Machine Learning in Finance Workshop 2020

Optimal, Truthful, and Private Securities Lending

Emily Diana
The Wharton School, University of Pennsylvania

COLUMBIA UNIVERSITY Data Science Institute

COLUMBIA ENGINEERING The Fu Foundation School of Engineering and Applied Science

Bloomberg

Overview

Acknowledgements

- Joint work with Michael Kearns, Seth Neel, and Aaron Roth
- To appear in 2020 ACM International Conference on AI in Finance
- Full paper available at `https://arxiv.org/abs/1912.06202`
- Questions and comments are welcome and can be directed to ediana@wharton.upenn.edu

Motivation

Motivated by challenges associated with **securities lending**, the mechanism underlying short selling of stocks in financial markets

We consider allocation of a **scarce commodity** in settings in which **privacy concerns** or **demand uncertainty** may be in conflict with **truthful reporting**

Goal is to construct a privacy protecting allocation mechanism that motivates truthful reporting without sacrificing too much utility

Model

- Lender distributes up to V shares to n clients over time horizon T at fixed price per unit
  - Thus, prices cannot be used as a tool to enforce truthfulness, as is standard in mechanism design
- Each client i has **non-strategic** distribution over usages, U_{it}
- Client has **strategic** distribution over requests, $Q_{it}(r_{it}|u_{it})$
- Together, these form a tabular, joint distribution:

$$Q_{it}(u_{it}, r_{it}) = Q_{it}(r_{it}|u_{it})U_{it}(u_{it})$$

- At each time t, client i draws $u_{it}, r_{it} \sim Q_{it}(u_{it}, r_{it})$, but **only request, r_{it}, is visible to lender**
- Client's payoff is number of shares actually used: if client i is allocated s_{it} shares in an allocation S_t, the payoff is

$$v_i(S_t) = min(s_{it}, u_{it})$$

Client Distributions

Definition

We consider a distribution for client i at time t truthful if $Q_{it}(r_{it}|u_{it}) = 1$ if $u_{it} = r_{it}$ and $Q_{it}(r_{it}|u_{it}) = 0$ otherwise.

Below are two strategic choices of $Q_{it}(r_{it}|u_{it})$ for client i at time t where

$$U_{it}(0) = U_{it}(1) = U_{it}(2) = \frac{1}{3}$$

Table: Sample Truthful Distribution

u_{it} \ r_{it}	0	1	2
0	$\frac{1}{3}$	0	0
1	0	$\frac{1}{3}$	0
2	0	0	$\frac{1}{3}$

Table: Sample Untruthful Distribution

u_{it} \ r_{it}	0	1	2
0	$\frac{1}{9}$	$\frac{1}{9}$	$\frac{1}{9}$
1	0	$\frac{1}{6}$	$\frac{1}{6}$
2	0	0	$\frac{1}{3}$

Lender's Setting

Lender's Goal: Choose *allocation rule* A to maximize lender's utility

Definition

An allocation rule A is a one-shot algorithm that maps a set of requests (r_{it}) and conditional distributions $Q_{it}(\cdot|u_{it})$ on r_{it} to an allocation S_t

- Specifically, $S_t = \{s_{it}\}$ s.t. $\sum_i s_{it} \leq V$
- Since the algorithm is one-shot, we can drop the subscript t
- Allocation rule assumes full knowledge of conditional distributions $Q_i(r_i|u_i)$, which could be estimated from a client's history
- Lender's utility for allocation rule A is:

$$v(A) = \sum_i \mathbb{E}_{Q_{it},A}[\min(A(r_1, \ldots, r_n; Q_1, \ldots, Q_n)_i, u_{it})]$$

- Client's utility for allocation A is:

$$v_A^i(Q_{it}) = \mathbb{E}_{r_{it} \sim Q_{it}(\cdot|u_{it})}[v_i(A(r_{it}, r_{-it}; Q_{it}, Q_{-it}))]$$

Lender's Setting

Given knowledge of Q_i, lender can compute the posterior distribution $Q_i(u_i|r_i)$ on the true demand u_i given r_i, via Bayes' rule:

$$Q_i(u_i|r_i) = \frac{Q_i(r_i|u_i)U_i(u_i)}{\sum_{u'} Q(r_i|u')U_i(u')}$$

Definition

Given $Q_i(u|r_i)$, we denote by $T_i(s|r_i)$ the tail probability $Pr_{u_i \sim Q(u|r_i)}[u \geq s] = \sum_{s' \geq s} Q_i(s'|r_i)$, or the probability of client i using *at least* s shares.

Optimal Allocation Rule

Algorithm maximizing lender's utility $v(S)$: The following algorithm operates by sequentially assigning shares $1 \ldots V$, where each share is assigned to the client i most likely to utilize one additional share.

Algorithm 1 Greedy Allocation Rule

Input: $n, V, \{Q_i(u_i|r_i)\}_{i\in[n]}, r$

Output: feasible allocation $S = \{s_i\}$.

procedure $\textsc{Greedy}(n, V, \{Q_i(u_i|r_i)\}_{i\in[n]}, r)$

  Initialize $s_i = 0,\ \forall i.$ ▷ number of shares allocated to client i

  for $t = 1 \ldots V$ **do**

    Let $i^* = \text{argmax}_i T_i(s_i + 1|r_i)$

    update $s_i \leftarrow s_i + 1$

Optimal Allocation Rule

Theorem: The allocation returned by *Greedy*, S, maximizes the expected payoff for the lender:

$$S \in \arg\max_{S:\sum_i s_i = V} v(S) = \sum_i \mathbb{E}_{Q_i(u|r_i)}[\min(s_i, u_i)]$$

Dominant-Strategy Truthfulness

Given that the lender is solving the allocation problem optimally for the reported Q distributions, **truth telling is a dominant strategy**:

Theorem: Fix a set of choices Q_{-i} and reports r_{-i} for all clients other than i, and a realization of client i's usage $u_i \sim U_i$. Let Q_i^T denote the truthful strategy $Q_i^T(r_i|u_i) = \mathbf{1}_{r_i}$, and let $Q_i(r_i|u_i)$ denote any other strategy. Let A denote the lender's optimal allocation. Then:

$$v_A^i(Q_i) \leq v_A^i(Q_i^T)$$

Dominant-Strategy Truthfulness

Intuition: Truthfulness is a dominant strategy, because it maximizes a client's tail probabilities for units up to his or her intended usage

- Consider our example of a truthful client i and untruthful client j, both of whom request one share
- $T_i(0|r_i = 1) = T_j(0|r_j = 1) = 1$, so lender knows both clients will trivially use at least zero units
- $T_i(1|r_i = 1) = 1 > T_j(1|r_j = 1) = \frac{2}{3}$, so lender has **higher confidence** that client i will use at least one unit compared to client j

Table: Sample Truthful Distribution

u_{it} \ r_{it}	0	1	2
0	$\frac{1}{3}$	0	0
1	0	$\frac{1}{3}$	0
2	0	0	$\frac{1}{3}$

Table: Sample Untruthful Distribution

u_{jt} \ r_{jt}	0	1	2
0	$\frac{1}{9}$	$\frac{1}{9}$	$\frac{1}{9}$
1	0	$\frac{1}{6}$	$\frac{1}{6}$
2	0	0	$\frac{1}{3}$

Dominant-Strategy Truthfulness

So, if lender has only one stock to allocate, it will go to the truthful client i

Auction Formulation

- We now seek to understand situations in which clients have **privacy concerns** and possibly an **adaptive request strategy**
- First, we re-conceptualize the problem of computing the optimal allocation for the lender given known distributions Q_i as computing the **social welfare maximizing allocation** with respect to a set of valuation functions for each client i
- We then give an algorithm that uses an **ascending price auction formulation** to compute an **approximately optimal allocation**, with can be adapted to satisfy **(joint) differential privacy**

Auction Setting

- Consider a more general setting in which V identical units of a good are being sold to n bidders
- Each bidder has arbitrary decreasing marginal valuation function for up to U units of each good
- **Goal**: Wish to find welfare maximizing allocation
- Note: We can map our problem onto this setting as follows:
  - For each agent i who requests r_i shares and has a posterior demand distribution $Q_i(u_i|r_i)$, we define the valuation function for agent i as a function of the quantity of the good they receive:

$$v_i(s) = \frac{1}{U}\mathbb{E}_{u\sim Q_i(u_i|r_i)}\min(s, u_i) = \sum_{j=1}^{s} \Pr_{u\sim Q_i(u_i|r_i)}[u \geq j]$$

  - Given an allocation $S = (s_1, \ldots, s_n)$ of V shares to n clients, we define the total social welfare to be $v(S) = \frac{1}{V}\sum_i v_i(s_i)$

Auction Setting

- Ascending price auction works by sequentially allowing bidders to claim an additional unit of the good if the current price is below their specified marginal utility for that unit
- Price increments by α after every V bids
- Auction terminates when there are no more bids

Auction Rule

Algorithm 2 Auction Rule

Input: $\alpha > 0, n, \{v_i\}_{i\in[n]}, U, V$ ▷ valuations $v_i : [U] \to [0,1]$ satisfy DMR property
Output: feasible allocation S.

```
procedure AUCTION(α, U, V)
    Initialize array S of length n, S[i] ← 0∀i        ▷ goods currently allocated to player i
    Initialize B ← n, T_B ← 0                          ▷ bids in current round, total bids
    Set the price p = 0, m = 1                         ▷ m is index of good currently being allocated
    while B ≠ 0 do                                     ▷ terminate if there are 0 bids in the round
        B ← 0
        for i = 1 . . . n do
            Let Δ_i = v_i(S[i] + 1) − v_i(S[i])        ▷ marginal utility of additional good
            if Δ_i ≥ p then
                B ← B + 1, S[i] ← S[i] + 1, m ← (m + 1) (mod V)
                S[i_m] ← S[i_m] − 1                    ▷ i_m is player holding good m
                if T_B (mod V) = 0 then                ▷ increment price every V bids
                    p ← p + α
    return S
```

Auction Rule Guarantee

Theorem

Auction(V, α, U) terminates after at most $\frac{V}{\alpha} + 1$ rounds. At termination, S constitutes an $\frac{\alpha V}{n}$-optimal allocation:

$$v(S) \geq \max_{S'} v(S') - \frac{\alpha V}{n}$$

Joint Differential Privacy

Definition (Joint Differential Privacy)

A mechanism $A : \mathcal{X}^n \to \mathcal{O}^n$ is (ϵ, δ)-jointly differentially private if for every i, every pair of i-neighboring datasets r, r', and for every subset $S_{-i} \subset \mathcal{O}^{n-1}$ of outputs corresponding to agents other than i:

$$\Pr[A(r)_{-i} \in S_{-i}] \leq \exp(\epsilon) \Pr[A(r')_{-i} \in S_{-i}] + \delta$$

If $\delta = 0$, we say A satisfies ϵ joint differential privacy (JDP).

- Here, $\mathcal{X}^n$ is dataset domain and $\mathcal{O}^n$ is set of n outputs (one per agent)
- Two datasets are i-neighboring if they differ in only the report of agent i
- Intuitively, this prevents adversaries from learning too much about agent i by observing the allocation to agents other than i (or from all other agents colluding)

Private Auction Formulation

We modify *Auction* as follows to make it jointly differentially private:

1. Running count T_B of total number of bids placed so far is computed approximately using a differentially private estimator – since price at each round is computed purely as a function of T_B, price trajectory is differentially private as well.
2. Rather than terminating when $\mathbf{B} = 0$, the algorithm terminates when $\mathbf{B} < \rho n$ (early stopping) for accuracy parameter ρ. This limits the maximum number of times any single buyer can place a bid, allowing us to bound the error of the differentially private bid count.
3. Rather than running the auction with a supply of V shares, we run the auction with a supply of $V - E$ shares, where E corresponds to the maximum error of our differentially private bid counter; this ensures that our computed allocation (which now may over or under allocate with respect to its target supply) is always feasible.

Private Auction Guarantees

Private Auction obtains the following results:

- For sufficiently large auctions, e.g. n sufficiently large relative to V, ϵ, we can **achieve privacy** while still outputting a **high-quality allocation** (near optimal welfare).
- Private auction remains **approximately dominant-strategy truthful**.

Allocation Mechanism

Finally, we form an approximately optimal and approximately private allocation mechanism that can handle **adaptive strategies** by clients

Definition

An allocation mechanism $\mathcal{A}$ maps the requests $r_t = (r_{it})$ at time t and the history H_t to allocations of shares: $\mathcal{A}(r_{1t}, \ldots, r_{nt}; H_t) = S_t$.

- Now, each client i has the freedom to (adaptively) choose an arbitrary mapping $L_i^t : \mathcal{H}_t^i \times [U] \to [U]$ that maps the realized history and demand H_t, u_{it} respectively, to a request r_{it}.
- The utility of client i is defined as:

$$v_{\mathcal{A}}^i(L_1^i, \ldots, L_T^i) = \sum_{t=1}^{T} \mathbb{E}[v_i(\mathcal{A}(r_{it}, r_{-it}; H_t))],$$

- Lender's utility now incorporates clients' histories:

$$v(A) = \sum_t \sum_i \mathbb{E}_{u_{it} \sim Q_{it}(u_{it} | r_{it}, H_t), A} \min(A(r_1, \ldots, r_n; Q_1, \ldots, Q_n)_i, u_{it})$$

Private Allocation Mechanism

Algorithm 3 Greedy Private Mechanism

procedure $\mathcal{A}$(Utility distributions $U_i \in \Delta([U])$ for n clients, V shares to allocate at each of T rounds, $\textsc{PrivAuc}, \epsilon, \alpha$)

  for $t = 1 \ldots T$ **do**

    for i $= 1 \ldots$ n **do**

      Client i draws $u_{it} \sim U_i$

      Client i picks request distribution $Q_{it} = L_t^i(\mathcal{H}_t^i, u_{it})$

      Client i draws $r_{it} \sim Q_{it}$, and submits r_{it}

    $\mathcal{A}$ updates its estimates $\hat{Q}_i(r_{it}) = \mathbf{1}_{r_{it}}$

    $\mathcal{A}$ computes allocation $S_t = \textsc{PrivAuc}(\hat{Q}_1(r_{1t}), \ldots \hat{Q}_t(r_{nt}), \epsilon, \alpha)$

    $\mathcal{A}$ observes the executed shares $v_i(S_t)$ for each client

    $\mathcal{A}$ updates its estimates of the conditionals $\hat{Q}_i(r_{it})$

    $\mathcal{A}$ updates the history: $H_{t+1} = H_t \cup (r_{it}, s_{it}, v_i(S_t))_{i=1}^n$

Approximate Optimality and Truthfulness

Joint differential privacy in our allocation mechanism enforces **truthfulness** as an approximately dominant strategy and guarantees **near optimality**

Theorem: Let A be a private auction with appropriate values of U, V, ϵ and ρ such that A is $(\epsilon', \beta/T)$-JDP with $\epsilon' = \tilde{O}(\epsilon/\sqrt{T})$ and outputs S such that $E[V(S)] \geq (1-\rho)OPT_V - \rho$. Take β, ρ such that $\sqrt{\beta + (1-\beta)\rho} \leq \beta^2/T$. Then for a $(1-\beta)$ fraction of the n clients i, let L^t_{i*} denote the truthful strategies, and let L^t_i be any other set of strategies. Then a private greedy allocation rule for the private auction satisfies:

$$v_i(L^1_i, \ldots, L^n_i) \leq e^{2\epsilon} v_i(L^1_{i*}, \ldots, L^n_{i*}) + 2\beta UT + e^{\epsilon}\frac{\beta^2}{1-\beta^2/T}$$

$$v_A(L^t_{i*}) \geq (1-\rho)OPT_V - \rho T,$$

where OPT_V denotes the lender's optimal utility.

Summary

1. Without privacy constraints, we construct an optimal greedy allocation for which truthfulness is a dominant strategy
2. In order to guarantee clients an appropriate notion of privacy, we reformulate the allocation rule as an ascending price auction in which clients cannot collude to infer too much information about any one bidder
3. Finally, we expand this into an allocation mechanism that can handle arbitrary and adaptive client request strategies while still providing privacy and near optimality and incentivizing truthfulness

Selected References

Ekkehart Boehmer and Juan Wu (2013), Short selling and the price discovery process, *Review of Financial Studies*.

Hubert Chan, Elaine Shi,and Dawn Song (2011), Private and continual release of statistics, *ACM Trans. Inf. Syst. Secur.*

Cynthia Dwork, Guy N. Rothblum, and Salil Vadhan (2010), Boosting and differential privacy, *Proceedings of the 2010 IEEE 51st Annual Symposium on Foundations of Computer Science, FOCS 10* 51-–60.

Kuzman Ganchev, Michael Kearns, Yuriy Nevmyvaka, and Jennifer Wortman Vaughan (2010), Censored Exploration and the Dark Pool Problem, *Proceedings of the Twenty-Fifth Conference on Uncertainty in Artificial Intelligence* 185–194.

Arpita Ghosh and Aaron Roth (2015), Selling privacy at auction, *Games and Economic Behavior* 91:334–346.

Justin Hsu, Zhiyi Huang, Aaron Roth, Tim Roughgarden, and Zhiwei Steven Wu (2014), Private matchings and allocations, *Proceedings of the Forty-sixth Annual ACM Symposium on Theory of Computing* 21–30.

Michael Kearns, Mallesh Pai, Aaron Roth, and Jonathan Ullman (2014), Mechanism design in large games: Incentives and privacy, *Proceedings of the 5th conference on Innovations in theoretical computer science* 403–410.

aehoon Lee and Sang-gyung Jun, After-hours block trading, short sales, and information leakage: Evidence from Korea, *The Journal of Applied Business Research* 33(2).

lexander S, Kelso and Vincent Crawford (1982), Job matching, coalition formation, and gross substitutes, *Econometrica* 50(6):1483–1504.

Hanwen Sun and Shuxing Yin (2017), Information leakage in family firms: Evidence from short selling around insider sales, *The Journal Corporate Finance*.