

TRAVEL MANAGEMENT OPTIMIZATION BASED ON AIR POLLUTION CONDITION USING MARKOV DECISION PROCESS AND GENETIC ALGORITHM (CASE STUDY: SHIRAZ CITY)

Mohammad BAGHERI¹, Hossein GHAFOURIAN², Morteza KASHEFIOLASL³,
Mohammad Taghi Sadati POUR⁴, Mohammad RABBANI⁵

¹ Environmental Engineering at Islamic Azad University, North Tehran branch, Iran

^{2, 3, 4, 5} Faculty of Marine Science Islamic Azad University, North Tehran branch, Iran

Abstract:

Currently, air pollution and energy consumption are the main issues in the transportation area in large urban cities. In these cities, most people choose their transportation mode according to corresponding utility including traveller's and trip's characteristics. Also, there is no effective solution in terms of population growth, urban space, and transportation demands, so it is essential to optimize systematically travel demands in the real network of roads in urban areas, especially in congested areas. Travel Demand Management (TDM) is one of the well-known ways to solve these problems. TDM defined as a strategy that aims to maximize the efficiency of the urban transport system by granting certain privileges for public transportation modes, Enforcement on the private car traffic prohibition in specific places or times, increase in the cost of using certain facilities like parking in congested areas. Network pricing is one of the most effective methods of managing transportation demands for reducing traffic and controlling air pollution especially in the crowded parts of downtown. A little paper may exist that optimize urban transportations in busy parts of cities with combined Markov decision making processes with reward and evolutionary-based algorithms and simultaneously considering customers' and trip's characteristics. Therefore, we present a new network traffic management for urban cities that optimizes a multi-objective function that related to the expected value of the Markov decision system's reward using the Genetic Algorithm. The planned Shiraz city is taken as a benchmark for evaluating the performance of the proposed approach. At first, an analysis is also performed on the impact of the toll levels on the variation of the user and operator cost components, respectively. After choosing suitable values for the network parameters, simulation of the Markov decision process and GA is dynamically performed, then the optimal decision for the Markov decision process in terms of total reward is obtained. The results illustrate that the proposed cordon pricing has significant improvement in performance for all seasons including spring, autumn, and winter.

Keywords: air pollution, dynamic optimization, genetic algorithm, Markov decision-making process

To cite this article:

Bagheri, M., Ghafourian, H., Kashefiolasl, M., Pour, M. T. S, Rabbani, M., 2020. Travel management optimization based on air pollution condition using Markov decision process and genetic algorithm (case study: Shiraz city). *Archives of Transport*, 53(1), 89-102. DOI: <https://doi.org/10.5604/01.3001.0014.1746>



Contact:

1) baghry.1398@gmail.com [<https://orcid.org/0000-0002-5901-7144>], 2) ghafourian.13912@gmail.com [<https://orcid.org/0000-0002-3859-732X>], 3) kashefiolasl.14581@gmail.com [<https://orcid.org/0000-0002-1165-3686>], 4) sadatipour.112381@gmail.com [<https://orcid.org/0000-0002-0518-9776>], 5) rabbanimohammad403@gmail.com [<https://orcid.org/0000-0003-2733-8282>]

1. Introduction

In the current age, the necessity of rapid transportation and reducing air pollution are the most complicated and important human issues (Kelly and Zhu 2016). The fact that most people consider is that if there is no effective solution in terms of population growth, urban space, and transportation demands, it will turn into complex and widespread problems for city dwellers and their environment. One of the well-known ways to solve these problems is travel demand management (TDM). TDM defined as a strategy that aims to maximize the efficiency of the urban transport system by granting certain privileges for public transportation modes, Enforcement on the private car traffic prohibition in specific places or times, increase in the cost of using certain facilities like parking in congested areas (Broaddus et. al 2009, Ferrarese 2016). The use of the latter approach (e.g. network pricing) will result in effective and flexible utilization of existing transport facilities to better travelers' demand. In network pricing could easily handle the costs of automobiles, the welfare cost of occupants and the environmental pollution-related cost. Thus, network pricing emerges as a useful TDM tool (Ferguson 2018). Although there have been many research works in network pricing and park-and-ride area to solve simultaneously congestion and air pollution issues, this paper presents an optimal solution for these problems. The proposed method starts with identifying and defining the congestion or pollution plagued area, then, price the boundary-crossing for all vehicles including public and private ones. The park-and-ride facilities are given in the study areas, and the price of them is associated with the cordon price. Also, the cordon price is calculated according to a stochastic Markovian process of air quality in the city such that the total reward of the transportation system is maximized.

There have been many papers which deal with pricing automobiles for air pollution control. A comprehensive review of literature related to network pricing for air pollution control issues is provided. Cavallaro et. al (2018) discussed the potential effect of the network pricing based congestion schemes on the reduction of carbon emissions. The results reveal that the role of these strategies in term of carbon emissions are significant, thus the techniques can be used by transport policy-makers. A recent paper on congestion pricing to control pollutions emission is

presented by Dai et. al (2015) who emphasize the air pollution in China. They extend a multimode-Logit (car and bus modes) pricing model in terms of delay and emissions of the vehicles. The numerical experiment results are illustrated that congestion pricing can increase the mode share of the bus. Lyons et. al (2003) proves that vehicle emission is proportional to vehicle-kilometers of travel, then, they presented a model for estimating vehicle pollutions in the urban cities in the United States (US). de Palma and Lindsey (2009) provided a review including pricing lanes, facilities, cordon, and zones based on distances time(at, time-of-day, responsive), and vehicle characteristics (type, weight, number of axles) schemes. They also mentioned that pricing technologies are also numerous including roadside recognition, dedicated short-range communications, and in-vehicle technologies that rely on satellite or cellular networks. They provide clues for the choice of appropriate methods and technologies for different purposes. One of the fascinating and popular methods to control vehicle flows in certain areas like central business districts (CBD) of a city, is cordon pricing.

This method is the second-best method for controlling traffic congestion method in the monocentric city (Verhoef 2005). Amirgholy et. al (2015) show that network pricing is a feasible and sustainable method for managing the traveler demands while relieving the network traffic effects (e.g. congestion and pollution). Their model results showed that cordon pricing results in a considerable reduction of the pollution level. Also, Amirgholy et. al (2015) mentioned that the pricing schemes make new decisions regarding trip purposes, characteristics of the transport modes, therefore this technique should be attractive from travelers' perspective.

In our investigated literature, we have not seen any research that considers modeling and optimization simultaneously with the interaction of operator-user-environment at a real traffic network in order to traveler choose an appropriate mode choice (correspond to the lowest utility) in the network. Therefore, the proposed approach helps to sustain urban air quality and encourage car people to use public transportation system from their origins or switch to public transport near a park-and-ride facility (Sawicki et al., 2016). The operator determines the optimal price for cordon districts based on the current level of the ambient air pollution and stochastic variations of the

environment condition. The total reward of the system includes transportation time, the monetary cost of the car drivers in choosing their next best modes, and the benefit of reduction in the pollution cost. Finally, the proposed strategy should be an essential part of the travel management system, especially in big cities. The rest of the paper is organized as follows: The problem is formulated as a Markov Decision making Process in Section 2, where the reward matrix components, and the state transition probability matrix and atmospheric conditions as states in the problem are defined. In section 3, the Genetic Algorithm (GA) optimization method is explained. The application of the Markov model to the large City of Shiraz in Iran is presented in section 4. Also, this Section will investigate the selection of decision level impact on the user and operator cost components and present optimizing results using GA. Finally, Section 5 summarizes and concludes the paper and suggests possible future research directions.

2. Material and methods

The problem is defined and then formulated as a Markov decision-making process and solved for the decision variables. The following sub-sections will present the problem definition and formulation.

2.1. The Markov process with reward

Markov decision problems (MDPs) are a general mathematical formalism for representing shortest path problems in stochastic environments. This formalism is based on the theory of Markov decision processes. A Markov decision process relies on the notions of state, describing the current situation of the agent, action (or decision), affecting the dynamics of the process, and reward, observed for each transition between states. Let $X_1 = \{x_n, n = 1, 2, \dots, N\}$ be a markov chain with finite state space E where, $|E| = m$. Therefore, x_n can take one of values 1 to m . Also, assume that transition probability from state i to j ($i, j \in E$) is equal to $p(i, j)$ and also is followed by a reward $r(i, j)$. Due to probability nature of the X_1 , the problem is stochastic. Now, let $v^n(i)$ be the expected value of the reward after n transition of the system, then:

$$v^n(i) = \sum_{j=1}^m p(i, j)[r(i, j) + v^{n-1}(i)] \quad (1)$$

Also, let $q(i) = \sum_{j=1}^m p(i, j)r(i, j)$. With performing Z-transform analysis on Eq. (1) for large values of n , we have:

$$v^n(i) = ng(i) + v(i), i \in E, n = 1, \dots, N \quad (2)$$

where $v(i)$ is the asymptotic value of $v^n(i)$ when $n \rightarrow \infty$ and $g(i)$ is defined as follows

$$g(i) = \sum_{j=1}^m s(i, j)q(j), i \in E \quad (3)$$

Also, the stochastic matrix $S = [s(i, j)]_{m \times m}$ represents limiting probability of the system that starting from state $i \in E$ and being in state $j \in E$, for now. $g(i)$ may be interpreted as the expected value of the reward of the system, given that it started from state i and has undergone many iterations, or the revenue of the i^{th} state, or the asymptote of $v^n(i)$. If the system is completely ergodic, all rows of S become equal, and all states would have the same revenue, say $g(i) = g$ for all i 's

$$g = \sum_{j=1}^m \Pi(j)q(j) \quad (4)$$

where $\Pi(j)$ is the limiting probability of the system to be in state j and for large n , and by using this, Eq. (2) is simplified as:

$$v^n(i) = ng + v(i), i \in E \quad (5)$$

Now, assume that the alternative $k = 1, 2, \dots, K$ are available action in each state $i \in E$. Also, suppose that for each action k , the transition matrix $P^k = [p^k(i, j)]_{m \times m \times K}$ and the corresponding reward $R^k = [r^k(i, j)]_{m \times m \times K}$ are known. Where $p^k(i, j)$ and $r^k(i, j)$ are the probability and reward of moving from state i to state j when decision k is taken. The decision makers want to find the best possible

decision in each stage (day) n and each state i , such that maximize the total reward during the time interval $[0, n]$. The set of all candidate decisions d_i^n for all $i \in E$ and $n = 1, 2, \dots, N$, is called a policy. d_i^n indicates the alternative decision to choose in stage n when system is in state i . Therefore, the problem is to find the set $\{d_i^n\}$ so that maximize Eq. (6) as

$$v^{n+1}(i) = \max_k \left(\sum_{j=1}^m p^k(i, j) [r^k(i, j) + v^{n-1}(j)] \right) = \max_k \left(q^k(i) + \sum_{j=1}^m p^k(i, j) v^n(j) \right), i \in E, n = 0, 1, \dots \quad (6)$$

where $q^k(i) = \sum_{j=1}^m p^k(i, j) r^k(i, j)$ is associated with the P^k and R^k , respectively. Given policy $\{d_i^n = k\}$ associated with known P^k and R^k matrices for a limiting behaviour of the system (n is large), Eqs. (2) and (5) yield:

$$g = q^k(i) + \sum_{j=1}^m p^k(i, j) (v(j) - v(i)), i = 1, \dots, m \quad (7)$$

This is a set of $m+1$ unknowns and m linear equations, that may be solved for g and the relative values of $v(i)$'s such that assuming $v(m) = 0$. With finding the relative values of $v(i)$'s and g , we may find the optimal action for state i and stage $n+1$. Also, by using Eq. (6) and $v^n(i) = ng + v(i)$ for large n , by maximization the right-hand side of the following equations over k as:

$$g + v(i) = \max_k \left(q^k(i) + \sum_{j=1}^m p^k(i, j) v^n(j) \right), i = 1, \dots, m \quad (8)$$

Thus, by solving Eq. (7) and substituting the $v(i)$ s that calculated in pervious iteration, again find new policy k^* that maximizing the right-hand side of the Eq. (8). If new policy k^* and pervious ones is same, then Howard-policy iteration method is stopped, otherwise the process is reiterative (Howard 1960). Eq. (8) is seen as objective function for Genetic Algorithm that is mentioned in the next section.

2.2. The Reward matrix

Consider the case where the O/D demands are fixed during several year. Further to assume, that the short-run action is only mode change. Then, the reward accrued by any air-pollution decision k is the sum of two outcomes: (a) the benefit of the positive air quality change from state i to j ($i, j \in E$) characterized by the ambient air CO level, and (b) the net benefit arises from a lower transportation cost (travel time). The action is associated to the change of tolls in private cars that entering to the restricted zone. This would impact on the mode choice selection of the drivers' actions. When the cost of the travel time is more than environment ones, because of the stagnant air of the city, increasing the tolls will force auto drivers toward using a public mode completely, park-and-ride transit or completely using public transportation to arrive final destinations. In the following sections, will describe the problem environment and its sensitivity to the operator's decisions on the prices (Sendek-Matysiak 2019). The model presented below is including: Mode choice models, Auto operating cost, Equivalent travel time monetary cost and Cost of reduction in car passenger's utilities.

2.2.1. Mode choice models

The mode choice model of the paper is of a multinomial logit type, as shown below:

$$Pr_{ks}^{pm} = \frac{e^{U_{ks}^{pm}}}{\sum_{m \in M} e^{U_{ks}^{pm}}}, m \in M, (k, s) \in P \quad (9)$$

where Pr_{ks}^{pm} is the probability of choosing mode m for a trip with purpose p and from origin k to destination s , which possesses a utility function U_{ks}^{pm} . Since the user can Travel according its utility, we assume that all the transport modes are available for the all of set O/D (Origin/Destination) pairs. The utility function U_{ks}^{pm} is a function of the independent variables of the modes (e.g., travel time), passenger features (e.g., car-ownership) and trip purpose (e.g., shop). The utility function is used in the paper is given in (Poozahedy et al 2016). The tolls that private cars must pay to cross the restricted zone of the city will have a negative effect on the utility of this mode. We also assume that only private cars trips

with originating from an origin (O) outside the restricted zone(s) and destinations (D) inside these zones can pay the tolls. In the paper, further more to Bus and auto private cars, extensive simulations have performed with the assumption that auto passengers can also choose the public transit or park-and-ride (P/R) system. The portion of the total auto passengers from origin k to destination s for trip purpose p who are willing to pay entrance the toll τ to the restricted area is equal to w_{ks}^p and assumed as:

$$w_{ks}^p = \exp(-a(\frac{\tau}{D_{ks}})^b); a, b > 0 \quad (10)$$

where D_{ks} is the shortest network distance from origin k to destination s . a and b are two parameters of the function. Now suppose that y_{ks}^p is the portion of the auto demand planning to change mode and choose public transit or park-and-ride system and $(1 - y_{ks}^p)$ choose taxi (which, in our case, is allowed to enter the restricted zone free of charge) as follows:

$$y_{ks}^p = \exp(-c(\frac{\tau}{D_{ks}})^d); c, d > 0 \quad (11)$$

where c and d are also two parameters of the function (Azari et al., 2013). Thus, by increasing τ per unit distance D_{ks} decreases w_{ks}^p , increases $(1 - w_{ks}^p)$, increases y_{ks}^p as well as $y_{ks}^p(1 - w_{ks}^p)$, and most likely increases $(1 - w_{ks}^p)(1 - y_{ks}^p)$, and the latter two being the share of transit and park-and-ride and the share of the taxi from the diverted auto demand, respectively (see the Fig. 1 (a) and (b) for pictorial views of these functions for the city of case study). The portion of the diverted auto demand for a purpose p , which is absorbed by public transit or park and-ride system, decides to use either mode based on the minimum of the transit time from origin k to s (t_{ks}^b), or equivalent park-and-ride time $t_{ks}^{P/R} = t_{k,P/R}^a + t_{P/R,s}^b + \beta c_{P/R}$. Where, In the above expression, P/R represents the nearest park-and-ride facility on the way from origin k to destination s , and $t_{k,P/R}^a$ is the shortest auto travel time from origin k to the nearest park-and-ride facility, $t_{P/R,s}^b$ is the shortest transit time from the park-and-ride facility to destination s , and $c_{P/R}$ is the parking and transit fare cost which is translated into equivalent travel time by the conversion factor β .

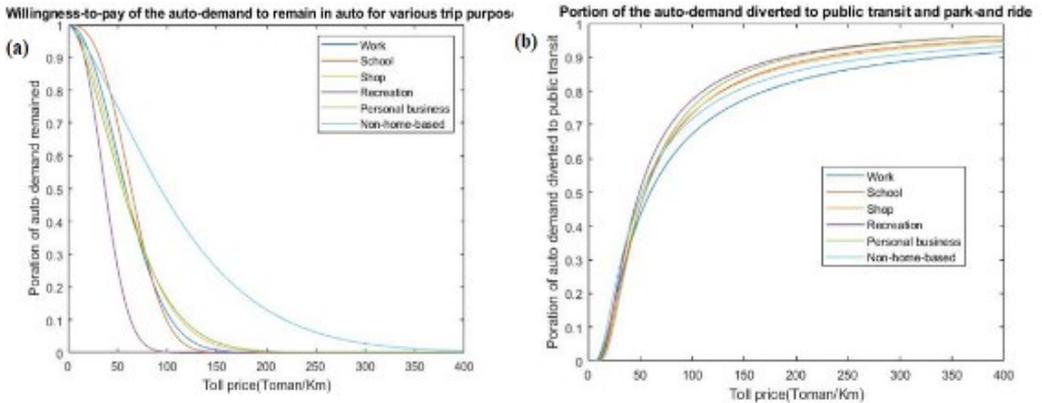


Fig. 1. Demand diversion models in Shiraz City (a): diagram of w_{ks}^p in term of $\frac{\tau}{D_{ks}}$ for various trip purposes

(b): diagram of y_{ks}^p in term of $\frac{\tau}{D_{ks}}$ for various trip purposes

We represent the effect of tolls τ upon the travel costs of the car drivers willing to pay τ to enter the priced zone, the travel times of all links heading inside the priced zone are increased by the travel time equivalent to the toll:

$$t_{ij} = t_{ij} + \gamma\tau \quad (12)$$

where t_{ij} is the average travel time of link (i, j) pointing into the priced zone boundary. The cost of parking at a P/R bus terminal ($c_{P/R}$) is simply assumed to be a linear function of the priced zone entrance fee (τ) as follows:

$$c_{P/R} = \alpha\tau \quad (13)$$

where α is a constant and assumed to be equal to 0.05 for the simulation study.

2.2.2. Auto operation cost

Auto operation cost is including of costs of auto investment, operation, maintenance, repair and as well as accident/insurance cost. Let \bar{c} represent an average ratio with money units per car-kilometer unit. So, auto operation cost per hour under decision k , C_{opa}^k , may be obtained as:

$$C_{opa}^k = \bar{c}(X^{r1} - X'^k) \quad (14)$$

2.2.3. Equivalent travel time monetary cost

Now, assume that c_i be the value of one person-hour travel time and C_{ta}^k be the equivalent monetary cost of total non-transit travel time under decision k . Then:

$$C_{ta}^k = c_i(Y^1 - Y^k) \quad (15)$$

where Y^k is the passenger-hour of auto vehicles under decision k ($k=1$ is the do-nothing alternative) and computed as

$$Y^k = \sum_{\substack{m \in M \\ m \neq bus}} \alpha^m Y^{mk} \quad (16)$$

where α^m is the average occupancy rate (passenger per vehicle) of mode m and Y^{mk} is the travel time of mode m under decision k . Also, average occupancy rate for private cars and taxis during simulation studies is considered as 2 and 3 passenger per car, respectively (Chen et al., 2017). Public transit (bus) total travel time (in passenger-hours) under decision k is viewed as the equivalent monetary cost of the public transit travel time spent in the network and is also calculated as:

$$C_s^k = c_i(\bar{Y}^1 - \bar{Y}^k) \quad (17)$$

2.2.4. Equivalent travel time monetary cost

We assume that the maximum number of demanded private cars for an O/D pair is equal to Q . Then, an estimate of the cost of the reduction in Consumers' Surplus (CS) is presented as follows. This cost is because at price p^k and q^k passengers were willing to drive their cars. When price is increased to $p^k + dp^k$, demand of automobiles decreases to $q^k - dq^k$ and dq^k is diverted to their best modes. Thus, dq^k passengers who were willing to pay price p^k to use their cars are now diverted to get off their cars by the increasing price from p^k to $p^k + dp^k$. An approximate measure of ΔC_{cs}^k may be analytically found by our proposed model that is described below, according to the following equations:

$$\begin{cases} \Delta C_{cs}^k = \sum_{k=1}^K \frac{(q^k + q^{k+1})(p^k - p^{k-1})}{2} \\ p^0 = 0, q^k = \frac{200Q}{1 + \exp(0.0005 \times |p^k - 250|)} \end{cases} \quad (18)$$

where q^k is the demand for private car (in this study, from an origin to a destination inside the cordon) when the price to enter the cordon is p^k under decision k .

2.3. Formulation of the Reward Matrix

For now, we can simulate the proposed dynamic pricing model to find optimal policy for the ambient air pollution control for the case study city. We do

this, first, by unifying the decision outcomes under a given weighting scheme for the different bodies involved to get the reward matrix. Then, we show how the decision would change if these weighting schemes and the surrounding forces (seasons) change. Changing of state from i to j in a period of time (say, a day), due to decision k , results in a reward of $R^k(i, j)$. In what follows we define an example for demonstration purposes. We assume the outcome of the decision k , in the reward function to be attributed to two major interest groups. Each outcome has an importance, or weight, that we show by ω :

- 1) The operator outcomes (changes in the consumption of the limited resources including gasoline G^k and diesel D^k fuels) with weight ω_{re} .
- 2) The public outcomes with weight ω_p : The public outcomes can be classified into two sub-groups:
 - 2.1) The general public outcomes (including the changes in the public health effects) in state j , H^j with weight ω_{ph} . Also, H^j define as a function of the state j as follows:

$$H_j = \begin{cases} 1; j = 1 \\ 0.5; j = 2 \\ 0; j = 3 \end{cases} \quad (19)$$

- 2.2) the users of the network outcomes with weight ω_u , that may be divided into new categories:
 - 2.2.1) the yearly operation cost, which include the distance traversed by the automobiles in the network (proportional to the monetary cost of travel), VK^k ; the travel time spent by the public vehicle passengers, Tp^k ; and the travel time spent by the non-public vehicle passenger, Tnp^k ; all of them with weight ω_{uo}
 - 2.2.2) the reduction in welfare of the automobile occupants, CS^k , which may be interpreted as users' capital cost, considered with weight

ω_{uc} . Therefore, the reward matrix $R^k(i, j)$ is given as Eq. 20:

$$R^k(i, j) = \omega_{re}(G^k + D^k) + \omega_{ph}(H^j) + \omega_{uo}(VK^k + Tp^k + Tnp^k) + \omega_{uc}(CS^k) \quad (20)$$

2.4. The transition probability matrix

There are three factors affecting the Markov transition probabilities that each having a stochastic effect. They are: (a) ambient air pollution (CO) level as a result of private vehicle emissions; (b) atmospheric conditions (winds, precipitations, etc.); and (c) the operator's decision (pricing automobiles entering the restricted zones of the city). At a regional level, there is a recursive equation postulated between the pollution concentration, P , pollutant emissions, E , and atmospheric condition, A , as follows:

$$P(n) = f(P(n-1), E(n), A(n)) \quad (21)$$

where n denotes day n . This equation is saying that, the concentration of the pollution level in day n is a function of the concentration of the pollution level in previous day, decision of the nature and the atmospheric condition. If there is no weather influence upon the ambient air, any state of system (state of the air pollution) will either remain as it is or turn into another worse condition by the car emissions. The transition probabilities of moving the state of system from state i to j ($j \geq i$), under decision k , $p^k(i, j)$ can be estimated by knowing the lifetime distribution of any state of the system. One suitable distribution for assessing lifetime of state T_i , is a negative exponential as:

$$f_{T_i}(t) = \theta_i \exp(-\theta_i t); t \geq 0, \theta_i > 0, i \in E \quad (22)$$

Also, let $E = \{1 = \text{Healthy}, 2 = \text{Tolerable}, 3 = \text{Harmful}\}$ be the state space of the ambient air quality. We want to obtain the probability of the daily change of system's state. In a meso-scale, total automobile emission of pollutants is a function of total vehicle-kilometers per hour. For simplicity, vehicles are divided into two public and private categories and for their

overwhelming roles in the travel market, buses and automobiles are taken as the representatives of these two categories, respectively. Thus, we focus on bus-km and car-km (or their equivalents). For $X(n)$ as the car-km and $X'(n)$ as the bus-km travelled in day n , one may write the parameters of the transition probability distributions as a function of these affecting factors as:

$$\theta_i(n-1, n) = G_i(X(n), X'(n)) \quad (23)$$

Where $\theta_i(n-1, n)$ is the parameter of lifetime probability function of state (air quality level) i of the system, influencing its change from (end of) day $n-1$ to (end of) day n , which is a function of the vehicle-kilometres travelled in day n , $X(n)$, and $X'(n)$. Assume that the changing in bus-kilometers traveled have negligible effect on the quality of ambient air. Thus, in the view, the decision-makers are willing to shift the demand from cars to buses (or completely public transport), is certainly friendlier to the environment than cars. So, Eq. (23) is simplified as follows:

$$\theta_i(n-1, n) = G_i(X(n)) \quad (24)$$

Or, in further simplified Eq. (24) as:

$$\theta_i = G_i(X) \quad (25)$$

By using A first-degree polynomial function of Eq. (25), we will obtain Eq. (26) for the three-state of the at atmospheric conditions:

$$\theta_i = a_i X + b_i; i = 1, 2, 3 \quad (26)$$

Where $a_i > 0$ and b_i are the parameters of the model for the states $i \in E$ ($E = \{1, 2, 3\}$). By combining the Eq. (22) and (26) we have:

$$f_{ij}(t) = (a_i X + b_i) \exp(-(a_i X + b_i)t) \quad (27)$$

By calculating of the transition probability of the state of the system turning from i to j and also, assuming that weather upon the ambient air and decision doesn't influence, we have:

$$\begin{aligned} p(1,1) &= \exp(-(a_1 X + b_1)) \\ p(1,2) &= \frac{(a_1 X + b_1)}{(a_2 - a_1)X + (b_2 - b_1)} [\exp(-(a_1 X + b_1)) \\ &\quad - \exp(-(a_2 X + b_2))] \\ p(1,3) &= 1 - p(1,1) - p(1,2) \quad (28) \\ p(2,2) &= \exp(-(a_2 X + b_2)) \\ p(2,3) &= 1 - \exp(-(a_2 X + b_2)) \\ p(3,3) &= 1 \end{aligned}$$

Note that $p(i, j)$ are functions of X , which is itself a function of the operator's decision $k = 1, 2, \dots, K$. For given value of k , $X(k)$ would be known, and so would be $p^k(i, j)$. Atmospheric condition are very important factors in ambient air quality. The complex cause-and-effect relationships among basically unknown variables and parameters governing the happenstance of these phenomena leave these incidents as stochastic processes. Let w_j be the probability that an atmospheric event would improve the ambient air quality by $j-1$ states regardless of the traffic condition. Then, for a three-state problem, w_1 , w_2 , and w_3 may be understood as light, medium and heavy atmospheric changes, and hence we have:

$$\sum_{j=1}^3 w_j = 1 \quad (29)$$

Now suppose that $p(i, j)$ is the transition probability from state i to j under the influence of the atmospheric incidents. Also, by assuming that independence of atmospheric changes and air quality, and prompt effect and as well as approximating happen at the beginning of each day, may write as:

$$\begin{aligned}
 P(1,1) &= p(1,1)(w_1 + w_2 + w_3) + p(1,2)(w_2 + w_3) \\
 &\quad + p(1,3)(w_3) \\
 P(1,2) &= p(1,2)(w_1) + p(1,3)(w_2) \\
 P(1,3) &= p(1,3)(w_1) \\
 P(2,1) &= p(2,2)(w_2 + w_3) + p(2,3)(w_3) \\
 P(2,1) &= p(2,2)(w_2 + w_3) + p(2,3)(w_3) \quad (30) \\
 P(2,2) &= p(2,2)(w_1) + p(2,3)(w_2) \\
 P(2,3) &= p(2,3)(w_1) \\
 P(3,1) &= p(3,3)(w_3) \\
 P(3,2) &= p(3,3)(w_2) \\
 P(3,3) &= p(3,3)(w_1)
 \end{aligned}$$

where $p(i, j)$ are derived as Eq. (28). Thus, the Transition Probabilities (TP) may be computed depending on the tolls level upon the automobile users in entering the congested zone(s), which changes the composition of private/public traffic in the network. This process is illustrated as follows: (a) Choose prices for entering the congested zone(s) and parking at a park-and-ride facility (choosing the action k). (b) Assign the O/D demand under the price of level k according to traffic assignment procedure. (c) Compute the policy k outcomes, including total car-kilometre x^k based on traffic assignment procedure. (d) Compute the parameters of the lifetime distribution function under the price of level k , $\theta_i^k = a_i x^k + b_i; i = 1, 2, 3$. compute the transition probabilities from state i to j in the study zone(s) under the price of level k , by using Eqs. (28-30). Table 1 shows the values for w_j , a_i , and b_i in Eq. (28) and for 3 seasons of the year.

Table 1. Values of w_j in Eq. (30) and a_i and b_i in Eq. (28).

	Spring			Autumn			Winter		
j	1	2	3	1	2	3	1	2	3
w_j	0.63	0.32	0.05	0.55	0.39	0.06	0.74	0.33	0.03
$a_1 = 1.86E - 6, a_2 = 7.74E - 05, b_1 = 0.9, b_2 = -0.005$ for all seasons.									

2.5. Demand Estimation and Traffic Assignment

Traffic assignment routine constitutes two parts: (1) demand estimation and (2) traffic assignment. The first part receives the zone's independent variables such as trip production and estimation, trip distribution and mode choice model. To use this model for this problem, the diversion mechanism of Eq. (10) is employed to compute the remaining trips of each trip purpose to be made by automobiles. Also, the diversion model of Eq. (11) is used to obtain the portion of the diverted auto demand that will be attracted by the transit mode, either directly from the origin of the trip or through the nearest P/R facilities, whichever has less travel time. In second ones, by applying a multi-class user equilibrium method (McCord, 1987); link volume, travel time components are updated and this process is reiterated until the travel time of network's link are convergence. These results (link volumes, travel time components, kilometers traveled, fuel consumed, emissions of CO and other pollutants, etc.) are summarized and reported in Fig. 2. Then, the outputs of the traffic assignment are used to compute non-transit passenger hour (Y^k), transit passenger-hour (\bar{Y}^k), passenger car-kilometer (X^k), and liters of fuel consumed (L^k), where k represents the decision on the price level. The first two quantities would give the equivalent monetary cost of travel times spent in the network (Eq. (15) and (17)). The third quantity would be used in the computations of auto operating cost (Eq. (14)), and the fourth may be transformed into its monetary value. Moreover, increasing the travel cost for car drivers would make them worse off by a reduction in consumers' surplus (Eq. (18)). Hence, this cost should also be added to the previous ones to account for the worse conditions of the car drivers as a result of our actions. The benefit accrued by our decisions would be the reduction of the emitted pollutants (here, CO) that is expected to outweigh the aforementioned costs. After obtaining the traffic assignment, Markovian decision-making procedure is performed. In the process, two important matrices must be calculated: reward matrix and the Transition Probabilities for the specified toll levels. Finally, by calculating the expected value of the reward of the Markov system as the objective function of the Genetic Algorithm, optimal values of the importance of the Reward matrix components are iteratively updated so that the objective function is maximized.

3. Genetic algorithm

Genetic algorithms belong to the class of stochastic search algorithms that inspired from Darwin's natural evolutionary process. As such, GAs present an intelligent approach of random searches for solving optimization problems. Due to these characteristics, GAs have been applied to a wide range of optimization problems, especially those with discrete design variables and discontinuous and nondifferentiable objective functions (Kramer, 2017). A typical GA-based optimization problem requires a genetic representation of the solution domain and a fitness function to evaluate the feasibility of the solution domain. The following fundamental procedures are involved in the GA optimization i.e., initialization, fitness evaluation, selection, crossover, and mutation. Fig. 2 shows the iterative process for the GA approach (Pal and Wang 2017). GA begins with a randomly generated population of individuals. The individuals in the initial population are characterized by genomes containing a string of chromosomes, which is composed of design variable values generated randomly. Then the evaluation process follows, where each individual is evaluated based on the fitness function expressed in terms of the objective function (MirJalili, 2019). At the end of the evaluation process, all the individuals are sorted according to their fitness values. The GA operators (i.e., selection, crossover and mutation) are then applied. The selection, which is usually the first operator to be applied, selects individuals with the best fitness values that can perform the crossover and mutation operators to produce new offspring. Typical techniques in the selection operator are elitism, roulette wheel and

ranking selection (Ahuja and Batra 2018). The crossover operator combines two selected individuals to produce a new offspring. The chromosomes of individuals with larger fitness values have a higher probability to reproduce, and the offspring, with combined traits from the selected parents, may achieve higher fitness values (Frances Buontempo 2019). Finally, the mutation operator creates new offspring through small random changes of the information contained in the design variables of a selected individual. The evolution process is not only determined by inherited traits but also genetic diversity is promoted. Typical mutation techniques include bit inversion, order changing, and adaptive mutation. Once the GA operator procedure finishes, a new generation with evolved individuals is produced and, if a stopping criterion is not met, the iteration process continues. Instead, the optimal solution is obtained when the stopping criterion is satisfied. When a stopping criterion (such as time, number of generations, etc.) has been verified then the GA can be converged (Eiben et al. 2003).

4. Results and Discussion

4.1. Case Study

The case under study is the City of Shiraz with a population of 1.8 million in 2020. The study area has 189 zones, of which 156 zones are internal and 33 zones external, related to the surrounding areas and outside world. These zones are grouped based on socioeconomic characteristics and administrative boundaries into 15 internal regions and region 16 constitutes the external zones. Fig. 3 shows the 156 internal zones and the respective 15 regions.

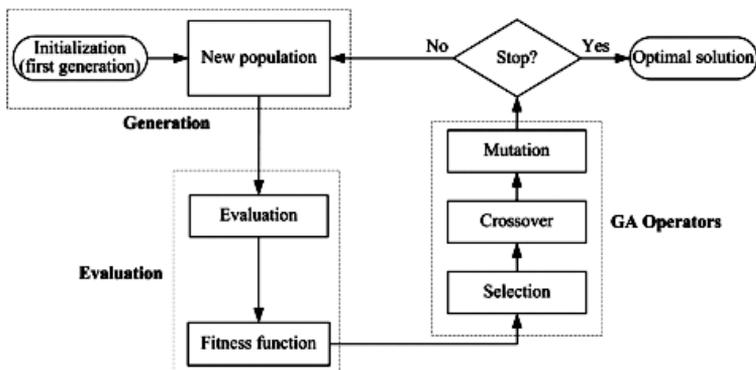


Fig. 2. Flow chart of the GA (Kramer 2017)

The O/D demand belongs to the design (morning peak) hour of a typical working day in 2006. The transportation system of the study area is comprised of two networks of roads and public transit. The road network has 1078 nodes and 1611 two-way and some one-way links. Fig. 3 shows this network. The public transit network is basically that of bus transport with 74 bus lines and 382 conventional buses. The street and road network have 17 types of links, with FHWA type of travel time functions (in the form of $t = a + bx^4$; a and b are two link-dependent constants). According to figure 3.b, there are the 10 park-and-ride bus terminals to facilitate the trips of those passengers who preferred to park at these terminals and use the bus transit systems to reach the central city (ITSR, 2001). Also, assume that the each of roads and streets are scheduled according to maximum and minimum of O/D matrices, e.g. for a trip, from zone i to j , maximum of the link capacity that connects these two zones, is assumed to be equal to the element in the i^{th} row and j^{th} column of the O/D matrix.

4.2. Traffic assignment results under different toll levels for the short-run duration

The outputs of the traffic assignment section for the case under study for different toll levels (k) would provide the reward matrix for the Markovian decision process and pave the way toward the application of the dynamic cordon pricing (Howard policy iteration) method mentioned at the of Section 2.1. Fig. 4 shows the results of the computation of the cost components of the Reward matrix for $\alpha = 0.05$. According to this, we observe the relative values of

the cost components for the transportation system. These values are computed as follows. Suppose that Z^i be the cost component Z e.g. gasoline cost for toll price level ($K = 1, 2, \dots, 10$), where toll level 1 is equal to zero (the do-nothing alternative). Let $DZ^i = Z^i - Z^1$. Normalize the values of DZ^i by dividing them by the respective maximum (minimum) values over i , $\frac{DZ^i}{Ext_j(DZ^i)}$, where

$$Ext_j(DZ^i) = \max(\min_j(DZ^i))$$

Z is a cost which increases (decreases) with the increase of the toll. The general trends of variation of the cost components of the system are shown in Fig. 4, which are s-shaped with respect to the increase in toll level.

4.3. Optimization Pricing Strategy using Markov Decision Process and Genetic Algorithm

In the research work, we present an approach in order to optimize pricing strategy using Markov Decision Process and Genetic Algorithm. According to Eq. (20), the optimal policy of the Markov Decision Process is dependent on the importance of the cost components and therefore we will obtain optimal values of the parameters that the expected value of the reward of the system is maximized. By performing simulation, in the Spring, Autumn and Winter, the optimal values are obtained through 12 iterations. The convergence criterion of the objective function is selected to select the maximum iteration. Also, the parameters of GA are set as Table 2.



Fig. 3. (a) The study area regions (bold lines) and zones (light lines). Region 1 boundary forms the cordon. (b): The study network, cordon area, and park-and-ride facility locations (dark dots) (ITSR, 2001)

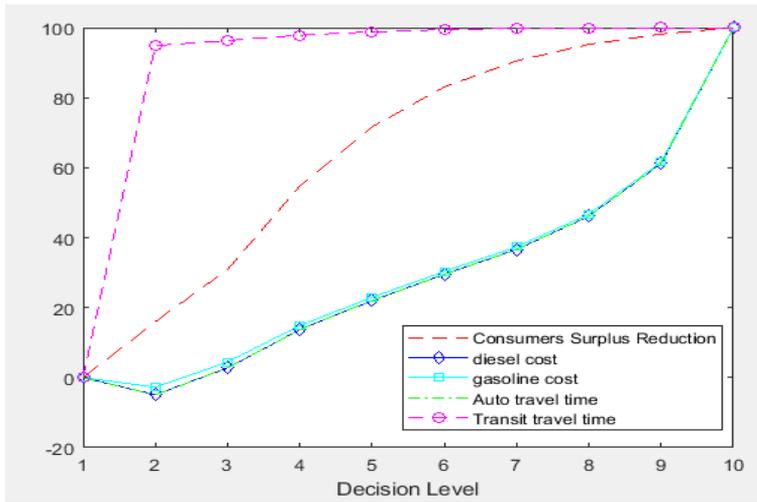


Fig. 4. Variation of the cost components of the total cost of the network.

Table 2. The values of the parameters of the GA in the simulation study.

<i>parameters</i>	<i>value</i>
<i>Variable variation range</i>	[0,1]
<i>Number of variables</i>	4
<i>Percentage of mutation</i>	0.3
<i>Mutation rate</i>	0.1
<i>Number of populations</i>	100
<i>Maximum iteration</i>	12

According to Fig. 5, values of objective function increase through iterations (for Spring, Autumn and Winter seasons), and therefore by using the GA, the expected value of the reward of the system is maximized. The important point of the results is that optimization has a significant effect on the optimizing performance of the system, especially for the Winter season.

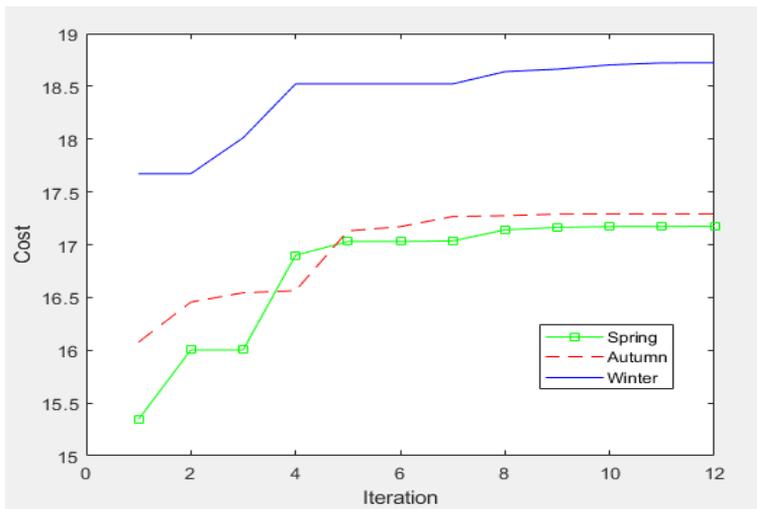


Fig. 5. Variation of the objective function through iterations.

5. Conclusions

The problem of congestion pricing in the Shiraz city's zones is investigated in this paper and Central Business District (CBD) considers as cordons that have priced for private cars entering these zones. In order to, the travelers could minimize their utility, consider several park-and-ride facilities around the cordon's boundaries, which are also priced proportional to the entering cordon's price. Both of these prices are related to the level of pollution in the city. Therefore, this paper proposes a model to decide on the price of using the park-and-ride facilities in congested zones, so that the optimal and feasible price level is calculated by using Markov Decision Process and Genetic Algorithm. Since each pollution level (also in different seasons) required different policies, so the Markov process with reward is considered in this paper. The reward of the process includes operator and general outcomes. The compromising between these outcomes can become a complex problem. Therefore, the Genetic Algorithm is applied to obtain the weighting of the outcomes. Also, in every iteration (correspond to every decision in Markov process), the multi-class user equilibrium method utilized to determine link volume, travel time, kilometer traveled, fuel consumed, and emission of CO. The optimal cordon pricing also applied to shiraz city in Iran, and the results show that the reward of the Markov process is maximized at the seasons including Spring, Autumn, and Winter. Further, the results are a significant effect on the total reward of the Markov process. Finally, the decision-making transportation management system is optimized in the case-study city during the numerical simulation. Suggestions for further research can be divided into two categories: making use of evolutionary optimization algorithms and fuzzy logic for the optimal dedication of traffic to transportation networks. Categorizing travelers based on their monthly income, having a driver's license, age, and other variables and providing a logistic regression model for each category and all modes of public and private transportation. The suggestions related to evolutionary optimization algorithms and fuzzy logic for the optimal dedication of traffic to transportation networks are as follows:

In the present study, genetic algorithms were used to acquire the optimal decision for the price level of CBDs in Shiraz. For future studies, it is recommended that other evolutionary algorithms such as

genetic and particle swarm optimization or synthetic optimization algorithms are used for optimizing the dedication of traffic to public and private roads.

Use of fuzzy logic in improving the performance of the proposed categories: Also, management of traffic in CBDs of a certain city can be improved using methods based on fuzzy logic and encourage travelers to commute through areas with less or light traffic.

Categorizing travelers based on their monthly income, having a driver's license, age, and other variables and providing a logistic regression model for each category and all modes of public and private transportation. In future studies, genetic algorithms can be used to provide a logistic regression model for each category of travelers based on source, destination, the goal of the journey, etc. and evaluate their tendency to use each mode of public and private transportation.

References

- [1] Ahuja, H., & Batra, U., 2018. Performance enhancement in requirement prioritization by using least-squares-based random genetic algorithm. In *Innovations in Computational Intelligence* (pp. 251-263). Springer, Singapore.
- [2] Amirgholy, M., Rezaeestakhrui, H., & Poorzahedy, H., 2015. Multi-objective cordon price design to control long run adverse traffic effects in large urban areas. *NETNOMICS: Economic Research and Electronic Networking*, 16(1-2), 1-52.
- [3] Azari, K.A., Arintono, S., Hamid, H., & Rahmat, R.A.O., 2013. Modelling demand under parking and cordon pricing policy. *Transport Policy*, 25, 1-9.
- [4] Broaddus, A., Litman, T. & Menon, G., 2009. Training document on transportation demand management. Sustainable Urban Transport Project (www.sutp.org) and GTZ (www.gtz.de).
- [5] Buontempo, F., 2019. Genetic algorithms and machine learning for programmers: create AI models and evolve solutions. The Pragmatic Bookshelf.
- [6] Cavallaro, F., Giaretta, F., & Nocera, S., 2018. The potential of road pricing schemes to reduce carbon emissions. *Transport Policy*, 67, 85-92.

- [7] Chen, Y., Lawell, C.Y.C.L., Muehlegger, E.J., & Wilen, J.E., 2017, Modeling Supply and Demand in the Chinese Automobile Industry. In 2017 Annual Meeting, July 30-August 1, Chicago, Illinois (No. 258557). Agricultural and Applied Economics Association.
- [8] Dai, H., Yao, E., & Zhao, R., 2015. Research on congestion pricing in multimode traffic considering delay and emission. *Discrete Dynamics in Nature and Society*.
- [9] De Palma, A., & Lindsey, R., 2009. Traffic congestion pricing methods and technologies.
- [10] Eiben, A.E. and Smith, J.E., 2003. Introduction to evolutionary computing (Vol. 53, p. 18). Berlin: springer.
- [11] Ferguson, E., 2018. Travel demand management and public policy. Routledge.
- [12] Ferrarese, M., 2016. A perceptual-behavioural approach with non-parametric experimental coefficient for urban parking business design. *Archives of Transport, 37(1), 15-30*.
- [13] Howard R. A., 1960, *Dynamic programming and markov processes*, (1th ed.). Cambridge: The MIT Press.
- [14] ITSR, (2001), The proposed bus transportation system for the City of Shiraz, Institute for Transportation Studies and Research of Tehran and Shiraz Comprehensive Transportation Studies, Report No. 80-08, Sharif University of Technology, Tehran (in Farsi).
- [15] Kelly, F.J., and Zhu, T., 2016. Transport solutions for cleaner air. *Science, 352(6288), 934-936*.
- [16] Kramer, O., 2017. Genetic algorithms. In *Genetic algorithm essentials* (pp. 11-19). Springer, Cham.
- [17] Lyons, T.J., Kenworthy, J.R., Moy, C. & Dos Santos, F., 2003. An international urban air pollution model for the transportation sector. *Transportation Research Part D: Transport and Environment, 8(3), 159-167*.
- [18] McCord, M.R., 1987. *Urban transportation networks: Equilibrium analysis with mathematical programming methods*: Prentice-Hall, Inc., Englewood Cliffs, NJ, USA.
- [19] Mirjalili, S. (2019). Genetic algorithm. In *Evolutionary algorithms and neural networks* (pp. 43-55). Springer, Cham.
- [20] Pal, S.K. & Wang, P.P., 2017. *Genetic algorithms for pattern recognition*. CRC press.
- [21] Poorzahedy H., AghaBabazadeh B., Babazadeh A., (2016), Dynamic network pricing to contain urban air pollution in stochastic environment, *Scientia Iranica, 23, 2005-2022*.
- [22] Sawicki, P., Kiciński, M., & Fierek, S., 2016. Selection of the most adequate trip-modelling tool for integrated transport planning system. *Archives of Transport, 37(1), 55-66*.
- [23] Sendek-Matysiak, E., 2019. Multi-criteria analysis and expert assessment of vehicles with different drive types regarding their functionality and environmental impact. *Scientific Journal of Silesian University of Technology. Series Transport, 102, 185-195*.
- [24] Verhoef, E.T., 2005. Second-best congestion pricing schemes in the monocentric city. *Journal of Urban Economics, 58(3), 367-388*.