Tabling on Campus

Operations Research II

Jae Eun Back
Vasudha Chanana
Kiyeun Kim
Kwan Kim

Introduction

Pittsburgh needs money. They need money to fill their budget gap. So who do they turn to? Students! Pittsburgh Mayor recently proposed 1 percent tax on students' yearly tuition. How ridiculous is this? How can the city prevent students from receiving education? There are students who make about $3,500 per year working part time jobs at night after studying during the day. These hardworking students are taxed from $200 up to $500. This amount of money can be better used to pay for rent or food for these students. Thus, our group decided to do our project on finding an efficient way to prevent the city from taxing us.

There are several ways to deal with this situation. One way to help students avoid this unreasonable tax is to collect signatures from them and bring the signatures to public hearings and let people know. This will not only let other people hear about this irrational tax but also find more people to fight with us against the Mayor of Pittsburgh. Since we, students, cannot find many options in collecting signatures, we thought tabling on campus is a good way to achieve our target. In reality, it might be the best that we can do. Therefore, we decided to find an efficient way to collect signatures on campus.

Tabling occurs for several reasons everyday on campus. It is the most effective way to interact with students as students walk to and from classes, food, and their residences. The three key reasons people table are to fundraise, to advertise events and keep the student body informed of ongoing happenings, and to get signatures (for petitions). Thus, to get signatures for the tuition tax repeal, tabling on campus appeared to be a viable option

Keeping the above in mind, the objective of our project is to find the optimal places to table on campus to collect the maximum number of signatures. We started with picking some possible locations for tabling. Intuitively, some areas are more suitable than others. For example, many students probably make visits to UC or Hunt Library more than they do to than other buildings for food and studying. Also, students might go to certain buildings more than to others depending on the time of the day. In order to gain a better understanding of when and where students are most likely to be, we conducted a campus survey.

Survey

Understanding the traffic pattern on campus is crucial in formulating a program to obtain a solution to our problem. Thus, our team created a survey and distributed to two hundred students on campus. In order to make a survey sheet, we first took the campus map from the school website and simplified it by choosing the buildings that we were interested in. We realized that including all buildings on campus is unnecessary since we only need to consider the buildings where we can put tables in front of. We selected the currently popular locations for tabling as well as the possible locations other than the current favorite spots. We also combined two buildings as one unit when two buildings are connected. Our final list of buildings for our map consists of Hunt Library, Baker and Porter Halls (BH&PH), Tepper building, Doherty and Wean Halls (DH&WH) , the University Center (UC), and Warner Hall. The aforementioned six buildings according to their relative locations to each other and paths from one building to another were set as vertices and edges respectively. The resultant map is an undirected graph with six vertices and fourteen edges as shown below. (Figure 3)

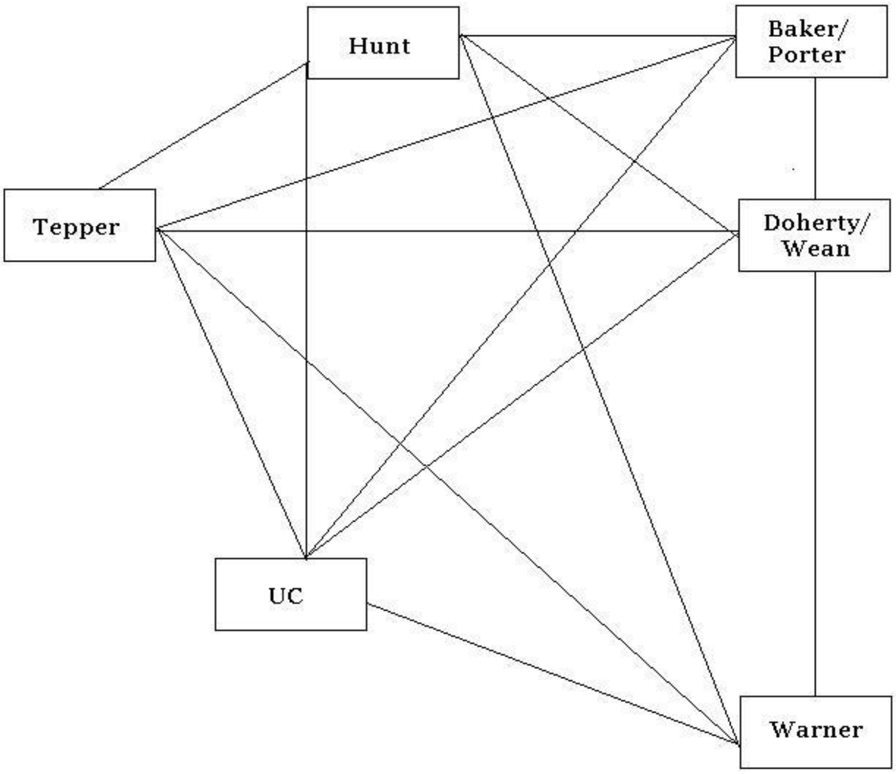

Figure 1: The Actual Survey Map

Survey participants were given the map and asked to mark the route they usually pass on regular school days in different times—Morning (8am – 11am), Around Noon (11am – 2pm), and Late Afternoon (2pm – 5pm). We distinguished three time slots because we suspected that the traffic patterns in each slot can possibly be different. When interpreting the survey, we assign a weight to each building by adding one to incident vertices if the edge was marked. We then constructed the table representing the number of students passing by each building at the certain time slot (Figure 4). The alphabets A,B, and C symbolize time slots, and variables $x_{1,} \cdots, x_6$ symbolize Hunt Library, Baker and Porter Halls (BH&PH), Tepper building, Doherty and Wean Halls (DH&WH) , the University Center (UC), and Warner Hall in order.

	A	B	C	Sum	Sum of Proportions
X1	56	80	88	224	134
X2	48	54	56	158	91
X3	66	68	58	192	109
X4	56	66	52	174	101
X5	58	110	68	236	146
X6	30	22	44	96	53

Figure 2: The Survey Result

We added the column totals because we were also interested in knowing the number of students walking by a building for the whole day, not only at a specific time slot. We created the last column under the assumption that the responding rate to tabling will be different depending on the time slot. We introduced responding rates 0.4, 0.8, and 0.6 to each time slot to make our model reflecting the more realistic situation. Based on our observation, 0.4 was assigned to the morning time, because people are usually still half sleeping around that time, and often do not have time to stop at a table. However, people are more willing to spare their time around noon after lunch. We still did not give a probability of 1.0 to this time slot because we have to eliminate the case where people passing by buildings have already signed for the petition. The Late Afternoon slot gets a probability of 0.6 with the assumption that some people will hesitate to answer to tabling after a long day, and that some of them have already participated earlier. The last column was calculated by adding the multiplication of the column A and 0.4, the column B and 0.8 and the column C and 0.6. This column approximates the number of students who will be signing the petition in a day.

Integer Programming Formulation

Using the data that we had obtained, we formulated four different models as an integer programming. The first model assumes that student organization will table at one location for a day as how people do currently. However, to get the large number of signatures from students in a short time, moving a table around to different spots can be a considerable option, and we put this as the second model. The third model considers the different responding rate, and the fourth model supposes that we can have more than one table at each location. The first model assumes that student organization will table at one location for a day as how people do currently. However, to get the large number of signatures from students in a short time, moving a table around to different spots can be a considerable option, and we put this as the second model. The third model considers the different responding rate, and the fourth model supposes that we can have more than one table at each location. There is no limit on how many tables we can have in total.

Regardless of which assumption we take, every model can be written in a form:

$$\text{maximize} \sum_{i=1}^{N} a_i x_i$$

$$\text{constraint to} \sum_{i=1}^{N} c_i\, x_i \leq K$$

and x_i is an integer

We want to maximize the number of signature which is represented by $\sum_{i=1}^{N} a_i x_i$. Although the number of people passing by a building will not be equal to the number of people signing for the petition at the table in front of the building in reality, we made an assumption that it will be. Thus the coefficients a_i's can be determined from the data that we already have from the survey. While the values of a_i's are different for each model, the coefficients c_i's and the constant K have the fixed values: $c_{1,} c_{2,} c_{3,} c_{4,} c_{5,} c_{6,}$ = 5, 4, 4, 3, 5, 2, K =11. The higher c_i indicates that it costs more to table in front of the building x_i. For our tabling problem, we chose c_i based on how competitive it is to book a table at the location, whether the location is indoor or outdoor, and how hard it is to advertise to non-CMU people.

The value 11 for the constant K was determined after a few experimental tries to get the combination of locations as the solution. For example, setting K equals to 5 will return a_5 most likely, because x_5 is the UC where the most students pass during lunch time and for a day, where our model becomes simple number-comparing problem. Setting K with a huge number such as 30 will tell us to put too many tables around the campus which is unrealistic. We concluded that setting K as 11 will gives us the most interesting solution while describing the reality pretty well.

Modeling

For Model 1, we considered that there were only a few tables available to the student organization, and they can only have one table per location due to location constraints. Additionally, a table had to be set up on a location for the entire day and could not be moved. Thus, we get the formulation below:

$$\text{Maximize } z = 224x_1 + 158x_2 + 192x_3 + 174x_4 + 236x_5 + 96x_6$$
$$\text{constraint to } \quad 5x_1 + 4x_2 + 4x_3 + 3x_4 + 5x_5 + 2x_6 \leq 11$$
$$0 \leq x_1, x_2, x_3, x_4, x_5, x_6 \leq 1, \text{ integers}$$

Note that this is like the knapsack problem with the weights corresponding to costs, and values corresponding to the number of people we want to target. Thus, possible algorithms to use could have been Dynamic programming, Branch and Bound, or Cutting Plane. We mostly used Matlab and Mathematica to solve integer programming problems.

The results from Model 1 were that we should have one table each at Tepper, U.C. and Warner for an entire day. These are interesting locations since only the U.C. has been used to table at historically.

For Model 2, we wanted to find the traffic based on different time slots, so we assumed that we can change locations at different time slots. However, we can still only have at most 1 table at each location. For the morning time slot (8 AM – 11 AM), we used the appropriate data and found the following integer programming formulation:

$$\text{Maximize } z = 56x_1 + 48x_2 + 66x_3 + 56x_4 + 58x_5 + 30x_6$$
$$\text{constraint to } \quad 5x_1 + 4x_2 + 4x_3 + 3x_4 + 5x_5 + 2x_6 \leq 11$$
$$0 \leq x_1, x_2, x_3, x_4, x_5, x_6 \leq 1, \text{ integers}$$

The results were that we should table outside Baker/Porter, Tepper, and Doherty/Wean in the morning. Baker/Porter is an interesting location since it is not historically used for tabling.

The noon slot (11 AM – 2 PM) yielded similar results:

$$\text{Maximize } z = 60x_1 + 54x_2 + 68x_3 + 66x_4 + 110x_5 + 22x_6$$
$$\text{constraint to } 5x_1 + 4x_2 + 4x_3 + 3x_4 + 5x_5 + 2x_6 \leq 11$$
$$0 \leq x_1, x_2, x_3, x_4, x_5, x_6 \leq 1, \text{ integers}$$

Note that for this model, 110 people pass by the U.C. and only 22 by Warner. This may be due to the fact that it is lunchtime, and most people are on campus already. For the above time slot, we found that we should have 1 table each around noon at Tepper, U.C., and Warner. Warner was selected due to its low costs, even though it had a low value attached to it.

For the late afternoon slot (2 PM- 5 PM), we formulated our problem as:

$$\text{Maximize } z = 88x_1 + 56x_2 + 58x_3 + 52x_4 + 68x_5 + 44x_6$$
$$\text{constraint to } 5x_1 + 4x_2 + 4x_3 + 3x_4 + 5x_5 + 2x_6 \leq 11$$
$$0 \leq x_1, x_2, x_3, x_4, x_5, x_6 \leq 1, \text{ integers}$$

We then found that we could table at that time, with one table each at Hunt, Tepper and Warner. Hunt seems reasonable since people go to study there after class.

For model 3, we considered not only the different time slots from model 2 but also the probability of each time. We assigned reasonable probabilities to each time slot based on the assumption that not everyone who passes by the table will respond to advertising or sign the petition. We then took the different characteristics of each time into consideration.

The first time slot is Morning from 8 AM to 11 AM. In the morning, people are likely to be still asleep and not as willing to participate as well in the early morning. We assigned the probability of 0.3 to the first time slot. Second time slot is Around noon from 11 AM to 2 PM. Around noon, people are more awake and more willing to respond to survey than they are in the morning, but it is likely that some people have already participated in the morning. So we assigned the probability of 0.8 to this time slot. Thirdly, Late afternoon slot ranges from 2 PM to 5 PM. We assumed that people are not as willing to respond to survey as they are around noon because they are tired after a long day and want to go home. It is also possible that some people have already participated in the morning or afternoon. We assigned the probability of 0.6 to Late afternoon. We further assume that we can have at most one table per location. Our maximization function is the following:

$$\text{Maximize } z = 134x_1 + 91x_2 + 109x_3 + 101x_4 + 146x_5 + 53x_6$$
$$\text{constraint to } \quad 5x_1 + 4x_2 + 4x_3 + 3x_4 + 5x_5 + 2x_6 \le 11$$
$$0 \le x_1, x_2, x_3, x_4, x_5, x_6 \le 1, \text{ integers}$$

As a result, the optimal solution is to table at Tepper, UC, and Warner.

For model 4, we assumed additionally that it is possible to have more than one table at each location to get more attention. This assumption is reasonable and practical because it actually happens frequently in real life and sometime it is better to have multiple tables at one location than to have one table at multiple locations. We covered the cases where we can locate 2 or 3 tables maximum at each place because it is impractical to have 4 or more tables at one location. We keep the same probability assumptions from Model 3. Our maximization function is the following:

$$\text{Maximize } z = 134x_1 + 91x_2 + 109x_3 + 101x_4 + 146x_5 + 53x_6$$
$$\text{constraint to } \quad 5x_1 + 4x_2 + 4x_3 + 3x_4 + 5x_5 + 2x_6 \le 11$$
$$0 \le x_1, x_2, x_3, x_4, x_5, x_6 \le 1, \text{ integers, maximum number of tables}$$

Intuitively, it makes sense that the only additional constraint from model 3 is to assign each x_i the maximum possible number of tables at each location.

The optimal result with maximum number of 2 tables at each location is that we can have 2 tables outside Doherty/Wean and table at UC.

Similarly, we found that with a maximum of 3 tables at each location, we should have 3 tables outside Doherty/Wean (mainly due to heavy traffic and moderate costs) and one outside Warner. Running similar analyses for higher number of tables did not make sense since we would then exceed our maximum cost of 11.

Conclusion:

In retrospect, we initially wanted to find out when and where to table shift on campus to maximize in collecting signatures from the students. After our project, we learned that Doherty and UC isn’t always the best place to table shift. So we recommend campus organizations to explore other buildings depending on how any tables they are able to reserve and how many tables they are willing to set up instead of just tabling in front of Doherty or UC, which they normally do.

Although we used our models to specifically maximize in number signatures on campus, our models can be used for so many other productive projects. For example, we can use our models to find out the best places to locate garbage cans or pay phones around the city, best locations for clinic locations or even best store locations when looking to open up a new store in the city.

In future studies, it will be definitely helpful if we had more data on more building since more data will always improve our model. Also, there can be more possible selections for tabling locations. If we had more time and funds available, it might be a better idea to set parents as our target since they are the ones who are paying for the tuition most of the time.