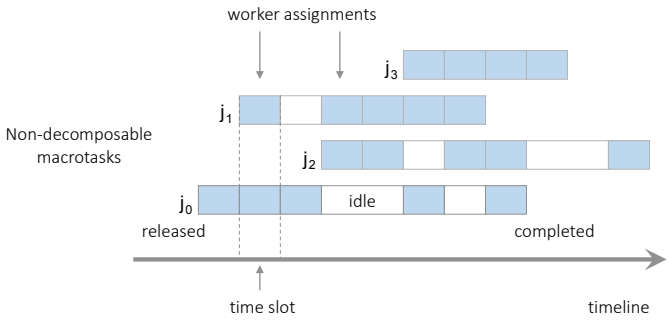
Online Sequencing of Non-Decomposable Macrotasks in Expert Crowdsourcing

Presentation of work originally published in ACM Transactions on Social Computing, Volume 1, Issue 1, Article No. 1, February 2018.

Heinz Schmitz[1], Ioanna Lykourentzou[2]

Keywords: crowdsourcing optimization; cooperative social computing; online scheduling decisions; macrotask scheduling

We introduce the problem of Task Assignment and Sequencing (TAS), which models online optimization in expert crowdsourcing settings that involve non-decomposable macrotasks. Non-decomposition is a property of certain types of complex problems, like the formulation of an R&D approach or the definition of a research methodology, which cannot be handled through the *divide and conquer* approach typically used in microtask crowdsourcing. In contrast to splitting the macrotask to multiple microtasks and allocating them to several workers in parallel, our model supports the sequential improvement of the macrotask one worker at a time, across distinct time slots of a given timeline, until a sufficient quality level is achieved. An online environment is assumed where expert workers are available only at specific time slots and worker/task arrivals are not known a priori.



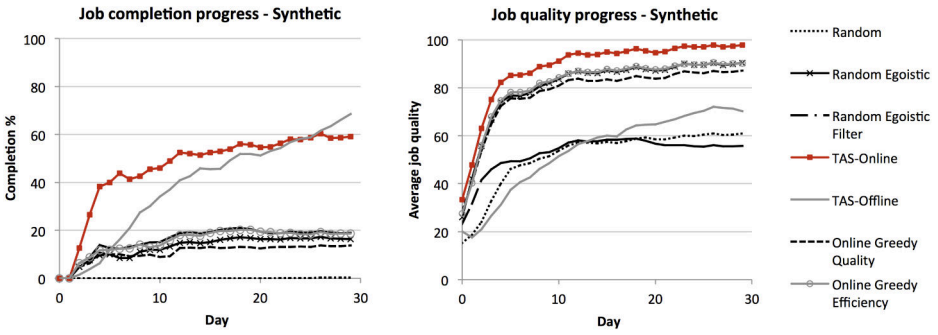
This new crowd work optimization model makes a conceptual shift from task decomposition and worker parallelization to task non-decomposition and worker sequencing. It combines

[1] Hochschule Trier, Schneidershof, 54293 Trier, Germany, h.schmitz@hochschule-trier.de
[2] Utrecht University, Princetonplein 5, 3584 CC Utrecht, The Netherlands, i.lykourentzou@uu.nl

doi:10.18420/inf2019_11

assignment decisions (per time slot and across tasks) with sequencing decisions (rolling out these assignments along a timeline) under reasonable constraints. We capture this setting with a mathematical formulation for TAS and prove its strong NP hardness.

With respect to this setting, we propose TAS-ONLINE, an online algorithm that aims to complete as many tasks as possible within budget, required quality and a given timeline, without any future input information regarding job release dates or worker availabilities. We illustrate, through synthetic and real-world experiments, that TAS-ONLINE can achieve more completed jobs, lower flow times and higher quality compared to five typical benchmarks.



Our small-scale real-world experiment on CrowdFlower[3] was intended to provide an initial and qualitative viewpoint of the model's performance in a real-world setting. The task we used was collaborative news article writing, where workers from an initial hiring pool were asked to build on each other's content sequentially, enriching a news article text on a given topic. At the end of the scheduling period, the competing benchmark algorithm achieved a successful completion of 3 out of 6 jobs, while our optimization algorithm successful completed 5 jobs. As it was expected the benchmark algorithm either allowed workers of the minimum necessary expertise to take a job, thus delaying the job's quality progress too much, or it starved the job of budget. On the other hand TAS-ONLINE selected workers in such a way as to improve job completion within the given time period.

Our results have implications for enhancing the Quality of Service of crowdsourcing platforms offering non-decomposable complex tasks, but also for allowing online expert crowdsourcing communities to make better use of their human capital and available expertise. Multiple future extensions can be foreseen. These include extending the proposed TAS model to handle requirements such as budget flexibility, the non-acceptance of assignments by the workers, different job quality aggregation mechanism, learning, varying modes of assignment order and number of assignments per worker, as well as forecasting.

[3] http://www.crowdflower.com