A Reputation-aware Decision-making Approach for Improving the Efficiency of Crowdsourcing Systems (Extended Abstract)

Han Yu, Zhiqi Shen, Chunyan Miao
School of Computer Engineering
Nanyang Technological University
Singapore 639798
{yuha0008, zqshen, ascymiao}@ntu.edu.sg

Bo An
Key Laboratory of Intelligent Information Processing
Institute of Computing Technology
Chinese Academy of Sciences, Beijing 100190, China
boan@ict.ac.cn

ABSTRACT
A crowdsourcing system is a useful platform for utilizing the intelligence and skills of the mass. Nevertheless, like any open system that involves the exchange of things of value, selfish and malicious behaviors exist in crowdsourcing systems and need to be mitigated. Trust management has been proven to be a viable solution in many systems. However, a major difference between crowdsourcing systems and existing trust models designed for multi-agent systems is that human trustees have limited task processing capacity per unit time compared to an intelligent agent program. This paper recognizes a problem in current trust-aware decision-making methods for task assignment in crowdsourcing platforms. On the one hand, trust-based methods over-assign tasks to trusted workers, while on the other hand, workload-based solutions do not give sufficient guarantees on the quality of work. The proposed solution, the social welfare optimizing reputation-aware decision-making (SWORD) approach, strikes a balance between the two and is shown through extensive simulations to significantly improve social welfare of crowdsourcing platforms compared to related work.

Categories and Subject Descriptors
I.2.11 [**Artificial Intelligence**]: Distributed Artificial Intelligence - *Intelligent Agents, Multiagent Systems*.

General Terms
Algorithms, Human Factor.

Keywords
Trust, reputation, decision-making.

1. INTRODUCTION
Crowdsourcing systems provide an environment where mass collaboration can be harnessed to quickly complete a large number of tasks that are challenging for computers. Crowdsourcing systems usually have a user base consisting of *requesters* who have tasks that need to be completed, and *workers* who are willing to offer their time and effort to complete tasks in exchange for payment. Requesters break down their project into small *human intelligence tasks* (HITs) and organize them as an *HIT group*. They then assign a monetary reward for each HIT. Usually, the payments are the same for HITs in the same HIT group. Workers can browse through the available HITs to work on those for which they are qualified. Similar to any system involving the exchange of things of value, ensuring the quality of HIT results is a challenging problem faced by crowdsourcing systems. Trust management mechanisms have been suggested by recent research [1] to be a viable way to address the problem of malicious workers in crowdsourcing.

Appears in: *Proceedings of the 12th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2013)*, Ito, Jonker, Gini, and Shehory (eds.), May, 6–10, 2013, Saint Paul, Minnesota, USA.

In contrast to the heavy focus on finding new ways to accurately estimate the reputation of trustees, there is a notable shortage of models dictating how the trusters should utilize such evaluations to guide their interaction decisions. The majority of existing approaches advocate that a truster should always engage the most trustworthy trustee it can find for interaction at any point in time [2]. However, in [3], the authors have demonstrated that such a self-interested decision-making approach negatively impacts social welfare in crowdsourcing systems where each worker (i.e., trustee) can only serve a limited number of requests per unit time. Uncoordinated interaction decisions may result in over-utilization of a small number of relatively more trustworthy workers and reduce the amount of business a crowdsourcing system can handle, thus jeopardizing its sustainable operation. In this paper, we propose a social welfare optimizing reputation-aware decision-making (SWORD) approach as a centralized HIT broker for crowdsourcing systems in order to mitigate such unintended consequences of existing trust management models. It provides a methodical way for trust-aware decisions to take the overall wellbeing of a system into account. Extensive simulations have shown that the proposed approach significantly outperforms related work in terms of enhancing the social welfare for an entire crowdsourcing system.

2. THE *SWORD* APPROACH
In a crowdsourcing system where the workers (i.e., trustees) are human beings and have limited capacity to serve incoming HIT requests per unit time, delegating HITs can be modeled as a *congestion game* [4]. In such a game, the payoff of each requester (i.e., truster) partially depends on the number of other requesters delegating HITs to the same workers as it does at the same time. If the HIT result is returned within specified deadline, the worker will be rated based on the quality of the result. Otherwise, the worker will receive a negative rating from the requester regardless of the quality of the result. Thus, SWORD aims to strike a balance between minimizing the requester's risk of receiving low quality HIT results from untrustworthy workers and minimizing the amount of delay experienced by the requesters before receiving the HIT results. Ideally, both of these two sub-objectives should be minimized together. Such an objective can be denoted by a *drift-minus-reward* expression:

$$\Delta(t) - V \times reward(t) \quad (1)$$

where $\Delta(t)$ represents the change in the level of congestion in the pending HIT queues of workers in a crowdsourcing system. Then, $V \in R^{+}$ is the only control parameter specifying the relative importance of the two sub-objectives. We adopt the *Lyapunov drift* [5] expression to measure the change in the level of congestion in the system:

$$\Delta(\boldsymbol{Q}(t)) \triangleq \mathbb{E}\{L(\boldsymbol{Q}(t+1)) - L(\boldsymbol{Q}(t)) | \boldsymbol{Q}(t)\} \quad (2)$$

where $\boldsymbol{Q}(t) = (Q_w(t)$ is the HIT backlog in the pending HIT queues of workers $w \in \boldsymbol{W}$. When workers complete HITs, the level of congestion in the system is pushed towards 0. However, in order to efficiently utilize workers capacities, the drift from the target queue length for each worker should be minimized in order to reduce idling time for the workers when there are enough HIT requests in the system. Thus, the Lyapunov function L is defined as:

$$L(\boldsymbol{Q}(t)) \triangleq \frac{1}{2}\sum_{w:w\in\boldsymbol{W}}(Q_w(t) - \theta_w)^2 \quad (3)$$

where θ_w is the target HIT queue length for worker *w*:

$$\theta_w \triangleq N\mu_w^{max} + V g_{max}\,\tau_w^{max} \quad (4)$$

where τ_w^{max} and μ_w^{max} are the maximum reputation and maximum rate of completing HITs of *w* observed over a given period of time. g_{max} is the maximum utility derived from an HIT being completed on time and with high quality. The values of the parameters *N* and *V* determine the trade-off between the expected waiting time and the expected quality of the HIT results and need to be selected by the crowdsourcing system administrators based on their objectives. Using $A_w(t)$ to denote the number of new HITs assigned to worker *w* at time *t*, the objective function (1) can be re-written as:

$$\sum_{w:w\in\boldsymbol{W}}[V(g_{max}\,\tau_w(t) - c)A_w(t) - (Q_w(t) - \theta_w)A_w(t)] \quad (5)$$

where *c* represents the cost incurred by a requester when delegating an HIT to a worker, and $\tau_w(t)$ is the reputation of worker *w* at time *t*. When new HITs are proposed by requesters, SWORD calculates an HIT allocation plan follows:

1) Re-evaluate the desirability score ($D_w(t)$) for all workers as:

$$D_w(t) = V(g_{max}\,\tau_w(t) - c) - Q_w(t) + \theta_w; \quad (6)$$

2) Rank the workers in descending order of their $D_w(t)$ values;
3) Assign HITs to the top ranked worker *w*, if $D_w(t)$ is positive and $\tau_w(t)$ is higher than the threshold value specified by the requester until his HIT queue length is equal to his target HIT queue length;
4) Repeat Step 3 for other workers following descending order of their desirability ranking;
5) The process ends if all new HITs have been allocated following the method in Step 3. If there are still more HITs unallocated, these HITs are randomly assigned to all workers with probabilities corresponding to their reputation values.

To approach the optimal social welfare in a given crowdsourcing system, average quality of the HIT results received by requesters should be improved. This requires concentrating HIT assignments to highly reputable workers by increasing the value of *V*. However, doing so will reduce the total number of HITs the system can complete per time step which have a negative influence on social welfare. Therefore, increasing the value of *V* can only increase the proximity of the solution to the optimal if the total number of HITs published is equal or less than that the system can complete per time step. Beyond that workload further increase the value of *V* will reduce social welfare.

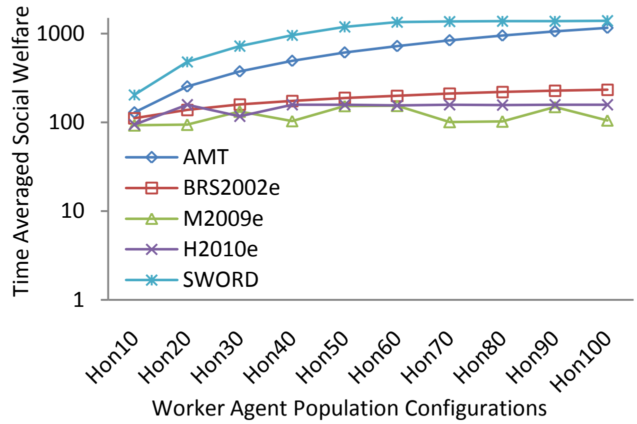

Figure 1. Time averaged social welfare.

To evaluate SWORD, we use the simulation test-bed environment and benchmarking approaches presented in [3]. As shown in Figure 1, the social welfare achieved by SWORD is significantly higher than other HIT allocation approaches under different worker agent population configurations.

3. DISCUSSIONS AND FUTURE WORK

In this paper, we propose a situation-aware trust decision-making approach to help crowdsourcing systems automate the assignment of HITs to workers. It enriches the system model used by existing trust-aware decision-making approaches by modeling the situation as a congestion game. The proposed approach, SWORD, continuously adjusts the distribution of HITs to workers using the difference between the target HIT queue lengths and the current HIT queue lengths of the workers to guide its HIT assignment decisions. It improves the social welfare of an entire crowdsourcing system by increasing the amount of business the system can efficiently process compared to existing reputation-aware decision-making approaches at a small sacrifice in the average quality of the HIT results received by the requesters.

Currently, the proposed approach operates in a centralized manner with the assumption that HITs require homogeneous effort level. In future work, we plan to enhance the proposed approach with HITs requiring heterogeneous effort levels and design new approaches that maximize social welfare through distributed decision-making.

4. REFERENCES

[1] Doan, A., Ramakrishnan, R., Halevy, A.Y., 2011. Crowdsourcing Systems on the World-Wide Web. Communications of the ACM 54, 86-96.

[2] Yu, H., Shen, Z., Miao, C., Leung, C., Niyato, D., 2010. A Survey of Trust and Reputation Management Systems in Wireless Communications. Proceedings of the IEEE 98, 1755-1772.

[3] Yu, H., Shen, Z., Miao, C., An, B., 2012. Challenges and Opportunities for Trust Management in Crowdsourcing. IEEE/WIC/ACM International Conference on Intelligent Agent Technology (IAT), 486-493.

[4] Monderer, D. and Shapley, L.S., 1996. Potential Games. Games and Economic Behavior, 14, 124-143.

[5] Neely, M.J., 2010. Stochastic Network Optimization with Application to Communication and Queueing Systems. Morgan and Claypool Publishers.