Discrete Algorithms Seminar

List of Presentations

Spring 2007

February 07, 2007

Title: "Aggregation of Partial Rankings, p-Ratings, and Top-m Lists"

Authors: Nir Ailon

Journal or Conference: SODA 2007: 415–424

Abstract: We study the problem of aggregating partial rankings. This problem is motivated by applications such as meta-searching and information retrieval, search engine spam fighting, e-commerce, learning from experts, analysis of population preference sampling, committee decision making and more. We improve recent constant factor approximation algorithms for aggregation of full rankings and generalize them to partial rankings. Our algorithms improved constant factor approximation with respect to all metrics discussed in Fagin et al's recent important work on comparing partial rankings. We pay special attention to two important types of partial rankings: the well-known top-m lists and the more general p-ratings which we define. We provide first evidence for hardness of aggregating them for constant m; p.

Link Paper: `http://www.math.ias.edu/~nailon/ranking_with_ties.pdf`

Speaker: Matthew P. Johnson

February 14, 2007

Title: "Fast Periodic Graph Exploration with Constant Memory"

Authors: Leszek Gasieniec, Ralf Klasing, Russell A. Martin, Alfredo Navarra, and Xiaohui Zhang

Journal or Conference: SIROCCO 2007: 26–40

Abstract: We consider the problem of periodic exploration of all nodes in undirected graphs by using a finite state automaton called later a robot. The robot, using a constant number of states (memory bits), must be able to explore any unknown anonymous graph. The nodes in the graph are neither labelled nor colored. However, while visiting a node v the robot can distinguish between edges incident to it. The edges are ordered and labelled by consecutive integers $1, \ldots, d(v)$ called port numbers, where $d(v)$ is the degree of v. Periodic graph exploration requires that the automaton has to visit every node infinitely many times in a periodic manner. Note that the problem is unsolvable if the local port numbers are set arbitrarily, see [8]. In this context, we are looking for the minimum function $\pi(n)$, such that, there exists an efficient deterministic algorithm for setting the local port numbers allowing the robot to explore all graphs of size n along a traversal route with the period $p(n)$. Dobrev et al. proved in [13] that for oblivious robots $\pi(n) \leq 10n$. Recently Ilcinkas proposed another port labelling algorithm for robots equipped with two extra memory bits, see [20], where the exploration period $\pi(n) \leq 4n - 2$. In the same paper, it is conjectured that the bound $4n - O(1)$ is tight even if the use of larger memory is allowed. In this paper, we disprove this conjecture presenting an efficient deterministic algorithm arranging the port numbers, such that, the robot equipped with a constant number of bits is able to complete the traversal period in $\pi(n) \leq 3.75n - 2$ steps hence decreasing the existing upper bound. This reduces the gap with the lower bound of $\pi(n) \geq 2n - 2$ holding for any robot.

Related papers:

- "Finding shortest right-hand-on-the-wall walks in graphs," SIROCCO 2005;
- "Setting port numbers for fast graph exploration," SIROCCO 2006;
- "Graph exploration by a finite automaton," TCS 345 (2005) 331-344;
- "Tree exploration with logarithmic memory," SODA 2007.
- "Tree exploration with little memory," J. Alg. 51 (2004) 38-63.

Speaker: Panagiotis Cheilaris (Panos Hilaris)

February 28, 2007

Title: "A Network Formation Game for Bipartite Exchange Economies"

Authors: Eyal Even-Dar, Michael J. Kearns, and Siddharth Suri

Journal or Conference: SODA 2007: 697–706

Abstract:

We introduce a natural new network formation game in which buyers and sellers may purchase edges representing trading opportunities between themselves, and then accrue wealth in the resulting exchange economy. Our main result is an exact characterization of the set of bipartite graphs G that are Nash equilibria for this game. This characterization provides sharp limits on the amount and structure of wealth variation that can occur, as well as on the allowable equilibrium exchange rates.

Link Paper: `http://www.cis.upenn.edu/~mkearns/papers/econform.pdf`

Related papers:

- A. Fabrikant, A. Luthra, E. Maneva, C.H. Papadimitriou, and S. Shenker. "On a network creation game." In Proc. 22nd Annual Symp. on Principles of Distributed Computing, 2003 http://www.cs.berkeley.edu/ alexf/papers/flmps03.pdf;
- S. M. Kakade, M. Kearns, L. E. Ortiz, R. Pemantle, and S. Suri. "Economic properties of social networks." In Lawrence K. Saul, Yair Weiss, and Leon Bottou, editors, Advances in Neural Information Processing Systems 17, Cambridge, MA, 2005. MIT Press;
- S. M. Kakade, M. Kearns, and L. E. Ortiz. "Graphical economics." In Proceedings of the 17th Annual Conference on Learning Theory (COLT), 2004;

Speaker: Joanna Klukowska

March 07, 2007

Title I: "Customized Newspaper Broadcast: Data Broadcast with Dependencies"

Authors I: Julien Robert and Nicolas Schabanel

Journal or Conference I: LATIN 2006: 362–373

Abstract I: Broadcasting has been proved to be an efficient means of disseminating data in wireless communication environments (such as Satellite, mobile phone networks; other typical broadcast networks are Videotext systems). Recent works provide strong evidence that correlationbased broadcast can significantly improve the average service time of broadcast systems. Most of the research on data broadcasting was done under the assumption that user requests are for a single item at a time and are independent of each other. However in many real world applications, such as web servers, dependencies exist among the data items, for instance: web pages on a server usually share a lot of items such as logos, style sheets, title-bar... and all these components have to be downloaded together when any individual page is requested. Such web server could take advantage of the correlations between the components of the pages, to speed up the broadcast of popular web pages. This paper presents a theoretical analysis of data dependencies and provides a polynomial time 4-approximation as well as theoretical proofs that our correlationbased approach can improve by an arbitrary factor the performances of the system. To our knowledge, our solutions are the first provably efficient algorithms to deal with dependencies involving more than two data items.

Link Paper I: `http://www.dim.uchile.cl/~schabanel/publications/2006/2006-DS-LATIN2006.pdf`

Title II: "Pull-Based Data Broadcast with Dependencies: Be Fair to Users, not to Items"

Authors II: Julien Robert and Nicolas Schabanel

Journal or Conference II: SODA 2007: 238–247

Abstract II: Broadcasting is known to be an efficient means of disseminating data in wireless communication environments (such as Satellite, mobile phone networks,...). It has been recently observed that the average service time of broadcast systems can be considerably improved by taking into consideration existing correlations between requests. We study a pull-based data broadcast system where users request possibly overlapping sets of items; a request is served when all its requested items are downloaded. We aim at minimizing the average user perceived latency, i.e. the average flow time of the requests. We first show that any algorithm that ignores the dependencies can yield arbitrary bad performances with respect to the optimum even if it is given arbitrary extra resources. We then design a $(4+\epsilon)$-speed $O(1+1/\epsilon^2)$-competitive algorithm for this setting that consists in 1) splitting evenly the bandwidth among each requested set and in 2) broadcasting arbitrarily the items still missing in each set into the bandwidth the set has received. Our algorithm presents several interesting features: it is simple to

implement, non-clairvoyant, fair to users so that no user may starve for a long period of time, and guarantees good performances in presence of correlations between user requests (without any change in the broadcast protocol). We also present a $(4+\epsilon)$-speed $O(1+1/\epsilon^3)$-competitive algorithm which broadcasts at most one item at any given time and preempts each item broadcast at most once on average. As a side result of our analysis, we design a competitive algorithm for a particular setting of non-clairvoyant job scheduling with dependencies, which might be of independent interest.

Link Paper II: `http://www.dim.uchile.cl/~schabanel/publications/2007/2007-SODA-RS.pdf`

Speaker: Yosef Alayev

March 14, 2007 & May 02, 2007

Title: "Speed Scaling for Weighted Flow Time"

Authors: Nikhil Bansal, Kirk Pruhs, and Clifford Stein

Journal or Conference: SODA 2007: 805–813

Abstract: Intel's SpeedStep and AMD's PowerNOW technologies allow the Win XP to dynamically change the speed of the processor to prolong batter life. In this setting, the OS must not only have a job selection policy to determine which job to run, but also a speed scaling policy to determine the speed at which the job will be run. We give the online speed scaling algorithm that is O(1)-competitive for the objective of weighted flow time plus energy. This algorithm also allows us to efficiently construct an O(1)-approximate schedule for minimizing weighted flow time subject to an energy constraint.

Link Paper: `http://domino.research.ibm.com/comm/research_people.nsf/pages/nikhil.pubs.html/$FILE/speedflow.ps`

Speaker: Ou Liu

March 21, 2007

Title: "Whole Genome Duplications, Multi- Break Rearrangements, and the Genome Halving Problem"

Authors: Max A. Alekseyev, Pavel A. Pevzne

Journal or Conference: SODA 2007: 665–679

Abstract: The Genome Halving Problem, motivated by the whole genome duplication events in molecular evolution, was solved by El-Mabrouko and Sankoff. The El-Mabrouko Sankoff algorithm is rather complex inspiring a quest for a simpler solution. An alternative approach to Genome Halving Problem based on the notion of the contracted breakpoint graph was recently proposed in [2]. This new technique reveals that while the El-Mabrouko Sankoff result is correct in most cases, it does not hold in the case of unichromosomal genomes. This raises a problem of correcting El-Mabrouko Sankoff analysis and devising an algorithm that deals adequately with all genomes. In this pa- per we efficiently classify all genomes into two classes and show that while the El-Mabrouko Sankoff theorem holds for the first class, it is incorrect for the second class. The crux of our analysis is a new combinatorial invariant de- fined on duplicated permutations. Using this invariant we were able to come up with a full proof of the Genome Halving theorem and a polynomial algorithm for Genome Halving Problem (for unichromosomal genomes). We also give the first short proof of the original El-Mabrouko Sankoff result for multichromosomal genomes. Finally, we discuss a generalization of Genome Halving Problem for a more general set of rearrangement operations (including transpositions) and propose an efficient algorithm for solving this problem.

Link Paper: `http://www- cse.ucsd.edu/~maxal/ap_soda07.pdf`

Speaker: Andres Varon

March 28, 2007

Title: "Paging Mobile Users Efficiently and Optimally"

Authors: Amotz Bar-Noy, Yi Feng, and Mordecai J. Golin

Journal or Conference: INFOCOM 2007: 1910–1918

Abstract: A mobile user is roaming in a zone composed of N cells in a cellular network system. When a call to the mobile user arrives, the system pages the mobile user in these cells since it never reports its location unless it leaves the zone. The N cells are associated with a probability vector $(p_1, ..., p_N)$ where p_i is the probability that the mobile user resides in the ith cell and all the probabilities are independent. A delay constraint paging strategy must find the mobile user within D $(1 <= D <= N)$ paging rounds; in each round a subset of the N cells is paged. The goal is to minimize the expected number of paged cells until the mobile user is found. Solutions based on dynamic programming that yield optimal strategies are known. The running time of the known implementations is $\Theta(N^2D)$. Our first contribution is to improve the running time to Theta(ND) by proving that the dynamic programming recursive formulation satisfies the Monge property, permitting us to use various dynamic programming speedup techniques. A $\Theta(N)$ heuristic solution is also known. Our second contribution is a heuristic whose running time is $\Theta(NlogD)$. Our heuristic outperforms the known heuristic while running faster for $D << N$. We compare the non-optimal heuristics with the optimal solution demonstrating the tradeoff between optimality and running time efficiency of various solutions.

Speaker: Yi Feng

April 11, 2007

Title: "A Lower Bound for Scheduling Mechanisms"

Authors: George Christodoulou, Elias Koutsoupias, and Angelina Vidali

Journal or Conference: SODA 2007: 1163–1170

Abstract: We study the mechanism design problem of scheduling tasks on n unrelated machines in which the machines are the players of the mechanism. The problem was proposed and studied in the seminal paper of Nisan and Ronen, where it was shown that the approximation ratio of mechanisms is between 2 and n. We improve the lower bound to $1 + \sqrt{2}$ for 3 or more machines.

Link Paper: `http://portal.acm.org/citation.cfm?id=1283508`

Speaker: Panagiotis Cheilaris (Panos Hilaris)

April 25, 2007

Title: "A 1.875-Approximation Algorithm for the Stable Marriage Problem"

Authors: Kazuo Iwama, Shuichi Miyazaki, and Naoya Yamauchi

Journal or Conference: SODA 2007: 288–297

Abstract: We consider the problem of finding a stable matching of maximum size when both ties and unacceptable partners are allowed in preference lists. This problem is known to be APX-hard, and the current best known approximation algorithm achieves the approximation ratio $2 - \frac{c}{\sqrt{N}}$, where c is some positive constant. In this paper, we give a 1.875 approximation algorithm, which is the first result on the approximation ratio better than two.

Link Paper: `www.lab2.kuis.kyoto-u.ac.jp/~iwama/papers/isaac2006-3.ppt`

Some Slides: `http://www.lab2.kuis.kyoto-u.ac.jp/~iwama/papers/soda2007-2.pdf`

Speaker: Simon Shamoun

May 16, 2007

Title: "Approximation Algorithms and LP-Duality"

Abstract:

Many NP-Complete problems can be solved approximately by solving a linear program optimally and rounding to an integer an solution. This is poly-time, but relatively expensive. Instead, some problems can be solved with the techniques of Primal-Dual or Dual Fitting, which both involve generating primal and dual solutions that satisfy certain conditions without actually solving the LP. Recently, Jain & Vazirani introduced Factor-Revealing LPs, for use in the analysis of approximation algorithms. The optimal value of such a (maximization) LP yields an upper bound on the approximation factor of the corresponding algorithm. (This LP is solved as part of the *analysis* of the algorithm, but not within the algorithm itself.) Finally, Vazirani et al. introduced Tradeoff-Revealing LPs, which helped them *derive* their algorithm for the Adwords problem. We will survey these techniques and their use in a linked series of papers, on Facility Location, inventory problems, online matching, and the Adwords problem.

References:

- "Approximation algorithms for metric facility location and k-Median problems using the primal-dual schema and Lagrangian relaxation," Kamal Jain and Vijay V. Vazirani, FOCS'99 and JACM'01.
- "Greedy facility location algorithms analyzed using dual fitting with factor-revealing L," Kamal Jain, Mohammad Mahdian, Evangelos Markakis, Amin Saberi, and Vijay V. Vazirani), STOC'02 and JACM'03.
- "Primal-dual algorithms for deterministic inventory problem," Retsef Levi, Robin Roundy, and David B. Shmoys, STOC'04.
- "AdWords and Generalized On-line Matching." Aranyak Mehta, Amin Saberi, Umesh V. Vazirani, and Vijay V. Vazirani, FOCS'05.

Speaker: Matthew P. Johnson