УДК 519.863
JEL C61

СПРАВЕДЛИВОЕ РАСПРЕДЕЛЕНИЕ СТУДЕНТОВ ПО БЛОКАМ ДИСЦИПЛИН ПО ВЫБОРУ

Глазунова Екатерина Валерьевна, асп.

Санкт-Петербургский государственный экономический университет, ул. Садовая, 21, Санкт-Петербург, Российская Федерация, 191023; e-mail: katarina.glazunova97@inbox.ru

Предмет: процесс распределения студентов университета по дисциплинам по выбору. *Цель*: разработка алгоритма поиска справедливого распределения студентов по дисциплинам, а также разбиение множества студентов на «подгруппы» для совместного изучения дисциплин. *Дизайн исследования*: задача рассматривается как задача о поиске распределения на двустороннем рынке, где сторонами рынка являются студенты и дисциплины. В предположении о том, что предпочтения агентов стороны рынка «дисциплины» одинаковы для всех, предлагается модификация алгоритма отложенного принятия для поиска распределения. Разбиение студентов на «подгруппы» осуществляется с использованием модели смешанного программирования. *Результаты*: разработан трехэтапный алгоритм поиска распределения студентов по дисциплинам, а также разбиения студентов по «подгруппам». Проведены расчеты на полномасштабных данных, алгоритм демонстрирует корректность результатов и эффективность в терминах быстродействия.

Ключевые слова: справедливое распределение, алгоритм отложенного принятия, смешанное программирование, дисциплины по выбору.

DOI: 10.17308/meps/2078-9017/2024/6/18-32

Введение

Дано множество дисциплин, которые присутствуют в учебных планах различных образовательных программ университета, и множество учебных групп, студенты которых должны быть распределены по этим дисциплинам. В результате из студентов, которые обучаются в таких учебных группах, формируются «подгруппы», из которых, в свою очередь, составляются «подпотоки».

Следуя учебному расписанию и распределению на «подгруппы», студенты посещают практические, семинарские или лабораторные занятия по соответствующей дисциплине в составе назначенной им «подгруппы», а лекционные – в составе одного «подпотока». Остальные занятия студенты посещают в составе своих академических групп.

Каждый студент характеризуется номером академической группы, а также результатами своей успеваемости (рейтингом). Для каждой академической группы известно подмножество дисциплин по выбору, на которые студенты данной группы могут быть распределены. Каждый студент ранжирует такие дисциплины, в результате чего определяются нестрогие предпочтения для каждого студента по дисциплинам. Нестрогие предпочтения дисциплины по множеству студентов, которые могут быть назначены на дисциплину, формируются в соответствии с академической успеваемостью студентов (рейтингом).

Все множество учебных групп университета, студенты которых подлежат распределению, разбито на пулы – непересекающиеся подмножества. Разбиение на пулы проводится, исходя из различных операционных ограничений и критериев, касающихся реализации учебного процесса (аудиторный фонд, учебное расписание, число преподавателей, задействованных для данной дисциплины по выбору).

Каждая учебная группа характеризуется множеством дисциплин по выбору, на которые могут быть распределены студенты этой группы. Пулы формируются так, что студенты из групп пула, распределенные на одну дисциплину по выбору, совместно посещают лекционные занятия в составе одного «подпотока». Множество дисциплин по выбору, на которые могут быть распределены студенты учебных групп пула, формирует блок дисциплин.

Для распределения студентов по дисциплинам необходимо найти распределение студентов групп каждого пула по дисциплинам соответствующего блока.

Для каждого пула и блока дисциплин задача распределения студентов по этим дисциплинам и «подгруппам» для их изучения решается в три этапа.

На первом этапе осуществляется поиск квот дисциплин, поскольку определенные ранее оценки квот могут привести к тому, что допустимого распределения студентов не существует из-за того, что студенты переводятся с факультета на факультет или из других университетов, отчисляются, уходят или восстанавливаются из академического отпуска. Поиск квот необходимо осуществлять, минимизируя отклонения от заданных оценок, при этом сумма квот должна равняться количеству студентов.

На втором этапе происходит распределение студентов, которые, расставили предпочтения (проголосовали) по дисциплинам, допустимым для них. Данная задача может быть представлена как задача о поиске распределения на двустороннем рынке [7, 9], где агентами рынка являются студенты и дисциплины. Модели на базе двусторонних рынков широко используются в схожих задачах. Например, в работе [8] представлена точная математическая постановка поиска распределения интернов по больницам, а в работе [10] описан алгоритм назначения студентов на проекты.

В данной работе для поиска распределения используется модификация алгоритма отложенного принятия [3].

Студенты, которые не расставили свои приоритеты по дисциплинам, распределяются на следующем этапе.

На третьем этапе решается задача определения количества студентов из каждой академической группы, которые входят в заданную «подгруппу», а также распределения студентов по «подгруппам». Задача формулируется как задача математического программирования [1].

Разработана и реализована оптимизационная модель для поиска распределения студентов на основе их предпочтений по дисциплинам, рейтинга, а также с учетом ограничений, связанных с правилами формирования «подгрупп». Проведены расчеты на полномасштабных данных, которые демонстрируют корректность и эффективность разработанного подхода и его программной реализации.

Методы и результаты исследования

Постановка задачи распределения студентов по дисциплинам по выбору

Заданы пул – набор академических групп, и блок соответствующих пулу дисциплин по выбору C. Определено множество студентов S академических групп рассматриваемого пула, которые должны быть распределены по дисциплинам по выбору из рассматриваемого блока. Каждый студент $s \in S$ характеризуется номером академической группы $n_s \in N$, где N – множество всех групп пула, а также результатами своей успеваемости.

Для изучения каждой дисциплины из блока C студенты должны быть распределены по «подгруппам», а «подгруппы» объединены в «подпотоки»:

- «подгруппа» – это подмножество студентов, которое формируется из студентов, принадлежащих множеству S (из различных академических групп пула) для посещения практических, семинарских или лабораторных занятий по данной дисциплине множества C;
- «подпоток» – это подмножество «подгрупп», которые совместно будут посещать лекции по данной дисциплине.

Для каждой академической группы $n \in N$ известен ее размер $size_n$, то есть количество студентов, которые входят в нее. Для пула групп можно выделить подмножество групп $N_c \subset N$ – академические группы, студенты которых могут изучать дисциплину $c \in C$.

Обозначим подмножество студентов, которые могут быть назначены на дисциплину $c \in C$ как S_c, а множество допустимых дисциплин для студента $s \in S$ как C_s. Каждый студент $s \in S$ ранжирует дисциплины множества C_s в порядке предпочтительности. Таким образом, определены нестрогие предпочтения $P_s(C_s)$ для каждого студента $s \in S$ по множеству дисциплин $C_s \subseteq C$. По результатам успеваемости студентов множества S_c ($c \in C$) формируются нестрогие предпочтения $P_c(S_c)$ дисциплины по множеству студентов.

Для проведения расчетов по поиску распределения студентов академических групп пула по соответствующему блоку дисциплин университет определяет:

1. для каждой дисциплины $c \in C$ и пула академических групп N количество g_c «подгрупп», которые, по результатам распределения студентов, необходимо создать для обучения по данной дисциплине;

2. оценку количества студентов q_c для каждой дисциплины $c \in C$ (во все «подгруппы» дисциплины c должно быть распределено такое количество студентов, чтобы оно минимально отличалось от q_c).

Для рассматриваемого пула групп необходимо найти распределение студентов по дисциплинам и «подгруппам» таким образом, чтобы оптимизировать значение следующих критериев:

- удовлетворенность студентов и дисциплин от распределения (максимизировать);
- разнообразие студентов из разных групп в одной «подгруппе» (минимизировать).

При этом должны быть выполнены следующие требования (ограничения):

- каждый студент $s \in S$ должен быть распределен ровно в одну «подгруппу» ровно одной дисциплины;
- для каждой дисциплины $c \in C$ и множества академических групп N – количество «подгрупп» должно быть равно g_c;
- для каждой дисциплины $c \in C$ и множества академических групп N – во всех «подгруппах» количество студентов близко к значению q_c;
- для изучения дисциплины $c \in C$ формируются «подгруппы» только из студентов групп N_c;
- распределение не должно содержать «конфликты зависти» [4, 11, 12]. Это означает, что в распределении не должно быть ни одной пары (дисциплина $c \in C$ – студент $s \in S$) такой, что на дисциплину c назначен студент $s' \in S$, но с более низкой, чем у студента s, успеваемостью, то есть $s \succ_c s'$, при этом студент s распределен на дисциплину $c' \in C$ менее предпочтительную для него, чем дисциплина c, то есть $c \succ_s c'$. Таким образом, распределение соответствует правилу: «Никто ниже тебя по успеваемости не попадет на ту дисциплину, куда ты хотел, но не попал».

Алгоритм распределения студентов по дисциплинам

Поиск квот

Задача поиска квот формулируется как многокритериальная задача смешанного программирования [2].

Переменные задачи имеют вид:

- $x_{n,c}, \forall c \in C, n \in N_c$ – количество студентов, которое может быть распределено из группы n на дисциплину c;
- $y_c, \forall c \in C$ – корректировка (в большую или меньшую сторону) оценки количества студентов, распределенных для изучения дисциплины c;
- $\alpha_c, \forall c \in C$ – абсолютное отклонение количества студентов, распределенных для изучения дисциплины c, от величины q_c;
- $\overline{\alpha}$ – максимальное значение среди всех переменных типа α;
- $\underline{\alpha}$ – минимальное значение среди всех переменных типа α.

При следующих критериях оптимизации:

$$\sum_{c \in C} \alpha_c \to \min \qquad (1)$$

$$\overline{\alpha} - \underline{\alpha} \to \min \qquad (2)$$

Ограничения задачи определяются следующим образом:

$$\sum_{n \in N_c} x_{n,c} + y_c = q_c \qquad \forall c \in C \qquad (3)$$

$$\sum_{c \in C} x_{n,c} = size_n \qquad \forall n \in C \qquad (4)$$

$$\alpha_c = |y_c| \qquad \forall c \in C \qquad (5)$$

$$\overline{\alpha} \geq \alpha_c, \quad \underline{\alpha} \leq \alpha_c \qquad \forall c \in C \qquad (6)$$

$$x_{n,c} \in \mathbb{Z}_{\geq 0} \qquad \forall c \in C, \ \forall n \in N_c \qquad (7)$$

$$y_c \in \mathbb{R} \qquad \forall c \in C \qquad (8)$$

$$\alpha_c \geq 0 \qquad \forall c \in C \qquad (9)$$

$$\overline{\alpha}, \underline{\alpha} \geq 0 \qquad (10)$$

Критерий оптимизации (1) минимизирует сумму отклонений квот от значений q_c ($\forall\, c \in C$), критерий оптимизации (2) минимизирует максимальную разность между квотой и значением q_c ($\forall\, c \in C$) от минимальной, критерий обеспечивает равномерность отклонений квот от значений q_c ($\forall\, c \in C$). Многокритериальная задача решается лексикографическим методом, где при минимальной сумме отклонений от значений q_c ($\forall\, c \in C$) достигается максимальная возможная равномерность таких отклонений.

Ограничения (3) определяют, насколько будет отличаться квота от заданной оценки количества студентов для каждой дисциплины. Ограничения (4) гарантируют, что все студенты будут распределены при найденных квотах. Ограничения (5) представляют собой определение абсолютных значений отклонений квот от значений q_c ($\forall\, c \in C$), формально данные ограничения являются нелинейными, однако могут быть линеаризованы стандартными подходами. Ограничения (6) определяют минимальную и максимальную разность между найденной квотой и оценкой количества студентов для каждой дисциплины. Ограничения (7)-(10) определяют область допустимых значений переменных.

В результате квоты для каждой дисциплины $c \in C$ определяются как $q_c^* = q_c + y_c$.

Распределение проголосовавших студентов по дисциплинам

Задача поиска распределения студентов по дисциплинам может быть рассмотрена, как задача о поиске распределения на двустороннем рынке, где агентами выступают студенты и дисциплины. Необходимо найти распределение, то есть множество пар вида студент-дисциплина, определяющие, какие студенты будут изучать рассматриваемые дисциплины.

Пусть распределение $\mu \subset S \times C$ называется допустимым, если выполняется следующее:

1. для любой пары $(s, c) \in \mu$ верно, что $s \in C_s$ и $c \in C_s$ – студент распределен на допустимую для него дисциплину;

2. $\mu_s \in C$, где μ_s означает дисциплину, на которую назначен студент s – каждый студент распределен ровно на одну дисциплину;

3. $|\mu_c| = q_c^*$, где μ_c означает всех студентов, назначенных на дисциплину c – на каждую дисциплину назначено ровно q_c^* студентов.

Предпочтения агентов содержат индифферентности, а именно не проголосовавшие студенты считаются безразличными, все дисциплины для них одинаково приоритетны. В случае наличия нестрогих предпочтений у агентов вводится несколько определений стабильности распределения на двустороннем рынке: слабо-стабильное, сильно-стабильное и суперстабильное [5].

Пусть распределение μ называется слабо-стабильным, если не существует конфликта обоснованной зависти, то есть пары $(s, c) \notin \mu$, такой что $s \in S_c$ и $c \in C_s$, $s \in S, c \in C$, и $s \succ_c s', s' \in \mu_c$, то есть студент s строго предпочтительнее некоторого назначенного дисциплине c студента s', а так же $c \succ_s \mu_s$, то есть дисциплина c строго предпочтительнее для студента s назначенной ему дисциплине μ_s.

Пусть распределение μ называется сильно-стабильным, если не существует пары $(s, c) \notin \mu$, такой что $s \in S_c$ и $c \in C_s$, $s \in S, c \in C$, и выполняется один из следующих пунктов:

1. $s \succsim_c s', s' \in \mu_c$, то есть студент s нестрого предпочтительнее некоторого назначенного дисциплине c студента s', а также $c \succ_s \mu_s$, то есть дисциплина c строго предпочтительнее для студента s назначенной ему дисциплине μ_s, или

2. $s \succ_c s', s' \in \mu_c$, то есть студент s строго предпочтительнее некоторого назначенного дисциплине c студента s', а также $c \succsim_s \mu_s$, то есть дисциплина c нестрого предпочтительнее для студента s назначенной ему дисциплине μ_s.

Пусть распределение μ называется супер-стабильным, если не существует пары $(s, c) \notin \mu$, такой что $s \in S_c$ и $c \in C_s$, $s \in S, c \in C$, и $s \succsim_c s', s' \in \mu_c$, то есть студент s нестрого предпочтительнее некоторого назначенного дисциплине c студента s', а так же $c \succsim_s \mu_s$, то есть дисциплина c нестрого предпочтительнее для студента s назначенной ему дисциплине μ_s.

Однако сильно-стабильного и суперстабильного распределения может не существовать, например, в случае, когда лучший по успеваемости студент не проголосовал. Тогда, куда бы ни был назначен этот студент, он нестрого предпочитает другую дисциплину, в то же время дисциплине этот студент более предпочтителен, чем любой другой, так как у него лучшая успеваемость. Поэтому в данной работе рассматривается алгоритм поиска слабо-стабильного распределения, которое существует всегда и может быть найдено с помощью алгоритма – модификации алгоритма отложенного принятия [5].

На данном этапе осуществляется только распределение студентов по дисциплинам, которые проголосовали, так как именно они не безразличны к тому, какую дисциплину изучать, остальные же не проголосовавшие студенты будут одинаково удовлетворены независимо от того, куда они будут распределены. Поэтому они могут быть распределены таким образом, чтобы достичь оптимального значения других критериев оптимизации задачи.

Пусть R_n – множество студентов из группы $n \in N$, которые упорядочили дисциплины в порядке предпочтительности. Тогда все множество проголосовавших студентов определяется как $\bigcup_{n \in N} R_n$. Особенностью данной задачи является то, что предпочтения всех дисциплин по студентам одинаковы (за исключением недопустимых студентов для дисциплин) и определяются на основе успеваемости студентов, поэтому все проголосовавшие студенты могут быть упорядочены по успеваемости, пусть R – упорядоченное множество проголосовавших студентов. Для каждой дисциплины $c \in C$ определим остаточные квоты q_c^{res} – количество свободных мест на дисциплине, с учетом уже назначенных студентов. На начальном этапе $q_c^{res} = q_c^*$, затем по мере распределения студентов остаточная квота будет уменьшаться.

Алгоритм распределения проголосовавших студентов определяется следующим образом. Для каждого студента $s \in R$:

1. $\mu_s = \max_{\succ_s}\{c \in C_s: q_c^{res} > 0\}$ – студент распределяется на наиболее предпочтительную дисциплину, остаточная квота которой позволяет его принять;

2. $q_{\mu_s}^{res} = q_{\mu_s}^{res} - 1$ – остаточная квота дисциплины уменьшается на единицу.

Поскольку на предыдущем этапе осуществлялся поиск квот, то гарантированно каждый студент сможет быть распределен на дисциплину. При этом полученное «частичное» распределение не будет содержать конфликты зависти, так как студенты распределяются на наиболее предпочтительную дисциплину, где есть свободное место, однако если на наиболее предпочтительной дисциплине нет свободного места, значит все места заняты более предпочтительными студентами с точки зрения дисциплины.

Распределение студентов по «подгруппам»

После того, как проголосовавшие студенты распределены по дисциплинам, необходимо определить размеры формируемых «подгрупп», а также определить количество проголосовавших и не проголосовавших студентов из каждой группы, которые будут входить в рассматриваемые «подгруппы».

Введем необходимые обозначения: $C^{(0)}$ – множество дисциплин, для которых не исчерпана остаточная квота, то есть $q_c^{res} > 0$, $\forall\, c \in C^{(0)}$. Пусть $C^{(1)}$ – множество дисциплин, для которых необходимо сформировать несколько «подгрупп», т.е. $g_c > 1$, $\forall\, c \in C^{(1)}$. Для $C^{(0)}$ и $C^{(1)}$ верно, что $C^{(0)} \cup C^{(1)} \subseteq C$.

Обозначим $N^=$ множество академических групп, таких, что в каждой группе $n \in N^=$ выполняется условие $size_n - |R_n| > 0$, т.е. хотя бы один студент из группы не расставил предпочтения по дисциплинам. Введем обозначение $size_n^= = size_n - |R_n|$.

Введем обозначение $\mu_c^n = \{s \in S: s \in \mu_c \wedge n_s = n\}$ – множество студентов из группы n назначенных на дисциплину c по результатам работы алгоритма распределения для проголосовавших студентов.

Рассмотрим следующую задачу смешанного программирования.

Неизвестные задачи имеют вид:

- $x_{c,i}^n$, $\forall\, c \in C^{(1)}: g_c > 1$, $n \in N_c$, $i \in \{1, \dots, g_c\}$ – количество проголосовавших студентов из академической группы n, которые назначены в «подгруппу» i дисциплины c.
- b_c^n, $\forall\, c \in C^{(0)}$, $n \in N_c \cap N^= \backslash \{n \in N^=: |\mu_c^n| > 0\}$ – индикаторная пе-

ременная, принимающая значение 1, если в единственную «подгруппу» дисциплины c назначен хотя бы один студент из группы n.

- $y_{c,i}^{n}$, $\forall\, c \in C^{(0)}$, $n \in N_c \cap N^{=}$, $i \in \{1, \dots, g_c\}$ – количество не проголосовавших студентов из учебной группы n, которые назначены в «подгруппу» i дисциплины c.
- $\ddot{b}_{c,i}^{n}$, $\forall\, c \in C^{(0)} \cup C^{(1)}: g_c > 1$, $n \in N_c \cap (N^{=} \cup N_c: |\mu_c^n| > 0\})$, $i \in \{1, \dots, g_c\}$ – индикаторная переменная, принимающая значение 1, если в «подгруппу» i дисциплины c назначен хотя бы один студент из группы n.
- $num_{c,i} \geq 0$ $\forall\, c \in C^{(0)} \cup C^{(1)}: g_c > 1$, $i \in \{1, \dots, g_c\}$ – количество студентов, назначенных в «подгруппу» i дисциплины C.

Критерии оптимизации:

$$\sum_{c \in C^{(0)}} \sum_{n \in N_c \cap N^{=}} \dot{b}_c^n + \sum_{c \in C^{(1)}} \sum_{n \in N_c} \sum_{i \in \{1, \dots, g_c\}} \ddot{b}_{c,i}^n \to \min \tag{11}$$

$$\sum_{c \in C^{(1)}: g_c = 2} |num_{c,1} - num_{c,2}| \tag{12}$$

$$\sum_{c \in C^{(1)}: g_c > 2} |\overline{num}_c - \underline{num}_c| \tag{13}$$

При ограничениях:

$$\sum_{n \in N_c \cap N^{=}} \sum_{i \in \{1, \dots, g_c\}} y_{c,i}^n = q_c^{res} \qquad \forall\, c \in C^{(0)} \tag{14}$$

$$\sum_{c \in C^{(0)}} \sum_{i \in \{1, \dots, g_c\}} y_{c,i}^n = size_n^{=} \qquad \forall\, n \in N_c \cap N^{=} \tag{15}$$

$$\sum_{i \in \{1, \dots, g_c\}} x_{c,i}^n = |\mu_c^n| \qquad \begin{array}{l} \forall\, c \in C^{(1)}: g_c > 1 \\ \forall\, n \in N_c: |\mu_c^n| > 0 \end{array} \tag{16}$$

$$num_{c,i} = \sum_{n \in N_c: |\mu_c^n| > 0} x_{c,i}^n \qquad \begin{array}{l} \forall\, c \in C^{(1)} \setminus C^{(0)}: g_c > 1 \\ \forall\, i \in \{1, \dots, g_c\} \end{array} \tag{17}$$

$$num_{c,i} = \sum_{n \in} x_{c,i}^n + y_{c,i}^n \qquad \begin{array}{l} \forall\, c \in C^{(0)} \cap C^{(1)}: g_c > 1 \\ \forall\, i \in \{1, \dots, g_c\} \end{array} \tag{18}$$

$$\dot{b}_c^n \geq y_{c,1}^n / M \qquad \begin{array}{l} \forall\, c \in C^{(0)}: g_c = 1 \\ \forall\, n \in N_c \cap N^{=} \setminus \{n \in N^{=}: |\mu_c^n| > 0\} \end{array} \tag{19}$$

$$\ddot{b}_{c,i}^n \geq x_{c,i}^n / M \qquad \begin{array}{l} \forall\, c \in C^{(1)} \backslash C^{(0)} \colon g_c > 1 \\ \forall\, n \in N_c \backslash N^{=} \colon |\mu_c^n| > 0 \\ \forall\, i \in \{1, \dots, g_c\} \end{array} \tag{20}$$

$$\ddot{b}_{c,i}^n \geq (x_{c,i}^n + y_{c,i}^n) / M \qquad \begin{array}{l} \forall\, c \in C^{(0)} \cap C^{(1)} \colon g_c > 1 \\ \forall\, n \in N_c \colon |\mu_c^n| > 0 \lor n \in N^{=} \\ \forall\, i \in \{1, \dots, g_c\} \end{array} \tag{21}$$

$$\begin{array}{l} \overline{num}_c \geq num_{c,i} \\ \underline{num}_c \leq num_{c,i} \end{array} \qquad \begin{array}{l} \forall\, c \in C^{(0)} \cup C^{(1)} \colon g_c > 1 \\ \forall\, i \in \{1, \dots, g_c\} \end{array} \tag{22}$$

$$x_{c,i}^n \in \mathbb{Z}_{\geq 0} \qquad \begin{array}{l} \forall c \in C^{(1)} \colon g_c > 1 \\ \forall\, n \in N_c \\ \forall\, i \in \{1, \dots, g_c\} \end{array} \tag{23}$$

$$\dot{b}_c^n \in \{0,1\} \qquad \begin{array}{l} \forall\, c \in C^{(0)} \\ \forall\, n \in N_c \cap N^{=} \backslash \{n \in N^{=} \colon |\mu_c^n| > 0\} \end{array} \tag{24}$$

$$y_{c,i}^n \in \mathbb{Z}_{\geq 0} \qquad \begin{array}{l} \forall c \in C^{(0)} \\ \forall\, n \in N_c \cap N^{=} \\ \forall\, i \in \{1, \dots, g_c\} \end{array} \tag{25}$$

$$\ddot{b}_{c,i}^n \in \{0,1\} \qquad \begin{array}{l} \forall\, c \in C^{(0)} \cup C^{(1)} \colon g_c > 1 \\ \forall\, n \in N_c \cap (N^{=} \cup N_c \colon |\mu_c^n| > 0\}) \\ \forall\, i \in \{1, \dots, g_c\} \end{array} \tag{26}$$

$$num_{c,i} \geq 0 \qquad \begin{array}{l} \forall\, c \in C^{(0)} \cup C^{(1)} \colon g_c > 1 \\ \forall\, i \in \{1, \dots, g_c\} \end{array} \tag{27}$$

Критерий оптимизации (11) минимизирует разнообразие групп, из которых студенты формируют «подгруппы». Критерий оптимизации (12) позволяет минимизировать разницу в размерах «подгрупп», если для дисциплины должно быть открыто всего две «подгруппы», критерий оптимизации (13) минимизирует разницу между максимальным и минимальным размерами «подгрупп», если их должно быть создано более двух. Формально ограничения являются нелинейными, однако могут быть линеаризованы стандартными подходами.

Ограничения (14) гарантируют, что на каждую дисциплину назначено столько непроголосовавших студентов, чему равна остаточная квота. Ограничения (15) позволяют гарантировать, что все непроголосовавшие студенты из каждой группы распределяются на дисциплины. Ограничения (16) требуют, чтобы все проголосовавшие студенты были распределены в «подгруппы» назначенной им дисциплины на предыдущем шаге. Ограничения (17), (18) определяют размер «подгрупп», в которые входят только проголосовавшие студенты и в которые входят как проголосовавшие, так и не проголосовавшие студенты соответственно. Ограничения (19)-(21) определяют, были ли назначены в «подгруппу» студенты из соответствующих «подгрупп». Ограничения (22) определяют минимальный и максимальный размеры «подгрупп» для дисциплин, для которых должно быть создано более двух «подгрупп». Ограничения (23)-(27) определяют области допустимых значений переменных.

Далее необходимо найти непосредственно распределение студентов по «подгруппам» с учетом найденных квот $x_{i,c}^{n}, \forall c \in C^{(1)}: g_c > 1,\ n \in N_c$, $i \in \{1, \dots, g_c\}$ на количество проголосовавших студентов в «подгруппе» i и $y_{c,i}^{n}, \forall c \in C^{(0)}, n \in N_c \cap N^{-}, i \in \{1, \dots, g_c\}$ на количество не проголосовавших студентов в «подгруппе» i. Распределять студентов по «подгруппам» необходимо, минимизируя разницу в среднем балле в «подгруппах». Эта задача эквивалентна multiway number partitioning problem [6]. Разбиение осуществляется следующим образом. Студенты упорядочиваются в порядке убывания их успеваемости, лексикографически упорядочиваются подгруппы, затем студенты поочередно назначаются в «подгруппы», студент назначается в «подгруппу» если только в ней еще не исчерпана соответствующая квота на количество проголосовавших и непроголосовавших студентов, иначе студент назначается в следующую «подгруппу» с неисчерпанной соответствующей квотой.

Числовые результаты

Представленный алгоритм был апробирован на полномасштабных данных Санкт-Петербургского государственного экономического университета при распределении студентов по дисциплинами блоков дисциплин по выбору.

Исходные данные обладают следующими характеристиками:

- 3970 студентов;
- 7 блоков дисциплин по выбору;
- 30 дисциплин;
- все множество групп разделено на 34 пула, для которых происходил расчет независимо.

Реализация описанного алгоритма была выполнена в Wolfram Mathematica 13.2[1] с использованием коммерческого оптимизатора Cardinal Optimizer[2].

В результате за 340 секунд было получено решение, где:

- 2636 студентов распределены на дисциплины по своему первому приоритету;
- 571 студент распределены на дисциплины по своему второму приоритету;
- 113 студентов распределены на дисциплины по своему третьему приоритету;
- 12 студентов распределены на дисциплины по своему четвертому приоритету;

[1] Wolfram Mathematica Documentation. – URL:https://reference.wolfram.com/language/
[2] Cardinal Optimizer (COPT) Documentation. – URL: https://arxiv.org/pdf/2208.14314.pdf

- 624 студента не приняло участия в голосовании по выбору приоритетов для дисциплин;
- 0 конфликтов зависти.

Таким образом, алгоритм демонстрирует возможность использования на реальных полномасштабных данных.

Заключение

В результате разработан трехэтапный алгоритм для поиска распределения студентов по дисциплинам по выбору.

На первом этапе осуществляется поиск квот, обеспечивающий минимальное изменение оценок квот, данных в задаче, при гарантии, что все студенты могут быть распределены в «подгруппы» для изучения дисциплин.

На втором этапе осуществляется назначение проголосовавшим студентам дисциплин при максимизации удовлетворенности студентов, назначение производится с помощью модификации алгоритма отложенного принятия; максимизация удовлетворенности осуществляется за счет того, что сторона рынка студенты «делает предложения» в алгоритме, то есть распределение студентов по дисциплинам является оптимальным с точки зрения предпочтения студентов.

Заключительный этап алгоритма позволяет сформировать «подгруппы» студентов для изучения дисциплин. Сначала определяется количество проголосовавших и непроголосовавших студентов в каждой формируемой «подгруппе», затем происходит назначение студентов в «подгруппы» в соответствии с критериями: минимизация количества исходных групп, студенты которых участвуют в формировании «подгрупп», минимизация отклонения размеров «подгрупп» и минимизация отклонения средних баллов в «подгруппах». При этом формирование «подгрупп» осуществляется с помощью решения задачи multiway number partitioning problem. Однако разбивать множество студентов на подгруппы можно исходя из других критериев оптимизации, например, можно максимизировать различие в среднем балле в «подгруппах», чтобы формировать «сильные» и «слабые подгруппы».

Предложенный алгоритм отличается гибкостью, поскольку на двух этапах решения задачи используется подход на основе смешанного программирования, что позволяет добавлять пользовательские ограничения. Кроме того, эксперт может варьировать иерархию важности критериев оптимизации, а также выбирать метод решения многокритериальной задачи, что также обеспечивает гибкость предлагаемого алгоритма.

Применение представленного алгоритма на полномасштабных исходных данных показывает корректность результата, а также высокую удовлетворенность студентов от распределения – большинство студентов распределены на дисциплины по своему первому приоритету. Модель позволяет получить результат за адекватное время (около пяти минут), что демонстрирует ее применимость при решении полномасштабных задач.

В дальнейшем планируется разработка алгоритма, в котором бы единовременно формировались пулы студентов, а также блоки дисциплин. Для формируемых пулов студентов необходимо определить количество «подгрупп», а также их размеры. Кроме того, необходимо назначить преподавателей для проведения занятий по дисциплинам у «подпотоков» и «подгрупп». Проведение таких занятий означает дополнительные требования к расписанию, которые также необходимо учесть.

Список источников

1. Мину М. *Математическое программирование. Теория и алгоритмы*. Москва, Наука, 1990.
2. Benson Harold. Nonlinear Multiobjective Optimization by Kaisa M. Miettinen // *SIAM Review*, 2000, no. 42, pp. 339-341.
3. Gale D., Shapley L.S. College Admissions and the Stability of Marriage // *The American Mathematical Monthly*, 1962, no. 69, pp. 9-15.
4. Heineke M.H., Ballering A.V., Jamin A., Ben Mkaddem S., Monteiro R.C., Van Egmond M. New insights in the pathogenesis of immunoglobulin A vasculitis (Henoch-Schönlein purpura) // *Autoimmun Rev*, 2017, no. 16(12), pp. 156-183.
5. Irving R.W. Stable Marriage and Indifference // *Discret. Appl. Math.*, 1994, no. 48, pp. 261-272.
6. Korf R. Multi-Way Number Partitioning // *IJCAI International Joint Conference on Artificial Intelligence*, 2009, pp. 538-543.
7. Manlove D. Algorithmics of Matching Under Preferences // *Bull. EATCS*, 2013.
8. Manlove D. Two Algorithms for the Student-Project Allocation Problem // *Journal of Discrete Algorithms*, 2007.
9. Roth A.E. Deferred Acceptance Algorithms: History, Theory, Practice, and Open Questions // *International Journal of Game Theory*, 2008, no. 36, pp. 537-569.
10. Shimada N., Yamazaki N., Takano Y. Multi-objective Optimization Models for Many-to-one Matching Problems // J*ournal of Information Processing*, 2020, no. 28. pp. 406-412.
11. Tadenuma K. Partnership, Solidarity, and Minimal Envy in Matching Problems // *Social Ethics and Normative Economics. Studies in Choice and Welfare*. Springer, Berlin, Heidelberg, 2011.
12. Yokoi Y. Envy-Free Matchings with Lower Quotas // *Algorithmica*, 2020, no. 82, pp. 188-211.

ASSIGNING STUDENTS TO BLOCKS OF ELECTIVE COURSES

Glazunova Ekaterina Valerevna, graduate student

Saint Petersburg State University of Economics, Sadovaya str., 21, St. Petersburg, Russian Federation, 191023; e-mail: katarina.glazunova97@inbox.ru

Importance: the process of assigning university students to blocks of elective courses. *Purpose*: the development of an algorithm to find a fair matching of students by elective courses and dividing a set of students into «subgroups» for joint study of courses. *Research design*: the problem is considered as a problem of finding matching in a two-sided market, where the sides to the market are students and elective courses. Assuming that the preferences of agents from the «elective courses» market side are the same for everyone, a modification of the deferred acceptance algorithm is proposed to find a matching. The division of students into «subgroups» is based on a mixed integer programming model. *Results*: a three-step algorithm is developed in the paper to find an optimal fair matching for assigning students to blocks of elective courses based on their preferences and ratings (academic performances). The numerical results are obtained for full-scale data. They demonstrate the correctness and effectiveness of the proposed approach compared to the procedure that currently used.

Keywords: fair matching, deferred acceptance algorithm, mixed integer programming.

References

1. Minu M. *Matematicheskoe programmirovanie. Teorija i algoritmy*. Moskva, Nauka, 1990. (In Russ.)

2. Benson Harold. Nonlinear Multiobjective Optimization by Kaisa M. Miettinen. *SIAM Review*, 2000, no. 42, pp. 339-341. (In Eng.)

3. Gale D., Shapley L.S. College Admissions and the Stability of Marriage. *The American Mathematical Monthly*, 1962, no. 69, pp. 9-15. (In Eng.)

4. Heineke M.H., Ballering A.V., Jamin A., Ben Mkaddem S., Monteiro R.C., Van Egmond M. New insights in the pathogenesis of immunoglobulin A vasculitis (Henoch-Schönlein purpura). *Autoimmun Rev*, 2017, no. 16(12), pp. 156-183. (In Eng.)

5. Irving R.W. Stable Marriage and Indifference. *Discret. Appl. Math.*, 1994, no. 48, pp. 261-272. (In Eng.)

6. Korf R. Multi-Way Number Partitioning. *IJCAI International Joint Conference on Artificial Intelligence*, 2009, pp. 538-543. (In Eng.)

7. Manlove D. Algorithmics of Matching Under Preferences. *Bull. EATCS,* 2013. (In Eng.)

8. Manlove D. Two Algorithms for the Student-Project Allocation Problem. *Journal of Discrete Algorithms*, 2007. (In Eng.)

9. Roth A.E. Deferred Acceptance Algorithms: History, Theory, Practice, and Open Questions. *International Journal of Game Theory*, 2008, no. 36, pp. 537-569. (In Eng.)

10. Shimada N., Yamazaki N., Takano Y. Multi-objective Optimization Models for Many-to-one Matching Problems. *Journal of Information Processing*, 2020, no. 28, pp. 406-412. (In Eng.)

11. Tadenuma K. Partnership, Solidarity, and Minimal Envy in Matching Problems. *(In Eng.) Social Ethics and Normative Economics*. Studies in Choice and Welfare. Springer, Berlin, Heidelberg, 2011.

12. Yokoi Y. Envy-Free Matchings with Lower Quotas. *Algorithmica*, 2020, no. 82, pp. 188-211. (In Eng.)