

**COST Action IC1205 on Computational Social Choice: STSM Report
COST-STSM-IC1205-28859**

Applicant: Mathijs de Weerd

Home institution: Delft University of Technology

Home country: The Netherlands

Host: Enrico Gerding

Host institution: University of Southampton

Host country: United Kingdom

Dates: 26/10/2015 to 30/10/2015

The goal of my STSM visit was to work with Enrico Gerding on *incremental aggregation of incomplete rankings*. We started with work on the (static problem of) aggregation of incomplete and noisy rankings: we formalised the objective (of minimising the expected error), defined both an optimal and a local search procedure, and performed many more experiments, both with synthetic and with real data. This work will be presented with a poster at the WINE Conference on Web and Internet Economics in Amsterdam (December 2015), and we expect to submit the full paper later this year.

Next we studied literature on incremental voting. We had several brainstorm sessions with Sebastian Stein (and Enrico and me of course) and one with Long Tran-Thanh. We identified a strong link with crowdsourcing as well as online learning: for ranking alternatives or selecting the best alternative (e.g. of a design, a paper, a solution to a problem, etc.) knowledge workers may be used, and an important decision is whether to ask another (costly) knowledge worker to make a better decision, or decide based on the information collected thus far.

State of the art approaches focus on using scores assigned to alternatives, but scores are typically biased, depending on the worker. Instead, we will contribute by considering ranks. There seems to be a gap between machine learning approaches which mostly focus on online binary classification problems, supported by experiments, and social choice theory with some work on incremental rank aggregation, but very little with experimental support. We aim to fill this gap by experimenting with known voting rules in realistic online settings.

We identified and formalised the initial problem we thus aim to solve: given anonymous crowd workers with independent noise (but from the same distribution), what is the expected improvement in the aggregate ranking of asking another worker to rank two given alternatives. This can then be used to decide on which pair to ask to be ranked, and this can be done repeatedly until the expected improvement does not outweigh the cost of another worker, or until the maximum number of workers is reached. This can be an important contribution to this very active and current issue in crowdsourcing. We will continue our collaboration on this remotely.