A Truthful Mechanism with Biparameter Learning for Online Crowdsourcing

(Extended Abstract)

Satyanath Bhat, Divya Padmanabhan, Shweta Jain, Yadati Narahari
Indian Institute of Science, Bangalore 560012, India
{satyanath.bhat, divya.padmanabhan, jainshweta, hari}@csa.iisc.ernet.in

General Terms

Algorithms, Economics, Theory

Keywords

Strategic Agents, Mechanism Design

1. INTRODUCTION

Consider the motivating example of crowdsourcing of a sequence of large translation jobs which arrive online. Each job has to be completed within a deadline and with an assured level of accuracy. To accomplish this, the requester could split such a job into tasks and allocate each task to a crowd worker. A worker, if employed for a long duration, might commit errors. We refer to the duration until which an agent works without committing any error as the time to failure (TTF). The time taken by a worker to complete the job all by himself is called the job completion time (JCT) of that worker. Each worker incurs a certain cost to complete the entire job. Note that the workers are heterogeneous in terms of their costs and stochastic parameters JCT and TTF. Importantly, the crowd workers are strategic and may misrepresent their costs in the hope of gaining higher utility.

In this work, we consider jobs which (a) arrive online, (b) are divisible (into tasks), (c) have strict completion deadlines, and (d) are to be completed with an assured accuracy. We propose a multi-armed bandit (MAB) mechanism which learns the two parameters (mean job completion time (MJCT) and mean time to failure (MTTF)) of the workers while eliciting their privately held costs truthfully [1].

2. THE MODEL

Let $N = \{1, \ldots, n\}$ denote the set of crowd workers (agents) available to the requester. A sequence of T homogeneous jobs arrives at the platform, one at a time. Following are some of the design issues of the requester.

Job Parameters: *(a) Deadline*: We use D to denote the deadline on each job starting from the arrival of that job, before which the job is required to be completed in expectation. *(b) Task creation*: The requester can divide a current job t $(t = 1, \ldots, T)$ into a certain number of tasks to meet deadline D. We use $x_i^{(t)}$ to denote the fraction of the job t assigned as a task to the worker i. We call an allocation x_i *feasible* when $0 \le x_i^{(t)} \le 1$ and $\sum_{i=1}^{n} x_i^{(t)} = 1$. *(c) Threshold on probability of failure*: We say a worker has failed when he commits an error. We use ε to denote (the common) threshold on probability of failure for any task.

Worker Parameters: *(a) Job Completion Time (JCT)*: A worker has a stochastic job completion time, which is the time he requires to complete the entire job by himself. JCT for a worker is a random variable with a fixed but unknown mean. We refer to the mean job completion time (MJCT) of worker i as ρ_i, which the requester wishes to learn. The task allocation $x_i^{(t)}$ will meet the deadline constraint in expectation if $x_i^{(t)} \times \rho_i \le D$. *(b) Time to Failure (TTF)*: A worker is also characterized by a stochastic time to failure, which denotes the duration for which a worker would work without a failure. Like JCT, TTF also has a fixed yet unknown mean (MTTF). Let F_i be the CDF of TTF for agent i, the requirement on threshold probability error dictates $F_i(x_i^{(t)} \times \rho_i) \le \varepsilon$. Like MJCT, a requester also seeks to learn MTTF. *(c) Cost Incurred*: Worker i has a privately held cost $c_i \in [\underline{c}, \bar{c}]$ which represents the cost incurred by worker i to complete the job entirely on his own. Therefore, the cost involved to complete $x_i^{(t)}$ fraction of the job by the worker i is $c_i x_i^{(t)}$.

Goal of the Optimization Problem: The constraints on deadline and task accuracy, which in turn depend on MJCT and MTTF, have to be met in a cost optimal way for every online job t. Thus, this problem is a biparameter learning problem.

We model the JCT of a worker as a log-normal distribution with unknown mean $\rho_i \in [\underline{\rho}, \bar{\rho}]$ while the TTF is modelled an exponential distribution with mean $\beta_i \in [\underline{\beta}, \bar{\beta}]$. If the parameters c_i, ρ_i, and β_i are known, the requester's optimization problem is given by eq. (1).

$$\min_{\substack{\text{Feasible } x_i \\ x_i \in [0,1]}} \sum_{i=1}^{n} c_i x_i, \text{ s.t. } , \rho_i x_i \le \min\left(D, \beta_i \ln\left(\frac{1}{1-\varepsilon}\right)\right) \ \forall i \in N. \tag{1}$$

In practice, ρ_i and β_i are unknown and need to be learnt.

2.1 Difficulty in Learning β_i

In learning of ρ_i, every allocation contributes a single sample. However, for estimating β_i, each input sample must correspond to a failure, but this is not practical as we do not observe failure at every instance of allocation. To handle this difficulty, we propose to use a surrogate random vari-

Appears in: *Proceedings of the 15th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2016), J. Thangarajah, K. Tuyls, C. Jonker, S. Marsella (eds.), May 9–13, 2016, Singapore.*

able. Consider the experiment where a worker i is allocated a task on which the worker spends a duration of at least δ. The experiment is deemed to have failed if the worker i fails in the first δ duration of allocation, otherwise it is deemed to be a success. Let $N_\delta^{(i)}$ be the number of such independent experiments until a failure is encountered. We propose to use the random variable $\beta'_{\delta,i} = \delta \times N_\delta^{(i)}$ to construct a sample from exponential(β_i). The expectation of the surrogate random variable in the limit coincides with β_i(MTTF).

2.2 Non strategic workers: SW-GREEDY

For learning MTTF and MJCT, we use the Robust UCB technique [2] with truncated empirical mean as the estimator. Let $\hat{\rho}_i^+$ and $\hat{\beta}_i^+$ denote the upper confidence indices while $\hat{\rho}_i^-$ and $\hat{\beta}_i^-$ denote the lower confidence indices of MJCT and MTTF, respectively, obtained from Robust UCB. $\hat{\rho}_i$ and $\hat{\beta}_i$ are the empirical estimates of MJCT and MTTF, respectively, for worker i. The workers are indexed in increasing order of their costs and each worker i is allocated the largest possible fraction x_i^{PST} which does not violate the constraints in eq. (1) until all tasks of the job are allocated. A higher value of ρ_i enforces a lower allocation to worker i compared to when a lower value of ρ_i is used. Hence we refer to $\hat{\rho}_i^+$ as a pessimistic estimate for ρ_i. Similarly, we refer to $\hat{\beta}_i^-$ as the pessimistic estimate for β_i. The use of pessimistic estimates ensures that even with the true values of underlying means, the constraint in eq. (1) is satisfied. We refer to the above allocation as SW-GREEDY (Algorithm 1). Every worker i is paid an amount equal to the cost incurred, i.e. $x_i^{\text{PST}}(t) \times c_i$.

3. STRATEGIC WORKERS: TD-UCB

Here, before an allocation is performed, the agents announce their bids. We denote the bid profile by (b_i, b_{-i}). We introduce a mechanism TD-UCB for which we use the same allocation rule as SW-GREEDY .

3.1 Payment Scheme

Let ξ_t denote a tuple of allocation and performance of the allocated workers for the job t. The learning until job t is captured in the history $h_t = \{\xi_k\}_{k=0}^t$. We denote the externality imposed by agent i on j as $x_{i,j}^{\text{EXT}}(b_i, b_{-i}; h_t, t)$, which signifies the additional fraction of the job allocated to the agent j in the absence of agent i. The externality for the job t depends on the bid profile (b_i, b_{-i}) as well as the history till job t. Let $\overline{k_t}$ be the agent with the largest reported bid in the worker set chosen by the allocation scheme.

$$x_{i,j}^{\text{EXT}}(b_i, b_{-i}; t) = \begin{cases} 0 & \text{if } j < \overline{k_t} \text{ or } i > \overline{k_t}, \\ Z_1 & \text{if } j = \overline{k_t}, \\ Z_2 & \text{if } j > \overline{k_t}, \quad \text{where,} \end{cases} \tag{2}$$

$$Z_1 = \min\left(\frac{1}{\hat{\rho}_j^+} \min\left(D, \hat{\beta}_j^- \log\left(\frac{1}{1-\epsilon}\right)\right) - x_j^{\text{PST}}(t), x_i^{\text{PST}}(t) \right),$$

$$Z_2 = \min\left(\frac{1}{\hat{\rho}_j^+} \min\left(D, \hat{\beta}_j^- \log\left(\frac{1}{1-\epsilon}\right)\right), x_i^{\text{PST}}(t) - \sum_{s=\overline{k}_t}^{j-1} x_{i,s}^{\text{EXT}}(t) \right)$$

We now propose a payment structure in eq. (3).

$$p_i(b_i, b_{-i}; t) = \begin{cases} 0 & \text{if } i > \overline{k_t}, \\ Z_3 & \text{otherwise}, \quad \text{where,} \end{cases} \tag{3}$$

ALGORITHM 1: SW-GREEDY Allocation Algorithm

Input: Set of workers N, number of jobs T, deadline D, accuracy level ε, input cost vector: $c_1 \le c_2 \le \ldots \le c_n$ (By re-indexing N)

$\forall i \in N$, $\hat{\rho}_i = \bar{\rho}$, $\hat{\rho}_i^+ = \bar{\rho}$, $\hat{\rho}_i^- = \underline{\rho}$, $N_{i,t=0}$
  $\hat{\beta}_i = \underline{\beta}$, $\hat{\beta}_i^+ = \bar{\beta}$, $\hat{\beta}_i^- = \underline{\beta}$, $N_{i,t}^\beta = 0$,
$\eta_\delta^{(i)} = 0$ Initialize

for *Online job arrival* $t = 1, \ldots, T$ **do**
  $x^{\text{PST}}(t) = \{x_1^{\text{PST}}(t), \ldots, x_n^{\text{PST}}(t)\} = \{0, \ldots, 0\}$
  $i = 1$ Pessimistic Selection
  while $\sum_{j=1}^n x_j^{\text{PST}}(t) < 1$ **do**
    $x_i^{\text{PST}}(t) = \frac{1}{\hat{\rho}_i^+} \min\left(D, \beta_i^- \ln\left[\frac{1}{1-\varepsilon}\right]\right)$
    if $1 - \sum_{j=1}^{i-1} x_j^{\text{PST}}(t) < x_i^{\text{PST}}(t)$ **then**
      $x_i^{\text{PST}}(t) = 1 - \sum_{j=1}^{i-1} x_j^{\text{PST}}(t)$
    $i = i + 1$
  Define $\bar{k}_t = \max\{i : x_i^{\text{PST}} > 0\}$
  Allocate the job t as per x_i^{PST}
  Observe $\tilde{\tau}_i$, the time of completion of x_i^{PST} by i
  $\forall i \in \{1, \ldots, \bar{k}_t\}, N_{i,t} = N_{i,t-1} + 1$
  $\hat{\rho}_i = \left(N_{i,t-1} \times \hat{\rho}_i + \frac{\tilde{\tau}_i}{x_i^{\text{PST}}(t)}\right) \times \frac{1}{N_{i,t}}$
  for $i \in \{1, \ldots, \bar{k}_t\}$ **do**
    if *Worker i made an error during δ* **then**
      $\hat{\beta}_i = \frac{\hat{\beta}_i \times N_{i,t-1}^\beta + (\delta \times \eta_\delta^{(i)})}{N_{i,t-1}^\beta + 1}$
      $N_{i,t}^\beta = N_{i,t-1}^\beta + 1$ Updates for Surrogate of β_i
      $\eta_\delta^{(i)} = 0$
    else
      $\eta_\delta^{(i)} = \eta_\delta^{(i)} + 1$
      $N_{i,t}^\beta = N_{i,t-1}^\beta$

$$Z_3 = \sum_{s=\overline{k}_t}^{n} \left[x_{i,s}^{\text{EXT}}(t) \times b_s\right] + \left(x_i^{\text{PST}}(t) - \sum_{s=\overline{k}_t}^{n} x_{i,s}^{\text{EXT}}(t) \right) \times \bar{c}$$

4. PROPERTIES OF TD-UCB MECHANISM

THEOREM 1. *The TD-UCB mechanism is Dominant Strategy Incentive Compatible and Individually Rational.*

DEFINITION 1. *Optimal worker set: For a problem instance with all the parameters known, in the solution to the optimization problem of eq. (1), we refer to the set of agents allocated a non-zero fraction of the job as the the optimal worker set.*

THEOREM 2. *The TD-UCB mechanism selects an optimal worker set after the job $t' \in O(\log T)$.*

THEOREM 3. *The average regret of TD-UCB mechanism approaches zero asymptotically.*

References

[1] S. Bhat, D. Padmanabhan, S. Jain, and Y. Narahari. A truthful mechanism with biparameter learning for online crowdsourcing. *arXiv preprint arXiv:1602.04032*, 2016.

[2] S. Bubeck, N. Cesa-Bianchi, and G. Lugosi. Bandits with heavy tail. *IEEE Transactions on Information Theory*, 59(11):7711–7717, 2013.