Welfare Maximization with Limited Interaction

Noga Alon, Noam Nisan, Ran Raz, Omri Weinstein

Presented by Pavel Veselý

Combinatorics and Graph Theory PhD Seminar, 2015/16, MFF UK

What is welfare? A large bipartite matching between players (bidders) and items (goods)!

Communication model

There is an bipartite graph G between a set of n players and a set of m items. Each player knows only a subset of items adjacent to him. There is a referee (central planner) that is supposed to compute a matching as large as possible, but cannot see G at all. The players communicate among each other using the following model.

D(Multiparty communication model with a shared blackboard): Players communicate in a fixed number of rounds r using a protocol π. In each round each player writes at most ℓ bits on the blackboard and they do it *simultaneously*.

If π is deterministic, then each message of a player depends only on the private input of the player and the content of the blackboard from previous rounds. In a randomized π the message may further depend on some random bits (private or public).

After the end of the r-th round, referee computes a matching M based *only* on the content of the blackboard (and some random bits if π is randomized). Referee may output illegal pairs, i.e., pairs that are not edges in G.

Let $M_\pi(G)$ be the output of protocol π on graph G. Then the size of the computed matching is $|M_\pi(G) \cap E(G)|$.

D(Approximate Matchings): We say that a protocol π computes an α-*approximate matching* for G ($\alpha \geq 1$) if $|M_\pi(G) \cap E(G)| \geq \frac{1}{\alpha} \cdot |M(G)|$ where $M(G)$ is a maximum matching in G.

Similarly, when the input graph G is distributed according to some distribution μ, we say that the *approximation ratio* of π is α if

$$\mathop{\mathbb{E}}_{G\sim\mu}[|M_\pi(G) \cap E(G)|] \geq \frac{1}{\alpha} \cdot \mathop{\mathbb{E}}_{G\sim\mu}[|M(G)|].$$

Parameters: $r = \#$ of round, and $\ell = \#$ of bits in a message (by a players in one round).

Result

T: For every $r \geq 1$, there exist a distribution μ_r such that the approximation ratio of any (deterministic or randomized) protocol is $\Omega(n^{1/5^{r+1}})$ if $\ell \leq n^{1/5^{r+1}}$.

By averaging (Yao's) principle, we may consider deterministic protocols only.

Hard distribution μ_r

D(Recursive definition of μ_r): Fix some ℓ. For $r = 0$, G^0 consists of a set of n_0 players $U^0 = \{b_1, \ldots, b_{n_0}\}$ and a set of m_0 items $V^0 = \{j_1 \ldots, j_{m_0}\}$, such that $n_0 = m_0 = \ell^5$. E^0 is then obtained by selecting a random permutation $\sigma \in_R S_{\ell^5}$ and connecting $(b_i, j_{\sigma(i)})$ by an edge.

For any $r \geq 0$, the distribution μ_{r+1} over $G^{r+1} = (U^{r+1}, V^{r+1}, E^{r+1})$ is defined as follows:

Vertices:

- The set of players is $U^{r+1} := \bigcup_{i=1}^{n_r^4} B_i$ where $|B_i| = n_r$. Thus, $n_{r+1} = n_r^5$.
- The set of items is $V^{r+1} := \bigcup_{j=1}^{n_r^4 + \ell \cdot n_r^2} T_j$ where $|T_j| = m_r$. Thus, $m_{r+1} = (n_r^4 + \ell \cdot n_r^2) \cdot m_r$.

Edges:

- Let d_r be the degree of each player in the graph G^r (it is the same for all).
- First choose $\ell \cdot n_r^2$ random indices $\{a_1, a_2, \ldots a_{\ell \cdot n_r^2}\}$ from $[n_r^4 + \ell \cdot n_r^2]$, and a random permutation $\sigma : [n_r^4] \longrightarrow [n_r^4 + \ell \cdot n_r^2] \setminus \{a_1, a_2, \ldots a_{\ell \cdot n_r^2}\}$.
- Each player $u \in B_i$ is connected to d_r random items *in each one* of the blocks $T_{a_1}, T_{a_2}, \ldots, T_{a_{\ell \cdot n_r^2}}$, using *independent* randomness for each of the blocks and for each player.
- The entire block B_i is further connected to the entire block $T_{\sigma(i)}$ *using an independent copy of the distribution μ_r.*

Main theorem

Since a graph generated by μ_r has a perfect matching, it suffices to prove the following:

T(Main): For every $r \geq 0$ expected size of matching produced by an r-round protocol π under distribution μ_r is at most $5n_r^{1-1/5^{r+1}}$.

D: The ℓ_1 (statistical) distance between two distributions in the same probability space is denoted $|\mu - \nu| := \frac{1}{2} \cdot \sum_a |\mu(a) - \nu(a)|$

Fact: Let μ and ν be two probability distributions over a non-negative random variable X, whose value is bounded by X_{max}. Then $\mathbb{E}_\nu[X] \leq \mathbb{E}_\mu[X] + |\mu - \nu| \cdot X_{max}$.

Notation

- For a vector random variable $X = X_1 X_2 \ldots X_s$, we use the shorthands $X_{\leq i}$ and X_{-i} to denote $X_1 X_2 \ldots X_i$ and $X_1 X_2 \ldots X_{i-1}, X_{i+1}, \ldots \ldots X_s$ respectively
- Each block B_i of players is connected to exactly $\ell \cdot n_r^2 + 1$ blocks of items whose indices we denote by
$$\mathcal{I}_i := \{\sigma(i), a_1, a_2, \ldots a_{\ell \cdot n_r^2}\}.$$
- For each B_i, let $\tau_i : \mathcal{I}_i \longrightarrow [\ell \cdot n_r^2 + 1]$ be the bijection that maps any index in $\mathcal{I}_i$ to its location in the sorted list of $\mathcal{I}_i$ (i.e., $\tau_i^{-1}(1)$ is the smallest index in $\mathcal{I}_i$, $\tau_i^{-1}(2)$ is the second smallest index in $\mathcal{I}_i$ and so forth).
- G_j^i is the (induced) subgraph of $G = G^{r+1}$ on the sets $(B_i, T_{\tau_i^{-1}(j)})$, for each $j \in [\ell \cdot n_r^2 + 1]$.
- For a player $u \in B_i$, let $G_j^u = (u, T_{\tau_i^{-1}(j)})$ denote the (induced) subgraph of G on the sets $(u, T_{\tau_i^{-1}(j)})$.
- Let $J_i := \tau_i(\sigma(i))$ denote the index (in $\mathcal{I}_i$) of the "hidden graph" $G_{J_i}^i = (B_i, T_{\sigma(i)})$. For brevity we write $G(J_i) := G_{J_i}^i$.
- We use the shorthands $\mathbf{J} := J_1, \ldots, J_{n_r^4}$ and $\mathcal{I} := \mathcal{I}_1, \mathcal{I}_2, \ldots, \mathcal{I}_{n_r^4}$.
- Let $M_{B_i} = M_{B_i}^1 M_{B_i}^2, \ldots, M_{B_i}^{n_r}$ denote the (concatenated) messages sent by all of the players in a block B_i in the first round of π
- Let $\psi_r^i := (G(J_i) \mid M_{B_i} = m_{B_i}, J_i = j_i, \mathcal{I}_i)$ denote the distribution of the "hidden graph" $G(J_i)$ conditioned on $M_{B_i}, \mathcal{I}_i$ and J_i.
- For every block B_i and every player $u \in B_i$, let
$$G(T_i) := \left(B_i, \bigcup_{j=1}^{\ell \cdot n_r^2} T_{a_j}\right) \quad , \quad G_T^u := \left(u, \bigcup_{j=1}^{\ell \cdot n_r^2} T_{a_j}\right)$$
denote the induced subgraph on the block B_i (on the player $u \in B_i$) and all "fooling blocks" respectively.
- For any subset $S \subseteq [n_r^4]$, we write $G(T_S) := \left(\bigcup_{i \in S} B_i, \bigcup_{j=1}^{\ell \cdot n_r^2} T_{a_j}\right)$ and use the convention $\mathbf{T} := T_{[n_r^4]}$.
- In what follows, $G(\mathbf{J}) := G(J_1) G(J_2) \ldots G(J_{n_r^4})$ denotes the (concatenation of the) "hidden" graphs.

Three key lemmas

L(Conditional subgraph decomposition):

1. $$((G_T^1, G_T^2, \ldots, G_T^{n_r}) \mid M_1, G(J_1), \mathbf{J}, \mathcal{I}) \sim \bigtimes_{u \in B_1} (G_T^u | M_1, G_{J_1}^u, \mathbf{J}, \mathcal{I}).$$
2. $$(G(\mathbf{J}), G(\mathbf{T}) \mid M_1, \mathbf{J}, \mathcal{I}) \sim \bigtimes_{i \in [n_r^4]} (G(J_i) G(T_i) \mid M_{B_i}, \mathbf{J}, \mathcal{I}).$$

L(ψ_r^i and μ_r are close): For every $i \in [n_r^4]$,

$$\mathop{\mathbb{E}}_{m_{B_i}, \mathcal{I}_i, j_i} \left[|\psi_r^i - \mu_r|\right] \leq \sqrt{\frac{1}{n_r}}.$$

L(r-round Embedding):

$$\mathop{\mathbb{E}}_{\substack{m_1 \\ \mathbf{J}, \mathcal{I}}} \; \mathop{\mathbb{E}}_{\substack{G(J_1) \sim \mu_r \\ G| \; (G(J_1), m_1, \mathbf{J}, \mathcal{I})}} \left[N_{\pi|m_1}(G, G(J_1))\right] \leq 5n_r \cdot \left(\sum_{k=0}^{r-1} \Delta_k^{1/2}\right) + 1.$$

A bit of information theory

D(Relative entropy): For two distributions μ and ν in the same probability space, the *Kullback-Leiber* divergence (or relative entropy) between μ and ν is defined as

$$\mathbb{D}\left(\mu(a) \| \nu(a)\right) := \mathop{\mathbb{E}}_{a \sim \mu}\left[\log \frac{\mu(a)}{\nu(a)}\right]. \tag{1}$$

L(Pinsker's inequality): For any two distributions μ and ν,

$$|\mu(a) - \nu(a)|^2 \leq \frac{1}{2} \cdot \mathbb{D}\left(\mu(a) \| \nu(a)\right).$$

D(Conditional Mutual Information): Let A, B, C be jointly distributed random variables. The *Mutual Information* between A and B conditioned on C is

$$I(A;B|C) := \mathop{\mathbb{E}}_{\mu(ca)} \mathbb{D}\left(\mu(b|ac) \| \mu(b|c)\right) = \sum_{a,b,c} \mu(abc) \log \frac{\mu(a|bc)}{\mu(a|c)}.$$

Fact: $I(A;C|D) \leq H(A|D) \leq H(A) \leq \log|\mathrm{supp}(A)|$,

Fact(Chain rule for mutual information): Let A, B, C, D be jointly distributed random variables. Then $I(AB;C|D) = I(A;C|D) + I(B;C|AD)$.

Fact(Conditioning on independent variables increases information): Let A, B, C, D be jointly distributed random variables. If $I(A;D|C) = 0$, then it holds that $I(A;B|C) \leq I(A;B|CD)$.

Fact: Let A, B, C, D be jointly distributed random variables such that $I(B;D|AC) = 0$. Then it holds that $I(A;B|C) \geq I(A;B|CD)$.

Fact(Data processing inequality): Let $X \to Y \to Z$ be a Markov chain ($I(X;Z|Y) = 0$). Then $I(X;Z) \leq I(X;Y)$.

An important special case is when Z is a deterministic function of Y.