Critique on Max Algorithms in Crowdsourcing Environments

Summary:
This paper focuses on the problem of finding max by crowdsourcing. The author developed a parameterized family of max algorithms. It investigates some heuristic and optimization strategies and evaluates them in experiments. They propose a way to optimize the performance of the max algorithm by combining factors like the execution time, the cost, and the quality of the results.

Comments:

Pros:
1. Clear explaining of the assumption, parameters and the logic flow of their proposed max algorithm.
2. Good abstraction and formulation of the problem in the problem definition section. It makes further discussion has a solid basis.

Cons:
1. In the experiment section, it need more explanation of the experiment settings. Or the experiments will be not repeatable.
2. This paper assumes each human have a specific probability providing the right answer and the distribution of that probability is available. But in practice, that distribution is not available.

Recommendations:
1. I would put the related work section after the Introduction. So readers not familiar with this area can get an idea about what the problem is and the current status of research in this area.

Critique on Human-powered Sorts and Joins

Summary:
The authors proposed a crowd sourcing system named Qurk (which has interface to Amazon crowdsourcing platform MTurk). This system uses crowds to compare items for
completing sorting and joining data tasks. With proper optimizations, Qurk can reduce the cost of join task from $70 to $3. In addition, it shows that batching in proper size can greatly reduce the complexity of the sorting task. In the step of assignments response combination, Majority voting method and Quality adjustment method are evaluated and compared in the experiments. Finally, the authors claimed that in the sorting task, hybrid methods of rating based method and comparison based method can reduce the HITs (Human Intelligence Tasks) from 1000 to 77.

Comments:

Pros:
1. This paper examined the sorting and joining task by crowdsourcing’s accuracy, latency and cost, which showed the value and potential of their crowdsourcing system by its convincing results.
2. The task to identify the celebrities is a good choice. Because it’s a typical problem easy for human but difficult for computers. Which implies that crowdsourcing is a great way to solve such kind of problems.
3. Sorting and Joining are basic but fundamental, the authors solid work on them open the field of crowdsourcing and will attract more and more researchers into this field.

Cons:
1. It could cost a lot time waiting for people finishing the assignments. So this method has it limitation that it can only be used to those query that don't need instant result.
2. For different specific tasks, the best batch size, the best batch size may varies, And for different people, the batch size may varies as well. So the smart batching idea relies highly on the specific sorting/joining task and different participants.

Recommendations:
1. As one of biggest problems of crowdsourcing is the result may be not reliable, I would do some analysis of the accuracy of the responses of their query on such database and try to feedback the user an estimated error rate.
2. I would do a profile of the proposed algorithm by a plot of time it take to design and accomplish such a crowdsourcing sorting or joining versus varies sizes of datasets