Budget-feasible Mechanisms for Representing Groups of Agents Proportionally

Xiang Liu[1] , **Hau Chan**[2] , **Minming Li**[3] and **Weiwei Wu**[1,*]
[1]Southeast University, China
[2]University of Nebraska–Lincoln, Nebraska, USA
[3]City University of Hong Kong, Hong Kong SAR, China
xiangliu@seu.edu.cn, hchan3@unl.edu, minming.li@cityu.edu.hk, weiweiwu@seu.edu.cn

Abstract

In this paper, we consider the problem of designing budget-feasible mechanisms for selecting agents with private costs from various groups to ensure proportional representation, where the minimum proportion of the selected agents from each group is maximized. Depending on agents' membership in the groups, we consider two main models: single group setting where each agent belongs to only one group, and multiple group setting where each agent may belong to multiple groups. We propose novel budget-feasible proportion-representative mechanisms for these models, which can select representative agents from different groups. The proposed mechanisms guarantee theoretical properties of individual rationality, budget-feasibility, truthfulness, and approximation performance on proportional representation.

1 Introduction

Selecting a proper number of agents from each group to fairly represent the population of each group has received increased attention in recent years. For instance, selecting a committee consisting of members from different groups or hiring a set of workers of diverse attributes require making selection decisions on a given set of population (*e.g.*, see [Bredereck *et al.*, 2018; Cahuc and Postel-Vinay, 2002; Lang and Manove, 2011]). Fair representations also can be applied to the context of political poll or survey sampling in which the organizer wishes to obtain a diverse set of responses from various groups of populations [Bradburn *et al.*, 2004; Jackson *et al.*, 2020]. In fact, the consequences of inadequate group representation can result in inaccurate accounts/analysis (*e.g.*, inaccurate poll predictions due to lacks of representative samples [Jackson *et al.*, 2020]) and discrimination (*e.g.*, group discrimination when hiring workers [Lang and Manove, 2011]) in various domains.

In addition to the challenge of achieving fair representation, in many settings, there is an inherited private agent cost associated with selecting each agent (*e.g.*, salary in job hiring or cost for conducting the survey) and the cost is internal or not visible to the social planner. Ideally, the planner elicits cost information from the agents, determines the agents to select, and derives appropriate compensation or payment to the selected agents. However, the agents can be strategic and do not necessarily report their true cost. As a result, the social planner must take into account whom to select to represent groups when taking costs into consideration to ensure that the total payment to all agents does not exceed the available budget (*e.g.*, the budget for hiring or conducting studies).

The problem in hand can be cast naturally into a budget-feasible mechanism design setting [Singer, 2010] where the social planner seeks to design a computationally efficient mechanism that elicits truthful cost information from the agents, selects representative agents to represent each group fairly, and ensures the total payment to the agents is no more than the budget. More specifically, *can we design a budget-feasible mechanism which fairly selects agents from different groups and guarantees desirable economic properties, while guaranteeing a bounded total payment from the planner?*

Our Contribution. We consider the problem of designing budget-feasible mechanisms for representing groups of agents proportionally satisfying standard properties (*i.e.*, individual rationality, budget feasibility, and truthfulness). We first adopt and formulate our objective to select agents that maximizes the minimum proportion ratio of the selected agents from each group, based on a well-studied notion of proportional representation (*e.g.*, in electoral systems).

We consider two general models depending on whether each agent belongs to (1) one group or (2) multiple groups (in which the agents can be (2a) counted exactly once or (2b) multi-counted) and design proportional representation budget-feasible mechanisms for the models under our objective. The proposed mechanisms guarantee desirable theoretical properties including budget feasibility, individual rationality, truthfulness, and approximation guarantee.

In particular, for (1), we construct a novel greedy mechanism that considers all possible proportion ratios and appropriate payment schemes that select agents from each group satisfying the ratios and ensuring budget feasibility. The proposed mechanism achieves approximation performance that depends on the size of the largest and smallest groups. Moreover, we show the asymptotic matching lower bound that no budget-feasible proportion-representative mechanisms can achieve better performance asymptotically.

*Corresponding author

For the multiple group setting (2a) or (2b), we construct a novel mechanism that leverages the Max-Flow algorithm to test the proportional representation in the maximum matching under a given agent set, in which we can find the candidate agent set with the greatest proportional representation within the budget constraint. We then apply the minimum weight matching to identify the final selected agents from candidate agents, whereby the tested maximum proportional representation can be obtained and the corresponding payment determined. The designed mechanisms in this setting can also achieve approximation performance that depends on the size of groups.

Related Work. Since the seminal work of [Singer, 2010], many research efforts have been invested to design budget-feasible mechanisms for various planner's valuation functions. [Chen *et al.*, 2011] further develop improved mechanisms with better approximation ratio for the submodular value function, while [Amanatidis *et al.*, 2017] consider symmetric submodular valuations, a prominent class of non-monotone submodular functions. [Anari *et al.*, 2014] design a constant-approximation budget-feasible mechanism for large markets where sellers' costs are far less than the buyer's budget and show that it is impossible to achieve bounded approximation ratio without large market assumption when sellers' items are divisible. [Singer and Mittal, 2013] focus on designing pricing mechanisms with the objective to maximize the number of tasks while guaranteeing budget feasibility. These mechanisms in existing literature do not perform well for our settings directly as they do not consider groups and ensure proportional representation. Several works take the group attributes into account and consider the diversity fairness when designing auction mechanisms [Ilvento *et al.*, 2020; Kuo *et al.*, 2020; Chawla and Jagadeesan, 2020]. However, they all ignore the planner's budget constraint and cannot guarantee the budget-feasibility. In this work, we want to design proportion-representative mechanisms while ensuring the budget constraint.

Rather than the incentive mechanism design setting addressed in this paper, we note that proportional representation has been studied from the optimization or algorithmic perspective in various areas such as voting and electoral systems. For example, [Procaccia *et al.*, 2008] focus on analyzing the complexity of achieving proportional representation. [Buisseret and Prato, 2020] consider the voter preferences in proportional representation systems to understand the candidate selection and behavior. In addition, there are works considering diversity fairness in matching/allocation problems [Benabbou *et al.*, 2020; Ahmadi *et al.*, 2019].

2 Preliminaries

In this section, we define the proportion-representative selection settings and the desirable properties of the mechanisms.

2.1 The Model

We consider a scenario with a planner a and a set of n agents $S = \{s_1, s_2, ..., s_n\}$. The agents have group attributes, specifying one or more groups the agent belongs to based on, *e.g.*, genders, ages, ethnicities, regions and educational levels. The agents are divided into m groups $G = \{G_1, G_2, ..., G_m\}$. Each group G_j is a subset of S, *i.e.*, $\emptyset \neq G_j \subseteq S$, and $G_1 \cup G_2 \cup \cdots \cup G_m = S$. Let $\mathcal{G}(s_i)$ denote the set of groups that agent s_i belongs to. The agents are to be selected by the planner for proportional representation.

The planner has a budget $B \in \mathbb{R}_+$ and each agent s_i has a private cost $c_i \in \mathbb{R}_+$ (*e.g.*, her required cost for time, privacy or fees) when selected to represent her group. We use $C = \{c_1, c_2, ..., c_n\}$ to denote agents' costs. Let C_{-i} denote all costs except s_i's cost c_i. Let n_j be the total number of agents in group G_j, *i.e.*, $|G_j| = n_j$. Denote by n_{min} and n_{max} the minimum and maximum total number of agents among all the groups respectively, *i.e.*, $n_{min} = \min_{1\leq j\leq m} n_j, n_{max} = \max_{1\leq j\leq m} n_j$. The agents may act strategically to maximize their own utilities by misreporting their costs. Each agent bids a cost b_i that may be different from her real cost c_i in order to maximize her utility (defined below). Let $b = \{b_1, b_2, ..., b_n\}$ denote agents' bid profile and b_{-i} denote all bids except s_i's bid b_i. We sometime use (b_i, b_{-i}) to represent b as to highlight s_i's bid.

2.2 The Mechanism

A mechanism $M = (\mathcal{X}, \mathcal{P})$ consists of an allocation rule $\mathcal{X}$ deciding the selected agents (who are chosen by the planner) and a payment scheme $\mathcal{P}$ deciding the payment to each agent. Denote by S_w the selected agent set. The allocation function $\mathcal{X}$ maps a set of bids b to the selected agent set $S_w = \mathcal{X}(b_1, ..., b_n) \subseteq S$. We use $x_i \in \{0, 1\}$ to indicate whether agent s_i is chosen by the planner and p_i to denote the payment to agent s_i. Let $X = \{x_1, x_2, ..., x_n\}$ and $P = \{p_1, p_2, ..., p_n\}$ denote the allocation and payment profile, respectively. Given a mechanism M, the utility of agent s_i is defined as the difference between the payment she receives and her true cost, *i.e.*,

$$u_i(b_i, b_{-i}, M) = p_i - x_i \cdot c_i. \tag{1}$$

We consider both the single group setting problem (SGP) where each agent only belongs to one group, and the multiple group setting problem (MGP) where agents may belong to multiple groups. Let Q_j denote the number of selected agents in group G_j. Next, we define two main group setting models.

Single Group Setting Problem (SGP): Since each agent belongs to only one group, we have $|\mathcal{G}(s_i)| = 1$. Then, the number of selected agents in group G_j is $Q_j = \sum_{s_i \in G_j} x_i$.

Multiple Group Setting Problem (MGP): In this setting, an individual agent may belong to multiple groups, *i.e.*, $1 \leq |\mathcal{G}(s_i)| \leq m$. Depending on whether each selected agent can be counted into all groups, we further consider two cases: Single Counting (MGP-SC) case where a selected agent is counted just once in one of the groups she belongs to and Multiple Counting (MGP-MC) case where a selected agent is counted in all groups she belongs to. For example, when forming a committee, the agent selected can only represent one of groups to which she belongs, or when she is selected, all the groups to which the agent belongs are happy.

(1) MGP-SC: Each selected agent is only included in the selected agents of the group she is matched to. Let $x_{ij} = 1$ indicate that agent s_i is matched to group $G_j \in \mathcal{G}(s_i)$, otherwise, $x_{ij} = 0$. We also have $x_i = \sum_{j:G_j \in \mathcal{G}(s_i)} x_{ij} \leq$

1 where $x_{ij} \in \{0,1\}, \forall i \leq n$. Thus, we have $Q_j = \sum_{s_i \in G_j} x_{ij}$. **(2) MGP-MC**: Each selected agent is counted in all groups she belongs to. Thus, we have $Q_j = \sum_{s_i \in G_j} x_i$.

To obtain a proportion-representative selection of agents, we define the *selection ratio* of group G_j as $\frac{Q_j}{n_j}$, representing the ratio between the number of selected agents and the total number of agents in group G_j. As such, the mechanisms we design aim to consider the objective of maximizing the minimum selection ratio of groups, *i.e.*, $\max \min_{1 \leq j \leq m} \frac{Q_j}{n_j}$.

We consider budget-feasible proportion-representative mechanism M that has the following properties:

- **Budget feasibility.** The total payment of the planner does not exceed her budget B, *i.e.*, $\sum_{1 \leq i \leq n} p_i \leq B$.
- **Individual rationality.** The utility of each agent s_i is non-negative, *i.e.*, $u_i(c_i, b_{-i}, M) \geq 0$.
- **Truthfulness.** Each agent achieves the maximum utility by bidding her real cost, *i.e.*, $u_i(c_i, b_{-i}, M) \geq u_i(b_i, b_{-i}, M)$.
- **Computational efficiency.** The output of the mechanism can be computed in polynomial time.
- **Approximation.** Let $ALG(I)$ be the minimum selection ratio among groups of the proposed mechanism M on input instance I. We compare the output of the mechanism with the optimal achievable selection ratio when agents' costs are known in advance. We say that a mechanism is α-approximate if $ALG(I) \geq \frac{1}{\alpha} OPT(I)$ for any instance I.

3 Mechanism for Single Group Settings

We introduce a **B**udget-feasible **P**roportion-representative mechanism for the **S**ingle **G**roup setting (BPSG) below.

The main idea of Mechanism BPSG is as follows. We first generate a *virtual ratio set* which contains all possible selection ratios for each group when selecting different number of agents from this group. In order to maximize the minimum selection ratio among all groups within the budget constraint, we find all feasible virtual ratios which ensure that the selection ratio for each group does not fall below that ratio and the current total payment does not exceed the budget. Specifically, the payments for the selected agents depend on the the bids of agents after the last selected agent in each group. Among all these feasible virtual ratios, we find the maximum one as the final selection ratio for all groups.

The detail of Mechanism BPSG is shown in Algorithm 1. In order to distinguish the agents belonging to different groups, we use b_i^j to denote the i-th agent's (agent s_i^j) bid in group G_j (We sort all agents in the same group $G_{j:1 \leq j \leq m}$ in the non-decreasing order of their bids, *i.e.*, $b_1^j \leq b_2^j \leq \cdots \leq b_{n_j}^j$). Denote by p_i^j the payment for agent s_i^j. We then generate a *virtual ratio set* $\mathcal{R}$ which consists of possible selection ratios among all groups, *i.e.*,

$$\mathcal{R} = \cup_{0 \leq i \leq n_j, 1 \leq j \leq m} \{i/n_j\} \tag{2}$$

We remove duplicate elements from $\mathcal{R}$ and sort all ratios in the non-decreasing order of their values where γ_l is the l-th element in $\mathcal{R}$, *i.e.*, $\gamma_1 < \gamma_2 < \cdots < \gamma_l < \cdots < \gamma_{|\mathcal{R}|}$. Denote by r_f the *final base selection ratio* as the minimum selection ratio among groups in the final solution.

To find the final base selection ratio, we iteratively consider ratios in $\mathcal{R}$ starting with the first ratio γ_1[1]. Suppose that we are now considering ratio γ_l. Let $I_j(\gamma_l)$ denote the minimum number of agents which ensures that the selection ratio in group G_j is at least γ_l, and we thus have $I_j(\gamma_l) = \lceil \gamma_l \cdot n_j \rceil$. Specifically, Mechanism BPSG will select up to $n_j - 1$ agents from each group G_j (as to ensure truthfulness). Thus, when trying ratio γ_l, BPSG will terminate and output γ_{l-1} as the final base selection ratio if there exists group G_j with $I_j(\gamma_l) = n_j$ (line 6). If we have $I_j(\gamma_l) < n_j, \forall j \leq m$, we compute the current payment for each of the first $I_j(\gamma_l)$ agents in group G_j as the bid of agent $s^j_{I_j(\gamma_l)+1}$, *i.e.*, $p_i^j = b^j_{I_j(\gamma_l)+1}, \forall i \leq I_j(\gamma_l)$. Thus, when all groups have a selection ratio of at least γ_l, the total payment, denoted by P_{γ_l}, is

$$P_{\gamma_l} = \sum_{1 \leq j \leq m} I_j(\gamma_l) \cdot b^j_{I_j(\gamma_l)+1}. \tag{3}$$

It is easy to see that P_{γ_l} is increasing with γ_l. If $P_{\gamma_l} \leq B$, we continue to try the next ratio γ_{l+1}. Otherwise, the *final base selection ratio* is $r_f = \gamma_{l-1}$.

Once deciding the final base selection ratio, we determine the final selected agents and corresponding payments. Let k_j denote the number of selected agents in group G_j, *i.e.*, $k_j = I_j(r_f) = Q_j$. In each group G_j, the first k_j agents are selected, *i.e.*, $s_i^j \in S_w, \forall i \leq k_j$, and we have $k_j < n_j$. Then we have

$$p_i^j = \begin{cases} b^j_{k_j+1}, & if\ s_i^j \in S_w \\ 0, & otherwise. \end{cases} \tag{4}$$

Running example: We now show a running example of Mechanism BPSG. Suppose there are nine agents who can be divided into two groups G_1 and G_2. Group G_1 has four agents $G_1 = \{s_1^1, s_2^1, s_3^1, s_4^1\}$ with costs $\{1, 1.5, 3, 4\}$ and group G_2 has five agents $G_2 = \{s_1^2, s_2^2, s_3^2, s_4^2, s_5^2\}$ with costs $\{0.5, 1.5, 2, 3, 5\}$. Thus, we have $n_1 = 4$ and $n_2 = 5$. The virtual ratio set is $\mathcal{R} = \{0, 0.2, 0.25, 0.4, 0.5, 0.6, 0.75, 0.8, 1\}$. The planner has budget $B = 10$. We now try virtual ratios by starting from the first non-zero ratio 0.2:

(1) Try ratio 0.2: We have $I_1(0.2) = 1$ and $I_2(0.2) = 1$. Thus, we will select s_1^1 from G_1 and pay her 1.5, and select s_1^2 from G_2 and pay her 1.5. Then, the total payment is $3 < B = 10$ and we will try next ratio 0.25.

(2) Try ratio 0.25: We have $I_1(0.25) = 1$ and $I_2(0.25) = 2$. Thus, we will select s_1^1 from G_1 and pay her 1.5, and select s_1^2, s_2^2 from G_2 and pay each of them 2. Then, the total payment is $5.5 < 10$ and we will try next ratio 0.4.

(3) Try ratio 0.4: We have $I_1(0.4) = 2$ and $I_2(0.4) = 2$. Thus, we will select s_1^1, s_2^1 from G_1 and pay each of them 3, and select s_1^2, s_2^2 from G_2 and pay each of them 2. Then, the total payment is $10 = B$ and we will try next ratio 0.5.

(4) Try ratio 0.5: We have $I_1(0.5) = 2$ and $I_2(0.5) = 3$. Thus, we will select s_1^1, s_2^1 from G_1 and pay each of them 3,

[1]Note that the first virtual ratio in $\mathcal{R}$ is 0. At such a ratio, we select no agents and pay each agent zero.

Algorithm 1: Mechanism **BPSG**(B, b, S, G)

Input: B, b, S, G.
Output: P, S_w
1 $P \leftarrow 0, S_w \leftarrow \emptyset$;
2 Sort agents in $G_j (\forall 1 \leq j \leq m)$, in the non-decreasing order of their bids $b_1^j \leq b_2^j \leq \cdots \leq b_{n_j}^j$ and generate the virtual ratio set $\mathcal{R}$ with value sorted and indexed by γ_l's;
3 **// Determine the final base selection ratio**;
4 **for** $1 \leq l \leq |\mathcal{R}|$ **do**
5   Compute $I_j(\gamma_l) = \lceil \gamma_l \cdot n_j \rceil$ for any $1 \leq j \leq m$;
6   **if** $I_j(\gamma_l) < n_j, \forall j \leq m$ **then**
7   Compute the payment P_{γ_l} according to (3);
8   **if** $P_{\gamma_l} \leq B$ **then**
9             $l \leftarrow l + 1$;
10   **else**
11   break;
12   **end**
13   **else**
14   break;
15   **end**
16 **end**
17 $r_f \leftarrow \gamma_{l-1}$;
18 **// Agent selection and payment scheme**;
19 Add agent $s_i^j (\forall j \leq m)$ with $i \leq k_j = I_j(r_f)$ into the selected agent set S_w;
20 Decide the payments to agents according to (4);
21 **return** P, S_w

and select s_1^2, s_2^2, s_3^2 from G_2 and pay each of them 3. Then, the total payment is $15 > B = 10$.

Then, Mechanism BPSG terminates with the final base selection ratio 0.4. The selected agent set is $S_w = \{s_1^1, s_2^1, s_1^2, s_2^2\}$ with payments $p_1^1 = 3, p_2^1 = 3, p_1^2 = 2, p_2^2 = 2$, while the payments for unselected agents are zero.

Next, we analyze the performance of Mechanism BPSG.

Theorem 1. *Mechanism BPSG guarantees individual rationality, budget feasibility, and computational efficiency.*

Proof. **1) Individual rationality:** Since Mechanism BPSG is truthful (proved below), we have $b_i^j = c_i^j$ where c_i^j is the true cost of agent s_i^j. For each selected agent s_i^j, we have $b_i^j \leq b_{k_j+1}^j$ where $i \leq k_j$ in group G_j, and her payment is $b_{k_j+1}^j$ which implies that her utility is $b_{k_j+1}^j - c_i^j = b_{k_j+1}^j - b_i^j \geq 0$ which is non-negative. **2) Budget feasibility:** After determining the selection ratio, it is easy to see that the total payment is $\sum_{1 \leq j \leq m} k_j \cdot b_{k_j+1}^j \leq B$ which is no greater than the budget B. **3) Computational efficiency:** The running time of Mechanism BPSG is dominated by the sorting (line 2) and the loop in determining the final base selection ratio (line 4-15) as shown in Algorithm 1. Therefore, the total computational complexity is $O(n^2)$. This completes the proof. □

We first provide a well-known Myerson's characterization for truthful mechanisms in the single parameter domain.

Theorem 2. ***(Monotone Theorem, [Myerson, 1981])*** *In the single parameter domains, a mechanism $M = (\mathcal{X}, \mathcal{P})$ guarantees sellers' truthfulness if and only if:*

*(1) $\mathcal{X}$ **is monotone:** $\forall s_i \in S$, if $b_i \leq c_i$, then $s_i \in \mathcal{X}(c_i, C_{-i})$ implies $s_i \in \mathcal{X}(b_i, C_{-i})$ for every C_{-i};*

*(2) **winners are paid threshold payments**: the payment to each winning bidder is the critical value $\inf\{c_i : i \notin \mathcal{X}(c_i, C_{-i})\}$.*

We prove the truthfulness of Mechanism BPSG by leveraging the theorem above. Suppose that the final base selection ratio is the l-th element in virtual set $\mathcal{R}$, *i.e.*, $r_f = \gamma_l$.

Theorem 3. *Mechanism BPSG guarantees truthfulness.*

Proof (Sketch). Depending on the final base selection ratio, we consider the following two cases:

Case 1: There exists no group G_j in which when trying the next virtual ratio γ_{l+1}, we select all its agents, *i.e.*, $I_j(\gamma_{l+1}) < n_j, \forall 1 \leq j \leq m$. 1) We first show that if the selected agent s_i^j in group G_j, *i.e.*, $i \leq k_j$, reports a lower cost $b_{i'}^j < b_i^j \leq b_{k_j+1}^j$, she will still be selected. Thus, BPSG is monotone. 2) If agent s_i^j with $i \leq k_j$ reports a cost higher than her payment $b_{i'}^j > b_{k_j+1}^j$, we prove that s_i^j will not be selected with zero utility. Thus, the selected agents are paid threshold payments.

Case 2: There exists a group G_j in which when trying the next virtual ratio γ_{l+1}, we select all its agents, *i.e.*, $\exists 1 \leq j \leq m, I_j(\gamma_{l+1}) = n_j$. We similarly prove that Mechanism BPSG is monotonic and the selected agents are paid threshold payments.

Therefore, Mechanism BPSG guarantees truthfulness. □

Next, we introduce a useful property of Mechanism BPSG. Let $r_j(\gamma_l)$ denote the selection ratio of group G_j after selecting k_j agents, *i.e.*, $r_j(\gamma_l) = \frac{k_j}{n_j}$. We use r_{max} and r_{min} to denote the maximum and minimum selection ratios among groups when the final base selection ratio is γ_l, *i.e.*, $r_{max} = \max_{1 \leq j \leq m}\{r_j(\gamma_l)\}$ and $r_{min} = \min_{1 \leq j \leq m}\{r_j(\gamma_l)\}$. Specifically, we have $r_{min} = \gamma_l = r_f$ since there must exist at least one group G_j whose selection ratio is $r_j(\gamma_l) = \gamma_l$ due to the generation of $\mathcal{R}$ in (2). Denote by α the ratio between n_{max} and n_{min}, *i.e.*, $\alpha = \frac{n_{max}}{n_{min}}$.

Lemma 1. *Mechanism BPSG has the following two properties: (1) $\gamma_{h+1} - \gamma_h \leq \frac{1}{n_{max}}, \forall 1 \leq h \leq |\mathcal{R}| - 1$; (2) $r_{max} - r_{min} < \frac{1}{n_{min}}$.*

Given the above lemma, we consider the approximation guarantee of BPSG. Let ALG and OPT denote the minimum selection ratio of Mechanism BPSG and the optimal solution, respectively. Our analysis considers two separate cases. In the first case when BPSG is able to select at least one agent from each group, we show that BPSG obtains an approximation ratio that depends on the ratio between the maximum and minimum number of agents among groups. In fact, the ratio is asymptomatically tight as we showed later. If Mechanism BPSG cannot select at least one agent from each group, we show that no budget-feasible proportion-representative mechanisms can output a better solution than BPSG.

Theorem 4. *(1) Mechanism BPSG achieves $(3 + \alpha)$-approximation ratio if BPSG selects at least one agent from each group (i.e., $ALG \geq \frac{1}{n_{max}}$) where $\alpha = \frac{n_{max}}{n_{min}}$.*

(2) No budget-feasible proportion-representative mechanism $\mathcal{M}$ can achieve $ALG_{\mathcal{M}} \geq \theta$ for any $\theta > 0$ where $ALG_{\mathcal{M}}$ is the solution of $\mathcal{M}$ if Mechanism BPSG cannot select at least one agent from each group, i.e., $ALG = 0$.

Proof (Sketch). Recall that the final base selection ratio is γ_l, *i.e.*, $ALG = \gamma_l$. We have $n_j \geq 2, \forall j \leq m$ since BFSG can select at least one agent from each group. Depending on the conditions at the termination of BPSG, we consider two cases: i) There exists group G_j in which we select all agents, *i.e.*, $\exists j \leq m, I_j(\gamma_{l+1}) = n_j$, when trying the next virtual ratio γ_{l+1}. We prove that $\gamma_l \cdot n_j = n_j - 1$ and $\frac{OPT}{ALG} \leq 1/(1 - \frac{1}{n_j}) \leq 2$. ii) Next, we focus on the case that not all agents are selected from any group G_j when trying γ_{l+1}, *i.e.*, $I_j(\gamma_{l+1}) < n_j, \forall j \leq m$. Let G' denote the set of all groups that can choose exactly $r_f \cdot n_j = k_j$ agents making the selection ratio in group G_j equal to r_f, $\forall G_j \in G'$. Depending on whether the next ratio γ_{l+1} is generated from a group in G' or not, we consider the following two sub-cases:

Sub-case 1: The virtual ratio γ_{l+1} is generated from the group $G_j \notin G'$. We divide agents into two parts: $\tilde{S} = \{s_i^j | i \leq k_j, \forall G_j \notin G'\} \cup \{s_i^j | i \leq k_j + 1, \forall G_j \in G'\}$ and $S \setminus \tilde{S}$ for further analysis.

For set $\tilde{S}$: The optimal solution can select all agents in $\tilde{S}$ with cost zero in the best case and spend the budget on the remaining agents in $S \setminus \tilde{S}$. We show that after OPT chooses all agents in $\tilde{S}$, the maximum selection ratio among all groups is $r_{min} + \frac{1}{n_{min}}$, and the minimum selection ratio is γ_{l+1}.

For set $S \setminus \tilde{S}$: We show that the optimal solution can select k_j agents from the remaining agents after s_{k_j} from group G_j for any $G_j \notin G'$, and $k_j + 1$ agents from the group $G_j \in G'$ with budget B. Thus, we prove that the minimum selection ratio among groups by selecting these numbers of agents is γ_{l+1} and the maximum selection ratio is $r_{min} + \frac{1}{n_{min}}$.

Thus, by combining the set $\tilde{S}$ and $S \setminus \tilde{S}$, we have

$$OPT \leq r_{min} + \frac{1}{n_{min}} + \gamma_{l+1} \leq 2r_f + \frac{1}{n_{max}} + \frac{1}{n_{min}} \quad (5)$$

due to Lemma 1.

Sub-case 2: The virtual ratio γ_{l+1} is generated from the group $G_j \in G'$. We can similarly prove that $OPT \leq 2r_f + \frac{1}{n_{max}} + \frac{1}{n_{min}}$.

Therefore, we have $OPT \leq 2ALG + \frac{1}{n_{min}} + \frac{1}{n_{max}}$. Specifically, if $ALG \geq \frac{1}{n_{max}}$, we have $\frac{OPT}{ALG} \leq 3 + \alpha$ which means BPSG achieves $(3 + \alpha)$-approximation ratio. □

Next, we provide an asymptotic tight lower bound for all of budget-feasible proportion-representative mechanisms.

Theorem 5. *No budget-feasible proportion-representative mechanism obtains an approximation ratio better than $\Omega(\alpha)$.*

In particular, it is worth stating that this lower bound also applies to the multiple group setting.

4 Mechanisms for Multiple Group Settings

In this section, we consider the multiple group setting where each agent s_i might belong to multiple groups, *i.e.*, $1 \leq |\mathcal{G}(s_i)| \leq m$. We distinguish two sub-cases according to whether the contribution of the agent is counted only once or not: single-counting case and multiple-counting case. In the single counting case, the selection ratio in one of the groups $\mathcal{G}(s_i)$, say G_j, would increase by $\frac{1}{n_j}$ where $n_j = |G_j|$ if s_i is selected and contributes to G_j; while in the multiple counting case, the selection ratios of all groups in $\mathcal{G}(s_i)$ would increase by $\frac{1}{n_j}$ for any $G_j \in \mathcal{G}(s_i)$, once s_i is selected.

Algorithm 2: Mechanism **BPMG-S**(B, b, S, G)

Input: B, b, S, G.
Output: P, S_w

$P \leftarrow 0, S_w \leftarrow \emptyset$;
Sort all agents in the non-decreasing order of their bids b_i, *i.e.*, $b_1 \leq b_2 \leq \cdots \leq b_n$;
// **Determine the candidate agent set**;
for $1 \leq i \leq n$ **do**
  Compute $r(s_i)$ according to (6);
  $F(s_i) \leftarrow \sum_{1 \leq j \leq m} \lceil r(s_i) \cdot n_j \rceil$;
  if $b_i \cdot F(s_i) \leq B$ **then**
    $i \leftarrow i + 1$;
  else
    break;
  end
end
$k \leftarrow i - 1$;
// **Agent selection and payment scheme**;
$X, S_w \leftarrow AgentSelect(S_k)$;
The payment for $s_i \in S_w$ is $p_i = \min\{\frac{B}{F(s_k)}, b_{k+1}\}$;

4.1 Single Counting Case

In this section, we introduce a **B**udget-feasible **P**roportion-representative mechanism for the **M**ultiple **G**roup setting in **S**ingle counting case, called BPMG-S.

Intuitively, we measure the supply of agents for a given fixed payment. Given all agents who have bids lower than this payment, we can compute the solution that maximizes the minimum selection ratio among groups. This minimum ratio will increase and reflect the supply as the number of agents increases. We then try to find the payment which can maximize such a ratio and ensure the budget feasibility, simultaneously.

In detail, we first sort all the agents according to their non-decreasing order of bids $b_1 \leq b_2 \leq b_3 \leq \cdots \leq b_n$. Let $x_{ij} = 1$ indicate that agent s_i is matched to group G_j, otherwise, $x_{ij} = 0$. Agent s_i can only be matched to at most one group, *i.e.*, $\sum_{j \leq m} x_{ij} \leq 1, \forall i \leq n$. Denote by $\mathcal{S}(s_h) = \{s_i | i \leq h\}$ the set containing agents before agent b_{h+1}.

Before introducing our mechanism, we first introduce an important component, an integer programming formulation to compute a matching result under a given agent set $\mathcal{S}(s_h)$, that maximizes the minimum selection ratio among groups while ignoring the costs and budget constraint, denoted by

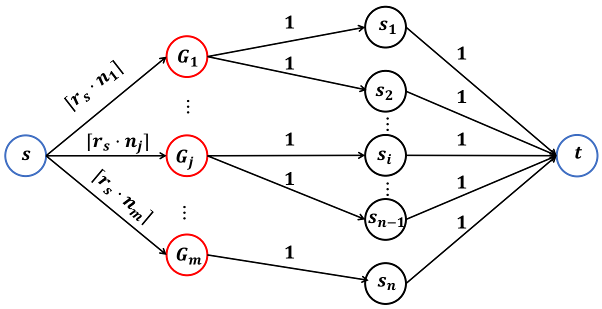

Figure 1: An example of a flow network under ratio r_s.

$ILP(s_h)$ as follows,

$$\begin{aligned} &\max \min_{1\le j\le m} \frac{\sum_{i\le h} x_{ij}}{n_j} \\ s.t., &\sum_{j\le m} x_{ij} \le 1, \forall i \le h \\ &x_{ij} \in \{0,1\}, \forall i \le h, G_j \in \mathcal{G}(S_i) \\ &x_{ij} = 0, \forall i \le h, G_j \notin \mathcal{G}(S_i) \end{aligned} \tag{6}$$

where these three conditions indicate that agents in $\mathcal{S}(s_h)$ can only be matched to at most one of the groups they belong to. Notice that the optimal solution in (6) can be computed in polynomial time by constructing Max-Flow networks as follows. In detail, given agent set $\mathcal{S}(s_h)$, all possible selection ratios for group G_j are in $\cup_{y\le|\mathcal{S}(s_h)\cap G_j|}\{\frac{y}{n_j}\}$. Thus, the optimal selection ratio in the solution of (6) must be one of the ratios in $R = \cup_{y\le|\mathcal{S}(s_h)\cap G_j|, j\le m}\{\frac{y}{n_j}\}$. Particularly, we remove duplicate elements in R. We say a ratio $r_s \in R$ is *feasible* if we can find a matching result ensuring that the selection ratio in each group is at least r_s and satisfying all conditions in (6). To find the optimal solution, we take $r_s \in R$ as an input and construct a flow network based on such a ratio. As shown in Fig. 1, we use blue, red and black circles to represent source node s/terminal node t, groups and agents, respectively. Specifically, there exists a directed edge from source s to each group G_j, an edge from each G_j to s_i if $s_i \in G_j$, and from each agent s_i to terminal node t. Each edge from s to G_j is assigned with a capacity $\lceil r_s \cdot n_j \rceil$, while the capacity of each directed edge from G_j to s_j and from s_i to t is 1. If the maximum flow on this network equals $\sum_{j\le m}\lceil r_s \cdot n_j \rceil$, then r_s is a feasible ratio. We can find the optimal solution in (6) by testing every possible input of r_s and validating its feasibility, *i.e.*, the maximum one among all feasible ratios is the desired optimal solution.

Let $X(s_h)$ denote the solution of $ILP(s_h)$, and $r(s_h)$ is the minimum selection ratio among groups under $X(s_h)$. Denote by $F(s_h)$ the total number of agents under ratio $r(s_h)$, *i.e.*,

$$F(s_h) = \sum_{1\le j\le m} \lceil r(s_h) \cdot n_j \rceil. \tag{7}$$

It is obvious that $F(s_h) \le |\mathcal{S}(s_h)|$.

Candidate agent selection. Now we are ready to select agents. We first decide on a set of candidate agents, from which we select agents. We iteratively test each agent's bid starting from the first agent's bid b_1. Suppose that we are now trying agent s_i and calculate the value of $F(s_i)$ by equation (7). If $b_i \cdot F(s_i) \le B$, we consider the next agent s_{i+1}, otherwise, we select agents from the previous $i-1$ agents $\mathcal{S}(s_{i-1})$. Assume that s_k is the last agent who satisfies $b_k \cdot F(s_k) \le B$ which implies $b_{k+1} \cdot F(s_{k+1}) > B$. After determining the value k, we define agent set $\mathcal{S}(s_k) = \{s_i : i \le k\}$ as the *candidate agent set*, where $|\mathcal{S}(s_k)| \ge F(s_k)$ follows.

Algorithm 3: Function **AgentSelect**(S, G, k)

Input: S, G, k.
Output: X, S_w

$X \leftarrow 0, S_w \leftarrow \emptyset$;
Assign each agent s_i a weight $w_i = 2^{z_i}$ where $z_i \in N_+$ is an arbitrary integer such that no two agents have the same weight, *i.e.*, $z_i \ne z_{i'}$ for any $i \ne i'$;
Compute the final allocation $X_w(s_k)$ according to (8);
$X \leftarrow X_w(s_k)$;
Add agent $s_i(\forall 1 \le i \le n)$ into S_w if $\sum_{j\le m} x_{ij} = 1$;

Agent selection and payment scheme. Now we select agents from the candidate agent set $\mathcal{S}(s_k)$. First, we assign each agent s_i a weight $w_i = 2^{z_i}$ where $z_i \in N_+$ is an arbitrary integer such that no two agents have the same weight, *i.e.*, $z_i \ne z_{i'}$ for any $i \ne i'$. Then, we use the function $AgentSelect(S, G, k)$ to select $F(s_k)$ agents from $\mathcal{S}(s_k)$. The detail of $AgentSelect(S, G, k)$ is shown in Algorithm 3. We try to find a minimum weight matching between agents and groups that can minimize the total weight of matched agents satisfying that each group G_j is matched $\lceil r(s_k) \cdot n_j \rceil$ agents from set $\mathcal{S}(s_k)$ as follows, denoted by $ILP_w(s_k)$[2],

$$\begin{aligned} &\min \sum_{1\le i\le k} \sum_{1\le j\le m} w_i x_{ij} \\ s.t., &\sum_{1\le j\le m} x_{ij} \le 1, \forall i \le k \\ &\sum_{1\le i\le k} x_{ij} = \lceil r(s_k) \cdot n_j \rceil, \forall j \le m \\ &x_{ij} \in \{0,1\}, \forall i \le k, G_j \in \mathcal{G}(S_i) \\ &x_{ij} = 0, \forall i \le k, G_j \notin \mathcal{G}(S_i). \end{aligned} \tag{8}$$

Since we assign agents exponential weights, there is a unique matching result in (8). Let $X_w(s_k)$ denote the solution of $ILP_w(s_k)$. If agent $s_{i:i\le k}$ is matched to one of the groups, *i.e.*, $\sum_{1\le j\le m} x_{ij} = 1$, then she is selected, *i.e.*, $s_i \in S_w$. The payment for each selected agent is $p_i = \min\{\frac{B}{F(s_k)}, b_{k+1}\}$ while the payments for unselected agents are zero[3].

Next we analyze the performance of Mechanism BPMG-S.

Theorem 6. *Mechanism BPMG-S guarantees individual rationality, budget feasibility and computational efficiency.*

We prove the truthfulness of Mechanism BPMG-S by showing that it satisfies Theorem 2.

[2]This problem can be similarly solved in polynomial time by constructing Max-Flow Min-Cost networks.

[3]The payment for each selected agent is $\frac{B}{F(s_k)}$ if $k = n$.

Theorem 7. *Mechanism BPMG-S guarantees truthfulness.*

Let OPT_s and ALG_s denote the optimal solution and the solution of BPMG-S, respectively. We show that BPFG-S generally achieves an approximation ratio with respect to the size of groups when at least one agent is selected in each group, while there is a small gap between BPFG-S and any other budget-feasible proportion-representative mechanism when BPFG-S cannot select at least one agent from each group.

Theorem 8. *(1) Mechanism BPMG-S achieves $(m\alpha(\alpha+2)+1)$-approximation ratio if BPMG-S can select at least one agent from each group (i.e., $ALG_s \geq \frac{1}{n_{max}}$) where $\alpha = \frac{n_{max}}{n_{min}}$.*

(2) No budget-feasible proportion-representative mechanism $\mathcal{M}'$ can obtain $ALG_{\mathcal{M}'} \geq \theta$ for any $\theta > \frac{1}{n_{min}}$ where $ALG_{\mathcal{M}'}$ is the solution of $\mathcal{M}'$ if Mechanism BPMG-S cannot select at least one agent from each group, i.e., $ALG_s = 0$.

4.2 Multiple Counting Case

In this section, we consider the multiple counting case where the selection ratio of each group G_j in $\mathcal{G}(s_i)$ can increase by $\frac{1}{n_j}$ for $G_j \in \mathcal{G}(s_i)$ when s_i is selected. We propose a modified version of Mechanism BPMG-S, called BPMG-M.

In general, Mechanism BPMG-M applies the methodological framework in Mechanism BPMG-S. We first sort all the agents according to their non-decreasing order of bids $b_1 \leq b_2 \leq b_3 \leq \cdots \leq b_n$. We then use an integer programming formulation to compute a result given an agent set $\mathcal{S}(s_h) = \{s_i : i \leq h\}$, that maximizes the minimum selection ratio among groups, denoted by $ILP^m(s_h)$ as follows,

$$\begin{aligned} &\max \min_{j \leq m} \frac{\sum_{s_i \in G_j} x_i}{n_j} \\ &s.t., x_i \in \{0,1\}, \forall i \leq h \\ &x_i = 0, \forall i > h \end{aligned} \tag{9}$$

where the two conditions mean that we only consider agents in $\mathcal{S}(s_h)$. Similar to that of (6), we can solve this problem by using the Max-Flow method. Let $X_m(s_h)$ denote the solution of $ILP^m(s_h)$, and $r_m(s_h)$ denote the minimum selection ratios under $X_m(s_h)$. We denote $F_m(s_h)$ as the total counted number of selected agents among groups by ratio $r_m(s_h)$, *i.e.*, $F_m(s_h) = \sum_{j \leq m} \lceil r_m(s_h) \cdot n_j \rceil$.

Candidate agent selection. We also decide on a set of candidate agents by iteratively testing each agent's bid starting from the first agent's bid b_1. That is, we find the last agent s_k who ensures $b_k \cdot F_m(s_k) \leq B$ and $b_{k+1} \cdot F_m(s_{k+1}) > B$. Then, the agent set $\mathcal{S}(s_k)$ is the candidate agent set.

Agent selection and payment scheme. Similar to the agent selection function $AgentSelect(S_k, G, B)$, we assign each agent s_i a weight $w_i = 2^{z_i}$ where $z_i \in N_+$ is an arbitrary integer such that no two agents have the same weight, *i.e.*, $z_i \neq z_{i'}$ for any $i \neq i'$. We try to select agents with the minimum total weight satisfying that each group G_j is selected at least $\lceil r_m(s_k) \cdot n_j \rceil$ agents from set $\mathcal{S}(s_k)$ as follows, denoted by $ILP_w^m(s_k)$,

$$\begin{aligned} &\min \sum_{i \leq k} w_i x_i \\ &s.t., \sum_{i \in G_j} x_i \geq \lceil r_m(s_k) \cdot n_j \rceil, \forall j \leq m \\ &x_i \in \{0,1\}, \forall i \leq k \\ &x_i = 0, \forall i > k \end{aligned} \tag{10}$$

Let $X_w^m(s_k)$ denote the solution of $ILP_w^m(s_k)$. If $x_i = 1$, agent s_i is selected and her payment is $p_i = \min\{\frac{B}{F_m(s_k)}, b_{k+1}\}$, otherwise, $p_i = 0$.

Theorem 9. *Mechanism BPMG-M guarantees individual rationality, budget feasibility, computational efficiency and truthfulness.*

Let ALG_m denote the solution of Mechanism BPMG-M.

Theorem 10. *(1) Mechanism BPMG-M achieves $(m\alpha(\alpha+2)+1)$-approximation ratio if BPMG-M can obtain at least one agent from each group (i.e., $ALG_m \geq \frac{1}{n_{max}}$) where $\alpha = \frac{n_{max}}{n_{min}}$.*

(2) No budget-feasible proportion-representative mechanism $\mathcal{M}''$ can obtain $ALG_{\mathcal{M}''} \geq \theta$ for any $\theta > \frac{2m}{n_{min}}$ where $ALG_{\mathcal{M}''}$ is the solution of $\mathcal{M}''$ if Mechanism BPMG-M cannot select at least one agent from each group, i.e., $ALG_m = 0$.

5 Conclusion

We consider the proportion representation budget-feasible mechanism design problem where agents may have diverse group attributes. We focus on designing budget-feasible mechanisms that can select appropriate proportions of agents from various groups satisfying individual rationality, budget feasibility, and truthfulness under several settings. For the single group setting, we propose Mechanism BPSG, which iteratively tests each virtual ratio generated by the distribution of agents across groups and finds the maximum one as the base selection ratio for all groups. For the multiple group setting, we first consider the single counting case by proposing Mechanism BPMG-S, which leverages Max-Flow solution to measure the supply of agents under a fixed payment and finds the payment which can maximize the minimum selection ratio among groups. We then extend BPMG-S to the multiple counting case.

Acknowledgements

This work was supported in part by the National Key Research and Development Program of China under grant No. 2019YFB2102200, Natural Science Foundation of China under Grant No. 61902062, 61672154, 61972086, the Project 11771365 supported by NSFC, the Postgraduate Research & Practice Innovation Program of Jiangsu Province of China (KYCX19_0089), and partially sponsored by Key Laboratory of Computer Network and Information Integration (Southeast University), Ministry of Education.

References

[Ahmadi *et al.*, 2019] Saba Ahmadi, Faez Ahmed, John P Dickerson, Mark Fuge, and Samir Khuller. An algorithm for multi-attribute diverse matching. *arXiv*, pages arXiv–1909, 2019.

[Amanatidis *et al.*, 2017] Georgios Amanatidis, Georgios Birmpas, and Evangelos Markakis. On budget-feasible mechanism design for symmetric submodular objectives. In *International Conference on Web and Internet Economics*, pages 1–15. Springer, 2017.

[Anari *et al.*, 2014] Nima Anari, Gagan Goel, and Afshin Nikzad. Mechanism design for crowdsourcing: An optimal 1-1/e competitive budget-feasible mechanism for large markets. In *2014 IEEE 55th Annual Symposium on Foundations of Computer Science*, pages 266–275. IEEE, 2014.

[Benabbou *et al.*, 2020] Nawal Benabbou, Mithun Chakraborty, Xuan-Vinh Ho, Jakub Sliwinski, and Yair Zick. The price of quota-based diversity in assignment problems. *ACM Transactions on Economics and Computation (TEAC)*, 8(3):1–32, 2020.

[Bradburn *et al.*, 2004] Norman M Bradburn, Seymour Sudman, and Brian Wansink. *Asking questions: the definitive guide to questionnaire design–for market research, political polls, and social and health questionnaires*. John Wiley & Sons, 2004.

[Bredereck *et al.*, 2018] Robert Bredereck, Piotr Faliszewski, Ayumi Igarashi, Martin Lackner, and Piotr Skowron. Multiwinner elections with diversity constraints. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 32, 2018.

[Buisseret and Prato, 2020] Peter Buisseret and Carlo Prato. Voting behavior under proportional representation. *Journal of Theoretical Politics*, 32(1):96–111, 2020.

[Cahuc and Postel-Vinay, 2002] Pierre Cahuc and Fabien Postel-Vinay. Temporary jobs, employment protection and labor market performance. *Labour economics*, 9(1):63–91, 2002.

[Chawla and Jagadeesan, 2020] Shuchi Chawla and Meena Jagadeesan. Fairness in ad auctions through inverse proportionality. *arXiv preprint arXiv:2003.13966*, 2020.

[Chen *et al.*, 2011] Ning Chen, Nick Gravin, and Pinyan Lu. On the approximability of budget feasible mechanisms. In *Proceedings of the twenty-second annual ACM-SIAM symposium on Discrete Algorithms*, pages 685–699. SIAM, 2011.

[Ilvento *et al.*, 2020] Christina Ilvento, Meena Jagadeesan, and Shuchi Chawla. Multi-category fairness in sponsored search auctions. In *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency*, pages 348–358, 2020.

[Jackson *et al.*, 2020] Natalie Jackson, Michael S Lewis-Beck, and Charles Tien. Pollster problems in the 2016 us presidential election: vote intention, vote prediction. *Quaderni dell Osservatorio elettorale QOE-IJES*, 83(1):17–28, 2020.

[Kuo *et al.*, 2020] Kevin Kuo, Anthony Ostuni, Elizabeth Horishny, Michael J Curry, Samuel Dooley, Ping-yeh Chiang, Tom Goldstein, and John P Dickerson. Proportionnet: Balancing fairness and revenue for auction design with deep learning. *arXiv preprint arXiv:2010.06398*, 2020.

[Lang and Manove, 2011] Kevin Lang and Michael Manove. Education and labor market discrimination. *American Economic Review*, 101(4):1467–96, 2011.

[Myerson, 1981] Roger B Myerson. Optimal auction design. *Mathematics of operations research*, 6(1):58–73, 1981.

[Procaccia *et al.*, 2008] Ariel D Procaccia, Jeffrey S Rosenschein, and Aviv Zohar. On the complexity of achieving proportional representation. *Social Choice and Welfare*, 30(3):353–362, 2008.

[Singer and Mittal, 2013] Yaron Singer and Manas Mittal. Pricing mechanisms for crowdsourcing markets. In *Proceedings of the 22nd international conference on World Wide Web*, pages 1157–1166, 2013.

[Singer, 2010] Yaron Singer. Budget feasible mechanisms. In *2010 IEEE 51st Annual Symposium on Foundations of Computer Science*, pages 765–774. IEEE, 2010.