For office use only	Team Control Number	For office use only
T1 ________________	**1922402**	F1 ________________
T2 ________________		F2 ________________
T3 ________________	Problem Chosen	F3 ________________
T4 ________________	**C**	F4 ________________

2019 Design for Taxi Driver Making Decisions at Airports

Abstract

Airport taxi scheduling is a classic and practical problem. Reasonable decision-making scheme is conducive to maximize the driver's profit, and also help the airport to allocate driving resources efficiently. In this paper, the decision-making mechanism of whether the driver leaving or staying at the airport is designed, and the scheme of airport's passenger-taxi boarding is planned.

For questions 1 and 2, considering the driver aims to make more profits, we take the taxi's net income in a certain time interval as the decision standard. we take the particular time period from the taxi entering the "storage pool" to the completion of the passenger delivery. In this way, various factors can be comprehensively analyzed. By cluster analysis of timestamp, taxi density, flight density, length of stay, we can predict the queueing time of taxi at a certain timestamp. Factors can be quantified by airport data, approximation, probability statistics and other methods. After substituting the model into the one-day data of Pudong airport, we come to the conclusion that the net income of taxi waiting in line during 22:00-24:00 and rush hour will be higher than that of returning to the city, and the decision-making model has the greatest dependence on flight density.

For the third problem, the setting of boarding points should consider the number and the interval of boarding points. Therefore, we set up a dynamic solution, which uses genetic algorithm. And in order to solve the problem of taxi and passenger arrangement, A queuing service model about how many taxis correspond to one boarding point and how many passengers correspond to one boarding point is setting up. After substitute the data into the model we get the setting schemes: before six o'clock in the evening, the pedestrian volume is small, three boarding points are set on one side and two parking points are set on the other side, the time interval of each parking point is 22m, three taxis are arranged for each batch to go to one boarding point. After 6 p.m., there is a large pedestrian volume. So there will be 28 boarding points on both sides of the road. Each boarding point is 122 meters apart, and for each batch two taxis are arranged to go to same boarding point.

For the fourth question, our priority scheme is to add a short distance Lane in the queuing area for the latest taxis after delivering short distance passengers. As to whether the scheme can make the taxis' revenue more balanced, we build a model to verify. In the verification, we find that when there is no priority, the short-distance vehicle revenue is indeed inferior to that of the long-distance vehicle. And the average revenue of the short-distance vehicle can be improved with our solution.

Contents

1 Restatement of the problems

1.1 Background

Taxi is an important way for passengers to go to the destination after getting off the plane. Passengers who want to take a taxi will go to the loading area for queuing, and then take a taxi. The airport's managers will allocate the number of passengers and taxis allowed. Domestic taxis often need to make decisions after delivering guests to the airport:

(A) go to the "car storage pool" and wait in line to take passengers and then return to the city. There is a time cost of queuing.

(B) return to the urban area without load, save the cost of queuing time, but pay the cost of no-load, and lose the potential passenger source of the airport.

For how to make decisions, drivers often judge empirically by the number of flights and passengers arriving in a certain period of time. In the real scene, there are many determinate and non determinate factors that affect the choice of drivers.

1.2 Tasks

In this paper, the mathematical model is established and the following problems are solved:

(1) find out the factors that affect the driver's judgment, including taxi revenue, number of airport passengers, etc. This paper analyzes the influence mechanism of various factors on taxi drivers, establishes the decision-making model of taxi drivers, and designs the selection strategy.

(2) collect the data of a domestic airport and related taxis, give the taxi driver's decision-making scheme in this situation, and analyze the rationality of the decision-making model built in the previous question, as well as the dependence on various factors.

(3) in the loading area, it is usually the taxi queuing to carry passengers or the passenger queuing for taxis. Now, there are two parallel lanes in the loading area. It is necessary to design the scheduling scheme of the taxi and the passenger and set the loading point so as to make the loading efficiency of the loading area the highest.

(4) there are short-distance passengers and long-distance passengers in the airport. Taxi drivers can't refuse to take them, but they can take them back and forth. In order to balance the income of all taxis, it is necessary to design a priority scheme for short-distance vehicles and verify the scheme works.

2 Model Analysis

2.1 Analysis of task 1

First, we have to determine the driver's decision criteria. Here it's the net income of the queuing car and the returning car in the time interval of the queuing taxi from start queuing to finish passengers delivering , so that important factors such as queuing time, airport passenger source, no-load cost can be taken into account. Here, the net income is calculated by subtracting the total fuel cost from the total fare income. For queuing vehicles, the net income is related to mileage, single ride fare and oil price. For return vehicles, the net income is also related to the saved queuing time, vehicle speed, probability of carrying urban passenger, weather, road conditions and other factors. Vehicle speed and mileage can be obtained by approximation; pricing rules and oil prices can be inquired; urban passenger carrying probability can be calculated by statistics; weather influence are related to visibility; road conditions can be estimated by the traffic flow obeying Poisson distribution; queuing time cannot be obtained directly, which is affected by timestamps, number of passengers and number of taxis, while drivers cannot determine the number of passengers directly, and what can be determined is the number of flights arrived in this period. Therefore, cluster analysis is carried out for (time point, flight density, taxi density, vehicle dwell time) to predict the taxi queuing time in the new situation.

2.2 Analysis of task 2

First of all, we need to determine an airport, take the relevant samples from the taxi and flight data, and set some parameters. Then bring the samples (time point, flight density, taxi density, vehicle dwell time) into the cluster model, adjust the cluster number k, label variable distance weight and other relevant parameters, find the best cluster number, and then bring the new samples without queuing time into the decision-making model to obtain the corresponding decision. This paper gives the selection scheme of taxi drivers in different situations.

As for the decision model's dependence on various factors. We combine the decision-making model with time period, flight density and taxi density. Observe and analyze their relations by drawing statistics.

Finally, tested by the actual situation, such as public transportation, urban roads and some official statistical data, the decision-making model is reasonable. Then, the sensitivity matrix is obtained by using sensitivity analysis on the decision-making model. And we can obtain the variables with the the model's highest dependence.

2.3 Analysis of task 3

For this question, we play the role of management department, need to set up boarding points, and arrange taxis and passengers to the corresponding boarding points, the goal is to make the total boarding efficiency the highest. There are only two lanes. In order to increase efficiency, boarding points are set on both sides of the road, and the number of passengers at each side of the boarding point should be as average as possible. It is also necessary to allocate the same number of taxis to the boarding point, as well as to find out how many taxis are dispatched to the same boarding point at a time.

In order to make the total boarding efficiency the highest, that is, the total waiting time of passengers is the least. The waiting time of passengers is divided into two parts, one is the time when the passengers arrive at the corresponding boarding point, the other is the time when the passengers wait in line at the boarding point until they become the first person of the line. When the passenger flow is large, and a certain flight arrives at the airport, the passengers of the previous flight are possibly still waiting for the taxis, so it is unnecessary to calculate when they arrive at the corresponding boarding point, because they are already in the waiting line. Therefore, the model needs to be classified into two types, which are respectively used in the scenarios with large and small passenger flow. Moreover, when the passenger flow is large, in order to meet the needs of passengers, the number of boarding points will inevitably increase, but not as much as possible, because there will be "ghost traffic jam" phenomenon affecting the average speed of taxis, so we use the logarithmic model of speed and traffic density to simulate this phenomenon when the traffic flow is large, and the setting of boarding points needs to consider the interval to ensure safety and convenience for passengers. Finally, we establish the relationship between the number of boarding points, the interval, the scheduling scheme and the total waiting time of passengers. We use genetic algorithm to find the number of boarding points, the interval and the scheduling scheme with the minimum waiting time of passengers, and use the data of Pudong Airport as an example to analyze.

2.4 Analysis of task 4

The purpose of establishing the short-distance vehicle priority scheme is to reduce the queuing time of short-distance vehicles. The effective method is to set up the short-distance vehicle priority channel, other vehicles can't enter the queue until the short-distance vehicle reaches the passengers.

The equilibrium of revenue is reflected in the variance of revenue of all taxis. Therefore, the average revenue model of taxis is established. It is assumed that the short-distance vehicle will return to the airport every time until it reaches the long-distance passenger, and the long-distance vehicle will not return to the airport after completing the passenger journey. The appropriate time period for the short-distance vehicle is from start queuing to finish delivering the long-distance passenger, and the appropriate time interval for the long-distance vehicle is from start queuing to completion of passenger delivering. According to this, we can calculate the hourly income of short-distance vehicles and long-distance vehicles in their respective time intervals. Finally, we can get the variance of hourly income of all taxis by considering the proportion of long-distance vehicles and short-distance vehicles. If the variance is reduced than before, it is proved that the scheme is feasible and effective.

3 Model Assumptions

- **Population**
  In general, a large population means high potential of disease's spread. People in a populous area have a greater frequency of contacting the others than those in a sparsely populated area. According to the route of transmission, it is obvious that the probability of getting infected would be larger.

- **Traffic**
  Convinient traffic encourages population mobility, which contributes to the spread of EVD. However, It also encourages freightage, including medicine.

- **Medical Level**
  A society will be less affected by EVD if proper measures are taken efficiently and promptly. These measures include segregating patients and strengthening the sanitary control of public places. The manufacture of drugs and vaccines aiming at EVD is also an important part.

- **Regional Custom**
  Funeral is considered solemn in the African culture. The dead should be cleaned, kissed and touched before buried. This kind of culture facilitates EVD infections.

- **Other Social Factors**
  The spread of diseases is also influenced by factors like social development, health situation, individual's living condition, etc. These factors are not considered to simplify our models.

4 Parameter Table

Symbols that newly appear in this section are listed in table 1. Some of the symbols in table **??** are also used in this section and are not listed here.

Symbol	Defination
$\mathcal{N}_i$	total volume of population in city i
$\mathcal{S}_i$	number of susceptible people in city i
$\mathcal{I}_i$	number of infected but not segregated people in city i
$\mathcal{G}_i$	number of segregated and also infected people in city i
$\mathcal{D}_i$	number of death caused by EVD in city i
$\mathcal{R}_i$	number of recovered people (also immunized) in city i
P_i	Total number of people in city i apart from the people who are dead and segregated
$t_{i,j}$	number of people transmit from city i to city j indeed
α	transmission coefficient
$d_{i,j}$	distance from city i to city j
$vacc_{tot}$	the number of shares of vaccine that can be provided to all the cities every day
$vacc_i$	the number of shares of vaccine that can be provided to city i every day

Table 1: The definition of the symbols.

5 Problem 1: Design for Decision Model

Taxi drivers always aim to make more profits, therefore, we use the net income of taxi drivers in a certain time interval as the standard of decision-making. Assuming that both taxis D_a and D_b arrive at the airport at time A, driver D_a decides to enter the "car storage pool" to queue for passengers and complete the delivery at time B, then time A to time B is the time period considered by our decision-making model, because this time period takes into account the main factors such as queuing time, no-load time and so on. In this period, the net income W_1 of D_a is the total income I_1 minus the fuel consumption O_1 when delivering passengers, i.e:

$$W_1 = I_1 - O_1 \tag{1}$$

Set $t(Y - X)$ as the time spent from X to Y, calculated in hours. If D_b decides to return to the urban area to carry passengers, then during the period of $T(B - A)$, D_b's net income W_2 is the total income I_2 minus the total fuel cost O_2 consumed in delivering passengers, and then minus the fuel cost O_e consumed in no-load time:

$$W_2 = I_2 - O_2 - O_e \tag{2}$$

If $W_1 > W_2$, it is a better choice for drivers who arrive at the airport at time A to enter the "car storage pool" to carry passengers. On the contrary, if $W_1 < W_2$, it is more cost-effective to return to the city without load. If the revenue is the same, both options are available.

5.1 Net Income for Queueing Taxis

Further analysis, for D_a: specifically speaking, I_1 is the fare paid by passengers. I_1 is related to the mileage L_1 for delivering passengers. In the real scene, most passengers from the airport will go to the city, which is also mentioned in the problem statement. Therefore, we set the distance from the airport to the city center as L_1. I_1 is also related to the charging standard of taxis. Generally speaking, L_1 exceeds the gradient charging mileage a and b of taxis, so there are:

$$I_1 = f_1 + f_2 \cdot (b - a) + f_3 \cdot (L_1 - b) \tag{3}$$

where $f_1, f_2$and f_3 are the gradient pricing fees for three sections of taxi mileage respectively. O_1 is L_1 multiplied by fuel consumption per kilometer o multiplied by oil price per volume k. Namely:

$$O_1 = L_1 \cdot o \cdot k \tag{4}$$

Therefore, for driver D_a, the net income can be specified as:

$$W_1 = f_1 + f_2 \cdot (b - a) + f_3 \cdot (L_1 - b) - L_1 \cdot o \cdot k \tag{5}$$

The time period $t(B - A)$ can be divided into "storage pool" queuing time t_q and delivery time t_s. t_s can be estimated by dividing average speed per mileage. The average speed v_1 here is approximated by the airport high speed limit minus $20km/h$. In real life, the speed will also be affected by the weather conditions, and the bad weather will reduce the safe driving speed. The impact factor is set as $\alpha(0 < \alpha \leq 1)$, and α is 1 when the weather is good. In other weather conditions, α is related to visibility. The lower the visibility is, the smaller α is and the slower the speed is. In addition to the weather, the traffic conditions in this period will also affect the driving speed, so that the influencing factor is $\beta, (0 < \beta \leq 1)$. Set the number of vehicles within the mileage L_1 section at time t as M, when $M \leq M_0$, the driving is smooth, β is 1. M_0 can be approximated by the average daily traffic flow of this section, M obeys the Poisson distribution $P(\lambda)$, so that:

$$\beta = \begin{cases} 1, if M \leq M_0 \\ P(M > M_0), if M > M_0 \end{cases} \tag{6}$$

And we can conclude that:

$$t_s = \frac{L_1}{\alpha \cdot \beta \cdot V_1} \tag{7}$$

But calculation of t_q is more complex, which is related to the number of passengers in this period and the vehicle density N_c in this period. Here, the number of passengers can be estimated by the flight density N_a. In order to better predict the t_q of a certain period, we divide 0:00-24:00 into 12 periods, each of which has a corresponding period label T. Using K prototype clustering[1], a set of sample points (T, N_a, N_c, t_q) is established, where T is the time period label. Therefore, we can find the corresponding classification for the specific (T, N_a, N_c), and then get t_q. The clustering steps are as follows:

$X = \{X_1, X_2, \ldots, X_n\}$ represents n such sample point sets. For each point set, there are four attributes (T, N_a, N_c, t_q) corresponding to $X_i = \{x_{iT}, x_{ia}, x_{ic}, x_{iq}\}$. For numerical attributes N_a, N_c, t_q, firstly normalization processing is carried out:

$${x_{ik}}' = \frac{x_{ik} - \bar{x_k}}{s_k} \tag{8}$$

where:

$$\bar{x_k} = \frac{1}{n} \cdot \sum_{i=1}^{n} x_{ik} \tag{9}$$

$$s_k = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (x_{ik} - \bar{x_k})^2} \tag{10}$$

we use Euclidean distance to define the different distance of numerical attribute:

$$D(X_i, X_j) = \sqrt{(x_{ia} - x_{ja})^2 + (x_{ic} - x_{jc})^2 + (x_{iq} - x_{jq})^2} \tag{11}$$

For classification attribute T, Hamming distance is used:

$$d(x_{ik}, x_{jk}) = \begin{cases} 1, if x_{ik} \neq x_{jk} \\ 0, if x_{ik} = x_{jk} \end{cases} \tag{12}$$

So the distance between the sample point set X_i and the cluster Q_l is:

$$D_s(X_i, Q_l) = D(X_i, X_l) + \mu \cdot d(x_{iT}, x_{lT}) \tag{13}$$

where X_l is the prototype (center) of the cluster Q_l. Then the total loss function of this classification is:

$$E = \sum_{l=1}^{k} \sum_{i=1}^{n} y_{il} D_s(X_i, X_l) \tag{14}$$

5.2 Net Income for Returning Taxis

For driver D_b, he takes time t_s to return to the urban area. Within the time of driver D_a's queuing time t_q, driver D_b can receive passengers many times, set it to be N_y times. The total income I_2 of driver D_b is the fare paid by a passenger per time (name it i_2) times N_y (assume the income of each passenger delivering in the urban area is the same). i_2 is related to the distance L_2. In reality, most of the passengers in the urban area are of short distance trip. Here, we use the data we collected to count the average mileage of each passenger delivering downtown as L_2, and then substitute it into the taxi's gradient charging standard to get I_2:

$$i_2 = \begin{cases} f_1, if L_2 \leq a \\ f_1 + (L_2 - a) \cdot f_2, if a < L_2 \leq b \\ f_1 + f_2 \cdot (b - a) + f_3 \cdot (L_2 - b), if L_2 > b \end{cases} \tag{15}$$

The calculation of O_2 is similar to that of O_1:

$$O_2 = L_2 \cdot o \cdot k \cdot N_y \tag{16}$$

The oil cost of O_e is divided into two parts, one is the no-load oil cost O_{e1} from the airport to the urban area, the other is the oil cost O_{e2} waiting for passengers in the urban area for N_y times. And the distance from the airport to the city is L_1, so we have:

$$O_{e1} = L_1 \cdot o \cdot k \tag{17}$$

On the other hand, we can use the collected data to count the average time taken by taxis to pick up a passenger in the urban area, at different time labels, and then take the urban speed limit minus $10km/h$ as the average speed of the taxis in the urban area, viz:

$$O_{e2} = t_w \cdot V_2 \cdot o \cdot k \cdot N_y \tag{18}$$

If the time of delivering passengers in urban area is t_p, then:

$$t_p = \frac{L_2}{\alpha \cdot \beta \cdot V_2} \tag{19}$$

In addition to the impact of weather on taxi speed, it will also affect t_w. For example, heavy rain or hot weather will make urban residents more inclined to take a taxi, and t_w will decrease, while in typhoon weather, people are not inclined to take a taxi, and t_w will increase. Set the influencing factor as $\gamma, (\gamma > 0)$. From this, we can get the approximate value of N_y:

$$N_y = \frac{t_q}{\gamma \cdot t_w + t_p} \tag{20}$$

Therefore, for driver D_b, the net income can be specified as:

$$W_2 = i_2 \cdot N_y - L_2 \cdot o \cdot k \cdot N_y - (L_1 \cdot o \cdot k + t_w \cdot V_2 \cdot o \cdot k \cdot N_y) \tag{21}$$

6 Problem 2: Verification for the Decision Model

1. **Our model is simple and easy to understand**
   Our model is the simplest model we can conceive to reflect the impact of concerned independent variables (factors regarding medication) and to solve the problem lifted by the question.

   Our single-city model is based on the most elegant model in the field of epidemiology - the SIR model, and we reconstruct the model (mainly add two clusters of people) in order to introduce concerned independent variable into our system.

   Our multi-city model is based on our single-city model and introduce only one 'people flow' to obtain the geographic characteristic of the spread of disease.

2. **Our model is effective and in good agreement with the reality**
   Simple as they are, they are effective in reflecting the complex relationships between numerous variables and parameters, and they not only reveal the intrinsic characteristics of the spread of disease itself but also successfully link factors of medication to the spread of disease.

   Comparing with the data we have find from several resources, the results of our model not only correspond the general trend of the records but also resemble the reality in some critical features.

3. **Good extensibility**
   Flow of people is a critical factor determining the spread of disease. Although our multi-city model only set the volume of people flow as a function of mere geographic distribution and population of cities, the determinants of people flow can be adjusted when other possible factors are considered. Then, the adjusted model can be applied to study the impact of other possible factors relating to epidemiology.

1. **Our model is just a rough model**
   For simplicity, we have neglected many potential parameters, variables or processes, and have made numerous assumptions. Eg. we did not consider the relationship between separate individuals and we did not dig deeper into the properties of social network which is a quite essential part determining the spread of disease. Some important general or specific factors are also neglected by us, a interesting example of which is a folk custom prevalent in the studied region that relatives kiss the death, which plays a significant role in the spread of disease and is categorized into *Super Spread Event*(SSE) academically.

2. **Our model is only a continuous model**
   Numbers of people, number of shares of drug/vaccine, etc. are important quantities in all the process of modeling and computation. For simplicity, we regard the numbers as directly real numbers instead of integers. It is justifiable when the numbers are large, since the decimal part of the number is negligible; when the system scales down, however, the statistics dose not work and the outcome deviates a lot from reality.

7 Problem 3: Design for Schedule Model

1. **Our model is simple and easy to understand**
   Our model is the simplest model we can conceive to reflect the impact of concerned independent variables (factors regarding medication) and to solve the problem lifted by the question.

Our single-city model is based on the most elegant model in the field of epidemiology - the SIR model, and we reconstruct the model (mainly add two clusters of people) in order to introduce concerned independent variable into our system.

Our multi-city model is based on our single-city model and introduce only one 'people flow' to obtain the geographic characteristic of the spread of disease.

2. **Our model is effective and in good agreement with the reality**
Simple as they are, they are effective in reflecting the complex relationships between numerous variables and parameters, and they not only reveal the intrinsic characteristics of the spread of disease itself but also successfully link factors of medication to the spread of disease.

   Comparing with the data we have find from several resources, the results of our model not only correspond the general trend of the records but also resemble the reality in some critical features.

3. **Good extensibility**
Flow of people is a critical factor determining the spread of disease. Although our multi-city model only set the volume of people flow as a function of mere geographic distribution and population of cities, the determinants of people flow can be adjusted when other possible factors are considered. Then, the adjusted model can be applied to study the impact of other possible factors relating to epidemiology.

1. **Our model is just a rough model**
For simplicity, we have neglected many potential parameters, variables or processes, and have made numerous assumptions. Eg. we did not consider the relationship between separate individuals and we did not dig deeper into the properties of social network which is a quite essential part determining the spread of disease. Some important general or specific factors are also neglected by us, a interesting example of which is a folk custom prevalent in the studied region that relatives kiss the death, which plays a significant role in the spread of disease and is categorized into *Super Spread Event*(SSE) academically.

2. **Our model is only a continuous model**
Numbers of people, number of shares of drug/vaccine, etc. are important quantities in all the process of modeling and computation. For simplicity, we regard the numbers as directly real numbers instead of integers. It is justifiable when the numbers are large, since the decimal part of the number is negligible; when the system scales down, however, the statistics dose not work and the outcome deviates a lot from reality.

8 Problem 4: Design for Priority Model

8.1 Construction of Priority Model

Generally speaking, the average income of long-distance taxis is higher than that of short-distance taxis. Therefore, the following priority schemes are formulated for short-distance taxis:

A short Lane will be added in the queuing area for taxis with short-distance passengers delivering last time. The other is the general queuing lane. The vehicles in the short Lane have priority to enter the loading area for reaching passengers, and the vehicles in the general lane can only enter the loading area for boarding passengers after all the vehicles in the short Lane have been arranged. Every time the taxi finishes loading passengers from the airport, there will be mileage record. When the airport staff determines that the mileage record is a short distance, the taxi can enter the short distance lane for priority.

The following analysis shows the change of taxi revenue after the implementation of the plan:

Assume the short-distance taxis returns to the airport every time after delivering short-distance passengers. They start to line up at time D, board at time E, finish delivering at time F, return to the airport at time G. The short-distance taxis continuously delivering the short-distance passengers i times in a cycle until get and deliver long-distance passenger, then the total revenue of the short-distance taxis is:

$$W_{st} = I_s \cdot i + I_l, i = 1, 2, 3, .., N_c \tag{22}$$

where I_s and I_l are the short-distance and long-distance taxi fares respectively, which are determined by the taxi charging standard and the mileage L_s and L_l, and the net income is:

$$W_s = W_{st} - 2 \cdot o \cdot i + O_l \tag{23}$$

O_s and O_l are respectively the fuel charge for short-distance and long-distance mileage:

$$O_s = o \cdot k + L_s \tag{24}$$

$$O_l = o \cdot k + L_l \tag{25}$$

The probability that the taxi will carry the short distance passengers for $i - 1$ times after the first time is:

$$P(x = i) = (1 - P_l)^{i-1} \cdot P_l \tag{26}$$

Obviously, the random variable i obeys the geometric distribution. Its expectation is:

$$E(X) = \frac{1}{P_l} \tag{27}$$

The total cycle time is the time of i times queues, plus the time of i round trips, plus the time of the last long-distance delivery:

$$t_{s2} = t_{DE} \cdot i + 2 \cdot i \cdot t_{EF} + t_l \tag{28}$$

t_{EF} and t_L are also related to mileage L_s and L_L, respectively:

$$t_{EF} = \frac{L_s}{\alpha \cdot \beta \cdot V_1} \tag{29}$$

$$t_l = \frac{L_l}{\alpha \cdot \beta \cdot V_1} \tag{30}$$

t_{DE} is obtained by clustering analysis of flight density and taxi density in this time period. So the average hourly revenue of short distance vehicles in the cycle becomes:

$$\bar{W_s} = \frac{I_s \cdot E(X) + I_l - 2 \cdot O_s \cdot E(X) - O_l}{t_{DE} \cdot E(X) + 2 \cdot E(X) \cdot t_{EF} + t_l} \tag{31}$$

It is set that the long-distance taxis will not return to the airport after delivering passengers. For the long-distance bus, one cycle is queuing, carrying and finishing delivering. Assume the long-distance taxis start to queue at time D, carry passengers at time E (the average queuing time of all the taxis is the same), and carry passengers at time H, then the average hourly income $\bar{W_l}$ of the long-distance taxi meets the requirements:

$$\bar{W_l} = \frac{I_l - O_l}{t_{DE} + t_{EH}} \tag{32}$$

where

$$O_l = o \cdot k + L_l \tag{33}$$

$$t_{EH} = t_l \tag{34}$$

Set the total number of taxis is N_C, and the proportion of long-distance vehicles is P_l. then, the average hourly income W_b for every taxi before the implementation of the plan is:

$$W_b = \bar{W_l} \cdot P_l + \bar{W_s} \cdot (1 - P_l) \tag{35}$$

The variance of the average revenue is:

$$S_b = N_c \cdot P_l \cdot (\bar{W_l} - W_b)^2 + N_c \cdot (1 - P_l) \cdot (\bar{W_s} - W_b)^2 \tag{36}$$

After the implementation of the plan, the cycle of the short distance vehicle is still i consecutive short distance cycle plus the last long distance cycle, and its net income is the same as before the implementation of the plan:

$$W_{s2} = W_s \tag{37}$$

After the first queue, the time of each queue is decreased, changes to:

$$t_{DE2} = (1 - P_l) \cdot t_{DE} \tag{38}$$

The total cycle time is the time of the first queue plus the time of the next $i - 1$ queues plus the time of i round trips plus the time of the last long-distance delivery:

$$t_{s2} = t_{DE} + t_{DE2} \cdot (i - 1) + 2 \cdot i \cdot t_{EF} + t_{EH} \tag{39}$$

So the average hourly revenue of short distance vehicles in the cycle becomes:

$$\bar{W_{s2}} = \frac{I_s \cdot E(X) + I_l - 2 \cdot O_s \cdot E(X) - O_l}{t_{DE} + t_{DE2} \cdot (E(X) - 1) + 2 \cdot E(X) \cdot t_{EF} + t_{EH}} \tag{40}$$

For a long-distance taxi, still assume the long-distance taxi does not turn back to the airport, the average revenue of one cycle of the long-distance taxi is:

$$\bar{W_{l2}} = \bar{W_l} \tag{41}$$

The average hourly revenue per taxi becomes:

$$W_a = \bar{W_{l2}} \cdot P_l + \bar{W_{s2}} \cdot (1 - P_l) \tag{42}$$

The variance of revenue becomes:

$$S_a = N_c \cdot P_l \cdot (\bar{W_{l2}} - W_a)^2 + N_c \cdot (1 - P_l) \cdot (\bar{W_{s2}} - W_a)^2 \tag{43}$$

If S_b > S_a, it shows that our priority scheme does make the revenue of all taxis in this time period more balanced.

8.2 Verification of Priority Model

Similarly, Pudong Airport is selected as the actual example. The scenario is the same as the previous example with the same parameter setting.

For the time being, the value of P_l cannot be obtained. Therefore, we have set multiple candidate values for P_l : $0.6, 0.7, 0.8$, (set $P_l > 0.5$, because most airport passengers are long-distance passengers) and we observed multiple experimental results. And from these results, we can draw some conclusions:

```
Average profit for short-distance taxi before adopting our plan 69.34274065712705
Average profit for short-distance taxi after adopting our plan 76.540957400783
Average profit for long-distance taxi 98.92494691567406
Average profit for all taxis before adopting our plan 87.09206441225525
Average profit for all taxis after adopting our plan 89.97135110971763
Square difference before adopting our plan:  210.02566250957221
Square difference after adopting our plan:  120.25031678466065
```

Figure 1: Results when $P_l = 0.6$

```
Average profit for short-distance taxi before adopting our plan 74.41487683227767
Average profit for short-distance taxi after adopting our plan 80.75062409690918
Average profit for long-distance taxi 98.92494691567406
Average profit for all taxis before adopting our plan 91.57192589065514
Average profit for all taxis after adopting our plan 93.4726500700446
Square difference before adopting our plan:  126.15614245353053
Square difference after adopting our plan:  69.36426208334234
```

Figure 2: Results when $P_l = 0.7$

```
Average profit for short-distance taxi before adopting our plan 79.03610508698397
Average profit for short-distance taxi after adopting our plan 83.82435144810836
Average profit for long-distance taxi 98.92494691567406
Average profit for all taxis before adopting our plan 94.94717854993604
Average profit for all taxis after adopting our plan 95.90482782216093
Square difference before adopting our plan:  63.2905646858644
Square difference after adopting our plan:  36.4844773560105
```

Figure 3: Results when $P_l = 0.8$

1. This scheme can significantly reduce the variance of average revenue and make the revenue of all taxis as balanced as possible.

2. It is true that airport long-distance passenger transport is more profitable than short-distance passenger transport, so it is a scientific and reasonable measure to set up a priority scheme for short-distance vehicles.

3. This scheme improves the average revenue of short distance vehicles and all taxis to a certain extent.

4. With the increase of P_l, almost all taxis take long-distance passengers, so the variance of the overall income decreases.

9 Model Appraisal

9.1 The strengths of the model

In this paper, we have constructed our models based on the biological features of EVD, social features of human society and several reasonable assumptions. Our models consist of two parts: one is considering the the spread of disease within a single city and serves as the base of the other; the other takes the people flow among the cities into account, the application of which gives an optimized plan regarding how should we allocate the resources of medication such as vaccine.

9.2 The weaknesses of the model

Both of the models are applied to specific cases separately, and the results of computation which are carefully studied justified our model. Through our analysis of the model, we explored and explained the complex relationship among numerous variables and parameters.

9.3 The improvement of the model

The effectiveness of medical treatment (including segregation, vaccination and pharmacotherapy) is verified by our model and the strategy to allocate vaccine and drug is revealed by our investigation.

9.4 The Extension of the model

10 Conclusion

In this paper, we have constructed our models based on the biological features of EVD, social features of human society and several reasonable assumptions. Our models consist of two parts: one is considering the the spread of disease within a single city and serves as the base of the other; the other takes the people flow among the cities into account, the application of which gives an optimized plan regarding how should we allocate the resources of medication such as vaccine.

Both of the models are applied to specific cases separately, and the results of computation which are carefully studied justified our model. Through our analysis of the model, we explored and explained the complex relationship among numerous variables and parameters.

The effectiveness of medical treatment (including segregation, vaccination and pharmacotherapy) is verified by our model and the strategy to allocate vaccine and drug is revealed by our investigation.

References

[1] Z. Huang, "Extensions to the k-means algorithm for clustering large data sets with categorical values," *Data mining and knowledge discovery*, vol. 2, no. 3, pp. 283–304, 1998.

Appendices

Appendix A First appendix

some text...

Here are simulation programmes we used in our model as follow.

Input matlab source:

```
function [t,seat,aisle]=OI6Sim(n,target,seated)
pab=rand(1,n);
for i=1:n
    if pab(i)<0.4
       aisleTime(i)=0;
    else
        aisleTime(i)=trirnd(3.2,7.1,38.7);
    end
end
```

Appendix B Second appendix

some more text **Input C++ source:**

```cpp
//==========================================================================
// Name        : Sudoku.cpp
// Author      : wzlf11
// Version     : a.0
// Copyright   : Your copyright notice
// Description : Sudoku in C++.
//==========================================================================

#include <iostream>
#include <cstdlib>
#include <ctime>

using namespace std;

int table[9][9];

int main() {

    for(int i = 0; i < 9; i++){
        table[0][i] = i + 1;
    }

    srand((unsigned int)time(NULL));

    shuffle((int *)&table[0], 9);

    while(!put_line(1))
    {
        shuffle((int *)&table[0], 9);
    }

    for(int x = 0; x < 9; x++){
        for(int y = 0; y < 9; y++){
            cout << table[x][y] << " ";
        }

        cout << endl;
    }

    return 0;
}
```