An Online Learning Approach to Improving the Quality of Crowd-Sourcing

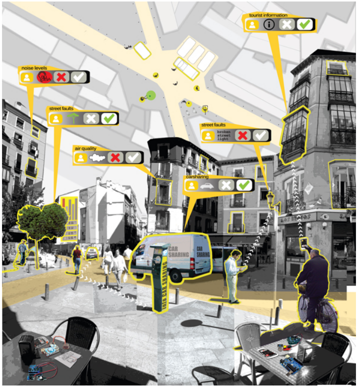
Yang Liu
(Joint work with Mingyan Liu)

Department of Electrical Engineering and Computer Science
University of Michigan, Ann Arbor, MI

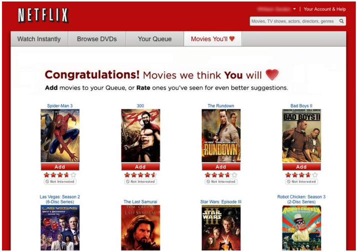
June 2015

The power of crowdsourcing

Data collection: participatory sensing, user-generated map :



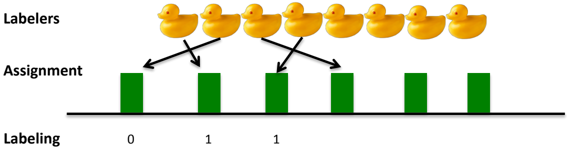
Recommendation: rating of movies, news, restaurants, services:



Social studies: opinion survey, the science of opinion survey:

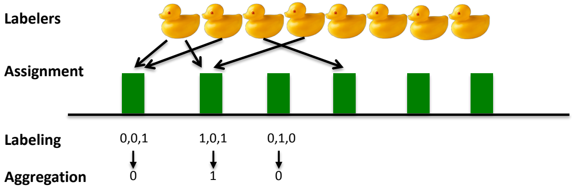
Crowd-sourcing market

Data processing: image labeling, annotation



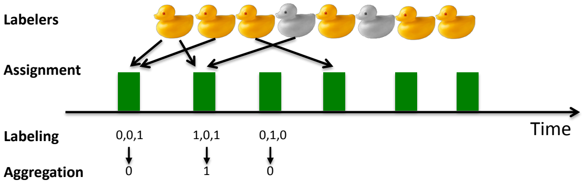
- Paid workers perform computational tasks.
- Hard to measure and evaluate quality objective: competence, bias, irresponsible behavior, etc.

One step further: Use the wisdom of crowd



- Redundant assignment.
- Label aggregation.

Our vision:



- Labeler selection. (first step)
- Adaptive learning. (second step)
- A weighted version of selection. (one more step)

Our objective

To make the most effective use of the crowdsourcing system

- Cost in having large amount of data labeled is non-trivial
- Time constrained machine learning tasks.

A sequential/online learning framework

- Over time learn which labelers are more competent, or whose reviews/opinion should be valued more.
- Quality control rather than random assignment
- Closed-loop, causal.

Multiarmed bandit (MAB) framework

A sequential decision making and learning framework:

- Objective: select the best of a set of choices ("arms")
- Principle: repeated sampling of different choices ("exploration"), while controlling how often each choice is used based on their empirical quality ("exploitation").
- Performance measure: "regret" – difference between an algorithm and a benchmark.

Challenges and key ideas

Main challenge in crowdsourcing: ground truth

- True label of data remains unknown
- If view each labeler as a choice/arm: unknown quality of outcome ("reward").

Key features:

- Mild assumption on the collective quality of the crowd; quality of an individual is estimated against the crowd.
- Online learning: Learning occurs as data/labeling tasks arrive.
- Comparing against optimal static selections.

Outline of the talk

Problem formulation

Online solution

- Simple/weighted majority voting

Extensions and discussions

Experiment results

- Numerical results & Results on AMT data

Conclusion and on-going works

Labeler selection

M labelers; labelers i has accuracy p_i (can be task-dependent).

- No two exactly the same: $p_i \neq p_j$ for $i \neq j$, and $0 < p_i < 1$, $\forall i$.
- Collective quality: $\bar{p} := \sum_i p_i / M > 1/2$.
- Probability that a simple majority vote over all M labelers is correct: $a_{\min} := P(\sum_i X_i / M > 1/2)$.
  - If $\bar{p} > 1/2$ and $M > \frac{\log 2}{\bar{p} - 1/2}$, then $a_{\min} > 1/2$.

Unlabeled tasks arrive at $t = 1, 2, \cdots$.

- User selects a subset S_t of labelers for task at t.
- Labeling payment of c_i for each task performed by labeler i.

Labeler selection

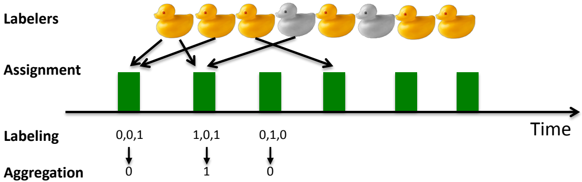
M labelers; labelers i has accuracy p_i (can be task-dependent).

- No two exactly the same: $p_i \neq p_j$ for $i \neq j$, and $0 < p_i < 1$, $\forall i$.
- Collective quality: $\bar{p} := \sum_i p_i / M > 1/2$.
- Probability that a simple majority vote over all M labelers is correct: $a_{\min} := P(\sum_i X_i / M > 1/2)$.
  - If $\bar{p} > 1/2$ and $M > \frac{\log 2}{\bar{p} - 1/2}$, then $a_{\min} > 1/2$.

Unlabeled tasks arrive at $t = 1, 2, \cdots$.

- User selects a subset S_t of labelers for task at t.
- Labeling payment of c_i for each task performed by labeler i.

Labeling outcome/Information aggregation



Aggregating results from multiple labelers:

- A task receives a set of labels: $\{L_i(t)\}_{i \in S_t}$.
- Use simple majority voting to compute the label output: $L^*(t)$. (extensible to weighted majority voting)

Probability of correct labeling outcome: $\pi(S_t)$.

- Optimal set of labelers: S^* that maximizes $\pi(S)$.

$$\pi(S_t) = \underbrace{\sum_{S:S\subseteq S_t, |S|\geq\lceil\frac{|S_t|+1}{2}\rceil} \prod_{i\in S} p_i \cdot \prod_{j\in S_t\setminus S}(1-p_j)}_{\text{Majority wins}}$$
$$+ \underbrace{\frac{\sum_{S:S\subseteq S_t, |S|=\frac{|S_t|}{2}} \prod_{i\in S} p_i \cdot \prod_{j\in S_t\setminus S}(1-p_j)}{2}}_{\text{Ties broken equally likely}} .$$

Assuming known $\{p_i\}$, S^* can be obtained using a linear search

Theorem

Under the simple majority voting rule, $|S^*|$ is an odd number. Furthermore, S^* is monotonic: if $i \in S^*$ and $j \notin S^*$, then we must have $p_i > p_j$.

Assuming known $\{p_i\}$, S^* can be obtained using a linear search

Theorem

Under the simple majority voting rule, $|S^*|$ is an odd number. Furthermore, S^* is monotonic: if $i \in S^*$ and $j \notin S^*$, then we must have $p_i > p_j$.

Objective of our online solution

Labeler expertise p_is being unknown a priori

- Goal: Gradually learn labelers' quality and make selections adaptively

Performance measure:

- Comparing with the optimal selection (static):

$$R(T) = T\pi(S^*) - E[\sum_{t=1}^{T} \pi(S_t)]$$

Outline of the talk

Problem formulation

Online solution

- Simple/weighted majority voting

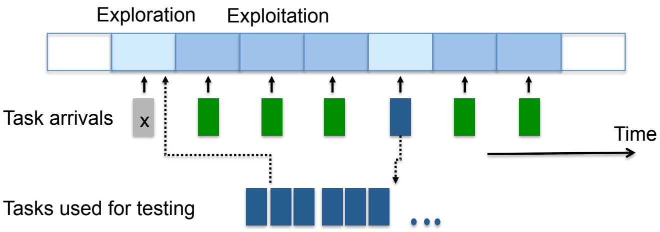
Extensions and discussions

Experiment results

- Numerical results & Results on AMT data

Conclusion and on-going works

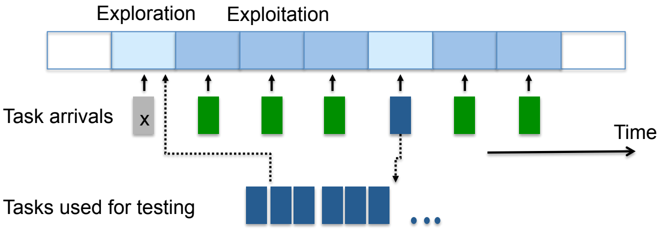
An online learning algorithm



Interlveaing explorations of MAB and Crowd-sourcing: Double exploration

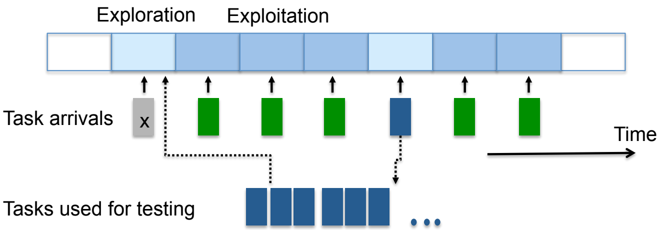
- There is a set of tasks $E(t)$ ($\sim \log t$) used for *testing* purposes.
- These or their independent and identical variants are repeatedly assigned to the labelers ($\sim \log t$). [1]

[1] More discussions follow later on independence.



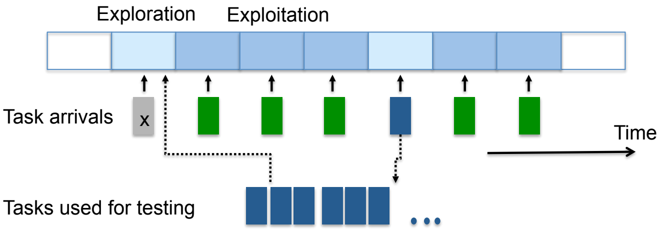
Two types of time steps:

- *Double exploration*: all M labelers are used. Exploration is entered if (1) the number of testers falls below a threshold ($\sim \log t$), or if (2) the number of times a tester has been tested falls below a threshold ($\sim \log t$).
- *Exploitation*: the estimated $\tilde{S}^*$ is used to label the arriving task based on the current estimated $\{\tilde{p}_i\}$.



Three types of tasks:

- Testers: those arriving to find (1) true and (2) false. These are added to $E(t)$ and are repeatedly used to collect independent labels whenever (2) is true subsequently.
- Throw-aways: those arriving to find (2) true. These are given a random label.
- Keepers: those arriving to find both (1) and (2) false. These are given a label outcome using the best estimated set of labelers.



Accuracy update

- Estimated label on tester k at time t: majority label over all test outcomes up to time t.
- $\tilde{p}_i$ at time t: the % of times i's label matches the majority vote known at t out of all tests on all testers.

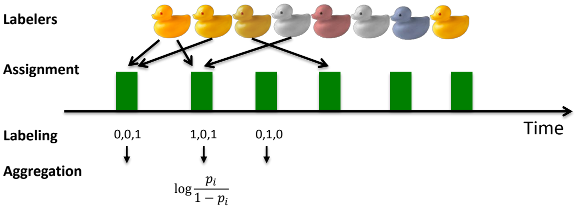
Regret

Main result:

$$R(T) \quad \leq \quad \text{Const}(S^*, \Delta_{\text{max}}, \Delta_{\text{min}}, \delta_{\text{max}}, \delta_{\text{min}}, a_{\text{min}}) \log^2(T) + \text{Const}$$

- $\Delta_{\text{max}} = \max_{S \neq S^*} \pi(S^*) - \pi(S)$, $\delta_{\text{max}} = \max_{i \neq j} |p_i - p_j|$.
- First term due to exploration; second due to exploitation.
- Can obtain similar result on the cost $C(T)$.

Weighted majority voting



- Each labeler i's decision is weighed by $\log \frac{p_i}{1-p_i}$.

Theorem

Under the weighted majority vote and assuming $p_i \geq 0.5, \forall i$, the optimal set S^* is monotonic, i.e., if we have $i \in S^*$ and $j \notin S^*$ then we must have $p_i > p_j$.

Main results on weighted majority voting:

- $R(T) \leq O(\log^2 T)$, but with strictly larger constants.
- Have to account for additional error in estimating the weights when determining label outcome.
- A larger constant: slower convergence to a better target.

Outline of the talk

Problem formulation

Online solution

- Simple/weighted majority voting

Extensions and discussions

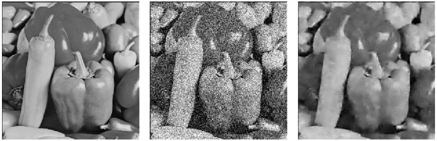
Experiment results

- Numerical results & Results on AMT data

Conclusion and on-going works

Re-assignment of the testers

IID noise insertion



Random delay

- Commonly adopted in survey methodology to ensure valid responses.
- Bounded delay $\tau_{\max}$ leads to $\tau_{\max}$-fold regret.

Other extensions

Prior knowledge on several constants

- Exploration length depends on several system parameters.
- Sub-logarithmic remedy without knowing prior knowledge.

Improve the bound by improving $a_{\min}$: weed out bad labelers.

- Ranking based on counting of disagreement.
- Start the weeding out from the end of list.
- Requires only $O(\log T)$ samples to achieve a bounded regret.

Other extensions cont.

Labelers with different type of tasks

- Finite number of types: Similar results.
- Infinite number of types.
  - Continuous MAB.
  - Sub-linear regret bound.

With delayed arrival of ground-truth

- $O(\log T)$ regret time uniformly.

Outline of the talk

Problem formulation

Online solution

- Simple/weighted majority voting

Extensions and discussions

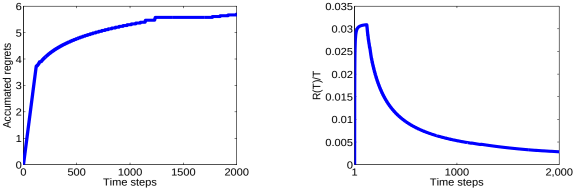
Experiment results

- Numerical results & Results on AMT data

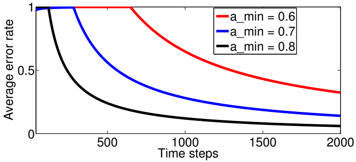
Conclusion and on-going works

Experiment I: simulation with $M = 5$

Accumulative regret & Average regret $R(T)/T$

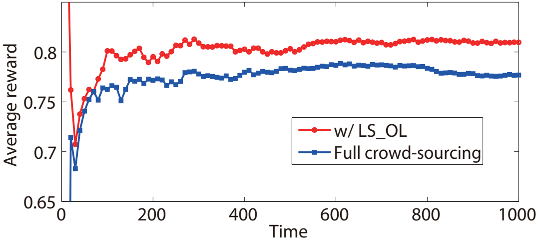


Effect of $a_{\min}$: higher $a_{\min}$ leads to much better performance.

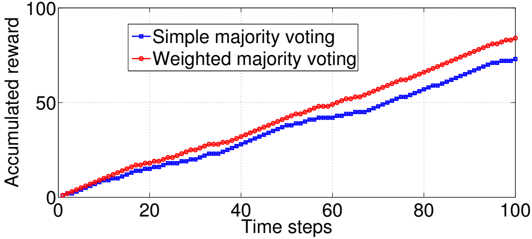


Performance comparison

labeler selection v.s. full crowd-sourcing (simple majority vote)



Comparing weighted and simple majority vote



M	5	10	15	20
Full crowd-sourcing (majority vote)	0.5154	0.5686	0.7000	0.7997
Majority vote w/ LS_OL	0.8320	0.9186	0.9434	0.9820
Weighted majority vote w/ LS_OL	**0.8726**	**0.9393**	**0.9641**	**0.9890**

Table: Average reward per labeler: there is a clear gap between with and without using LS_OL.

Experiment II: on a real AMT dataset

The dataset

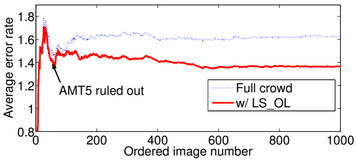
- Contains 1,000 images each labeled by the same set of 5 AMTs.
- Labels are on a scale from 0 to 5, indicating how many scenes are seen from each image.
- A second dataset summarizing keywords for scenes of each image: use this count as the ground truth.

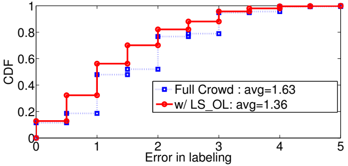
Counting number of disagreement (online):

	AMT1	AMT2	AMT3	AMT4	AMT5
# of disagree	348	353	376	338	441

Table: Total number of disagreement each AMT has

Performance comparison





(L) AMT 5 was quickly weeded out; eventually settled on the optimal set of AMTs 1, 2, and 4 for most of the time.

(R) CDF of all images' labeling error at the end of this process.

Conclusion

We discussed a quality control problem in labeler market

- How to select the best set of labelers over a sequence of tasks.
  - An algorithm that estimates labeler's quality by comparing against (weighted) majority vote; new regret bound.

Currently under investigation

- Lower bound on the regret in the labeler selection problem.
- Hypothesis testing & coupling argument.

Q & A

Thank you. Any question?

http://www.umich.edu/~youngliu