Melding Mechanism Design with Machine Learning: A Multi-Armed Bandit Crowdsourcing Mechanism with Incentive Compatible Learning

Y. Narahari

Sept 10, 2014

Joint Work With: Shweta Jain, Sujit Gujar, Satyanath Bhat, Onno Zoeter

Mechanism Design (MD)

Given: A set of utility maximizing (strategic) agents with private information and a social choice function that captures desirable (social) goals.

MD provides a game theoretic setting to explore if the given social choice function can be implemented as an equilibrium outcome of an induced game.

Example: Vickery Auction[1]

[1] Y.Narahari: Game Theory and Mechanism Design. IISc Press and WSPC,2014

Machine Learning (ML)

In a multiagent setting, ML seeks to learn the preferences or types of the agents through some available data or through intelligent exploration.

Example: UCB algorithm[2]

[2] Peter Auer, Nicolò Cesa-Bianchi, and Paul Fischer. Finite-time Analysis of the Multiarmed Bandit Problem. Machine Learning, 2002.

Melding ML and MD

- Modern problems involve strategic agents, private information, unknown information, some opportunities to explore and interact with agents, etc.
- Examples:
  - Sponsored Search Auctions on the Web
  - Crowdsourcing
  - Online Auctions/Internet Markets
  - etc.
- ML and MD are extremely well investigated as individual problems. Interesting research questions arise when you try to meld them.

Melding ML and MD (Continued)

- There have been several efforts in the literature in the past decade seeking a fusion of ML and MD[3]
  - Mostly special settings (in particular sponsored search auctions)
- A general approach is elusive
- We have made some baby steps [4]

[3]Nicolo Cesa-Bianchi and Gabor Lugosi. Prediction, Learning, and Games. Cambridge University Press, 2006

[4]Akash Das Sharma, Sujit Gujar, and Y. Narahari. Truthful multi-armed bandit mechanisms for multi-slot sponsored search auctions. Current Science, 2012

Motivation

- Providing investment advice on financial stocks (Goldman Sachs)
- Classifying patents (Infosys)
- Classifying legal documents (Xerox)
- Categorizing video clips (HP)

These problems belong to the realm of crowdsourcing or expertsourcing

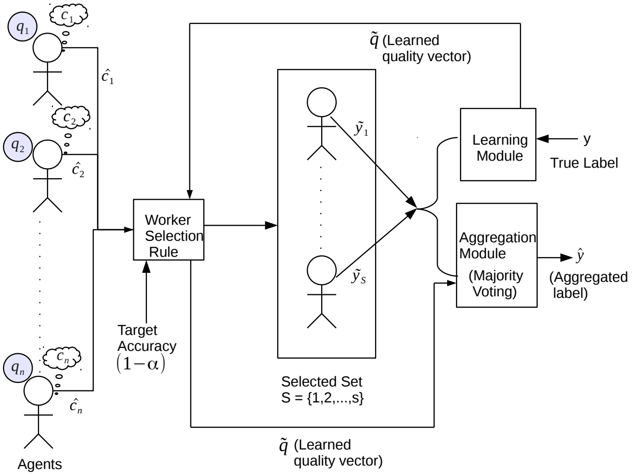
Challenges in Crowdsourcing

- Assuring quality
- Aggregating labels that may be incorrect with certain probability
  - Limit number of workers
  - Minimize error probability
- Learning qualities with high confidence
- Eliciting costs truthfully

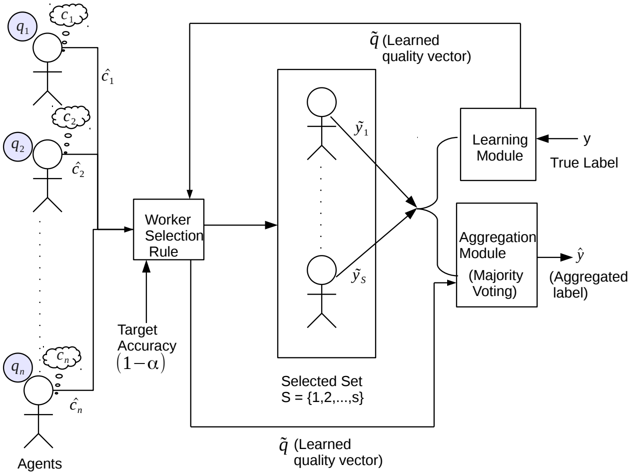
The Model

- Abstraction of the problem:
- Sequence of documents to be labeled (0 or 1), $t = 1, 2, \ldots, T$
- Worker i can provide the document label $\{0, 1\}$ with accuracy q_i (unknown)
- Worker i incurs a cost c_i to label a document (private information)
- Need to select $S^t \subset \mathcal{N}$ for $t = 1, 2, \ldots, T$, such that the requester has confidence of at least $1 - \alpha$ on aggregated answer $\forall\ t = 1, 2, \ldots, T$
- Multi Armed Bandit Mechanism with subset selection.

Our Setting


q_1
c_1
$\hat{c}_1$
q_2
c_2
$\hat{c}_2$
q_n
c_n
$\hat{c}_n$
Agents
Worker Selection Rule
Target Accuracy $(1-\alpha)$
$\tilde{q}$ (Learned quality vector)
$\tilde{y}_1$
$\tilde{y}_S$
Selected Set S = {1,2,...,s}
Learning Module
y
True Label
Aggregation Module (Majority Voting)
$\hat{y}$
(Aggregated label)
$\tilde{q}$ (Learned quality vector)

Our Setting



$P(\hat{y} \neq y) < \alpha$ (Accuracy Constraint)

Our Work

- Non-Strategic Version
  1. An adaptive exploration separated algorithm: Non-Strategic Constrained Confidence Bound (CCB-NS)
  2. Though the true qualities are not known, the algorithm makes sure that the accuracy constraint is satisfied with high probability
  3. We provide an upper bound on the number of exploration steps

[5] Moshe Babaioff, Robert D. Kleinberg, and Aleksandrs Slivkins. Truthful mechanisms with implicit payment computation. Journal of ACM 2014, EC-2010

Our Work

- Non-Strategic Version
  1. An adaptive exploration separated algorithm: Non-Strategic Constrained Confidence Bound (CCB-NS)
  2. Though the true qualities are not known, the algorithm makes sure that the accuracy constraint is satisfied with high probability
  3. We provide an upper bound on the number of exploration steps
- Strategic Version
  1. A modification of the algorithm CCB-NS: strategic constrained confidence bound (CCB-S)
  2. We prove that the allocation rule provided by CCB-S is ex-post monotone
  3. Given this ex-post monotone allocation rule, we invoke the technique from Babaioff et. al. [5] to design an ex-post truthful and ex-post individually rational mechanism

[5] Moshe Babaioff, Robert D. Kleinberg, and Aleksandrs Slivkins. Truthful mechanisms with implicit payment computation. Journal of ACM 2014, EC-2010

The Optimization Problem (to be solved for each task)

- Suppose $q = (q_1, q_2, \ldots, q_n)$ is a vector of qualities of workers
- Let $f_S(q)$ represent some measure of error probability
- $1 - f_S(q)$ represents the accuracy with quality profile q
- For each task, $t = 1, 2, \ldots, T$, we wish to solve:

.

$$\min_{S^t \subseteq N} \sum_{i \in S^t} c_i \tag{1}$$

$$\text{s.t. } f_{S^t}(q) < \alpha \quad \text{(Accuracy Constraint)} \tag{2}$$

Properties of Error Probability Function

- *Monotonicity:* $f_S(q)$ is monotone if for all quality profiles q and q' such that $\forall i \in \mathcal{N},\ q'_i \leq q_i,$ we have,

$$f_S(q') < \alpha \implies f_S(q) < \alpha \ \ \forall S \subseteq \mathcal{N},\ \forall \alpha \in [0,1]$$

Properties of Error Probability Function

- *Monotonicity:* $f_S(q)$ is monotone if for all quality profiles q and q' such that $\forall i \in \mathcal{N},\ q'_i \leq q_i,$ we have,

$$f_S(q') < \alpha \implies f_S(q) < \alpha \ \ \forall S \subseteq \mathcal{N},\ \forall \alpha \in [0, 1]$$

- *Bounded smoothness:* $f_S(q)$ satisfies bounded smoothness property if there exists a monotone continuous function h such that if

$$\max_i |q_i - q'_i| \leq \delta \implies |f_S(q) - f_S(q')| \leq h(\delta) \ \ \forall S \subseteq \mathcal{N}, \forall q, q' \in [0.5, 1]$$

Assumptions in the Model

- The error probability function satisfies the assumptions of monotonicity and bounded smoothness
- The true label is observed once the task is completed
- If all workers are selected, then the constraint is satisfied with respect to true qualities (enough number of workers are available)

Learning Qualities in Our Model

- Extension of UCB algorithm[6]
  - Suppose $\hat{q}_i(t)$ is mean success observed so far till t tasks
  - Let $n_{i,t}$ be the number of tasks assigned to worker i till t tasks
  - Let μ be the confidence parameter for estimating q_i
- Do the following for $t = 1, 2, \ldots, T$
- Maintain upper confidence bound and lower confidence bound on qualities
  - UCB: $\hat{q}_i^+(t) = \hat{q}_i(t) + \sqrt{\frac{1}{2n_{i,t}} ln(\frac{1}{\mu})}$
  - LCB: $\hat{q}_i^-(t) = \hat{q}_i(t) - \sqrt{\frac{1}{2n_{i,t}} ln(\frac{1}{\mu})}$
- Solve the optimization problem with respect to the UCB
- Check if the constraint is satisfied with respect to LCB
  - If not, add a minimal set of additional workers so as to satisfy the constraint
  - If yes, allocate this optimal set for every future task

[6] Peter Auer, Nicolò Cesa-Bianchi, and Paul Fischer. Finite-time Analysis of the Multiarmed Bandit Problem. Machine Learning, 2002

CCB-NS Algorithm

Input: Task error tolerance α, confidence level μ, tasks $\{1, 2, \ldots, T\}$, workers $\mathcal{N}$, costs c
Output: Worker selection set S^t, Label $\hat{y}_t$ for task t
Initialization: $\forall i,\ \hat{q}_i^+ = 1,\ \hat{q}_i^- = 0.5,\ k_{i,1} = 0,\ S^1 = \mathcal{N}$, and $\hat{y}_1 = \text{AGGREGATE}(\tilde{y}(S^1))$
Observe true label y_1
$\forall i \in \mathcal{N},\ n_{i,1} = 1,\ k_{i,1} = 1$ if $\tilde{y}_i = y_1$ and $\hat{q}_i = k_{i,1}/n_{i,1}$

for $t = 2$ to T

Let $S^t = \arg\min_{S \subseteq \mathcal{N}} \sum_{i \in S} c_i$ s.t. $f_S(\hat{q}^+) < \alpha$

% Explore

if $f_{S^t}(\hat{q}^-) > \alpha$ **then**

$S^t = S^t \cup minimal(\mathcal{N} \setminus S^t),\ \hat{y}_t = \text{AGGREGATE}(\tilde{y}(S^t))$

Observe true label y_t; $\forall i \in S^t$: $n_{i,t} = n_{i,t} + 1,\ k_{i,t} = k_{i,t} + 1$ if $\tilde{y}_i = y_t,\ \hat{q}_i = k_{i,t}/n_{i,t}$,

$\hat{q}_i^+ = \hat{q}_i + \sqrt{\frac{1}{2n_{i,t}} ln(\frac{2}{\mu})},\ \hat{q}_i^- = \hat{q}_i - \sqrt{\frac{1}{2n_{i,t}} ln(\frac{2}{\mu})}$

else

$t^* = t,\ \hat{y}_t = \text{AGGREGATE}(\tilde{y}(S^t))$

Break

%Exploit

for $t = t^* + 1$ to T

$S^t = S^{t^*},\ \hat{y}_t = \text{AGGREGATE}(\tilde{y}(S^t))$

Properties of CCB-NS

- CCB-NS is an adaptive exploration separated learning algorithm

Theorem

CCB-NS satisfies the accuracy constraint with probability at least $(1 - \mu)$ *at every round* t

Lemma

Set S^{t^*} *returned by the CCB-NS algorithm is an optimal set with probability at least* $1 - \mu$. *That is,* $C(S^{t^*}) = C(S^*)$ *w.p.* $(1 - \mu)$

Can be proved using monotonicity properties of error probability function

Properties of CCB-NS

- Let $\Delta = \min_{S \subseteq \mathcal{N}} |f_S(.) - \alpha|$
  (minimum difference between error tolerance α and error probability value of any set)

Theorem

The number of exploration rounds by the CCB-NS algorithm is bounded by $\frac{2n}{(h^{-1}(\Delta))^2} ln(\frac{2n}{\mu})$ *with probability* $(1 - \mu)$

where h is the bounded smoothness function

Strategic Version

- Costs c_i are private information of workers
- Valuations: $v_i = -c_i, \forall i \in \mathcal{N}$
- Utilities (Quasilinear):
$$\sum_{t=1}^{T} u_i(\tilde{q}(t), \hat{c}, c_i) = -c_i \sum_{t=1}^{T} \mathcal{A}_i^t(\tilde{q}(t), \hat{c}) + \mathcal{P}_i^t(\tilde{q}(t), \hat{c})$$
- $\mathcal{A}_i^t$ represents whether the t^{th} task is allocated to worker i
- $\mathcal{P}_i^t$ is the monetary transfer to worker i for task t

Some Definitions

- Success Realization: A success realization is a matrix s.t.,

$$\rho_{it} = \begin{cases} 1 \text{ If } \tilde{y}_i^t = y_t \\ 0 \text{ if } \tilde{y}_i^t \neq y_t \\ -1 \text{ if worker } i \text{ is not selected for task } t \end{cases}$$

- Ex-post Monotone Allocation: An allocation rule $\mathcal{A}$ is ex-post monotone if $\forall \rho \in \{0, 1, -1\}^{n \times T}, \;\; \forall i \in \mathcal{N}, \;\; \forall \hat{c}_{-i} \in [0, 1]^{n-1}$,

$$\hat{c}_i \leq \hat{c}_i' \Rightarrow \mathcal{A}_i(\hat{c}_i, \hat{c}_{-i}; \rho) \geq \mathcal{A}_i(\hat{c}_i', \hat{c}_{-i}; \rho)$$

  $\mathcal{A}_i(\hat{c}_i, \hat{c}_{-i})$ is the number of tasks assigned to worker i with bids $\hat{c}_i$ and $\hat{c}_{-i}$

- Ex-post Truthful Mechanism: A mechanism $\mathcal{M} = (\mathcal{A}, \mathcal{P})$ is ex-post truthful if $\forall \rho \in \{0, 1, -1\}^{n \times T}, \;\; \forall i \in \mathcal{N}, \;\; \forall \hat{c}_{-i} \in [0, 1]^{n-1}$,

$$-c_i \mathcal{A}_i(c_i, \hat{c}_{-i}; \rho) + \mathcal{P}_i(c_i, \hat{c}_{-i}; \rho) \geq -c_i \mathcal{A}_i(\hat{c}_i, \hat{c}_{-i}; \rho) + \mathcal{P}_i(\hat{c}_i, \hat{c}_{-i}; \rho) \; \forall \hat{c}_i \in [0, 1]$$

Deriving an Ex-Post Truthful Mechanism from an Ex-Post Monotone Allocation

- Work by Babaioff, Kleinberg, and Slivkins[7] reduces the problem of designing truthful mechanisms to that of monotone allocations
- If $\mathcal{A}$ is ex-post monotone MAB allocation rule, then one can compute payment scheme to obtain a MAB mechanism $\mathcal{M}$ such that it is ex-post truthful and ex-post individually rational

[7] Babaioff, Moshe and Kleinberg, Robert D. and Slivkins, Aleksandrs. Truthful mechanisms with implicit payment computation, Journal of ACM 2014 (EC 2010)

An Ex-post Monotone Algorithm, CCB-S

- Suppose $\hat{q}_i(t)$ is mean success observed so far till t tasks
- Let $n_{i,t}$ be the number of tasks assigned to worker i till t tasks
- Let μ be the confidence parameter for estimating q_i
- Do the following for $t = 1, 2, \ldots, T$
- Maintain upper confidence bound and lower confidence bound on qualities
  - UCB: $\hat{q}_i^+(t) = \hat{q}_i(t) + \sqrt{\frac{1}{2n_{i,t}} ln(\frac{1}{\mu})}$
  - LCB: $\hat{q}_i^-(t) = \hat{q}_i(t) - \sqrt{\frac{1}{2n_{i,t}} ln(\frac{1}{\mu})}$
- Solve the optimization problem with respect to the UCB
- Check if the constraint is satisfied with respect to LCB
  - If not, select all workers
  - If yes, allocate this optimal set for every future task

CCB-S algorithm

Input: Task error tolerance α, confidence level μ, tasks $\{1, 2, \ldots, T\}$, workers $\mathcal{N}$, costs c
Output: Worker selection set S^t, Label $\hat{y}_t$ for task t
Initialization: $\forall i$, $\hat{q}_i^+ = 1$, $\hat{q}_i^- = 0.5$, $k_{i,1} = 0$, $S^1 = \mathcal{N}$, and $\hat{y}_1 = \text{AGGREGATE}(\tilde{y}(S^1))$
Observe true label y_1
$\forall i \in \mathcal{N}$, $n_{i,1} = 1$, $k_{i,1} = 1$ if $\tilde{y}_i = y_1$ and $\hat{q}_i = k_{i,1}/n_{i,1}$
for $t = 2$ to T

Let $S^t = \arg\min_{S \subseteq \mathcal{N}} \sum_{i \in S} c_i$ s.t. $f_S(\hat{q}^+) < \alpha$

% Explore

if $f_{S^t}(\hat{q}^-) > \alpha$ **then**

$S^t = \mathcal{N}$, $\hat{y}_t = \text{AGGREGATE}(\tilde{y}(S^t))$

Observe true label y_t; $\forall i \in S^t$: $n_{i,t} = n_{i,t} + 1$, $k_{i,t} = k_{i,t} + 1$ if $\tilde{y}_i = y_t$, $\hat{q}_i = k_{i,t}/n_{i,t}$,

$\hat{q}_i^+ = \hat{q}_i + \sqrt{\frac{1}{2n_{i,t}} ln(\frac{2}{\mu})}$, $\hat{q}_i^- = \hat{q}_i - \sqrt{\frac{1}{2n_{i,t}} ln(\frac{2}{\mu})}$

else

$t^* = t$, $\hat{y}_t = \text{AGGREGATE}(\tilde{y}(S^t))$

Break

%Exploit

for $t = t^* + 1$ to T

$S^t = S^{t^*}$, $\hat{y}_t = \text{AGGREGATE}(\tilde{y}(S^t))$

Properties of CCB-S

- Satisfies all the properties that were satisfied by CCB-NS algorithm

Theorem

Number of exploration rounds by the CCB-S algorithm is bounded by $\frac{2}{(h^{-1}(\Delta))^2} ln(\frac{2n}{\mu})$ *with probability* $(1 - \mu)$

Properties of CCB-S

- Satisfies all the properties that were satisfied by CCB-NS algorithm

Theorem

Number of exploration rounds by the CCB-S algorithm is bounded by $\frac{2}{(h^{-1}(\Delta))^2} ln(\frac{2n}{\mu})$ *with probability* $(1 - \mu)$

Note:

CCB-S algorithm only provides an allocation rule. For the payment rule we use transformation given by Babaioff et. al. [a] as a black box

[a] Babaioff, Moshe and Kleinberg, Robert D. and Slivkins, Aleksandrs. Truthful mechanisms with implicit payment computation, Journal of ACM 2014 (EC 2010)

Properties of CCB-S (continued)

Theorem

Allocation rule given by the CCB-S algorithm ($\mathcal{A}^{CCB-S}$) is ex-post monotone and thus produces an ex-post incentive compatible and ex-post individual rational mechanism

Proof:

We need to prove:

$$\mathcal{A}_i^t(\hat{c}_i, c_{-i}; \rho) \leq \mathcal{A}_i^t(c_i, c_{-i}; \rho)$$
$$\forall \rho \forall i \in \mathcal{N},\ \forall t \in \{1, 2, \ldots, T\},\ \forall \hat{c}_i \geq c_i$$

For notation convenience, assume ρ is fixed and denote $\mathcal{A}_i^t(\hat{c}_i, c_{-i}; \rho)$ as $\mathcal{A}_i^t(\hat{c}_i, c_{-i})$

Proof Continued

- Prove by induction:

Proof Continued

- Prove by induction:
- $\mathcal{A}_j^1(\hat{c}_i, c_{-i}) = \mathcal{A}_j^1(c_i, c_{-i}) = 1 \ \forall j$ (since task 1 is given to all workers)

Proof Continued

- Prove by induction:
- $\mathcal{A}_j^1(\hat{c}_i, c_{-i}) = \mathcal{A}_j^1(c_i, c_{-i}) = 1 \; \forall j$ (since task 1 is given to all workers)
- Let t be the largest time step such that, $\forall j$, $\mathcal{A}_j^{t-1}(\hat{c}_i, c_{-i}) = \mathcal{A}_j^{t-1}(c_i, c_{-i}) = t - 1$ (Exploration round with $\hat{c}_i$ and c_i)

Proof Continued

- Prove by induction:
- $\mathcal{A}_j^1(\hat{c}_i, c_{-i}) = \mathcal{A}_j^1(c_i, c_{-i}) = 1 \; \forall j$ (since task 1 is given to all workers)
- Let t be the largest time step such that, $\forall j$, $\mathcal{A}_j^{t-1}(\hat{c}_i, c_{-i}) = \mathcal{A}_j^{t-1}(c_i, c_{-i}) = t-1$ (Exploration round with $\hat{c}_i$ and c_i)
- And $\exists i$ such that,

$$\mathcal{A}_i^t(\hat{c}_i, c_{-i}) \neq \mathcal{A}_i^t(c_i, c_{-i})$$

Proof Continued

- Prove by induction:
- $\mathcal{A}_j^1(\hat{c}_i, c_{-i}) = \mathcal{A}_j^1(c_i, c_{-i}) = 1 \; \forall j$ (since task 1 is given to all workers)
- Let t be the largest time step such that, $\forall j$, $\mathcal{A}_j^{t-1}(\hat{c}_i, c_{-i}) = \mathcal{A}_j^{t-1}(c_i, c_{-i}) = t-1$ (Exploration round with $\hat{c}_i$ and c_i)
- And $\exists i$ such that,

$$\mathcal{A}_i^t(\hat{c}_i, c_{-i}) \neq \mathcal{A}_i^t(c_i, c_{-i})$$

- Since the costs and quality estimates are the same for all the workers till tasks t, this can happen only when in one case worker i is selected, while in the other case worker i is not selected

Proof Continued

- Prove by induction:
- $\mathcal{A}_j^1(\hat{c}_i, c_{-i}) = \mathcal{A}_j^1(c_i, c_{-i}) = 1 \; \forall j$ (since task 1 is given to all workers)
- Let t be the largest time step such that, $\forall j$, $\mathcal{A}_j^{t-1}(\hat{c}_i, c_{-i}) = \mathcal{A}_j^{t-1}(c_i, c_{-i}) = t-1$ (Exploration round with $\hat{c}_i$ and c_i)
- And $\exists i$ such that,

$$\mathcal{A}_i^t(\hat{c}_i, c_{-i}) \neq \mathcal{A}_i^t(c_i, c_{-i})$$

- Since the costs and quality estimates are the same for all the workers till tasks t, this can happen only when in one case worker i is selected, while in the other case worker i is not selected
- Let the two sets selected with c_i and $\hat{c}_i$ be $S(c_i)$ and $S(\hat{c}_i)$ respectively

Proof Continued

- Since the optimization problem involves cost minimization and quality updates are the same, we have,

$$\mathcal{A}_i^t(\hat{c}_i, c_{-i}) = t - 1 \text{ which implies } i \notin S(\hat{c}_i)$$
$$\mathcal{A}_i^t(c_i, c_{-i}) = t \text{ which implies } i \in S(c_i)$$

Proof Continued

- Since the optimization problem involves cost minimization and quality updates are the same, we have,

$$\mathcal{A}_i^t(\hat{c}_i, c_{-i}) = t - 1 \text{ which implies } i \notin S(\hat{c}_i)$$
$$\mathcal{A}_i^t(c_i, c_{-i}) = t \text{ which implies } i \in S(c_i)$$

- Since $i \notin S(\hat{c}_i)$, selected set $S(\hat{c}_i)$ satisfies the lower confidence bound too (exploitation round with bid $\hat{c}_i$)

Proof Continued

- Since the optimization problem involves cost minimization and quality updates are the same, we have,

$$\mathcal{A}_i^t(\hat{c}_i, c_{-i}) = t - 1 \text{ which implies } i \notin S(\hat{c}_i)$$
$$\mathcal{A}_i^t(c_i, c_{-i}) = t \text{ which implies } i \in S(c_i)$$

- Since $i \notin S(\hat{c}_i)$, selected set $S(\hat{c}_i)$ satisfies the lower confidence bound too (exploitation round with bid $\hat{c}_i$)
- Thus for the rest of the tasks, only $S(\hat{c}_i)$ is selected and thus we have, $\mathcal{A}_i^t(\hat{c}_i, c_{-i}) \leq \mathcal{A}_i^t(c_i, c_{-i})$

Summary So Far

Our Contributions[8]

- Proposed a novel framework Assured Accuracy Bandit (AAB)
- Developed an adaptive exploration separated algorithm, Constrained Confidence Bound (CCB-S)
- Provided an upper bound on the number of exploration steps
- CCB-S algorithm leads to an ex-post truthful and ex-post individual rational mechanism

[8] Shweta Jain and Sujit Gujar and Y Narahari and Onno Zoeter, "A Quality Assuring Multi-Armed Bandit Crowdsourcing Mechanism with Incentive Compatible Learning". To *appear in AAMAS'14*

Directions for Future Work

- Non exploration-separated algorithms satisfying desirable mechanism properties with lower regret
- An approximate mechanism that solves the optimization problem efficiently
- Provide lower bounds on the regret in this setting
- Computational Issues
- Extension to more general task settings

Conclusion

- Melding ML and MD is an interesting problem with many exciting challenges ahead
- The ultimate goal is to evolve a general framework to address a wide class of problems
  - Will be a symphony of GT, ML, and Optimization

Thank You

Learning in Crowdsourcing

- A setting of assured quality per task is considered uniform cost workers [9]
- The goal of selecting a single optimal crowd for a single task is considered by aggregating the answers in a sequential way until a certain accuracy is achieved for each task with Homogeneous workers having same quality in a crowd [10]
- Efficient selection of a single worker for each task based on MAB algorithms by formulating a knapsack problem [11]

[9] Chien ju Ho, Shahin Jabbari, and Jennifer W. Vaughan. Adaptive task assignment for crowdsourced classification. In International Conference on Machine Learning, volume 28, pages 534–542, 2013

[10] Ittai Abraham, Omar Alonso, Vasilis Kandylas, and Aleksandrs Slivkins. Adaptive crowd-sourcing algorithms for the bandit survey problem. In Conference On Learning Theory, volume 30 of JMLR Proceedings, pages882–910. 2013

[11] Long Tran-Thanh, Archie C. Chapman, Alex Rogers, and Nicholas R. Jennings. Knapsack based optimal policies for budget-limited multi-armed bandits. In Twenty-Sixth Conference on Artificial Intelligence (AAAI 2012), 2012

Learning in Crowdsourcing

- A setting of assured quality per task is considered uniform cost workers [9]
- The goal of selecting a single optimal crowd for a single task is considered by aggregating the answers in a sequential way until a certain accuracy is achieved for each task with Homogeneous workers having same quality in a crowd [10]
- Efficient selection of a single worker for each task based on MAB algorithms by formulating a knapsack problem [11]

Note:

None of the above literature had costs as strategic parameter!

[9] Chien ju Ho, Shahin Jabbari, and Jennifer W. Vaughan. Adaptive task assignment for crowdsourced classification. In International Conference on Machine Learning, volume 28, pages 534–542, 2013

[10] Ittai Abraham, Omar Alonso, Vasilis Kandylas, and Aleksandrs Slivkins. Adaptive crowd-sourcing algorithms for the bandit survey problem. In Conference On Learning Theory, volume 30 of JMLR Proceedings, pages882–910. 2013

[11] Long Tran-Thanh, Archie C. Chapman, Alex Rogers, and Nicholas R. Jennings. Knapsack based optimal policies for budget-limited multi-armed bandits. In Twenty-Sixth Conference on Artificial Intelligence (AAAI 2012), 2012

Mechanism Design in Crowdsourcing

- MAB mechanism to determine an optimal pricing mechanism for a crowdsourcing problem within a specified budget (bandits with knapsack) [12]
- A posted price mechanism to elicit true costs from the users using MAB mechanisms while maintaining a budget constraint[13]

[12] Moshe Babaioff, Shaddin Dughmi, Robert Kleinberg, and Aleksandrs Slivkins. Dynamic pricing with limited supply. In Thirteenth ACM Conference on Electronic Commerce, pages 74–91. ACM, 2012

[13] Adish Singla and Andreas Krause. Truthful incentives in crowdsourcing tasks using regret minimization mechanisms. In Twenty Second International World Wide Web Conference, pages 1167–1178, 2013

[14] Satyanath Bhat, Swaprava Nath, Onno Xoeter, Sujit Gujar, Yadati Narahari, and Chris Dance. A mechanism to optimally balance cost and quality of labeling tasks outsourced to strategic agents. In Thirteenth International Conference on Autonomous Agents and Multiagent Systems, pages 917–924, 2014

Mechanism Design in Crowdsourcing

- MAB mechanism to determine an optimal pricing mechanism for a crowdsourcing problem within a specified budget (bandits with knapsack) [12]
- A posted price mechanism to elicit true costs from the users using MAB mechanisms while maintaining a budget constraint[13]

Note:

Homogeneous qualities are considered!

[12] Moshe Babaioff, Shaddin Dughmi, Robert Kleinberg, and Aleksandrs Slivkins. Dynamic pricing with limited supply. In Thirteenth ACM Conference on Electronic Commerce, pages 74–91. ACM, 2012

[13] Adish Singla and Andreas Krause. Truthful incentives in crowdsourcing tasks using regret minimization mechanisms. In Twenty Second International World Wide Web Conference, pages 1167–1178, 2013

[14] Satyanath Bhat, Swaprava Nath, Onno Xoeter, Sujit Gujar, Yadati Narahari, and Chris Dance. A mechanism to optimally balance cost and quality of labeling tasks outsourced to strategic agents. In Thirteenth International Conference on Autonomous Agents and Multiagent Systems, pages 917–924, 2014

Mechanism Design in Crowdsourcing

- MAB mechanism to determine an optimal pricing mechanism for a crowdsourcing problem within a specified budget (bandits with knapsack) [12]
- A posted price mechanism to elicit true costs from the users using MAB mechanisms while maintaining a budget constraint[13]

Note:

Homogeneous qualities are considered!

- Work by Bhat et al. considers cost of the workers to be public and qualities to be private strategic quantity of the workers [14]

[12] Moshe Babaioff, Shaddin Dughmi, Robert Kleinberg, and Aleksandrs Slivkins. Dynamic pricing with limited supply. In Thirteenth ACM Conference on Electronic Commerce, pages 74–91. ACM, 2012

[13] Adish Singla and Andreas Krause. Truthful incentives in crowdsourcing tasks using regret minimization mechanisms. In Twenty Second International World Wide Web Conference, pages 1167–1178, 2013

[14] Satyanath Bhat, Swaprava Nath, Onno Xoeter, Sujit Gujar, Yadati Narahari, and Chris Dance. A mechanism to optimally balance cost and quality of labeling tasks outsourced to strategic agents. In Thirteenth International Conference on Autonomous Agents and Multiagent Systems, pages 917–924, 2014

MAB Algorithms

- A recent survey by Bubeck and Cesa-Bianchi compiles various variations on stochastic and non-stochastic MAB problem [15]
- A general bandit problem with concave rewards and convex constraints is solved [16]
- The probably Approximately Correct (PAC) learning framework for single pull and multiple pull MAB is considered by Even Dar et al. [17] and by Kalyanakrishnan et al. [18] respectively

[15] Sebastien Bubeck and Nicolo Cesa-Bianchi. Regret analysis of stochastic and nonstochastic multi-armed bandit problems. Foundations and Trends in Machine Learning, 5(1):1–122, 2012

[16] Shipra Agrawal and Nikhil R. Devanur. Bandits with concave rewards and convex knap- sacks. In Fifteenth ACM Conference on Economics and Computation, To appear, 2014

[17] Eyal Even-Dar, Shie Mannor, and Yishay Mansour. Action elimination and stopping conditions for the multi-armed bandit and reinforcement learning problems. Journal of Machine Learning, 7:1079–1105, 2006

[18] Shivaram Kalyanakrishnan and Peter Stone. Efficient selection of multiple bandit arms: Theory and practice. In International Conference on Machine Learning, 2010

MAB Mechanisms

- Characterization of truthful single pull MAB mechanism in forward setting [19] [20]
- Extension to multiple pull MAB mechanism in forward setting [21] [22]
- Algorithms with improved regret bounds [23] [24]

---

[19] Moshe Babaioff, Yogeshwer Sharma, and Aleksandrs Slivkins. Characterizing truthful multi-armed bandit mechanisms: extended abstract. In Tenth ACM Conference on Electronic Commerce, pages 79–88. ACM, 2009

[20] Nikhil R. Devanur and Sham M. Kakade. The price of truthfulness for pay-per-click auctions. In Tenth ACM Conference on Electronic Commerce, pages 99–106, 2009

[21] Nicola Gatti, Alessandro Lazaric, and Francesco Trov'o. A truthful learning mechanism for contextual multi-slot sponsored search auctions with externalities. In Thirteenth ACM Conference on Electronic Commerce, pages 605–622, 2012

[22] Akash Das Sharma, Sujit Gujar, and Y. Narahari. Truthful multi-armed bandit mechanisms for multi-slot sponsored search auctions. Current Science, Vol. 103 Issue 9, 2012

[23] Moshe Babaioff, Robert D. Kleinberg, and Aleksandrs Slivkins. Truthful mechanisms with implicit payment computation. In Eleventh ACM Conference on Electronic Commerce, pages 43–52. ACM, 2010

[24] Debmalya Mandal and Y. Narahari. A Novel Ex-Post Truthful Mechanism for Multi-Slot Sponsored Search Auctions. Pages 1555-1556, AAMAS, 2014

MAB Mechanisms

- Characterization of truthful single pull MAB mechanism in forward setting [19] [20]
- Extension to multiple pull MAB mechanism in forward setting [21] [22]
- Algorithms with improved regret bounds [23] [24]

Note:

Only forward setting is considered.

[19] Moshe Babaioff, Yogeshwer Sharma, and Aleksandrs Slivkins. Characterizing truthful multi-armed bandit mechanisms: extended abstract. In Tenth ACM Conference on Electronic Commerce, pages 79–88. ACM, 2009

[20] Nikhil R. Devanur and Sham M. Kakade. The price of truthfulness for pay-per-click auctions. In Tenth ACM Conference on Electronic Commerce, pages 99–106, 2009

[21] Nicola Gatti, Alessandro Lazaric, and Francesco Trov'o. A truthful learning mechanism for contextual multi-slot sponsored search auctions with externalities. In Thirteenth ACM Conference on Electronic Commerce, pages 605–622, 2012

[22] Akash Das Sharma, Sujit Gujar, and Y. Narahari. Truthful multi-armed bandit mechanisms for multi-slot sponsored search auctions. Current Science, Vol. 103 Issue 9, 2012

[23] Moshe Babaioff, Robert D. Kleinberg, and Aleksandrs Slivkins. Truthful mechanisms with implicit payment computation. In Eleventh ACM Conference on Electronic Commerce, pages 43–52. ACM, 2010

[24] Debmalya Mandal and Y. Narahari. A Novel Ex-Post Truthful Mechanism for Multi-Slot Sponsored Search Auctions. Pages 1555-1556, AAMAS, 2014

Research Gaps

- Need true costs for optimal selection of workers
- Need to ensure the target accuracy that depends on unknown qualities

× No previous work on learning qualities and eliciting true costs in crowd-sourcing environment

- This calls for developing a new approach for MAB