems called Ferry problems, which may be viewed as generalizations of this familiar puzzle.

In all Ferry problems we are given a set of items and a graph with edges connecting items that must not be left together unattended. We present the Ferry Cover problem (FC), where the objective is to determine the minimum required boat size and demonstrate a close connection with Vertex Cover which leads to hardness and approximation results. We also completely solve the problem on trees. Then we focus on a variation of the same problem with the added constraint that only 1 round-trip is allowed (FC1). We present a reduction from MAX-NAE-3-SAT which shows that this problem is NP-hard and APX-hard. We also provide an approximation algorithm for trees with a factor asymptotically equal to 4/3 . Finally, we generalize the above problem to define FC_m, where at most m round-trips are allowed, and MFT_k, which is the problem of minimizing the number of round-trips when the boat capacity is k. We present some preliminary lemmata for both, which provide bounds on the value of the optimal solution, and relate them to FC.

Speaker: Valia Mitsou.

October 3, 2007

Title: "Digraph Measures: Kelly Decompositions, Games, and Orderings."

Authors: Paul Hunter and Stephan Kreutzer.

Journal or Conference: SODA 2007.

Abstract: We consider various well-known, equivalent complexity measures for graphs such as elimination orderings, k-trees and cops and robber games and study their natural translations to digraphs. We show that on digraphs all these measures are also equivalent and induce a natural connectivity measure.We introduce a decomposition for digraphs and an associated width, Kelly-width, which is equivalent to the aforementioned measure. We demonstrate its usefulness by exhibiting a number of potential applications including polynomial-time algorithms for NP-complete problems on graphs of bounded Kelly-width, and complexity analysis of asymmetric matrix factorization. Finally, we compare the new width to other known decompositions of digraphs.

Link Paper: `http://www.cl.cam.ac.uk/~pwh21/stuff/kellywidth_full.pdf`

Speaker: Michael Lampis.

October 10, 2007

Title: "Learning with Attribute Costs."

Authors: Haim Kaplan, Eyal Kushilevitz, and Yishay Mansour.

Journal or Conference: STOC 2005.

Abstract: We study an extension of the "standard" learning models to settings where observing the value of an attribute has an associated cost (which might be different for different attributes). Our model assumes that the correct classification is given by some target function f from a class of functions F; most of our results discuss the ability to learn a clause (an OR function of a subset of the variables) in various settings:

Offline: We are given both the function f and the distribution D that is used to generate an input x. The goal is to design a strategy to decide what attribute of x to observe next so as to minimize the expected evaluation cost of f(x). (In this setting there is no "learning" to be done but only an optimization problem to be solved; this problem turns out to be NP-hard and hence approximation algorithms are presented.)

Distributional online: We study two types of "learning" problems; one where the target function f is known to the learner but the distribution D is unknown (and the goal is to minimize the expected cost including the cost that stems from "learning" D), and the other where f is unknown (except that f F) but D is known (and the goal is to minimize the expected cost while limiting the prediction error involved in "learning" f).

Adversarial online: We are given f, however the inputs are selected by an adversary. The goal is to compare the learners cost to that of the best fixed evaluation order (i.e., we analyze the learners performance by a competitive analysis).

Link Paper: `http://www.math.tau.ac.il/~mansour/papers/05stoc.pdf`

Speaker: Yi Feng.

October 17, 2007

Title: "Improved Recommendation systems."

Authors: Baruch Awerbuch, Boaz Patt-Shamir, David Peleg, and Mark Tuttle.

Journal or Conference: SODA 2005.

Abstract: We consider a model of competitive recommendation systems proposed by Drineas et al.[4]. In recommendation systems (e.g., for books or movies), the system tracks which product each user chose in the past, and tries to deduce which other products an asking user is likely to be satisfied with. Obviously, recommendation systems can be effective only for users who share preferences with many other users. Such users are said to belong to a "dominant type." Current approaches to on-line recommendation systems involve using Singular Value Decomposition (SVD), which is computationally intensive and, more important, often applicable only under additional strong conditions. Specifically, correctness is guaranteed in [4] only if users of different dominant types essentially do not share a product they like "type separability"), and only if the number of users in non-dominant types is significantly smaller than the number of users in dominant types ("gap assumption"). The complexity of that algorithm is $O(mn)$, where m and n denote the number of users and products, respectively. In this paper, we show that in fact, very simple combinatorial algorithms can make good recommendations without using SVD. Our algorithms require neither the type separability nor the gap assumption, they are naturally amenable to distributed computation, and their complexity is lower. In particular, the paper presents an $O(m+n)$ time centralized algorithm and a distributed algorithm that can be implemented in a peer-to-peer model even in the presence of adaptively colluding malicious players, with only logarithmic over-head.

Link Paper: `www.markrtuttle.com/papers/appt05-soda.ps`

Speaker: Yosef Alayev.

October 24, 2007

Title: "The Hiring Problem and Lake Wobegon Strategies."

Authors: Andrei Z. Broder, Adam Kirschy, Ravi Kumarz, Michael Mitzenmacherx, Eli Upfal, and Sergei Vassilvitskiik.

Journal or Conference: SODA 2008.

Abstract: We introduce the hiring problem, in which a growing company continuously interviews and decides whether to hire applicants. This problem is similar in spirit but quite different from the well-studied secretary problem. Like the secretary problem, it captures fundamental aspects of decision making under uncertainty and has many possible applications. We analyze natural strategies of hiring above the current average, considering both the mean and the median averages; we call these Lake Wobegon strategies. Like the hiring problem itself, our strategies are intuitive, simple to describe, and amenable to mathematically and economically significant modifications. We demonstrate several intriguing behaviors of the two strategies. Specifically, we show dramatic differences between hiring above the mean and above the median. We also show that both strategies are intrinsically connected to the lognormal distribution, leading to only very weak concentration results, and the marked importance of the first few hires on the overall outcome.

Link Paper: `http://www.eecs.harvard.edu/~kirsch/pubs/hiring/hiring-submit.pdf`

Speaker: Lijun Feng.

October 31, 2007

Title: "Better Online Buffer Management."

Authors: Fei Li, Jay Sethuraman, and Clifford Stein.

Journal or Conference: SODA 2007.

Abstract: As the Internet becomes more mature, there is a realization that improving the performance of routers has the potential to substantially improve Internet performance in general. Currently, most routers forward packets in a First-In-First-Out (FIFO) order. However, the diversity of applications supported by modern IP-based networks has resulted in un-predictable packet flows, and heterogeneous network traffic. Thus, it is becoming more reasonable to consider differentiating between different types of packets, and perhaps to consider allowing packets to specify a deadline by which it must be processed. These issues have made buffer management at routers a critical issue in providing effective quality of service to the various applications that use the network. In this paper, we study an online problem in which each packet is described by its discrete arrival time, non-negative weight and discrete deadline; arriving packets are buffered for delivery and all packets have the same processing time. The packets arrive online, and our objective is to maximize the sum of weights of those packets that are sent by their deadlines. We describe an online deterministic algorithm with a competitive ratio of 1.854, improving the best previous known competitive ratio of 1.939 (Bartal et al. STACS 2004). The algorithmic framework we use has several interesting features. First, we do not use a potential function. Instead, after each step we modify the adversary's buffer. Second, we introduce "dummy packets" to facilitate the decision making.

Link Paper: `http://www1.cs.columbia.edu/~lifei/papers/SODA07.pdf`

Speaker: Omer Demir.

November 07, 2007

Title I: "On the Competitiveness of On-Line Scheduling of Unit-Length Packets with Hard Deadlines in Slotted Time".

Authors I: B. Hajek.

Journal or Conference I:

Abstract I: It is shown that the competitive factor for online scheduling of unit-length packets with hard deadlines in slotted time is in the interval $[0.5, \phi]$, where ϕ is the inverse of the golden ratio, $\phi = (\sqrt{5}+1)/2 = 0.618034$. Moreover, any static priority policy, that in each slot schedules a maximum value packet irrespective of deadlines, achieves competitive factor 0.5.

Link Paper I: `http://www.ifp.uiuc.edu/~hajek/Papers/Hajek01_OL.pdf`

Title II: "An optimal online algorithm for packet scheduling with agreeable deadlines."

Authors II: Fei Li, Jay Sethuraman and Clifford Stein.

Journal or Conference II: SODA 2005.

Abstract II: An important issue in IP-based QoS networks is the effective management of packets at the router level. Specifically, if the arriving packets cannot all be stored in a buffer, or if the packets have deadlines by which they must be delivered, the router needs to identify the packets that should be dropped. Suppose each packet has a delivery deadline to meet and an associated weight. The goal of a packet scheduling algorithm is to drop and deliver the arriving packets so that the combined weight of the delivered message can be maximized. The paper designs and analyzes a simple deterministic packet scheduling algorithm, with which a competitive ratio $\phi = (\sqrt{5}+1)/2$ can be achieved. The authors claim that the competitive ratio is at the time the best one achieved by a deterministic algorithm.

Link Paper II: `http://www1.cs.columbia.edu/~lifei/papers/SODA05.pdf`

Link Slides II: `http://www1.cs.columbia.edu/~lifei/papers/SODA05_talk.pdf`

Speaker: Jiang Wu.

November 14, 2007

Title: "Online Primal-Dual Algorithms for Covering and Packing Problems."

Authors: Niv Buchbinder and Seffi Naor.

Journal or Conference: ESA 2005.

Abstract: We study a wide range of online covering and packing optimization problems. In an online covering problem a linear cost function is known in advance, but the linear constraints that define the feasible solution space are given one by one in an online fashion. In an online packing problem the profit function as well as the exact packing constraints are not fully known in advance. In each round additional information about the profit function and the constraints is revealed. We provide general deterministic primal-dual schemes for online fractional covering and packing problems. We also provide deterministic algorithms for several integral online covering and packing problems. Our scheme is designed via a novel primal-dual technique that extends the scheme used for many offline optimization problems. Recently, it was shown that this general primal-dual framework is useful for capturing many more online problems. We list several of the later results that follow from this approach.

Link Paper: `http://www.cs.technion.ac.il/~nivb/`

(one of the papers in the list).

Other Papers: "The Online Set Cover Problem," (STOC 03); "Online Make-to-Order Joint Replenishment Model: Primal Dual Competitive Algorithms," (SODA 08); By Buchbinder and Naor.

Speaker: Matthew P. Johnson.

November 28, 2007

Title: "Online scheduling of equal length jobs on parallel machines."

Authors: Jihuan Ding, Tomays Ebenlendr, Jiri Sgall, and Guochuan Zhang

Journal or Conference: ESA 2007.

Abstract: We study online scheduling of equal length jobs on parallel machines. Our main result is an algorithm with competitive ratio decreasing to $e/(e-1) = 1.58$ as the number of machine increases. For $m \geq 3$, this is the first algorithm better than 2-competitive greedy algorithm. Our algorithm has an additional property called immediate decision: at each time, it is immediately decided for each newly released job if it will be scheduled, and if so, then also the time interval and machine where it is scheduled is fixed and cannot be changed later. We show that for two machines, no deterministic algorithm with immediate decision is better than 1.8-competitive; this lower bound shows that our algorithm is optimal for $m = 2$ in this restricted model. We give some additional lower bounds for algorithms with immediate decision.

Link Paper: `http://www.math.cas.cz/~sgall/ps/eqpar.ps`

Speaker: Simon Shamoun.

December 05, 2007

Title: "To Fill or not to Fill: The Gas Station Problem."

Authors: Authors: Samir Khuller, Azarakhsh Malekian, and Julian Mestre.

Journal or Conference: ESA 2007.

Abstract: In this paper we study several routing problems that generalize shortest paths and the Traveling Salesman Problem. We consider a more general model that incorporates the actual cost in terms of gas prices. We have a vehicle with a given tank capacity. We assume that at each vertex gas may be purchased at a certain price. The objective is to find the cheapest route to go from s to t, or the cheapest tour visiting a given set of locations. Surprisingly, the problem of finding the cheapest way to go from s to t can be solved in polynomial time and is not NP-complete. For most other versions however, the problem is NP-complete and we develop polynomial time approximation algorithms for these versions.

Link Paper: `http://www.cs.umd.edu/projects/gas/`

Speakers: Michael Lampis and Valia Mitsou.

December 12, 2007

Title: "Non-Preemptive Min-Sum Scheduling with Resource Augmentation."

Authors: Nikhil Bansal, Ho-Leung Chan, Rohit Khandekar, Kirk Pruhs, Baruch Schieber, and Cliff Stein.

Journal or Conference: FOCS 2007.

Abstract: We give the first $O(1)$-speed $O(1)$-approximation polynomial-time algorithms for several nonpreemptive minsum scheduling problems where jobs arrive over time and must be processed on one machine. More precisely, we give the first $O(1)$-speed $O(1)$-approximations for the nonpreemptive scheduling problems $1\,|r_j|\sum w_jF_j$ (weighted flow time), $1\,|r_j|\sum T_j$ (total tardiness), the broadcast version of $1\,|r_j|\sum w_jF_j$, an $O(1)$-speed, 1-approximation for $1\,|r_j|\sum \bar{U}_j$ (throughput maximization), and an $O(1)$-machine, $O(1)$-speed $O(1)$-approximation for $1\,|r_j|\sum w_jT_j$ (weighted tardiness). Our main contribution is an integer programming formulation whose relaxation is sufficiently close to the integer optimum, and which can be transformed to a schedule on a faster machine.

Link Paper:

`http://csdl.computer.org/dl/proceedings/focs/2007/3010/00/30100614.pdf`

Speaker: Ou Liu.