Mô hình tối ưu cho việc phân tuyến học sinh vào lớp 6

Minh Hoàng, Vũ Hùng, Minh Hiếu

Tháng 8 năm 2016

Tổng quan

Giáo dục và đào tạo là một nhân tố quan trọng, đóng vai trò nồng cốt của một quốc gia, hay nói cách khác giáo dục chính là chìa khoá mở ra cho một quốc gia như Việt Nam có một cơ hội hoà nhập với toàn thế giới. Nước ta đã và đang từng ngày đẩy mạnh phát triển giáo dục và đạt được nhiều thành tựu nhất định Song, vẫn còn những vấn đề bất cập nhức nhối vẫn đang diễn ra âm thầm từng ngày mạnh mẽ đã tạo sự tiêu cực trong môi trường giáo dục nước nhà. Tiêu biểu nhất là vấn đề "chạy đua"để đưa các con vào các trường "điểm". Trong một cuộc khảo sát trực tuyến của báo Dân Trí với gần 20,000 người, hơn 62% phụ huynh thừa nhận đã dùng tiền hoặc quan hệ cá nhân để "chạy"trường tốt cho con mà trong đó có tới 31% học sinh không đạt tiêu chuẩn tuyển[1]. Vậy việc đưa ra một phương án phân tuyến phù hợp là một vấn đề cấp thiết và cũng như một bài câu hỏi khó cho Bộ giáo dục và đào tạo. Một cách phân tuyến phải đảm bảo được các yếu tố như số học sinh phân vào các trường phải phù hợp với điều kiện của các trường cũng như sự công bằng và các quyền lợi của học sinh. Vì vậy chúng tôi đã phát triển và đưa ra một số mô hình, sử dụng các thuật toán như cặp ghép hay quy hoạch tuyến tính, để có thể tự động đưa ra cách phân tuyến phù hợp với các dữ liệu về học sinh và trường học.

Từ khóa: phân tuyến; ghép cặp; quy hoach tuyến tính; ghép cặp ổn định

---

Mục lục

1 Bài toán đặt ra

Chúng ta cần tìm phương án cho phân tuyến các học sinh tốt nghiệp tiểu học vào các trường cấp 2 trong cùng khu vực quận. Ta có thể phân một số lượng học sinh nhất định từ các trường tiểu học bất kì đến trường trung học bất kì. Do trong thực tế không gian của mỗi trường và cần một kinh phí nhất định để hoạt động nên chúng ta sẽ có giới hạn trên và giới hạn dưới của của học sinh từng trường.Do sự di chuyển của các học sinh tới trường học cũng có ảnh hưởng không nhỏ tới cách chọn phương án nên chúng ta có dữ liệu về thời gian di chuyển từ các trường tiểu học đến các trường trung học. Vậy chúng ta cần chọn trọng số các cạnh một cách hợp lý đề thỏa các điều kiện của trường và phải tối ưu theo các yếu tố khác.

2 Giả thiết

1. Do trường tiểu học không phân biệt trường chuyên hoặc thường nên ta giả sử các trường đều có các quyền lợi giống nhau và không có sự phận biệt giữa các trường

2. Để việc phân bố trở nên hiệu quả ta nên coi học sinh của mỗi trường là một nhóm có vị trí địa lý giống nhau, nên ta giả sử học sinh sống gần trường tiểu học.

3. Trong trường hợp xấu nhất, các học sinh tốt nghiệp đều được phân vào một trường trung học cơ sở.

4. Số học sinh không học các trường trung học cơ sở trực thuộc sở là thấp và không đáng kể. .

5. Các học sinh được chia thành 4 nhóm theo xếp loại học tập: Giỏi, Khá, Trung Bình, Yếu.

3 Xây dựng mô hình

3.1 Mô hình phân tuyến theo thời gian di chuyển và sự phân bố của học sinh

Khi phân tuyến, đội tượng quan trọng nhất cần phải tối ưu là các học sinh. Một yếu tố quan trọng là thời gian di chuyển của các em đến trường. Vì vậy mô hình đầu tiên của chúng tôi được phát triển dựa trên vị trí của các học sinh vào các trường.

3.1.1 Các biến số

Biến số	Ý nghĩa
M	Số lượng trường tiểu học trong khu vực.
N	Số lượng trường trung học cơ sở trong khu vực.
G_i	Số lượng học sinh tốt nghiệp từ trường tiểu học i.
L_i	Biểu diễn số học sinh tối thiểu mà trường trung học cơ sở i nhận.
H_i	Biểu diễn số học sinh tối đa mà trường trung học cơ sở i nhận.
S_i	Biểu diễn số học sinh tối ưu của trường trung học cơ sở i.
$D_{i,j}$	Thời gian di chuyển từ trường tiểu học i đến trường trung học cơ sở j.
$X_{i,j}$	Số học sinh được từ trường tiểu học i được phân vào trường trung học cơ sở j.
Sum	Tổng số học sinh tốt nghiệp từ tất cả các trường tiểu học trong khu vực.

Ngoại trừ $X_{i,j}$ thì tất cả biến số ở trên đều là dữ liệu nhập vào từ thực tế.

Chúng tôi sẽ dùng tài liệu thống kê "Số liệu tiếp nhận học sinh tiểu học vào trường Trung học cơ sở năm 2016-2017"[3] được trích từ báo "Thanh Niên"để phân tuyến học sinh trong khu vực quận Tân Bình, thành phố Hồ Chí Minh. Việc áp dụng mô hình trên vào thực tế sẽ cho ta thấy khả năng ứng dụng của mô hình trong thực tế.

stt	Trường THCS	Số lớp				Số học sinh				SS bquân		Trường Tiểu học	Số học sinh				Số liệu phân bố															
		Tổng	Thường	TC Anh	TC Pháp	Tổng	Thường	TC Anh	TC Pháp	Thường	TC Anh		Tổng	Thường	TC Anh	TC Pháp	P.1	P.2	P.3	P.4	P.5	P.6	P.7	P.8	P.9	P.10	P.11	P.12	P.13	P.14	P.15	Ng.q
1	THCS Ngô Sĩ Liên	14	9	4	1	665	451	157	57	50.1	39.3	Lê Văn Sĩ	365	289	76		92	85	26	7	7	4						2	2	1	7	56
												Tân Sơn Nhất	34	34				34														
												Nguyễn Thanh Tuyền	209	128	81		1	86	6	6	3						2	3	3	1	1	16
												Trần Quốc Toản	57			57																
2	THCS Âu Lạc	10	6	4		430	267	163		44.5	40.8	Hoàng Văn Thụ	280	199	81		1	1	5	116	11		2	1	1	1		7	3	1	3	46
												Tân Sơn Nhất	150	68	82					47								4	1		1	15
3	THCS Nguyễn Gia Thiều	10	5	5		394	226	168		45.2	33.6	Chi Lăng	98	98						1	1	77	1	1	1	2		1	1			12
												Đống Đa	168		168																	
												Trần Quốc Toản	128	128																128		
4	THCS Trần Văn Đang	6	4	2		222	169	53		42.3	26.5	Bạch Đằng	87	34	53							23		1					2	1		7
												DL Thanh Bình	13	13																		13
												Bành Văn Trân	122	122				3	4	3	24	41										47
5	THCS Tân Bình	14	12	2		646	573	73		47.8	36.5	Bình Giã	127	127			2	10	104	4	1				1	1						4
												Phạm Văn Hai	119	119					1	2	91	4	1		1				2	1		16
												Bành Văn Trân	288	215	73								89	47	12	38	18	4	1	4	2	
												Thái Bình Dương	14	14				2		1												11
												QT Á Châu	50	50																		50
												Lê Anh Xuân	48	48					1			5	34	1			1		1			5
6	THCS Quang Trung	8	8			383	383			47.9		Trần Văn Ơn	227	227					1	2				150	31	16	7					20
												Lạc Long Quân	74	74						2	1			2	3	46						20
												Trần Quốc Toản	82	82						2				3		14	18	24	16		5	

stt	Trường THCS	Số lớp				Số học sinh				SS bquân		Trường Tiểu học	Số học sinh				Số liệu phân bố															
		Tổng	Thường	TC Anh	TC Pháp	Tổng	Thường	TC Anh	TC Pháp	Thường	TC Anh		Tổng	Thường	TC Anh	TC Pháp	P.1	P.2	P.3	P.4	P.5	P.6	P.7	P.8	P.9	P.10	P.11	P.12	P.13	P.14	P.15	Ng.q
7	THCS Lý Thường Kiệt	8	8			314	314			39.3		Lý Thường Kiệt	110	110					1			1		3	85	5	1	1				13
												Ngọc Hồi	65	65											56	2						7
												Trần Quốc Toản	102	102																		102
												DL Lương T. Vinh	32	32			1			1							2		1	2	2	23
												DL.Rạng Đông	5	5																1		4
8	THCS Phạm Ngọc Thạch	8	8			389	389			48.6		Phú Thọ Hòa	215	215											1	173				1		40
												Lê Thị Hồng Gấm	174	174				1							1	128	1		9	24		10
9	THCS Võ Văn Tần	11	8	3		517	366	151		45.8	50.3	Lê Thị Hồng Gấm	88		88																	
												Nguyễn Văn Trỗi	323	260	63		1						2	3	1	29	146	6	6	7	16	43
												Lạc Long Quân	106	106													101	1		2	2	
10	THCS Ngô Quyền	13	9	4		612	426	186		47.3	46.5	Cách Mạng Tháng Tám	178	178					2	2	2	1			1	3	3	67	21	7	9	60
												Nguyễn Khuyến	229	182	47											4	3	108	34	20	1	12
												Tân Trụ	205	66	139																66	
11	THCS Trường Chinh	12	10	2		576	477	99		47.7	49.5	Nguyễn Văn Kịp	252	252			1		1			1			1		1				240	7
												Cách Mạng Tháng Tám	43	43														43				
												Trần Quốc Toản	60	60																60		
												Trần Quốc Tuấn	221	122	99			1	1	2	2			2		1	3	72		13	25	
12	THCS Hoàng Hoa Thám	17	12	5		783	555	228		46.3	45.6	Trần Quốc Tuấn	125	125															38			87
												Yên Thế	279	188	91							1		1		1		2	145	6	10	22
												Sơn Cang	131	44	87												1	1	30		4	8
												Thân Nhân Trung	72	72			1		1									3	52	1	3	11
												Tân Trụ	126	126																	100	26
												Nguyễn Văn Kịp	50		50																	
		131	99	31	1	5931	4596	1278	57				5931	4596	1278	57	100	223	154	198	143	158	129	215	196	464	308	349	368	281	497	813

Từ bảng số liệu trên, chúng tôi chỉ xét 6 trường THCS theo thứ tự từ 1 đến 6 đơn giản hóa việc tính toán cũng như giúp bài toán được gọn và dễ hiểu hơn. từ đó, lập được ma trận dưới đây và dùng Matlab để giải:

$$X_{i,j} = \begin{pmatrix} 365 & 0 & 0 & 0 & 0 & 0 \\ 34 & 150 & 0 & 0 & 0 & 0 \\ 209 & 0 & 0 & 0 & 0 & 0 \\ 57 & 0 & 128 & 0 & 0 & 0 \\ 0 & 208 & 0 & 0 & 0 & 0 \\ 0 & 0 & 98 & 0 & 0 & 0 \\ 0 & 0 & 168 & 0 & 0 & 0 \\ 0 & 0 & 0 & 87 & 0 & 0 \\ 0 & 0 & 0 & 13 & 0 & 0 \\ 0 & 0 & 0 & 122 & 288 & 0 \\ 0 & 0 & 0 & 0 & 127 & 0 \\ 0 & 0 & 0 & 0 & 119 & 0 \\ 0 & 0 & 0 & 0 & 14 & 0 \\ 0 & 0 & 0 & 0 & 50 & 0 \\ 0 & 0 & 0 & 0 & 48 & 0 \\ 0 & 0 & 0 & 0 & 0 & 227 \\ 0 & 0 & 0 & 0 & 0 & 74 \end{pmatrix}$$

Chúng tôi sẽ chỉ xét 3 trường Tiểu học là Ngô Sĩ Liên, Âu Lạc và Nguyễn Gia Thiều để đơn giản hóa việc tính toán cũng như giúp các bạn hiểu rõ nội dung mô hình. Dưới đây là ma trận thời gian di chuyển giữa 6 trường Tiểu học và 3 trường THCS (phút)

$$D_{i,j} = \begin{array}{l} \\ \text{Lê Văn Sĩ} \\ \text{Tân Sơn Nhất} \\ \text{Nguyễn Thanh Tuyền} \\ \text{Hoàng Văn Thụ} \\ \text{Chi Lăng} \\ \text{Đống Đa} \end{array} \begin{array}{c} \begin{array}{ccc} \text{Ngô Sĩ Liên} & \text{Âu Lạc} & \text{Nguyễn Gia Thiều} \end{array} \\ \begin{pmatrix} 3 & 6 & 17 \\ 6 & 7 & 15 \\ 10 & 14 & 12 \\ 5 & 1 & 9 \\ 11 & 7 & 5 \\ 13 & 10 & 3 \end{pmatrix} \end{array}$$

$$G_i = \begin{array}{l} \text{Lê Văn Sĩ} \\ \text{Tân Sơn Nhất} \\ \text{Nguyễn Thanh Tuyền} \\ \text{Hoàng Văn Thụ} \\ \text{Chi Lăng} \\ \text{Đống Đa} \end{array} \begin{pmatrix} 365 \\ 34 \\ 209 \\ 280 \\ 98 \\ 168 \end{pmatrix}$$

$$L_i = \begin{array}{l} \text{Ngô Sĩ Liên} \\ \text{Âu Lạc} \\ \text{Nguyễn Gia Thiều} \end{array} \begin{pmatrix} 558 \\ 380 \\ 266 \end{pmatrix} \quad R_i = \begin{array}{l} \text{Ngô Sĩ Liên} \\ \text{Âu Lạc} \\ \text{Nguyễn Gia Thiều} \end{array} \begin{pmatrix} 658 \\ 480 \\ 366 \end{pmatrix}$$

3.1.2 Giải quyết mô hình

Đầu tiên hãy tiếp cận bài toán theo một cách trực tiếp nhất. Cho các danh sách các trường tiểu học và số lượng học sinh tốt nghiệp từ mỗi trường, danh sách các trường trung học cơ sở và các giới hạn trên và dưới của chỉ tiêu các trường đó., lập ra một kế hoạch phân tuyến sao cho các học sinh đều được phân vào một trường, và tổng thời gian để đi tới trường của các em là nhỏ nhất. Do chúng ta đã giả sử học sinh sống ở gần trường tiểu học, khoảng cách từ nhà một học sinh tiểu học trường i bất kì đến một trường trung học cơ sở j là $D_{i,j}$.

Từ bài toán trên ta có thể lập một phương pháp giải theo quy hoạch tuyến tính qua những điều kiện sau:

1. $\sum_{j=1}^{M} X_{i,j} = G_i, \forall\ i \leq N$
2. $\sum_{i=1}^{N} X_{i,j} \geq L_j, \forall\ j \leq M$
3. $\sum_{i=1}^{N} X_{i,j} \leq H_j, \forall\ j \leq M$
4. *minimize* $\sum_{i=1}^{N} \sum_{j=1}^{M} X_{i,j} \times D_{i,j}$.

Dưới đây chúng tôi xin đưa ra một hướng tiếp cận khác với bài toán. Bài toán không còn đưa về việc tối ưu thời gian nữa mà là một yếu tố quan trọng không kém trong công việc phân tuyến.

Sau khi chúng tôi đã thử nghiệm mô hình trên với những dữ liêu từ quận Tân Bình, cụ thể hơn là trong 3 trường trung học cơ sở và 6 trường tiểu học bằng MATLAB. Chương trình trả ra kết quả.

$$X_{i,j} = \begin{pmatrix} 365 & 0 & 0 \\ 184 & 0 & 0 \\ 125 & 0 & 84 \\ 0 & 185 & 0 \\ 0 & 0 & 98 \\ 0 & 0 & 168 \end{pmatrix}$$

Có thể thấy kết quả trả ra cũng khá tương đồng với số lượng học snh được phân ra bởi sở. Với số lượng học sinh từ các trương Lê Văn Sĩ, Chi Lăng và Đống Đa là bằng nhau. Vậy ta có thể thấy do các yếu tố là các lớp tăng cường ngoại ngữ tạo thêm một số ràng buộc cho mô hình của chúng ta.

Mở rộng 1

Với mô hình trên, chúng ta có thể phân tuyến học sinh phù hợp cho việc di chuyển đến trường. Song, không thể tối đa được chỉ tiêu thích hợp cho từng trường và học lực học sinh sẽ không đồng đều dẫn đến chênh lệch về chất lượng học sinh giữa các trường cũng như không đảm bảo được chất lượng giảng dạy. Nhận ra được vấn đề đó, chúng tôi đã phát triển thêm các mô hình mở rộng nhằm khắc phục những điểm chưa phù hợp của mô hình chính.Trong trường hợp tất cả các học sinh học sinh tiểu học đã tốt nghiệp không vượt quá tất cả số học sinh cần tuyển vào trường trung học cơ sở:

$\sum_{j=1}^{N} G_j \leq \sum_{i=1}^{M} S_i$

Ta đặt $\omega = \sum_{i=1}^{M} S_i - \sum_{j=1}^{N} G_j$. Dễ thấy ω là số lượng học sinh thừa thiếu trong năm đó, vậy để công bằng, ta chia đều số lượng học sinh đó cho mỗi trường.

Đặt P_i là số nguyên gần nhất với $S_i - \frac{\omega}{M}$ vậy theo điều kiện đề bài, bài toán đưa về:

1. $\sum_{i=1}^{N} X_{ij} = P_j$

2. $\sum_{j=1}^{M} X_{ij} = G_i$

3. minimize $\sum_{i=1}^{N} \sum_{j=1}^{M} X_{ij} D_{ij}$

Sau khi thử nghiệm với mô hình

Mở rộng 2

Một tình huống khác sẽ gặp trong thực tế chính là số học sinh tốt nghiệp không đủ dẫn đến thiếu chỉ tiêu các trường hoặc vượt quá chỉ tiêu nhưng vẫn theo một tỉ lệ nhất định . Chúng tôi xây dựng mô hình này để bố trí số lượng học sinh hợp lí cho từng trưởng trong trường hợp này. Đặt k_i là số học sinh trường trung học cơ ở i thiếu, ta có:

- $\sum_{i=1}^{M} k_i = \omega$

- $\frac{k_1}{S_1} = \frac{k_2}{S_2}=...=\frac{k_M}{S_M}$

Vậy $\frac{k_1}{S_1} = \frac{k_2}{S_2}=...=\frac{k_M}{S_M}=\frac{\sum_{i=1}^{M} k_i}{\sum_{i=1}^{M} S_i} = \frac{\omega}{Sum}$
$\rightarrow k_i = \frac{\omega}{Sum} S_i$

Mặt khác gọi K_i là số học sinh mà trường trung học cơ ở i thực sự thiếu ta cần tìm K_i sao cho $P = \sum_{i=1}^{M}(K_i - k_i)^2$ nhỏ nhất. Đặt $r_i = K_i - k_i$. Bài toán đưa về:

1. $S_i - L_i - k_i \geq r_i \geq S_i - H_i - k_i$

2. $\sum_{i=1}^{M} r_i = 0$

3. minimize $P = \sum_{i=1}^{M}(r_i)^2$

Các điều kiện này có thể được giải bằng một biến thể của thuật toán Criss-Cross, một mở rộng của phương pháp đơn hình nhưng với hàm mục tiêu không cần phải là một hàm tuyến tính. Hoặc ta có thể giải bằng hàm $fmincon()$ nếu sử dụng MATLAB.

3.1.3 Kết luận

Với mô hình và các mở rộng trên ta có thể thấy ở mỗi phiên bản của bài toán ta chỉ tập trung tối ưu một yếu tố duy nhất của bài toán (thời gian di chuyển ở mô hình gốc và độ phân bố của các học sinh ở các trường ở trong phần mở rộng). Đây là hai hướng giải quyết hoàn toán khác nhau điều này dẫn đến khi chúng ta đang cố tối ưu một yếu tố sẽ dẫn đến việc yếu tố còn lại sẽ bị tệ đi rất nhiều. Trong thực tiễn nhưng mô hình này là không thực tế vì không kiểm soát được một yếu tố có thể đưa đến những tình huống bất khả thi. Vì vậy dựa vào những mô hình lý thuyết trên, chúng tôi đưa ra một mô hình thực tế hơn.

Trên thực tế, chúng ta không cần cố đưa một yếu tố về cực tiểu mà chỉ cần nó bé hơn một giới hạn nào đó. Trong trường hợp này việc cố gắng tối ưu thời gian di chuyển là cực kì không càn thiết, chúng ta có thể hi sinh một vài giây hoặc vài phút để có thể cải thiển sự phân bố học sinh trong các trường một cách đáng kể. Vì vậy thay vì đặt tư tưởng là phải tối ưu cả hai hàm ta có thể thay một yếu tố thành điều kiện. Nếu vậy thì bài toán trên hoàn toàn có thể đặt lại như sau:

1. $S_i - L_i - k_i \geq r_i \geq S_i - H_i - k_i$
2. $\sum_{i=1}^{M} r_i = 0$
3. $\sum_{i=1}^{N} \sum_{j=1}^{M} X_{ij} D_{ij} \leq TCap$
4. minimize $P = \sum_{i=1}^{M} (r_i)^2$

Với $TCap$ là chặn trên của thời gian.

Đây là bài toán hoàn toàn giống như bài mở rộng 2 với điều kiện thêm là 3. $TCap$ sẽ là một con số mà người dùng sẽ nhập tùy theo nhu cầu của họ. Theo chúng tôi đây là một mô hình thực tế vì người dùng có thể thử nhiều trường hợp của $TCap$ để đưa ra các cách phân chia hợp lý tùy với các điều kiện khác của môi trường chứ không phải một kết quả cụ thể cho một tập input nhất định.

3.1.4 Mô hình trong trường hợp các học sinh có xếp loại học sinh

Do chúng ta phân bô học sinh theo khoảng cách và độ phân bố nên có thể các trường đưa đến tình trạng ở các trường sẽ không đồng đều về số lượng học sinh ở các phân loại khác nhau. Ví dụ ở một trường số lượng học sinh giỏi sẽ trở nên quá nhiều và ít hơn ở các trường khác. Điều nãy dẫn đến sự thiếu cân bằng trong môi trường học tập. Vậy nên với các số liệu có sự chênh lệch lớn trong số liệu, chúng ta nên tìm một mô hình có thể tối ưu điều này nhưng đồng thời không được gây thiệt thòi cho các học sinh.

Mô hình mà chúng tôi sau đây dựa trên ý tưởng xếp các học sinh vào theo tỉ lệ đầu ra của mỗi trường như tỉ lệ xếp loại học sinh lớp 9 tốt nghiệp từ trường. Chúng ta cần đưa các tỉ lệ này xấp xỉ nhau thê nên ta có thể đưa ra mô hình này:

Biến số	**Ý nghĩa**
G_{i1}	Số học sinh giỏi tốt nghiệp trường tiểu học i
G_{i2}	Số học sinh khá tốt nghiệp trường tiểu học i
G_{i3}	Số học sinh trung bình tốt nghiệp trường tiểu học i
G_{i4}	Số học sinh yếu tốt nghiệp trường tiểu học i
F_{i1}	Số học sinh giỏi vào trường trung học cơ sở i
F_{i2}	Số học sinh khá vào trường trung học cơ sở i
F_{i3}	Số học sinh trung bình vào trường trung học cơ sở i
F_{i4}	Số học sinh yếu vào trường trung học cơ sở i
X_{ij1}	Số học sinh giỏi từ trường tiểu học i đến trường trung học cơ sở j
X_{ij2}	Số học sinh khá từ trường tiểu học i đến trường trung học cơ sở j
X_{ij3}	Số học sinh trung bình từ trường tiểu học i đến trường trung học cơ sở j
X_{ij4}	Số học sinh yếu từ trường tiểu học i đến trường trung học cơ sở j
R_{i1}	Tỉ lệ trung bình học sinh giỏi tốt nghiệp trường trung học cơ sở i trong 5 năm gần đây
R_{i2}	Tỉ lệ trung bình học sinh khá tốt nghiệp trường trung học cơ sở i trong 5 năm gần đây
R_{i3}	Tỉ lệ trung bình học sinh trung bình tốt nghiệp trường trung học cơ sở i trong 5 năm gần đây
R_{i4}	Tỉ lệ trung bình học sinh yếu tốt nghiệp trường trung học cơ sở i trong 5 năm gần đây

Và dựa vào những nhận xét đã đặt ra ở trên:

$\sum_{j=1}^{M} X_{ijk} = G_{ik}$

$\sum_{i=1}^{N} X_{ijk} = F_{ik}$

$F_{i1} : F_{i2} : F_{i3} : F_{i4} = R_{i1} : R_{i2} : R_{i3} : R_{i4}$

minimize $F = \sum_{k=1}^{4} \sum_{i=1}^{N} \sum_{j=1}^{M} X_{ijk} D_{ij}$

Với mô hình trên nghiệm của chúng ta có thể là những số không nguyên, do khi áp dụng vào thực tế ta có thể bỏ qua những sai số nhỏ nên ta có thể làm tròn các kết quả đến số nguyên gần nhất nhưng vẫn bảo đảm được kết quả ứng dụng được vào thực tiễn.

3.2 Mô hình theo nguyện vọng học sinh

Hai mô hình trên được phát triển từ ý tưởng tối ưu hóa thời gian di chuyển của các học sinh. Nhưng sau khi tham khảo một số tài liệu của một số quốc gia khác, điều này dẫn đến sự bất công do các học sinh không được vào những ngôi trường mà mình yêu thích mà bị phân tuyến một cách cưỡng bức vào các trường gần nhà. Điều này tuy có tối ưu được các yếu tố như thời gian di chuyển hay tỉ lệ học sinh của các trường, nhưng lại làm gia tăng số lượng học sinh học những trường không mong muốn hoặc những ngôi trường không phù hợp. Hệ quả của việc này là sự gia tăng trong tham nhũng trong giáo dục, cha mẹ sử dụng tiền bạc và quan hệ để đưa con vào các trường mình muốn. Chính vì vậy, chúng tôi đã phát triển một mô hình mới dựa trên các nguyện vọng của học sinh, giảm thiểu tối đa việc học sinh bị phân vào các trường trái với mong muốn cũng như sự xuất hiện của các hiện tượng tiêu cực. Sau khi tìm hiểu thì chúng tôi tìm ra một mô hình tương tự để giải quyết được tình trạng này, đó chính là mô hình hôn nhân ổn định của Gale-Shapley[2].

3.2.1 Thuật toán cặp ghép ổn định

Với mô hình này học sinh sẽ đánh giá các trường theo độ yêu thích. Do sự phân bố yêu thích của các học sinh sẽ không đồng đều với nhau nên chắc chắn không phải lúc nào cũng sắp xếp được các học sinh vào những trường mà họ muốn. Do vậy để cho công bằng chúng ta còn phải xếp hạng học sinh theo các tiêu chí, cụ thể trong đây có thể là điểm thi tốt nghiệp tiểu học của các em. Do các trường cũng không hề có thông tin gì khác về các học sinh nộp vào trường mình, ta có thể coi thứ tự điểm của các em cũng là thứ tự ưu tiên của các trường. Chúng ta sử dụng thuật toán sau vào mô hình:

Bước 1: Xét học sinh từ điểm cao tới thấp, nếu trường xếp hạng cao nhất của họ còn chỉ tiêu thì xếp học sinh đó vào ngôi trường đó không thì coi như các học sinh đó bị loại ở bước 1.

Bước k: Chỉ xét các học sinh bị loại ở bước $k-1$. Xét học sinh từ điểm cao tới thấp, nếu trường xếp hạng cao thứ k của họ còn chỉ tiêu thì xếp học sinh đó vào ngôi trường đó không thì coi như các học sinh đó bị loại ở bước k.

Sau khi đã thực hiện thuật toán trên thì tất cả học sinh đều đã tìm được một trường thích hợp cho mình. Thuật toán này đảm bảo các học sinh được phân bô theo trường mình thích và các trường cũng sẽ tuyển được những học sinh thích hợp. Thuật toán này bảo đảm được hai điều sau:

1. Nếu số học sinh không bé hơn tổng số chỉ tiêu của các trường thì các học sinh luôn có trường. Tức nếu $\sum_{i=1}^{M} Ct_i \leq Num$ thì ở bước cuối của thuật toán số học sinh bị loại bằng 0(với Ct_i là chỉ tiêu của trường i và Num là tổng số học sinh).

2. Đảm bảo sự phân chia luôn đảm bảo tính công bằng. Tính công bằng được tính là không tồn tại một cặp học sinh i và j vào trường Sc_i và trường Sc$_j$ mà $score_i < score_j$ nhưng $q_{i,Sc_i} < q_{j,Sc_j}$(với $score_i$ là điểm tốt nghiệp học sinh i và $q_{i,j}$ là hạng độ yêu thích của trường j với học sinh i.

Với mô hình như vậy chúng ta sẽ bảo đảm các học sinh sẽ có cơ hội vào được trường yêu thích của mình. Xếp hạng các em sẽ cũng là động lực để các em cố gắng hơn nữa để có thể vô được ngôi trường mình yêu thích.

3.2.2 Ứng dụng trong thực tế

'

Với mô hình trên nếu muốn biến đổi thành một cơ chế áp dụng được sẽ phải qua rất nhiều thay đổi do những yếu tố khác ảnh hưởng. Sau đây chúng tôi sẽ nói qua một số những thay đổi nhất định để phù hợp với những điều kiện trong thực tế:

- Do trong thực tế số lượng trường trung học cơ sở có thể rất lớn nên việc cho mỗi em học sinh xếp hạng tất cả các trường theo thứ tự là không khả thi. Vậy một biến đổi có thể là cho các học sinh đưa ra một danh sách các trường có giới hạn và theo thứ tự rồi giả sử độ yêu thích của các trường còn lại là bằng nhau. Do các trường còn lại có độ ưu tiên băng nhau, sau khi áp dụng thuật toán thì các học sinh dư còn lại có thể phân tối ưu theo mô hình 1.

- Do số lượng điểm thi phân biệt là hữu hạn và nhỏ hơn rất nhiều so với số học sinh nên có thể khi các trường gần tới giới hạn một số học sinh bằng điểm nhau có thể lại có những kết quả khác nhau vậy đem lại kết quả không công bằng. Do vậy có thể một trường có thể không cố định số lượng tuyển sinh mà có thể có một phần trăm học sinh có thể nhận thêm nếu có trường hợp có những học sinh bằng điểm. Điều này sẽ công bằng hơn trong thực tế và tránh sự ưu tiên về các yếu tố khác.

- Do không phải trường nào cũng muốn nhấn mạnh về học tập nên việc chỉ xét điểm tốt nghiệp là điều không hợp lý. Vì thế ta có thể để các trường quyết định về một số cách để xếp hạng hồ sơ để có thể đa dạng hóa các chỉ tiêu cũng như cho các em có định hướng tốt hơn về một số yếu tố khác ngoài chỉ việc học thuật.

4 Đánh giá mô hình và ứng dụng vào thực tế

Ở mô hình 1, chúng ta tập trung vào việc tối ưu hai yếu tố là thời gian và sự phân bố học sinh trong các trường trung học cơ sở. Điều này giúp cho tỉ lệ phân bố các học sinh giữa các trường là đồng đều nhau, không làm cho một số trường dư hoặc thùa chỉ tiêu tối ưu một cách quá đáng. Ngoài ra việc tối ưu hóa thời gian bằng biến chặn có thể cho người sử dụng mô hình có thể linh hoạt điều chỉnh thời gian giới hạn để có thể tối ưu các yếu tố khác một cách dễ dàng hơn ở các trường hợp cụ thể khác nhau. Ngoài ra chúng tôi còn đưa ra một mô hình biến thể của mô hình một nhằm có thể phân chia các học sinh, sao cho chất lượng học sinh của mỗi trường là phù hợp với các yếu tố của trường đó nhất. Tuy nhiên, tất cả các biến thể và mở rộng này đều không dựa trên các yếu tố đến từ phía học sinh. Điều này phần nào làm cho việc phân tuyến trở nên ép buộc và sẽ khiến cho sẽ mất cơ hội phát triển ở những môi trường phù hợp với các em học sinh. Ngoài ra với các điều kiện thêm vào như các lớp ngoại ngữ hay năng khiếu sẽ làm mô hình trở nên khó kiểm soát.

Ở mô hình 2, chúng ta tập trung giải quyết những bất cập mà mô hình 1 có thể mang lại. Việc không được học trong môi trường phù hợp có thể dẫn đến việc các em học sinh và cha mẹ các em không hài lòng với việc phân tuyến và cảm thấy bất công. Điều này sẽ dẫn đến các vấn đề tiêu cực như sử dụng các quan hệ để đưa con em mình vào những ngôi trường phù hợp. Vì vậy, việc đưa ra một mô hình dựa trên sự yêu thích của học sinh là một giải pháp của chúng tôi với vấn đề này. Mô hình tối ưu các lựa chọn của các em học sinh sẽ phần nào đặt sự yêu thích của học sinh lên trên cùng giúp cho các em có cơ hội nhiều hơn để vô được những trường phù hợp với mục tiêu của từng bạn. Ngoài ra sự ưu tiên việc xét tuyển theo năng lực sẽ phần nào giúp các học sinh có động lực để trau dồi và hoàn thiện bản thân để có thể vô được môi trường mình thích. Ngoài ra các trường được đặt các chỉ tiêu khác ngoài năng lực học vấn cũng giúp trường dễ dàng nhận những học sinh phù hợp với những hướng phát triển riêng. Tuy nhiên, việc phân theo các năng lực và nguyện vọng như vậy sẽ gây ra sự phân bố không đồng đều giữa các trường về học sinh. Trong trường hợp tệ nhất, một trường có thể nhận các học sinh hoàn toàn khác với tiêu chí đề ra của trường.

Khi áp dụng vào thực tế khó có một mô hình cụ thể nào có thể luôn đưa ra được cách sắp xếp tối ưu cho bài toán. Tuy nhiên, với tình hình hội nhập hiện nay, các học sinh được tiếp cận với những thứ mình yếu thích dễ dàng hơn bao giờ hết. Điều này khiến cho việc đưa ra một mô hình có thể tối ưu về niềm đam mê và sự yêu thích của các em là quan trọng hơn cả. Tuy nhiên để có thể áp dụng được mô hình này một cách tối ưu trong thực tế thì những cải tiến và biến thể là cần thiết. Các điều kiện như giới hạn lại các nguyên vọng, hay kết hợp với mô hình 1 như chúng tôi đã nói ở trên là một hướng hoàn toàn khả thi cho một phương pháp phân tuyến phù hợp với tình hình hiện nay.

5 Tài liệu tham khảo

[1] : https://towardstransparency.vn/wp-content/uploads/2014/07/GCR_Vietnam-Case-Study_VN_FINAL.pdf

[2] : http://cramton.umd.edu/market-design/gale-shapley-college-admissions.pdf

[3] : http://static.thanhnien.com.vn/Uploaded/congthang/2016_05_30/khtsphantuyenq tanbinh_UOBK.pdf

6 Phụ lục

Mô hình tuyến tính

Các biến của hàm:

1. Mảng F

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1	3	6	17	6	7	15	10	14	12	5	1	9	11	7	5	13	10	3

2. Mảng Beq

	1	2	3	4	5	6
1	365	184	209	185	98	168

3.Mảng B

	1	2	3	4	5	6	7	8	9	10	11	12
1	700	200	450	-600	-100	-350	0	0	0	0	0	0

13	14	15	16	17	18	19	20	21	22	23	24
0	0	0	0	0	0	0	0	0	0	0	0

4.Mảng Aeq

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2	0	0	0	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0
3	0	0	0	0	0	0	1	1	1	0	0	0	0	0	0	0	0	0
4	0	0	0	0	0	0	0	0	0	1	1	1	0	0	0	0	0	0
5	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	0	0	0
6	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1

5.Mảng A

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1	1	0	0	1	0	0	1	0	0	1	0	0	1	0	0	1	0	0
2	0	1	0	0	1	0	0	1	0	0	1	0	0	1	0	0	1	0
3	0	0	1	0	0	1	0	0	1	0	0	1	0	0	1	0	0	1
4	-1	0	0	-1	0	0	-1	0	0	-1	0	0	-1	0	0	-1	0	0
5	0	-1	0	0	-1	0	0	-1	0	0	-1	0	0	-1	0	0	-1	0
6	0	0	-1	0	0	-1	0	0	-1	0	0	-1	0	0	-1	0	0	-1
7	-1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
8	0	-1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
9	0	0	-1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
10	0	0	0	-1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
11	0	0	0	0	-1	0	0	0	0	0	0	0	0	0	0	0	0	0
12	0	0	0	0	0	-1	0	0	0	0	0	0	0	0	0	0	0	0
13	0	0	0	0	0	0	-1	0	0	0	0	0	0	0	0	0	0	0
14	0	0	0	0	0	0	0	-1	0	0	0	0	0	0	0	0	0	0
15	0	0	0	0	0	0	0	0	-1	0	0	0	0	0	0	0	0	0
16	0	0	0	0	0	0	0	0	0	-1	0	0	0	0	0	0	0	0
17	0	0	0	0	0	0	0	0	0	0	-1	0	0	0	0	0	0	0
18	0	0	0	0	0	0	0	0	0	0	0	-1	0	0	0	0	0	0
19	0	0	0	0	0	0	0	0	0	0	0	0	-1	0	0	0	0	0
20	0	0	0	0	0	0	0	0	0	0	0	0	0	-1	0	0	0	0
21	0	0	0	0	0	0	0	0	0	0	0	0	0	0	-1	0	0	0
22	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	-1	0	0
23	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	-1	0
24	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	-1

Code của thuật toán cặp ghép ổn định

Đây là code của thuật toán tìm cặp ghép ổn định ở trên. Với $choose[i]$ là trường của học sinh i, $p[i]$ là sắp xếp các thứ tự các trường của học sinh i, $ct[i]$ là chỉ tiêu học sinh trường thứ i.

```
for (int i=1;i<=n;i++)
    choose[i]=0;
for (int i=1;i<=m;i++)
{
    for (int j=1;j<=n;j++)
        if (!choose[i])
    {
        if (num[p[j][i]]<ct[p[j][i]])
        {
            num[p[j][i]]++;
            choose[i]=p[j][i];
        }
    }
}
```