Differentially Private and Incentive Compatible Recommendation System for the Adoption of Network Goods

Kevin He and Xiaosheng Mu

Harvard University

June 12, 2014

Introduction

- Motivation: rising popularity of friendship-based recommendation systems
- Many platforms include a social networking feature. Platforms use friendship in these networks as a tool towards better product recommendation.
- Examples
  - Netflix Social recommends TV shows that friends have marked as favorites.
  - Last.fm's music recommendation takes into account the musical taste of friends.
  - Facebook's "People You May Know" feature suggests new "products" (new Facebook friends) based on the friends lists of current friends

Social recommendation systems and privacy

- An example: some users may view their Facebook friends list as private information. However, the "People You May Know" feature gives a user recommendations (regarding new friendships) using the private data of others (friends lists of her current friends).
- A further problem: users may have prior belief about the quality of product, not compelled to follow website's recommendation
- Roughly speaking, the problem is to design an optimal recommendation system subject to the constraints of (i) privacy; (ii) incentive compatibility. Later, will see the tension between these two constraints.

Outline for rest of this talk

- A model to capture the privacy and incentive compatibility constraints
  - adoption of a network good
  - myopic agents
- A toy example to highlight the geometry of the constraints
- Feasibility and optimality results

Basic model setup

"Agents living on a graph face the introduction of a new network good."

Why network good?

Let $G = (V, E)$ denote a graph where:

- $V = \{v_1, ..., v_n\}$ represents agents
- Edges represent a symmetric relationship along which externality effect of the network good takes form.
- Agents have incomplete information regarding value of the good, since adoption status of other agents in the network are private information.

STAGE I: initial adoption

Each node independently adopts the good with probability p.
Denote adoption status of node v_i as $\theta_i \in \{0, 1\}$, private to v_i.
Summarize the STAGE I adoption state of the network as
$\theta = (\theta_1, ..., \theta_n) \in \Theta$.

The ideas behind STAGE I:

- Models a decentralized adoption process before intervention of recommendation system
- From technical perspective, STAGE I gives rise to a common prior

Stage II: recommendation

Recommendation system sends a recommendation to a subset of Stage I non-adopters. Each non-adopter then decides whether to adopt or not. Non-adopter v_i gets payoff

$$\phi\left(\frac{\sum_{w \in N(v_i)} \theta_w}{d(v_i)}\right) - c$$

if they adopt in Stage II, 0 otherwise, where

- $N(v_i)$ is the set of neighbors of v_i
- $d(v_i)$ is the degree of v_i
- $c > 0$ is a parameter representing cost of adoption
- ϕ is a continuous, increasing function normalized to $\phi(0) = 0$, $\phi(1) = 1$.
  - special case: $\phi = \mathrm{id}$.

Remark: agents are myopic.

The recommendation system

The recommendation system is a stochastic function $\mathcal{A} : \Theta \to \Theta$ which the designer announces before $\textsc{Stage I}$.

- The function maps a profile of $\textsc{Stage I}$ adoption status to a profile of recommendations.
- For each i with $\theta_i = 0$, $\mathcal{A}^i(\theta) = 1$ means $\mathcal{A}$ sends a recommendation to i, while $\mathcal{A}^i(\theta) = 0$ means $\mathcal{A}$ does not.

Upon receiving or not receiving a recommendation, agent v_i with $\theta_i = 0$ computes conditional expected payoff from adopting the good:

$$\pi_i(\mathcal{A}^i) := \mathbb{E}_\theta \left[\phi \left(\frac{\sum_{w \in N(v_i)} \theta_w}{d(v_i)} \right) - c \;\middle|\; \mathcal{A}^i \right]$$

and adopts if and only if $\pi_i \geq 0$.

The designer faces a constrained optimization problem.

Social welfare (SW) objective

$$\mathbb{E}_{\theta}\left[\mathbb{E}_{\mathcal{A}(\theta)}\left[\sum_{v_i \in V} 1_{\{\theta_i=0\}} \cdot 1_{\{\pi_i \geq 0\}} \cdot \left(\phi\left(\frac{\sum_{w \in N(v_i)} \theta_w}{d(v_i)}\right) - c\right)\right]\right]$$

For every realization of $\mathcal{A}(\theta)$, each agent with $\theta_i = 0$ either receives a recommendation or receives no recommendation.

Based on this signal, such agents compute conditional expected payoff to adoption and uses the rule $\pi_i \geq 0$ to make adoption decision.

Incentive compatibility (IC) constraint

Incentive compatibility (**IC**) says each agent must find it beneficial to do as recommended[1].

$$\begin{cases} \mathbb{E}_{\theta,\mathcal{A}}\left(\phi\left(\frac{\sum_{w\in N(v_i)}\theta_w}{d(v_i)}\right)-c \;\middle|\; \mathcal{A}^i(\theta)=1\right) & \geq 0 \\ \mathbb{E}_{\theta,\mathcal{A}}\left(\phi\left(\frac{\sum_{w\in N(v_i)}\theta_w}{d(v_i)}\right)-c \;\middle|\; \mathcal{A}^i(\theta)=0\right) & \leq 0 \end{cases} \qquad (\textbf{IC})$$

[1] Some authors like Myerson also call this the "obedience constraint".

Differential privacy (DP) constraint

ϵ-differential privacy (**DP**) requires that $\mathcal{A}$ is not too sensitive to small changes in θ.

In the context of database privacy, **DP** was originally proposed by Dwork:

Definition

Write $\mathbb{D}$ for the set of possible states of the database. Then the $\mathbb{S}$-valued randomized algorithm $\mathcal{A} : \mathbb{D} \to \mathbb{S}$ is said to be ϵ-differentially private if for every $S \subseteq \mathbb{S}$ and every $D, D' \in \mathbb{D}$ where D' differs from D in only a single row,

$$\Pr[\mathcal{A}(D) \in S] \leq \exp(\epsilon) \cdot \Pr[\mathcal{A}(D') \in S]$$

Differential privacy (DP) constraint

In the context of our problem, for any two profiles of $\textsc{Stage I}$ adoption states θ, θ' differing in only one component and for $\chi \in \{0, 1\}$,

$$\Pr\left[\mathcal{A}^i(\theta) = \chi\right] \leq \exp(\epsilon) \cdot \Pr\left[\mathcal{A}^i(\theta') = \chi\right] \qquad \textbf{(DP)}$$

Here, Θ plays the role of possible states of database.

Technically, this is a relaxation of the usual **DP** definition. The "output" of $\mathcal{A}$ is really a vector of recommendations. However, we are only asking that recommendation given to v_i is not too sensitive to the types of v_j for $j \neq i$. This relaxation can be motivated by a no-collusion assumption.

Summary of designer's problem

Call a stochastic function $\mathcal{A}$ **feasible** if it satisfies both **IC** and **DP**.

The designer's problem is to pick a feasible $\mathcal{A}$ to maximize **SW**.

The key feature of the model is the tension between **IC** and **DP**:

- **DP** says recommendations must not be too informative of the adoption status of neighbors
- However, **IC** requires the recommendations to be sufficiently informative
- In two extreme cases:
  - Uniformly random recommendations satisfy **DP**, but violates **IC** if c is large
  - Socially optimal recommendations satisfy **IC**, but violates **DP** if ϵ is small
- the classic tension between preserving data privacy and producing useful outputs

A toy example

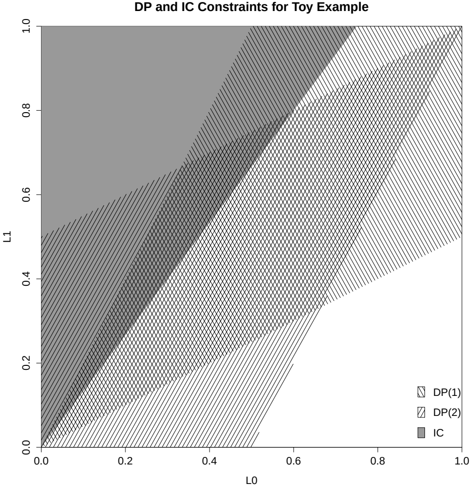
Consider a simple network with only two vertices connected with an edge. Set parameters:

- **DP** parameter $\epsilon = \ln(2)$
- probability of $\textsc{Stage I}$ adoption $p = 0.2$
- cost of adoption $c = 0.25$
- utility functional form $\phi = \mathrm{id}$

A recommendation system in this toy example is characterized by $l_0, l_1 \in [0, 1]$.

- l_0 is probability of receiving recommendation when neighbor is a $\textsc{Stage I}$ non-adopter
- l_1 is probability of receiving recommendation when neighbor is a $\textsc{Stage I}$ adopter
- Think of the set of all recommendation systems as $[0, 1] \times [0, 1]$

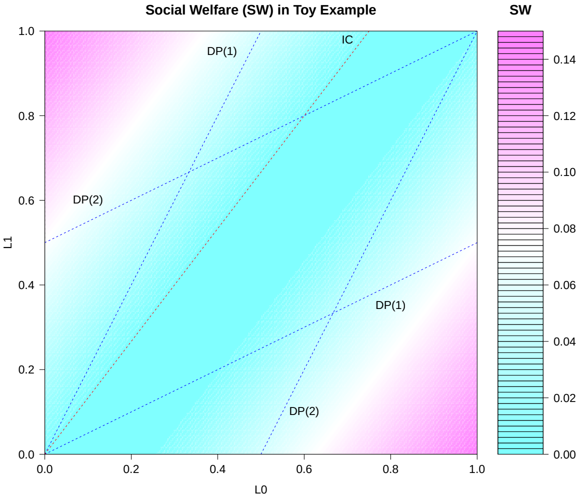
A toy example



- Rec. systems in the central diamond satisfy **DP**
- Rec. systems in upper left corner satisfy **IC**

A toy example

In the case where **DP** and **IC** intersect at points other than $(0,0)$, the problem of the designer is to maximize **SW** over this intersection.



General results

Assume throughout cost of adoption is not too small relative to its benefits, else there is no role for a recommendation system.

$$c \geq \sum_{k=0}^{d} \phi(d/k) \cdot \binom{d}{k} \cdot p^k (1-p)^{d-k} \qquad \text{(Non-Trivial Cost)}$$

With this one additional assumption, we derive a feasibility condition for the constrained optimization problem and explicitly characterize the optimal recommendation system.

Feasibility condition

Proposition

Fix a $\textsc{Stage I}$ *non-adopter agent* v *and write* $d := d(v)$*. There exists a non-trivial recommendation system* $\mathcal{A}$ *satisfying* ***IC*** *and* ***DP*** *for* v *if and only if the cost of adoption satisfies* $c \leq \bar{c}(d, \epsilon, p)$*, where:*

$$\bar{c}(d, \epsilon, p) := \frac{1}{(1 - p + p \cdot \exp(\epsilon))^d} \sum_{k=0}^{d} \phi\left(\frac{k}{d}\right) \cdot \binom{d}{k} (p \cdot \exp(\epsilon))^k (1-p)^{d-k}$$

To get a feasibility condition for the entire network, take min of $\bar{c}(d, \epsilon, p)$ across all vertices.

Remarkably, in the special case where $\phi = \text{id}$, the feasibility condition independent of d, making the result topology-free.

Corollary

If $\phi = id$*, then* $\bar{c}(d, \epsilon, p) = \frac{\exp(\epsilon) \cdot p}{1 - p + \exp(\epsilon) \cdot p}$

The optimal recommendation system

Proposition

Fix a STAGE I *non-adopter agent v and suppose $c \leq \bar{c}(d, \epsilon, p)$. There exists integer $\bar{k}$ with $0 \leq \bar{k} \leq d$ such that v's interim utility is maximized when:*

$$I_k = \begin{cases} \frac{\exp(\epsilon)}{\exp(\epsilon)+1} \cdot \exp(\epsilon \cdot (k - \bar{k})) & \text{if } k \leq \bar{k} \\ 1 - \frac{1}{\exp(\epsilon)+1} \cdot \exp(\epsilon \cdot (\bar{k} - k)) & \text{if } k > \bar{k} \end{cases}$$

To interpret, the optimal rec. system uses a cutoff $\bar{k}$, dependent on d, ϵ, p, c, so that:

- Rec is sent with probability $\frac{\exp(\epsilon)}{\exp(\epsilon)+1}$ to an agent with $\bar{k}$ adopter neighbors
- Rec probability decreases at rate $\exp(\epsilon)$ for agents with fewer than $\bar{k}$ adopter neighbors
- Non-rec probability decreases at rate $\exp(\epsilon)$ for agents with more than $\bar{k}$ adopter neighbors

Summary

- Model: A network good adoption model that captures the tension between incentive compatibility and privacy
- Results:
  - Feasibility condition: $c \leq \bar{c}(d, \epsilon, p)$.
  - Optimal recommendation system: There is some cutoff $\bar{k}$ associated with $\frac{\exp(\epsilon)}{\exp(\epsilon)+1}$ probability of recommendation. Designer fully exploits privacy constraint around this cutoff.

The End

Thank you!