dr hab. Marcin Mucha
Instytut Informatyki,
Uniwersytet Warszawski

Warszawa, 30 sierpnia 2022

Review of the PhD dissertation of mgr Artur Kraska

The price of delays in on-line algorithms

In his PhD dissertation, mgr Artur Kraska presents results concerning three different problems in the area of on-line algorithms. The common theme for all three is the ability of the algorithm to delay serving any request arbitrarily, which incurs a certain cost. Other than that the problems are not really connected in any way. The presented results have already been published (as conference proceedings at ICALP, SIROCCO, 2xWAOA). However, the dissertation is not merely a collection of three research papers. It does contain a well-written introduction that presents the main themes and emphasizes the common feature of the three parts, as well as an afterword.

Evaluation of the contributions

Section 2 presents the results concerning the Travelling Repairman Problem (TRP), Dial-A-Ride Problem (DARP) and the Unrelated Machines Scheduling (UMS) problem, and more generally the class of Resettable Scheduling (RS) problems, which contains all of them. The key feature of the RS class is the ability to simulate any (partial) solution with an arbitrary delay. The obvious way to exploit this feature in an on-line algorithm is to choose a sequence of time points, and then for each time t in the sequence compute the optimum off-line solution for the prefix $[0, t]$, and then simulate it with a delay of t. The actual algorithms proposed in the dissertation are just slightly more complicated (e.g. instead of just using the optimum for $[0, t]$, an optimum for a certain proxy cost is used, to facilitate analysis). This idea is, of course, not new. The main contribution, however, lies in the analysis, which leads to improved competitive ratios for (variants of) all three problems mentioned above. The author proposes a very clever model, that one can view as a sort of discretization of the algorithmic idea above, and an LP that describes this model, and whose maximum is an upper bound on the competitive ratio. This is a so-called factor revealing LP, an approach used before, but never for this problem family. The said maximum is then upper-bounded by presenting a low-cost solution to the dual LP. This part of the dissertation is by far the most interesting. Every step of the way, from choosing the right discretization/model to solving the dual LP and clever modeling and analyzing the randomized variant of the algorithm, is highly non-trivial are surely required significant insight and a lot of work. I was particularly impressed by the fact that the author managed to give an explicit general solution to the dual LP, which at first sight seems to be a nearly impossible task, given the complexity of the LP in question. Typically, one would use an automated solver to produce numerical solutions for small-sized duals, and then argue that the values do not blow up significantly with the LP size.

Section 3 presents results for the Online Service with Delay (OSD) problem. In OSD, the algorithm is presented with requests in a metric space and processes them with arbitrary delays, balancing the movement cost and delay cost. An $O(\log n)$-competitive algorithm is presented for the case of equidistant points on a line. The algorithm and its analysis seem to follow the ideas of the $O(\log n)$-competitive algorithm for the Reordering Buffer Management problem on the line, due to Gamzu and Segev. I think this is the most routine part of the dissertation. The authors clearly understood the GS algorithm and analysis well and adapted the key parts to the different setting.

MM

Section 4 presents results for the Min-cost (Bipartite) Perfect Matching with Delays problem (M(B)PMD). Here, the algorithm is given a sequence of $2m$ points in a metric space and needs to match them, possibly with a delay. The algorithm needs to balance the (metric) matching cost and the cost of delays. In the bipartite version, the points are signed and only points of opposite signs can be matched. The current best-known bounds for randomized algorithms are fairly tight. In the case of deterministic algorithms, however, the gap is rather large. In particular, the best competitive ratio known used to be $O(m^{2.46})$. The authors improve this to $O(m)$ (and this has subsequently been improved to $O(m^{0.58})$). The overall approach is primal-dual, and more specifically its moat-growing interpretation. However, while adapting the traditional Goemans-Williamson framework to approximation algorithms is often a relatively straightforward task, using it for online algorithms is rather tricky and typically requires making several careful adjustments in a coordinated way. This case is no different. Based on my own experience with moat-growing algorithms I believe this is a very solid contribution and making all the pieces of the algorithm fit together must have required significant insight and a lot of work, especially since no previous attempts were made to apply the moat-growing approach to this particular problem.

Evaluation of the write-up

The dissertation is very well planned out, written, and presented. Multiple figures simplify the understanding of the more subtle points. The dissertation also contains an extensive and well-selected bibliography.

Conclusions

I believe this is a strong PhD thesis. It tackles natural and difficult problems and makes significant contributions to each of them. Obtaining these contributions required both a very good understanding of several technical tools in the area of on-line algorithms, as well as a lot of ingenuity - this is especially true in the case of Section 2 concerning Resettable Scheduling problems. Moreover, the results have already been published at good conferences, and the dissertation itself is very well written and edited. With this in mind, I believe that this thesis meets the requirements for a PhD dissertation.

....
dr hab. Marcin Mucha