



# **CUỘC THI TOÁN MÔ HÌNH 2024 VÒNG 1**

**02 - 28 tháng 7, 2024**

Tên đội thi: CTO  
Ngày thực hiện: (21/07/2024)

# Mục lục

<b>Giới thiệu</b>	<b>1</b>
<b>Nội dung chính</b>	<b>1</b>
<b>1 Bài toán xếp lịch</b>	<b>1</b>
1.1 Phát biểu bài toán . . . . .	1
1.2 Mô hình bài toán . . . . .	1
1.2.1 Thiết lập các kí hiệu . . . . .	2
1.2.2 Mô hình tối ưu . . . . .	3
1.2.3 Mô hình điều hoà . . . . .	3
<b>2 Thí nghiệm</b>	<b>4</b>
2.1 Công cụ . . . . .	4
2.2 Đánh giá kết quả . . . . .	4
2.3 Bảng phân công lịch trực . . . . .	5
<b>3 Nhận xét mô hình</b>	<b>5</b>
3.1 Kết quả từ mô hình . . . . .	6
3.2 Mô hình có thể tổng quát . . . . .	6
<b>4 Hướng phát triển</b>	<b>7</b>
4.1 Trong phạm vi bài toán . . . . .	7
4.2 Ứng dụng thực tiễn . . . . .	7
<b>Tài liệu tham khảo</b>	<b>9</b>
<b>Phụ lục</b>	<b>10</b>
<b>A Bảng phân công lịch trực</b>	<b>10</b>
<b>B Động lực tìm tham số</b>	<b>10</b>
<b>C Tính tổng quát của mô hình</b>	<b>13</b>
<b>D Các bảng kết quả</b>	<b>14</b>

# Giới thiệu

Bài toán Nurse Scheduling Problem (NSP) là một vấn đề phổ biến trong việc xếp lịch trực cho các bác sĩ. Một bài toán điển hình yêu cầu phải đáp ứng được các ràng buộc chính (**hard constraint**) và một số ràng buộc phụ khác bám sát tối đa nhất (**soft constraint**). Việc thiết lập được lịch trực tốt phải thoả mãn tối đa các ràng buộc trên, đồng thời đưa ra số liệu định lượng nhằm đánh giá và phát triển mô hình.

## Nội dung chính

### 1. Bài toán xếp lịch

#### 1.1. Phát biểu bài toán

Tại bệnh viện Môn Hoà, mỗi bác sĩ sẽ thuộc một trong ba loại chính (bác sĩ cộc I, bác sĩ cộc II, điều dưỡng), có điểm thâm niên và trực thuộc một khoa (cấp cứu, nội, ngoại, lâm sàng).

Bệnh viện mong muốn sắp xếp lịch trực cho các nhân viên của mình, với điều kiện phải đáp ứng được các ràng buộc sau:

- **Hard constraints.** Mỗi kíp trực phải có 1 bác sĩ cộc I, 1 bác sĩ cộc II và 2 điều dưỡng.
- **Soft constraints.** Bao gồm hai ràng buộc sau
  - Mỗi kíp trực nên có đa dạng thành viên từ 4 khoa, và nên có thành viên nhiều thâm niên.
  - Cần có tính hợp lý trong việc xoay vòng nhân sự, đảm bảo các nhân viên làm việc đều nhau.

Để thuận tiện cho việc nhắc lại, ta gọi các **soft constraints** trên với tên gọi lần lượt là **ràng buộc khoa**, **ràng buộc thâm niên** và **ràng buộc ngày**.

#### 1.2. Mô hình bài toán

Sau khi xác định yêu cầu, nhóm tác giả quyết định sử dụng mô hình Mixed Integer Linear Programming (MILP), nhằm xây dựng lịch trực cho nhân viên

ở bệnh viện Môn Hoà.

### 1.2.1. Thiết lập các kí hiệu

Trước hết ta định nghĩa các tập hợp được miêu tả ở trên dưới dạng toán học.

Kí hiệu	Định nghĩa
$H$	tập hợp tất cả nhân viên bệnh viện Môn Hoà
$T$	tập hợp các loại nhân viên
$D$	tập hợp các khoa
$T_1, T_2, T_3$	tập hợp nhân viên cộc I, cộc II và điều dưỡng
$D_1, D_2, D_3, D_4$	tập hợp nhân viên khoa nội, ngoại, cấp cứu và lâm sàng
$SE$	tập hợp các nhân viên có điểm thâm niên cao
$S$	tập hợp các ca trực

Bảng 1: Bảng kí hiệu các tập hợp

Lưu ý rằng các tập hợp  $T_1, T_2, T_3, D_1, D_2, D_3, D_4$  đều là tập hợp con của  $H$ .

Bảng dưới đây mô tả biến quyết định của bài toán.

Biến quyết định	Định nghĩa
$x_{i,j}$	1 nếu nhân viên $i \in H$ trực ca $j \in S$ , 0 nếu ngược lại

Bảng 2: Bảng kí hiệu biến quyết định

Bảng bên dưới thiết lập các tham số còn lại của bài toán.

Tham số	Định nghĩa
$sen_j$	số nhân viên có điểm thâm niên cao yêu cầu ở mỗi ca trực $j \in S$
$type_{id}$	số nhân viên loại $id \in T$ yêu cầu ở mỗi ca trực
$dept_{id}$	số nhân viên khoa $id \in D$ yêu cầu ở mỗi ca trực
$lower_{id}, upper_{id}$	chặn dưới và chặn trên cho số ca trực của các nhân viên khoa $id$ , với $id \in D$
$w_1, w_2, w_3, w_4$	các hệ số của hàm mất mát

Bảng 3: Bảng tham số của mô hình

### 1.2.2. Mô hình tối ưu

Sở dĩ nhóm tác giả đặt tên cho mô hình dưới đây là "mô hình tối ưu" vì mô hình này giải quyết được triệt để **hard constraint** và hai trên ba **soft constraints** một cách tối đa. Việc đánh giá **ràng buộc ngày** cũng có thể gọi là tốt vì độ chênh lệch về ca trực của các nhân viên của bệnh viện không quá nhiều.

Loại	Phương trình
Phương trình tối ưu	$\min \text{Sat} = \sum_{j \in S} \sum_{i \in D_4 \cap SE} x_{i,j} - \sum_{j \in S} \sum_{i \in D_4 \setminus SE} x_{i,j}$
Ràng buộc	$\sum_{i \in id} x_{i,j} = \text{type}_{id} \quad \forall j \in S, id \in T \quad (1)$ $\sum_{i \in id} x_{i,j} = \text{dept}_{id} \quad \forall j \in S, id \in D \quad (2)$ $\sum_{i \in SE} x_{i,j} = \text{sen}_j \quad \forall j \in S \quad (3)$ $\text{lower}_{id} \leq \sum_{j \in S} x_{i,j} \leq \text{upper}_{id} \quad \forall i \in id, id \in D \quad (4)$

Bảng 4: Ràng buộc mô hình tối ưu

Ràng buộc (1) thể hiện **hard constraint**, ràng buộc (2-4) lần lượt thể hiện **ràng buộc khoa**, **ràng buộc thâm niên** và **ràng buộc ngày**.

Mục tiêu của hàm tối ưu là giảm tối đa khoảng cách về số ca làm việc giữa nhân viên có điểm thâm niên cao và không cao của khoa  $D_4$  (khoa lâm sàng). Lí giải cho việc này đến từ việc ở khoa lâm sàng, số nhân viên có điểm thâm niên cao chiếm ưu thế hơn hẳn số nhân viên có điểm thâm niên không cao, dẫn đến độ lệch chuẩn của các nhân viên lớn.

### 1.2.3. Mô hình điều hoà

Mô hình dưới đây giảm nhẹ việc đáp ứng tối đa các **soft constraint**, thay vào đó cố gắng điều hoà các ràng buộc với nhau để thu được hàm mất mát tối ưu nhất.

Loại	Phương trình
Phương trình tối ưu	$\max \mathbf{Sat} = \sum_{x \in \{1,2,3,4\}} w_x \left( \sum_{i \in D_x} \sum_{j \in S} x_{i,j} \right)$
Ràng buộc	$\sum_{i \in id} x_{i,j} = \text{type}_{id} \quad \forall j \in S, id \in T \quad (1)$ $\sum_{i \in id} x_{i,j} = \text{dept}_{id} \quad \forall j \in S, id \in D' \quad (2)$ $\sum_{i \in SE} x_{i,j} = \text{sen}_j \quad \forall j \in S \quad (3)$ $\text{lower}_{id} \leq \sum_{j \in S} x_{i,j} \leq \text{upper}_{id} \quad \forall i \in id, id \in D \quad (4)$

Bảng 5: Ràng buộc mô hình điều hoà

Trong đó  $D' \subset D$  nghĩa là ta sẽ chọn  $|D'|$  khoa sao cho trong mỗi kíp thì số lượng nhân viên thuộc một trong các khoa đó đúng bằng  $\text{dept}_{id}$ .

Ý nghĩa của phương trình tối ưu là sẽ điều hoà số tổng số khoa trong một kíp bằng cách đặt hệ số phạt hợp lý. Cụ thể, nhận thấy số nhân viên khoa nội và khoa ngoại ít hơn, ta cố gắng đặt hệ số phạt lớn nếu có nhân viên ở khoa đó bị giảm số ca làm việc.

## 2. Thí nghiệm

### 2.1. Công cụ

Qua việc sử dụng các công cụ khác nhau, nhóm tác giả quyết định sử dụng công cụ CP-SAT ([2]) được phát triển bởi nhóm Google Optimization Developers. Đây là công cụ được đánh giá là tốt nhất trong việc giải quyết các bài toán chỉ bao gồm các ràng buộc tuyến tính có tham số là biến nguyên.

### 2.2. Đánh giá kết quả

Nhóm tác giả đánh giá độ hiệu quả của mô hình dựa trên các tiêu chí sau:

1. Số **hard constraint** bị vi phạm.
2. Số khoa trung bình của mỗi kíp.

3. Tỷ lệ mỗi kíp có thành viên có thâm niên cao.
4. Độ lệch chuẩn về số ngày làm việc của các nhân viên. Bao gồm độ lệch chuẩn của các nhân viên trong toàn bệnh viện và ở từng khoa.

Các kết quả dưới đây được tính toán bằng cách lấy giá trị trung bình của 5 lần chạy ngẫu nhiên. Sở dĩ có sự thay đổi vì mặc dù hàm tối ưu có giá trị không đổi nhưng cách sắp xếp nhân sự khác nhau sẽ thu được số liệu khác nhau. Ta gọi **R(a)** và **R(b)** lần lượt là kết quả của mô hình tối ưu và mô hình điều hoà.

Tiêu chí	<b>R(a)</b>	<b>R(b)</b>
Số <b>hard constraint</b> bị vi phạm	0	0
Số khoa trung bình của mỗi kíp	4.0	3.60
Tỷ lệ mỗi kíp có thành viên có thâm niên cao	100%	100%
Độ lệch chuẩn về số ngày làm việc của các nhân viên	0.85	0.80
○ Khoa nội	0.48	0.41
○ Khoa ngoại	0.29	0.0
○ Khoa cấp cứu	0.47	0.47
○ Khoa lâm sàng	0.90	0.89

Bảng 6: Bảng kết quả của hai mô hình

Từ bảng kết quả trên, ta nhận thấy rằng mô hình tối ưu đúng như mong muốn là thoả mãn tối đa hai trên ba ràng buộc của **soft constraints**, nhưng có thể thấy độ lệch chuẩn về số ca trực của các nhân viên khá lớn. Về mô hình điều hoà, tuy đã giảm về số ca trung bình mỗi kíp, nhưng bù lại các nhân viên làm việc đều nhau hơn.

### 2.3. Bảng phân công lịch trực

Nhóm tác giả quyết định sử dụng kết quả từ mô hình tối ưu để sắp xếp lịch trực cho nhân viên. Kết quả của mô hình có thể được tìm thấy tại **phụ lục A**, cụ thể tại hình 1 và hình 2.

## 3. Nhận xét mô hình

Ở phần này nhóm tác giả đưa ra một số nhận xét thấy được từ kết quả của mô hình, và đưa ra quan điểm vì sao mô hình này tốt và có thể tổng quát lên.

### 3.1. Kết quả từ mô hình

**Nhận xét 1.** Phần lớn các nhân viên có số ca trực bằng 2 tập trung phổ biến ở nhân viên điều dưỡng.

Điều này có thể giải thích được do số nhân viên điều dưỡng của bệnh viện (44) nhiều hơn gấp 2 lần so với hai loại còn lại (lần lượt là 17 và 19). Dẫn đến việc phân bố điều dưỡng dễ được thoả mãn hơn nên không phải trực nhiều. (xem hình 4)

**Nhận xét 2.** Phần lớn nhân viên có số ca trực lớn hơn 3 tập trung ở khoa Nội.

Điều này giải thích được bằng việc nhân viên khoa Nội ít hơn hẳn so với nhân viên ở từng ba khoa còn lại. Do đó, để thoả mãn **ràng buộc khoa**, các nhân viên khoa Nội phải trực nhiều hơn. (xem hình 5)

**Nhận xét 3.** Độ lệch chuẩn về số ngày làm của nhân viên khoa Lâm Sàng lớn.

Điều này có thể đến từ việc có sự chênh lệch lớn về số nhân viên có điểm thâm niên cao và không cao trong khoa Lâm Sàng. Các nhân viên có điểm thâm niên cao phải trực nhiều hơn.

### 3.2. Mô hình có thể tổng quát

Để đánh giá đây là mô hình tổng quát, ta đánh giá các tiêu chí sau:

1. **Các suy luận được đưa ra từ tính toán số liệu.** Các tham số chặn trên và chặn dưới của bài toán xuất phát từ việc quan sát số liệu và đưa ra phép tính. Do đó có thể uyển chuyển thay đổi tham số. Động lực và ý nghĩa tìm ra các tham số có thể xem tại **phụ lục B**.
2. **Độ hiệu quả của mô hình được đánh giá độc lập.** Việc đánh giá các tiêu chí hoàn toàn được đánh giá dựa trên các số liệu trong điều kiện cụ thể. Do đó không có chuẩn chính xác cho mọi mô hình, mà có thể dễ dàng thay đổi dựa trên tùy điều kiện.
3. **Mô hình vẫn chạy tốt trên tập dữ liệu mất cân bằng.** Nhóm tác giả còn chủ động kiểm tra tính tổng quát của mô hình bằng việc thí nghiệm trên bộ dữ liệu riêng, và thu được kết quả rất khả quan có thể xem tại **phụ lục C**.



## 4. Hướng phát triển

### 4.1. Trong phạm vi bài toán

Mô hình này vẫn còn nhiều hướng có thể mở rộng thêm, cũng như tạo được lịch tốt trong điều kiện thực tế. Cụ thể

1. Mỗi nhân viên phải có thời gian nghỉ sau khi làm việc ([1], [5]). Để đáp ứng điều này, sau khi khảo sát số liệu, nhóm tác giả cho rằng mỗi nhân viên có thể trực chỉ 1 lần trong một tuần, từ đó dẫn đến ràng buộc sau.

$$\sum_{k=j}^{j+7} x_{i,k} \leq 1, j \in [0, |S| - 7], i \in H$$

2. Để đảm bảo khối lượng kinh nghiệm mỗi kíp trực được rải đều, dẫn đến nếu có trường hợp xấu xảy ra, các bác sĩ có thể xử lý tốt như nhau, nhóm tác giả gợi ý việc trải đều tổng điểm mỗi ca không chênh lệch quá lớn.

$$\text{score}_l \leq \sum_{i \in H} x_{i,j} \leq \text{score}_u, \forall j \in S$$

Trong đó  $\text{score}_l, \text{score}_u$  lần lượt là chặn dưới và trên cho tổng điểm một kíp.

Áp dụng các ràng buộc trên vào mô hình, ta thu được kết quả sau, kí hiệu  $\mathbf{R}(\mathbf{c})_a$  và  $\mathbf{R}(\mathbf{c})_b$  là kết quả khi thêm các ràng buộc trên vào mô hình tối ưu và mô hình điều hoà (xem hình 9).

### 4.2. Ứng dụng thực tiễn

Trên thực tế, không phải lúc nào nhân viên được xếp lịch cũng có mặt, do vậy đề ra hai vấn đề cần được giải quyết:

1. Sắp xếp lịch sao cho khả năng nhân viên đó không đi trực nhỏ nhất.
2. Nếu có nhân viên trực bị nghỉ, cần tìm nhân viên thay thế.

Như vậy, ta bổ sung thêm

Tham số	Định nghĩa
$p_{i,j}$ $lower'_{id}, upper'_{id}$	xác suất để nhân viên $i \in H$ đi trực vào ngày $j \in S$ chặn trên và dưới cho tổng số ca trực bổ sung của một nhân viên

Bảng 7: Bảng tham số bổ sung

Ta có thể mở rộng mô hình như sau

Loại	Phương trình
Phương trình tối ưu	$\max \mathbf{Sat} = \sum_{i \in H} \sum_{j \in S} x_{i,j} \cdot p_{i,j}$
Ràng buộc	$\sum_{i \in id} x_{i,j} \geq \text{type}_{id} \quad \forall j \in S, id \in T \quad (1)$ $\sum_{i \in id} x_{i,j} = \text{dept}_{id} \quad \forall j \in S, id \in D' \quad (2)$ $\sum_{i \in SE} x_{i,j} = \text{sen}_j \quad \forall j \in S \quad (3)$ $\text{lower}_{id} \leq \sum_{j \in S} x_{i,j} \leq \text{upper}_{id} \quad \forall i \in id, id \in D \quad (4)$ $\text{lower}'_{id} \leq \sum_{j \in S'} x_{i,j} \leq \text{upper}'_{id} \quad \forall i \in id, id \in D \quad (5)$

Bảng 8: Mô hình chính trong bài toán thực tiễn

Phương trình tối ưu của bài toán này nhằm đạt được giá trị lớn nhất về tổng xác suất các nhân viên của bệnh viện đi trực.

Mô hình trên thoạt nhìn không khác mô hình tối ưu nhiều, nhưng thực chất ta bổ sung thêm ràng buộc (5) nhằm ý nghĩa ta sắp xếp các nhân viên vào lịch trực "bổ sung" nhằm thoả mãn lúc nào cũng có nhân viên thay thế.

## Tài liệu tham khảo

- [1] Emmons, Hamilton, *WorkForce Scheduling with Cyclic Requirements and Constraints on Days Off, Weekends Off, and Work Stretch*, Iie Transactions, 17. 8-16. 10.1080/07408178508975266, 1985.
- [2] Laurent Perron and Frédéric Didier, OR-Tools CP-SAT v9.10, [https://developers.google.com/optimization/cp/cp\\_solver/](https://developers.google.com/optimization/cp/cp_solver/), Google, May 7, 2024
- [3] P. De Bruecker, J. Van den Bergh, J. Belien, E. Demeulemeester, Workforce planning incorporating skills: State of the art, *European Journal of Operational Research*, 243(1), 1-16, 2015.
- [4] R. Soto, B. Crawford, E. Monfroy, W. Palma, F. Paredes, Nurse and paramedic rostering with constraint programming: A case study, *Romanian Journal of Information Science and Technology*, 16(1), 52-64, 2013.
- [5] Yamamoto, Masahito & Ohuchi, Azuma, *Evolutionary Algorithms for Nurse Scheduling Problem*, 2000.

Phụ lục

Phụ lục A. Bảng phân công lịch trực

Lịch trực tháng 8																			
Chủ Nhật				Thứ Hai				Thứ Ba				Thứ Tư				Thứ Năm			
1				2				3				4				5			
Trang				Hà				Khánh Linh				Huyền				Bằng			
Phương				Tiến				Len				Thương				Thắng			
Thu				Phi				1				2				3			
Long				Dung				Tơ				Nam				Ngân Hà			
Minh Phương				Linh				Dung				Huỳnh				Hùng			
Hỏi				Vương				Thư				Tâm				Tú			
Quýnh				1				2				3				4			
Linh				Hưng				Sang				Sơn				Gần			
Thu Hằng				Ngọc				Bắc				Hưng				Tuyền			
Cộng				11				12				13				14			
Bà Tuân				Vân Anh				Chi Đức				Dũng				Long			
Diệp				Thuận				Tiến				Long				Quân			
Hưng				Thu				Tú				Sơn				Giảng			
Chung				Thánh				Thư				Gần				Vương			
Len				18				19				20				21			
Bằng				Dũng				Hoài				Dũng				Trang			
Minh Phương				An				Lương				Xuân				Hưng			
Linh				Thống				Bà Tuân				Phượng				Tuyền			
Minh				Tuân				Linh				Thảo				Hoài			
Sơn				Vân Anh				Kim Anh				Yên				Thương			
Xuân				Dương				Dung				25				26			
Long				Thắng				Hương				Tú				Thủy			
Minh Phương				Huê				Dũng				Huỳnh				Giảng			
Dũng				Sang				Thư				Hà				Oanh			
Vân Đức				Xuân				Linh				An				Quân			
Dung				Thống				Huyền				Sơn				Xuân			
Anh				Quýnh				29				30				31			

Hình 1: Lịch tháng 8

Lịch trực tháng 9																			
Chủ Nhật				Thứ Hai				Thứ Ba				Thứ Tư				Thứ Năm			
1				2				3				4				5			
Ngân Hà				Út				Tơ				Thu Hằng				Bà Tuân			
Thắng				Khánh Linh				Cộng				Giảng				Hưng			
Huê				Thảo				Tâm				Hương				Hoài			
Dũng				Thủy				Hà				Linh				Vân Đức			
Thương				Dũng				Dương				Thánh				Thánh Hà			
Hưng				Ngân				Tuyền				8				9			
Phi				10				11				12				13			
Vân Anh				Anh				An				Linh				Út			
Vương				Linh				Thắng				Chi				Phi			
15				16				17				18				19			
Tuân				Hưng				Dương				Thảo				Tâm			
Hùng				Thu				Dũng				Thủy				Thuận			
Ngân				Huyền				Huỳnh				Linh				Thống			
Long				22				23				24				25			
Bắc				Dũng				Thánh Hà				Út				An			
Thu Hằng				Quân				Minh Phương				Huê				Tơ			
Hưng				Dũng				Huyền				Thu				Tâm			
Hưng				Kim Anh				Cộng				Diệp				Thuận			
Huê				Phi				29				30				31			
Bắc				Thắng				Chi				Vân Đức				Giảng			
Gần				Huê				Sang											

Hình 2: Lịch tháng 9

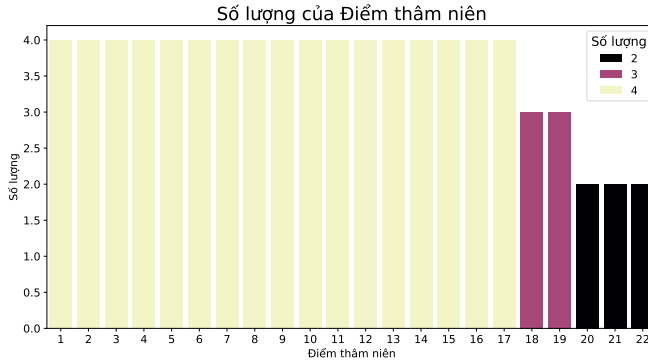
Trong hình trên, mỗi ngày trong tháng được chia làm 4 ô nhỏ, theo thứ tự tương ứng với bác sĩ cộc I, bác sĩ cộc II và hai điều dưỡng.

Phụ lục B. Động lực tìm tham số

Ở phần này, nhóm tác giả đưa ra động lực tìm ra các tham số của bài toán được liệt kê tại bảng 3 thông qua suy luận và thực hành.

Về điểm được xem là có thâm niên cao.

Xuất phát từ yêu cầu của bài toán rằng mỗi kíp trực nên có ít nhất một nhân viên có điểm thâm niên cao, mà mỗi kíp trực gồm 4 thành viên, điều này tạo tiền đề cho việc chọn mốc điểm mà tại đó có khoảng 25% nhân viên có điểm cao hơn. Dựa vào tập dữ liệu, ta có thể chọn mốc 16 điểm.



Hình 3: Biểu đồ phân bố điểm của nhân viên

### Về số nhân viên yêu cầu ở mỗi loại trong một kíp.

Vì đây là ràng buộc về **hard constraint**, các tham số phải được đưa ra theo đúng yêu cầu. Cụ thể ta có

$$\text{type}_{T_1} = 1, \text{type}_{T_2} = 1, \text{type}_{T_3} = 2$$

### Về số nhân viên có điểm thâm niên cao yêu cầu ở mỗi kíp.

Để tối ưu mô hình, nhóm tác giả chọn

$$\text{sen}_j = 1, \forall j \in S$$

Với ý nghĩa mỗi kíp chỉ cần một nhân viên có điểm thâm niên cao. Lý giải cho việc sử dụng dấu "=" thay vì " $\geq$ " vì để tập dụng vừa đủ số nhân viên có điểm thâm niên cao.

### Về số nhân viên yêu cầu ở mỗi khoa trong một kíp.

Do đây là trường hợp bài toán với số liệu lý tưởng, có nghĩa là số khoa trong bệnh viện bằng đúng số nhân viên trong một kíp, nên ta có thể đặt các tham số như sau

$$\text{dept}_{D_1} = \text{dept}_{D_2} = \text{dept}_{D_3} = \text{dept}_{D_4} = 1$$

Tham số này có thể thay đổi tùy vào mục tiêu của mô hình, và dấu "=" ở ràng buộc (2) tại bảng 4 có thể thay bằng dấu " $\leq$ " trong trường hợp số khoa nhiều hơn số nhân viên yêu cầu trong một kíp.

### Về chặn trên và chặn dưới cho số ca trực của các nhân viên.

Với mỗi tập hợp  $X$ , ta kí hiệu  $|X|$  là lực lượng của  $X$ , và cũng chính là số phần tử của  $X$ .

Rõ ràng trong trường hợp lý tưởng nhất, mỗi nhân viên thuộc khoa  $id \in D$  sẽ trực đúng  $\frac{|S|}{|id|}$  ngày. Chính vì vậy, với mỗi  $id \in D$ , ta sẽ chọn

$$(\text{lower}_{id}, \text{upper}_{id}) = \left( \left\lceil \frac{|S|}{|id|} \right\rceil - \varepsilon_l, \left\lceil \frac{|S|}{|id|} \right\rceil + \varepsilon_u \right)$$

Với kí hiệu  $[x]$  là phép làm tròn  $x$  đến số nguyên gần nhất, và  $\varepsilon_l, \varepsilon_u$  thường sẽ được điều chỉnh tùy theo số liệu để bài toán đạt điều kiện tốt nhất.

Sau khi thực hành, nhóm tác giả quyết định sử dụng các số sau làm khoảng chặn trên và chặn dưới

$$\begin{aligned} (\text{lower}_{D_1}, \text{upper}_{D_1}) &= (4, 5) \\ (\text{lower}_{D_2}, \text{upper}_{D_2}) &= (2, 3) \\ (\text{lower}_{D_3}, \text{upper}_{D_3}) &= (2, 3) \\ (\text{lower}_{D_4}, \text{upper}_{D_4}) &= (2, 5) \end{aligned}$$

### Về các hệ số của hàm mất mát.

Ta quan tâm đến hệ số trong mô hình điều hoà được đưa ra ở bảng 5. Sau khi chạy mô hình tối ưu ở bảng 4, ta thu được độ lệch chuẩn về số ca trực của mỗi nhân viên như xấp xỉ như sau.

$$0.4 \ 0.2 \ 0.4 \ 0.9$$

Vấn đề đặt ra là cố gắng điều hoà các số này lại. Để đạt được điều đó, ở bước đầu tiên, ta chọn các hệ số phạt thoả mãn

$$w_1 \cdot 0.4 = w_2 \cdot 0.2 = w_3 \cdot 0.4 = w_4 \cdot 0.9$$

Như vậy, ở khoa nào có độ lệch chuẩn thấp, ta sẽ cố gắng "tăng nhẹ" độ lệch chuẩn, đồng thời "giảm bớt" độ lệch chuẩn ở các khoa có giá trị cao. Việc làm như vậy sẽ "điều hoà" độ lệch chuẩn về số ca trực của nhân viên các khoa.

Sau khi thực thi ý tưởng và thực hành, nhóm tác giả đưa ra bộ hệ số như sau

$$(w_1, w_2, w_3, w_4) = (9, 18, 16, 4)$$

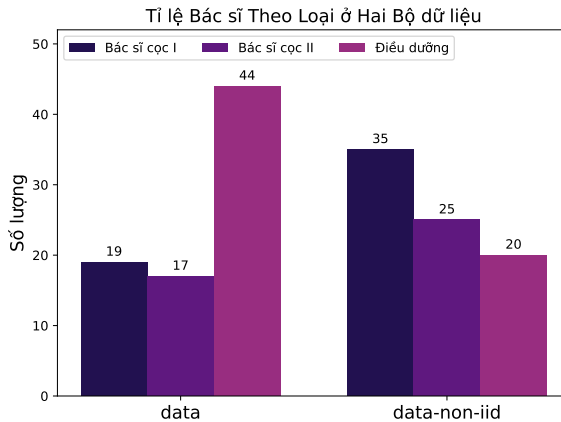
Và thu được các kết quả khá tốt được thể hiện ở bảng 6 và bảng 9.

## Phụ lục C. Tính tổng quát của mô hình

Để kiểm tra tính tổng quát của mô hình, nhóm tác giả đã chủ động tạo bộ dữ liệu mới (`data-non-iid.csv`) và thí nghiệm trên mô hình đề xuất. Gọi  $\mathbf{R}(\mathbf{n})_a$  và  $\mathbf{R}(\mathbf{n})_b$  là hai kết quả tương ứng với mô hình tối ưu và mô hình điều hoà đối với bộ dữ liệu mới, ta thu được bảng kết quả (xem tại bảng 10).

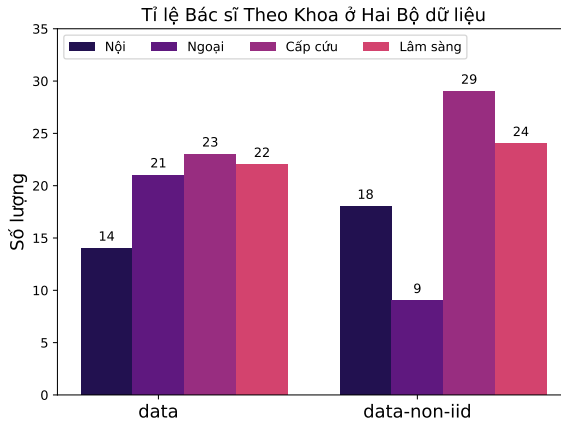
Mô hình vẫn giữ được việc có tổng số khoa trong một kíp ở mức cao, và ta đã đạt được mục tiêu mong muốn ở hàm mất mát, vì có thể thấy độ lệch chuẩn về ca trực của các khoa ở mô hình điều hoà là khá đều so với cùng tiêu chí ở mô hình tối ưu.

Để hiểu rõ hơn, ta xem xét biểu đồ phân bố bác sĩ cộc I, bác sĩ cộc II và điều dưỡng ở hai bộ dữ liệu.



Hình 4: Phân bố số lượng loại ở hai bộ dữ liệu

Từ biểu đồ trên, ta thấy bộ dữ liệu ban đầu cho sự phân bố lý tưởng là số bác sĩ cộc I, cộc II và điều dưỡng xấp xỉ tỉ số 1:1:2, tạo điều kiện thuận lợi cho việc sắp xếp thỏa mãn bài toán. Còn ở bộ dữ liệu mới, tỉ lệ này xấp xỉ 4:3:2, dẫn đến có sự khó khăn nhất định khi cố gắng sắp xếp nhân sự thỏa mãn các **soft constraints**.



Hình 5: Phân bố số lượng khoa ở hai bộ dữ liệu

Nhìn vào sự phân bố về khoa ở hai bộ dữ liệu, ta cũng có thể thấy khó khăn nhất định khi cố gắng cho các nhân viên làm việc đều nhau ở `data-non-iid`, vì rõ ràng giữa các khoa có số nhân viên chênh lệch rất lớn. Việc thu được kết quả trình bày ở bảng 10 là một kết quả tốt.

## Phụ lục D. Các bảng kết quả

Ở phần này nhóm tác giả liệt kê lại các bảng kết quả thu được từ mô hình mở rộng và mô hình chính khi áp dụng với bộ dữ liệu mới.



Tiêu chí	$R(c)_a$	$R(c)_b$
Số <b>hard constraint</b> bị vi phạm	0	0
Số khoa trung bình của mỗi kíp	4.0	3.65
Tỉ lệ mỗi kíp có thành viên có thâm niên cao	100%	100%
Độ lệch chuẩn về số ngày làm việc của các nhân viên	0.85	0.80
○ Khoa nội	0.48	0.41
○ Khoa ngoại	0.29	0.0
○ Khoa cấp cứu	0.47	0.47
○ Khoa lâm sàng	0.90	0.89

Bảng 9: Kết quả của mô hình mở rộng

Tiêu chí	$R(n)_a$	$R(n)_b$
Số <b>hard constraint</b> bị vi phạm	0	0
Số khoa trung bình của mỗi kíp	4.0	3.5
Tỉ lệ mỗi kíp có thành viên có thâm niên cao	100%	100%
Độ lệch chuẩn về số ngày làm việc của các nhân viên	2.12	2.62
○ Khoa nội	2.61	2.18
○ Khoa ngoại	1.99	2.06
○ Khoa cấp cứu	1.67	1.63
○ Khoa lâm sàng	2.35	2.36

Bảng 10: Kết quả của mô hình trên tập dữ liệu mới