Game Theory for Data Science: Eliciting High-Quality Information

Boi Faltings[1] Goran Radanovic[2]

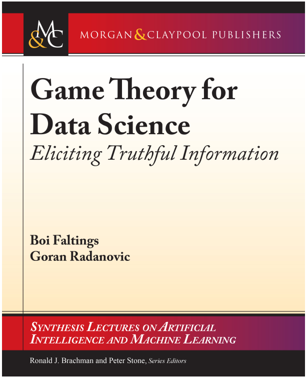
[1]AI Laboratory, EPFL
boi.faltings@epfl.ch

[2]EconCS, Harvard
gradanovic@seas.harvard.edu

December 11th, 2017

Background Material

Boi Faltings and Goran Radanovic:
Game Theory for Data Science:
Eliciting Truthful Information,
Morgan & Claypool Publishers, 2017.
15% discount with code: authorcoll



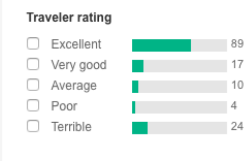
Big data

Organizations and individuals base decisions on data rather than principles:

- financial markets
- medicine
- choosing a restaurant/hotel/spouse
- law enforcement
- ...

Often, data must be obtained from others.

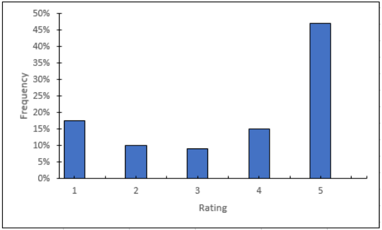
Product reviews

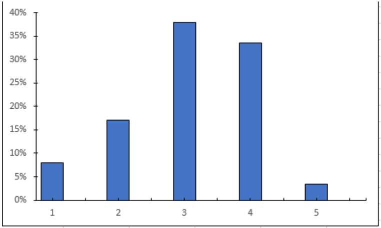


Epinions.com
a Shopping.com company

- Reviews and ratings great to avoid poor products.
- Having reviews is essential for selling the product.

Why do we need to worry about quality?



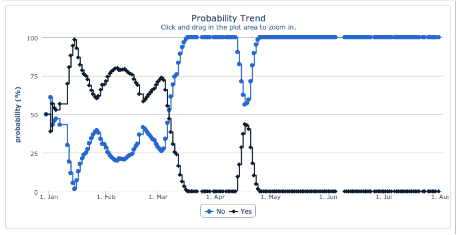


Amazon ratings distribution for music CD Mr. A-Z, reported by Hu et al. (2006), and empirical observation.

- Laziness: most people do not write reviews.
- Self-selection: most reviews are written for ulterior motives, e.g. reviews paid for by hotel, push your own opinion, etc.
- Malicious participants: paint a fake picture of reality.

Forecasting polls

Will Scotland become independent?



- Internet can be used to collect forecasts of important events.
- Important for many high-stakes decisions.
- Need to encourage knowledgeable participants and accurate estimates.

Crowdwork

- Human computation: tasks solved by workers recruited through the internet (e.g. Amazon Mechanical Turk).
- Peer grading: students grade each others' homework.
- Huge benefits for knowledge acquisition, online courses, etc.
- Need to make workers invest effort to obtain quality results.

Quality control options

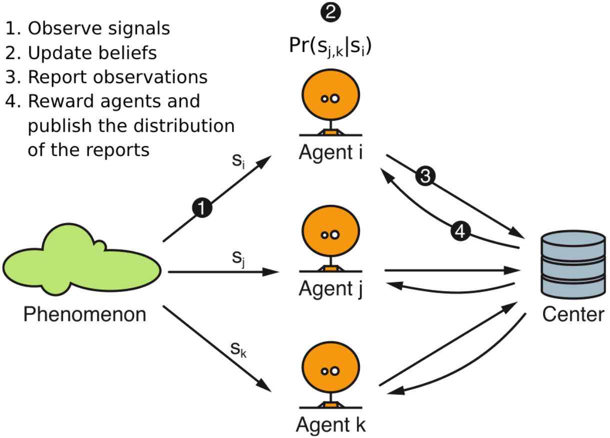
- Filtering: eliminate outliers, inconsistent data.
- Reputation: eliminate bad providers.
- Incentives: encourage participation and effort of good data providers.

All 3 can be used together - we focus on incentives and reputation.

The promise of incentives

- Filtering, reputation need to throw away data $\Rightarrow$ wasteful.
- Incentives can also increase the amount of good data.
- Incentives can be inaccurate as long as participants *believe* that they are right on average.
- However, participants may misunderstand or not care.

Setting



Limit to variables with **discrete** values $x_1, .., x_k$.

Choosing a strategy

- Agent has to choose strategy:
  - heuristic:
    - report a constant value.
    - report a random number.
    - report ...
  - honest/truthful: perform accurate measurement and report truthfully.
- Rational agent: chooses strategy with highest payoff.
- Mechanism: influence choice through *payment rule*.

Principle underlying truthful mechanisms

Reward reports according to *consistency* with a *reference*:

- verifiable information: ground truth g will become known and can be used as a reference.
- unverifiable information: ground truth will never be known. Reference is constructed from *peer* reports.

Roadmap

- Verfiable information
- Unverifiable, objective information
- Parametric mechanisms for unverifiable information
- Non-parametric mechanisms for unverifiable information
- Distributed machine learning

Eliciting verifiable, objective information

Forecasting, estimation, cumulative phenomena: truth can be verified later.
$\Rightarrow$ payment can use verification.

- eliciting a value: reward if report is accurate prediction.
- eliciting a probability distribution: scoring rules.
- eliciting a consensus probability distribution: prediction markets.

Agent beliefs

Agent i has beliefs about what others observe:

1. *prior* probability distribution $\Pr_i(x)$.
   We abbreviate $\Pr_i(x) = p_i(x)$ or just $p(x)$.
   Often common to all agents (e.g. current review scores).
   Maximum likelihood: agent *endorses* $x^p = argmax_x p(x)$.
2. measures signal s_i and forms a *posterior* distribution $\Pr_i(x|s_i)$.
   We abbreviate $\Pr_i(x|s_i) = q_i(x)$ or $q(x)$.
   Update prior $\rightarrow$ posterior often different for each agent.
   Agent *endorses* $x^q = argmax_x q(x)$.

Beliefs motivate agent actions: crucial for incentives.

Eliciting a value

Truth Matching mechanism:

- t_1 agent makes observation and forms posterior belief q.
- t_2 agent reports a value v to the center.
- t_3 ground truth g observed; center pays reward only if $v = g$.

Expected reward $E[pay] = q(v)$
$\Rightarrow$ maximized by choosing $v = x^q = argmax_x q(x)$:
Rational agent reports its best estimate truthfully.

Costly measurements

- Measurement has a cost m.
- If $q \simeq p$, agent may decide to skip the measurement!
- m should not exceed

$$\underbrace{q(x^q)}_{E_{post}[pay]} - \underbrace{p(x^p)}_{E_{prior}[pay]}$$

- Scale payment by $\alpha \geq \frac{m}{q(x^q)-p(x^p)}$ to ensure this condition!
- α depends on measurement technology - could also be inferred from agent behavior.

Discouraging random reports

- Even without measurement, agent still gets a reward by reporting its prior most likely value x^p.

$\Rightarrow$ system will be polluted with many uninformed reports!

- Subtract $E_{prior}[pay] = p(x^p)$.
- Designer needs to estimate $p(x^p)$, $q(x^q)$; can use some background constraints, e.g. $p(x^p) \geq 1/N$.

Components of payment schemes

Final payment rule:

$$pay(x, g) = \alpha \left[-p(x^p) + \begin{cases} 1 & \text{if } x = g \\ 0 & \text{otherwise} \end{cases} \right]$$
$$= \alpha \left[\mathbf{1}_{x=g} - p(x^p) \right]$$

Components:

- incentive for truthful report (1 if ground truth is matched)
- offset to make expected reward of random reports $= 0$
- scale to compensate cost of measurement

Focus on incentives for truthful reports.

Reporting probability distributions

Report is not a value, but probability distribution A.
Proper scoring rule = payment function $pay(A, g)$ such that:

$$(\forall \underline{q'} \neq \underline{q}) \sum_{x} q(x) \cdot pay(\underline{q}, x) > \sum_{x} q(x) \cdot pay(\underline{q'}, x)$$

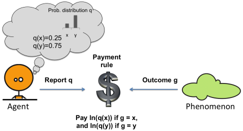
Examples:

- quadratic scoring rule:

$$pay(\underline{A}, g) = 2 \cdot A(g) - \sum_{x \in X} A(x)^2$$

- logarithmic scoring rule:

$$pay(\underline{A}, g) = C + \ln A(g)$$



Gneiting T., and Raftery A. Strictly proper scoring rules, prediction, and estimation. JASA 2007.

Eliciting forecasts with proper scoring rules

"What will be the weather next Sunday: Rain, Cloudy or Sunny?"

- Agent prior = historial averages (for example):

$$p = \begin{array}{c|c|c} \textit{Rain} & \textit{Cloud} & \textit{Sun} \\ \hline 0.2 & 0.3 & 0.5 \end{array}$$

- Agent studies data $\Rightarrow$ posterior belief:

$$q = \begin{array}{c|c|c} \textit{Rain} & \textit{Cloud} & \textit{Sun} \\ \hline 0.8 & 0.15 & 0.05 \end{array}$$

- Agent reports the entire posterior belief.
- On Sunday, using the logarithmic scoring rule he gets paid $C + \ln 0.8 = C - 0.22$ if it rains, and $C + \ln 0.05 = C - 3$ if it is sunny.

Why is this truthful?

Expected reward using log scoring rule:

$$E[pay(\underline{A}, g)] = \sum_x q(x) \cdot pay(\underline{A}, x) = \sum_x q(x) \cdot [C + \ln(A(x))]$$

and the difference between truthful/non-truthful reporting:

$$\begin{aligned} E[(pay(\underline{q}, g)] & -E[pay(\underline{A}, g)] \\ = & \sum_x q(x) \cdot [C + \ln q(x)) - (C + \ln A(x))] \\ = & \sum_x q(x) \cdot \ln \frac{q(x)}{A(x)} \\ = & D_{KL}(\underline{q}||\underline{A}) \end{aligned}$$

By Gibbs' inequality, $D_{KL}(q||A) \geq 0$, so reporting an $\underline{A} \neq \underline{q}$ can only get a lower payoff!

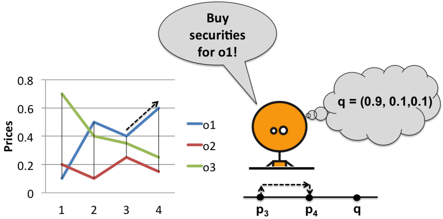
Prediction markets

Model on a financial market.

Market = trade securities $\sigma(x_i)$ for predictions x_i that pay \$1 if $g = x_i$ and \$0 otherwise.

Every security has a market price $\pi(x_i)$.
Competitive equilibrium = $\pi(x_i)$ is a consensus probability estimate for $\Pr(g = x_i)$.

Bigger investment $\Leftrightarrow$ bigger influence, but also risk.



Hanson R. Logarithmic market scoring rules for modular combinatorial information aggregation. JPM 2007.
Chen Y., and David P. A utility framework for bounded-loss market makers. UAI 2007.

Automated market makers

- Consider security R that pays 1 if it rains at any time next Saturday, and 0 if it doesn't rain.
- To ensure that there is always a way to buy and sell this security, we introduce an automated market maker agent.
- Market maker maintains a net balance n of securities R that have been bought or sold to it (negative means sold short).
- Q: how to choose a price function $\pi(n)$ for buying or selling the next share, so that it shows the estimated probability of the event?
- A: Agents should be paid for how much their actions improved or deteriated the predicted probability, according to a proper scoring rule.

Example

- Security R: Rain(Saturday) pays 1 if and only if it rains at any time on Saturday, 0 otherwise. Initially $\pi(0) = p(R) = 0.5$.

t_1 Agent 1 uses prior probability (historical average) $p(R) = 0.2$, short sells n_1 securities R until price becomes $\pi = 0.2$.

t_2 Agent 2 studies data and concludes: $q(r) = 0.8$
$\Rightarrow$ buys n_2 securities R as long as $\pi < 0.8$.

t_3 next, Agent 3 studies the data and concludes $q(r) = 0.5$
$\Rightarrow$ short sells n_3 securities R as long as $\pi > 0.5$.

t_e on Saturday, it rains and agents should get the following rewards (using $\log_2$ scoring rule $Sr(p, x) = b \log_2 p(x)$):

1. $Sr(0.2, 1) - Sr(0.5, 1) = b(\log_2 0.2 - \log_2 0.5) = -1.322b$
2. $Sr(0.8, 1) - Sr(0.2, 1) = b(\log_2 0.8 - \log_2 0.2) = 2b$
3. $Sr(0.5, 1) - Sr(0.8, 1) = b(\log_2 0.5 - \log_2 0.8) = -0.678b$

Market makers with a logarithmic scoring rule

- Assume participant believes that true probability of outcome x_i is $\pi^*(x_i) > \pi(x_i)$.

$\Rightarrow$ buys m securities and makes the price increase to some $\pi(n+m) = \pi' > \pi(n)$.

$\Rightarrow$ he should make a profit of $Sr(\pi', 1) - Sr(\pi, 1)$ if the outcome is indeed x_i:

$$
\begin{aligned}
m - \int_n^{n+m} \pi(\mu) d\mu &= Sr(\pi(n+m), 1) - Sr(\pi(n), 1) \\
(1 - \pi(n)) &= \frac{dSr(\pi(n))}{dn} = \frac{dSr}{d\pi}\frac{d\pi}{dn}
\end{aligned}
$$

LMSR: $Sr(\pi) = b \log_2 \pi \implies \pi(n) = \frac{2^{n/b}}{2^{n/b}+1}$

Example

In the example, assuming $b = 1$: Securities bought/sold:

$$t_1 : \pi(n) = \frac{2^n}{2^n + 1} = 0.2 \quad \Rightarrow \quad n(t_1) = -2 : \; n_1 = -2$$

$$t_2 : \pi(n) = \frac{2^n}{2^n + 1} = 0.8 \quad \Rightarrow \quad n(t_2) = 2 : \; n_2 = 4$$

$$t_3 : \pi(n) = \frac{2^n}{2^n + 1} = 0.5 \quad \Rightarrow \quad n(t_3) = 0 : \; n_3 = -2$$

1. 2 shares sold that drive the price from 0.5 to 0.2: price/share $= (\log_2 0.2 - \log_2 0.5)/2 = -0.339$
2. 4 shares bought that drive the price from 0.2 to 0.8: price/share $= (\log_2 0.8 - \log_2 0.2)/4 = 0.5$
3. 2 shares sold that drive the price from 0.8 to 0.5: price/share $= (\log_2 0.5 - \log_2 0.8)/2 = -0.661$

Example..

At the end:

- Agent 1 has to pay for 2 borrowed shares, net gain $= 2 \cdot 0.339 - 2 = -1.322$.
- Agent 2 gets 4 for this shares, net gain $= 4 - 4 \cdot 0.5 = 2$.
- Agent 3 has to pay for 2 borrowed shares, net gain $= 2 \cdot 0.661 - 2 = -0.678$

Correspondence with scoring rules makes it optimal for agents to act according to their true beliefs about the posterior probability q!

"Irrational" agents

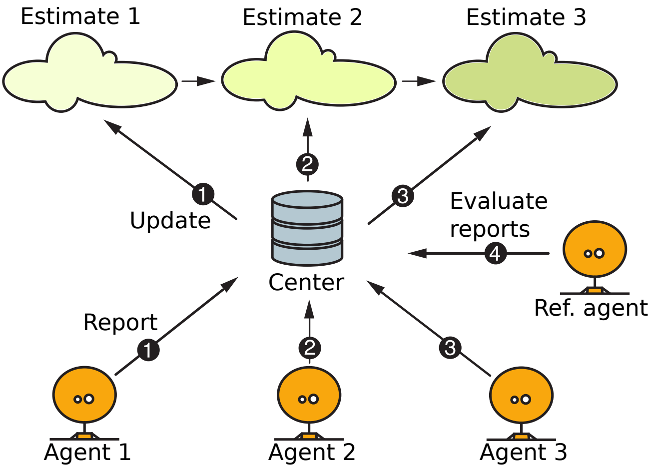
Some agents do not respond to incentives

- faulty agents, who do not consider the incentive or who are unable to provide correct data.
- malicious agents, who want to insert fake data for ulterior motives, for example to hide pollution.

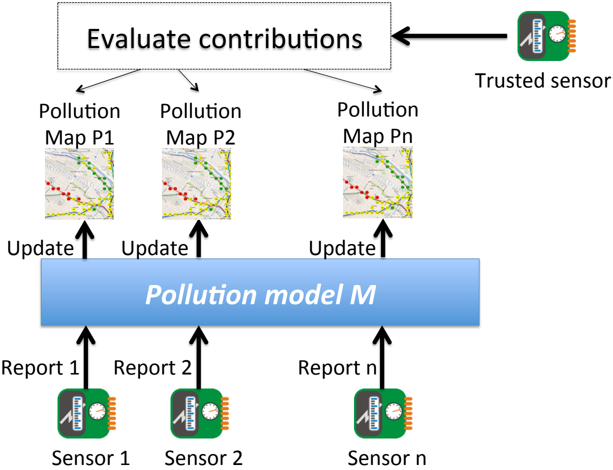
Approach: limit their negative *influence* on the learned model through reputation.

Resnick P. and Sami R. The influence limiter: provably manipulation-resistant recommender systems. RecSys 2007.

Information fusion



Information fusion in community sensing



Reputation principle

- Agents interact with the system over time t.
- Assign a reputation score that determines if an agent is misbehaving.
- The reputation of an agent is based on the agent's *influence*.
- The report of an agent changes prediction for location of reference measurement from $p(x)$ to $q(x)$.
- Evaluate the quality by proper scoring rule Sr on reference measurement g_t:

$$score_t = Sr(q, g_t) - Sr(p, g_t) \in [-1, +1]$$

- Use $score_t$ to update the reputation rep_t.

Simple reputation system

- Thresholding: submitting data requires a minimal reputation.
- Most common representation — β reputation system:

$$rep_t = \frac{\alpha_t}{\alpha_t + \beta_t}$$

where $\alpha_t = \alpha_0 + \sum_{s \in \{scores_\tau > 0\}} |s|$ and
$\beta_t = \beta_0 + \sum_{s \in \{scores_\tau < 0\}} |s|$.

- However, allows manipulation:
  - Provide good data that does not change the model.
  - $\Rightarrow$ build up reputation.
  - Use reputation to insert bad data that *does* change the model.

Josang A., and Ismail R. The beta reputation system. EC 2002.

Stochastic influence limiter

- Stochastic information fusion: with probability $\frac{rep_t}{rep_t+1}$ accept report.
- Exponential reputation update:

$$rep_{t+1} = rep_t \cdot (1 + \frac{1}{2} \cdot score_t)$$

- Main properties:
  1. Negative influence is upper bounded by $2 \cdot init.rep$.
  2. Information loss is upper bounded by a constant.
- Empirical performance often much better.

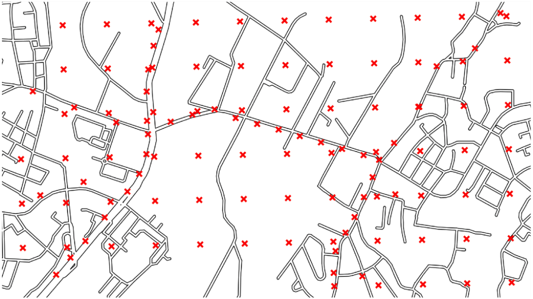
Radanovic G., and Faltings B. Limiting the influence of low quality information in community sensing. AAMAS 2016.

Empirical evaluation (Pollution Sensing)

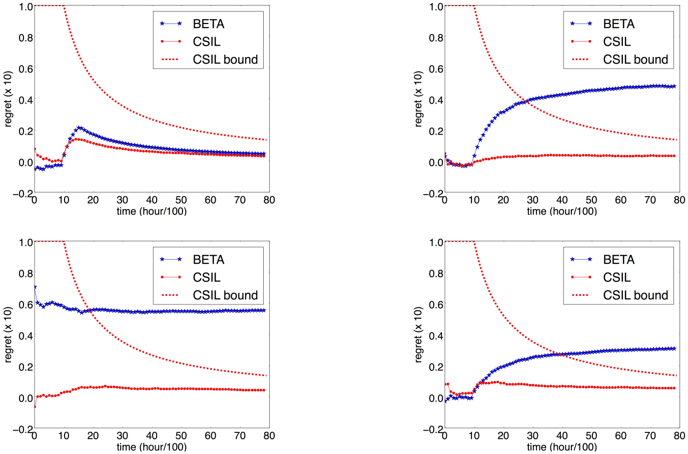
Reputation systems:

- CSIL - stochastic influence limiter
- BETA - beta reputation system

Pollution model of Strassbourg (France):



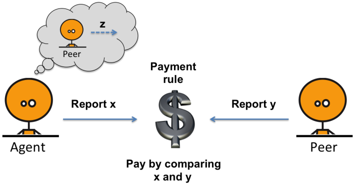
Performance results



Ground truth is never known

In many cases, ground truth is never known:

- product reviews
- community sensing
- predictions about hypothetical questions



Peer consistency: evaluate consistency with peer reports.

Peer consistency mechanisms

- Reporting information becomes a *game* among agents: reward depends on actions of agent *and* peer agent.
- Optimal strategy = equilibrium of the game.
- Truthtelling becomes an equilibrium: if peers are truthful, truthtelling is the best response.
- Equilibria also depend on agents' *beliefs* about others.
- Agents need to have similar beliefs so that the same mechanism works for all of them!

Output agreement

Term coined by von Ahn:

- Ask 2 people to solve a task.
- Pay a constant reward if two people give the same answer.

Q: When does this incentivize truthfulness/maximum effort?
A: In objective tasks: agents believe that honest peers are most likely to obtain the same answer.

Truthful reporting is an *equilibrium*.

von Ahn, L. and Dabbish, L. Designing games with a purpose. Communications of the ACM 2008.

ESP game

- Guess keywords to label image.
- Matching guess of an (unknown) partner gives points.
- Taboo words to exclude trivial choices.

von Ahn, L. and Dabbish, L. Labeling images with a computer game. HFCS 2004.

Objective vs. Subjective

- *objective* data: all agents observe a noisy version of the *same* realization of the phenomenon. Example: temperature at point x and time t. Center wants to know *ground truth*.
- *subjective* data: each agent observes a *different* realization of the phenomenon. Example: service in a restaurant, quality of a product. Center wants to know *distribution*.

Goal: predict *observations* of a new agent.

Subjective observations

- Reporting on the quality of service of Blue Star Airlines, with very high reputation.
- My plane is late and baggage lost. Should I report poor service?
- A: no, because most people enjoy good service, so my report is less likely to match the peer!
- Not categorical: agents do not believe that the same bad service is most likely for others.

Types of peer consistency mechanisms for subjective tasks

There is no truthful peer consistency mechanism without assumptions about agent beliefs!
Known mechanisms make assumptions about prior and posterior beliefs about the distribution of peer agent reports:

- Homogeneous agent population with identical and known prior *and* posterior beliefs, example: peer prediction
- Common and known prior beliefs, but belief updates can be heterogeneous as long as they satisfy a self-predicting condition, for example: Peer truth serum (PTS)
- Utilize the multi-task structure of crowdsourcing to accomodate subjective beliefs and provide stronger incentives, for example: PTS for crowdsourcing

Peer Truth Serum

- Assume center maintains and publishes distribution R of all prior reports (initially uniform). When agent reports $u_i = x_j$, center updates:

$$\hat{R}_{t+1} = (1-\delta)R_t + \delta \underline{x}_j$$

- Reward by impact of x_j on accuracy of model, evaluated using proper scoring rule SR for predicting a random peer report x_p:

$$\begin{aligned} pay(R_t, x_j) &= SR(\hat{R}, x_p) - SR(R, x_p) \\ &= SR((1-\delta)R + \delta \underline{x}_j, x_p) - SR(R, x_p) \end{aligned}$$

Derivative of reward function

- Assume we use log scoring rule:

$$SR(R, x_p) = \ln r(x_p)$$

- with derivative:

$$\frac{\partial SR(R, x_p)}{\partial r(x)} = \begin{cases} 1/r(x) & x = x_p \\ 0 & x \neq x_p \end{cases} = \frac{\mathbf{1}_{x=x_p}}{r(x)}$$

- Partial derivatives of $\hat{R}$ with respect to parameter δ are as follows:

$$\frac{d\hat{r}(x)}{d\delta} = \begin{cases} 1 - r(x) & x = x_j \\ -r(x) & x = x_k \neq x_j \end{cases} = \mathbf{1}_{x=x_j} - r(x)$$

Approximation by Taylor expansion

- Approximate model improvement on a random peer report x_p by the first term of the Taylor expansion:

$$\begin{aligned} SR(\hat{R}, x_p) - SR(R, x_p) &\approx \delta \sum_z \frac{\partial SR(\hat{R}, x_p)}{\partial \hat{r}(z)} \frac{d\hat{r}(z)}{d\delta} \\ &= \delta \sum_z \left(\frac{\mathbf{1}_{z=x_p}}{r(z)} \right) \left(\mathbf{1}_{z=x_j} - r(z) \right) \\ &= \delta \left(\sum_z \frac{\mathbf{1}_{z=x_j} \mathbf{1}_{z=x_p}}{r(z)} - \sum_z \mathbf{1}_{z=x_p} \frac{r(z)}{r(z)} \right) \\ &= \delta \left(\frac{\mathbf{1}_{x_j=x_p}}{r(x_j)} - 1 \right) \end{aligned}$$

Reward for quadratic scoring rule

- Using the quadratic scoring rule:

$$SR(\hat{R}, x_p) = \hat{r}(x_p) - 0.5 \sum_x \hat{r}^2(x)$$

- with derivative:

$$\frac{\partial SR(R, x_p)}{\partial r(x)} = \begin{cases} 1 - r(x) & x = x_p \\ -r(x) & x \neq x_p \end{cases} = \mathbf{1}_{x=x_p} - r(x)$$

- we obtain:

$$\begin{aligned} pay(\underline{\hat{q}}, x_p) &= \delta \sum_z \left(\mathbf{1}_{z=x_p} - r(z)\right)\left(\mathbf{1}_{z=x_j} - r(x_j)\right) \\ &= \delta \left(\mathbf{1}_{x_j=x_p} - r(x_j)\right) \end{aligned}$$

Payment

- Agent i reproduces the identical calculation $\Rightarrow$ incentive for optimally improving the center's estimate.
- Assume Bayesian update; δ for agent i is unknown...
- ...but reward is proportional to δ: δ is just a scaling factor!

$\Rightarrow$ choose payment proportional to the improvement:

$$pay(x_j, x_p) = \frac{\mathbf{1}_{x_j = x_p}}{r(x_j)} - 1$$

Faltings, B., Jurca, R., and Radanovic, G. Peer Truth Serum: Incentives for Crowdsourcing Measurements and Opinions. CoRR abs/1704.05269 2017.

Incentive Compatibility

- Incentive compatibility condition for $x_l \neq x_j$:

$$\begin{aligned} & E_{P(x|x_j)}[pay(x_j, x)] = q(x_j) \cdot pay(x_j, x_j) = q(x_j)/r(x_j) \\ > \quad & E_{P(x|x_j)}[pay(x_l, x)] = q(x_l) \cdot pay(x_l, x_j) = q(x_l)/r(x_j) \end{aligned}$$

- Assume (for now) that agent adopts R as its prior P.

$\Rightarrow$ translates to self-predicting condition:

$$\frac{p(x_j|x_j)}{p(x_j)} > \frac{p(x_l|x_j)}{p(x_l)}, l \neq j$$

- Satisfied for Bayesian belief updates.

Belief updates

- Assume agent receives a signal s_i:

$$\underline{s}_i = (Pr(obs|x_1), Pr(obs|x_2), ..., Pr(obs|x_k))$$

- Bayesian update:

$$u_i(x|obs) = \alpha p_i(x) Pr(obs|x)$$

where $\alpha = 1/Pr(obs)$ set so that $\sum u = 1$.

- Objective update: agent trusts its measurement

$$q_i(x) = u_i(x)$$

- Subjective update: observation is one of many data points

$$q_i(x) = (1-\delta)p_i(x) + \delta u_i(x)$$

with $\delta = 1/n$ if p is formed by $n-1$ other observations.

Sensing error: agent might treat objective data as subjective.

Self-predicting belief updates

- Maximum likelihood: Agent *endorses* $x_i = argmax_x Pr(obs|x)$
- Bayesian update for subjective data:

$$q_i(x) = p_i(x)(1 - \delta + \delta\alpha Pr(obs|x))$$

- $\Rightarrow$ update is *self-predicting*:

$$q_i(x_i)/p_i(x_i) > q_i(x_j)/p_i(x_j), x_j \neq x_i$$

since $x_i = argmax_x Pr(obs|x) = argmax_x \delta\alpha Pr(obs|x)$

Helpful reporting

What if $R \neq P$ (for example, on initializing the mechanism)?
Consider that P is more *informed*, i.e. closer to true distribution P^* than R (in the interval between P^* and R).

$\Rightarrow$ Agents partition values into:

- under-reported: $r(x) < p(x) \Leftrightarrow r(x) < p^*(x)$
- over-reported: $r(x) \geq p(x) \Leftrightarrow r(x) \geq p^*(x)$

Non-truthful strategy: report x instead of y:

- May be profitable if x under-reported or y over-reported.
- Never profitable if x over-reported and y under-reported.

Helpful strategy: never report over-reported x for under-reported y.

Jurca R., and Faltings B. 2011. Incentives for answering hypothetical questions. SCUGC 2011.

Asymptotic accuracy

- Assume center maintains R as an aggregate over reports received over time (for example histogram).
- Asymptotically accurate: R converges to true distribution P^*.
- Any mechanism that induces helpful reporting is asymptotically accurate.
- Peer truth serum admits equilibria in helpful strategies.

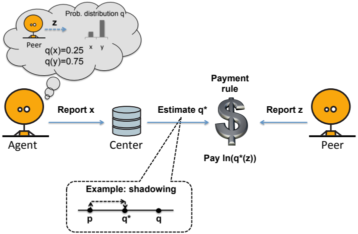
Properties of the Peer Truth Serum

- Optimal: when loss function is logarithmic scoring rule, incentive for agent reports is to minimize loss function for center.
- Unique: any payment function that incentivizes truthful reporting with only the self-predicting condition must have the form $f = 1/p(x_j) + g(-x_j)$ where $g(-x_j)$ is a function independent of the report x_j.
- Maximal: weakening conditions leads to impossibility results.

Peer prediction method

Rather than reward the most likely value...
Peer prediction method:

- Each value for answer x_i is associated with an assumed posterior distribution $\hat{q}(x) = \hat{\Pr}(x|x_i)$.
- $\hat{q}$ is skewed so that x_i is more likely than in prior.
- Use a proper scoring rule to score this posterior against a random peer report.



Miller N., et al. Eliciting informative feedback: the peer prediction method. Management Science 2005.

Other equilibria...

- All agents report x with smallest $r(x)$.

$\Rightarrow$ equilibrium with highest possible payoff.

- Will lead to uniformative, uniform distribution.
- Can be detected: distribution of reports varies a lot over time.

$\Rightarrow$ penalize agents for such behavior.

- More elegant solution: do not publish distribution R, but derive it from multiple answers: PTSC.

Knowing Agent Beliefs

Mechanism design requires knowledge of agent beliefs (prior, prior + posterior).

- *Elicit* beliefs through additional reports, example: Bayesian Truth Serum.
- *Learn* distributions from data; agents believe that mechanism has correct observation $\Rightarrow$ beliefs are identical to measured distribution, example: Peer Truth Serum for Crowdsourcing.

Bayesian Truth Serum

- obtain agent beliefs through an additional *prediction report*: estimate of probability distribution of values in other agents' reports.
- prediction report indicates agents' beliefs.

Prelec D. 2004. A bayesian truth serum for subjective data. Science, 34(5695): 462466.

Decomposable BTS mechanisms

Keep the decomposable structure of the score. Example:

$$\tau_{decomp}(x_i, F_i, x_j, F_j) = \underbrace{\frac{\mathbf{1}_{x_i = x_j}}{F_j(x_i)}}_{\textit{information score}} \underbrace{+ F_i(x_j) - \frac{1}{2}\sum_z F_i(z)^2}_{\textit{prediction score}}$$

Requires additional constraint on agents' beliefs:

$$y = argmax_z Pr(x_i = y | x_j = z)$$

Witkowski J., and Parkes D. 2012. A robust Bayesian truth serum for small populations. AAAI 2012.

Radanovic G., and Faltings B. 2013. A robust Bayesian truth serum for non-binary Signals. AAAI 2013.

Divergence-based BTS mechanism

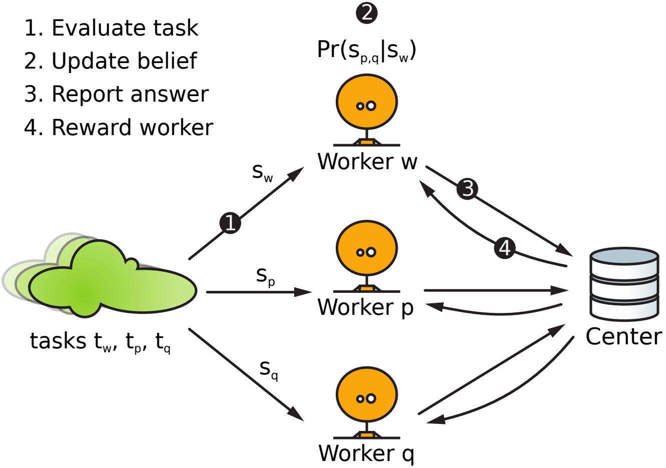
- Drawback of the original BTS:
  - requires a large number of agents
  - robust (decomposable) versions require additional constraints
- Alternative approach: penalize agents for inconsistencies
  - Information score: penalize agents who have the same information reports, while significantly different predictions.
  - Prediction score: score an agent's posterior against a peer report with a proper scoring rule.

Radanovic G., and Faltings B. Incentives for truthful information elicitation of continuous signals. AAAI 2014.

Kong Y., and Schoenebeck G. Equilibrium selection in information elicitation without verification via information monotonicity. Working paper 2016.

Multi-task crowdsourcing

1. Evaluate task
2. Update belief
3. Report answer
4. Reward worker



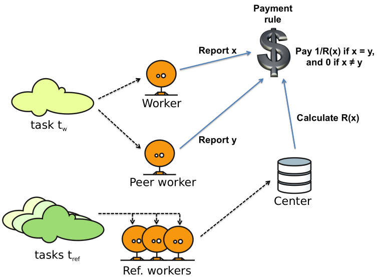
Peer truth serum for crowdsourcing (PTSC)

- Idea: collect R from agents' reports, but keep it private.
- $R =$ histogram of reports from a set of many *similar* tasks, e.g. multiple agents evaluate different airlines.
- Peer report is chosen from reports on the *same* task.
- Agent should believe that:
    - $P \simeq R$ (in the limit of infinitely many tasks).
    - for its own task, $q(x)/r(x)$ is maximized for its own observation x_i.

Radanovic G., et al. Incentives for effort in crowdsourcing using the peer truth serum. ACM TIST 2016.

Algorithm (PTSC)

1. Collect answers to a set of similar tasks $\mathcal{T}$ from crowdworkers.
2. For worker w, calculate $R_w(x) = \frac{num(x)}{\sum_y num(y)}$, where reports by worker w are excluded.
3. For each task t_w carried out by worker w, select a peer worker p that has solved the same task. If they gave the same answer x, reward w with $\alpha \cdot (1/R_w(x) - 1)$, otherwise charge α.



Example (PTSC)

Task	Answers	g
t_1	b , a , a , c	a
t_2	b , b , b , a	b
t_3	a , a , b , a	a
t_4	a , d , a , a	a
t_5	c , c , a , b	c
t_6	d , a , d , d	d
t_7	a , a , c , a	a
t_8	b , b , a , b	b
t_9	a , a , a , a	a
t_{10}	b , b , a , b	b

Probability of different answers across all tasks:

Answer	a	b	c	d
Count	20	12	4	4
R	0.50	0.30	0.1	0.1

Example (PTSC)

Consider an agent a_i who solves t_7 and has $x_i = a$.
Suppose $p(x) \leftarrow R(x)$ and $q(x) \leftarrow freq(x|a)$: self-predicting condition satisfied!
Expected payoffs:

- honest, report a:
  $E[pay(a)] = \frac{0.75}{0.5} - 1 = \frac{1}{2}$
- strategic, report c:
  $E[pay(a)] = \frac{0.1}{0.1} - 1 = 0$
- random, report according to r:
  $E[pay([0.5, 0.3, 0.1, 0.1])] =$
  $0.5 \cdot \frac{0.75}{0.5} + 0.3 \cdot \frac{0.1}{0.3} + 0.1 \cdot \frac{0.1}{0.1} + 0.1 \cdot \frac{0.05}{0.1} - 1 = 0$

Example (PTSC)

Probability of different answers across tasks with the same answer:

Correct answer		Observed answer a	b	c	d
a	$Count(a)$	15	2	2	1
	$freq(\cdot\|a)$	**0.75**	0.1	0.1	0.05
b	$Count(b)$	3	9	0	0
	$freq(\cdot\|b)$	0.25	**0.75**	0	0
c	$Count(c)$	1	1	2	0
	$freq(\cdot\|c)$	0.25	0.25	**0.5**	0
d	$Count(d)$	1	0	0	3
	$freq(\cdot\|d)$	0.25	0	0	**0.75**
	$Count$	20	12	4	4
	R	0.5	0.3	0.1	0.1

$\Rightarrow$ for each task, reporting correct answer has highest prob. of matching peer and payoff!

Properties (PTSC)

- Truthful equilibrium when agents' beliefs satisfy self-predicting condition.
- Expected payoff $= 0$ for heuristic reporting, e.g., random answers according to R
- Truthful equilibrium has the highest payoff.
- Agents do not need to have common prior distribution.

Large number of peers - Log PTS

- If each agent has a large number of peers, the self-predicting condition is not needed.
- Logarithmic Peer Truth Serum:

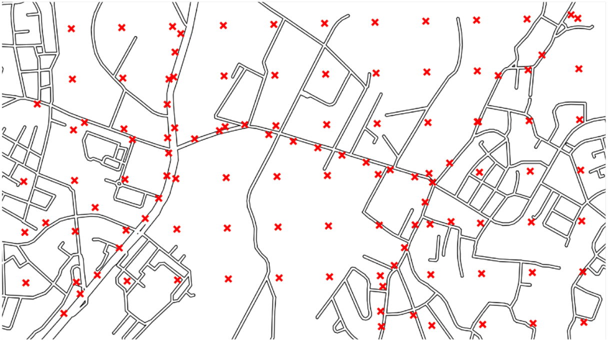
$$\tau(x_{agent}, \ldots) = \ln \frac{freq_{local}(x_{agent})}{R_w(x_{agent})}$$

where $freq_{peers}(x_{agent})$ is a (normalized) frequency of reports equal to x_{agent} among the peers who solve the same task.

Radanovic G. and Faltings B. Incentive Schemes for Participatory Sensing. AAMAS 2015.

Empirical evaluation

Pollution model of Strassbourg (France):



Based on actual measurements, discretized to 4 values.

Dependence on number of Tasks

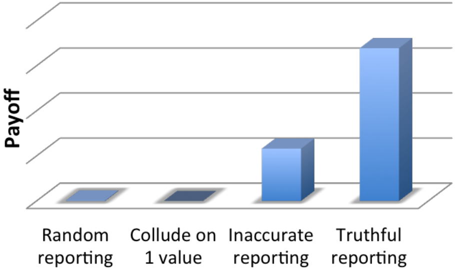
- Few tasks $\Rightarrow$ distribution R_W is noisy.

$\Rightarrow$ truthful incentive may be weakened.

- Min. number of tasks depends on confidence of worker:

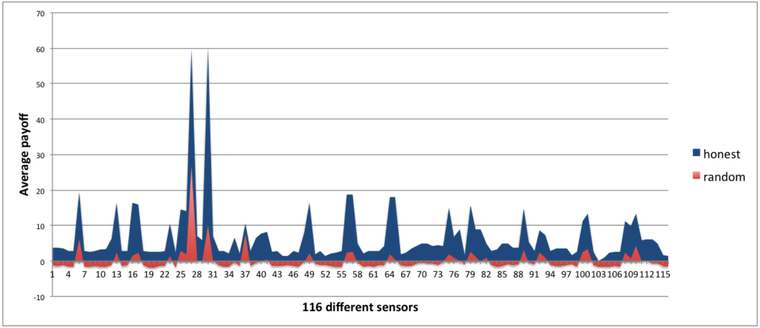
$$min_x \left[\frac{q(x)}{p(x)} - 1 \right]$$

Accurate reports pay off



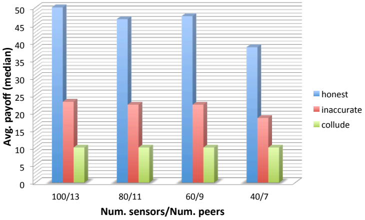
- Collusive and inaccurate reporting strategies are worse than accurate reporting.
- Random reports carry no payoff.

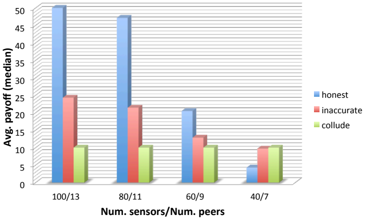
Incentives per sensor



- Sensors have different payoffs (depending on how much pollution varies).
- For each and every sensor, reporting accurate and truthful data is better than other strategies.

Robustness to small populations





Payoffs for different strategy profiles: PTSC (left) vs. Log PTS (right).

- Both PTSC and Log PTS encourage truthful reporting
- Log PTS is more sensitive to the decrease in the number of sensors/peers

Heterogenous Agent Beliefs

- Agents answer multiple tasks and use the same strategy everywhere.
- Agents and center know and agree on sign of correlation among each answer pair for different agents/same task.
- Nothing else is known about agent beliefs.
- Distinguishing correlated values is not important.

Correlated Agreement

Rewards are given through comparison of report x with a randomly chosen peer's answer y.

- Idea 1: base payment on correlation matrix Δ of signals: $\Delta(x, y) = Pr(x, y) - Pr(x)Pr(y)$.
- Define score for agent report x, peer report y as:

$$S(x, y) = \begin{cases} 1 & \text{if } \Delta(x, y) > 0 \\ 0 & \text{otherwise} \end{cases}$$

- Idea 2: compare scores x, y for *same* task t_1 with score for randomly chosen *different* tasks using reports v of agent for t_2 and w of peer agent for t_3:

$$Pay(x, y) = S(x, y) - S(v, w)$$

Correlated Agreement (2)

- Expected payment for truthful reporting is the sum of all positive entries in Δ:

$$E[pay] = \sum_{i,j} \Delta(x_i, x_j) S(x_i, x_j) = \sum_{i,j,\Delta(x_i,x_j)>0} \Delta(x_i, x_j)$$

- Non-truthful strategies would sum other elements: can only achieve smaller sum.
- Truthful strategies result in highest-paying equilibrium!

Dasgupta, A. and Ghosh, A.. Crowdsourced judgement elicitation with endogenous proficiency. WWW 2013
Shnayder, V., Agarwal, A., Frongillo, R. and Parkes, D. Informed Truthfulness in Multi-Task Peer Prediction. EC 2016

Managing the Information Agents

- Group dynamics: learning in repeated applications.
- Self selection.
- Low-quality signals.
- Agent selection.

Group Dynamics

- Peer-based mechanisms assume that argents coordinate through a signal observed from the phenomenon.
- In repeated elicitations with the same peers, agents may learn other heuristic strategies (e.g. always report the same value).
- Can be studied using *replicator dynamics*.
- Output agreement/peer prediction vulnerable, but CA and PTSC are not.

Gao, A., Mao, A., Chen, Y/ and Adams,R. Trick or Treat: Putting Peer Prediction to the Test, EC 2014
Shnayder, V., Frongillo, R. and Parkes, D. Measuring performance of peer prediction mechanisms using replicator dynamics, IJCAI-16

Self-selection

- Center can create incentives for measurements of uncertain signals.
- However, center does not know what it doesn't know!
- Self-selection: agents decide themselves what to measure and contribute.
- Requires that mechanism gives an incentive to measure the interesting (uncertain) signals, and to provide accurate values.

2 Scenarios for Comparison

- Novelty: posterior indicates a different value from the prior:

$$P_1 = (0.1, 0.8, 0.1), Q_1 = P_1 \text{ vs.}$$
$$P_2 = (0.1, 0.8, 0.1), Q_2 = (0.8, 0.1, 0.1)$$

- Precision: lower precision (3 values) vs. higher precision (5 values):

$$\begin{aligned} P_3 &= (0.3, 0.4, 0.3), & Q_3 &= (0.1, 0.8, 0.1)\text{vs.} \\ P_4 &= (0.1, 0.2, 0.4, 0.2, 0.1) & Q_4 &= (0.05, 0.1, 0.7, 0.10.05) \end{aligned}$$

Expected Payments

Mechanism	Expected Payment	Novelty	Precision
Truth Matching (value)	$max_x q(x) - max_x p(x)$	0 vs. 0	0.4 vs. 0.4
Truth Matching (log rule)	$H(P) - H(Q)$	0 vs. 0	0.648 vs. 0.728
Truth Matching (quadratic rule)	$\lambda(Q) - \lambda(P)$	0 vs. 0	0.32 vs. 0.28
Output Agreement	$max_x q(x) - max_x p(x)$	0 vs. 0	0.4 vs. 0.4
Peer Prediction (log rule)	$H(P) - H(Q)$	0 vs. 0	0.648 vs. 0.728
Peer Prediction (quadratic rule)	$\lambda(Q) - \lambda(P)$	0 vs. 0	0.32 vs. 0.28
Peer Truth Serum	$max_x \gamma(x)$	0 vs. 7	1 vs. 1.33
Correlated Agreement	$max_x[q(x) - p(x)]$	0 vs. 0.7	0.4 vs. 0.4
PTS for Crowdsourcing	$max_x \gamma(x)$	0 vs. 7	1 vs. 1.33
Logarithmic PTS	$D_{KL}(Q\|\|P)$	0 vs. 2.1	0.483 vs. 0.492
Bayesian Truth Serum	$H(P) - H(Q)$	0 vs. 0	0.648 vs. 0.728
Divergence-based BTS (log)	$H(P) - H(Q)$	0 vs. 0	0.195 vs. 0.221
Divergence-based BTS (quadratic)	$\lambda(Q) - \lambda(P)$	0 vs. 0	0.32 vs. 0.28

$H(P) = -\sum_x p(x) \log p(x)$ (Shannon Entropy),
$\lambda(P) = \sum_x p(x)^2$ (Simpson's diversity index),
$\gamma(x) = q(x)/p(x) - 1$ (Confidence).

Novelty scenario

- Value does not change (P_1/Q_1): all schemes have expected reward $= 0$.
- Value changes (P_2/Q_2): only PTS, CA, PTSC and LPTS provide an expected incentive!

$\Rightarrow$ with other mechanisms, agents would not want to measure novel data!

- Center has to provide extra incentives $\Rightarrow$ center has to know what it doesn't know.

Precision scenario

- Mechanisms with constant rewards (truth matching, output agreement, correlated agreement) are incentive-neutral.
- Mechanisms based on quadratic scoring rule *discourage* precision!
- Mechanisms based on log. scoring rule (including PTS) give incentive for higher precision.

Incentives for precision are also important to discourage reporting low-quality signals.

Exploiting self-selection

- Self-selection is an important idea: only information agents know what the center does not know.
- However, only some of the schemes provide the right incentives.
- Further research required to focus on this design criterion as well.

Avoiding low-quality signals

- Agents could collude to report something else than the true signal.
- Example: hash task description into the answer space $\Rightarrow$ answers depend on the task, but not in the right way.
- Requires coordination among information agents; feasible only in some scenarios.
- Best counter: spot-check with trusted reports and penalize disagreement (as in influence limiter).

Scaling incentives...

Traditional approach — scale so that:

1. Expected payments for uninformed reporting $= 0$.
2. Expected payments for accurate reporting > 0.
3. Possible to learn a scaling parameter.*

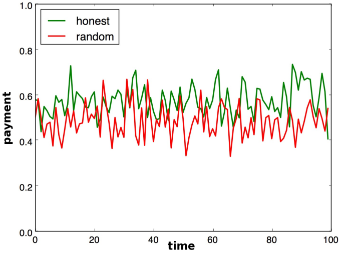
Two drawbacks:

1. Requires negative payments.
2. Susceptible to large noise.

*Liu, Y. and Chen Y. Learning to incentivize : eliciting effort via output agreement. IJCAI 2016

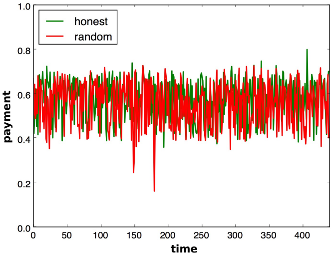
Traditional approach — susceptibility to noise

Crowdsourcing



noisy peer answers

Community sensing



noisy measurements

How to boost the difference between the payments? — apply the reputation based approach!

Radanovic, G. and Faltings, B. Learning to scale payments in crowdsourcing with PropeRBoost. HCOMP 2016

Learning to scale payments

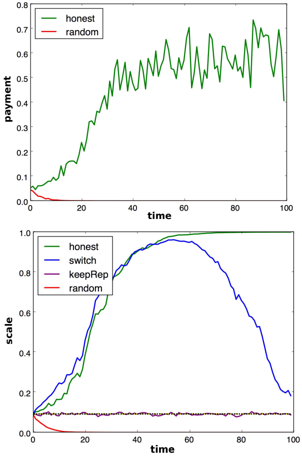
- Agents interact with a mechanisms over time.
- Use reputations to track the quality of reported information.
- The reputation is based on the quality score calculated by a peer consistency approach.
- Peer can be an output of a "truth estimator" — $\hat{\theta}_{\mathcal{F}}$.
- Quality score:

$$\pi_t(x) = \mathbf{1}_{\hat{\theta}_{\mathcal{F}}=x} - Pr(\hat{\theta}_{\mathcal{F}} = x)$$
$$score_t(x) = (1-\alpha)\cdot\pi_t(x) - \alpha$$

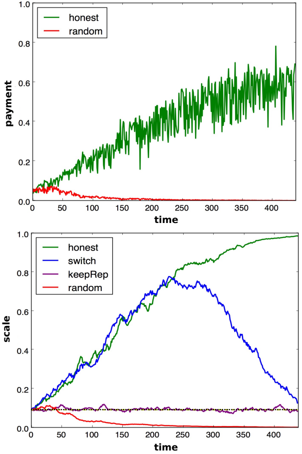
$Pr(\hat{\theta}_{\mathcal{F}} = x)$ can be estimated, α determines the minimal acceptable quality.

ProperBoost — performance results

Crowdsourcing



Community sensing



Agent and peer selection

For distributed agents, important to define:

1. Possible peers of each agent:
   - Peer can be an output of a "truth estimator"
   - Applying machine learning to obtain an unbiased estimator
2. Agent selection under limited budget:
   - Each selected agent must have a "good" peer
   - Leads to a constrained subset selection problem

Liu, Y. and Chen Y. Machine-learning aided peer prediction. EC 2017

Radanovic, G., Singla, A., Krause, A., and Faltings, B. Information Gathering with Peers: Submodular Optimization with Peer-Prediction Constraints. AAAI 2018

Conclusions

- Game theory allows to make payments for data depend on accuracy:
  1. Dominant strategies for verifiable information.
  2. Strongly truthful equilibria for unverifiable information.
- Non-parametric mechanisms allow heterogeneous agent beliefs, under some conditions.
- Managing agents to avoid collusion can be important.

To read more

Boi Faltings and Goran Radanovic:
Game Theory for Data Science:
Eliciting Truthful Information,
Morgan & Claypool Publishers, 2017.
15% discount with code: authorcoll