Optimistic Knowledge Gradient Policy for Optimal Budget Allocation in Crowdsourcing

Presentation:

Kamran Ghasedi

Spring 2015

www.fppt.info

Outline

- Crowdsourcing
- Problem definition
- K-coin tossing problem
- Markov Decision Process (MDP) and Optimal Policy
- Bayesian Setup
- Accuracy Maximization
- Optimization problem
- Stage-wise expected reward
- Optimistic Knowledge Gradient
- Experiment on simulated data
- Experiment on real dataset
- Conclusion

www.fppt.info

Paper

Optimistic Knowledge Gradient Policy for Optimal Budget Allocation in Crowdsourcing

Xi Chen XICHEN@CS.CMU.EDU
Machine Learning Department, Carnegie Mellon University, Pittsburgh, PA 15213, USA

Qihang Lin QIHANGL@ANDREW.CMU.EDU
Tepper School of Business, Carnegie Mellon University, Pittsburgh, PA 15213, USA

Dengyong Zhou DENGYONG.ZHOU@MICROSOFT.COM
Machine Learning Group, Microsoft Research, Redmond, WA 98052, USA

Abstract

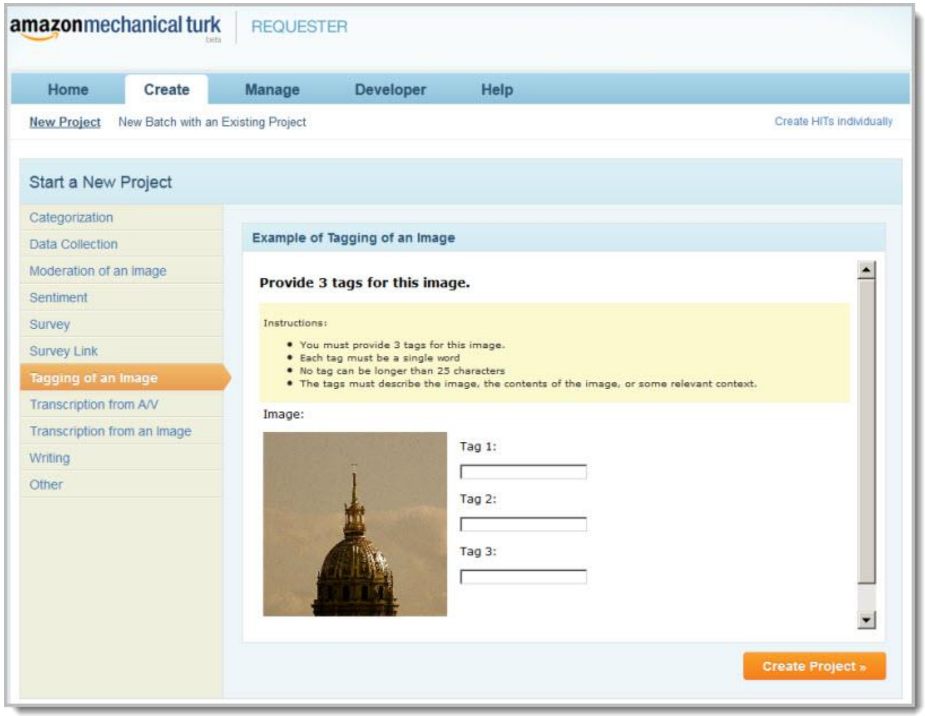
We consider the budget allocation problem in binary/multi-class crowd labeling where each label from the crowd has a certain cost. Since different instances have different ambiguities and different workers have different reliabilities, a fundamental challenge is how to allocate a pre-fixed amount of budget among instance-worker pairs so that the overall accuracy can be maximized. We start with a simple setting where all workers

ing good classifiers. Due to the flourish of many online crowdsourcing services (e.g., Amazon Mechanical Turk), an effective way of collecting training labels is to ask a crowd of workers to label training instances. Since low-paid class labels provided by the crowd could be very noisy, each instance has to be labeled several times by different workers so that we have a higher chance to infer its true class correctly after aggregating these labels. Each labeling from the crowd usually has a certain cost (e.g., 10 cents). Given the limited amount of budget, it is important to wisely allocate

www.fppt.info

Crowdsourcing

- Need training labels for building good classifiers
- Crowdsourcing services (e.g., Amazon Mechanical Turk)
- Collecting training labels from crowd of workers to label training instances
- Noisy labels
- Inferring true labels



Problem definition

Given the limited amount of budget, it is important to wisely allocate the budget among instances so that the overall accuracy is maximized.

1. How to accurately estimate the labeling difficulty for each instance ?
2. Whether to spend more budget on ambiguous instances ?
3. How to estimate the reliability of workers?

www.fppt.info

K-coin tossing problem

- Binary labeling task
- workers are identical and noiseless (perfectly reliable)

- We are allowed to sequentially specify a coin to toss.
- Then observe the outcome of the toss.
- We note that each coin can be chosen multiple times.
- After the coin toss budget T runs out, we decide whether a coin is biased more towards the head or the tail for each coin.

www.fppt.info

Markov Decision Process (MDP) and Optimal Policy

- *K* instances
- True label $Z_i \in \{-1, 1\}$ for $1 \leq i \leq K$.
- Positive set $H^* = \{i : Z_i = 1\} = \{i : \theta_i \geq 0.5\}$
- Labeling difficulty of each instance $\theta_i \in [0, 1]$
- Total budget T $0 \leq t \leq T-1$
- Action set $i_t \in \mathcal{A} = \{1, \ldots, K\}$
- Predicted label y_{i_t}
- Optimal allocation sequence $(i_0, \ldots, i_{T-1})$

Illustration Example

θ_1	1	1	
θ_2	1	-1	?
θ_3	1	?	

Table 1. Labeling matrix

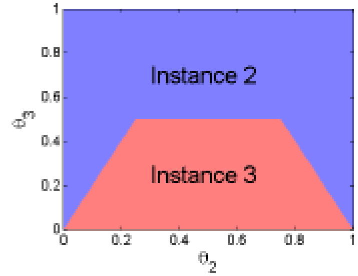


Figure 1. Decision Boundary

Blue region:

$$\max(\theta_2 - 0.5, 0.5 - \theta_2) > 0.5(1 - \theta_3) \text{ or } \theta_3 > 0.5$$

Acc	Cur.	$y = 1$	$y = -1$	Expected Acc	Improvement
$\theta_1 > 0.5$	1	1	1	1	0
$\theta_1 < 0.5$	0	0	0	0	0
$\theta_2 > 0.5$	0.5	1	0	θ_2	$\theta_2 - 0.5 > 0$
$\theta_2 < 0.5$	0.5	0	1	$1 - \theta_2$	$0.5 - \theta_2 > 0$
$\theta_3 > 0.5$	1	1	0.5	$\theta_3 + 0.5(1 - \theta_3)$	$0.5(\theta_3 - 1) < 0$
$\theta_3 < 0.5$	0	0	0.5	$0.5(1 - \theta_3)$	$0.5(1 - \theta_3) > 0$

Table 2. Calculation of the expected improvement. The 2nd column is the current accuracy. The 3rd and 4th are accuracies if the next label is 1 and -1.

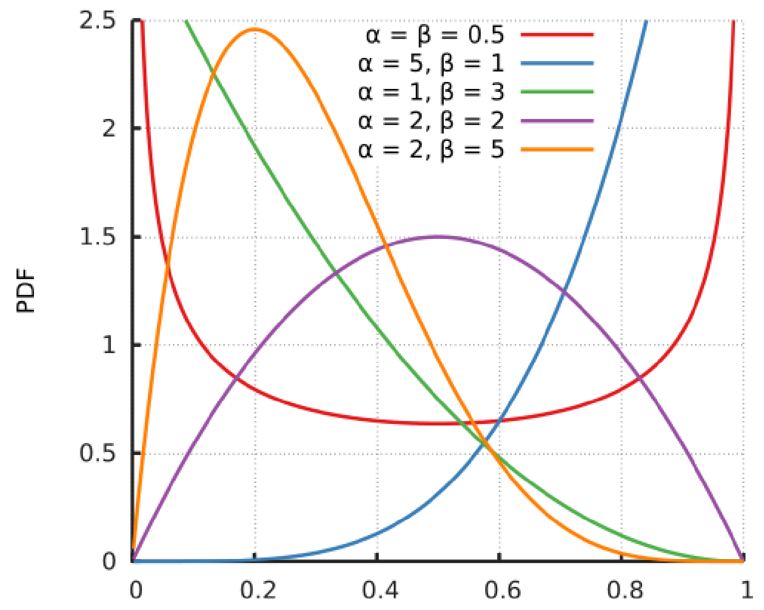
www.fppt.info

Bayesian Setup

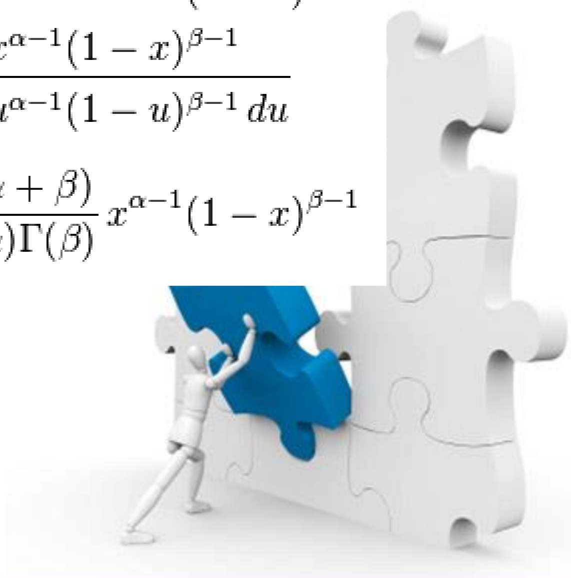
$$y_{i_t} \sim \text{Bernoulli}(\theta_{i_t})$$

Beta is the conjugate prior of the Bernoulli:

θ_i is drawn from a known Beta prior distribution $\text{Beta}(a_i^0, b_i^0)$



$$\begin{aligned} f(x;\alpha,\beta) &= \text{constant} \cdot x^{\alpha-1}(1-x)^{\beta-1} \\ &= \frac{x^{\alpha-1}(1-x)^{\beta-1}}{\int_0^1 u^{\alpha-1}(1-u)^{\beta-1}\,du} \\ &= \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}\,x^{\alpha-1}(1-x)^{\beta-1} \end{aligned}$$



www.fppt.info

Bayesian Setup

$$y_{i_t} \sim \text{Bernoulli}(\theta_{i_t})$$

Beta is the conjugate prior of the Bernoulli:

θ_i is drawn from a known Beta prior distribution $\text{Beta}(a_i^0, b_i^0)$

$$\begin{cases} \text{Beta}(a_{i_t}^{t+1}, b_{i_t}^{t+1}) = \text{Beta}(a_{i_t}^t + 1, b_{i_t}^t) \text{ if } y_{i_t} = 1 \\ \text{Beta}(a_{i_t}^{t+1}, b_{i_t}^{t+1}) = \text{Beta}(a_{i_t}^t, b_{i_t}^t + 1) \text{ if } y_{i_t} = -1 \end{cases}$$

$$S^t = \{a_i^t, b_i^t\}_{i=1}^K \qquad K \times 2 \text{ matrix}$$

$$S^{t+1} = \begin{cases} S^t + (\mathbf{e}_{i_t}, \mathbf{0}) & \text{if } y_{i_t} = 1; \\ S^t + (\mathbf{0}, \mathbf{e}_{i_t}) & \text{if } y_{i_t} = -1, \end{cases}$$

$$\Pr(y_{i_t} = 1 | S^t, i_t) = \mathbb{E}(\theta_{i_t} | S^t) = \frac{a_{i_t}^t}{a_{i_t}^t + b_{i_t}^t},$$

$$\Pr(y_{i_t} = -1 | S^t, i_t) = 1 - \Pr(y_{i_t} = 1 | S^t, i_t)$$

www.fppt.info

Accuracy Maximization

$$\{\mathcal{F}_t\}_{t=0}^{T} \longleftarrow (i_0, y_{i_0}, \ldots, i_{t-1}, y_{i_{t-1}})$$

$$H_T = \underset{H \subset \{1,\ldots,K\}}{\arg\max} \, \mathbb{E}\left(\sum_{i \in H} \mathbf{1}(i \in H^*) + \sum_{i \notin H} \mathbf{1}(i \notin H^*) \middle| \mathcal{F}_T\right)$$

Conditional distribution $\theta_i | \mathcal{F}_t$ $\longrightarrow$ Posterior distribution $\mathrm{Beta}(a_i^t, b_i^t)$

$$I(a, b) = \Pr(\theta \geq 0.5 | \theta \sim \mathrm{Beta}(a, b)),$$
$$P_i^t = \Pr(i \in H^* | \mathcal{F}_t) = \Pr(\theta \geq 0.5 | S_i^t) = I(a_i^t, b_i^t),$$

Proposition 2.1 $H_T = \{i : P_i^T \geq 0.5\}$ *solves* (3) *and the expected accuracy on RHS of* (3) *can be written as* $\sum_{i=1}^{K} h(P_i^T)$, *where* $h(x) = \max(x, 1 - x)$.

Corollary 2.2 $I(a, b) > 0.5$ *if and only if* $a > b$ *and* $I(a, b) = 0.5$ *if and only if* $a = b$. *Therefore,* $H_T = \{i : a_i^T \geq b_i^T\}$ *solves* (3).

www.fppt.info

Optimization problem

$\{\mathcal{F}_t\}_{t=0}^{T} \longleftarrow (i_0, y_{i_0}, \ldots, i_{t-1}, y_{i_{t-1}})$

$$H_T = \underset{H \subset \{1,\ldots,K\}}{\arg\max}\ \mathbb{E}\left(\sum_{i\in H} \mathbf{1}(i \in H^*) + \sum_{i \notin H} \mathbf{1}(i \notin H^*) \middle| \mathcal{F}_T\right)$$

Conditional distribution $\theta_i|\mathcal{F}_t \longrightarrow$ Posterior distribution $\text{Beta}(a_i^t, b_i^t)$

$$I(a,b) = \Pr(\theta \geq 0.5 | \theta \sim \text{Beta}(a,b)),$$
$$P_i^t = \Pr(i \in H^* | \mathcal{F}_t) = \Pr(\theta \geq 0.5 | S_i^t) = I(a_i^t, b_i^t),$$

Proposition 2.1 $H_T = \{i : P_i^T \geq 0.5\}$ *solves* (3) *and the expected accuracy on RHS of* (3) *can be written as* $\sum_{i=1}^{K} h(P_i^T)$, *where* $h(x) = \max(x, 1-x)$.

$$V(S^0) \doteq \sup_\pi \mathbb{E}^\pi\left[\mathbb{E}\left(\sum_{i\in H_T} \mathbf{1}(i \in H^*) + \sum_{i \notin H_T} \mathbf{1}(i \notin H^*) \middle| \mathcal{F}_T\right)\right]$$
$$= \sup_\pi \mathbb{E}^\pi\left(\sum_{i=1}^{K} h(P_i^T)\right), \qquad (6)$$

www.fppt.info

Stage-wise expected reward

Proposition 2.3 *Define the stage-wise expected reward as:*

$$R(S^t, i_t) = \mathbb{E}\left(h(P_{i_t}^{t+1}) - h(P_{i_t}^t)|S^t, i_t\right), \qquad (7)$$

$$V(S^0) \doteq \sup_{\pi} \mathbb{E}^{\pi}\left[\mathbb{E}\left(\sum_{i\in H_T} \mathbf{1}(i \in H^*) + \sum_{i\notin H_T} \mathbf{1}(i \notin H^*)\Big|\mathcal{F}_T\right)\right]$$
$$= \sup_{\pi} \mathbb{E}^{\pi}\left(\sum_{i=1}^{K} h(P_i^T)\right), \qquad (6)$$

$$V(S^0) = G_0(S^0) + \sup_{\pi} \mathbb{E}^{\pi}\left(\sum_{t=0}^{T-1} R(S^t, i_t)\right), \qquad (8)$$

where $G_0(S^0) = \sum_{i=1}^{K} h(P_i^0)$

With Proposition 2.3, the maximization problem (6) is formulated as a T-stage MDP (8).

www.fppt.info

Stage-wise expected reward

$$V(S^0) = G_0(S^0) + \sup_\pi \mathbb{E}^\pi \left(\sum_{t=0}^{T-1} R(S^t, i_t) \right), \qquad (8)$$

since $S_{i_t}^t = (a_{i_t}^t, b_{i_t}^t) \in \mathbb{R}_+^2 \longrightarrow R(a_{i_t}^t, b_{i_t}^t) = R(S^t, i_t)$

$$R_1(a,b) = h(I(a+1,b)) - h(I(a,b)),$$
$$R_2(a,b) = h(I(a,b+1)) - h(I(a,b)).$$

Expected reward:

$$R(a,b) = p_1 R_1 + p_2 R_2$$

$$p_1 = \frac{a}{a+b} \qquad p_2 = \frac{b}{a+b}$$

Compute exact optimal policy

Dynamic programming (DP) algorithm
(Puterman, 2005; Powell, 2007) (a.k.a. backward induction)

computation is intractable,
since size of the state space grows exponentially in t

need some computationally efficient approximate policies

www.fppt.info

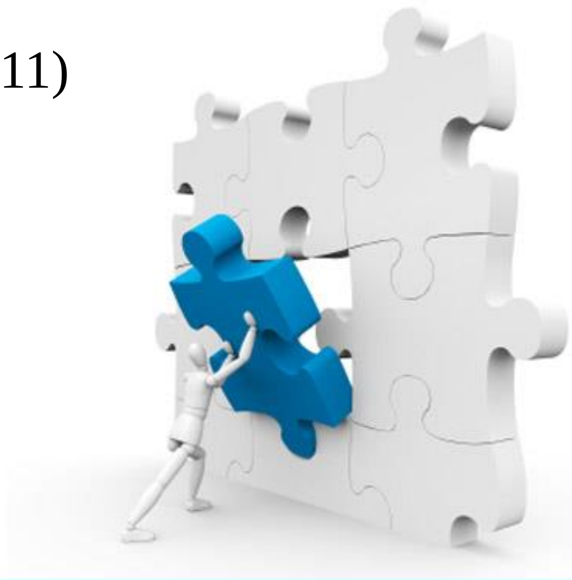
Compute approximate optimal policy

Problem: finite-horizon Bayesian multi-armed bandit (MAB)

- calibration method (Gittins, 1989; Nino-Mora, 2011)
  O(T^3) time and space complexity

- the state-of-the-art exact method (Nino-Mora, 2011)
  O(T^6) time and space complexity

Knowledge gradient (KG) method:



www.fppt.info

Knowledge Gradient

Knowledge Gradient:

$$i_t = \arg\max_i \left(R(a_i^t, b_i^t) \doteq \frac{a_i^t}{a_i^t + b_i^t} R_1(a_i^t, b_i^t) + \frac{b_i^t}{a_i^t + b_i^t} R_2(a_i^t, b_i^t) \right). \tag{12}$$

Deterministic KG is NOT a consistent policy,

where the consistent policy refers to the policy that will achieve 100% accuracy almost surely when T goes to infinity.

randomized KG policy's empirical performance is undesirable.

Conditional Value-at-Risk (CVaR)

$$\mathrm{CVaR}_\alpha(X) = \max_{\{q_1 \geq 0, q_2 \geq 0\}} q_1 R_1 + q_2 R_2,$$

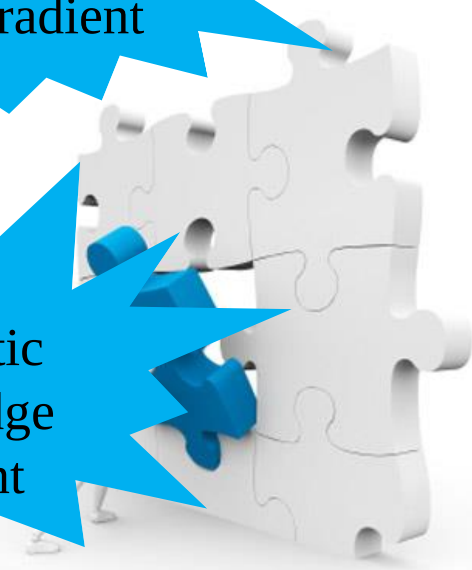
$$\text{s.t.} \quad q_1 \leq \frac{1}{\alpha} p_1, \; q_2 \leq \frac{1}{\alpha} p_2, \; q_1 + q_2 = 1.$$

when $\alpha = 1 \longrightarrow \mathrm{CVaR}_\alpha(X) = p_1 R_1 + p_2 R_2$

Knowledge Gradient

when $\alpha \to 0 \longrightarrow \mathrm{CVaR}_\alpha(X) = \max(R_1, R_2)$

Optimistic Knowledge Gradient



Optimistic Knowledge Gradient

The stage-wise reward $R(a, b)$ can be viewed as a random variable with a two point distribution:

$$\begin{cases} \text{probability } p_1 = \frac{a}{a+b} \text{ of being } R_1(a,b) \\ \text{probability } p_2 = \frac{b}{a+b} \text{ of being } R_2(a,b) \end{cases}$$

A simple idea is to select the instance based on the optimistic outcome of the reward, i.e., the instance with the largest:

$$R^+(a,b) = \max(R_1(a,b), R_2(a,b)).$$

Time complexity: $\mathcal{O}(KT)$
Space complexity: $\mathcal{O}(K)$

www.fppt.info

Optimistic Knowledge Gradient

Algorithm 1 Optimistic Knowledge Gradient

Input: Parameters of prior distributions for instances $\{a_i^0, b_i^0\}_{i=1}^K$ and the total budget T.

for $t = 0, \ldots, T-1$ **do**

Select the next instance i_t to label according to:

$$i_t = \arg\max_{i \in \{1,\ldots,K\}} \left(R^+(a_i^t, b_i^t) \doteq \max(R_1(a_i^t, b_i^t), R_2(a_i^t, b_i^t)) \right).$$

Acquire the label $y_{i_t} \in \{-1, 1\}$.

if $y_{i_t} = 1$ **then**

$a_{i_t}^{t+1} = a_{i_t}^t + 1, b_{i_t}^{t+1} = b_{i_t}^t;\ a_i^{t+1} = a_i^t, b_i^{t+1} = b_i^t$ for all $i \neq i_t$.

else

$a_{i_t}^{t+1} = a_{i_t}^t, b_{i_t}^{t+1} = b_{i_t}^t + 1;\ a_i^{t+1} = a_i^t, b_i^{t+1} = b_i^t$ for all $i \neq i_t$.

end if

end for

Output: The positive set $H_T = \{i : a_i^T \geq b_i^T\}$.

www.fppt.info

Incorporate Workers' Reliability

M workers

reliability of the *j-th* worker $\rho_j \in [0, 1]$

$$\rho_j = \Pr(Z_{ij} = Y_i | Y_i)$$

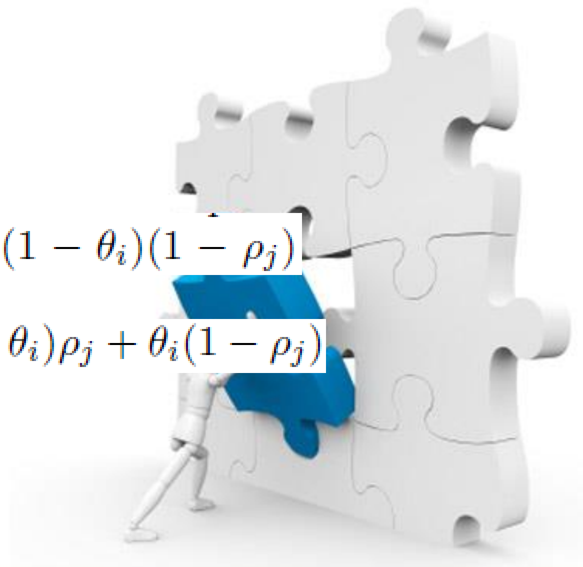
$$\begin{aligned}\Pr(Z_{ij} = 1) =& \Pr(Z_{ij} = 1 | Y_i = 1) \Pr(Y_i = 1) + \\ & \Pr(Z_{ij} = 1 | Y_i = -1) \Pr(Y_i = -1) \\ =& \rho_j \theta_i + (1 - \rho_j)(1 - \theta_i). \qquad (13)\end{aligned}$$

If we assume: $\rho_j \sim \mathrm{Beta}(c_j^0, d_j^0)$

When we observe label 1: $\Pr(Z_{ij} = 1 | \theta_i, \rho_j) = \theta_i \rho_j + (1 - \theta_i)(1 - \rho_j)$

When we observe label -1: $\Pr(Z_{ij} = -1 | \theta_i, \rho_j) = (1 - \theta_i)\rho_j + \theta_i(1 - \rho_j)$

$$p(\theta_i, \rho_j | Z_{ij} = z) \approx p(\theta_i | Z_{ij} = z) p(\rho_j | Z_{ij} = z)$$



www.fppt.info

Extensions

- Incorporate feature information of instances
- Address multi-class labeling problems

www.fppt.info

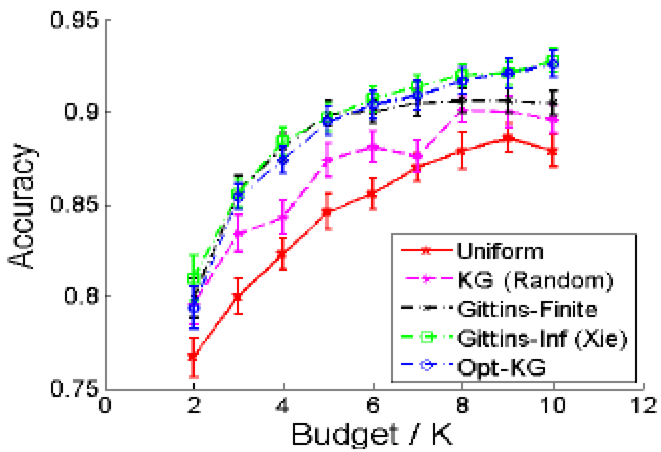
Experiment on simulated data

$K = 50$

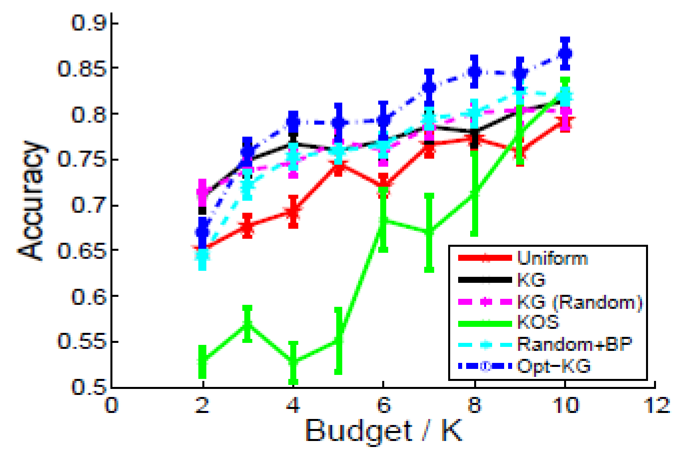
generate 20 different sets of $\{\theta_i\}_{i=1}^{K}$.

total budget $T = 2K, 3K, \ldots, 10K$

workers' reliability $\rho_j \sim \text{Beta}(4,1)$
$j = 1, \ldots, 10$



(a) Majority Vote



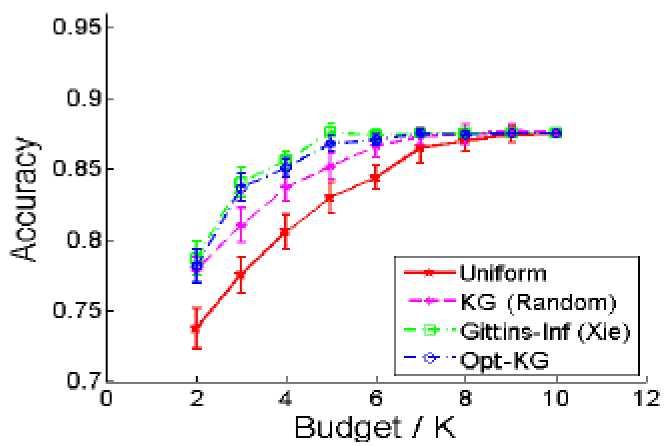
(b) One-Coin Model

www.fppt.info

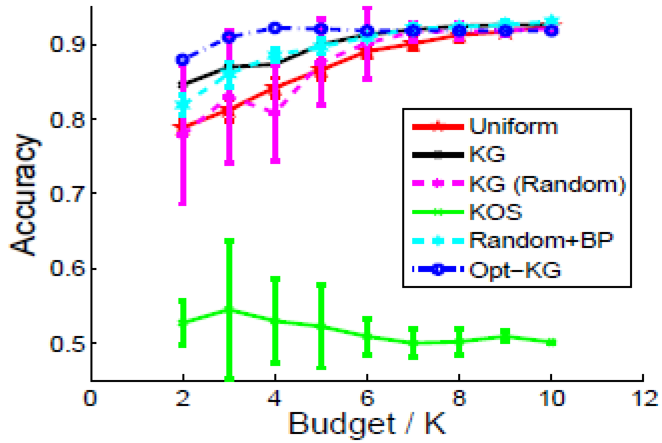
Experiment on real dataset

Recognizing textual entailment (RTE)
800 instances and each instance is a sentence pair
10 workers
Whether the second hypothesis sentence can be inferred from the first one

Workers' reliability: *Beta(4, 1)*



(a) RTE: Majority Vote



(b) RTE: One-Coin Model

www.fppt.info

Conclusion

- Formulate the budget allocation in crowdsourcing into a MDP and characterize the optimal policy using DP.
- Computationally propose an approximate policy, optimistic knowledge gradient.
- MDP formulation can be used as a general framework to address various budget allocation problems in crowdsourcing.

www.fppt.info

Thanks for your attention

www.fppt.info