ADVANCES IN TRANSPORTATION STUDIES
An International Journal

Editor in Chief: Alessandro Calvi

Vol. LXVII November 2025

Contents

aracne

Advances in Transportation Studies: an international Journal 67 (2025)

ISBN 979-12-218-2194-9 – ISSN 1824-5463 – DOI 10.53136/97912218219491 – pag. 3-16

An algorithm for the discrete network design problem without traffic assignment iterations

T. Xu Y. Cui A. Cheng Y. Mao

School of Civil and Transportation Engineering at Henan University of Urban Construction, China
email: selecxtz@126.com

subm. 18th December 2024 *approv. after rev. 12nd February 2025*

Abstract

Though the transportation network design problem (TNDP) has been studied widely due to its importance, the majority of past studies need to solve the traffic assignment problem many times in the solution process, which makes the methods inefficient. In this paper, an algorithm for the TNDP of adding a lane/lanes to a link on a one-to-one network without traffic assignment iterations is presented. First, for each link, the travel time drop (TTD) of the network when adding a lane to each link is estimated. The estimation of the TTD does not need solving the traffic assignment problem on the network. It just need to adjust flows on the used paths of the OD. Since it does not involve finding the shortest path, it is far more efficient than the methods that need traffic assignment iterations. Then the link with the maximal TTD is selected for capacity expansion by adding a lane or lanes to it. At the end of each step, the TTD of the network is estimated again and a link is selected again for expansion. The process proceeds until the budget is met. The method is applied to a network and the validity of it is shown. The algorithm has two pros: 1) It can find the exact optimal solution for a one-to-one network; 2) It avoids UE iterations. The con of the algorithm is that it can only find the approximate optimal solution for multiple-origin-multiple-destination network. It helps to develop the approximate algorithm for a multiple-origin-multiple-destination (many-to-many) network.

Keywords – capacity expansion, discrete network design problem, travel time drop

1. Introduction

The transportation network design problem (TNDP) is a very important question in transportation engineering. It can be divided into three categories: continuous network design problem (CNDP) [1] - [42], discrete network design problem (DNDP) [43] - [62], and mixed network design problem (MNDP) [63] - [66]. The CNDP requires the decision variable to be continuous, such as setting signal time and link tolls. The DNDP deals with selecting links from a potential link set to add to an existing network, adding a lane/lanes to a link of the network, or selecting a subset of links to expand their capacity, etc., under certain budget. The MNDP combines the former two. Since the CNDP requires the decision variable to be continuous, the method for the CNDP does not apply to the DNDP and MNDP.

According to the Highway Capacity Manual (HCM) [68], the capacity of a road is defined and measured by lane. Other factors, including road condition (lane width, lateral clearance, heavy vehicle, sight distance, environment) and traffic condition, can only be adjustors influencing the capacity at a limited level. Thus the fundamental way to increase the capacity of a street is to add lanes. Accordingly, the expansion of the capacity of a street cannot be continuous. In this sense, all the studies of capacity expansion by the CNDP do not apply to adding lanes to a link. In addition, for the DNDP, adding links to an existing network will change the topology of the network whereas adding a lane/lanes to a link will not change the topology. When the topology of the network is changed, the problem of the Braess's paradox may occur [69]. So the methods for them are different. As is covered in the literature review, the studies on the CNDP are a lot, the studies on adding links to an existing network are fewer, but there is only one study on adding a lane/lanes to a link [64]. The study in [64] needs to solve the user-equilibrium (UE) problem many times in the solution process, which makes the method inefficient.

The biggest challenge of the TNDP lies in how to select the links to do work on (adding lanes, setting signal, toll, etc.) efficiently especially for real size networks. In the majority of past studies on the TNDP, selecting the links needs to solve the traffic assignment problem many times in the solution process. It is well known that solving the traffic assignment problem is very time-consuming especially for large networks. To increase the efficiency in selecting the links, many studies have been conducted to decrease the number of the iterations for solving the traffic assignment problem (mostly the user-equilibrium problem and is abbreviated as the UE iterations hereafter). However, none of the past studies on the DNDP can avoid the multiple UE iterations in the solution process. In this paper, we presented an algorithm for the DNDP of adding a lane/lanes to links without UE iterations. It can find the exact optimal solution for a one-origin-one-destination (one-to-one) network.

The rest of the paper is arranged as follows: Section 2 covers the literature review of this study. Section 3 presents the method for the DNDP for one-to-one networks without multiple UE iterations. Section 4 shows the validity of the algorithm by applying it to a simulation network. Section 5 presents the conclusion of the paper.

2. Literature review

Most studies on the CNDP used a bi-level model, where the upper level is to minimize total travel cost of road network plus construction costs or other indexes, and the lower level is a traffic assignment model. Different algorithms were used to solve the models. They include Genetic algorithm [1] - [4], Perturbation based approach [5], Gradient Projection (GP), Conjugate Gradient projection (CG), Qusai-Newton projection (QNEW), and PARATAN version of gradient projection (PT)[6], Hooke–Jeeves algorithm [7], Equilibrium Decomposed Optimization (EDO) algorithm [8], path based mixed-integer linear program [9], Modified Differential Evolution (MODE) algorithm [10], The enhanced differential evolution algorithm based on multiple improvement strategies (EDEMIS) [11], Norm-relaxed method of feasible direction (NRMFD) algorithm integrated with the Euler-based approximation (EBA) method [12], partial linearized subgradient methods [13], Particle swarm optimization (PSO) [14], Sensitivity analysis [15] - [25], Iterative optimization-assignment algorithm [26] - [29], Simulated annealing (SA) and Sensitivity analysis [30] - [31], SA and GA [32] , Hybrid GASA algorithm [33], Global optimization method [34], etc. Some studies on the CNDP used a one-level model and different algorithms to solve the models. They include Relaxation algorithm [35], Multi-cutting plane method [36], Frank–Wolfe algorithm based method [37], LINDOGLOBAL in GAMS [38], Augmented Lagrangian method [39], optimization method [40], CPLEX solver [41], Sensitivity analysis with sequence average algorithm [42], etc. Since the CNDP requires the decision variable to be continuous, the method for the CNDP does not apply to the DNDP and MNDP. In addition, most of them require multiple UE iterations in the solution process.

Most studies on the DNDP used a bi-level model and different algorithms to solve the models. They include Multi-objective genetic algorithm (MOGA) [43], Genetic algorithm (GA) [44], Augmented Frank–Wolfe based algorithm [45], Branch-and-bound techniques [46] - [50], Lagrange relaxation and dual ascent process [51], Decomposition with quasi-optimization [52], Support function concept [53], Ant system method [54], Genetic algorithm [55] - [56], Hybrid meta-heuristic algorithm [57], Parsimonious heuristic [58], Global optimization methods [59] - [61]. One study on the DNDP used a one-level model and Outer Approximation algorithm [62] to solve the model. However, all the above methods on the DNDP require multiple UE iterations in the solution process. None of them can avoid the multiple UE iterations in the solution process.

Different algorithms were used to solve the MNDP. They include Dimension-down iterative algorithm [63], Hill Climbing, Simulated Annealing, Tabu Search, Genetic Algorithm, Hybrids of Tabu Search [64], Scatter Search [65], and mixed-integer linear programming approach [66]. However, all the above methods on the MNDP require multiple UE iterations in the solution process. None of them can avoid the multiple UE iterations in the solution process.

The literature review for this study indicated that there was no study on the DNDP that can avoid the multiple UE iterations. Our paper is the first to study the DNDP on a one-to-one network without the multiple UE iterations.

3. The DNDP model and the algorithm for it

3.1. The DNDP

To present the algorithm for the DNDP, the following notations are defined.

A: the set of links in the network.
R: the set of origins.
S: the set of destinations.
Q: the set of OD demand, $Q = [q_{rs}], \forall\, r \in R, s \in S$.
K_{rs}: the set of used paths between OD pair rs, $\forall\, r \in R, s \in S$.
q_{rs}: the total demand of OD pair rs.
f_k^{rs}: the flow on the kth path of OD pair rs, $k \in K_{rs}$, $\forall\, r \in R, s \in S$.
$\mathbf{f}$: the vector of path flows, $\mathbf{f} = [f_k^{rs}]$, $k \in K_{rs}$, $\forall\, r \in R, s \in S$.
p_k^{rs} is the kth used path of OD pair rs, $k \in K_{rs}$, $\forall\, r \in R, s \in S$.
P: the set of used paths, $P = [p_k^{rs}]$.
x_a: the volume on link a, $\forall\, a \in A$.
$\mathbf{x}$: the vector of equilibrium link flows, $\mathbf{x} = [x_a]$, $\forall\, a \in A$.
t_a: the travel time on link a, $\forall\, a \in A$.
$\mathbf{t}$: the vector of link travel times, $\mathbf{t} = [t_a(x_a)]$ $\forall\, a \in A$.
c_k^{rs}: the travel time on the kth path connecting origin r and destination s, $k \in K_{rs}$, $\forall\, r \in R, s \in S$.
$\delta_{a,k}^{rs}$: the link-path incidence indicator, $\delta_{a,k}^{rs} = 1$ if link a is on the kth route connecting origin r and destination s, and $\delta_{a,k}^{rs} = 0$ otherwise.
c_a: the capacity of link a, $\forall\, a \in A$.
d_a: the capacity of a lane on link a, $\forall\, a \in A$.
n_a: the number of lanes to be added to link a, $n_a = 0, 1, 2, \dots \dots$, $\forall\, a \in A$.
n: the vector of the number of lanes added to the links, $\mathbf{t} = [n_a]$, $\forall\, a \in A$.
w_a: the maximal number of lanes allowed to be added to link a, $n_a = 0, 1, 2, \dots \dots$, $\forall\, a \in A$.
B_a: the cost of adding a lane to link a, $\forall\, a \in A$.
B: the total budget for the capacity expansion.

g_k^{rs}: the subtracted flow on the kth path connecting origin r and destination s, $k \in K_{rs}$, $\forall\, r \in R, s \in S$.
h_k^{rs}: the added flow on the kth path connecting origin r and destination s, $k \in K_{rs}$, $\forall\, r \in R, s \in S$.
m_k^{rs}: the change of travel time on the kth path connecting origin r and destination s incurred by the infinitesimal change of path flow f_k^{rs}, $k \in K_{rs}$, $\forall\, r \in R, s \in S$.
m_a: the change of travel time on link a incurred by the infinitesimal change of link flow x_a, $\forall\, a \in A$.
H_k^{rs}: the set of the links on a used path p_k^{rs}, $k \in K_{rs}$, $\forall\, r \in R, s \in S$.
H^{rs}: the set of the links on all the used paths of OD rs, $\forall\, r \in R, s \in S$.
H: the set of the links on all the used paths of all ODs.

Since only a one-to-one network is studied in this paper, the set R and S contain an origin and a destination, respectively.

A link with nonzero link volume at the UE status is a used link. Likewise, a path with nonzero path flow at the UE status is a used path. Obviously, H^{rs} is the union of H_k^{rs}, or $H^{rs} = \bigcup_k H_k^{rs}$. H^{rs} is a subset of A, where A is the set of links on the network. Similarly, H is the union of H^{rs}, or $H = \bigcup_{r,s} H^{rs}$.

The link travel time is estimated by the following popularly used BPR function

$$t_a(x_a) = t_a^0(1 + \beta(x_a/c_a)^\alpha) \tag{1}$$

where t_a is travel time, x_a is volume, c_a is capacity, t_a^0 is free flow travel time, β is 0.15, and α is 4.

The bi-level DNDP model of adding lanes to links is defined as follows:
The upper level model is given as

$$\min_{\mathbf{x},\mathbf{n}} Z_1\,(\mathbf{x},\mathbf{n}) = \sum_a x_a t_a^0(1 + \beta(x_a/(c_a + d_a n_a))^\alpha) \tag{2}$$

subject to

$$\sum_a B_a n_a \le B \qquad \forall\, a \in A \tag{3}$$

$$n_a \le w_a \qquad \forall\, a \in A \tag{4}$$

and the lower level model which is given as

$$\min_{\mathbf{x}} Z_2\,(\mathbf{x}) = \sum_a \int_0^{x_a} t_a^0(1 + \beta(\omega/(c_a + d_a n_a))^\alpha)\, d\omega \tag{5}$$

subject to (5), (6), and (7) as follows:

$$\sum_k f_k^{rs} = q_{rs} \qquad \forall\, r \in R, s \in S \tag{6}$$

$$x_a = \sum_{rs} \sum_{k \in K_{rs}} f_k^{rs}\, \delta_{a,k}^{rs} \qquad k \in K_{rs}, \forall\, r \in R, s \in S \tag{7}$$

$$f_k^{rs} \ge 0 \qquad k \in K_{rs}, \forall\, r \in R, s \in S \tag{8}$$

(2) is to minimize the total travel time of the network, which is also the objective function of the system optimum (SO) of traffic assignment problem. (3) states that the total cost of capacity expansion should not exceed the total budget B. B_a is the cost of adding a lane to link a. It is the length of link a times the unit cost of adding a lane to link a. The unit cost of adding a lane to a link can be different for different links but it is set equal in this study without loss of generality. In the real world, there are different types of roads (local, collectors, etc.) and capacity issues in road links under various hierarchical categories are treated differently. For example, capacity increase in a collector road is not given the same priority as increasing the capacity of a highway. (4) is used to reflect such prioritization in adding lanes to the links. (5) – (8) is the UE traffic assignment problem. (5) is the objective function of the user optimum (UE) of traffic assignment problem. c_a is the capacity of link a. d_a is the capacity of adding a lane on link a. Both c_a and d_a are given in practice. x_a is the volume on link a. n_a is the number of lanes to be added to link a. x_a and n_a are

variables that needs to be solved for. ω is the integral variable. (6) is the OD demand conservation constraint. It states that the sum of the flows on all the paths of OD pair rs is the demand of the OD pair. (7) means that the link flow on the link a equals the sum of the flows of all the paths passing through the link. (7) indicates that all the path flows are nonnegative.

Note that $\mathbf{n}$ is not variables in the lower level UE model. The UE program is solved for the link or path flows given $\mathbf{n}$. Note again that Z_1 is a non-increasing function of $\mathbf{n}$ because more lanes always increase the link capacity and lower the travel time on the link and the network as long as the topology of the network is fixed (No links are added to or deleted from the transportation network. In this paper, we always assume the topology of the network is fixed). Thus, if we can add a lane one by one and guarantee the drop of Z_1 is the maximum when each lane is added, the accumulation of the drop will be the maximum under the budget. For this purpose, we need to estimate the drops of Z_1 when a lane is added to all links and select the link with the maximal drop of Z_1 to add a lane.

Since n_a is a non-negative integer and is not continuous, the gradient of Z_1 with respect to n_a does not exist. To express the change of the value of Z_1 with respect to the change of n_a, the travel time drop (TTD) of Z_1 with respect to n_a is given as follows

$$\frac{\Delta Z_1}{\Delta n_a} = \sum_a \left[x_a t_a^0 (1 + \beta(x_a/(c_a + d_a n_a))^\alpha) - \bar{x}_a t_a^0 \left(1 + \beta\left(\bar{x}_a/\left(c_a + d_a(n_a + 1)\right)\right)^\alpha\right)\right] \quad (9)$$

where x_a and $\bar{x}_a$ are the link flows at the UE status when the added lanes on link a are n_a and $(n_a + 1)$, respectively. In other words, x_a and $\bar{x}_a$ are the link flow solution for the lower level UE program (5) – (8) when the number of lanes to be added to link a is set as n_a and $(n_a + 1)$, respectively. Since adding a lane to a used link on a one-to-one network always lowers the travel time or the value of Z_1, we have $\Delta Z_1/\Delta n_a \geq 0, \forall\, a \in A$. The TTD $\Delta Z/\Delta n_a$ indicates the drop of the total travel time on the network at the UE status when a lane is added to link a.

To guarantee the drop of Z_1 is the maximum when each lane is added, we just need to select the link with the maximal value of the TTD to add a lane. However, estimating the TTD needs to know the value of Z_1 when the number of lanes to be added is taken as n_a and $(n_a + 1)$, respectively. The only way to estimate the value of Z_1 is to solve the UE problem (5) – (8) given n_a and $(n_a + 1)$, respectively, and substitute the resultant x_a and $\bar{x}_a$ and the corresponding link travel time in (2). Solving the UE problem is very time-consuming even for one-to-one networks because it needs to find the shortest path many times in the solution process. It is better if the TTD can be estimated without the need of solving the UE problem.

3.2. The exact algorithm for the DNDP on a one-to-one network

For a one-to-one network, we have the following statement.

Proposition 1: Adding a lane to a link in set H^{rs} decreases the value of Z_1. Meanwhile, it does not change H^{rs}, where H^{rs} is the set of the links on all the used paths of OD rs. In other words, all the unused links will keep unused and still have zero-flow on them.

Proof: Adding a lane to link a in H^{rs} always increase the capacity of the link and decrease the travel time on the link. At the UE status, all the used paths have equal travel time. A drop of the travel time on link a will lower the travel time on the paths consisting of link a. These paths will have lower travel time than other used paths that do not consist of link a. To keep equilibrium, a portion of the flow on the latter will shift to and only to the former. The travel time on the latter will be lowered too. Since the flow on the latter shift to and only to the former, all the unused links will keep unused and still have zero-flow on them. At the new UE status, the travel time on all the used paths are lower than their original travel time. The demand is the sum of the path flows and is the same before and after a lane is added. Since Z_1 equals the sum of the product of the path flow and path time, the value of Z_1 will decrease.

Based on proposition 1, if we can adjust the flow of the used paths to reach a new UE status when a lane is added to a link in H^{rs}, we can obtain the value of Z_1 and estimate the TTD without finding the shortest paths. Below we present an algorithm for estimating the TTD by adjusting the flow of the used paths on a one-to-one network without UE iterations.

For the convenience of illustration, denote c_{k,n_a}^{rs} and c_{k,n_a+1}^{rs} as the travel time on path p_k^{rs} at the UE status when the number of lanes to be added to link a are n_a and (n_a+1), respectively. Likewise, denote $t_b^{n_a}$ and $t_b^{n_a+1}$ as the travel time on link b for n_a and (n_a+1), $\forall\, b \in H^{rs}$, respectively.

Suppose the link volume and path flow of the UE problem given n_a are known. The algorithm of adjusting the flow of the used paths to reach a new UE status for (n_a+1) is given as follows.

Algorithm 1

Step 1: Select a link in H^{rs}, for example, link a, $a \in H^{rs}$, and add a lane to link a. Calculate the travel time on it based on the link travel time function $t_a^{n_a+1} = t_a^0\left(1+\beta\left(\omega/\left(c_a+d_a(n_a+1)\right)\right)^{\alpha}\right)$.

Step 2: Update the path travel time on all the paths that consist of link a based on $t_a^{n_a+1}$ to obtain the updated path travel time $c_{k,mid}^{rs}$, i.e., update the path travel time for p_k^{rs} as $c_{k,mid}^{rs}$ where $k \in K_{rs}$ and $\delta_{a,k}^{rs}=1$. Pick any one of them and set it as the minimum path p_{kmin}^{rs}. Pick any one of the used paths that do not consist of link a based on $t_a^{n_a}$ and set it as the maximum path p_{kmax}^{rs}.

Step 3: Set an counter $m:=1$. Denote p_{kmin}^{rs} and p_{kmax}^{rs} as $p_{kmin}^{rs(m)}$ and $p_{kmax}^{rs(m)}$, respectively. Let $c_{kmin}^{rs(m)}$ and $c_{kmax}^{rs(m)}$ be the path cost for the minimum path $p_{kmin}^{rs(m)}$ and the maximum path $p_{kmax}^{rs(m)}$, respectively.

Step 4: Calculate the partial derivative of the travel time functions for links not common on both $p_{kmin}^{rs(m)}$ and $p_{kmax}^{rs(m)}$, i.e., links where $\delta_{b,kmin}^{rs}=1$ or $\delta_{b,kmax}^{rs}=1$ but $\delta_{b,kmin}^{rs}\delta_{b,kmax}^{rs}=0$. Sum up the partial derivatives for $p_{kmin}^{rs(m)}$ and $p_{kmax}^{rs(m)}$, respectively, as follows

$$s_{kmin}^{rs(m)} = \sum_{b\in H_{kmin}^{rs}} \frac{\mathrm{d}t_b}{\mathrm{d}x_b}\left(\delta_{b,kmin}^{rs}-\delta_{b,kmax}^{rs}\right)^2 \tag{10}$$

and

$$s_{kmax}^{rs(m)} = \sum_{b\in H_{kmmx}^{rs}} \frac{\mathrm{d}t_b}{\mathrm{d}x_b}\left(\delta_{b,kmin}^{rs}-\delta_{b,kmax}^{rs}\right)^2 \tag{11}$$

Step 5: Calculate the flow to be shifted from the maximum path $p_{kmax}^{rs(m)}$ to the minimum path $p_{kmin}^{rs(m)}$.

Step 5.1: Set an counter $l:=1$.

Step 5.2: Calculate the initial flow to be shifted from the maximum path $p_{kmax}^{rs(m)}$ to the minimum path $p_{kmin}^{rs(m)}$ as follows

$$\Delta f_{rs}^{(l)} = \left(c_{kmax}^{rs(m)}-c_{kmin}^{rs(m)}\right)/s_{kmin}^{rs(m)} \tag{12}$$

Step 5.3: Shift $\Delta f_{rs}^{(l)}$ of flow from the maximum path $p_{kmax}^{rs(m)}$ to the minimum path $p_{kmin}^{rs(m)}$ and update the link volumes on all the links on path $p_{kmax}^{rs(m)}$ and $p_{kmin}^{rs(m)}$ as follow: Set $x_b := x_b + \Delta f_{rs}^{(l)}$, $\forall\, b \in H_{kmin}^{rs}$; $x_b := x_b - \Delta f_{rs}^{(l)}$, $\forall\, b \in H_{kmmx}^{rs}$.

Step 5.4: Estimate the link travel time on all the links on path $p_{kmax}^{rs(m)}$ and $p_{kmin}^{rs(m)}$ based on the

updated link volumes. Compare the path travel time on path $p_{kmax}^{rs(m)}$ and $p_{kmin}^{rs(m)}$. If $\left(c_{kmax}^{rs(m)} - c_{kmin}^{rs(m)}\right) < \varepsilon$, where ε is a preset small figure, go to Step 7; If $c_{kmax}^{rs(m)} < c_{kmin}^{rs(m)}$, update the counter $l := l + 1$.let $\Delta f_{rs}^{(l)} := 0.5\Delta f_{rs}^{(l-1)}$ and go to Step 5.3; If $c_{kmax}^{rs(m)} > c_{kmin}^{rs(m)}$, proceed ;
Step 6: Based on the updated link volume, do the same thing as Step 4 and obtain $s_{kmin}^{rs(m)}$ and $s_{kmax}^{rs(m)}$. Go to Step 5.
Step 7: Update the path travel time on all the used paths. Update the counter $m := m + 1$.
Step 8: Pick a used path with the lowest travel time (Pick any one of them if there are multiple used paths with the lowest travel time) and set it as the minimum path $p_{kmin}^{rs(m)}$. Pick a used path with the largest travel time (Pick any one of them if there are multiple used paths with the largest travel time) and set it as the maximum path $p_{kmax}^{rs(m)}$.
Step 8: Convergence test. If $\left(c_{kmax}^{rs(m)} - c_{kmin}^{rs(m)}\right) < \varepsilon$, stop; otherwise, go to Step 4 and repeat the process.

Proposition 2: The path flow found by Algorithm 1 is a solution of the UE problem.

Proof: To prove the new path flow found by Algorithm 1 is a solution of the UE problem, we only need to show the path time on the used paths are equal. Since there are very limited number of used paths in the solution of the UE problem, as long as the travel time of $p_{kmax}^{rs(m)}$ and $p_{kmin}^{rs(m)}$ can be equalized, the travel time of all the used paths will be equalized sequentially. Therefore, we only need to show the travel time of $p_{kmax}^{rs(m)}$ and $p_{kmin}^{rs(m)}$ can be equalized. Since $p_{kmax}^{rs(m)}$ and $p_{kmin}^{rs(m)}$ have equal travel time at convergence, we only need to show the algorithm will converge, or $\lim_{m\to\infty}\left(c_{kmax}^{rs(m)} - c_{kmin}^{rs(m)}\right) = 0$.

When the shifted flow $\Delta f_{rs}^{(l)}$ is taken as the variable, all the links on path $p_{kmax}^{rs(m)}$ or $p_{kmin}^{rs(m)}$ have the single variable $\Delta f_{rs}^{(l)}$. The path time is the summation of the travel times on all the links on the path. Since the BPR link travel time functions are convex functions and the summation of multiple convex functions is still a convex function, the path travel time function is also a convex function with variable $\Delta f_{rs}^{(l)}$. The partial derivative of the link travel time functions with respect to link volume is the same as that with respect to the shifted flow $\Delta f_{rs}^{(l)}$. $s_{kmin}^{rs(m)}$ is the tangent or slope of the path travel time function for path $p_{kmin}^{rs(m)}$ at $\Delta f_{rs}^{(l)} = 0$. $s_{kmax}^{rs(m)}$ is the tangent or slope of the path travel time function for path $p_{kmax}^{rs(m)}$ at $\Delta f_{rs}^{(l)} = 0$. Since the path travel time functions for both $p_{kmax}^{rs(m)}$ and $p_{kmin}^{rs(m)}$ are convex functions, both $\left(c_{kmax}^{rs(m)} - c_{kmin}^{rs(m)}\right)/s_{kmax}^{rs(m)}$ and $\left(c_{kmax}^{rs(m)} - c_{kmin}^{rs(m)}\right)/s_{kmin}^{rs(m)}$ are likely to overestimate the flow to be shifted from $p_{kmax}^{rs(m)}$ to $p_{kmin}^{rs(m)}$. (If the link travel time functions are linear, $\frac{dt_b}{dx_b}$ is a constancy. $\left(s_{kmax}^{rs(m)} + s_{kmin}^{rs(m)}\right)$ is the decrease of the difference of the travel time on $p_{kmax}^{rs(m)}$ and $p_{kmin}^{rs(m)}$ when a unit of flow is shifted from $p_{kmax}^{rs(m)}$ to $p_{kmin}^{rs(m)}$. Since the original difference of the travel time on $p_{kmax}^{rs(m)}$ and $p_{kmin}^{rs(m)}$ is $\left(c_{kmax}^{rs(m)} - c_{kmin}^{rs(m)}\right)$, dividing $\left(c_{kmax}^{rs(m)} - c_{kmin}^{rs(m)}\right)$ by $s_{rs}^{(m)}$ gives the value of the flow that should be shifted from $p_{kmax}^{rs(m)}$ to $p_{kmin}^{rs(m)}$ to vanish the difference of the travel time on $p_{kmax}^{rs(m)}$ and $p_{kmin}^{rs(m)}$. The link travel time functions in this paper are nonlinear. We just show this fact to help explain the proof.). We take either $\left(c_{kmax}^{rs(m)} - c_{kmin}^{rs(m)}\right)/s_{kmax}^{rs(m)}$ or $\left(c_{kmax}^{rs(m)} - c_{kmin}^{rs(m)}\right)/s_{kmin}^{rs(m)}$ as the initial flow to be shifted from $p_{kmax}^{rs(m)}$ to $p_{kmin}^{rs(m)}$. If $c_{kmax}^{rs(m)} < c_{kmin}^{rs(m)}$ after shifting the flow by $\Delta f_{rs}^{(l)}$, it means overestimating the flow to be shifted. We then halve the flow to be shifted by letting $\Delta f_{rs}^{(l)} := 0.5\Delta f_{rs}^{(l-1)}$. And go to Step 5.3 to try if $c_{kmax}^{rs(m)} > c_{kmin}^{rs(m)}$ holds or not. If it does not, we continue to halve the flow to be shifted until $c_{kmax}^{rs(m)} > c_{kmin}^{rs(m)}$ holds. If it does, it means the amount of the flow to be shifted is suitable and is the maximal possible. We then update the path flow by $\Delta f_{rs}^{(l)}$, update the volume and travel time on the links on $p_{kmax}^{rs(m)}$ to $p_{kmin}^{rs(m)}$, compare the updated travel time $c_{kmax}^{rs(m)}$ and $c_{kmin}^{rs(m)}$, and iterate the process. Since the difference of the path travel time on $p_{kmax}^{rs(m)}$ and $p_{kmin}^{rs(m)}$ will be reduced at a pace close to half of $\left(c_{kmax}^{rs(m)} - c_{kmin}^{rs(m)}\right)$ at each iteration, we have $\lim_{l\to\infty}\left(c_{kmax}^{rs(m)(l)} - c_{kmin}^{rs(m)(l)}\right) = \lim_{l\to\infty}\left(\frac{1}{2}\right)^{l}\left(c_{kmax}^{rs(m)(0)} - c_{kmin}^{rs(m)(0)}\right) = 0$. In other words, the algorithm will converge. This completes the proof.

Based on Proposition 2, the new path flow found by the Algorithm 1 is the solution of the UE problem when the number of lane to be added to link a is set as $(n_a + 1)$. With the path flow and path time, we can easily get the travel time on the network or the value of Z_1 for $(n_a + 1)$. Since we have already known the solution of the UE problem when the number of lane to be added to link a is n_a, i.e., we have known the value of Z_1 for n_a, the difference of the values of Z_1 for $(n_a + 1)$ and n_a is the TTD.

Given the solution of the UE problem when no link has been added to any link, the TTD for each link can be estimated by using Algorithm 1. The first lane should be added to the link with the largest TTD. Only by doing so can the value of Z_1 be lowered maximally. After a lane has been added to the selected link, the TTD for each link can be estimated again, the second lane should be added to the link with the largest TTD, etc. This process iterates until the budget is exceeded. The idea is summarized as the following algorithm for the DNDP of adding lanes to links on a one-to-one network.

Algorithm 2

Step 1: Solve the UE problem when $n_a = 0, \forall\, a \in A$. This gives $x_a, \forall\, a \in A$, and $f_k^{rs},\ k \in K_{rs}$, $\forall\, r \in R, s \in S$.

Step 1: Set an counter $m := 1$. Set the accumulated cost $D = 0$.

Step 3: Pick a link in H^{rs}, add a lane to the link, and apply Algorithm 1 to the network with the added link to find the UE solution.

Step 4: Estimate the TTD of the link selected in Step 3.

Step 5: Remove the link added in Step 3. Go to Step 3 to repeat the process until the TTD of all the links in H^{rs} are estimated.

Step 6: Select the link with the highest TTD and add a lane to the link. If the number of the added lanes to the link exceeds the maximal number of lanes allowed to be added to the link, i.e., if $n_a > w_a$, deselect the current link and select the link with the second highest TTD and add a lane to the link, etc. Estimate the cost of adding the lane (it is simply the length of the link times the unit cost of adding the lane) and add the cost to D. If $D >$ B, stop; otherwise, set $m := m + 1$, go to Step 3 and repeat the process.

Note that the added link in Step 6 should not be removed in the subsequent process. Note also that the link to be picked in Step 3 includes the links that have been added a link in Step 6. This means a link may be added multiple lanes.

Proposition 3: The solution found by Algorithm 2 is a the minimum of the DNDP.

Proof: Step 6 of the algorithm selects the link with the highest TTD to add a lane. This means that the drop of Z_1 is maximal possible when each lane is added. Since Z_1 is a decreasing function of n_a, $\forall\ a \in H^{rs}$, the value of Z_1 decreases with the increase of n_a. Based on Algorithm 2, the value of Z_1 decreases with the maximal possible ladder-wise drop when each lane is added until no more lanes can be added to any link due to the limitation of the budget. Thus the solution found by Algorithm 2 is a the minimum of the DNDP.

It is obvious that Algorithm 2 will find the optimal solution of the DNDP model.

4. Numerical example

In this section, the above DNDP model and algorithm are applied to the network shown in Fig. 1. There are 14 links and 8 nodes in the auto network. The length of the links and the capacity of the links are shown in Table 1. The origin is node 1 and the destination is node 8. The OD demand is 8000 vehicles. The total budget is B =50,500,000 ($). The unit cost of adding a lane is 500,000 ($/mile/lane). The free flow speed is 40 miles/hour for all links. The capacity of a lane is set as 1400 vehicles/hour/lane. Suppose the maximal number of lanes to be added to all the links are 5. Two cases are considered. In case 1, the method presented in this paper is applied to the problem. In case 2, Algorithm 1 in the method is replaced by the UE iterations.

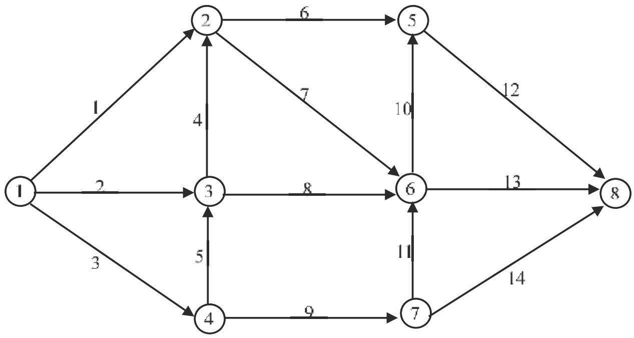


Fig. 1 - Network for the example

Tab. 1 - The length and the capacity for each link

a	l_a	c_a	a	l_a	c_a	a	l_a	c_a
1	8	2800	6	7	4200	11	7	2800
2	7	5600	7	6	2800	12	8	4200
3	5	4200	8	5	2800	13	7	5600
4	5	4200	9	8	4200	14	6	5600
5	3	2800	10	9	4200			

l_a is the link length. c_a is the link capacity.

The algorithm was programmed in Matlab 2022 and run on a computer with 2.8G processor and 8G RAM. The convergence criterion is set as 0.0001. The total computation time for case 1 and case 2 are 0.73814 seconds and 1.6032 seconds, respectively. The latter is two times more than the former. This indicates that Algorithm 1 is more efficient than the UE iterations.

The objective function value at each iteration in both cases is the same and is shown in Fig. 2. As can be seen in the figure, the objective function value, or the total travel time, decreases gradually from iteration to iteration. The link marginal cost in both cases is shown in Fig. 3. The link Volume/Capacity ratio in both cases is shown in Fig. 4. The number of lanes added to a link in both cases is shown in Fig. 5. As can be seen in the figures, the link marginal cost, the link Volume/Capacity ratio, and the number of lanes added to a link in both cases are all the same. This indicates that Algorithm 1 and the UE iterations are equivalent.

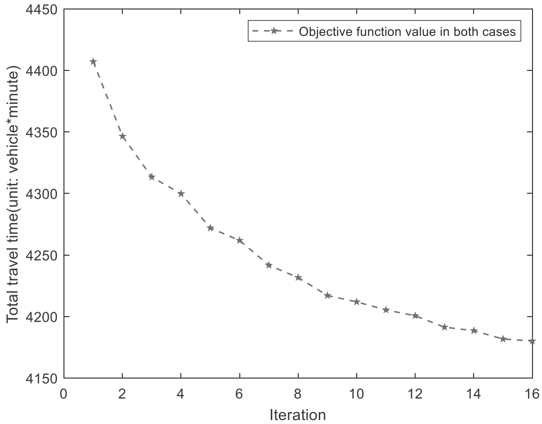


Fig. 2 - The objective function value in both cases

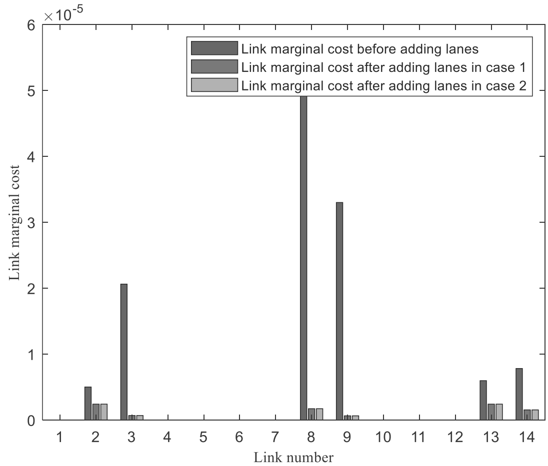


Fig. 3 - The link marginal cost in both cases

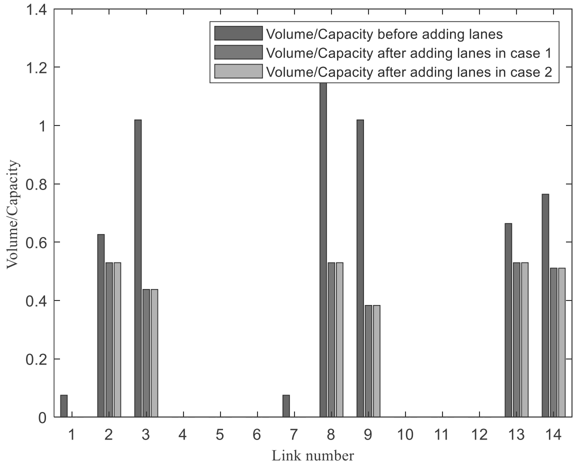


Fig. 4 - The link Volume/Capacity ratio in both cases

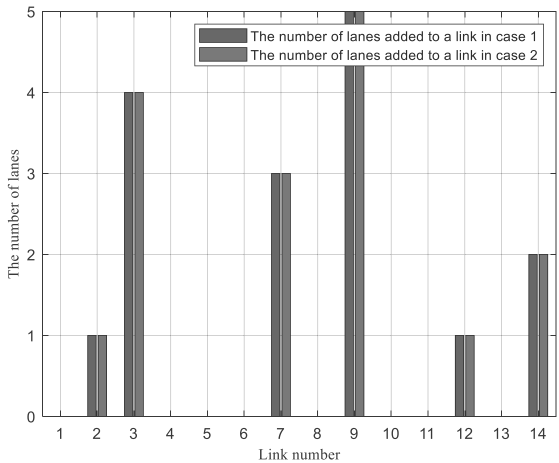


Fig. 5 - The number of lanes added to a link in both cases

5. Conclusions

The fundamental way to increase the capacity of a street is to add lanes. Accordingly, the expansion of the capacity of a street cannot be continuous. The method for the CNDP do not apply to adding lanes to a link. Current study on adding a lane/lanes to a link needs many rounds of UE iterations in the solution process, which makes the method inefficient.

In this paper, an algorithm for the DNDP of adding a lane/lanes to a link on a one-to-one network without UE iterations is presented. It does not involve finding the shortest path and is far more efficient than the methods that need UE iterations. The algorithm has two pros: 1) It can find the exact optimal solution; 2) It avoids UE iterations. The con of the algorithm is that it applies to a one-to-one network only. The method presented in this paper can be extended to a one-to-many network or to a multiple-origin-multiple-destination (many-to-many) network. Based on it, we can further develop the approximate algorithm for the DNDP on a many-to-many transportation network.

References

1. Zhou, Z.; Yang, M.; Sun, F.; Wang, Z.; Wang, B. (2021) A Continuous Transportation Network Design Problem with the Consideration of Road Congestion Charging. Sustainability 2021, 13, 7008. https://doi.org/10.3390/su13137008
2. Li, X.; Lang, M. (2014) Continuous network design based on the paired combinatorial logit stochastic user equilibrium model(Article). Open Electrical and Electronic Engineering Journal. 8(1): 316-322.
3. Sun, ZC (Sun, Zhichen). (2016) Continuous Transportation Network Design Problem Based on Bi-level Programming Model. Green intelligent transportation system and safety. 2016: 277-282.
4. Sumalee A (2007). Multi-concentric optimal charging cordon design. Transportmetrica 3: 41–71.
5. Msigwa, RE (Msigwa, Robert Ebihart); Lu, Y (Lu, Yue); Xiao, XT (Xiao, Xiantao); Zhang, LW (Zhang, Liwei). (2015) A perturbation-based approach for continuous network design problem with emissions(Article). Numerical Algebra, Control and Optimization. 5(2): 135-149.
6. Chiou SW (2005). Bilevel programming for the continuous transport network design problem. Transportation Research Part B 39: 361–383.
7. Abdulaal M, LeBlanc LJ (1979). Continuous equilibrium network design models. Transportation Research Part B 13: 19–32.
8. Suwansirikul C, Friesz TL, Tobin RL (1987). Equilibrium decomposed optimization: a heuristic for the continuous equilibrium network design problems. Transportation Science 21: 254–263.
9. Wang DZW, Lo HK (2010). Global optimum of the linearized network design problem with equilibrium flows. Transportation Research Part B 44: 482–492.
10. Baskan, Ozgur; Ceylan, Huseyin. (2014) Modified Differential Evolution Algorithm for the Continuous Network Design Problem. Procedia - Social and Behavioral Sciences. 111(0): 48-57.
11. Baskan, O; Ozan, C; Dell'Orco, M; Marinelli, M. (2018) Improving the performance of the bilevel solution for the continuous network design problem. Promet – Traffic & Transportation, 30(6): 709-720
12. Wang, Jian; He, Xiaozheng; Peeta, Srinivas; Wang, Wei. (2022) Globally convergent line search algorithm with Euler-based step size-determination method for continuous network design problem. Transportation Research Part B: Methodological. 163: 119-144.
13. Kalantari, Navid; Mirzahossein, Hamid; Najafi, Pooyan; Waller, Travis; Zhang, Xiang. (2023) Continuous Network Design Using Partial Linearized Subgradient Methods. Transportation Research Record. 2677(8): 633-643.
14. Wu, JJ (Wu, Jianjun); Guo, X (Guo, Xin); Sun, HJ (Sun, Huijun); Wang, B (Wang, Bo). (2014) Topological Effects and Performance Optimization in Transportation Continuous Network Design. Mathematical Problems in Engineering. 2014(1): 8.
15. Lim, Yongtaek; Heydecker, Benjamin G; Lee, Seungjae. (2005) A Continuous Network Design Model in Stochastic User Equilibrium Based on Sensitivity Analysis. *Journal of Advanced Transportation*. 39(1): 63-79.
16. Connors RD, Sumalee A, Watling DP (2007). Sensitivity analysis of the variable demand probit stochastic user equilibrium with multiple user-classes. Transportation Research Part B 41: 593–615.
17. Sumalee A, Watling DP, Nakayama S (2006). Reliable network design problem: case with uncertain demand and total travel time reliability. Transportation Research Record 1964: 81–90.
18. Yang H, Yagar S (1995). Traffic assignment and signal control in saturated road networks. Transportation Research Part A 29: 125–139.
19. Gao Z, Song YF (2002). A reserve capacity model of optimal signal control with user-equilibrium route choice. Transportation Research Part B 36: 313–323.
20. Gao Z, Sun H, Shan L (2004). A continuous equilibrium network design model and algorithm for transit systems. Transportation Research Part B 38: 235–250.
21. Wong SC, Yang H (1997). Reserve capacity of a signal-controlled road network. Transportation Research Part B 31: 397–402.
22. Chiou SW (1999). Optimization of area traffic control for equilibrium network flows. Transportation Science 33: 279–289.

23. Friesz TL, Tobin RL, Cho HJ, Mehta NJ (1990). Sensitivity analysis based heuristic algorithms for mathematical programs with variational inequality constraints. Mathematical Programming 48: 265–284.
24. Yang H, Yagar S (1994). Traffic assignment and traffic control in general freeway-arterial corridor systems. Transportation Research Part B 28: 463–486.
25. Yang H, Yagar S, Iida Y, Asakura Y (1994). An algorithm for the inflow control problem on urban freeway networks with user-optimal flows. Transportation Research Part B 28: 123–139.
26. Allsop, R.E. (1974). Some possibilities of using traffic control to influence trip distribution and route choice. In: Buckley, D.J. (Ed.), Proceedings of the 6th International Symposium on Transportation and Traffic Theory. Elsevier, Amsterdam, pp. 345–374.
27. Marcotte P (1986). Network design problem with congestion effects: a case of bi-level programming. Mathematical Programming 34: 142–162.
28. Friesz TL, Harker PT (1985). Properties of the iterative optimization-equilibrium algorithm. Civil Engineering Systems 2: 142–154.
29. Marcotte P, Marquis G (1992). Efficient implementation of heuristic for the continuous network design problems. Annals of Operation Research 34: 163–176.
30. Friesz TL, Hsun-jung C, Mehta NJ, Tobin RL, Anandalingam G (1992). A simulated annealing approach to the network design problem with variational inequality constraints. Transportation Science 26: 18–26.
31. Yang, J.; Xu, M.; Gao, Z. (2009) Sensitivity Analysis of Simulated Annealing for Continuous Network Design Problems. Journal of Transportation Systems Engineering and Information Technology. 9(3): 64-70.
32. Xu, T; Wei, H; Wang, Z. (2009) Study on Continuous Network Design Problem Using Simulated Annealing and Genetic Algorithm. *Expert System and its Application*. 36 (2009): 2735–2741.
33. Jiang, Shan; Fu, Ling-feng; Liu, Wei-ming; Liu, Yu-yin. (2013) A Bi-level Programming Model for Multi-vehicle-type Freeway Continuous Equilibrium Network Design Problem. Journal of Highway and Transportation Research and Development (English Edition). 7(1): 90-97.
34. Liu H, Wang DZW (2015). Global optimization method for network design problem with stochastic user equilibrium. Transportation Research Part B 72: 20–39.
35. Ban, Jeff X.; Liu, Henry X.; Ferris, Michael C.; Ran, Bin. (2005) A general MPCC model and its solution algorithm for continuous network design problem. Mathematical and Computer Modelling. 43(5): 493-505.
36. Li, Changmin; Yang, Hai; Zhu, Daoli; Meng, Qiang. (2012) A global optimization method for continuous network design problems. Transportation Research Part B: Methodological. 46(9): 1144-1158.
37. Gao, ZY (Gao, Ziyou); Sun, HJ (Sun, Huijun); Zhang, HZ (Zhang, Haozhi). (2007) A globally convergent algorithm for transportation continuous network design problem. Optimization and engineering. 8(3): 241-257.
38. Wang, GM (Wang, Guangmin); Yu, JW (Yu, Junwei); Li, SB (Li, Shubin). (2017) An MPCC formulation and its smooth solution algorithm for continuous network design problem. Promet – Traffic & Transportation. 29(6): 569-580.
39. Meng, Q (Meng, Q); Yang, H (Yang, H); Bell, MGH (Bell, MGH). (2001) An Equivalent Continuously Differentiable Model and a Locally Convergent Algorithm for the Continuous Network Design Problem. Transportation Research Part B: Methodological. 35(1): 83-105.
40. Du, Bo; Wang, Z. W., David. (2016) Solving Continuous Network Design Problem with Generalized Geometric Programming Approach. Transportation Research Record. 2016, 2567(1): 38-46.
41. Wang, H (Wang, Hua); Xiao, GY (Xiao, Gui-Yuan); Zhang, LY (Zhang, Li-Ye); Ji, YBB (Ji, Yangbeibei). (2014) Transportation Network Design considering Morning and Evening Peak-Hour Demands. Mathematical Problems in Engineering. 2014(1): 806916. http://dx.doi.org/10.1155/2014/806916
42. Sun, Hua; Gao, Ziyou; Long, Jiancheng. (2011) The Robust Model of Continuous Transportation Network Design Problem with Demand Uncertainty. Journal of Transportation Systems Engineering and Information Technology. 11(2): 70-76.
43. Sohn, Keemin. (2011) Multi-objective optimization of a road diet network design. Transportation Research Part A: Policy and Practice. 45(6): 499-511.
44. Cao, Jin Xin; Wang, Yang; Wei, Zheng Mao; Wu, Jun. (2013) Solve the Discrete Network Design Problem Under Construction cost Uncertainties with the Stochastic Programming Approach. Procedia - Social and Behavioral Sciences. 2013: 1039-1049.
45. Gao, Z., Wu, J., & Sun, H. (2005). Solution algorithm for the bi-level discrete network design problem. Transportation Research Part B, 479-495.
46. LeBlanc LJ (1975). An algorithm for the discrete network design problem. Transportation Science 9: 183–199.
47. Chen MY, Alfa AS (1991). A network design algorithm using a stochastic incremental traffic assignment

approach. Transportation Science 25: 214–224.
48. Drezner Z, Wesolowsky GO (1997). Selecting an optimum configuration of one- way and two-way routes. Transportation Science 31: 386–394.
49. Farvaresh H, Sepehri MM (2013). A branch and bound algorithm for bi-level discrete network design problem. Networks & Spatial Economics 13: 67–106.
50. Poorzahedy H, Turnquist MA (1982). Approximate algorithms for the discrete network design problem. Transportation Research Part B 16: 45–55
51. Cantarella GE, Pavone G, Vitetta A (2006). Heuristics for urban road network design: lane layout and signal settings. European Journal of Operational Research 175: 1682–1695.
52. Magnanti TL, Wong RT (1984). Network design and transportation planning: models and algorithms. Transportation Science 18: 1–55.
53. Solanki RS, Gorti JK, Southworth F (1998). Using decomposition in large-scale highway network design with a quasi-optimization heuristic. Transportation Research Part B 32: 127–140.
54. Gao Z, Wu J, Sun H (2005). Solution algorithm for the bi-level discrete network design problem. Transportation Research Part B 39: 479–495.
55. Poorzahedy H, Abulghasemi F (2005). Application of ant system to network design problem. Transportation 32: 251–273.
56. Drezner Z, Wesolowsky GO (2003). Network design: selection and design of links and facility location. Transportation Research PartA 37: 241–256.
57. Wu J, Lu H, Yu X (2012). Genetic Algorithm for Multiuser Discrete Network Design Problem under Demand Uncertainty. Mathematical Problems in Engineering 686272
58. Poorzahedy H, Rouhani OM (2007). Hybrid meta-heuristic algorithms for solving network design problem. European Journal of Operational Research 182: 578–596.
59. Haas I, Bekhor S (2016). A parsimonious heuristic for the discrete network design problem. Transportmetrica A-Transport Science 12: 43–64.
60. Wang S, Meng Q, Yang H (2013). Global optimization methods for the discrete network design problem. Transportation Research Part B 50: 42–60
61. Wang DZW, Liu H, Szeto WY (2015). A novel discrete network design problem formulation and its global optimization solution algorithm. Transportation Research Part E 79: 213–230.
62. Riemann R, Wang DZW, Busch F (2015). Optimal location of wireless charging facilities for electric vehicles: flow-capturing location model with stochastic user equilibrium. Transportation Research Part C 58: 1–12
63. Asadi Bagloee S, Sarvi M (2018) An outer approximation method for the road network design problem. PLoS ONE 13(3): e0192454. https://doi.org/10.1371/journal.pone.0192454
64. Liu X, Chen Q. (2016) An Algorithm for the Mixed Transportation Network Design Problem. PLoS ONE 11(9): e0162618. doi:10.1371/journal.pone.0162618
65. Dimitriou L, Tsekeris T, Stathopoulos A (2008). Genetic computation of road network design and pricing Stackelberg games with multi-class users. In: Giacobini M. et al. (Eds.), Applications of Evolutionary Computing. Springer, Berlin, Heidelberg, pp. 669–678.
66. Gallo M, D' Acierno L, Montella B (2010). A meta-heuristic approach for solving the urban network design problem. European Journal of Operational Research 201: 144–157.
67. Luathep P, Sumalee A, Lam WHK, Li Z, Lo HK (2011). Global optimization method for mixed transportation network design problem: A mixed-integer linear programming approach. Transportation Research Part B 45: 808–827.
68. Highway Capacity Manual, 4th Edition, Transportation Research Board, National Science Foundation, Washington DC, 2000.
69. Sheffi, Y. (1985). Urban transportation networks: Equilibrium analysis with mathematical programming methods. Englewood Cliffs, N.J.: Prentice-Hall.

Advances in Transportation Studies: an international Journal 67 (2025)

ISBN 979-12-218-2194-9 – ISSN 1824-5463 – DOI 10.53136/97912218219492 – pag. 17-28

Vehicle dynamics-based approach for superelevation transition in reverse consecutive horizontal curves

S. Mavromatis[1] V. Matragos[1] K. Amiridis[1] A. Kontizas[1] A.A Kordani[2]

[1]*National Technical University of Athens, Department of Transportation Planning and Engineering*
5, Iroon Polytechniou str., GR-15773 Athens, Greece
email: stemavro@central.ntua.gr
[2]*Imam Khomeini International University, Department of Civil-Transportation Planning, Qazvin, Iran*

subm. 11st November 2024 *approv. after rev. 3rd March 2025*

Abstract

The design of consecutive reverse horizontal curves is a common practice in road design with many environmental and economic benefits. In order for the driver to avoid instantaneous lateral acceleration variation during the curvature direction shifting, nearly all design guidelines adopt spiral curves between the opposite turning directions. During the design of such curves, most European road design guidelines treat the spiral curves on both sides of the reverse curvature point (point where end of spiral on first curve concurs with beginning of spiral on second curve), as the boundaries within where the necessary superelevation transition takes place as well. Through this concept, the point with level (horizontal) superelevation rate usually coincides with the point of reverse curvature, thus creating a breakpoint at the superelevation transition when the reverse curves have different either superelevation values or spiral lengths. The paper investigates the impact of utilizing a continuous (linear) superelevation transition between the points where the first circular curve ends and the second circular curve begins. This assessment is carried out by quantifying the safety margins in terms of demanded friction values for both approaches. The investigation is based on the German RAL, 2012 design guidelines, tailored for unfavorable cases, assuming a poor friction pavement of high superelevation demand (sharp curves), through the utilization of an existing vehicle dynamics model. The analysis revealed that the vehicle undergoes an immediate but rather moderate lateral friction demand variation. However, before introducing the proposed approach in road design practice, there are certain issues that necessitate further research.

Keywords – reverse horizontal curves, superelevation transition, side friction

1. Introduction and problem statement

Road layouts with consecutive reverse horizontal curves are a common practice in road design. Although the proper design of such arrangements offers many environmental and economic benefits, reverse horizontal curves impose numerous safety issues to overcome. For example, besides sight distance limitations, drivers experience difficulties in sustaining their vehicle inside the driving lane due to the direction swap of the centrifugal force.

Excessive centrifugal force generated by high design speed values, or sharp horizontal curvature, may cause abrupt, uncomfortable and unsafe maneuvers, or even lateral drifting of the vehicle. The effect of such unfavorable conditions can be reduced by the superelevation requirement, which is critical in terms of counterbalancing the instantaneous lateral acceleration variation caused by the centrifugal force during the curvature direction shifting.

Although lateral friction force developed between the tires and road surface interaction also assists such a counterbalance of the lateral acceleration, in order a vehicle during the negotiation of horizontal curves to optimize the reduction impact of the centrifugal force, it is very important the outer to the curve edge-line of the pavement to be the superelevated one; higher than the one inside.

Figure 1(a,b) shows a simplification of the lateral force distribution during a right turned curve applied on the well know mass-point, widely accepted in road design for the determination of the minimum horizontal radius [1, 2, 3, 4].

During the lateral equilibrium, the following Equation applies:

$$\frac{m\,V^2}{R}\cos e = m\,g\,\sin e + S \quad (1)$$

$$\rightarrow \frac{m\,V^2}{R}\cos e = m\,g\,\sin e + f_R\left(m\,g\,\cos e + \frac{m\,V^2}{R}\sin e\right) \quad (2)$$

Assuming that ($e \ll \rightarrow \cos e \approx 1,\ \sin e \approx e$)

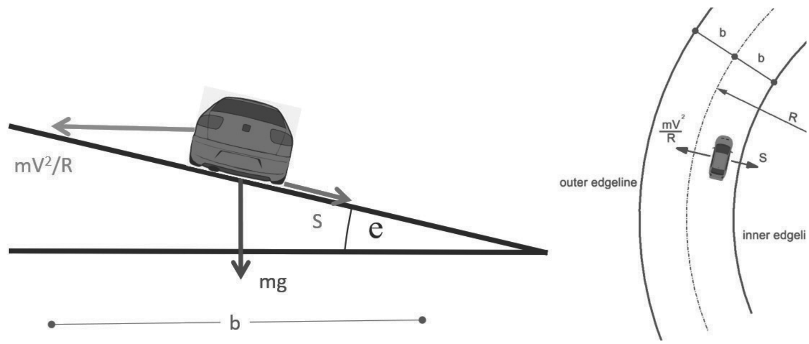
$$\rightarrow \frac{V^2}{g\,R} = \frac{e+f_R}{1-e\,f_R} \approx e + f_R \quad (3)$$

$$\rightarrow R_{min} = \frac{V^2}{g(f_R+e)} \quad (4)$$

where:
g: gravitational constant ($g=9.81 m/s^2$)
m: vehicle mass (kg),
e: superelevation rate (%/100)
V: vehicle speed (m/s),
R: curve horizontal radius (m),
S: lateral friction force (N),
f_R: lateral friction coefficient

From Equation 4 it can be seen that the use of superelevation allows a vehicle to travel through curves more safely, since there is a safety margin either to increase speed or to reduce the curve radius by maintaining the alignment and vehicle speed respectively.

Therefore, between succeeding reverse circular curves, in order the superelevation to be designed in accordance with Figure 1, many design guidelines [1, 2, 3, 4] request an adequate tangent length between, or preferably, an equivalent length with spiral curves and a maximum rate between the radii of both circular curves.



(a) lateral direction of travel (b) plan view

Note: "b" stands for the lane with.

Fig. 1(a,b) – Lateral force distribution during a right turned curve on the point-mass model

Spiral curves provide gradual change on the centrifugal force as well as more comfortable alignment between tangents and circular curves. More specifically in the AASHTO, 2018 design guidelines [1] it is stated that transition (spiral) curves simulate the natural turning path of a vehicle, minimize encroachment on adjoining traffic lanes and tend to promote uniformity in speed.

During the design of consecutive curves, most European road design guidelines [2, 3, 4] treat the spiral curves on both sides of the reverse curvature point (point where end of spiral on first curve concurs with beginning of spiral on second curve), as the boundaries within where the necessary superelevation transition takes place as well.

Through this conventional approach, the point with level (horizontal) superelevation rate usually coincides with the point of reverse curvature. As far as the two opposite curves have the same superelevation rates and spiral lengths the rotation rate of the superelevated pavement transition within the spiral curves is uniform (continuous). However, in case the reverse curves have different superelevation rates or spiral lengths, a breakpoint of the rotation rate is created at the point of reverse curvature. Such a case, where the reverse curve is formed by left curve 1 and right curve 2 (Figure 2a), is shown through Figure 2b.

The paper investigates the impact of generalizing a continuous superelevation transition (proposed rotation rate in Figure 2b) between the points where the first circular curve ends and the second circular curve begins for arrangements of reverse curves. Such an alternative approach simplifies the superelevation rotation process as is far more construction-friendly (linear rotation rate without breaks).

However, the added value of the proposed approach is that by delivering a smoother rotation rate to one of the involved spirals, with respect to the evasion of areas with poor drainage, lower grade values along the alignment axis could be utilized.

For example, in Figure 2b, the left and right edge-line rotation rates (Δs_{2L} and Δs_{2R} respectively) of the second spiral curve appear to be sharper when the conventional approach is utilized compared to the proposed one. As a result, the maximum rotation rate (Δs_{max}) cannot be implemented in the proposed approach.

Based on the German RAL,2012 design guidelines, the following (algebraic) Equations apply:

$$s_{\text{edge line}} = s_{\text{axis}} + \Delta s \tag{5}$$

$$p_{\text{edge line}} = \sqrt{s_{\text{edge line}}^2 + e_i^2} \tag{6}$$

where:

$s_{\text{edge line}}$: longitudinal grade along the pavement edge line (%)

s_{axis}: longitudinal grade along the alignment's rotation axis (%),

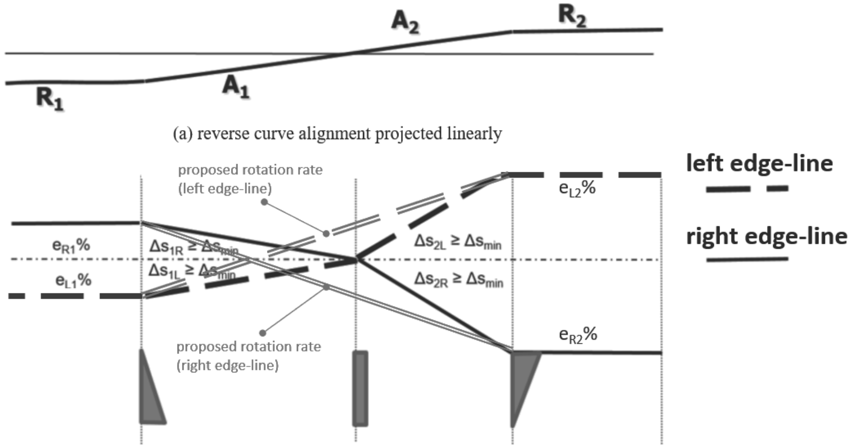
Δs: rotation rate which stands for relative grade at spiral area between longitudinal gradient along the carriageway edge line and longitudinal gradient along the alignment's rotation axis (%)

$p_{\text{edge line}}$: compound grade at the edge line (%)

e_i: instantaneous superelevation rate along the edge line area (e is variable along spiral curves) (%)

In order the compound grade p along the pavement edge line to be always greater or equal to 0.50% (avoidance of poor drainage), at the area where the superelevation rate along the spiral curves is practically level, the longitudinal grade along the left or right edge line ($s_{\text{edge line}}$) is equivalent to the respective compound grade ($p_{\text{edge line}}$). As a result, the following algebraic Equation is demanded per left or right edge line:

$$s_{\text{axis}} + \Delta s \geq 0.50 \tag{7}$$



Note (i=1,2): Ri, Ai, Li : curve i radius, spiral parameter and spiral length (m).
e_{Li}, e_{Ri}, Δs_{iL}, Δs_{iR}: curve i left and right superelevation rates and superelevation rotation rates (%),
Δs_{min}: minimum superelevation rotation rate (%).

Fig. 2(a,b) – Superelevation rotation breakpoint at the point of reverse curvature

Based on the above Equation and considering Figure 2b where a smoother rotation rate can be proposed to one of the involved spirals, it is evident that when s_{axis} and $\Delta s_{proposed}$ have opposite algebraic signs (e.g. $\Delta s_{proposed\ left\ edge\text{-}line} > 0$ and $s_{axis} < 0$, or $\Delta s_{proposed\ right\ edge\text{-}line} < 0$ and $s_{axis} > 0$), with respect to drainage adequacy, lower grade values along the alignment axis (s_{axis}) could be utilized.

This assessment of the proposed approach is performed by quantifying the safety margins in terms of demanded friction values for both conventional and continuous approaches, utilizing an existing vehicle dynamics model.

2. Methodology

The proposed investigation is based on a realistic representation of the forces acting on a moving vehicle during tangent and curved alignments. Such an investigation has already been addressed in previous research of the authors [5-7] where aiming to assess vehicle safety from the interaction between road geometry, tire - pavement friction and vehicle parameters, a vehicle dynamics model was developed. The following sub-section provides a brief discussion on how the model was structured, where more details regarding the full equations description as well as the model's validation process are available through references [5-7].

2.1. Vehicle dynamics approach

All forces and moments applied to the vehicle were analyzed into a moving three-dimensional coordinate system, coinciding at the vehicle gravity center and formed by the vehicle's longitudinal (X), lateral (Y) and vertical (Z) axis respectively. The vehicle's coordinate system applied for a front wheel driven (FWD) vehicle is shown in Figure 3. Through these axes, the impact of certain vehicle technical characteristics, road geometry and tire friction were expressed, such as: vehicle speed/ wheel drive/ sprung and unsprung mass and it's position of gravity center/ aerodynamic drag/ vertical lift/ track width/ wheel-base/ roll center/ suspension roll stiffness/ cornering stiffness/ grade/ superelevation rate/ rolling resistance tire-road adhesion values and horsepower supply.