サーチ活動

Kyohei Okumura*

June 21, 2018

1 Kremer, Mansour and Perry (2014, JPE) "Implementing the Wisdom of the Crowd"

- Recommendation Mechanism
  - Trip Advisor・AirBNB・食べログ・car navigation system・(Amazon) etc.
- Information Design, Bandit with Incentives

1.1 Model

- One central planner + T agents
- two arms: $A := \{a_1, a_2\}$
- The arm a_i has its reward R_i.
  - Before exploration, no one knows the exact value of R_i; once someone explores a_i, CP knows its value. (deterministic reward)
  - $R_i \sim \pi_i$, $\mathbb{E}_{\pi_i}[R_i] = \mu_1$, $\mu_1 \geq \mu_2$.
  - π: the joint distribution, common knowledge.
- Each agent t knows his position in a line.
- CP knows entire history: CP's recommendation, agents' choices, and their rewards.
- CP commits to a message policy.
- Each period, CP sends a message m^t to agent t, agent t decides the arm to choose, and agent t receives his reward R^t.
- Agent t chooses the arm $a_i^* \in \arg\max_{a_i}\{\mathbb{E}[a_i \mid m^t]\}$
- CP wants to maximize the social welfare:

$$\mathbb{E}\left[\frac{1}{T}R^t\right]$$

- What is the best policy?

*E-mail: kyohei.okumura@gmail.com

1.2 Example

- $R_1 \sim U[-1,5],\ R_2 \sim U[-5,5].\ \mu_1 = 2,\ \mu_2 = 0.$
- Each agent arrives sequentially: agent t arrives in period t.
- Suppose for simplicity that CP wants to explore both arms as soon as possible.
- NB: Agent 1 always chooses a_1.

1.2.1 Full transparency

- If CP discloses all information, agent 2 explores a_2 only if $R_1 \leq 0$.

1.2.2 Threshold policy

- Consider the following recommendation policy:

$$m^2 := \begin{cases} a_2 & (R_1 \leq 1) \\ a_1 & (R_1 > 1) \end{cases}$$

$$\mathbb{E}[R_1 \mid m^2 = a_2] = 0 \leq m_2$$

- Whenever CP recommends a_2 to agent 2, agent 2 follows the recommendation.
- Pr(exploration) $\uparrow$
- As for agent 3 we can construct the threshold policy:

$$m^3 := \begin{cases} a_2 & ([R_1 \leq 1, R_2 > R_1] \text{ or } [R_1 \in (1, 1+x]]) \\ a_1 & \text{o.w.} \end{cases}$$

- In case $R_1 \in [-1, 1]$, agent 2 explores a_2.
- In case $R_1 \in (1, 1+x]$, agent 3 explores a_3.
- ...

1.3 Results

- We can focus on the specific class of policies: *recommendation policy*
  - Revelation-principle-like argument
  - Myerson (1986), **Sugaya and Wolitzky (2017)**
- The optimal policy takes the form of threshold policy.
- Extension: Imperfect information about location, **stochastic reward**

1.4 Related Literature

1.4.1 Frazier et al. (2014, EC'14), "Incentivizing Exploration"

- Use monetary transfer to resolve exploration-exploitation tradeoff.

1.4.2 Bergemann and Valimaki (1996), "Learning and Strategic Pricing"

- One consumer v.s. Many firms, firms = arms
- Arms themselves can set strategic prices for begin pulled.

1.4.3 Abraham et al.(2013), "Adaptive Crowdsourcing Algorithms for the Bandit Survey Problem"

- crowd-sourcing of tasks
- arms = agents
- learn the quality of agents' work from observing them while providing enough incentives for them to work

1.4.4 Mansour et al. (2018, EC'15), "Bayesian Incentive-Compatible Bandit Exploration"

- Generalization of KMP (2014)
- K arms
- stochastic rewards
- detail-free
- near-optimal performance(?)

1.4.5 Bimplins et al. (Management Science, forthcoming), "Crowdsourcing Exploration"

- Generalization of KMP
- two arms
- stochastic rewards
- less-than-fully informative policy
- information design + monetary transfer
- optimal (not first-best) policy can be obtained by solving large-scale LP (intractabel)
- also propose a heuristic solution that is near-optimal in numerical experiments

1.5 Possible Extension

- Bounded rationality?: It is not realistic to assume that agents are perfectly Bayesian-rational.
  - ignorance of selection-bias?
  - Agents may not recognize the possibility that some firms incentivize reviewers by monetary payoffs.

2 Others

2.1 Bimpikins et al. (Operations Research, forthcoming) "Designing Dynamic Contests"

- contests, learning, dynamic competition, information design
- Innovation contests by firms.
  - outsource innovation to the crowd
  - winners are awarded by prize
- How to best design a contest?

- What is the best information disclosure policy?:
  Whether and when should the contest designer disclose the information regarding the competitors' partial progress?
- What is the role of intermediate awards?

- 成功するかどうかわからない.
- 最初に成功した人のみが報酬をもらう.
- Tradeoff: encouragement effect v.s. competition effect
  - 他人が成功していることがわかる → 成功しやすいプロジェクトなので自分も参加したい.
  - 他人が成功していることがわかる → 自分が遅れていると，今から参入しても勝てなさそうなので参加したくない.

2.2 Weyl and Zhang (2018), "Depreciating Licenses"

- simple DMD

2.3 Kleinberg, Waggoner and Weyl (2016), "Descending Price Optimally Coordinates Search"

- information acquisition cost
- Descending auctions are "better" than ascending auctions?