Intelligent Passenger-Flow-Control Scheme at Subway Stations in Smart City

Yuan Zhang¹, Congduan Li¹*, Jinyang Zhang² and Yiqian Zhang¹

¹ School of Electronics and Communication Engineering, Sun Yat-sen University, Guangzhou, China
² School of Data and Computer Science, Sun Yat-sen University, Guangzhou, China
* Corresponding Email: licongd@mail.sysu.edu.cn

Abstract. With the rapid development of urbanization, more and more cities have built the subway systems. However, the super cities like Beijing, Shanghai and Guangzhou, have experienced problems such as overloading and insufficient capacity, especially during peak hours, due to the tremendous increase of passengers. A scientific and efficient optimization scheme is important to alleviate the passenger flow congestion problem and improve the transportation efficiency. This paper investigates the passenger-flow-control problem, aiming to shorten the average waiting time for passengers. An intelligent flow-control scheme is proposed to optimally alleviate the congestion problem in the subway system. Experiments show that the proposed flow-control scheme works well in alleviating the congestion and passengers’ waiting time at subway stations.

Keywords. Subway; passenger flow; path optimization; flow-control strategy.

1. Background
With the rapid development of urbanization, more and more cities have built their subway systems. Such a system usually consists of several railway lines operating underground. The first subway in the world was the Metropolitan Underground in London, UK [1], which was built in 1863. After that, cities like New York (1869) [2], Paris (1900) [3], Berlin (1902) [4], Tokyo (1927) [5] also started to build their subway system. The first subway in China is Line 1 of Beijing Subway, which started in 1971 [6]. Then, Hong Kong and other cities have successively built subway lines. With the development of the city, the construction of subways has sprung up in the early 21st century. Its main purpose is to alleviate ground traffic congestions. In many cities, subway is the main transportation for residents and tourists [7].

However, the rising number of passengers in subways has made some lines overloaded. For instance, the yearly increase rate of passengers in Guangzhou is about 10% in the past decade. The average daily passengers in 2018 has reached 8 million, and even exceeded 10 million in some holidays. [8] The subway system of Guangzhou is shown in figure 1. The queuing time at some stations is even more than 30 minutes in peak hours, which causes much inconvenience for passengers.

One commonly used solution is to reduce the time interval between two trains in peak hours. But it has limited improvement and costs too much, since the number of trains running on a line are limited.

Another economic solution is to control the passenger flow in some stations to reduce the average waiting time and at the meantime, to reduce the risk of accidents to ensure subway safety. For instance, almost all stations of the entire Batong line in Beijing use the passenger-flow-control scheme in peak hours [9].
The main cause of congestion in a subway station is that too many people get aboard in previous stations so that the ability to take passengers at the current station is limited. Due to the different functions near each station, there may be more people queuing at some stations and few people queuing at some other stations. It also depends on what day or what time it is. However, we should treat every passenger equally regardless of what station he/she is in. It is unfair for passengers in the later stations if they cannot get on the train simply due to the train is occupied by the passengers from previous stations. Hence, we would like to propose an optimization model to serve each passenger fairly.

In Ref. [10], it was proposed to analyze and forecast the passenger flow based on experience, which will encourage citizens to use other transportsations, like buses, instead of subway. Bus lines that are parallel to the subway lines should be able to effectively alleviate subway congestion. Ref. [11] proposed three levels of control, including the station control, line control and network control. It was found that station control can reduce the passenger flow, but only qualitative analysis was given. Our later analysis will be also based on the line control, while the quantitative analysis of each station will be provided as well.

Ref. [12] proposed that authorities can improve publicity and improve the efficiency of passenger transport organization. This is also feasible, but it cannot be quantitatively analyzed. So it is beyond our discussion. Refs. [13] and [14] proposed a coordinated flow-control optimization model for unexpected large passenger flows. Based on the analysis of passenger OD dataset (a dataset includes passengers’

**Figure 1.** Guangzhou subway map (2019).
origin and destination) and relevant parameters of various subways, they build a mathematical model with the constraints of passenger flow demand, the capacity of inter-regional transportation, platform safety and so on. Their goal is to minimize the average waiting time and maximize the train carrying capacity. It reduces the problem to a multiple linear regression problem. However, it needs a lot of data support, which brings more difficulties.

In our work, we also need to build an optimization model. Data can be obtained through passenger OD dataset or on-the-spot investigation. And we can list the passenger feature matrix M, which indicates the number of people getting on and off from one station to another on a certain line. What's more, we should know the maximum capacity of the subway. After that, we could calculate the number of people in the queue, the number of people getting on the train, the number of people getting off the train and the number of delayed passengers. What we should optimize is average waiting time for each passenger.

The rest of this paper is organized as follows. The optimization model to minimize the average waiting time is formulated in Section 2, with the constraints and an algorithm to solve it proposed in Section 3. Tests on real data are presented in Section 4 and the validity of the proposed solution is shown. Section 5 concludes the paper.

2. System Model

2.1. Relationship Analysis

Subway line can be simply divided into 2 types, straight lines and loop lines. Most of the lines are straight and some of the lines are loop lines, such as Beijing Line 10. The loop line can be also viewed as straight lines. So, for ease of discussion, we consider the straight lines in the following discussions. Some notations are shown in table 1.

We assume that there are N subway stations on a line, and there are two directions. As shown in figure 2. We can calculate the number of people queuing in each subway station, the number of people getting on the train when it comes. We can also count out the number of people getting on and off each subway station.

For ease of discussions, we list some notations as follows.

When consider the number of people in a train, we believe that a train has capacity (P), and the number of people in a train cannot exceed it. Matrix M includes information on number of people getting on and off. Suppose we can estimate the value of the matrix M from OD dataset, then we know the number of people traveling from any station and to another station. We can calculate the ideal number of people getting on a train and getting off the train in each station through the matrix M. Since we only count one direction at a time, the value in the other direction is 0.

\[
M = \begin{bmatrix}
0 & m_{12} & m_{13} & \cdots & m_{1n} \\
0 & 0 & m_{23} & \cdots & m_{2n} \\
0 & 0 & 0 & \cdots & m_{3n} \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
0 & 0 & 0 & \cdots & 0
\end{bmatrix}
\] (1)
Table 1. Notations.

| Name                                                                 | Symbol |
|----------------------------------------------------------------------|--------|
| Number of Stations in a line                                        | $n$    |
| Passenger Feature Matrix $m_{ab}$ means the number of people getting on in station $a$ and getting off in station $b$ | $M = \begin{bmatrix} m_{11} & m_{12} & \cdots & m_{1n} \\ m_{21} & m_{22} & \cdots & m_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ m_{n1} & m_{n2} & \cdots & m_{nn} \end{bmatrix}$ |
| No. of People queuing in a station                                  | $X = [x_1, x_2, \ldots, x_n]$ |
| No. of People getting off a train (Ideal)                           | $Y = [y_1, y_2, \ldots, y_n]$ |
| No. of People getting on a train                                    | $A = [a_1, a_2, \ldots, a_n]$ |
| No. of People getting off a train (Real)                            | $B = [b_1, b_2, \ldots, b_n]$ |
| No. of People in a train when it passes                             | $C = [c_1, c_2, \ldots, c_n]$ |
| No. of People getting stuck in a station (A passenger who queues in line, but can’t get on the train when it comes.) | $D = [d_1, d_2, \ldots, d_n]$ |
| Rate of people getting on and queueing                              | $F = [f_1, f_2, \ldots, f_n]$ |
| Initial waiting time (If passengers can get on the first train, the time they wait for. Usually, it is about 2-5 minutes.) | $T$ |
| Subway interval the interval of 2 trains on a line. Usually, 3-8 minutes. | $T_c$ |
| Average waiting time (station)                                      | $T = [t_1, t_2, \ldots, t_n]$ |
| Total average waiting time                                           | $\bar{t}$ |
| Maximum passenger capacity in a train                               | $P$    |

From 1 to $n$, the ideal number of passengers getting on at each station (or queueing) is:

$$x_i = in_i = \sum_{j=1}^{n} m_{ji} \quad (2)$$

From 1 to $n$, the ideal number of people getting off at each station is:

$$\text{out}_i = \sum_{j=1}^{n} m_{ij} \quad (3)$$

The vector $A$ can be controlled by the flow controlling scheme. If we do not have flow control, passengers will get on the train as long as the train is occupied. The number of people getting on is

$$a_i = \min(x_i, P - c_{i-1} + b_i) \quad (4)$$

The number of people in the train is the number of people at the last station, plus the number of people getting on, minus the number of people getting off.

$$c_i = c_{i-1} + a_i - b_i$$

$B$ can only be obtained through statistical methods.

The number of people getting stuck is the number of people queuing minus the number of people getting on the train.

$$d_i = x_i - a_i \quad (5)$$

The ratio of the number of passengers getting on the train to the number of people queuing at the station.
\[ F_i = a_i / x_i \]  \hfill (6)

The average waiting time at a station can be regarded as the initial waiting time (if you can get on the first train when it comes) and the interval waiting time. We assume:

\[ t_i = T_c (1 / F_i - 1) + T_s \]  \hfill (7)

We can use iterative relationships to get relevant information about subway operation without optimization.

**Algorithm 1**: Data analysis of passengers in a train without optimization.

**Require**: N, M, T_s, T_c

**Ensure**: X, Y, A, B, C, D, E, F, T

**Function** WITHOUT_OPTIMIZATION(N, M, T_s, T_c)

```
For i = 0 \rightarrow N do
    X[i], Y[i], A[i], B[i], C[i], D[i], E[i], F[i], T[i] \rightarrow 0
end For

For i = 0 \rightarrow N - 1 do
    For j = 0 \rightarrow N - 1 do
        X[i+1] \leftarrow X[i+1] + M[i][j]
        Y[i+1] \leftarrow Y[i+1] + M[j][i]
    end For

    B[i+1] \leftarrow \min(N, Y[i+1])
    A[i+1] \leftarrow \min(X[i+1], N - C[i] + B[i+1])
    C[i+1] \leftarrow C[i] + A[i+1] - B[i+1]
    D[i+1] \leftarrow X[i+1] - A[i+1]
    E[i+1] \leftarrow \frac{A[i+1]}{X[i+1]}
    F[i+1] \leftarrow \frac{A[i+1]}{X[i+1]}
    T[i+1] \leftarrow T_c \times (1 / X[i+1] - 1) + T_s
end For

end Function
```

### 2.2. The Lines Should be Optimized

We can analyze the flow changes at 3 consecutive situations to know about the situation. It can be roughly divided into 3 different cases.

The first case is that only a few passengers aboard at 3 consecutive stations. At this time, the space in the train was not fully occupied and these stations will not be congested. If passengers are planned on such a route, it can divert traffic and reduce the pressure on congested lines.

In this case, we do not need to control the flow. For example, for the line 9 of Guangzhou Subway, we will have table 2.

\[
M_1 = \begin{bmatrix}
0 & 5 & 5 & 5 & 5 \\
0 & 0 & 5 & 5 & 5 \\
0 & 0 & 0 & 5 & 5 \\
0 & 0 & 0 & 0 & 5 \\
0 & 0 & 0 & 0 & 0
\end{bmatrix}
\]  \hfill (8)

We can analyze it and get the following information, as shown in table 3.
Table 2. Situation low-low-low.

| Passengers | In  | Out  |
|------------|-----|------|
| Station 1  | Low | High/Low |
| Station 2  | Low | High/Low |
| Station 3  | Low | High/Low |

Table 3. Data analysis (situation low-low-low).

| No. | Queue | On | Off | Subway | Delay | Rate   | Wait |
|-----|-------|----|-----|--------|-------|--------|------|
| 0   | 0     | 0  | 0   | 0      | 0     | 100.00% | 4    |
| 1   | 25    | 25 | 0   | 25     | 0     | 100.00% | 4    |
| 2   | 20    | 20 | 5   | 40     | 0     | 100.00% | 4    |
| 3   | 15    | 15 | 10  | 45     | 0     | 100.00% | 4    |
| 4   | 10    | 10 | 15  | 40     | 0     | 100.00% | 4    |
| 5   | 5     | 5  | 20  | 25     | 0     | 100.00% | 4    |
| 6   | 0     | 0  | 25  | 0      | 0     | 100.00% | 4    |

The second case is that a lot of people aboard in the 1st station and many people get off at the 2nd station, then a mass of people gets on at the 3rd station. The second station frees up space in the train and relieves the loading pressure on the third station. It will not cause line congestion, as shown in table 4.

\[
M_2 = \begin{bmatrix}
0 & 50 & 5 & 5 & 5 & 5 \\
0 & 0 & 50 & 10 & 10 & 10 \\
0 & 0 & 0 & 45 & 5 & 5 \\
0 & 0 & 0 & 0 & 50 & 10 \\
0 & 0 & 0 & 0 & 0 & 60 \\
0 & 0 & 0 & 0 & 0 & 0 \\
\end{bmatrix}
\]  

(9)

Table 4. Situation high-low-high.

| Passengers | In    | Out  |
|------------|-------|------|
| Station 1  | High  | Low  |
| Station 2  | Low   | High |
| Station 3  | High  | Low  |

In this case, the train carrying capacity is high, as shown in table 5.

Table 5. Data analysis of situation high-low-high.

| No. | Queue | On  | Off | Subway | Delay | Rate   | Wait |
|-----|-------|-----|-----|--------|-------|--------|------|
| 0   | 0     | 0   | 0   | 0      | 0     | 100.00% | 4    |
| 1   | 70    | 70  | 0   | 70     | 0     | 100.00% | 4    |
| 2   | 80    | 80  | 50  | 100    | 0     | 100.00% | 4    |
| 3   | 50    | 50  | 55  | 100    | 0     | 100.00% | 4    |
| 4   | 60    | 60  | 60  | 100    | 0     | 100.00% | 4    |
| 5   | 60    | 60  | 70  | 90     | 0     | 100.00% | 4    |
| 6   | 0     | 0   | 90  | 0      | 0     | 100.00% | 4    |
The third case is that there are large passenger flows at 3 consecutive stations, and the demanded space is large. Passengers in the 1st station can get on the train. Therefore, part of passengers in the 2nd station can get on and the train is full. At the 3rd station, the number of passengers who can get on the train depends on the number of passengers getting off at the 3rd station. It causes many people getting stuck at the 3rd station. It’s an ill situation. The number of people will increase like a snowball if this situation cannot be improved in time, as shown in table 6.

| Table 6. Situation high-high-high. |
|------------------------------------|
| Passengers | In | Out |
| Station 1   | High | Low |
| Station 2   | High | Low |
| Station 3   | High | Low |

Here is an example. We could assume the initial waiting time ($T_1$) is 4 minutes, interval waiting time ($T_2$) is 6 minutes. And we can find the waiting time at each station ($T$), as shown in table 7.

| Table 7. Data analysis of situation high-low-high. |
|--------------------------------------------------|
| No. | Queue | On | Off | Subway | Delay | Rate | Wait |
|-----|-------|----|-----|--------|-------|------|------|
| 0   | 0     | 0  | 0   | 0      | 0     | 100.00% | 4    |
| 1   | 100   | 100| 0   | 100    | 0     | 100.00% | 4    |
| 2   | 100   | 10 | 10  | 100    | 90    | 10.00%  | 49   |
| 3   | 100   | 20 | 20  | 100    | 80    | 20.00%  | 49   |
| 4   | 100   | 30 | 30  | 100    | 70    | 30.00%  | 49   |
| 5   | 80    | 40 | 40  | 90     | 40    | 50.00%  | 49   |
| 6   | 0     | 0  | 100 | 0      | 0     | 100.00% | 20.875 |

The problem we need to solve is to optimize the number of people getting on the train by flow control. Flow control means to reduce the number of people getting on the train at the previous station so that more people at the later station can get on the train, as shown in figures 3 and 4.

![Figure 3](image1.jpg)

**Figure 3.** Passenger flow at several successive stations would lead to the congestion of the following stations.

![Figure 4](image2.jpg)

**Figure 4.** Limit the flow to several successive stations can alleviate the flow of the following stations.

In this way, we can consider the interests of passengers at the later stations and achieved relative fairness. We can calculate the average waiting time by the data above:
\[ t = \frac{\sum_{i=1}^{n} x_i T_i}{\sum_{i=1}^{n} x_i} = \frac{\sum_{i=1}^{n} x_i [T_i (1/F_i - 1) + T_f]}{\sum_{i=1}^{n} x_i} = \frac{\sum_{i=1}^{n} x_i [T_i (x / \min(x_i, P - c_{i+1} + b_i) - 1) + T_f]}{\sum_{i=1}^{n} x_i} \]  

(10)

The average time in last example is 20.875 mins. The queuing time in station 1 is short, but in station 2 and 3 is long. We want to optimize it. The goal is to minimize average waiting time.

We need to minimize \( t \) by adjusting the value of \( C \). Minimize the average time can reduce passenger’s average waiting time and ensure the fairness of travel of passengers at different subway stations.

3. Constraints and Optimization Algorithms

Among the model above, we can get some relevant constraints. We will use the idea of “short-distance passengers first”. In every station, short-distance passengers can get priority to the train. It can make more space for the train and satisfy more passengers in the later stations. We have

\[
\begin{align*}
0 \leq B_1 & \leq B_1 \\
0 \leq B_2 & \leq B_2 \\
0 \leq B_3 & \leq B_3 \\
0 \leq B_4 & \leq B_4 \\
0 \leq B_5 & \leq B_5 \\
0 \leq B_6 & \leq B_6
\end{align*}
\]

(11)

If the subway is not full, \( B_i \leq B_i \) will be satisfied. The constraints are

\[
\begin{align*}
0 \leq A_1 & \leq \min(X_1, N - C_0 + B_1) \\
0 \leq A_1 & \leq \min(X_1, N - C_0 + B_2) \\
0 \leq A_1 & \leq \min(X_1, N - C_0 + B_3) \\
0 \leq A_1 & \leq \min(X_1, N - C_0 + B_4) \\
0 \leq A_2 & \leq \min(X_2, N - C_0 + B_1) \\
0 \leq A_2 & \leq \min(X_2, N - C_0 + B_2) \\
0 \leq A_2 & \leq \min(X_2, N - C_0 + B_3) \\
0 \leq A_3 & \leq \min(X_3, N - C_0 + B_1) \\
0 \leq A_3 & \leq \min(X_3, N - C_0 + B_2) \\
0 \leq A_3 & \leq \min(X_3, N - C_0 + B_3) \\
0 \leq C_1 & \leq A_1 - B_1 \leq N \\
0 \leq C_1 & \leq A_1 + A_2 + A_3 + A_4 + A_5 + A_6 - B_1 - B_2 - B_3 - B_4 - B_5 \leq N \\
0 \leq C_1 & \leq A_1 + A_2 + A_3 + A_4 + A_5 + A_6 - B_1 - B_2 - B_3 - B_4 - B_5 \leq N \\
0 \leq C_1 & \leq A_1 + A_2 + A_3 + A_4 + A_5 + A_6 - B_1 - B_2 - B_3 - B_4 - B_5 \leq N \\
0 \leq C_1 & \leq A_1 + A_2 + A_3 + A_4 + A_5 + A_6 - B_1 - B_2 - B_3 - B_4 - B_5 \leq N
\end{align*}
\]

(12)

We can solve those inequalities and draw a conclusion. Average waiting time relates to each element in \( A \). Here is the result:

\[
\begin{align*}
\hat{t} = \frac{50000(1/A_1 + 1/A_2 + 1/A_3 + 1/A_4 + 32000/A_5 - 480}{A_1 + A_2 + A_3 + A_4 + A_5 + A_6}
\end{align*}
\]

This is a multi-variate and multi-constrained nonlinear model. There are several algorithms for this kind of problems. Currently, intelligent optimization algorithms are commonly used. Intelligent optimization algorithms include Simulated Annealing Algorithm (SSA), Genetic Algorithm (GA), Ant Colony Algorithm (ACO), etc. [15]. Here, we use Simulated Annealing Algorithm (SSA) to solve this problem.
SSA is based on probability and ideas to steelmaking in nature. The solid is heated to a high level and the particles inside the solid become disordered. When gradually cooled, the particles tend to become ordered and finally reach the ground state at low temperature. Some algorithms are easy to fall into the local optimal solution, but SSA has time-varying probability jump in search process, which eventually approaches 0, which can effectively avoid falling into the local minimum [16].

We fix B and optimize each value of A. It is assumed that the initial temperature is 1000, the lower temperature is $10^{-12}$, the temperature drop rate is 0.99, the iteration times is 1000. Here is the Pseudocode for the proposed algorithm.

**Algorithm 2**: Use SSA to achieve reasonable flow control

Require: X, M, n
Ensure: A

Function F(A)
- $f_A \leftarrow$ a function of A
Return $f_A$
end Function

Function getP(c,t)
- $p \leftarrow \exp(-c/t)$
Return $p$
end Function

Function getA(X,B,C)
For $j=1 \rightarrow N$
do
- $A[i] \leftarrow B[i+1] + \text{floor}(\min((N-C[i]+B[i],X[i]-B[i])) \cdot \text{random}(0,1))$
- $C[i] \leftarrow A[i]-B[i+1]$
end For
Return A,B
end Function

Function SSA(X,B)
- $T \leftarrow 1000$
- $T_{\text{min}} \leftarrow 10^{-12}$
- $\alpha \leftarrow 0.98$
- $k \leftarrow 1000$
- $[A,C] \leftarrow \text{getA}(X,B,C)$
While $T>T_{\text{min}}$ do
For $j=1 \rightarrow k$
do
- $fA \leftarrow F(A)$
- $[A_{\text{new}},C_{\text{new}}] \leftarrow \text{getA}(X,B,C)$
- $f(A_{\text{new}}) \leftarrow F(A_{\text{new}})$
- $\Delta \leftarrow f(A_{\text{new}})-f(A)$
If $\Delta < 0$ then
For $j=1 \rightarrow N$
do
- $A[j] \leftarrow A_{\text{new}}[j] + \text{randint}(0,1)$
end For
Else
- $P \leftarrow \text{getP(}\Delta,T\text{)}$
- If $P > \text{random}(0,1)$ then
- $A \leftarrow A_{\text{new}}$
end If
end While
end Function
Substitute to the above situation, we found that $\bar{t}$ is 11.292 minutes. It is reduced by 45.91% compared with before flow control. The result is shown in table 8.

| No. | Queue | On | Off | Subway | Delay | Rate   | Wait |
|-----|-------|----|-----|--------|-------|--------|------|
| 0   | 0     | 0  | 0   | 0      | 0     | 100.00%| 4    |
| 1   | 100   | 34 | 0   | 34     | 66    | 34.00% | 14   |
| 2   | 100   | 39 | 10  | 63     | 61    | 39.00% | 12   |
| 3   | 100   | 38 | 20  | 81     | 62    | 38.00% | 12   |
| 4   | 100   | 49 | 30  | 100    | 51    | 49.00% | 9    |
| 5   | 80    | 40 | 40  | 100    | 40    | 50.00% | 9    |
| 6   | 0     | 100| 0   | 0      | 0     | 100.00%| 11.292|

4. Case Analysis-Guangzhou Subway Line 3

Guangzhou Subway has been starting new lines every year for the past few years; there are also some detailed lines in the future. Now, Guangzhou subway mileage is about 450 km, but the mileage in 2023 will reach about 800 km [17-24]. As the development of Guangzhou Subway for more than 20 years, although the number of lines and the mileage have been grown, people’s demand and dependence on the subway also raised.

Now, let’s take Guangzhou Subway Line 3 as an example. The schematic diagram of Line 3 is shown in figure 5. Line 3 runs from north to south, Panyu Square in the south and Baiyun Airport in the north. This line has interchange stations with Line 1, Line 2, Line 5, Line 6, Line 7, Line 8, Line 9, Line 14, Line APM and Line GF. The places this line passes are mostly transportation hubs, dense residential area, city centers or tourist attractions. It’s also main line for many people who live in Huadu District, Conghua District or Panyu District but work at city center. This line has the most passenger flow in China. Therefore, this line is extremely crowded especially during morning and evening peak hours, as shown in figure 6.

Line 3 trains are type B, with a total passenger capacity of approximately 1400 passengers [25-27], so $P = 1400$.

The data at each station of line 3 is shown in table 9. If the train is really large and carry every passenger, and we can calculate that the train needs to carry 2462 passengers, which is far beyond the number of seats the subway contains. We know that the starting stations 001, 002, 003, 004 and 005 are all stations with a large number of passengers getting on but few people getting off. Passengers in those stations will make the train occupied, which will affect passengers in stations 006, 007, 008, 009, 010, 011 and 012. In stations 013, 014, 015 and 016, more people will get off the train and the congestion eased. And there is no problem in stations 017, 018 and the rest. The main reason is because the passengers need to go to work by subway in early rush hour. Therefore, lots of people would go from the suburbs to the urban area.

We find that station 014 and 015 have the largest number of people getting off the train. In station 016, 017 and so on, there is no need to limit because people can all get on the train. Therefore, we can use 014 as the end point of the line we analyzed. In real life, Line 3 was built in different periods. Line 3 North, started operation in 2010, is from 001 to 015. The part from 016 to 025 in line 3 was started in 2005 or 2006.
Figure 5. Guangzhou Subway Line 3.

Figure 6. OD Table for Passengers of Guangzhou Subway Line 3 in one train in Rush hours in a morning (estimated).
Table 9. Passenger flow of each station on Line 3.

| No. | Queue | On | Off | Subway | Delay | Rate | Wait |
|-----|-------|----|-----|--------|-------|------|------|
| 000 | 0     | 0  | 0   | 0      | 0     | 100.00% | 4    |
| 001 | 208   | 208| 0   | 208    | 0     | 100.00% | 4    |
| 002 | 175   | 175| 1   | 382    | 0     | 100.00% | 4    |
| 003 | 324   | 324| 8   | 698    | 0     | 100.00% | 4    |
| 004 | 374   | 374| 12  | 1060   | 0     | 100.00% | 4    |
| 005 | 366   | 366| 27  | 1399   | 0     | 100.00% | 4    |
| 006 | 360   | 203| 202 | 1400   | 157   | 56.39%  | 8    |
| 007 | 235   | 23 | 23  | 1400   | 212   | 9.79%   | 50   |
| 008 | 177   | 19 | 19  | 1400   | 158   | 10.73%  | 46   |
| 009 | 199   | 21 | 21  | 1400   | 178   | 10.55%  | 46   |
| 010 | 211   | 25 | 25  | 1400   | 186   | 11.85%  | 41   |
| 011 | 131   | 29 | 29  | 1400   | 102   | 22.14%  | 22   |
| 012 | 99    | 30 | 30  | 1400   | 69    | 30.30%  | 16   |
| 013 | 113   | 113| 145 | 1368   | 0     | 100.00% | 4    |
| 014 | 78    | 78 | 78  | 1368   | 0     | 100.00% | 4    |
| 015 | 429   | 429| 884 | 913    | 0     | 100.00% | 4    |
| 016 | 168   | 168| 440 | 641    | 0     | 100.00% | 4    |
| 017 | 95    | 95 | 132 | 604    | 0     | 100.00% | 4    |
| 018 | 43    | 43 | 162 | 485    | 0     | 100.00% | 4    |
| 019 | 38    | 38 | 134 | 389    | 0     | 100.00% | 4    |
| 020 | 29    | 29 | 113 | 305    | 0     | 100.00% | 4    |
| 021 | 15    | 15 | 86  | 234    | 0     | 100.00% | 4    |
| 022 | 32    | 32 | 105 | 161    | 0     | 100.00% | 4    |
| 023 | 53    | 53 | 131 | 83     | 0     | 100.00% | 4    |
| 024 | 25    | 25 | 40  | 68     | 0     | 100.00% | 4    |
| 025 | 0     | 0  | 68  | 0      | 0     | 100.00% | 4    |

There are 15 subway stations. The average waiting time is calculated as:

\[ t = \frac{\sum_{k=1}^{15} X_k (5X_k / A_k - 1)}{\sum_{k=1}^{15} A_k} \]

In the same way, we use SSA for analysis, and the results are as follows.

\[ A = [134 \ 95 \ 239 \ 300 \ 238 \ 240 \ 135 \ 96 \ 98 \ 95 \ 50 \ 77 \ 113 \ 78 \ 429] \]

Then, we can find the waiting time for each station and the average waiting time.

After optimization, the average time was reduced by 55.93%. It can be seen that our proposed algorithm can effectively improve the average waiting time. The result is shown in Table 10.

In this case, the “flow control” is mainly to limit the passengers in station 001, 002, 003, 004, 005 to ensure the interests of passengers in 006, 007, 008, 009, 010. It’s relatively fair for passengers and reduces the average travel time for passengers across the entire line.

SSA is a heuristic algorithm, so the result we get might not be the optimal solution theoretically. However, there are so many variables and we should make a timely decision when the subway system
is working. Obtaining the optimal solution in a short time requires a lot of storage, which is not realistic. Here, the SSA can be easily implemented and applied.

| No.  | Queue | On  | Off | Subway | Delay | Rate     | Wait |
|------|-------|-----|-----|--------|-------|----------|------|
| 000  | 0     | 0   | 0   | 0      | 0     | 100.00%  | 4    |
| 001  | 208   | 134 | 0   | 134    | 74    | 64.42%   | 7    |
| 002  | 175   | 95  | 1   | 228    | 80    | 54.29%   | 8    |
| 003  | 324   | 239 | 8   | 459    | 85    | 73.77%   | 6    |
| 004  | 374   | 300 | 12  | 747    | 74    | 80.21%   | 5    |
| 005  | 366   | 238 | 27  | 958    | 128   | 65.03%   | 7    |
| 006  | 360   | 240 | 202 | 996    | 120   | 66.67%   | 7    |
| 007  | 235   | 135 | 23  | 1108   | 100   | 57.45%   | 8    |
| 008  | 177   | 96  | 19  | 1185   | 81    | 54.24%   | 8    |
| 009  | 199   | 98  | 21  | 1262   | 101   | 49.25%   | 9    |
| 010  | 211   | 95  | 25  | 1332   | 116   | 45.02%   | 10   |
| 011  | 131   | 50  | 29  | 1353   | 81    | 38.17%   | 12   |
| 012  | 99    | 77  | 30  | 1400   | 22    | 77.78%   | 5    |
| 013  | 113   | 113 | 145 | 1368   | 0     | 100.00%  | 4    |
| 014  | 78    | 78  | 78  | 1368   | 0     | 100.00%  | 4    |
| 015  | 429   | 429 | 884 | 913    | 0     | 100.00%  | 4    |
|      |       |     |     |        |       |          |      |
| Average time | 6.754 min | Before limit | 15.324 min |

This model addresses the issue of “relative fairness to passengers at different stations”. The precondition of our research is “ensure fairness to passengers in each station”. It might be unfair for passengers who really need long distance subway travel. For short-distance passengers, maybe they can take a bus; ride a bicycle and so on. But for long-distance passengers, they can only go by subway. Therefore, if we need to pay more attention to the long-distance passengers, we can have different weights for different stations as the coefficients in the multiplication above.

5. Conclusion
This paper investigates the way to reduce the average waiting time for subway passengers, especially in peak hours. An optimization model is formulated and an algorithm to solve it is proposed. Experiments on real data validate the effectiveness of the proposed solution.

Acknowledgment
This work is supported in part by the National Science Foundation of China (NSFC) with grant no. 61901534, in part by the Guangdong Basic and Applied Basic Research Foundation under Key Project 2019B1515120032, in part by the Science, Technology and Innovation Commission of Shenzhen Municipal Government with grant no. JCYJ20190807155617099, and in part by the SYSU with grant no. 76150-18841214 and 76150-18845101.

References
[1] Da P 2003 The best of subway Construction Workers (11) 52.
[2] Li X 1987 History of New York subway development International Science and Technology Exchange (05) 41-42.
[3] Zheng Z 1996 Centennial development history of Paris subway Urban Public Utilities (02) 10-12.
[4] W Q 2010 Research on the Development of Berlin Urban Rail Transit (1897-1929) (East China Normal University).

[5] Su X 2018 World subway and Japan subway Modern Urban Rail Transit (09) 72-7.

[6] Yu X 2015 The history of Beijing subway (Part 1) Urban Public Transport (08) 67-68.

[7] Du Q and Wang W 2019 Impact of urban subway construction on housing prices along the line Science and Technology Management of Land and Resources 36 (04) 93-103.

[8] Cai C 2018 Co-development with the city to build a beautiful new life: 20 years of practice and exploration of Guangzhou subway operation Urban Rail Transit Research 21 (7) 3-6.

[9] Kang N 2011 Right and wrong of subway current limit Economy (12) 86-89.

[10] Li J 2016 Analysis on the current situation and strategy of the large passenger flow of Guangzhou Subway China High-tech Enterprises (31) 92-94.

[11] Han W 2014 Research on the organization strategy of large passenger flow of Guangzhou subway Line 3 Enterprise Technology Development 33 (20) 148-149.

[12] Li Y 2014 Analysis of the current situation of Guangzhou subway passenger flow and research on countermeasures Science and Technology Wind 2014 (19) 255-257.

[13] Feng S, Chen Y and Xin M 2019 Optimization model of subway coordinated flow controlation under sudden large passenger flow Journal of Harbin Institute of Technology 51 (02) 179-185.

[14] Yang A, Chen Y, Huang J, Xiong J and Wang S 2018 Research on subway dynamic limiting warning model based on short-term passenger flow forecasting Urban Rail Transit Research 21 (10) 29-33.

[15] Jiang N 2016 Overview of intelligent path optimization methods Information Technology (04) 187-189.

[16] Tao G Yuan L, Ma X 2006 Subway Vehicles of Guangzhou Subway Line 3 Electric Drive for Locomotives 2006 (4) 53-59.

[17] Guangzhou Subway 2012 Annual Report The Official Website of Guangzhou Subway.

[18] Guangzhou Subway 2013 Annual Report The Official Website of Guangzhou Subway.

[19] Guangzhou Subway 2014 Annual Report The Official Website of Guangzhou Subway.

[20] Guangzhou Subway 2015 Annual Report The Official Website of Guangzhou Subway.

[21] Guangzhou Subway 2016 Annual Report The Official Website of Guangzhou Subway.

[22] Guangzhou Subway 2017 Annual Report The Official Website of Guangzhou Subway.

[23] Guangzhou Subway 2018 Annual Report The Official Website of Guangzhou Subway.

[24] Guangzhou Urban Rail Transit Phase III Construction Plan (2017-2023) The official website of Guangzhou Subway.

[25] Zhang W 2018 Research on Collaborative Passenger Flow Control of Urban Rail Transit Network (Beijing Jiaotong University).

[26] Chang Y, Zhang Z, Luo J and Qu J 2017 Summary of optimal solution algorithms Civil Engineering Information Technology 9 (03) 115-118.

[27] Zhou X 2015 Optimization analysis of transportation organization during peak periods of Guangzhou Subway Line 3 (including North Three Lines) Enterprise Technology Development 34 (20) 72-74.