Analysis of subway station capacity with the use of queueing theory

Xin-yue Xu a,b,⇑, Jun Liu b, Hai-ying Li a, Jian-Qiang Hu b

a State Key Lab of Rail Traffic Control and Safety, Beijing Jiaotong University, Beijing 100044, China
b School of Traffic and Transportation, Beijing Jiaotong University, Beijing 100044, China

ARTICLE INFO

Article history:
Received 18 September 2012
Received in revised form 15 October 2013
Accepted 30 October 2013

Keywords:
Subway station capacity
Response surface methodology
Queuing network
Remaining passengers
Incoming passenger control

ABSTRACT

A new concept of subway station capacity (SSC) is defined according to the gathering and scattering process. A queuing network analytical model of station is created for calculating SSC, which is built by M/G/C/C state dependent queuing network and discrete time Markov chain (DTMC). Based on the definition and the analytical queuing network, a SSC optimization model is developed, whose objective function is to optimize SSC with a satisfactory rate of remaining passengers. Besides, a solution to the model is proposed integrating response surface methodology with iterative generalized expansion method (IGEM) and DTMC. A case study of Beijing Station in Beijing subway line 2 is implemented to verify the validity and practicability of the proposed methods by comparison with simulation model in different experiments. Finally, some sensitivity analysis results are provided to identify the nodes that have the greatest impact on SSC.

© 2013 Elsevier Ltd. All rights reserved.

1. Introduction

Congestion and pollution problem in China’s large cities has created an urgent need for construction and expansion of the subway system. Many subway stations have been suffering very high levels of pedestrian density. Indeed, several statistical analyses of accident data performed in China, Europe and United States of America (Abril et al., 2008; Baysari et al., 2008; Ben-Elia and Ettema, 2011; Kyriakidis et al., 2012) show that a lot of injuries occur during the boarding and alighting process and that these injuries are closely linked to the design of station platform (Zhang and Han, 2008; Shi et al., 2012). Moreover, transit platforms have critical passenger holding capacities, which if exceeded, could result in passengers being pushed onto the tracks.

As a result, many researches focus on the field of the subway station. Capacity, whose definition is a classical problem, has long been a significant issue in the railway industry. As the passenger traffic demand rapidly increases during recent years, capacity is the hot topic in the subway domain (Kittelson & Associates et al., 2003; Jiang et al., 2009; Hu, 2011; Chen and Liu, 2012). Also, the capacity problem is the key to the station passenger flow organization, station device schemes and train operation. Thus, the method to calculate the SSC is mainly studied in this paper. And, the question will be researched in two steps.

One is to define the SSC. There are a few definitions involving SSC in existing literature. Person capacity, which means the number of people that can be served in a given amount of time, is discussed by Kittelson & Associates et al. (2003). SSC is the maximum passenger flow through various facilities of one station (Washington Metropolitan Area Transit Authority,
2008). However, the facilities are not described in detail because it is difficult to express and present. However, the facilities are not described in detail because it is difficult to express and present. Therefore, SSC should be defined further in practice.

The other is to develop a method to solve the problem of SSC based on the definition this paper has proposed. The methods that are close to this issue can be classified into two main categories according to the suggested modelling approach: simulation and mathematical modelling.

Simulation models (Teknomo, 2006; Fateh et al., 2007; Yalçınkaya and Bayhan, 2009; Asanoa et al., 2010; Fernández, 2010) are suitable for describing station in different perspectives, but few consider the capacity calculation problem. For example, a simulation model based on hybrid Petri nets, which is able to help transit authorities to carry out performance evaluation procedures, was presented by Fateh et al. (2007). A modelling approach based on discrete-event simulation and response surface methodology (RSM) dealt with average passenger travel time optimization problem inherent to the metro planning process (Yalçınkaya and Bayhan, 2009). These studies focused on how to model the station, evaluate the usage of device or level of service (LOS), but existing literatures could not provide a method to calculate capacity integrating simulation models.

Mathematical models, often called analysis models, are generally designed to model the subway station system by means of mathematical formulae or algebraic expressions. Most literatures studied the problem of SSC from a local point of view. The Transit Capacity and Quality of service Manual (Kittelson & Associates et al., 2003) and Code for design of metro (GB, 2003) were some of the works which deal with the capacity of each element (facility) of a subway station, but they were not in a systematic point of view. The relationships between the dwelling time on trains and the crowding situations at the Light Rail Transit (LRT) stations in Hong Kong were firstly determined, and regression models were established for the dwelling delays of train by William et al. (1999). Measurement of boarding and alighting times for different train types has been studied by Harris and Anderson (2006). Based on the field data of passenger boarding time, the time characteristics of boarding passengers were analyzed and a piecewise linear mathematical model for the average boarding time was presented by Cao and Yuan (2009). Arrival process was considered to be continuous and steady and can be assumed to follow a Poisson distribution (Yalçınkaya and Bayhan, 2009; Fernández, 2010). A M/G/C/C-based capacity model of staircases and corridors was proposed for passenger evacuation in consideration of space facilities in metro stations through analysis of passenger movements by Chen and Liu (2012). Remaining passengers (passengers arriving into platform will not be able to depart (board) within the same cycle and should wait for the next) at the subway station platform was studied by probability model (Xu et al., 2013a). However, there was little interest in the evaluation or calculation of SSC in a holistic fashion.

Hence, in order to systematically model all aspects of subway transit stations operating under certain conditions and calculate the capacity, it is urgent to present a new definition of SSC, to develop a systematic station model and then to provide a useful method to solve the calculation problem of SSC.

This paper presents an available definition of SSC, and then a queuing network model of station is created based on M/G/C/C state dependent queuing network and DTMC. Finally, the SSC optimization model is developed and a solution algorithm is proposed the use of RSM integrated with IGEM and DTMC. This paper extends from our previous work (Xu et al., 2013a) which only considered remaining passengers at platforms. In addition, the solution algorithm is inspired by the procedure suggested by Yalçınkaya and Bayhan (2009). This paper is structured as follows. Section 2 presents the problem definition. Section 3 describes the queuing network model for one subway station. Section 4 presents the multi-objective optimization problem and develops a solving algorithm. Later in Section 5 it is shown through a case study, how the proposed approach can be applied to calculate SSC. Finally Section 6 concludes this paper.

2. Problem definition

In this section, after analyzing the gathering and scattering process at the station, a suitable definition to SSC of the subway station is presented. Then, some SSC-related issues including incoming passenger control and alighting and boarding behavior are discussed.

2.1. Gathering and scattering process

Gathering and scattering process at the station usually consists of three sub-procedures: arrival, departure and alighting–boarding process, as shown in Fig. 1. The arrival process begins when passengers enter the gates and ends when passengers leave from the staircase and arrive at the platform. The process contains these key devices: gates, staircases, escalators, walkways, hall. Departure process is similar to arrival process, but different in direction. Alighting–boarding process which includes platforms and trains is a bridge connecting the arrival and departure process.

A single station is denoted by $S$. Considering a single-track alighting–boarding process, inflow passengers arrive continuously, but alighting and boarding passengers get up and down only the time when train stops at the station. Due to the cyclic timetable, the alighting–boarding process has a cyclic characteristic and the definition of the cycle is shown in Fig. 2. In this figure, $t_k$ and $t_k^e$ denote arrival and departure time of the $k$th train at and from $S$, respectively. Index $k$ is a positive integer. Three main parameters are used for building a timetable. The first one is the train headway denoted $T_k$ and defined by the time interval between two successive arrivals of the train at any station of the line ($T_k = t_{k+1}^e - t_k^e$). The second one is the dwell time $\tau_{dk}$ which corresponds to the time the train spends at the station ($\tau_{dk} = t_{k+1}^e - t_{k+1}$). The third one is the
separation time \( \tau_{Sk} \) defined by the time interval between the train departure and the next train arrival \( \tau_{Sk} = t_{k+1} - t_{k} \). Actually, \( T_k \) represents the period of the timetable and, consequently, it also corresponds to the operating period of the station as shown in Fig. 2.

### 2.2. Definition of SSC

According to the gathering and scattering process, SSC can be defined as the maximum number of passengers that can be served by one station in a given period of time, typically 1 h, under specified operating conditions, without unreasonable delay, hazard, or restriction, and with reasonable certainty.

Note that, there are some important points that should be focused on. One is the words “number of passengers”, which not only means the passengers boarded on the train from the station, but also the outbound passengers from alighted passengers. “Specified operating conditions” denotes that the number of people that can be served depends on the number of vehicles operated and the size of those vehicles. It is important to specify whether the reported capacity reflects the current timetable (number of trains getting through), vehicle capacity (number of passengers could be carried in the maximum loading factor), or some other conditions. Another named “without unreasonable delay, hazard, or restriction” indicates that all trains run strictly by timetable and station operates in a normal safety state. Finally, “with reasonable certainty” suggests that SSC should reflect LOS of station for comfort and security purposes, usually considering the maximal density of travelers into public transports or on waiting/queuing areas of station platforms (Kittelson & Associates et al., 2003), average service time, remaining passengers and so on.

So, the number of served passengers \( N^S \) in 1 h, is equal to the outbound passenger volume in 1 h \( N^{out} \), plus the sum of passengers getting on from trains of class \( i \) \( N^b_i \). This can be formulated as follows:

\[
N^S = N^{out} + \sum_{i=1}^{2n} N^b_i
\]  

(1)

in which \( 2n \) is the number of trains passing through the station in two directions in 1 h.

Besides, the total number of passengers getting on the train from the station should not exceed the incoming passenger volume during the unit time, as follows:

\[
N^S \leq N^{out} + N^p \leq N^{out} + N^{in}
\]  

(2)

in which \( N^p \) is the total number of passengers arriving at platform in 1 h, \( N^{in} \) is the incoming passenger volume in 1 h.
2.3. SSC-related issues

Based on previous sections, SSC may be determined by incoming passenger volume, incoming passenger characteristics, timetable, capacity of one single device, device utility control strategy of the whole station and passenger control strategy of the station and so on.

In Beijing subway, incoming passenger volume is so huge that major devices are overloaded in some stations in the rush hour of a day (Beijing Subway, 2013) and \( N^m \) will be controlled by station managers if \( N^m \geq 0.7 \times N^s \) (Beijing Municipal Administration of Quality and Technical Supervision, 2011). Moreover, the headway time of some lines is close to their minimum design time and timetable is usually fixed in different time period of different days for convenience. So, passenger control strategy of the station is used by lots of stations in the rush hour. Incoming passenger control is one kind of passenger control strategy widely used in China as shown in Fig. 3, which aims at keeping incoming passenger volume and incoming passenger flow rate under certain level, respectively. The incoming passenger usually queued through moveable handrails in the square of the subway station when the incoming passenger control is used. Moreover, Fig. 3 shows that the queue length and service time can be adapted by changing the width and length of handrails.

According to observations of passenger alighting and boarding behavior and an analysis of field video data (see Appendix A) collected in some subway stations in Beijing between 7:00 and 9:00 in 1 week, some useful characteristics about alighting and boarding passengers were found. Firstly, alighting and boarding processes take place consecutively rather than simultaneously. Secondly, the number of alighting passengers of each train in the same direction may be either the same or follow a Poisson distribution. Finally, all alighting passengers could leave the platform when the next train arrives in the rush hour. Some of similar features in Beijing Subway were obtained by other authors (Hu, 2011; Zhao et al., 2011).

Thus, in this paper we consider how to optimize SSC with a satisfactory rate of remaining passengers on station platforms under fixed timetable, fixed routing probability at the station and changeable arrival rate of incoming passenger.

2.4. Assumptions

Assume that: (1) alighting and boarding passengers should follow the rule: alighting first. Moreover, the arrival distribution is considered to be invariant with time; (2) the running timetable in two directions are the same; (3) all alighting passengers could leave the platform when the next train arrives and total number of alighting passengers in each cycle is deterministic and independent of time.

3. Queuing network analytical model for a subway station

In this section, gathering and scattering process of the subway station is modelled by a queuing network, while the performance evaluation method is also developed. We begin by constructing a single node modelling based on queueing theory. Then, we present a queuing network for subway station. Finally, DTMC and IGEM are introduced for queuing network performance evaluation.

3.1. Single node modelling

In this subsection we present a methodology for modelling a single node (staircase, walkways, gate, hall, escalator) with uni-directional pedestrian flows as an M/G/C/C state dependent or M/M/C/C queue. Suppose that the arrival time of one node for each pedestrian is independent. Therefore, it is assumed that pedestrians arrive at a Poisson process with rate \( \lambda \). The service time of nodes modelling elements like gates, escalators and halls conforms to the exponential distribution, so the node (gate, escalator, hall) can be modelled by the M/M/C/C model.

Fig. 3. Passenger control strategy of Longze Station in rush hour.
The passenger movement area of a node (staircase, walkways) may be seen as \( c \) parallel servers to its occupants, which is also the total number of users allowed in a system where there is no buffer or waiting space. Secondly, the service time of the queuing model of a node is equal to the time for a pedestrian to traverse the entire node, which depends on the number of users currently in the system. As a consequence, a M/G/C/C state dependent queuing model seems to be a reasonable tool to describe a single node (Yuhaski and Smith, 1989; Mitchell and Smith, 2001; Chen and Liu, 2012).

The limiting probabilities for the random number of entities \( N \) in an M/G/C/C queuing model, \( p_n = \text{Pr} \{ N = n \} \), are as follows (Yuhaski and Smith, 1989):

\[
p_n = \left\{ \frac{[\lambda E(T_1)]^n}{n!f(n)f(n-1)\ldots f(1)} \right\} p_0 \quad n = 0, 1, \ldots, c
\]

in which \( n = 1, 2, \ldots, c, p_0 \) is the empty system probability, given by

\[
p_0^{-1} = 1 + \sum_{i=1}^{c} \left\{ \frac{[\lambda E(T_1)]^i}{i!f(i)f(i-1)\ldots f(1)} \right\}
\]

\[
c = [\rho W]
\]

In this model, \( E(T_1) = l/V_1 \) represents the expected service time of a lone passenger in a node of length \( l \). \( V_1 \) is the speed of a lone passenger, and \( c \) is the capacity of node space, \( \lfloor x \rfloor \) is the integer part of \( x \); \( w \) is the width in number of node, and \( \rho \) is the max density. The service rate \( f(n) = V_n/V_1 \) is considered to be the ratio of the average walking speed of \( n \) people in the node to that of a lone occupant. The average walking speed of \( n \) pedestrians in the node is calculated using the exponential model for pedestrian walking speeds versus crowd density within the node, following the expression:

\[
f(n) = \exp\left[ -\left( \frac{n-1}{\beta} \right)^\gamma \right]
\]

with

\[
r = \ln\left( \frac{\ln(V_n/V_1)}{\ln(V_1/V_1)} \right) / \ln\left( \frac{a-1}{b-1} \right) \quad \beta = \frac{a-1}{\ln(V_1/V_0)}^{1/\gamma}
\]

in which the values \( a \) and \( b \) are arbitrary points used to adjust the exponential curve.

From Eq. (3), important performance measures can be derived:

\[
p_c = \text{Pr} \{ N = c \}
\]

\[
\theta = \lambda(1 - p_c)
\]

\[
L = E(N) = \sum_{i=1}^{c} ip_i
\]

\[
W = L/\theta
\]

in which \( p_c \) is the blocking probability, \( \theta \) is the throughput in per/hour, \( L \) is the expected number of customers in the node, and \( W \) here derived from Little’s formula, is the expected service time in hours.

### 3.2. Queuing network modelling

In this model, the gathering and scattering process is a mixed continuous and discrete queuing system, which consists of arrival-departure process and alighting-boarding process. On one hand, arrival-departure process can be considered as a multilevel queuing network system with blocking, and the devices like gates, stairs, walkways and train doors can be regarded as the network nodes which can be described by M/G/C/C state dependent queuing model or M/M/C/C model; on the other hand, alighting-boarding process is a discrete queuing process, which can be modelled as a DTMC (Xu et al., 2013a). In summary, the gathering and scattering process model is divided into two subsystems, of which the first is built by M/G/C/C state dependent queuing network and the second is constructed by probability theory of remaining passengers. Moreover, the output of the first subsystem is the input of the second subsystem, as shown in Fig. 4. In the following, how to develop the queuing network model will be discussed.

#### 3.2.1. First subsystem network

In the arrival (departure) process, passengers receive service at one node (device), and then leave the station or arrive at another node to continue to receive service in accordance with a given routing probability. If each node is modelled by a queuing system, the routing network of pedestrian in the arrival (departure) process can be considered as an open queuing network.
Proposition 1. In the M/G/C/C state dependent model, the departure process (including both customers completing service and those are lost) is a Poisson process (Cheah and Smith, 1994).

The departure process (including both customers completing service and those are lost) of the M/M/C/C model is also a Poisson process at the rate. Hence, the arrival (departure) process can be modelled by a M/G/C/C state dependent queuing network, where some node developed by M/M/C/C can be uniformly labeled M/G/C/C, as shown in Fig. 5. Take for example the passenger arrival process, passengers arriving at the in-gate of the station’s entrance are considered as the input of the queuing network and the number of passengers arriving at platform is the system output.

Deriving performance measures for M/G/C/C state dependent queues configured in networks is a task considerably more complex because of the inter-blocking effects. An algorithm available is generalized expansion method (GEM), which was successfully used to estimate performance measures for finite queuing networks (Kerbache and Smith, 2000; Cruz and Smith, 2007). The GEM is basically a combination of repeated trials and node-by-node decomposition in which each queue is analyzed separately and then corrections are made in order to take into account the interrelation between the queues in the network. The GEM uses blocking after service, which is prevalent in most production and manufacturing, transportation, and other similar systems.

A new iterative algorithm presented in Fig. 6 is recently proposed by Cruz and Smith (2007), which is different from GEM by means of computing the service times in M/G/C/C queues straightforward. In this paper, we refer to the algorithm as IGEM and present an overview of the method by Fig. 6. Important performance measures are easily computed with high accuracy by this algorithm, such as the blocking probability, throughput, expected number of customers in the system, and expected waiting time.

3.2.2. Second subsystem network

Considering a single-track arriving and boarding process, the whole service procedure can be considered as a cyclical discrete and continuous queuing system, but the number of passengers changing over time does not have the Markov property. So traditional queueing theory cannot be directly used for the mixed queuing system. However, if the variation of number of

![Fig. 4. Queuing network of one subway station.](image)

![Fig. 5. M/G/C/C state dependent queuing network of passenger inflow process.](image)
passengers at the end time of each cycle is analyzed, remaining passengers process at the end of each cycle can be considered as a Markov chain process. Thus, remaining passengers length at the end of each cycle is calculated using DTMC.

Let $Q_{\text{max}}$ be the assumed maximum value of the queue length, which can be accommodated in the considered platform, $a$ be the average rate of passengers arriving at platforms, $s$ be the average service rate or departure rate, $s_{\text{max}}$ be the maximum service capacity within one cycle and $a_{\text{max}}$ be the maximum number of passengers arriving at platforms within a cycle. The number of passengers arriving at platforms, and the number of served passengers during the $k$th cycle, are denoted as $a_k$ and $s_k$ respectively. $Q_k$ is the number of remaining passengers at the end of the $k$th cycle. The queue process can be computed in a stochastic fashion by first computing the transition matrix $p_{ij}(k)$, which gives the probability that the queue length moves from queue state $i$ at the time $k-1$ to state $j$ at the time $k$. Assuming that arrivals are stationary and independent with known probability distribution functions $P_a$ and service probability functions $P_s$ is also known (e.g., normal distribution), the transition probability is expressed as follows:

$$p_{ij}(k) = \sum_{s_k = 0}^{a_{\text{max}}} P_a(n = j - i + s_k) P_s(s_k) \quad \forall j - i + s_k \leq a_{\text{max}}$$

Since queues are constrained to be non-negative, when the departures are larger than the sum of the arrivals and the queue at the start of the cycle, the queue at the end of the dwell phase will be zero. Obviously, part of this dwell phase will not be used by any passenger. According to this consideration the chance of a queue $i$ to become zero (no remaining passenger) is computed with the following condition:

$$p_{i0}(k) = \sum_{n=0}^{s_{\text{max}} - i} P_a(n \leq s_k - i), \forall s_k - i \leq a_{\text{max}}$$
An analogous consideration holds for the probability of queues that are larger than the maximum number of passengers allowed in a platform:

\[ p_{Q_{\text{max}}}(k) = \sum_{n=Q_{\text{max}}-i+s_k}^{a_{\text{max}}} P_i(n > Q_{\text{max}} - i + s_k) p_{i}(s_k) \forall Q_{\text{max}} - i + s_k \leq a_{\text{max}} \]  

(13)

Every time step \( k \) is uniquely determined once and an initial condition \( Q_0 \) is assumed. Since the queue probability distribution at every time \( k - 1 \) and the transition matrix \( p_{ij} \) are as defined, independent, the probability of each state \( j \) observed at the end of cycle \( k \) is given by:

\[ \Pr(Q_0 = j, k) = \sum_{i=0}^{Q_{\text{max}}} \Pr(Q_0 = i, k - 1) p_{ij}(k) \]  

(14)

Expected value of remaining queue is computed with the following equation:

\[ E[Q_0(k)] = \sum_{j=0}^{Q_{\text{max}}} j \Pr(Q_0 = j, k) \]  

(15)

According to the attribute of remaining passengers, the expected value at any time \( t \), \( E[Q_0(t)] = E[Q_0(k)] \), where \( \sum_{i=0}^{k} T_i \leq t \leq \sum_{i=1}^{k+1} T_i \) and \( T_0 = 0 \).

We can assume that the arriving and boarding process in different directions have little interaction and can be considered as independent with each other. In fact, few passengers having alighted from one direction of the platform will get aboard the other. Due to the same timetable, we can easily model the whole arriving and boarding process similar to the single-track (Xu et al., 2013a).

3.3. Limitations of the model

The proposed model has two main limitations. Firstly, we simply model the bi-directional facility (stair, walkway, hall, etc.) by two independent parts. In fact, some parts of the subway station, which serve bi-directional movement, are physically separated as much as possible by mobile facilities as shown in Fig. 7. Capacities of the two small split facilities can be dynamically adapted to passenger flow change by administrators of subway station. But, halls are usually shared by two different passenger flows. It is possible to raise the model accuracy if a hall is simplified to two smaller independent halls. Fortunately, the capacity of hall in station of Beijing subway is bigger than other devices enough to omit the interaction of the two different passenger flows from a macroscopic point of view (Chen and Wu, 2009; Jiang et al., 2010; Hu, 2011).

Secondly, the suggested model assumes that passenger departure process at platform is a Poisson process and all alighting passengers can get out. So \( \lambda_0 \) can be estimated by the average number of alighting passengers in each cycle. Although there are some references are relative to the similar assumption in evacuation scene (Smith, 2011; Chen and Liu, 2012) and rush hours (Cao et al., 2009), the characteristic of passenger departure process will be studied and further calibrated by more research data. It is an interesting future work.

4. Multi-objective optimization of SSC

In this section, we consider the multi-objective optimization of SSC where queuing methods (IGEM and DTMC) are incorporated in RSM.

Fig. 7. Passenger flow organization of Xi ZhiMen Station at staircase or walkway in rush hour.
4.1. Model formulation

Based on the queueing network model, the total passengers getting on train is the number of passengers arriving at the platform, minus the number of remaining passengers. Meanwhile, the outbound passenger volume is equivalent to the total passengers getting off, which is the sum of passengers getting off from each train. These can be formulated as follows:

\[
\sum_{i=1}^{2n} N_i^b = N^p - E[Q_0(3600)] = N^p - g(N^p)
\]

\[
N_{out} = \sum_{i=1}^{n} N_i^p
\]

\[
N_{in} = 3600 \sum_{i=1}^{m} \lambda_i
\]

\[
N^p = 3600 \lambda^* = \Theta(\lambda)
\]

in which \(\lambda_i\) is the Poisson rate arriving at the entrance \(i\), \(\lambda^*\) is the Poisson rate arriving into platforms, \(\Theta(\lambda)\) is the throughput of the first subsystem model calculated by IGEM, \(E[Q_0(3600)] = g(N^p)\) is the remaining passengers at the end of 1 h calculated from the second subsystem model developing by probability theory, \(m\) is the number of entrances.

So, the number of served passengers can be reconstructed by:

\[
N^s = N^p - g(N^p) + \sum_{i=1}^{n} N_i^p
\]

\[
= \Theta(\lambda) - g[\Theta(\lambda)] + \sum_{i=1}^{n} N_i^p
\]

\[
\triangle f(\lambda_1, \lambda_2, \ldots, \lambda_m)
\]

where \(f(\lambda_1, \lambda_2, \ldots, \lambda_m)\) is the function relation between \(N^s\) and \((\lambda_1, \lambda_2, \ldots, \lambda_m)\). Although the expression in Eq. (20) can be calculated by analytical model presented in Section 3, the function is too complicated for us to hardly obtain the explicit expression.

The problem described can be formulated as a multi-objective optimization problem with two conflicting objectives, minimizing the number of remaining passengers at station platform and maximizing the number of served passengers. The following mathematical formulation represents the multi-objective optimization problem:

\[
\begin{align*}
\min & \quad F(\lambda) = (-f(\lambda_1, \lambda_2, \ldots, \lambda_m), h(\lambda_1, \lambda_2, \ldots, \lambda_m))^T \\
\text{s.t.} & \quad \lambda_i \leq \lambda_i^{\text{max}}, \forall i \in \{1, 2, \ldots, m\} \\
& \quad \lambda_i \geq 0, \forall i \in \{1, 2, \ldots, m\}
\end{align*}
\]

in which the objective can be solved by IGEM and DTMC, \(\lambda^{\text{max}}\) is the maximum Poisson rate arriving at the entrance \(i\), which is determined by maximum capacity of single node in the streamline of the queueing network.

4.2. Solution algorithm for SSC problem

To solve the SSC problem formulation, Eq. (21), we use a powerful class of optimization method called RSM. RSM was chosen because it is quite fit for the multi-objective optimization problems where the form of the relationship between independent variables and the response is unknown, and has to be approximated. RSM is a collection of mathematical and statistical techniques that are useful for the modelling and analysis of problems in which a response of interest is influenced by several quantifiable variables (or factors), with the objective of optimizing the response (Montgomery, 2005).

Integrating with the IGEM and DTMC, overall construction of our algorithm is illustrated in Fig. 8, which is an implementation of the procedure suggested by Yalcinkaya and Bayhan (2009). The first step in RSM is to find a suitable approximation to the true relationship between response and the set of independent variables. The most common forms are low-order polynomials (first or second-order). If response varies in a linear manner, the response can be represented by this linear function equation:

\[
y_{\varphi} = \beta_0 + \sum_{i=1}^{m} \beta_i \lambda_i + \sum_{j=1}^{m} \sum_{i=1}^{l} \beta_{ij} \lambda_i \lambda_j + e_{\varphi} \quad \varphi = 1, 2
\]

where \(y_{\varphi}\) are the output variables (i.e. number of served passengers and remaining passengers), \(\beta_i\) refer to unknown coefficients, and their values are, on the basis of experiments, determined by the method of least squares. \(e_{\varphi}\) is a random value that cannot be explained by the model.
But if curvature is there in the system, a higher order polynomial like quadratic model is used which can be stated in the form of the following equation:

\[ y = \beta_0 + \sum_{i=1}^{m} \beta_i x_i + \sum_{i=1}^{m} \sum_{j=1}^{1} \beta_{ij} x_i x_j + \varepsilon \quad \phi = 1 \ 2 \]

For each experiment, the output responses of interest, number of served passengers and remaining passengers are measured using IGEM and DTMC. The regression model, Eq. (23), is a relationship that shows the responses (number of served passengers and remaining passengers) according to the values of \( x_i \). Therefore, this model can be used to find the input variables (Poisson rate) that produce the desired responses.

Lastly, in order to solve this optimization problem, that is, for determining the input variables to optimize two responses, the Derringer–Suich multi-response optimization procedure is used. The Derringer–Suich method uses a desirability function in which the priorities and desires on the response values are built into one optimization procedure (Derringer and Suich, 1980).

Let \( y, A, B \) and \( C \) be predict value, the lower, target, and upper values, respectively, that are desired for response, with \( A \leq B \leq C \). If a response is of the “target is best” kind, then its individual desirability function \( d_i \) is:

\[ d_i = \begin{cases} 0 & \hat{y} < A \quad \text{or} \quad \hat{y} > C \\ \left(\frac{\hat{y} - A}{B - A}\right)^s & A \leq \hat{y} < B \\ \left(\frac{C - \hat{y}}{C - B}\right)^t & B \leq \hat{y} \leq C \end{cases} \]

with the exponents \( s \) and \( t \) determining how important it is to hit the target value. For each response, we can select a weight from 0.1 to 10 to emphasize or de-emphasize the target. Less, equal, or more emphasis is given on the target if a weight is (a) less than one (minimum is 0.1), (b) equal to one, or (c) greater than one (maximum is 10) respectively.

For each response, desirability function \( d_i \) is assigned numbers between 0 and 1 to the possible values, with \( d_i = 0 \) representing a completely undesirable value and \( d_i = 1 \) representing a completely desirable or ideal response value. The individual desirabilities are then combined using the geometric mean, which gives the overall desirability (\( D \)):

\[ D = (d_1 \cdot d_2 \ldots d_m)^{1/m} \]
After the overall desirability function is defined, input variables to maximize the overall desirability function by the optimizing algorithm are determined.

5. Case study

In this section, we illustrate the queueing network analytical model and the solving algorithm using RSM in a subway station. The accuracy of the method is studied using a simulation model developed by Xu et al. (2013b). Moreover, some sensitivity analysis results are provided to identify the nodes who have the greatest impact on SSC.

5.1. Station description and model development

The proposed case study, illustrated in Fig. 9, is a typical subway station in Beijing subway line 2 named “Beijing” and located in the town center of Beijing. The station involves a multi-modal interchange between subway and railway (Beijing railway station) and thus supports a lot of connections of passengers. The Station has four entrances, two of them are bi-directional (in-outbound), and the others are uni-directional. Like most of the stations in the subway network of Beijing, the station has an island platform, whose length equals to 100 m, effective width equals to 8 m, and one input–output gate (stair) with 5.8 m width. The subway train serving this station consists of six vehicles. Each vehicle is fitted out with four shared alighting–boarding gates. The headway time and dwell time running by the current timetable are 165 s and 30 s, the maximum load factor is less than 130% during the peak time. Other key devices belong to arrival network and their pedestrian characteristics are described in Table 1. The purpose of this case study is to evaluate the volume of passengers who are able to be served under the current design of this station during rush hours and to ensure a good LOS.

Observations on passenger arrivals were taken from the records of Automatic Fare Collection (AFC) and flow investigation data, more than 1 month, located in every entrance. Input analyser module of Minitab was used to determine which distribution best fits ($\alpha = 0.05$) the data. The Exponential distribution and Poisson distribution provided the best fit for the time between the arrival of passengers and number of passengers, respectively. From the real video data gathered at platform and from the literature (Zhang and Han, 2008; Cao and Yuan, 2009; Chen and Liu, 2012; Xu et al., 2013b), it seems that the Normal distribution best fits the number of passengers getting on trains. Passenger routing probabilities were estimated as the ratio of the passengers who arrived at an entrance and followed the specific route, by the total number of passengers who entered the system from this entrance. These calculations were made using the data obtained from the questionnaire conducted by the metro company with 1000 passengers between 8.00 and 9.00 h within 1 week. Total number of alighting passengers in each cycle is deterministic and equal to 570, that is, alighting rates is approximately 0.2.

The simplified model of subway station can be developed based on the method mentioned in Section 3, which is shown in Fig. 10.

5.2. Experimental design and optimal solution

In the present study, three input parameters (Poisson rate arriving at the entrance) were chosen as independent factors in the experimental design. Two-level full factorial designs with central runs were designed to fit first order regression models for both responses (number of served passengers and remaining passengers). The low and high level of input factors (Poisson rate) for two-level full factorial designs, determined by experiences and simple network analysis, are: $\lambda_1 \in [1, 3]$, $\lambda_2 \in [1, 3]$, $\lambda_3 \in [1, 3]$. The most popular class of second order designs called central composite design (CCD) was used for the response surface methodology in the experimental design. The design matrix contained a $2^3$ factorial design augmented by six axial points ($a = 1$, and design was face-centered) and six replications at the center point (all factors at zero

Fig. 9. Beijing subway station network.
level) to evaluate the pure error. This design leading to a total number of 20 experiments was carried out in random order as required in many design procedures, as shown in Table 2. Experimental data were analyzed using the Minitab 14 software. The variables were coded according to the following equation:

\[ k_{ci} = \frac{k_i}{C0_k/C0_i D_{ki}} \]

where \( k_{ci} \) is the coded value of \( k_i \), \( k_i/C0_i \) is the value of \( k_i \) at the center point of the investigated area, and \( D_{ki} \) is the step size. The quality of fit of the polynomial model equation was expressed by the coefficient of determination (\( R^2 \)) and the responses were completely analyzed using analysis of variance (ANOVA).

The obtained responses in Table 2 were correlated with the three independent variables using two polynomial equations, Eqs. (26) and (27). Least squares regression was used to fit the obtained data to Eqs. (26) and (27). The best-fit models where the coefficient of multiple determinations and the output variable (\( y_1 \): number of served passengers, and \( y_2 \): remaining passengers) are as follows:

\[ R^2_{y_1} = 98.31\% \quad R^2_{y_2} = 77\% \]

\[ y_1 = 30657.5 + 931.4 k_{c1} + 1643.5 k_{c2} + 1213.3 k_{c3} - 308.9 k_{c1}^2 - 308.9 k_{c2}^2 - 514.9 k_{c3}^2 - 409.6(k_{c1} k_{c2})^2 - 971.8(k_{c1} k_{c3})^2 - 706.1(k_{c2} k_{c3})^2 \]  

(26)

\[ y_2 = 200.2 + 101.8 k_{c1} + 315.2 k_{c2} + 279 k_{c3} + 235.6 k_{c1}^2 + 91.4(k_{c1} k_{c2})^2 + 188.6(k_{c1} k_{c3})^2 \]  

(27)

Since the computed \( F \)-values (64.57, 10.08) are much greater than the tabular \( F \)-value (2.76) at the 5% level, the best-fit models were applied for prediction, and the Derringer–Suich multi-response optimization procedure was used for optimization.

In this case study, the maximum value of remaining passengers in single direction is not greater than 240 persons for comfort and security purpose, which means that number of remaining passengers at each gate of one train is less than 10. In fact, remaining passengers at the end of each cycle will always stay in the waiting/queuing areas of station platforms until they get aboard the next train and number of passengers on the platforms including remaining passengers may reach to
maximum point when the train arrives. So the number of remaining passengers has important effect on maximum density of passengers on station platforms, which determines the safety and comfort of subway station (Kittelson & Associates et al., 2003). The limit number of served passengers is the total number of passengers getting off plus maximum total number of passengers getting on. We use Eq. (24) to solve the optimization procedure and the composite desirability values are found by using the following equation:

$$D(k) = \left( d_1(Y_1(k)) d_2(Y_2(k)) \right)^{1/2}$$  \hspace{1cm} (28)

The current and proposed (optimum) Poisson rate that maximize the overall desirability ($D$) and SSC are $\lambda^* = (1.77, 2.07, 2.48)^T$, $Y^* = (30988, 324)^T$, respectively. An additional experiment was done under the optimum conditions to confirm the agreement of the model and experimental results. The experimental value $(30824, 288)^T$ closely agreed with the predicted result from RSM and hence validated the findings of response surface optimization.

5.3. Consistency with simulation model

To simulate passenger flow at subway station under the same modelling assumptions of the queuing model we developed a simulation program based on Simulink (see Xu et al., 2013b, for an overview of the features in this model). Some reference (Karris, 2012) may be useful to readers who are interested in how to use the Simulink tool to simulate real system.

The simulation program includes some key objects (staircase, walkways, gate, hall, escalator, platform and trains), which are connected by the gathering and scatter process. The framework of this program is hierarchical and reusable, and each module represents one object only. Moreover, the service process of device is implemented by queue model and the process of train is modelled by Stateflow module of Simulink. Based on the structure and data presented in previous section, the simulation program of Beijing subway station was built. And the program is proved to be effective comparing with real data by Xu et al. (2013b).

We test the consistency and performance of the queuing network model under different demand scenarios which were designed by Section 5.2. In Table 3 a comparison is given between the analysis model (queuing network model) and simulation model. The column labeled “Experiment No” indicates the experiment point that are in accord with the column labeled “run order” of Table 2. For each experiment we generated 10 repetitions for simulation method, the simulation results obtained from the average of 10 runs. The deviation between analysis and simulation mode is less than 4%, which do not increase dramatically with the change of variables. It is very encouraging and confirms the efficiency of analysis model for capacity calculation.

SSC is $Y^* = (32012, 312)^T$, which is obtained by simulation model, compared with the one computed using the queuing network analytical model. It is shown that the analytical result is consistent with that of the simulation, which requires more computational efforts and parameters as shown in Table 3, however.
Table 3
Comparison between analysis and simulation model.

<table>
<thead>
<tr>
<th>Experiment no</th>
<th>$N^*$</th>
<th>Queuing</th>
<th>Simulation</th>
<th>$E[Q(t)]$</th>
<th>Queuing</th>
<th>Simulation</th>
<th>CPU times (s)</th>
<th>Queuing</th>
<th>Simulation</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>28373</td>
<td>28320</td>
<td>0</td>
<td>0</td>
<td>4</td>
<td>757</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>29089</td>
<td>28882</td>
<td>2</td>
<td>0</td>
<td>4</td>
<td>672</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>27965</td>
<td>27324</td>
<td>0</td>
<td>0</td>
<td>4</td>
<td>774</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>30989</td>
<td>32189</td>
<td>560</td>
<td>539</td>
<td>5</td>
<td>724</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>30796</td>
<td>31961</td>
<td>60</td>
<td>54</td>
<td>5</td>
<td>720</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>30988</td>
<td>30554</td>
<td>1122</td>
<td>1161</td>
<td>5</td>
<td>731</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>23990</td>
<td>22378</td>
<td>0</td>
<td>0</td>
<td>5</td>
<td>724</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 4
Changeable parameters for each type node.

<table>
<thead>
<tr>
<th>Type node</th>
<th>Node</th>
<th>Changeable parameter</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gate</td>
<td>1, 4, 8</td>
<td>Amount</td>
</tr>
<tr>
<td>Staircase</td>
<td>2, 5, 9, 12, 14</td>
<td>Width</td>
</tr>
<tr>
<td>Walkway</td>
<td>3, 7, 10</td>
<td>Width</td>
</tr>
<tr>
<td>Escalator</td>
<td>6</td>
<td>Velocity</td>
</tr>
<tr>
<td>Hall, platform</td>
<td>11, 13</td>
<td>Area</td>
</tr>
</tbody>
</table>

Table 5
SSC resulting from a unit increase of node capacity (changeable parameters).

<table>
<thead>
<tr>
<th>Node</th>
<th>Solution details</th>
<th>$Y^*$</th>
<th>Node</th>
<th>Solution details</th>
<th>$Y^*$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>$\lambda' = (1.65, 2.05, 2.48)$</td>
<td>(29657, 284)</td>
<td>8</td>
<td>$\lambda' = (1.77, 2.07, 2.97)$</td>
<td>(31601, 358)</td>
</tr>
<tr>
<td>2</td>
<td>$\lambda' = (1.77, 2.07, 2.48)$</td>
<td>(30988, 324)</td>
<td>9</td>
<td>$\lambda' = (1.77, 2.07, 2.48)$</td>
<td>(30988, 324)</td>
</tr>
<tr>
<td>3</td>
<td>$\lambda' = (1.77, 2.07, 2.48)$</td>
<td>(30988, 324)</td>
<td>10</td>
<td>$\lambda' = (1.77, 2.07, 2.48)$</td>
<td>(30988, 324)</td>
</tr>
<tr>
<td>4</td>
<td>$\lambda' = (2.02, 1.16, 2.53)$</td>
<td>(30493, 316)</td>
<td>11</td>
<td>$\lambda' = (1.77, 2.07, 2.48)$</td>
<td>(30988, 324)</td>
</tr>
<tr>
<td>5</td>
<td>$\lambda' = (1.77, 2.07, 2.48)$</td>
<td>(30988, 324)</td>
<td>12</td>
<td>$\lambda' = (2.29, 2.21, 2.51)$</td>
<td>(33469, 418)</td>
</tr>
<tr>
<td>6</td>
<td>$\lambda' = (1.77, 2.07, 2.48)$</td>
<td>(30988, 324)</td>
<td>13</td>
<td>$\lambda' = (1.77, 2.07, 2.48)$</td>
<td>(30988, 324)</td>
</tr>
<tr>
<td>7</td>
<td>$\lambda' = (1.77, 2.07, 2.48)$</td>
<td>(30988, 324)</td>
<td>14</td>
<td>$\lambda' = (1.77, 2.07, 2.51)$</td>
<td>(31325, 356)</td>
</tr>
</tbody>
</table>

5.4. Sensitivity analysis

Sensitivity analysis of the SSC is conducted to identify the nodes that have the greatest impact on the SSC in terms of a satisfactory rate of remaining passengers. In the sensitivity analysis, the values for these parameters including timetable, alighting rate and routing probability are the same as those in the case study. We consider the changeable parameters for each node type for all test instances in terms of related research and specification (Mitchell and Smith, 2001; Kittelson & Associates et al., 2003; GB, 2003). The solutions in terms of changeable parameters and their base values, which is the information needed for the sensitivity analysis, are provided in Table 4.

Sensitivity analysis results are provided in Table 5. The results indicate that a unit change in the parameter of node 8, 12 and node 14 affects SSC much more than a unit change in the parameters of other nodes. In terms of a satisfactory rate of remaining passengers, change in the capacity parameter of node 12 affects the changes of SSC much more than node 8 or 14 does. If the decision to be made is to increase the capacity of one node but not all, then the expansion should be done at the node 12. However, node 8 will be selected due to practical and economic reasons.

Notice that an increase of one unit capacity parameter in node 1or 4 will decrease the overall SSC to 29,657. According to Yang and Bell (1998), the decrease value indicates a capacity paradoxical behavior. That is, adding capacity to node 1 or 4 would actually reduce the potential capacity of the station network. If the purpose of capacity expansion is to accommodate future travel demand, selection of node 1 or 4 for improvement would be detrimental.

6. Conclusions

In this paper, a definition of SSC of the subway station is presented according to the gathering and scattering process. An analytical queuing network model for subway station including two subsystems is built, of which the first subsystem is built by M/G/C/C state dependent queuing network and the second is constructed by the probability theory. The proposed queuing network, integrating IGEM and DTMC, allows to master passengers’ state at station from a macro-point of view. A SSC optimization model is developed based upon the definition and the queuing network, and a solution to this model is also proposed according to RSM integrating with DTMC and IGEM. Lastly, a case study of Beijing Station in Beijing subway line 2
shows that the proposed model and solution are capable of calculating SSC and then sensitivity analysis for identifying bottleneck of SSC is done.

Main contribution of the study is to develop an analytical queuing network model for station in a macro-point of view, to solve the problem of how to compute the number of gathering and scattering passengers with a certain rate of remaining passengers, and to offer some directions for subway managers to find out optimum solutions by using an integration of IGEM, DTMC and RSM.

There are a number of directions possible with this research. For example, a more general capacity calculation framework may be considered, in which the number of decision variables is larger than those tested in this article. More work can be done to enhance capacity bottleneck according to the queuing network model, especially in subway transfer station.

Acknowledgments

This work is financially supported by the National Key Technologies R&D Program of China (2011BAG01B01) and the State Key Lab of Rail Traffic Control and Safety, China (RCS2012ZZ002). The video data in the subway station is provided by the State Key Lab of Rail Traffic Control and Safety, Master. Zhao Jun-Duo. The authors acknowledge with gratitude the suggestions by anonymous referees that helped to improve the presentation of the paper.

Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.trc.2013.10.010.

References


