Advanced Optimization Project: The Hipster Tourist Problem

Ji Yihong, Li Tiantian, Kaviyan, Wan Ying

March 2020

1 Data Collection

We assumed that the Hipster Tourist will be travelling by public transport, as it is the most economical mode of transport in Singapore. Thus, we used Google Maps to capture the time and distance required to travel between each location. The shortest travelling time taken by public transport was taken and rounded up or down to the nearest multiple of 15 except for a few exceptions:

1. Pulau Ubin

We added 15 minutes of additional travelling time to account for the ferry ride as Google Maps only had a route to Changi Village.

2. Sentosa and Singapore Zoo/Night Safari/River Safari

The shortest travelling time to these places sometimes included the Sentosa Monorail/Mandai Express. However, considering that an additional fee had to be paid for these services and to maintain consistency in travelling cost, we chose the shortest paths which did not include them.

Since we were unable to extract the exact distance travelled by public transport on Google Maps, we estimated this distance with the distance travelled by car instead. Furthermore, we collected this data at a time when the traffic was non-existent across Singapore, to ensure consistency. Then, we calculated the transportation cost between each location by roughly following Singapore's MRT train fare model,0.92 dollar for the first 3.2km and 0.1 dollar for every 1km after.

Assumptions

To ease data collection, we clustered similar locations together as the same location. The locations in this combined location will thus have zero transportation cost and zero travelling time between each other. The general rule of thumb we followed was that the locations were 5-10 min of walking distance apart. The locations that were clustered together were:

Clustered Location	Sub-Location
Mandarin Orchard Singapore	Mandarin Orchard Singapore, Jamie's Italian
Raffles Place	Odette, 1 Altitude
Tian Tian Hainanese Chicken Rice Bedok	Tian Tian Hainanese Chicken Rice Bedok, Simpang Bedok
Marina Bay Sands	ArtScience Museum, Gardens by the bay
Singapore Zoo	Singapore Zoo, Night Safari, River Safari
Punggol Waterway Park	Coney Island, Punggol Waterway Park
Jurong East	Miam Miam French and Japanese, Science Centre Singapore, The Rink

Figure 1: Location Clustering

2 Task 2 Model Construction

2.1 Assumption

- Revisiting the same place is not allowed(except for hotel)
- Have discrete time intervals(15 minutes for each unit time interval)

2.2 Definition of Variables

Sets

A	Set of attractions
R	Set of restaurants
H	Set of hotels
P	Set of all nodes, i.e. $P = A \cup R \cup H$

Parameters

o_i	$i \in A$	$\mathbb{1}\{\text{Attraction i is outdoor}\}$
t_{ij}	$i, j \in P$	Time taken to travel from place i to j
c_i	$i \in A$	Cost of attraction i
π_{ij}	$i, j \in P$	Transportation cost between places i and j
p_i	$i \in P$	Processing time at places i
D		Diner time period
L		Lunch time period
T		Time span

Variables

x_{it}	$i \in P, t \in T$	$\mathbb{1}\{\text{Arrive place } i \text{ at time } t\}$
e_{ij}	$i, j \in P$	$\mathbb{1}\{\text{Transport from } i \text{ to } j\}$

2.3 Objective

$$\min \quad \sum_{i,j \in P} \pi_{ij} e_{ij} + \sum_{i \in A} \sum_{t \in T} x_{it} c_i$$

2.4 Constraints

$$\sum_{t\in T}\sum_{i\in P} x_{it}p_i + \sum_{i,j\in P, i\neq j} t_{ij}e_{ij} \leq |T| \tag{1}$$

$$\sum_{i\in T} x_{it} \leq 1 \qquad \forall i \in A \cup R \tag{2}$$

$$\sum_{i\in T} x_{it} = 2 \qquad \forall i \in H \tag{3}$$

$$x_{i0} = 1 \qquad \forall i \in H \tag{4}$$

$$\sum_{t\in L}\sum_{i\in R} x_{it} = 1 \qquad \forall Q \in \{L, D\} \tag{5}$$

$$\sum_{t\in F}\sum_{i\in A} x_{it}o_i = 0 \tag{6}$$

$$\sum_{t\in T}\sum_{i\in A} x_{it} \geq 8 \tag{7}$$

$$\sum_{i\in A}\sum_{t\in T} x_{it}o_i \geq 3 \tag{8}$$

$$\delta_W + \delta_E + \delta_N + \delta_S + \delta_C \geq 3 \tag{9}$$

$$\sum_{t\in T}\sum_{i\in R\cup A} U_i x_{it} \leq 1 + M\delta_U \qquad \forall U \in \{W, E, S, N, C\} \tag{10}$$

$$\sum_{t\in T}\sum_{i\in R\cup A} U_i x_{it} \geq 1 - M(1-\delta_U) \qquad \forall U \in \{W, E, S, N, C\} \tag{11}$$

$$M = |P| \tag{12}$$

$$\sum_{t\in T} x_{it}t + p_i + t_{ij} - \sum_{t\in T} x_{jt}t \leq N(1-e_{ij}) \qquad \forall i, j \in P \tag{13}$$

$$N = |T| \tag{14}$$

$$\sum_{t\in T} x_{it} + \sum_{t\in T} x_{jt} \geq 2e_{ij} \qquad \forall i, j \in P \tag{15}$$

$$\sum_{j\neq i} e_{ij} = \sum_{t\in T} x_{it} \qquad \forall i \in P \tag{16}$$

$$\sum_{j\neq i} e_{ij} + \sum_{j\neq i} e_{ji} = 2K_i \qquad \forall i \in P \tag{17}$$

$$x_{it} = 0 \qquad \forall t \notin L \cup D,\ \forall i \in R \tag{18}$$

$$K_i \in \{0, 1\} \tag{19}$$

$$\delta_l \in \{0, 1\} \qquad \forall l \in \{W, E, S, N, C\} \tag{20}$$

$$x_{it} \in \{0, 1\} \qquad \forall i \in P,\ \forall t \in T \tag{21}$$

$$e_{ij} \in \{0, 1\} \qquad \forall i, j \in P \tag{22}$$

2.5 Explanation

The detailed description of constrains are in appendix. We would pick up several important constraints to demonstrate the correctness of our modeling.

- Constraint (24) ensures that each place is either inside a cycle or it is not visited at all.
- Constraint (22) and (23) ensures that if there is an edge between two places, then these two places must be visited at sometime.
- Constraint (20) is understood as if there is a transportation from i to j two, then the visiting time of j must not be earlier than the visiting time of i plus the processing time of i plus the transportation time between theses two places.
  This is to guarantee that we could assign each visited node to the time slot while maintaining the correct chronological order.
- Constraint (25) guarantees that the cycle starts from hotel, thus there does not exist multiple disjoint loops in our graph.
- Constraints (20)-(25) could be visualized as Figure 2.
- Big M method is used to linearise constraints (10)-(18), which are to make sure at least two regions other than central region is visited.

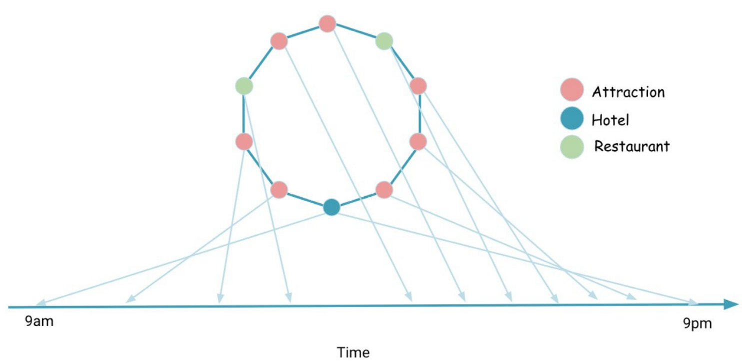

Figure 2: Visualization of Modelling

3 Interpretation of Output

Our program terminates after a few hours with an optimal solution. We plot the solution both using the edge variables and time-index variables. As we can see in the table below, the solution satisfies all of Task 2's constraints:

- **Hotel constraint**
  Tourist leaves hotel no earlier than 9am and reaches no later than 9pm.
- **Meals constraint**
  Tourist has meals at different locations and has 1hr lunch and dinner from 12pm to 2pm and 6pm to 8pm respectively.
- **Attraction constraint**
  Tourist visits at least 8 attractions and spends 30 min at each attraction
- **Outdoor time constraint**
  Tourist visits at least 3 outdoor attractions and is unable to visit outdoor attractions from 11am to 4pm.
- **Region constraint**
  at least 3 regions visited.

The total cost of the day amounts to SGD 90.47

Time	Location	Type of Location	Region
9:00 a.m	Mandarin Orchard	Hotel	Central
9:15 a.m	Botanic Gardens	Outdoor Attraction	Central
10:15 a.m	Chinese Garden	Outdoor Attraction	West
11:00 a.m	The Rink	Indoor Attraction	West
12:00 p.m	Miam Miam French and Japanese Café Kitchen	Dining	West
1:00 p.m	Science Centre Singapore	Indoor Attraction	West
2:00 p.m	Singapore Discovery Centre	Indoor Attraction	West
4:15 p.m	Sungei Buloh Wetland Reserve	Outdoor Attraction	North
6:00 p.m	Buey Tahan See-Food CCK	Dining	West
7:15 p.m	Coney Island	Outdoor Attraction	North-East
7:45 p.m	Punggol Waterway Park	Outdoor Attraction	North-East
9:00 p.m	Mandarin Orchard	Hotel	Central

Figure 3: Task 2 Solution

The solution shows that the tour is indeed a cycle and covers most regions of Singapore. The tourist did have a wonderful tour around the nation!

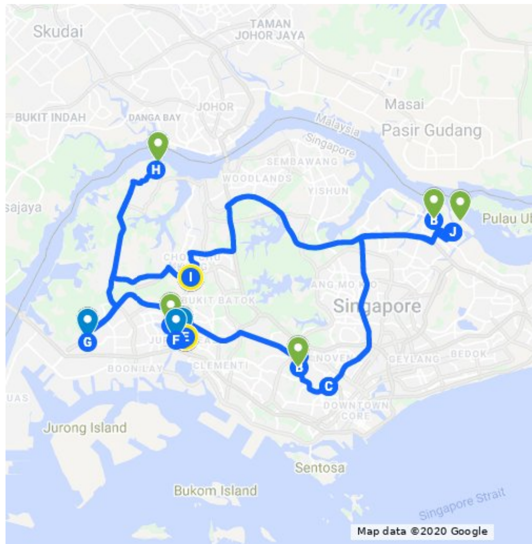

Figure 4: Map View of Optimal Solution

4 Task 3 Model Construction

The definition of diversity is to go to at least half of the total attractions, meanwhile minimizing the difference between number of outdoor attractions and indoor attractions visited. We also introduce constraints to balance the workload on two days.
We inherit the model from task 2, the modifications are now we need to have two intersecting cycles at chosen hotel.

4.1 Assumptions

- It is unnecessary to switch between hotels.

  Switching hotels does not positively influence the objective function that we will define in next section.
  Instead of spending time on switching hotels, it is actually better if the tourist visits a new attraction. However, we would initialize our tour at five hotels and compare the quality of solutions

- Extension of lunch time and dinner time to our hour earlier and our hour later.

  Another assumption we made is that in order to facilitate our search for a feasible solution, we extend the dining time for lunch and dinner for both days in task 3.

- Drop the constraint about no outdoor activity during 11am to 4 pm .

Since the forbidden outdoor time constraint exponentially increased the run time for our model in task 2, we removed this constraint in task 3 to find an optimal solution more quickly.

4.2 Definition of Diversity

Objective

$$\min \quad \sum_{i\in P}\sum_{j\in P}(e_{ij}+l_{ij})o_i + \sum_{i\in P}\sum_{j\in P}(e_j+l_j)(1-o_i) \tag{23}$$

$$3 \le \sum_{i\in P}\sum_{j\in P} l_{ij}o_i \le 6 \tag{24}$$

$$3 \le \sum_{i\in P}\sum_{j\in P} e_{ij}o_i \le 6 \tag{25}$$

$$2 \le \sum_{i\in P}\sum_{j\in P} l_{ij}(1-o_i) \le 4 \tag{26}$$

$$2 \le \sum_{i\in P}\sum_{j\in P} e_{ij}(1-o_i) \le 4 \tag{27}$$

and inherit other necessary constraints from task 2.

- Instead of going to as many places as possible, we Set lower bounds on (24) to (27) of each day's attraction to make sure that we visit at least half of the recommended attractions.
- The upper bounds on (24) to (27) are to guarantee that our tourist does not go to too many places in a single day.
  The situation of going all places in one day and and be idle in another day is not desirable.
  We introduce a new variable l_{ij} and this variable serves the same function as the e_{ij} variable in task 2, except that this time it represents transport between two locations on Day 2.
- The lower bound and upper bound on (24) to (27) are adjustable parameters, which could be set by Tourist' personal preference.

4.3 Estimation of parameters

According to Task 3, we need to develop three plans with different budgets.

- Lower bound = Optimal value of cost in task 2 multiplied by two

  As the cost of travel plan in a single day is minimised in task 2, it is reasonable to infer that we need to spend twice the minimum daily cost on a two day travel.

- Upper bound = Cost of feasible solution obtain in equations (23) to (27).

  Without any constraint on the budget, we let the programme spend as much as it needs and then calculate the cost of this solution

- Middle Cost = Average of upper bound and lower bound.

Below is an example of the budget constraint to compute the low budget plan. "Tall" is the set of time interval for two days (from 0 to 97)

$$\sum_{i \in P} \sum_{j \in P: j \neq i} (e_{ij} + l_{ij}) u_{ij} + \sum_{i \in P} \sum_{t \in Tall} x_{it} c_i \leq budget \tag{28}$$

4.4 Outline of Algorithm

Algorithm 1 Task 3 algorithm

R: a set of all possible paths
F: a set for feasible paths, empty at first
H: a set of all hotels
s: best feasible solution so far
procedure FINDDIVERSEPATH(H, R)
    for each item p in R **do**
        $X_{h_0} \leftarrow 1$
        $X_{h_4 9} \leftarrow 1$
        Find all feasible paths given initial states and store them in F
        for each path p in F **do**
            if p is more diverse than s **then**
                if p and s are equally diverse **then**
                    Pick the one with more attractions visited
                $s \leftarrow p$
    return s

4.5 Interpretation of Output

After running the model through the algorithm, we were able to get the following solutions for low budget, medium budget and high budget. Since the models only took around a maximum of 5 minutes to run, we were able to iterate through the algorithm easily.

Time	Location	Type of Location	Region
9:00 a.m	Crowne Plaza Hotel @Changi	Hotel	East
10:15 a.m	East Coast Park	Outdoor Attraction	East
12:00 p.m	Fort Canning Art Centre	Indoor Attraction	Central
12:45 p.m	ArtScience Museum	Indoor Attraction	Central
1:30 p.m	Odette	Dining	Central
3:00 p.m	Chinese Garden	Outdoor Attraction	West
3:45 p.m	The Rink	Indoor Attraction	West
5:00 p.m	Gardens by the bay	Indoor Attraction	Central
6:15 p.m	Punggol Nasi lemak	Dining	North-East
7:45 p.m	Punggol Waterway Park	Outdoor Attraction	North-East
9:00 p.m	Crowne Plaza Hotel @Changi	Hotel	East

Figure 5: Low Budget Day 1

Time	Location	Type of Location	Region
9:00 a.m	Crowne Plaza Hotel @Changi	Hotel	East
9:15 a.m	Jewel Changi Airport	Indoor Attraction	East
11:45 a.m	Sungei Buloh Wetland Reserve	Outdoor Attraction	North
1:15 p.m	Science Centre Singapore	Indoor Attraction	West
1:45 p.m	Miam Miam French and Japanese	Dining	West
4:45 p.m	Simpang Bedok	Dining	East
6:30 p.m	Botanic Gardens	Outdoor Attraction	Central
7:15 p.m	Haw Par Villa	Outdoor Attraction	Central
9:00 p.m	Crowne Plaza Hotel @Changi	Hotel	East

Figure 6: Low Budget Day 2

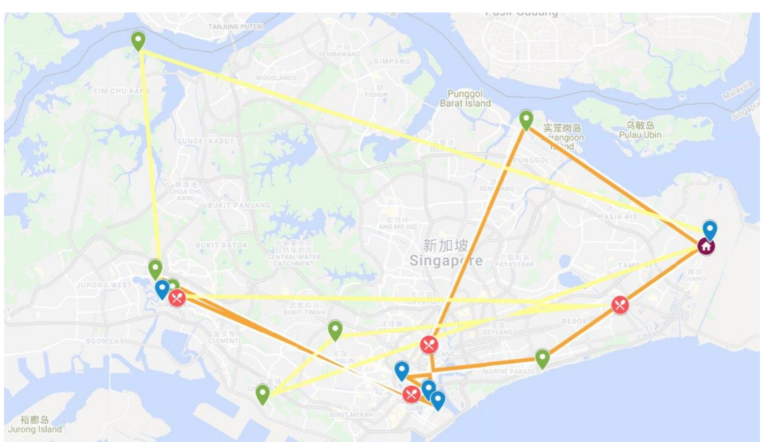

Figure 7: Low Budget Visualisation

Time	Location	Type of Location	Region
9:00 a.m	Hilton Garden Inn Singapore Serangoon	Hotel	North-East
9:30 a.m	Botanic Gardens	Outdoor Attraction	Central
10:45 a.m	River Safari	Indoor Attraction	North
11:15 a.m	Singapore Zoo	Outdoor Attraction	North
12:45 p.m	Orchid Live Seafood	Dining	North
3:45 p.m	The Rink	Indoor Attraction	West
5:15 p.m	Simpang Bedok	Dining	East
7:30 p.m	Pulau Ubin	Outdoor Attraction	East
9:00 p.m	Hilton Garden Inn Singapore Serangoon	Hotel	North-East

Figure 8: Medium Budget Day 1

Time	Location	Type of Location	Region
9:00 a.m	Hilton Garden Inn Singapore Serangoon	Hotel	North-East
9:30 a.m	Fort Canning Art Centre	Indoor Attraction	Central
11:00 a.m	Singapore Discovery Centre	Indoor Attraction	West
12:00 p.m	Miam Miam French and Japanese	Dining	West
1:00 p.m	Science Centre Singapore	Indoor Attraction	West
2:30 p.m	Sungei Buloh Wetland Reserve	Outdoor Attraction	North
4:30 p.m	Haw Par Villa	Outdoor Attraction	Central
5:00 p.m	Timbre+	Dining	Central
6:30 p.m	Gardens by the Bay	Indoor Attraction	Central
7:45 p.m	Punggol Waterway Park	Outdoor Attraction	North-East
9:00 p.m	Hilton Garden Inn Singapore Serangoon	Hotel	North-East

Figure 9: Medium Budget Day 2

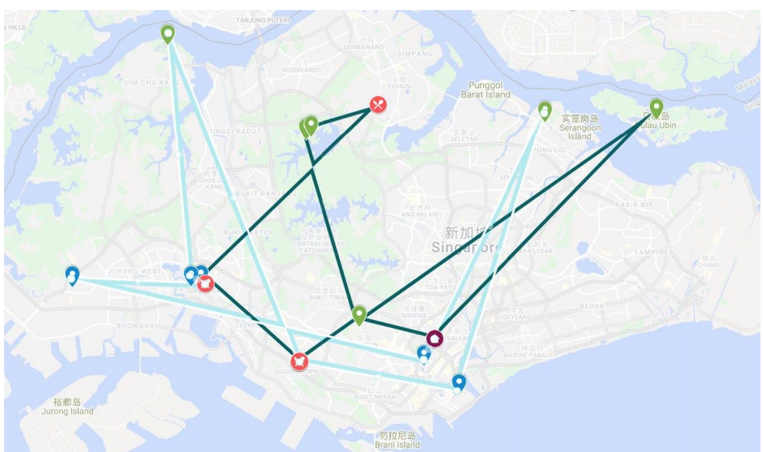

Figure 10: Medium Budget Visualisation

Time	Location	Type of Location	Region
9:00 a.m	Crowne Plaza Hotel @Changi	Hotel	East
10:15 a.m	Fort Canning Art Centre	Indoor Attraction	Central
11:30 a.m	Miam Miam French and Japanese	Dining	West
12:45 p.m	Chinese Garden	Outdoor Attraction	West
1:45 p.m	Singapore Discovery Centre	Indoor Attraction	West
3:30 p.m	Art Science Museum	Indoor Attraction	Central
4:15 p.m	1-Altitude	Indoor Attraction	Central
5:30 p.m	Chomp Chomp Food Centre	Dining	North-East
7:15 p.m	Coney Island	Outdoor Attraction	North-East
8:00 p.m	Punggol Waterway Park	Outdoor Attraction	North-East
9:00 p.m	Crowne Plaza Hotel @Changi	Hotel	East

Figure 11: High Budget Day 1

Time	Location	Type of Location	Region
9:00 a.m	Crowne Plaza Hotel @Changi	Hotel	East
9:45 a.m	Jewel Changi Airport	Indoor Attraction	East
10:45 a.m	Wild Wild Wet	Outdoor Attraction	East
12:00 p.m	Pulau Ubin	Outdoor Attraction	East
1:45 p.m	Jamie's Italian	Dining	Central
4:45 p.m	Odette	Dining	Central
6:00 p.m	Gardens by the bay	Indoor Attraction	Central
7:15 p.m	Haw Par Villa	Outdoor Attraction	Central
9:00 p.m	Crowne Plaza Hotel @Changi	Hotel	East

Figure 12: High Budget Day 2

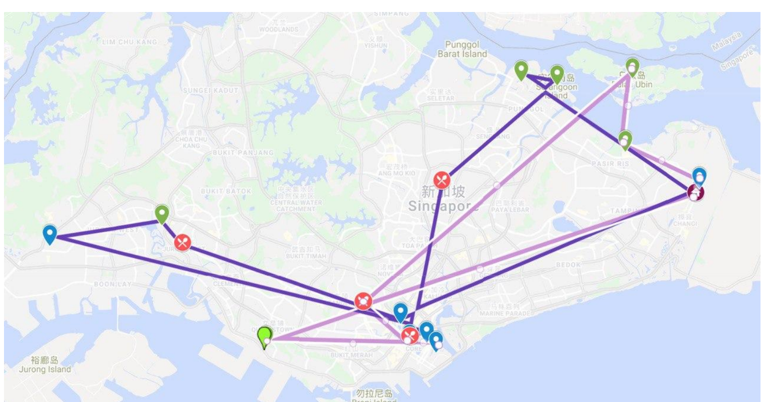

Figure 13: High Budget Visualisation

5 Reflections/Limitations of our Solution

For limitations, we have four limitations in this problem.

First, when collecting data, we gave the general rule of thumb that for those locations that are 5-10 min walking distance apart, such as, Singapore Zoo, Night Safari, River Safari which are clustered together, we set the transportation cost and travelling time be both zero. In reality, it's fine that the transportation cost is zero, however, the travelling time shouldn't be ignored.
Second, we only considered public transportation which includes MRT and buses and we didn't consider about taxis or other transportation tools, which made our travelling cost low in general. This low cost was insignificant compared to the cost of attractions and thus transportation cost did not play a big role in our solutions. Furthermore, the discrete time interval of 15 minutes which is not that precise may also give a loose solution.

Third, according to our logical assumption referred in the Problem Statement part, we concluded that we do not need to change hotel since it does not help with diversity, and changing hotels in day 1 may cause more complicated situations. Although the assumption is reasonable, we don't have any real data to support our assumption.

Last but not least, our constraints are not that flexible. According to our assumption mentioned in the third limitation, the parameters and constraints need to be changed when we are running for different hotels and have to introduce and algorithm to test for the different hotels.

6 Appendix

The detailed descriptions of constraints are as follows:
(1) describes the sum of cost due to attraction and cost due to transportation.
(2) describes that the time spent on all the places cannot exceed total timespan.
(3) describes that every restaurant and attraction can only be visited at most once.
(4) describes that the hotel should be visited exactly twice.
(5) describes the first time the hotel is visited at the beginning.
(6) describes exactly one restaurant which is visited for lunch time.
(7) describes exactly one restaurant which is visited for dinner time.
(8) describes that no outdoor attraction should be visited during the forbidden period.
(9) describes that at least eight attractions are visited.
(10) describes that at least three of the attractions are outdoors.
(11) describes that at least two other regions should be visited.
(12) and (13) describes the indicator variable of whether West region is visited.
(14) and (15) describes the indicator variable of whether East region is visited.
(16) and (17) describes the indicator variable of whether South region is visited.
(18) and (19) describes indicator variable of whether North region is visited.

(20) describes the upper bound of big M, which is the total number of places.
(21) describes if place j is visited after i, then the duration between start time of i and j should be greater or equal to the process time and transportation time .
(22) describes the upper bound of big N, which is the timespan.
(23) describes that if there is a transportation between i and j, then both i and j must be visited.
(24) describes that if there is a transportation from i and j, then i must be visited.
(25) describes that every place is either inside the cycle or not visited at all.
(5),(25),and (23) together ensure that the hotel is the start point and end point of the cycle.
(26) describes that restaurants are only allowed to be visited during lunch time or dinner time.
(27)-(30) together describes the domain of our decision variables.