Infusing Human Factors into Algorithmic Crowdsourcing

Han Yu[1], **Chunyan Miao**[1], **Zhiqi Shen**[1], **Jun Lin**[1,2], **Cyril Leung**[1,3] and **Qiang Yang**[4]
[1]Joint NTU-UBC Research Centre of Excellence in Active Living for the Elderly, Nanyang Technological University, Singapore
[2]College of Software Engineering, Beihang University, Beijing, China
[3]Department of Electrical and Computer Engineering, the University of British Columbia, Vancouver, Canada
[4]Department of Computer Science and Engineering, Hong Kong University of Science and Technology, Hong Kong
{han.yu, ascymiao, zqshen, junlin}@ntu.edu.sg, cleung@ece.ubc.ca, qyang@cse.ust.hk

Background

Crowdsourcing offers a new way for mobilizing the collective intelligence and efforts of a large group of people to tackle tasks which cannot be efficiently performed by machines (e.g., video transcription, monitoring road conditions at specific locations) (Doan, Ramakrishnan, and Halevy 2011). Many commercial crowdsourcing platforms are now available. Their business depends on providing satisfactory services to both the *crowdsourcers* (i.e. *task requesters*) and the *workers* involved. From the crowdsourcers' perspective, they expect to receive high quality results for their crowdsourcing tasks in a timely manner. From the workers' perspective, they want to earn as much as possible while committing limited productive effort.

As the crowdsourcers are the main source of revenue for crowdsourcing platforms, their requirements tend to take precedence over those of the workers. Many commercial crowdsourcing platforms have implemented some variants of reputation-based mechanisms (Yu et al. 2013a) to gauge the workers' competence based on their past performance, and allow only those with good reputations to access tasks.

While this simple reputation-based task delegation method is intuitive and has its own merits, a different, albeit related problem, has not received much attention. As crowdsourcing workers are human beings, they have limited availability and productive capacities to work on tasks delegated to them (Yu et al. 2012). Concentrating requests to workers with good reputations may result in details. In addition, as a small portion of reputable workers become overloaded with task requests while others remain relatively idle, attrition may occur in the worker population, thereby leaving the crowdsourcing platform with a shortage of workers.

The potentially conflicting objectives between crowdsourcers and workers can be formalized under the *congestion game* framework (Monderer and Shapley 1996). In congestion games, the payoff for each player depends on the resources it selects and the number of other players selecting the same resources. For instance, the morning commute to work places by many people can be modeled as a congestion game. The time taken by a traveller on a given day depends on how many others are taking the same route as him.

Challenges

Algorithmic Crowdsourcing (AC) is an emerging field in which computational methods are proposed to automate certain aspects of crowdsourcing. A number of AC methods have proposed recently in an attempt to address this problem. In (Yu et al. 2014a; 2013b), decision support methods based on heuristics and queuing theory have been proposed for reputation aware task delegation in crowdsourcing systems where workers are not assumed to be able to sub-delegate tasks to other workers. In (Heidari and Kearns 2013; Nath and Narayanaswamy 2014; Yu et al. 2015b), graph theory, game theory, and queuing theory based decision support methods have been proposed to allow workers to sub-delegate tasks to other workers. In order to help crowdsourcing systems efficiently harness human resources, AC technologies will play an important role.

However, existing AC approaches are based on highly simplified models of worker behaviour which limit their practical applicability. To make efficient utilization of human resources for crowdsourcing tasks, the following technical challenges remain open:

- *Fairness of the solution*: As workers in a crowdsourcing system may come from various backgrounds, their trustworthiness, competence in performing certain tasks, and committed productive effort may differ. Task allocation strategies need to make workers feel that they have been treated fairly given their perceived performance and the available task requests. Alternative metrics for measuring the fairness of task allocation plans can be tested using the provided dataset to identify those most likely to be accepted by workers.
- *Temporal changes in behaviour*: Another challenge for crowdsourcing task allocation strategies is that the performance of workers may change over time. The changes can be either positive (workers' skills might improve) or negative (workers might become careless), and the period of such changes can also vary.
- *Optimizing wellbeing*: Workers may take elements (such as mood, work/life balance, wellbeing and altruism) which improve their happiness into account when gauging the value they can derive from crowdsourcing tasks. These qualities need to be infused into workers' objective functions when designing task allocation strategies.

- *Non-compliance by users*: As crowdsourcers and workers are humans, they may not always make the most rational decisions. Such behaviours will introduce additional complexity and uncertainty into the problem. Models for human behaviour/choices and the corresponding task allocation strategies need to be examined.

For AI researchers to propose effective solutions to these challenges, labelled datasets reflecting various aspects of human decision-making related to task allocation in crowdsourcing are needed. There are a number of crowdsourcing datasets currently available to AI researchers.However, these datasets contain only the results from crowdsourcing. The process of finding suitable workers and determining how tasks are allocated has not been captured.

Decisions Dataset

To bridge this gap, we construct an anonymized dataset[1] based on player behavior trajectories captured by a multi-agent game platform - *Agile Manager* (AM) (Yu et al. 2014b). It allows players to demonstrate their task delegation strategies under different scenarios based on key characteristics involved in crowdsourcing task allocation (Yu et al. 2015a). The game adopts implicit human computation (Quinn and Bederson 2011) in which players contribute data which are valuable for research through informal games.

The game is presented as a virtual environment for players to understand the challenges facing a software engineering team manager who needs to efficiently delegate tasks to team members with diverse skills and productivity. It is designed for university level software engineering courses teaching agile software development (ASD) methodologies. The Scrum-based ASD methodology (Lin et al. 2014) is used to construct the game play as it is very similar to the task delegation problem in crowdsourcing. A player manages a virtual ASD team consisting of ten programmers each controlled by a programmer agent (PA). Each level of the game consists of multiple Sprints of development activities. Each task is characterized by its value, difficulty, required effort, and deadline. The player must delegate many tasks to PAs in each Sprint with the objective of matching tasks to suitable PAs and ensure they can be completed on time.

The AM game platform has been used by a university in Singapore and a university Beijing as a coursework tool for undergraduate software engineering students. Over 450 people have played it. They completed 3,439 game sessions. People played the AM game in two different modes: 1) *familiar PA mode*: in which the skill level and productivity of each PA in a player's team does not change across different game sessions; and 2) *unfamiliar PA mode*: in which the skill level and productivity of each PA in a player's team changes across different game sessions. 1,159 game sessions were played in the familiar PA mode, and 2,280 game sessions in the unfamiliar PA mode.

Close to 200,000 player task allocation decisions are included in the dataset. Each record contains a comprehensive snapshot of the situation related to a PA (including its reputation, ground truth regarding its competence and productivity, workload, the list of tasks in its backlog, time taken for the player to make each decision, and the player's current mood). The decision data forms a time series from which the variations in each player's task allocation strategies and the consequence of each decision can be inferred.

With this dataset, efforts by researchers from many fields of AI can now be supported to design methods for efficiently harnessing human resource through crowdsourcing.

[1] `http://www.agelesslily.org/demo_agilemanager/`

Acknowledgements

This research is supported, in part, by the National Research Foundation, Prime Minister's Office, Singapore under its IDM Futures Funding Initiative and administered by the Interactive and Digital Media Programme Office; and the Lee Kuan Yew Post-Doctoral Fellowship Grant.

References

Doan, A.; Ramakrishnan, R.; and Halevy, A. Y. 2011. Crowdsourcing systems on the world-wide web. *Communications of the ACM* 54(4):86–96.

Heidari, H., and Kearns, M. 2013. Depth-workload tradeoffs for workforce organization. In *Proc. HCOMP*, 60–68.

Lin, J.; Yu, H.; Shen, Z.; and Miao, C. 2014. Studying task allocation decisions of novice agile teams with data from agile project management tools. In *Proc. ASE*, 689–694.

Monderer, D., and Shapley, L. S. 1996. Potential games. *Games and Economic Behavior* 14:124–143.

Nath, S., and Narayanaswamy, B. 2014. Productive output in hierarchical crowdsourcing. In *Proc. AAMAS*, 469–476.

Quinn, A. J., and Bederson, B. B. 2011. Human computation: A survey and taxonomy of a growing field. In *Proc. CHI*, 1403–1412.

Yu, H.; Shen, Z.; Miao, C.; and An, B. 2012. Challenges and opportunities for trust management in crowdsourcing. In *Proc. IAT*, 486–493.

Yu, H.; Shen, Z.; Leung, C.; Miao, C.; and Lesser, V. R. 2013a. A survey of multi-agent trust management systems. *IEEE Access* 1(1):35–50.

Yu, H.; Shen, Z.; Miao, C.; and An, B. 2013b. A reputation-aware decision-making approach for improving the efficiency of crowdsourcing systems. In *Proc. AAMAS*, 1315–1316.

Yu, H.; Miao, C.; An, B.; Shen, Z.; and Leung, C. 2014a. Reputation-aware task allocation for human trustees. In *Proc. AAMAS*, 357–364.

Yu, H.; Yu, X.; Lim, S. F.; Lin, J.; Shen, Z.; and Miao, C. 2014b. A multi-agent game for studying human decision-making. In *Proc. AAMAS*, 1661–1662.

Yu, H.; Lin, H.; Lim, S. F.; Lin, J.; Shen, Z.; and Miao, C. 2015a. Empirical analysis of reputation aware task delegation by people from a multi-agent game. In *Proc. AAMAS*, 1687–1688.

Yu, H.; Miao, C.; Shen, Z.; and Leung, C. 2015b. Efficient task sub-delegation for resource constrained trustee agents in multi-agent trust networks. In *Proc. AAAI*, 1305–1311.