Multi-objective optimization of cordon sanitaire with vehicle waiting time constraint

Hongzhi Lin

School of Economics and Management, Southeast University, Nanjing 211189, China

Correspondence
Hongzhi Lin, School of Economics and Management, Southeast University, Nanjing 211189, China.
Email: linhz@seu.edu.cn

Funding information
Natural Science Foundation of Jiangsu Province, Grant/Award Number: BK20201277; Fundamental Research Funds for the Central Universities, Grant/Award Number: 2242021820022

Abstract
The outbreak of COVID-19 has disrupted our regular life. Many local authorities have enforced a cordon sanitaire for the protection of sensitive areas. Travelers can only travel across the cordon after getting qualified. This paper aims to propose a method to determine the optimal deployment of cordon sanitaire in terms of the number of parallel checkpoints at each entry link. A bi-level multi-objective programming model is formulated where the lower-level is the transportation system equilibrium with queueing to predict link traffic flow, and the upper-level is queueing network optimization that is a multi-objective integer non-linear programming. The primary objective is to minimize the total operation cost of checkpoints with a predetermined waiting time constraint and the secondary objective is to minimize system total travel time. A heuristic algorithm is designed to solve the proposed bi-level model where the method of successive averages is adopted for the lower-level model, and a hybrid genetic algorithm is designed for the upper-level model. An experimental study is conducted to demonstrate the effectiveness of the proposed methods. The results show that the methods can find a good heuristic optimal solution. These methods are useful for operators to determine the optimal deployment of cordon sanitaire.

1 | INTRODUCTION

The COVID-19 pandemic is an ongoing pandemic of the coronavirus disease caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2). As of 29 September 2020, over 33.3 million cases of COVID-19 have been reported in more than 188 countries, resulting in more than 1 million deaths.

To contain the rapid spread of COVID-19, authorities worldwide have responded by implementing travel restrictions, facility closures, and lockdowns while increasing the testing capacity and tracing contacts of infected people. The first cordon sanitaire was set up on 23 January 2020, to control entry to and exit from the City of Wuhan, China, known as the Wuhan lockdown, and then it is extended to almost all cities in China. As the outbreak expanded, many cities in other countries enacted similar restrictions. Cordon sanitaire is used to restrict the movement of people and freights into and/or out of a defined geographic area, such as a community, city, or region. The term denoted a barrier used to stop the spread of infectious diseases.

The enforced cordon sanitaire has been demonstrated to be an effective way to prevent the infectious virus from spreading into a protected area [1–3]. All travellers crossing the cordons are imposed to test body temperature to ensure they are not infected. However, it is reported that the queue length is too long, and the waiting cost is too high at the cordon sanitaire. Therefore, there is an urgent need to optimize the queueing system to improve the service level of testing. This paper aims to propose a method to deploy checkpoints at the cordon sanitaire to ensure the maximum waiting time. It is not only helpful to control and prevent epidemics but also beneficial to check other kinds of dangers such as drunk driving, terrorists, criminals, smuggling etc.

The optimal deployment of cordon sanitaire, including the location and the number of checkpoints, has not been investigated yet. However, an analogous problem, namely cordon pricing, has been explored a lot. Cordon pricing is a toll paid by private vehicles to enter a restricted area, usually within a city centre, as part of a travel demand management strategy.
to relieve traffic congestion within that area. It has been successfully implemented in some cities in the world, for example, London, Stockholm, and Singapore. Earlier studies have demonstrated that the performance of cordon schemes is critically dependent on toll location and toll levels. Therefore, the achievement of cordon pricing could inspire the research of cordon sanitaire.

Both cordon pricing and cordon sanitaire operate at the entry links around a restricted area. The difference is that cordon pricing aims to determine the optimal toll levels while cordon sanitaire seeks to determine the optimal number of checkpoints. In general, queue length on a road network is highly dependent on toll levels and toll points. It is usually assumed that, in cordon pricing, the operator sets tolls to maximize social welfare, defined as total benefits minus total costs according to the Marshallian measure. The problem has been well studied by Verhoef [4], May et al. [5], Mun et al. [6], Santos [7], Shepherd and Sumalee [8], Zhang and Yang [9], and Yang et al. [10] in the early 21st century.

The principle of social welfare maximization concerns the system efficiency that tends to harm equity. The service level of some travellers could be deteriorated significantly, which could cause cordon schemes to be inaccessible. In the past, researchers have explored the equity issues related to charging schemes. To identify the optimal toll location and toll levels, Sumalee et al. [11], as well as Maruyama and Sumalee (2007), developed a methodology based on genetic algorithms to achieve social welfare maximization with and without constraints on the desired level of revenue and spatial equity impact. Abubidbeh et al. [12] used origin-destination data to assess the vertical equity effects of a hypothetical cordon pricing scheme in Canada’s largest city. Souche et al. [13] and Souche et al. [14] simulated the impact of cordon pricing on equity using both social and spatial indicators (Gini, Theil, and Atkinson indices) to make a city sustainable. They concluded that introducing a toll would increase inequalities, and it was crucial to investigate any changes among different categories of network users. Afandizadeh and Abdolmanafi [15] examined the issue of environmental equity in cordon pricing since air pollutants may transfer from inside to outside the cordon. Camporeale et al. [16] explored the distributional effects using the Theil index, considering an elastic demand associated with the cordon pricing strategy. A bi-level optimization model with an equity constraint is usually developed to reduce all kinds of inequities.

Besides the equity issues, environmental considerations are often included for multi-objective optimization. Amigholy et al. [17] presented an equilibrium model to estimate the long-run adverse traffic effects of a cordon pricing scheme on consumer welfare and air pollution. Günthermann et al. [18] analysed the application of cordon pricing for the reduction of local air pollution in addition to social welfare maximization. Li et al. [19] focused on an environment-friendly cordon pricing design problem, where an acceptable road network performance is promised.

Cordon pricing is still a hot topic in transportation research, and there are several directions for future research. Firstly, a logsum-based welfare computation before and after a change in the travel environment could be used as an essential policy evaluation measure [20, 21]. Given that various logit choice models are widely adopted to simulate the response of consumers to a policy, consumer surplus can be easily computed using the model’s logsum. Therefore, utilizing a logsum-based objective function in the design of a cordon pricing scheme would likely be feasible for most planning agencies. Secondly, stochastic user equilibrium with elastic demand could be used to represent route choice behaviour [22–24]. It is not surprising that the stochastic user equilibrium is more realistic than deterministic user equilibrium. However, it is a challenge in computation. Thirdly, mixed traffic flow is considered in a cordon tolling model, including motorcycle and automobile trips [15, 25, 26]. Lastly, combined models are proposed to address cordon pricing and other decisions, such as road capacity expansion choice [27] and land-use regulation [28].

Although cordon sanitaire is analogous to cordon pricing, it has different characteristics that deserve to be studied further. The contribution of this paper lies in four aspects. Firstly, the queueing theory is used to quantify the phenomenon of waiting at the sanitary cordon, but it is rarely applied to the pricing cordon. With electronic toll collection (ETC) being widely implemented, there is no need for the inflow vehicles to queue to pay tolls. Secondly, a stochastic model is used to obtain representative measures of waiting at the sanitary cordon, such as average queue length, average waiting time in queue etc. From the standpoint of analysing queues, the arrival of vehicles is represented by the inter-arrival time (time between successive arrivals), and the testing service is measured by the service time per vehicle. The inter-arrival and service times are assumed to be probabilistic due to various disturbance factors, such as traffic accidents, bad weather, and occupancy rate. Thirdly, the queueing network is considered at the sanitary cordon other than a single queueing system. The testing is conducted at all the entry links of one cordon around a city such that there could be many queues. Therefore, it is necessary to study the entire queueing network. Lastly, multi-objective optimization is conducted where the primary objective is to minimize operation investment with an acceptable level of queueing delay and the secondary objective is to minimize system total travel time.

The structure of this paper is organized as follows. Section 2 proposes a bi-level multi-objective programming model, where the lower-level is transportation system equilibrium with queueing, and the upper-level is a multi-objective network optimization. An explicit heuristic algorithm is designed in Section 3 to solve the proposed model. Section 4 demonstrates the effectiveness of the model and algorithm through an experimental study. Section 5 concludes this paper.

## 2 | METHODOLOGY

This paper aims to propose a method and an algorithm to design a cordon sanitaire in terms of checkpoint deployment at each entry link around a geographical area. It is a Stackelberg game with a leader–follower decision structure, which is usually formulated as a bi-level programming model. The operator
in the upper-level aims to minimize the operation cost of cordon sanitaire with queueing delay constraints and minimize system total travel time meanwhile. It is a multi-objective decision. The operator can predict, but cannot control network users’ travel behaviours including destination choice and route choice, while all users make their decisions in a user optimal manner. The users’ decisions in the lower-level are made after the upper-level decisions. However, the operator must take the behavioural responses of the users into consideration to adjust decisions. The conceptual framework is shown in Figure 1. It can explicitly capture the leader–follower nature of the relationship between the operator and users. Note that the lower-level is a feedback procedure between trip distribution and traffic assignment with queueing. It is usually termed as transportation system equilibrium. The detailed models are elaborated in the followed subsections.

2.1 Network equilibrium with queueing

The lower-level model is a transportation system equilibrium that combines trip distribution and traffic assignment models with fixed travel demand and a given road network. It has long been criticized that travel times are inconsistent in the conventional four-step sequential model because travel times are determined endogenously in fact. Generally speaking, two ways can be used to solve the inconsistent problem to achieve transportation system equilibrium in literature. One way is to combine several steps to an equivalent mathematical programming, which can be proved to be a well-converged and consistent result [29, 30]. The other is to feedback the sequential models iteratively until travel times meet the consistency criteria [31, 32]. Although the former is commonly adopted in literature, the latter one is more flexible at each step [33, 34]. Therefore, a combined model with feedback is adopted here.

Note that traffic assignment is not a traditional one here as queueing delay at cordon sanitaire is accommodated. The determination of queueing delay time is a critical problem. Generally, queueing theory is an excellent tool to analyse the cost of waiting experienced by vehicles. In most traffic situations, inter-arrival and service times are described randomly by the exponential distribution. This stage adopts a stochastic queueing model that combines both arrivals and departures based on the Poisson assumptions. That is, the inter-arrival and the service times follow the exponential distribution. The derivation of the specialized queueing model is based on the steady-state behaviour of the queueing situation, achieved after the system has been in operation for a sufficiently long time.

According to the conventional traffic flow theory [35], the waiting line at each freeway toll booth can be formulated as a fundamental $M/M/c$ queueing model, where $M$ means Markovian (or Poisson) arrivals or departures distribution or equivalently exponential inter-arrival or service time distribution, and $c$ means the number of identical parallel servers with same service rate per unit time. There could be one or more parallel checkpoints (i.e. servers) at each entry link. Suppose that there are $m$ entrances at a sanitary cordon. It is necessary to study the entire queueing network performances. Assume that vehicles arrive at each entrance $i$ ($i = 1, 2, \ldots, m$) according to a Poisson process with predicted inflow $\lambda_i$ and that each entrance $i$ has an exponential service time distribution with an identical parameter $\mu$ for its $c_i$ parallel checkpoints, where $c_i \mu > \lambda_i$. Therefore, the elementary $M/M/c$ queueing model can be used to analyse each entrance.

Analogous to a single service facility, the most commonly used measures of queueing situation in a given entrance $i$ are the expected number of vehicles in the queue ($l_i$) and expected waiting time in the queue ($d_i$). The relationship between $l_i$ and $d_i$ is known as Little’s formula, and it is given as $l_i = \frac{\lambda_i}{\mu}$, the expression $l_i$ can be determined as follows:

$$l_i = \frac{\rho_i^{c_i+1}}{(c_i - 1)!c_i^{\rho_i} \rho_i \mu}, \forall i \in A^*$$  \hspace{1cm} (1)

$$p_i = \frac{\rho_i^{c_i+1}}{c_i!} + \frac{\rho_i^{c_i+1}}{(c_i - 1)!c_i^{\rho_i} \rho_i \mu}, \forall i \in A^*$$  \hspace{1cm} (2)

$$\frac{\rho_i}{c_i} < 1, \forall i \in A^*$$  \hspace{1cm} (3)

where $p_i$ is the steady-state probability of none customers in an entrance $i$. $A^*$ is the set of entry links, Equation (3) is a steady-state condition. The queueing delay $d_i$ is determined through dividing $l_i$ by $\lambda_i$ according to Little’s formula. It can be formulated in detail as follows:
FIGURE 2  The iterative process at the lower-level model

\[
q_{rs} = O_r \frac{\exp(\beta_i t_{rs})}{\sum_{s' \in S_r} \exp(\beta_i t_{s's})} \quad \min Z(x) = \sum_{rs} \int_{t_{rs}}^{t_{rs}^*} t_s(x, c_s) \exp(-\rho t_s) \, dx \\
\sum_k f^m_k = q_{rs}, \forall k, rs \\
s.t., \quad v_a = \sum_{rs} f^m_k \delta_{rs}^k, \forall a \\
f^m_k \geq 0, \forall k, rs
\]

where these notations are defined as follows:
- \( q_{rs} \): travel demand between origin \( r \) and destination \( s \);
- \( O_r \): travel demand at zone \( r \);
- \( S_r \): the set of destinations departed from origin \( r \);
- \( \beta_i \): traveller preference for destination \( i \);
- \( t_{rs} \): path travel time between origin \( r \) and destination \( s \);
- \( \rho \): the set of entry links that is a subset of all links \( A \);
- \( \beta_i^* \): coefficient of travel time \( t_{rs} \);
- \( \gamma \): traffic flow at link \( \gamma \);
- \( d_{\delta_i} \): number of checkpoints at entry link \( \delta_i \);
- \( t_{\delta_i} \): travel time at link \( \delta_i \) which is a function of traffic flow \( v_{\delta_i} \) and number of checkpoints \( \gamma \);
- \( f^m_k \): traffic flow on path \( k \) connecting origin \( r \) and destination \( s \);
- \( \delta_{rs}^k \): link-path incidence relationship which is expressed as:
- \( \delta_{rs}^k = \{1, \text{if link } \delta_i \text{ is on path } \gamma \}; 0, \text{if not} \).

2.2  Bi-level multi-objective optimization model

There is a transportation system equilibrium at the lower level for a given checkpoint deployment decision at the upper level. The results of queueing can be incorporated in a cost minimization mode that seeks to minimize the sum of the cost of offering checkpoints and the cost of queueing time. It is straightforward that the cost of service increases with the increase in the level of service (e.g. the number of checkpoints). At the same time, the waiting time decreases with the increase in the level of service. The cost-based model attempts to balance two conflicting costs: the cost of offering the service and vehicle waiting time. An increase in one cost automatically causes a decrease in the other. As the cost of waiting time is difficult to be determined in dollars, the waiting time is adopted as a constraint while the cost of offering checkpoints is used as a primary objective. Therefore, the primary objective function in the upper-level programming model can be expressed as

\[
\text{Min } E = \sum_{j=1}^m \xi_j \quad (5)
\]

where \( E \) is the expected cost of operating the facilities per unit time and \( \xi_j \) is the marginal operation cost per checkpoint per unit time. The expected total cost \( E \) in the entire system
is obtained by merely summing the corresponding quantities obtained at the respective entries.

The objective of operator could be multiple. Besides the total operation cost of checkpoints, the operator usually consider some other aspects, for example, system total travel time, system total queuing time, queuing equity etc. The system total travel time is adopted for secondary objective here for demonstration. Therefore, the secondary objective function in the upper-level programming model can be formulated as:

$$\text{Min} F = \sum_{d \in D} t_d(p_d, c_d)p_d$$ (6)

Acceptable control should consider queuing time constraint. The aspiration level model works directly with the performance measures of the queueing situation. The idea is to determine an acceptable range for the service level by specifying reasonable limits on measures of performance. Such limits are the aspiration levels the decision-maker wishes to reach. Note that the service level in a given entrance \( i \) is a function of the number of parallel checkpoints, \( c_i \). The upper-level presents a decision model to determine \( c_i \) for acceptable service levels in terms of waiting time \( d_i \). The model recognizes that higher service levels reduce the waiting time in the system. The goal is to strike a balance between operation cost and waiting time. The problem reduces to determining the number of servers \( c_i \) with identical operation cost such that

$$d_i \leq T, \forall i \in A^*$$

The constant \( T \) is the level of aspiration specified by the decision-maker, for example, \( T = 5 \) min. Note that \( d_i \) is a function of \( c_i \). According to Equation (4), the aspiration level of average waiting time \( d_i \) can be specified as

$$d_i(c_i) = \frac{\rho_i^{c_i+1}}{\lambda_i(c_i - 1)[\rho_i - \rho_i^c]} \sum_{n=0}^{c_i-1} \frac{\rho_i^n}{n!} + \frac{\rho_i^c}{(c_i - 1)[\rho_i - \rho_i^c]} \leq T, \forall i \in A^*$$ (7)

In conclusion, an integer non-linear programming model can be proposed where the primary objective is to minimize the expected total cost of testing station and the secondary objective is to minimize the system total travel time, the constraint is an aspiration level of the vehicle waiting time at each entrance, and the decision variables are the number of parallel checkpoints in each entrance. The upper-level programming model is formulated as:

$$\text{Min} E = \sum_{i=1}^{n} \lambda_i p^{c_i}$$ (8)

$$\text{Min} F = \sum_{d \in D} t_d(p_d, c_d)p_d$$ (9)

where traffic flows \( \lambda_i, \forall i \in A^* \) are determined by the lower-level model, i.e. network equilibrium model with queueing. Equation (8) is the primary objective and Equation (9) is the secondary objective. Equation (10) is the aspiration level of the vehicle waiting time for each entrance where \( T \) is a constant determined by the policymaker. Equation (11) is the steady-state condition. Equation (12) enforces that the number of checkpoints \( c_i \) is not more than the capacity \( c_i^* \) for each entrance. Equation (13) makes sure that the decision variables are non-negative integers. It should be noted that the multi-objective integer non-linear programming model is hard to solve so that a heuristic algorithm is designed in the followed section.

3 | SOLUTION ALGORITHM

3.1 | Equilibrium algorithm with queueing

To solve the proposed bi-level programming model, it is always beneficial to solve the lower-level model first as it is embedded in the upper-level model. With a fixed travel demand and a built road network, there will be a stable flow pattern in the lower-level for a given decision from the upper-level. Note that the lower-level is a feedback procedure between trip distribution and traffic assignment with queueing. The method of successive averages (MSA) can be used to achieve system equilibrium. An initial trip distribution matrix can be produced by a multinomial logit model with initialized origin-destination (OD) pair travel times. The trips are then assigned the road network by the state-of-the-art Reduced Gradient (RG) algorithm [36] which is demonstrated to be more efficient than the conventional Frank–Wolfe algorithm. The link travel flows and link travel times can be generated. In addition, queueing delay times can be determined with predicted traffic flows at entry links. The generalized travel time of each entry link includes link travel time and queuing delay time. According to Wardrop’s first principle of route choice, also known as user equilibrium, traffic arranges itself in congested networks such that all used paths between an OD pair have an equal and minimum cost. Therefore, Dijkstra’s algorithm for the shortest path problem is used.
to update OD pair travel times. These times are then fed back to the multinomial logit model to generate a new trip distribution matrix. However, this matrix cannot be assigned to the road network directly. The convergence of direct or naive feedback is usually impossible. An averaging of successive trip distribution matrix is necessary. Although there are some successful applications of constant weights, the convergence is usually not guaranteed. Therefore, the MSA with decreasing weight is used here to update the trip distribution matrix, which is the reciprocal of the iteration number. The updated matrix is further assigned to the road network. The iteration process continues until the successive matrices are quasi-equal. The convergence is generally measured by the squared root of the relative gap. If a predetermined tolerance is achieved, terminate the iteration. The stable state is termed as the transportation system equilibrium with queueing. The resultant traffic flows at all links then go into the upper-level model. Figure 3 shows the flowchart of the equilibrium algorithm with queueing.
The detailed MSA algorithm is specified step by step as follows:

Step 1: Input a checkpoint deployment decision with fixed travel demand and built road network.
Step 2: Initialize trip distribution matrix \( q^0_{rs} \) with initial OD pair travel time \( t^0_{rs} \). Besides, let \( n = 1 \) be the number of iterations.
Step 3: Traffic assignment with queueing. The trip distribution matrix \( q^1_{rs} \) is assigned to the road network by the efficient Reduced Gradient algorithm. The link travel flows \( v_r \) and generalized link travel times \( t^1_r \) are generated. Note that queuing delay times can be determined with flows at entry links.
Step 4: Update the shortest path travel time between an OD pair \( rs \), namely \( t^1_{rs} \), by Dijkstra’s algorithm.
Step 5: Trip distribution. The multinomial logit model is used to update the trip distribution matrix \( q^1_{rs} \):

\[
q^1_{rs} = \frac{O_r \exp(\beta_r + \beta_s t^1_{rs})}{\sum_{s' \in s} \exp(\beta_r + \beta_s t^1_{rs})} \quad (14)
\]

Step 6: Average trip distribution matrices \( q^1_{rs} \) and \( q^0_{rs} \) using decreasing weight

\[
q^1_{rs} = q^0_{rs} + \frac{1}{n} (q^1_{rs} - q^0_{rs}) \quad (15)
\]

Step 7: Convergence identification. Check the convergence of trip distribution matrix using the squared root of the relative gap:

\[
\sqrt{\sum_{rs} \left( \frac{q^1_{rs} - q^0_{rs}}{q^0_{rs}} \right)^2} < \varepsilon \quad (16)
\]

where \( \varepsilon \) is a predetermined tolerance. If the convergence condition is satisfied, terminate the iteration and turn to Step 9, otherwise turn to Step 8.

Step 8: Let \( q^0_{rs} := q^1_{rs} \) and \( n := n + 1 \). Then turn to Step 3.
Step 9: The outputs are the trip distribution matrix \( q^1_{rs} \) and the link traffic flow \( v_r \).

### 3.2 Hybrid genetic algorithm

The bi-level programming problem is a well-known NP-hard problem that is hard to be solved by classical optimization algorithms. It is challenging even the upper-level and lower-level are both linear programming, let alone the upper-level is a multi-objective non-linear integer programming model. The traditional gradient-based approaches to solve the optimal cordon toll problem usually fail to converge for larger-scale problems due to multiple optima. This failure led to the development of a heuristic algorithm to determine the optimal toll level and toll location problem. The heuristic algorithm was shown to be successful in solving the cordon toll optimization problem, although it is found to be time-consuming, and there is no proof of global optimum. However, the successful applications of heuristic methods, especially genetic algorithms, have been growing to generate high-quality cordon schemes in the literature [8, 22, 37, 38].

Therefore, a hybrid genetic algorithm is proposed here. It is a combination of genetic algorithm and lexicographic method. The lexicographic method is an effective method for multi-objective optimization problems. It assumes that the objectives can be ranked in the order of importance. It solves a sequence of single-objective optimization problems. It first finds the optimal solutions of the most important objective, and then finds the optimal solutions of the next objective in turn on the premise of ensuring the optimal solution of the previous objective, until the last objective. There are two ranked objectives in the upper-level optimization model. Therefore, the primary objective function is solved first by genetic algorithm with elite strategy. The best chromosomes are kept and the worst chromosomes are eliminated. The population evolves generation by generation. The best chromosomes are kept until the last generation. There could be some optimal solutions according to the primary objective at the last generation. Then the performances of these optimal solutions are evaluated further in terms of the secondary objective. The optimal one can be determined as the solution of the bi-level multi-objective optimization problem. Figure 4 shows its flowchart.

To be more specific, the detailed hybrid genetic algorithm is specified in steps as follows:

Step 1: Initialization. Set the parameters used in the genetic algorithm, including population size \( M \), the maximum number of generations \( Gen \), crossover probability \( p_c \), mutation probability \( p_m \), the notation of generation \( gen = 1 \), the portion for elitist strategy \( p_e \). Note that the population size depends on the nature of the problem, but typically contains hundreds of possible solutions.
Step 2: Generate a feasible initial population randomly. A chromosome is a solution that consists of \( m \) genes. Integer encoding technology is used where a gene stands for the number of parallel checkpoints at an entry link. Generate a chromosome randomly. If it is not feasible, generate another one until it is feasible. A total number of \( M \) viable chromosomes are generated, scattering the entire range of possible solutions.
Step 3: Selection operation. The primary objective function of the upper-level model is used to work as a fitness function to evaluate the performance of all chromosomes in the population. Note that it is to minimize the total number of checkpoints, the best \( p_e \) are labelled for elitists, and the worst \( p_e \) is discarded.
Step 4: Crossover operation. The remaining \((1 - p_e)M\) chromosomes are used for crossover operation. These parent chromosomes are matched in pairs randomly. The probability of carrying out the crossover is \( p_c \). If it is chosen for the crossover, a random gene is identified.
If new-born chromosomes are not feasible according to constraints in the upper-level model, try another gene location until they are feasible. These new solutions typically share many of the characteristics of their parents.

Step 5: Mutation operation. The probability of carrying out mutation is $p_m$. A random gene is identified for mutation within the domain of definition. If the new chromosome is not feasible, try another gene location until it is feasible.

Step 6: Generate the next generation population. After genetic operators, there are still $(1 - p_e)M$ feasible chromosomes. The labelled $p_eM$ elitists are added to ensure the population size $M$. This allows the best chromosomes from the current generation to carry over the
next unaltered. It guarantees that the solution quality will not decrease from one generation to the next. Let the notation of generation be \( \text{gen} := \text{gen} + 1 \).

Step 7: Termination judgment. If the maximum number of generations is achieved, that is \( \text{gen} \geq \text{Gen} \), terminate the iteration process and output the optimal solutions. Otherwise, turn to Step 3.

Step 8: Lexicographic method. Calculate the performances of the optimal solutions in terms of the secondary solution. The best one is determined as the final optimal checkpoint deployment scheme.

4 | EXPERIMENTAL STUDY

An experimental study is conducted to verify the effectiveness of the proposed method and algorithm. The Nguyen–Dupuis road network, as shown in Figure 5, is commonly used in transportation research to demonstrate various methods. The link characteristics, including free-flow travel time, link capacity, and link length, are shown in Table 1. Note that the road characteristics are set to be as a freeway network as the cordon sanitaire is usually set around a city where freeway is used. For example, the link capacity is set as 2000 pcu/h and the free flow speed is about 70 km/h.

There are two origins and two destinations in the Nguyen–Dupuis network. The predicted travel demands at origin zones 1 and 4 are 1500 and 1000 pcu/h, respectively. That is, \( O_1 = 1500 \) pcu/h and \( O_4 = 1000 \) pcu/h. All the existing links are labelled
from 1 to 19. The destination zones are 2 and 3 so that the entry links are 11, 15, 16, 19. The problem is to determine the number of checkpoints at each entry link to minimize the testing cost with a certain level of service and to minimize the system total travel time.

The parameters used in the lower-level model are summarized as follows. The multinomial logit model for destination choices is simplified as

$$q^t_s = O\frac{\exp(\beta_s + \beta_t t^t_s)}{\sum_{s \in V} \exp(\beta_s + \beta_t t^t_s)}$$

where $\beta_s$ is traveller preference on destination $s$ and $\beta_t$ is the coefficient of path travel time between O-D pair $rs$. The values of $\beta_s$ and $\beta_t$ can be calibrated empirically. Here we set $\beta_2 = 0.5$, $\beta_3 = 0$, and $\beta_t = -0.1$. That is, the traveller preference on destination zone 2 is 0.5, and on destination zone 3 is 0, which means that the travellers traditionally prefer destination 2. The coefficient of travel time is $-0.1$, which means that the travel time is a negative utility. Besides, a traditional link impedance function, named BPR function, is used to accommodate the congestion effect in traffic assignment with the following formulation:

$$t_a(v_a) = t_a^0\left[1 + \alpha\left(\frac{v_a}{v_a^c}\right)^{\beta}\right], a \in A$$

where $t_a^0$ is the free-flow travel time of link $a$; $v_a$ is traffic volume of link $a$ and $v_a^c$ is traffic capacity of link $a$; $\alpha$ and $\beta$ are volume/delay coefficients which can be calibrated empirically, they are usually set as $\alpha = 0.15$ and $\beta = 4$ conventionally. The convergence criteria for MSA is set as $\varepsilon = 0.05$. A stable transportation system can be achieved for a built road network.

The parameters used in the upper-level model are listed as follows. They depend on the scale of the experimental network. The population size is $M = 300$. The maximum number of generations is $Gen = 10$. The portion for elitists is $p_c = 0.1$. The crossover probability is $p_c = 0.5$, and the mutation probability is $p_m = 0.1$. Although these parameters are conventionally used in genetic algorithms, it is worth tuning parameters to find appropriate settings for the problem. The allowed maximum waiting time at each entry link is set as $T = 5$ min. The maximum number of checkpoints can be set up at each entry link is assumed to be $l^t_e = 9, \forall e \in A^r$. This is just for simplicity, and they can be diverse. It is reported that the cordon sanitaire is usually located at highway toll stations with plaza. The number of checkpoints could be more than the number of corresponding lanes. The upper bound of number of checkpoints 9 is reasonable just as toll booths. The marginal cost per checkpoint per unit time $\mu$ is set to be one without affecting the decision.

The calculation is programmed using a popular open-source language R 3.6.3 in a personal computer with Intel Core i7-4790 CPU @ 3.60 GHz. The running time is 43.5 min. The predicted traffic volumes at entry links work as the average arrival rates of the queuing model. Note that traffic volumes are usually measured in hours while the arrival rates are typically measured in minutes. Therefore, unit conversion is needed. The average service rate for a single checkpoint is assumed to be $\mu = 2$ pcu/min. That is, the checkpoint averages tests two passenger car units per minute. The optimal solutions for primary objective are shown in Table 2.

It is shown that there are no less than 5 optimal solutions in terms of the primary objective. They have the same total number of checkpoints 22. The problem is which one to take. According to lexicographic method, the performance of the secondary objective is assessed which is the system total travel time. It can be found that the system total travel time, including queuing delay time, is diverse. The checkpoint deployment does have impact on transportation system performance. The minimum one is 89,378 min and the checkpoint deployment is 9, 1, 7, 5 for links 11, 15, 16, 19 respectively. The queuing performance at this scheme is shown in Table 3.

It is shown that traffic inflow volumes vary considerably at entry links. The maximum one is 1067 pcu/h at entry link 11, while the minimum one is 60 pcu/h at entry link 15. As a result, the number of checkpoints needed at each entry link will be various too. The maximum number is 9 at links 11, and the minimum one is 1 at links 15. The total number of checkpoints needed is 22, which is the minimum cost to maintain a certain level of waiting time. In this way, the maximum waiting time at each entry link will not exceed 5 minutes. However, attention should be paid to the issue of choosing the best parameters for the GA based methods, i.e. generation number, population number, the probability of crossover, and the probability of mutation.

Although a simple network is adopted for demonstration, the proposed methods could be expected to work well for large-scale networks. The traffic assignment problem (TAP) is the heart of the proposed algorithm because it is required to be solved so many times to predict the flow pattern. In fact, reducing the running time of TAP algorithm is the key to speed-up the computational efficiency. The conventional Frank–Wolfe (FW) algorithm is a kind of link-based algorithms and has been used in most of software packages because of its simple implementation. However, it cannot gain precise solutions and has a slow convergence rate, especially for large-scale networks. Its poor performance is attributed to the severe zigzag movement that occurs close to the optimal solution. Fortunately, recent advances in computer science and advent of random-access memories (RAM) with larger storage capacities have opened the way for implementing path-based algorithms on large-scale networks. The path-based algorithms are demonstrated to be very effective and can extremely accelerate the convergence speed of traffic assignment. Many researchers have devoted to enhance the path-based algorithms. The Reduced Gradient (RG) algorithm for user equilibrium TAP is the state-of-the-art path-based algorithm. It is claimed to be superior than -scale networks. Therefore, the RG algorithm is adopted here for traffic assignment module. The performance of RG algorithm is verified in the Philadelphia and Chicago test networks with about 40 thousand links. It can reach relative gap of 1.0$^{-4}$ in 2 min while FW algorithm will take 50 min. It can be expected that the proposed algorithm will work well in large-scale networks even the traffic assignment is called for hundreds of times.
TABLE 2 The optimal solutions for primary objective

| Scheme | Link 11 | Link 15 | Link 16 | Link 19 | Total number of checkpoints | System total travel time (min) |
|--------|---------|---------|---------|---------|-----------------------------|------------------------------|
| 1      | 8       | 3       | 5       | 6       | 22                          | 89,772                       |
| 2      | 9       | 5       | 3       | 5       | 22                          | 90,197                       |
| 3      | 9       | 1       | 7       | 5       | 22                          | 89,378                       |
| 4      | 9       | 5       | 4       | 4       | 22                          | 90,666                       |
| 5      | 9       | 5       | 2       | 6       | 22                          | 89,753                       |

TABLE 3 The predicted traffic inflow volume and queueing performance

| Entry link | Traffic flow volume (pcu/h) | Number of checkpoints | Steady condition | Queueing delay (min) |
|------------|-----------------------------|-----------------------|------------------|---------------------|
| 11         | 1067                        | 9                     | 0.988            | 4.86                |
| 15         | 60                          | 1                     | 0.500            | 1.00                |
| 16         | 787                         | 7                     | 0.937            | 1.43                |
| 19         | 586                         | 5                     | 0.976            | 4.48                |

5 | CONCLUSION

The COVID-19 pandemic has caused global social and economic disruption. Many authorities set up a cordon sanitaire around an area with high population density or an area experiencing an epidemic to prevent the spreading of infectious disease. Once the cordon is established, people are no longer allowed to enter or leave the area without testing. In the most extreme form, the cordon is not lifted until the infection is extinguished. It is much more efficient to keep the disease from being introduced from the very beginning. Public health specialists have included cordon sanitaire along with quarantine and medical isolation as “non-pharmaceutical interventions” designed to prevent the transmission of microbial pathogens, especially COVID-19.

The line around a cordon sanitaire is usually set up at each entry link. The cordon sanitaire is demonstrated to be an effective way to battle the affliction. However, some severe problems turn up, such as the long waiting time for testing and the massive investment for checkpoints. This is a resource allocation problem in principle that deserves to be investigated systematically. Although the cordon pricing problem is analogous to this problem, it has different characteristics and it has not been explored yet in literature.

This paper proposed a method to determine the optimal deployment of checkpoints at each entry link around the cordon sanitaire with an acceptable queueing delay time constraint. A bi-level multi-objective programming model is formulated where the lower-level model is transportation system equilibrium with queueing, and the upper-level model is a multi-objective queueing network optimization. To be more specific, the lower-level is a feedback procedure between trip distribution and traffic assignment with queueing. The iteration continues until the transportation system equilibrium is achieved. Note that a multinomial logit model is used for trip distribution, and the user equilibrium model is used for traffic assignment with queueing. The traffic flow and waiting time at each entry link can be predicted. The upper-level is queueing network optimization based on the predicted traffic flow and waiting time from the lower-level. The traffic inflow at each entry link is regarded as an average arrival rate in the queueing model. Note that there could be many entry links, and the $M/M/c$ queueing model is used to analyse each entry independently of the others. A multi-objective integer non-linear programming model is built where the primary objective is to minimize the total operation cost of checkpoints, the secondary objective is to minimize the system total travel time, and the constraint is to make the maximum waiting time within a predetermined service level.

A heuristic algorithm is proposed to solve the bi-level multi-objective optimization model. The method of successive averages (MSA) is used to achieve a transportation system equilibrium with queueing at the lower-level model. Note that the state-of-the-art Reduced Gradient algorithm is used for traffic assignment with queueing. Based on the predicted traffic flow and queueing delay time at each entry link from the lower-level model, a hybrid genetic algorithm is designed to solve the multi-objective integer non-linear programming model at the upper-level. It is a combination of genetic algorithm with elite strategy and lexicographic method. The proposed algorithm with efficient traffic assignment module is expected to work well for large-scale networks.

An experimental study is conducted, which is the well-known Nguyen–Dupuis road network, to demonstrate the effectiveness of the proposed method and algorithm. The results show that the proposed methods can find at least a satisfying optimal heuristic solution. However, it is worth tuning parameters to find appropriate settings for the problem.

The limitation of this research lies in the simplified network, i.e. the Nguyen–Dupuis road network. Without a doubt, a large-scale network with real data will be more convincing. However, the models and algorithms explored in this research are demonstrated to be valid and ready to be applied. In future research, the authors would like to study further a variety of related problems, including the queueing equity problem with a given investment budget, the on-ramp traffic control problem, and other multi-objective optimization problems.
ACKNOWLEDGMENTS
This work was supported by the Natural Science Foundation of Jiangsu Province (Grant No. BK20201277) and the Fundamental Research Funds for the Central Universities (Grant No. 2242021820022). The efforts and comments from editors and anonymous reviewers are significantly appreciated.

REFERENCES
1. Goslin, O., Wiley, J.L: Governmental public health powers during the COVID-19 pandemic: Stay-at-home orders, business closures, and travel restrictions. JAMA - J. Am. Med. Assoc. 323(21), 2137–2138 (2020)
2. Pan, A., et al.: Association of public health interventions with the epidemiology of the COVID-19 outbreak in Wuhan, China. JAMA - J. Am. Med. Assoc. 323(19), 1915–1923 (2020)
3. Quilty, B.J., et al.: The effect of travel restrictions on the geographical spread of COVID-19 between large cities in China: A modelling study. BMC Med. 18(1), 259 (2020)
4. Verhoef, E.T: Second-best congestion pricing in general networks. Heuristic algorithms for finding second-best optimal toll levels and toll points. Transp. Res. Part B: Methodol. 36(8), 707–729 (2002)
5. May, A.D., et al.: The impact of cordon design on the performance of road pricing schemes. Transp. Policy 9(3), 209–220 (2002)
6. Mun, S.I., Konishi, K.J., Yoshikawa, K: Optimal cordon pricing. J. Urban Econ. 54(1), 21–38 (2005)
7. Santos, G: Urban congestion charging: A second-best alternative. J. Transp. Econ. Policy 38(3), 345–369 (2004)
8. Shepherd, S., Sumalee, A: A genetic algorithm based approach to optimal toll level and location problems. Networks Spatial Econ. 4(2), 161–179 (2004)
9. Zhang, X., Yang, H: The optimal cordon-based network congestion pricing problem. Transp. Res. Part B: Methodol. 38(6), 517–537 (2004)
10. Yang, H., Zhang, X., Meng, Q: Modeling private highways in networks with entry–exit based toll charges. Transp. Res. Part B: Methodol. 38(3), 191–213 (2004)
11. Sumalee, A., May, T., Shepherd, S: Comparison of judgmental and optimal road pricing cordons. Transp. Policy 12(5), 384–390 (2005)
12. Abulideh, A., Andrey, J., Melnik, M: Insights into the fairness of cordon pricing based on origin-destination data. J. Transp. Geogr. 49, 61–67 (2015)
13. Souche, S., Mercier, A., Ovtracht, N: Income and access inequalities of a cordon pricing. Res. Transp. Econ. 51, 20–30 (2015)
14. Souche, S., Mercier, A., Ovtracht, N: The impacts of urban pricing on social and spatial inequalities: The case study of Lyon (France). Urban Stud. 53(2), 373–399 (2016)
15. Afandizadeh, S., Abdolmanafi, S.E: Cordon pricing considering air pollutants emission. Promet - Traffic Transp. 28(2), 179–189 (2016)
16. Camporeale, R., et al.: Study of the accessibility inequalities of cordon-based pricing strategies using a multimodal Theil index. Transp. Plan Technol. 42(5), 498–514 (2019)
17. Amirgholy, M., Rezaeestalthrue, H., Poorzahedy, H: Multi-objective cordon price design to control long run adverse traffic effects in large urban areas. NETNOMICS: Econ. Res. Electron. Networking 16(1), 1 (2015)
18. Güthmann, A., Koh, A., Shepherd, S: Optimal charging strategies under conflicting objectives for the protection of sensitive areas: A case study of the Trans-Pennine Corridor. Networks Spatial Econ. 16(1), 199–226 (2016)
19. Li, X., et al.: Cordon- or link-based pricing: Environment-oriented toll design models development and application. Sustainability (Switzerland) 11(1), 258 (2019)
20. Gupta, S., Kalmane, S., Kockelman, K.M: Road pricing simulations: Traffic, land use and welfare impacts for Austin. Texas Transp. Plan. Technol. 29(1), 1–23 (2006)
21. Rodriguez-Roman, D., Ritchie, S.G: Surrogate-based optimization for the design of area charging schemes under environmental constraints. Transp. Res. Part D Transp. Environ. 72, 162–186 (2019)
22. Liu, Z., Meng, Q, Wang, S: Speed-based toll design for cordon-based congestion pricing scheme. Transp. Res. Part C Emerging Technol. 31, 83–98 (2013)
23. Liu, Z., Meng, Q, Wang, S: Variational inequality model for cordon-based congestion pricing under side constrained stochastic user equilibrium conditions. Transportmetrica A Transp. Sci. 10(8), 693–704 (2014)
24. Liu, Z., Wang, S., Meng, Q: Optimal joint distance and time toll for cordon-based congestion pricing. Transp. Res. Part B Methodol. 69, 81–97 (2014)
25. Afandizadeh, S., Abdolmanafi, S.E: Development of a model for a cordon pricing scheme considering environmental equity: A case study of Tehran. Sustainability (Switzerland) 8(2), 192 (2016)
26. Johari, M., Haghshenas, H: Modeling the cordon pricing policy for a multimodal transportation system. Case Stud. Transp. Policy 7(3), 531–539 (2019)
27. Guo, Q., et al.: An integrated model for road capacity choice and cordon toll pricing. Res. Transp. Econ. 62, 68–79 (2017)
28. Kono, T., Kawaguchi, H: Cordon pricing and land-use regulation. Scand. J. Econ. 119(2), 405–434 (2017)
29. Sheffi, Y: Urban Transportation Networks. Prentice-Hall, Englewood Cliffs, NJ (1985)
30. Oppenheim, N: Urban Travel Demand Modeling: From Individual Choices To General Equilibrium. John Wiley and Sons, Hoboken, NJ (1995)
31. Boyle, D.E., Zhang, Y.-F., Lupa, M.R: Introducing “feedback” into four-step travel forecasting procedure versus equilibrium solution of combined model. Transp. Res. Record 1443, 65–74 (1994)
32. Boyle, D., Zhang, Y.-F: Calibrating combined model of trip distribution, modal split, and traffic assignment. Transp. Res. Record 1607, 1–5 (1997)
33. Lin, H: An accessibility-oriented optimal control method for land use development. J. Urban Plann. Dev. 145(4), 04019011 (2019)
34. Lin, H.-Z., Wei, J: Optimal transport network design for both traffic safety and risk equity considerations. J. Cleaner Prod. 218, 738–745 (2019)
35. Gartner, N.H., Messer, C.J., Rath, A: Revised Monograph on Traffic Flow Theory. Federal Highway Administration, Washington, DC (1999)
36. Babazadeh, A., et al.: Reduced gradient algorithm for user equilibrium traffic assignment problem. Transportmetrica A Transp. Sci. 16(3), 1111–1135 (2020)
37. Sumalee, A: Optimal road user charging cordon design: A heuristic optimization approach. Comput.-Aided Civ. Infrastruct. Eng. 19(5), 377–392 (2004)
38. Sumalee, A: Multi-concentric optimal charging cordon design. Transportmetrica 3(1), 41–71 (2007)}