University Library



A gateway to Melbourne's research publications

Minerva Access is the Institutional Repository of The University of Melbourne

Author/s:
Bagloee, SA;Sarvi, M;Patriksson, M;Rajabifard, A

Title:
A Mixed User-Equilibrium and System-Optimal Traffic Flow for Connected Vehicles Stated as a Complementarity Problem

Date:
2017-07-01

Citation:
Bagloee, S. A., Sarvi, M., Patriksson, M. & Rajabifard, A. (2017). A Mixed User-Equilibrium and System-Optimal Traffic Flow for Connected Vehicles Stated as a Complementarity Problem. Computer Aided Civil and Infrastructure Engineering, 32 (7), pp.562-580. https://doi.org/10.1111/mice.12261.

Persistent Link:
https://hdl.handle.net/11343/292779

A mixed user-equilibrium and system-optimal traffic flow stated as a complementarity problem

Saeed Asadi Bagloee , Majid Sarvi
Melbourne School of Engineering, The University of Melbourne, Victoria 3010, Australia

Michael Patriksson
Chalmers University of Technology and University of Gothenburg, SE-412 96 Gothenburg, Sweden,
&
Abbas Rajabifard
Melbourne School of Engineering, The University of Melbourne, Victoria 3010, Australia

Author Manuscript

Abstract: *Connected vehicles (CVs), be they autonomous vehicles or a fleet of cargo carriers or Uber, are a matter of when they become a reality and not if. It is not unreasonable to think that CV technology may have a far-reaching impact, even to the genesis of a completely new traffic pattern. To this end, the literature has yet to address the routing behavior of the CVs, namely traffic assignment problem (TAP) (perhaps it is assumed, they ought to follow the traditional shortest possible paths, known as user equilibrium - UE). It is possible that real-time data could be derived from the vehicles' communications that in turn could be used to achieve a better traffic circulation. In this article, we propose a mathematical formulation to ensure the CVs are seeking the system optimal (SO) principles, while the remainder continue to pursue the old-fashion UE pattern. The model is formulated using the nonlinear complementarity problem (NCP). This paper contributes to the literature in three distinct ways: (i) mathematical formulation for the CVs' routing, stated as a mixed UE-SO traffic pattern, is proposed; (ii) a variety of realistic features are explicitly considered in the solution to the TAP including road capacity, elastic demand, multiclass and asymmetric travel time; and (iii) formal proof of the existence and uniqueness of the solutions are also presented. The proposed methodology is applied to the networks of Sioux-Falls and Melbourne.*

1. Introduction

New technologies in communication and robotics have had a substantial influence in our daily lifestyle of which transportation is no exception. These technologies have given rise to the prospect of the connected vehicles (CVs) in general and autonomous vehicle (AV) technology in particular which aim to reducing crashes (NHTSA, 2012), energy consumption (Li et al., 2015), pollution and congestion (Dresner and Stone, 2007) while at the same time increasing the reach of transport, mobility and accessibility. In light of looming new ideas and technologies such as social networks, smartphones, CVs and AVs, the landscape of transportation is rapidly changing (Anderson et al., 2014; Folsom, 2011; Piao and McDonald, 2008). An example is Uber which is sweeping cities to the extent that taxi companies are struggling to retain business and to remain competitive. Significant efforts have been thrown behind the idea of the CV/AV to fruition, the speed of which has increased enormously in the last decade to the extent, it is now a matter of "how soon" and not "if". Therefore, it is of the highest importance to be prepared for such eventualities and to understand the challenges that lie ahead whilst embracing and welcoming the ensuing opportunities.

Though, ad hoc deployments of autonomous vehicles such as Google's car have reached to operational level and are occasionally seen on the road, they are astronomically expensive. For instance, sensing devices such as LIDAR camera alone costs around $75,000. Prices as to the whole setup cost around $150,000 (see (Google, 2017)). An alternative and affordable deployment is to install less expensive devices on the AVs, while connecting them to share data. As the result, connectedness –to increase efficiency and lower expenses- is one of the indispensable features for a mass production and deployment of the autonomous vehicle.

This is the author manuscript accepted for publication and has undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1111/mice.12261.

This article is protected by copyright. All rights reserved.

With the same rationale, the idea of vehicle to infrastructure communication emerges (Sarvi and Kuwahara, 2008; Sarvi et al., 2004). In this context, extensive efforts have been made towards determining traffic incident detection based on artificial intelligent methods (Adeli and Karim, 2005) that have been shown to be highly effective (Karim and Adeli, 2003; Samant and Adeli, 2000). Generally speaking, when provided with speed, volume, and occupancy data from detector(s) (Ghosh-Dastidar and Adeli, 2003), neural network techniques are employed for pattern recognition and classification (Adeli and Samant, 2000). These models have been significantly improved by integrating initiatives such as wavelet (to denoise the traffic data (Karim and Adeli, 2002b; Samant and Adeli, 2001)) fuzzy (Adeli and Karim, 2000; Karim and Adeli, 2002a). This information can be then relayed to the cars to assist them in their navigation.

As of the opportunities, communications between vehicles and infrastructure also provide a fertile ground to push for a more efficient traffic circulation known as system optimal (SO).

It is to cater an answer to the traffic assignment problem (TAP) where CVs follow SO while human-driven vehicles still seek shortest paths or user equilibrium (UE). In this study, given the presence of connected vehicles, we provide a solution to the TAP in which a mixture of SO and UE traffic pattern coexist (called SO-UE-TAP).

It is postulated that the connected vehicles follow the SO. The questions that may arise are: (i) are connected vehicles something to be considered in the distant future and not relevant to today's reality?; and (ii) what guarantees are there that the connected vehicle will follow the SO traffic pattern.

To answer these questions, it is important to note that connectedness is not limited only to the AV. A fleet of carriers such as DHL, UPS and FedEx, as well as Uber, can also be connected; that is, the vehicles share their traveler information (congestion levels, traffic incidents, weather conditions, etc). Considering the emerging technologies related to smart cities, the internet of things and the pervasive use of sensors in vehicles and transport infrastructure, the connectedness is becoming an inevitable feature of transport.

Consequently, cooperative routing is highly likely. Furthermore, debates on regulations, standards and protocols for the connected vehicle and AV are still underway. As a result, it is within the realm of traffic authorities, license agencies and regulators to make cargo fleets carriers as well as car companies to program their AVs or fleets to follow the SO. A similar paradigm is being practiced for decades in the aviation sector. In light of smart cities moving toward such schemes for the land transport is very likely. For AVs, it is easier, the authorities can trade the added-value that AVs bring to their drivers (being off the wheel) with SO principles. Nevertheless, the willingness of AV drivers to accept such a tradeoff is a question remains unanswered.

However, as the market share of the SO-seeking vehicles increases (more AVs and CVs on the roads), their losses diminish and turn around to report a net gain, so, the entire transport system moves towards a complete SO (Bagloee et al., 2016b). Therefore, in this manuscript AV and CV are used synonym for referring to SO-seeking vehicles (note that, not any connected vehicle seeks SO. We assume centrally-dispatched connected vehicles seek SO flow).

Solving the TAP itself is a center of attention in the literature for which a variety of algorithms have been proposed. Therefore the contribution of this study can be scrutinized with regards to the way it solves the TAP. For the SO-UE-TAP we propose a solution method based on nonlinear complementarity problem (NCP). The solution algorithm is coded in GAMS (an optimization software (Brooke et al., 1996)) and it is numerically applied to the benchmark network of Sioux-Falls as well as a network of Melbourne, Australia.

This paper contributes to the literature in five folds:

I. The concept of mixed SO and UE traffic pattern is proposed for the combination of connected vehicles and other vehicles.

II. In order to maintain the TAP mathematically and computationally tractable -in the past studies- roads physical capacities were not explicitly considered. Therefore many realistic features being represented by capacity constraints in the TAP were largely overlooked. In this paper, the roads capacities are explicitly embedded in the NCP formulation using their Lagrangian representations.

III. "Multiclass" is a feature of the TAP to represent the situation in which different types of vehicles (say car, truck, heavy vehicles) with different characteristics coexist. Due to computational complexities, the multiclass is –traditionally- relaxed or compromised (Florian and Morosan, 2014). We consider two distinct traffic flow patterns (SO and UE) as two vehicle classes and hence it can be generalized to more vehicle classes.

IV. The TAP is conventionally formulated as an optimization problem using Beckmann formulation (Beckmann et al., 1956) in which the roads delay functions must abide by some restrictive properties (Each road is associated with a function called delay or performance function which specifies the travel time of the respective road based on the traversing traffic volumes). Of such restrictions the delay function must be separable that is the travel time on link "a" depends solely on traffic flow of link "a". It is obvious that the

This article is protected by copyright. All rights reserved.

travel time may depend on traffic flow of the downstream roads or crossing roads at junctions. The proposed NCP formulation can accommodate a more general type of functions namely asymmetric delay functions. As the result, travel time depends on a vector of the links' traffic flows, and not reciprocally (Aashtiani, 1979; Aashtiani and Magnanti, 1981).

V. A valid concern in the TAP is the fact that the travel demand can itself be sensitive to the changes in travel time. For instance, in response to high travel time, some commuters might decide to cancel the trip or to change their usual travel mode or their departure times. The mode choice between the CVs and non-CVs are formulated using a logit model and it is embedded in the proposed formulation.

In the remainder of the article, Section 2 provides a review of the literature of the mix/multiclass, capacity, asymmetric delay function and variable demand; Section 3 presents the NCP formulation of the SO-UE; Section 4 is dedicated to the numerical results; Section 5 concludes the article.

2. Literature on mix/multiclass, capacity, asymmetric delay function, variable demand

Capacity constraints: As noted before, in order to keep the TAP mathematically and computationally tractable, roads' capacities are ignored (no capacity constraint). Moreover, capacity constraints can also represent a wide range of realistic features such as (i) refinement of the traffic equilibrium (Ferrari, 1997; Larsson and Patriksson, 1999), (ii) environmental constraint (Chen et al., 2011a), (iii) replicating traffic count (Bell et al., 1997), (iv) traffic control (Yang and Bell, 1997), (v) congestion pricing (Yang and Bell, 1997), (vi) queuing effects (Larsson et al., 2004), (vii) combined/integrated modelling (Ryu et al., 2014b). The concept of capacity can also be utilized in the network design problem (NDP) that is, in the Lagrangian-based algorithms such as Benders decomposition or Lagrangian relaxation, one needs to solve a capacitated traffic assignment (Bagloee et al., 2016a; Gao et al., 2005). Therefore capacity binding traffic assignment (C-TAP) has always attracted interest from both scholars as well as practitioners (FHWA, 2002; Larsson and Patriksson, 1999). The prominent methods developed for the C-TAP can be classified as Lagrangian multipliers (Hearn and Ribera, 1980; Larsson and Patriksson, 1992, 1995; Larsson et al., 2004; Nie et al., 2004) or a penalty function (Hearn, 1980; Inouye, 1987; Morowati-Shalilvand and Mehri-Tekmeh, 2013; Nie et al., 2004; Prashker and Toledo, 2004; Ryu et al., 2014b; Shahpar et al., 2008; Yang and Yagar, 1994, 1995). The mathematical complexities involved in the C-TAP have resulted in solution algorithms laden with a number of parameters for which calibration is a significant concern. In addition, arriving at an initial feasible solution at the outset of the algorithm's computation is also a nontrivial challenge (for more details, the interested reader can consult with (Bagloee and Sarvi, 2015)). The formulation provided in this study obviates these two concerns.

Variable demand: Elasticity of the travel demand has been extensively studied in the past (Aashtiani, 1979; Babonneau and Vial, 2008; Dafermos, 1982; Fisk and Boyce, 1983; Ryu et al., 2014a). The demand elasticity stands for the effort of considering the changes in the travel demand in response to changes in the supply side. In particular, the travel demand is sensitive to travel time in the network. There are two main approaches: (i) the network is expanded to represent the variation of travel demand, (ii) the inverse format of variable demand function is brought to the Beckmann objective function. A comprehensive review on the demand elasticity is provided by (Gartner, 1980a, b; Patriksson, 2004; Patriksson, 1994; Sheffi, 1985). The former, though intuitive, is not computationally appealing. Nevertheless, the latter calls for a very restrictive type of inverse variable demand function similar to the requirements of the travel demand function. Consequently, we employ the latter approach, meanwhile, the NCP formulation imposes much less restrictive conditions on the inverse travel demand function.

Asymmetric delay function: The other widely neglected feature is the fact that the roads' travel times may be a function of flow on the respective road as well as some other roads -especially conflicting approaches at the intersections. This leads to a more comprehensive view which suggests the impact of traffic volume of road "a" on the delay of road "b" may not reciprocal. The delays of primary and secondary approaches at intersections are of such examples. These phenomena are discussed under asymmetric delay functions for which the interested reader can consult with (Muijlwijk, 2012; Noriega and Florian, 2007). Why are the asymmetric functions neglected in the conventional models for the TAP? The answer is: inclusion of the asymmetric functions will strip the TAP off the convenience of being a convex optimization problem. In other words, in the presence of the asymmetric delay functions, the Beckmann optimization model for the TAP is no longer applicable. Alternatively, non-optimization methods such as variational inequality, complementarity and fixed-points are shown to be up to this job (Aashtiani, 1979; Florian and Morosan, 2014; Nagurney and Dong, 2002).

Mixed and multiclass flow: The concept of mixed traffic equilibrium is not new in the literature. Some pioneer studies have flagged situations in which both competition and cooperation traffic flow coexist (Haurie and Marcotte, 1985). Harker (1988) provides a formulation in which the UE is mixed with some other traffic equilibrium patterns. In the advent of "advanced traveler information system the idea of route guidance -aiming towards the SO principles- has also been introduced in the literature (Maher and Hughes, 1995; Van Vuren et al., 1990; Van Vuren and Watling, 1991). The concept of mixed traffic equilibriums has also extended to the

This article is protected by copyright. All rights reserved.

stochastic traffic assignment (Lo and Szeto, 2002; Yang, 1998; Yin and Yang, 2003). Yang et al. (2007) propose a Stackelberg leader-follower game for improving network performance in the presence of mixed traffic patterns. To this particular line of research, the current study contributes to the literature by taking the roads capacities, multiclass, asymmetric travel times and variable demand into a single formulation.

Nevertheless, the mixture of traffic patterns can also be generalized as multiclass traffic assignment problem (MC-TAP) which is computationally more intensive than a single class TAP. To this end, there exists a variety of methods such as variational inequality, complementarity method, fixed-points and entropy maximization as well as origin-based (or bush based) methods have been proposed (Aashtiani, 1979; Bar-Gera and Boyce, 1999; Boyce, 2014; Chen et al., 2011b; Dafermos, 1972; Florian and Morosan, 2014; Nagurney, 2000; Nagurney and Dong, 2002; Xie and Xie, 2014, 2015; Zhang and Chen, 2010). In some path-based methods, an exhaustive and computationally expensive path enumeration task is inevitable (Aashtiani, 1979). However, in a number of studies the MC-TAP has permuted to the extent it obviates exhaustive path enumeration (Castillo et al., 2013b). Castillo et al. (2013a) transformed the MC-TAP as an NCP and developed a solution method which does not require path enumeration –instead, it sticks to the already used paths. Castillo et al. (2013c) present a dynamic model based on the link travel time function in which different class users interact in the traffic network. However, these formulations come at the cost of a number of parameters to be calibrated.

Given the structure of the mixed UE-SO assignment, solving the problem in two levels (that is a bi-level method) seems to be an intuitive approach. The main obstacle in bi-level approaches is the computational time. It has been mathematically proven that any bi-level problem is NP-hard (Ben-Ayed and Blair, 1990; Colson et al., 2005, 2007; Dempe, 2003); that is, as the size of the problem increases the solution becomes unattainable.

Consequently, a capacitated SO-UE-TAP subject to variable demand and asymmetric delay functions is formulated as an NCP and it is encoded in GMAS.

Though the primary motive of this study is to develop a mathematical model for the mixed traffic equilibrium –in the anticipation of the connected vehicles we also pursue two more credits: (i) to address the above-mentioned shortcomings in the TAP as a single and unifying platform and (ii) to make a variety of optimization tools available in the GAMS handy for transport planners enabling them to address problems such as network design and road (or congestion) pricing etc. These kinds of problems can be cast as bilevel programing problem with a leader problem at the upper level while accounting for the traffic assignment in the lower bound. Hence these problems can then be solved using mpec (mathematical programming with equilibrium constraints) module of the GAMS. In the sequel, we present the mathematical formulation.

3. SO-UE traffic assignment problem (SO-UE-TAP)

In this section, we first elaborate on the concept of mixed SO-UE traffic equilibrium using Beckmann formulation. We then provide an NCP formulation for the SO-UE-TAP. Hereinafter, in contrast to the connected vehicles (CV) we referred to the other vehicles as nCV (not connected vehicle). In other words, CV vs nCV or equivalently SO vs UE. Consider $G(N,A)$ a road network as a graph consists of N,A sets of nodes and roads respectively on which $D \subset N$ is a set of destinations. Since a set of roads are defined based on nodes (i.e. $A \subset N \times N$), we represent roads using a single character $a \in A$ as well as start and end nodes: $a = ij = (i,j) \in A$.

3.1. Formulation based on Beckmann formulation

At equilibrium conditions, both UE and SO arrive at a stable situation in which no car changes its route. For the UE part, consider x_a is nCVs traffic flow on road a while $\bar{x}_a$ denotes CVs as the background traffic volume. Therefore the UE traffic pattern can be formulated as a nonlinear programming problem (Patriksson, 1994; Sheffi, 1985) (throughout the manuscript, all terms are non-negative unless otherwise stated):

$$[\text{UE-TAP}]: \min z(x) = \sum_{a \in A} \int_0^{x_a} t_a(x_a + \bar{x}_a)\, dx \qquad (1)$$

$$\text{s.t.}: \sum_p f_{p,i}^k = q_i^k \quad i \in N,\ k \in D \qquad (2)$$

$$f_{p,i}^k \geq 0 \quad p \in P_i^k, i \in N, k \in D \qquad (3)$$

$$x_a = \sum_i \sum_k \sum_p f_{p,i}^k . \delta_{a,p,i}^k \quad a \in A,\ p \in P_i^k, i \in N, k \in D \qquad (4)$$

where z: the Beckmann objective function to be minimized; q_i^k: nCVs travel demand from i to k; $f_{p,i}^k$:the flow of nCVs on path p from i to k; P_i^k: set of all paths available to nCVs from i to k; $\delta_{a,p,i}^k$: the link-path incidence (1: if link a belongs to path p from i to k available to nCVs, and 0 otherwise). Similarly, at equilibrium, CVs volumes $(\bar{x}_a)$ in the context of background traffic of selfish volume (x_a) can be formulated as follows (selfish driving is a term used in the literature to highlights the fact that vehicles non-cooperatively seek their individual benefit in choosing the shortest paths. Moreover the words of cooperative and non-cooperative

This article is protected by copyright. All rights reserved.

are used for SO and UE traffic pattern respectively). In other words $\bar{x}_a$ and x_a represent SO traffic volume (or CVs) and UE (or nCVs as background traffic) on road a (Patriksson, 1994; Sheffi, 1985):

$$[\text{SO-TAP}]: \min z(\bar{x}) = \sum_{a\in A} \bar{x}_a . \bar{t}_a(\bar{x}_a + x_a) \quad (5)$$

$$\text{s.t.}: \sum_p \bar{f}^k_{\bar{p},i} = \bar{q}^k_i \qquad i \in N, k \in D \quad (6)$$

$$\bar{f}^k_{\bar{p},i} \geq 0 \qquad p \in \bar{P}^k_i, i \in N, k \in D \quad (7)$$

$$\bar{x}_a = \sum_i \sum_k \sum_p \bar{f}^k_{p,i} . \bar{\delta}^k_{a,p,i} \quad a \in A, p \in \bar{P}^k_i, i \in N, k \in D \quad (8)$$

The notations are similar as the bar on top of the terms represents the CVs. Both CVs and nCVs' travel demand would share the same network ($a \in A$). It is important to highlight the difference of the objective functions: while the UE-TAP rests on the Beckmann formulation (Eq.(1)), the SO-TAP rests on the total travel time spent in the network (Eq.(5)).

There are a plethora of methods that have been proposed to solve the UE-TAP efficiently. Hence, an easy way to solve an SO-TAP is to transform it to a UE-TAP. To do so, one simply needs to replace the delay function $\bar{t}_a(\bar{x}_a + x_a)$ with the marginal delay function: $\bar{t}_a(\bar{x}_a + x_a) = t_a(\bar{x}_a + x_a) + \bar{x}_a . \partial t_a(\bar{x}_a + x_a) / \partial \bar{x}_a$ (Dafermos and Sparrow, 1969). Therefore both UE-TAP and SO-TAP can be combined as a multiclass TAP but with two different delay functions and travel demand matrices.

There is a delicate difference between the real or actual travel cost (or time) and the marginal cost. Let us consider a single road associated with a delay function t(x). The x denotes the respective traffic volume, connecting origin A to destination B for which there are two travel demands - d_ue and d_so- both of which pertain to the UE and SO respectively. It is apparent that the cost of travel (or time) that UE demand experiences are: t(d_ue + d_so). However, for the SO demand, only for the purpose of routing, the marginal cost will be charged, which is t(d_ue + d_so) + d_so . t'(d_ue + d_so). It is also apparent that the SO demand chooses the same road and they experience an "actual" travel cost identical to that of the UE: t(d_ue + d_so). In other words, subjecting the SO demand to the charge of the marginal cost is merely a mathematical technique to make them follow the SO pattern. Furthermore, for the TAP, it is proven that under mild conditions, the link solution is unique while the corresponding path solutions are not. In other words, for a given origin-destination one may, find a number of different SO paths

The SO traffic pattern is the most desirable traffic pattern when the total travel time spent on the network (as a network performance index) is minimized. In reality, people follow the shortest path which leads to the UE. In terms of the total travel time (also an index for the traffic congestion), the gap between UE and SO can reach as high as 2.15 (Roughgarden and Tardos, 2002). In other words, by enforcing SO rather than UE traffic pattern, one can significantly improve the congestion level by up to 2.15 times. This gap has been a driving force and motive in a variety of traffic management (or control) measures and policies such as parking planning, congestion pricing, and ramp metering, etc. The advent of CVs can also be added to these schemes.

3.2. Formulation based on nonlinear complementarity problem (NCP)

The nonlinear complementarity problem (NCP) is to find a vector satisfying the system of equations and inequalities $x \geq 0, F(x) \geq 0, x^T . F(x) = 0$ where $F(x)$ is a given function defined on a subset containing at least the nonnegative orthant (Ferris and Munson, 2014). More precisely x is a vector of variables and so $F(x)$ is a vector function. The aforementioned formulation is also denoted with a more compact notation as follows: $0 \leq x \perp F(x) \geq 0$ (note, symbol "$\perp$" represents the projection of x over $F(x)$). The NCP was first introduced by (Cottle, 1964) followed by the introduction of the variational inequality (Hartman and Stampacchia, 1966). For more details, the interested reader can consult with (Facchinei and Pang, 2003; Ferris and Pang, 1997; Gabriel et al., 2012; Harker and Pang, 1990; Yong, 2010).

Let us consider $X = \{X^1, X^2, ..., X^k, ..., X^{|D|}\}$ the flow vector destined to a destination $D = \{1, 2, ..., k, .., |D|\}$ such that x^k_{ij} (or x^k_a) is flow on road $a = (i, j)$ destined to the destination k. Note: $X^k = \{x^k_a \mid a \in A\}$ Similarly, $T = \{T^1, T^2, ..., T^k, ..., T^{|D|}\}$ is the travel time vector composed of t^k_i minimum travel time from node i to destination k, while t_a is the delay on the link $a \in A$ as a function of vector flow: $t_a(X)$. The wardropian traffic flow is enforced by upholding the following complementarity conditions (Ferris et al., 1999):

$$\begin{cases} t^k_j + t_{ij}(X) - t^k_i \geq 0 \\ x^k_{ij} \geq 0 \qquad (i, j) \in A, k \in D \\ x^k_{ij} . (t^k_j + t_{ij}(X) - t^k_i) = 0 \end{cases} \quad (9)$$

This article is protected by copyright. All rights reserved.

According to (9): if the link (i, j) is found to be on the shortest path from i to k (i.e. $t_j^k + t_{ij}(X) - t_i^k = 0$) we then expect to see some traffic volume on it (i.e. $x_{ij}^k \geq 0$). Otherwise (i.e. $t_j^k + t_{ij}(X) - t_i^k \geq 0$) no such traffic volume will be found (i.e. $x_{ij}^k = 0$). The above formulation (9) can be expressed in a more compact format as follows:

$$0 \leq x_{ij}^k . \perp \{t_j^k + t_{ij}(X) - t_i^k\} \geq 0 \quad (i, j) \in A, k \in D \qquad (10)$$

It is important to note that in this equation travel time of a road is not merely depended on traffic flow of the respective road, rather it is a function of the vector of traffic flows across the network (i.e. $t_{ij}(X)$). Moreover, as noted before, it is enough to replace delay function with the marginal delay function (i.e. $t_{ij}(X) + x_{ij}.\partial t_{ij}(X)/\partial x_{ij}$) to ensure SO traffic flow. In the Beckmann formulation, the capacity constraints can be easily added to the constraints (c_{ij} capacity of road (i, j)):

$$\sum_k (x_{ij}^k + \bar{x}_{ij}^k) \leq c_{ij} \qquad (i, j) \in A \qquad (11)$$

In order to adhere to the general format of the NCP, a capacity constraint cannot be added directly to the system of the equations as side constraint. To this end, the capacity constraints can be relaxed and their corresponding lagrangian values are added to their respective travel times (Larsson and Patriksson, 1995, 1999; Larsson et al., 2004). In other words, these lagrangian values are then interpreted as inflated travel time for the respective roads hence equation (11) can be rewritten as:

$$0 \leq x_{ij}^k \perp \{t_j^k + t_{ij}(X) + \lambda_{ij} - t_i^k\} \geq 0 \quad (i, j) \in A, k \in D \qquad (12)$$

where λ_{ij} is the lagrangian value of the capacity constraint pertaining to road (i, j). As such, complementarity format of the capacity constraints can be expressed as:

$$0 \leq \lambda_{ij} \perp \{c_{ij} - \sum_k (x_{ij}^k + \bar{x}_{ij}^k)\} \geq 0 \quad (i, j) \in A \qquad (13)$$

It is evident that if the capacity constraint is found not binding the corresponding lagrangian value will take the value of zero, otherwise, it is a positive value.

The flow conservation constraints can be expressed as:

$$\sum_{j|(i,j)\in A} x_{ij}^k - \sum_{j|(j,i)\in A} x_{ji}^k = q_i^k \quad i \in N, k \in D \qquad (14)$$

where q_i^k is travel demand from node $i \in N$ to destination $k \in D$. The flow conservation constraints must also be written as per NCP format. One intuitive way is to tie Eq.(14) to its respective travel time as follows:

$$0 \leq t_i^k \perp \{\sum_{j|(i,j)\in A} x_{ij}^k - \sum_{j|(j,i)\in A} x_{ji}^k - q_i^k\} \geq 0 \quad i \in N, k \in D \qquad (15)$$

where it is obvious that the travel time is always non-zero hence the equality condition (14) is always upheld.

As delay function in (12) depends on the traffic volumes of the roads, the system of equations must also have constraints to derive roads flow x_{ij} from path flows x_{ij}^k:

$$0 \leq x_{ij} \perp \{x_{ij} - \sum_k x_{ij}^k\} \geq 0 \qquad (i, j) \in A \qquad (16)$$

The demand elasticity can be defined as:

$$q_i^k = q_i^k(T) \qquad i \in N, k \in D \qquad (17)$$

where $q_i^k(T)$ is a smooth function of the vector of the origin-destination (OD) pairs' travel time (T). It is worth noting that unlike convex optimization formulation for the TAP, the NCP does not require any special characteristic for the demand function such as being invertible or integrable (Aashtiani, 1979; Ferris et al., 1999). Nevertheless, the elasticity of the travel demand is widely formulated using the well-behaved logit models which are thought to be more realistic and widely used (Ferris et al., 1999; Huang, 2002) as follows (In Appendix A, we prove that to ensure existence/uniqueness of the solution the demand function must be a strictly monotone function):

$$q_i^k(T) = \hat{q}_i^k \frac{\exp(-\rho.t_i^k)}{\exp(-\rho.t_i^k) + \exp(-\bar{\rho}.\bar{t}_i^k)} \quad i \in N, k \in D \qquad (18)$$

where $q_i^k(T)$ is travel demand of the UE and $\rho, \bar{\rho}$ are disutility factors to be calibrated using field survey data pertaining to the UE and SO respectively; $\hat{q}_i^k$ is a constant value representing total travel demand (i.e. $0 \leq q_i^k \leq \hat{q}_i^k$) and $\bar{t}_i^k$ denotes travel time on alternative transport mode SO. Since $\hat{q}_i^k = q_i^k(T) + \bar{q}_i^k(T)$ (where $\bar{q}_i^k$ is travel demand of SO), the equality sign in (18) can be relaxed to "$\geq$" and the equation can be rewritten as:

$$(q_i^k(T) - \hat{q}_i^k).\exp(-\rho.t_i^k) + q_i^k(T).\exp(-\bar{\rho}.\bar{t}_i^k) \geq 0 \quad i \in N, k \in D \qquad (19)$$

This article is protected by copyright. All rights reserved.

Now, the variable demand can be formulated as an NCP:

$$0 \le q_i^k(T) \perp \{(q_i^k(T) - \hat{q}_i^k).\exp(-\rho.t_i^k) + q_i^k(T).\exp(-\bar{\rho}.\bar{t}_i^k)\} \ge 0 \;\; i \in N, k \in D \quad (20)$$

That is, if $0 \le q_i^k(T)$ holds, logit Eq.(18) will be binding. Now we can establish an NCP formulation of the SO-UE-TAP:

$$0 \le x_{ij}^k \perp \{t_j^k + t_{ij}(X) + \lambda_{ij} - t_i^k\} \ge 0 \; (i,j) \in A, \; k \in D \quad (12)$$

$$0 \le \lambda_{ij} \perp \{c_{ij} - \sum_k (x_{ij}^k + \bar{x}_{ij}^k)\} \ge 0 \qquad (i,j) \in A \quad (13)$$

$$0 \le t_i^k \perp \{\sum_{j|(i,j)\in A} x_{ij}^k - \sum_{j|(j,i)\in A} x_{ji}^k - q_i^k\} \ge 0 \; i \in N, k \in D \quad (15)$$

$$0 \le x_{ij} \perp \{x_{ij} - \sum_k x_{ij}^k\} \ge 0 \qquad (i,j) \in A \quad (16)$$

$$0 \le q_i^k(T) \perp \{(q_i^k(T) - \hat{q}_i^k).\exp(-\rho.t_i^k) + q_i^k(T).\exp(-\bar{\rho}.\bar{t}_i^k)\} \ge 0 \; i \in N, k \in D \quad (20)$$

$$0 \le \bar{x}_{ij}^k \perp (\bar{t}_j^k + \bar{t}_{ij}(X) + \bar{x}_{ij}.\partial\bar{t}_{ij}(X)/\partial\bar{x}_{ij} + \lambda_{ij} - \bar{t}_i^k) \ge 0 \; (i,j) \in A, \; k \in D \quad (21)$$

$$0 \le \bar{t}_i^k \perp \{\sum_{j|(i,j)\in A} \bar{x}_{ij}^k - \sum_{j|(j,i)\in A} \bar{x}_{ji}^k - \bar{q}_i^k\} \ge 0$$

$$i \in N, k \in D \quad (22)$$

$$0 \le \bar{x}_{ij} \perp \{\bar{x}_{ij} - \sum_k \bar{x}_{ij}^k\} \ge 0 \qquad (i,j) \in A \quad (23)$$

$$0 \le \bar{q}_i^k(T) \perp \{(\bar{q}_i^k(T) - \hat{q}_i^k).\exp(-\bar{\rho}.\bar{t}_i^k) + \bar{q}_i^k(T).\exp(-\rho.t_i^k)\} \ge 0 \qquad i \in N, k \in D \quad (24)$$

In the above NCP formulation, SO and UE traffic flows are represented by equations (12, 15, 16, 20) and (21, 22, 23, 24) respectively while equation (13) is common and it ties UE and SO flows sharing a common road together subject to the respective capacity (Note that delay function in equations (12, 21) also include traffic volumes of SO and UE).

It is also important to note that, the above formulation can similarly be expanded to accommodate more vehicle classes, therefore, a multiclass formulation is attainable. Before we progress to the numerical tests, in the sequel we discuss the feasibility, existence, and uniqueness of solutions (A formal proof is provided in Appendix A).

3.3. Capacity feasibility, existence, and uniqueness

Capacity feasibility: It is worth noting that, one of the chronic issues in the capacitated traffic assignment is to deal with the infeasibility concern. There might be a situation in which the total demand far exceeds the road capacity and the algorithm is puzzled how to handle the excess demand. One intuitive and easy remedy are to add dummy capacity to the network. To this end, we added a dummy node to which all origin and destination nodes are connected via dummy roads. The dummy roads are associated with a large value (infinity value) for their respective capacities and free-flow-travel-times (i.e. $C_{ij} = +\infty$, $\alpha_{ij} = +\infty$). By doing so, the excessive demand will find the dummy links as the last resort.

Existence and uniqueness: Existence and uniqueness of NCP solution under some mild conditions which are the case for the TAP have been proven in the literature (Aashtiani, 1979; Simsek et al., 2005).

Existence: the proof of the existence of a solution is generally put forward based on Brouwer's fixed-point theorem (Brouwer, 1912) as provided by Aashtiani and Magnanti (1981). The set of feasible solutions is closed and bounded for two reasons: (i) link flows lie in a compact set (owed much to the capacity constraints); and (ii) the feasible set is convex because the constraints are linear. It is also non-empty provided the link capacities are not too low, which is always the case when dummy links exist. The only requirement for the delay function is to be integrable and positive. If so, then the corresponding marginal delay function becomes positive and continuous. As the result, the second term in the compact format will be found continuous, non-empty, convex and compact feasible set. As for the conditions of the variable demand function, Aashtiani and Magnanti (1981) prove that the only requirements for the function are: being non-negative, continuous and bounded from above on the set of travel time which all holds for the logit formula (18). Hence the problem has a feasible solution.

Uniqueness: Aashtiani and Magnanti (1981) prove that establishing strictly monotone functions for the delay and variable demand over a strongly connected network is enough to guarantee uniqueness of links traffic volumes as well as their corresponding travel times.

4. Numerical results

The GAMS is a widely used optimization software to succinctly model and solve complicated problems. It is worth noting that many of transport planning problems can be formulated as optimization problems subject to a

This article is protected by copyright. All rights reserved.

Author Manuscript

TAP for which commercial optimization software can be found very "handy". As for the delay functions, we comply with a format proposed by the U.S. Bureau of Public Road (BPR) (Spiess, 1990) as:

$$t_{ij}(X) = \alpha_{ij} + \beta_{ij}.((x_{ij} + \bar{x}_{ij} + 0.1.(x_{ji} + \bar{x}_{ji}))/C_{ij})^4 \qquad (25)$$

where α_{ij}, β_{ij} are parameters to be calibrated by field survey data (α_{ij} can be interpreted as free flow travel time)? Note that, 0.1 is an arbitrary (positive) value to include the delay due to the traffic volume in the opposite directions which is the case at intersections. As can be seen, it is an asymmetric delay function where the travel time depends on the traffic flow of the respective link as well as that of the opposite direction. The marginal delay function can be computed as (an in-depth discussion is provided in Appendix B):

$$\tilde{t}_{ij}(X) = \alpha_{ij} + \beta_{ij}((x_{ij} + \bar{x}_{ij} + 0.1(x_{ji} + \bar{x}_{ji}))/C_{ij})^4 + 4\beta_{ij}\,\bar{x}_{ij}/C_{ij}{}^4.(x_{ij} + \bar{x}_{ij} + 0.1(x_{ji} + \bar{x}_{ji}))^3 \quad (i,j) \in A \ (26)$$

A desktop PC with a 3.70GHz CPU and 64 GB of RAM was employed. In this section, we report on two case studies, Sioux Falls and Melbourne. The Sioux Falls is still traditionally being used in the literature, especially when the papers present a new method. The Sioux-Falls network is perceived to be a common platform for scholars to communicate their results over a benchmark case study. Nevertheless, we've applied our method to the network of the Melbourne CBD and show the computational expenses.

4.1. Sioux-Falls

The Sioux-Falls network has 76 links, 24 nodes and a demand matrix of 528 nonzero entries. The GAMS code was executed and the computations elapsed around 6 seconds while a very fine convergence tolerance (1e-6) is upheld. We started the numerical analysis using identical disutility factors: $\rho = \bar{\rho} = 0.01$.

The travel demand of SO and UE were found 196.179 and 200.244 respectively, out of which 69 and 55 were found as excessive demand. The resulting roads' traffic volumes for both UE and SO traffic flows are presented in Table 1 and 2. It is clear that the SO is associated with marginal travel time which is higher than the travel time itself, therefore the disutility of SO become greater than the UE and so the mode share. There is a delicate observation: it is the UE trips who selfishly occupy the road space so much so less space is left to the SO which resulted in more of SO trips over the dummy links (69 of SO versus 55 of UE). Nevertheless, the total travel time experienced by SO versus UE is considerably lower. Figures 1 and 2 also depict traffic volumes distinguished based on SO and UE and saturated roads. As shown, the bottlenecks are shaped in the middle of the network in which roads reach their respective capacities.

We also carried out a sensitive analysis over disutility factor for the UE of $0.0101 \le \rho \le 0.011$, with the incremental pace of 0.0101 while the SO's disutility value remains the same $\bar{\rho} = 0.01$. The numerical results are presented in Table 3. Figures 3 and 6 shows some of the information provided in Tables 3 graphically. In Figure 3 (a) the variation of total time spent on the real road network (excluding the dummy roads) over incremental values of the disutility factors (rho) pertaining to UE flow are shown. It is obvious that as the disutility of UE increases the travel demand shifts to the SO. As the result, the total travel time of the SO increases while the trend for the UE is reverse. In this figure total travel time (i.e UE+SO) has also been depicted, in which it increases slightly in the second pace, while it gradually declines afterward. The gradual decrease of the total travel time on the network is an endorsement for pushing more demand towards the SO traffic pattern. For the abnormal behavior in the beginning of this graph, one possible answer can be sought in the principles of Braess Paradox, since the total travel time unexpectedly increases. Some studies have shown the variation of travel demands can significantly contribute to the occurrence or disappearance of the BP (Aashtiani and Poorzahedy, 2004; Nagurney, 2010).

Figure 3 (b) illustrates a variation of excessive travel demand with respect to the roads' capacities. As more demand shifts towards the SO, share of the SO versus UE on the dummy link constantly increases.

Figure 4 depicts average travel time per trips for both SO and UE. In the beginning of the graph, a similar abnormality is observed such that the average travel time for both SO and UE increases. From that point onward something interesting occurs: almost halfway through the graph, the average travel time of the SO is higher than the UE, though the gap is constantly closing. It is the situation in which the SO make a sacrifice by cooperatively seeking what is best for the entire road network not for them. Therefore the CVs (SO) incur slightly higher travel time compared to the selfish nCVs (UE). As the mode share of the SO increase and they dominate the road network, the trend tends to reverse such that from $\rho = 0.105$ onwards the CVs -though are in majority- incur much less travel time compared to their nCV counterparts.

4.2. Melbourne

The Melbourne dataset used in this study covers central business district (CBD) and surrounding areas which include 100 zones and 362 directional roads. The total travel demand consists of 9990 origin-destination pairs (non-zero values) which account for up to 59981.7 hourly trips. The roads capacities vary from 1200 to 4000 vehicle per hour while free-flow speeds vary from 40 to 80 kilometers per hour. We first run the GAMS code for a fixed and single travel demand matrix subject to UE traffic pattern and the same asymmetric BPR delay function as proposed in equation (25). By adopting convergence tolerance of 0.0001 (as termination criteria) the

This article is protected by copyright. All rights reserved.

computation lasted 9.1 minutes and successfully ended while no violation in the capacity constraints occurred. The traffic volume plot is shown in Figure 5 where at-capacity roads are red colored. Only 155 trips were found excessive due to the roads' capacities. The results indicate that average travel time is 3.29 minutes.

We also run the GAMS code for mixed SO and UE using the same disutility factors: $\rho = \bar{\rho} = 0.01$ which lasted almost 25 min. The split in the travel demand was found to be 49.5% SO and 50.5% UE. The average travel time of SO and UE cars were found to be 3.15 and 3.21 minutes respectively, which shows vehicle (of mixed SO+UE) are better off compared to their counterparts in the UE traffic pattern.

What is noticeable is the fact that the computational time is not promising at all. It is because the GAMS code is not a customized code for the traffic assignment (compared to the state-of-the-art methods such as TAPAS alike (Bar-Gera, 2010; Xie and Xie, 2014) (Note that, we are comparing our results with a method (TAPAS) developed for a simple TAP: no capacity constraint, no multiclass, no asymmetric delay function, no variable demand). To this end, (Aashtiani, 1979) proposed an efficient linearized base algorithm for multiclass traffic assignment which can be considered as a worthy line of research.

5. Conclusion

Autonomous and connected vehicle technology have seen an unprecedented surge of interest from a variety of disciplines such as computer science, electrical engineering, robotics, transportation science etc.

Connectedness is defined as communications amongst vehicles and the infrastructure which results in the precious real-time dataset. Accordingly, we've tapped into such dataset to draw a real-time map of traffic situations across the network: where are the congestion bottlenecks? what are the alternative routes? etc. The CVs are then supplied with reliable and real-time traffic advice aiming to push for a much more efficient traffic circulation. To this end, we've employed system optimal (SO) principles in which AVs are cooperatively seeking what is best for the entire road network. This idea is formulated using the nonlinear complementarity problem (NCP) in which a mixture of SO and the user equilibrium (UE) representing non-connected vehicles is enforced. In fact, we've proposed a new solution to the famous traffic assignment problem (TAP) which itself is a center of research in the literature associated with a large body of different algorithms.

Nevertheless, there are a number of realistic features that are largely overlooked in the literature to mathematically keep the TAP a tractable problem. To this end, in our NCP formulation, the following features are explicitly considered: (i) the roads' capacities, that is, the traffic load cannot exceed the capacity of the respective road, (ii) demand elasticity, that is, the travel demand is responsive to the changes of travel time (more precisely, we've employed logit formulation as a mode choice model between SO and UE), (iii) the NCP bears an inherent capability to consider multiclass traffic flows which otherwise remains a mathematically daunting task, (iv) in the NCP formulation, roads can be associated with a more comprehensive and general travel time function in which the delay is not merely a function of the flow of the respective road, but rather a function of the flow vector. As the result, one can better model delay function at junctions.

The proposed NCP is coded using GAMS and it is applied to the network of Sioux-Falls and Melbourne. It is clear that GAMS is not a customized algorithm for the TAP in general. In other words, from computational efficiency, the proposed methodology in this study cannot compete with some of the leading algorithms in the literature, notably TAPAS (Bar-Gera, 2010). This research seeks to achieve two main goals: (i) to highlight the potential benefit of the cooperative routing (SO) in the new era of the connected vehicle; and (ii) to present a mathematical formulation and a general solution algorithm for the TAP. Nevertheless, the next step to for this research is to customize an efficient solution algorithm tailored to large sized road networks. The linearization method proposed by Aashtiani and his colleague (Aashtiani, 1979; Babazadeh and Aashtiani, 2005; Shahpar et al., 2008) is a worthwhile effort (it is basically a path-based algorithm- that has recently shown a very promising performance, even comparable with origin-based methods (Olarte, 2009). Furthermore, path generation in our method is inevitable which is why the computational time of the Melbourne network became significant. As we have previously stated, the next step up from this current study is to customize an efficient solution algorithm for which linearization has proven to be highly efficient (Aashtiani, 1979). One important component of linearization is the path generation. Though the number of paths between an origin-destination pair can tend to infinity, only a few of them actually are considered by travelers (Aashtiani and Magnanti, 1981). Therefore, the linearization algorithm doesn't need to generate all the paths (Aashtiani, 1979)).

Since the ultimate goal is to guide the connected vehicles through the road network, future efforts are needed to be directed towards path-based methods to advice the vehicles on the best SO route (note, in the proposed method the routes of the SO and UE are not exclusively determined by equilibrium conditions alone; additional requirements, such as maximum entropy are required). To this end, again the linearization methods proposed for the complementarity problem have shown to be highly effective.

Moreover, it is important to acknowledge that the CVs are ought to suffer from higher travel time compared to their counterparts of the nCVs. However, as the market share of the CVs increases, these losses turns to net

Author Proof

This article is protected by copyright. All rights reserved.

gains. Investigation on the minimum market share of the SO-seeking vehicle to ensure an al-out SO traffic pattern is a worthy line of research.

The numerical results reported for the Sioux-Falls suggest that in the capacitated assignment, the connected vehicles take most of the effort (in terms of the delay time) to ensure equilibrium for the other vehicles. It is due to the fact that the CVs are exposed to additional travel time compared to their non-connected counterparts. As the results, in the case of capacity-infeasibility, the CVs are more likely to be seen in the dummy roads. To this end, it is worth noting that the above-mentioned concern is a fundamental drawback of the static traffic assignment methods. As a concrete example, consider a single road connected A to B with a capacity of 1000 vehicle per hour while the travel demand is 1500 for a peak hour time. By solving this example, the first 1000 vehicles drive through while the remainder (500 cars) are assigned to the dummy roads. In fact, the excess demand is forced to wait until the traffic load is gradually discharged. Such deficiencies have been attributed to the definition of the traffic equilibrium in the capacity constraints condition for which some scholars call a thorough revisit in the definition (Correa et al., 2004; Marcotte et al., 2004). Given such an obvious drawback, the dynamic traffic assignment (DTA) is gaining momentum among scholars and practitioners in which traffic assignment is carried out over a finer slice of time, rather than at peak hour. The inclusion of the dummy roads serves the purpose of providing a feasible solution for the problem at all times.

Accordingly, in light of the mandate of this study, a decent and valuable thread of research could be devoted to the DTA. The idea is to arrive at a dynamic model considering real-time data and being able to act promptly in response to sudden traffic incidents (note that Unfortunately, the DTA comes at a significantly higher computational cost compared to that of the static methods. For example, for a real sized network of Melbourne (+3000 zones), it takes couple of days to execute only a few iterations, whereas the static methods take a couple of minutes (say 3 minutes for the same Melbourne network) while it reaches a very fine convergence. An additional capacity constraint increases the computational costs fourfold (Larsson and Patriksson, 1999), which is still a tolerable cost (say 12 minutes for the Melbourne network). Furthermore, unlike the static methods, there is little consensus in the literature. In other words, the DTA is still in its infancy. As a result, in the absence of an efficient DTA model, the static methods are still the main model to reckon.). Such approach would unequivocally contribute to the safety aspect of autonomous driving as well traffic control practices.

Acknowledgement

The authors are very grateful to Prof Adeli and six anonymous reviewers for their insightful and constructive comments which greatly contributed to the content and presentation of the article.

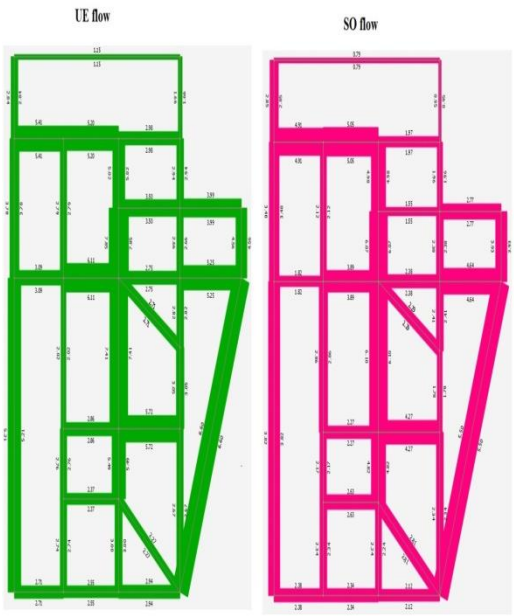


Figure 1, Sioux-Falls numerical results UE and SO traffic flow

Author Manuscript

This article is protected by copyright. All rights reserved.

Total flow: UE+SO flow
Saturated roads are red colored

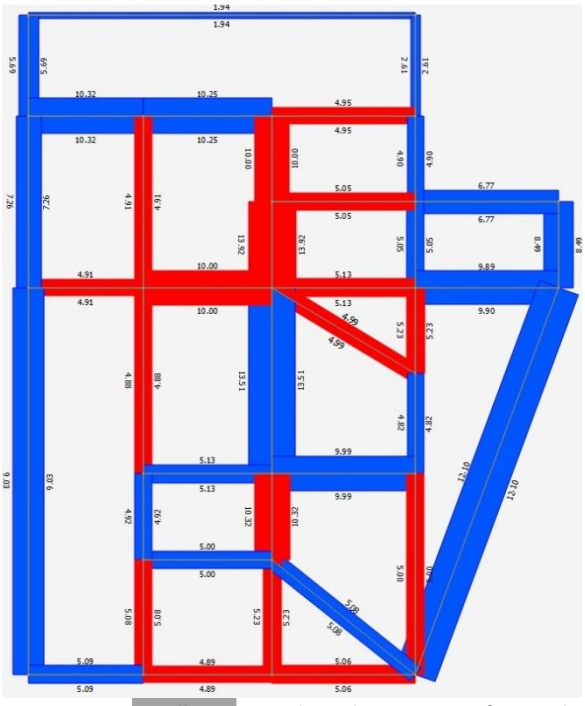


Figure 2, Sioux-Falls numerical results: mixture of UE and SO traffic flow

Author Manuscript

This article is protected by copyright. All rights reserved.

Table 1, Numerical results pertaining to disutility factors $\rho = \bar{\rho} = 0.01$ (excluding the dummy links)

node i	node j	α_{ij}	β_{ij}	Cap	vol-UE-ij	vol-SO-ij	vol-Total-ij	time-ij	vol-UE-ji	vol-SO-ji	vol-Total-ji	time-ji
1	2	6	0.9	25.9	1.152	0.79	1.943	6.00004	1.152	0.79	1.942	6.00004
1	3	4	0.6	23.403	2.839	2.848	5.688	4.00306	2.839	2.848	5.687	4.00306
2	6	5	0.75	4.958	1.664	0.95	2.614	5.08485	1.664	0.95	2.614	5.08485
3	4	4	0.6	17.11	5.411	4.905	10.316	4.11608	5.411	4.905	10.316	4.11608
3	12	4	0.6	23.403	3.78	3.475	7.255	4.00811	3.78	3.475	7.255	4.00811
4	5	2	0.3	17.783	5.203	5.045	10.249	2.04844	5.203	5.045	10.248	2.04844
4	11	6	0.9	4.909	2.794	2.115	4.909	7.31769	2.794	2.115	4.909	7.31769
5	6	4	0.6	4.948	2.975	1.973	4.948	4.87846	2.975	1.973	4.948	4.87846
5	9	5	0.75	10	5.019	4.981	10	6.09808	5.019	4.981	10	6.09808
6	8	2	0.3	4.899	2.942	1.956	4.899	2.43887	2.942	1.956	4.898	2.43887
7	8	3	0.45	7.842	3.994	2.772	6.766	3.36509	3.994	2.772	6.766	3.36509
7	18	2	0.3	23.403	4.563	3.929	8.491	2.00761	4.563	3.929	8.492	2.00761
8	9	10	1.5	5.05	3.495	1.555	5.05	12.19615	3.495	1.555	5.05	12.19615
8	16	5	0.75	5.046	2.663	2.383	5.046	6.09808	2.663	2.383	5.046	6.09808
9	10	3	0.45	13.916	7.851	6.065	13.916	3.65885	7.851	6.065	13.916	3.65885
10	11	5	0.75	10	6.106	3.894	10	6.09808	6.106	3.894	10	6.09808
10	15	6	0.9	13.512	7.409	6.103	13.512	7.31769	7.409	6.103	13.512	7.31769
10	16	5	0.75	5.133	2.749	2.385	5.134	6.09893	2.749	2.385	5.134	6.09893
10	17	8	1.2	4.993	2.706	2.287	4.993	9.75692	2.706	2.287	4.993	9.75692
11	12	6	0.9	4.909	3.086	1.823	4.909	7.31769	3.086	1.823	4.909	7.31769
11	14	4	0.6	4.876	2.017	2.86	4.877	4.87918	2.017	2.86	4.877	4.87918
12	13	3	0.45	25.9	5.206	3.82	9.026	3.00972	5.206	3.82	9.026	3.00972
13	24	4	0.6	5.091	2.707	2.384	5.091	4.87846	2.707	2.384	5.091	4.87846
14	15	5	0.75	5.128	2.857	2.27	5.127	6.09722	2.857	2.27	5.127	6.09722
14	23	4	0.6	4.925	2.757	2.168	4.925	4.87846	2.757	2.168	4.925	4.87846
15	19	4	0.6	15.651	5.725	4.27	9.995	4.14611	5.725	4.27	9.995	4.14611
15	22	4	0.6	10.315	5.491	4.824	10.315	4.87846	5.491	4.824	10.315	4.87846
16	17	2	0.3	5.23	2.824	2.406	5.23	2.43923	2.824	2.406	5.23	2.43923
16	18	3	0.45	19.68	5.253	4.643	9.895	3.04212	5.253	4.642	9.895	3.04211
17	19	2	0.3	4.824	3.046	1.778	4.824	2.43923	3.046	1.778	4.824	2.43923
18	20	4	0.6	23.403	6.6	5.5	12.1	4.06277	6.6	5.5	12.1	4.06277
19	20	4	0.6	5.003	2.667	2.336	5.003	4.87846	2.667	2.336	5.003	4.87846
20	21	6	0.9	5.06	2.94	2.12	5.06	7.31769	2.94	2.12	5.06	7.31769
20	22	5	0.75	5.076	2.267	2.809	5.076	6.09808	2.267	2.809	5.076	6.09808
21	22	2	0.3	5.23	2.995	2.235	5.23	2.43923	2.995	2.235	5.23	2.43923
21	24	3	0.45	4.885	2.549	2.337	4.885	3.65938	2.549	2.337	4.886	3.65938
22	23	4	0.6	5	2.37	2.63	5	4.87846	2.37	2.63	5	4.87846
23	24	2	0.3	5.078	2.738	2.341	5.078	2.43958	2.738	2.341	5.079	2.43958

Author Manuscript

This article is protected by copyright. All rights reserved.

Table 2, Numerical results pertaining to disutility factors $\rho = \bar{\rho} = 0.01$ (the dummy links)

node i	node j	α_{ij}	β_{ij}	Cap	vol-UE-ij	vol-SO-ij	vol-Total-ij	time-ij	vol-UE-ji	vol-SO-ji	vol-Total-ji	time-ji
1	25	999		10000	1.875	2.308	4.183	999	1.875	2.308	4.183	999
2	25	999		10000	0.605	0.923	1.528	999	0.605	0.923	1.528	999
3	25	999		10000	0.44	0.494	0.934	999	0.44	0.494	0.934	999
4	25	999		10000	1.87	2.408	4.278	999	1.87	2.408	4.278	999
5	25	999		10000	0.77	0.933	1.703	999	0.77	0.933	1.703	999
6	25	999		10000	0.77	0.77	1.54	999	0.77	0.77	1.54	999
7	25	999		10000	2.375	3.061	5.436	999	2.375	3.061	5.436	999
8	25	999		10000	2.789	4.745	7.534	999	2.789	4.745	7.534	999
9	25	999		10000	1.265	1.914	3.179	999	1.265	1.914	3.179	999
10	25	999		10000	4.081	7.226	11.307	999	4.081	7.226	11.307	999
11	25	999		10000	3.797	3.798	7.596	999	3.797	3.798	7.595	999
12	25	999		10000	3.22	3.38	6.6	999	3.22	3.38	6.6	999
13	25	999		10000	3.135	4.151	7.286	999	3.135	4.151	7.286	999
14	25	999		10000	2.584	3.049	5.633	999	2.584	3.049	5.633	999
15	25	999		10000	3.128	3.249	6.378	999	3.128	3.249	6.377	999
16	25	999		10000	4.125	4.167	8.292	999	4.125	4.167	8.292	999
17	25	999		10000	4.407	6.286	10.693	999	4.407	6.286	10.693	999
18	25	999		10000	0.482	0.927	1.41	999	0.482	0.927	1.409	999
19	25	999		10000	1.327	1.482	2.809	999	1.327	1.482	2.809	999
20	25	999		10000	3.653	4.225	7.878	999	3.653	4.225	7.878	999
21	25	999		10000	1.944	2.034	3.978	999	1.944	2.034	3.978	999
22	25	999		10000	2.86	3.753	6.613	999	2.86	3.753	6.613	999
23	25	999		10000	2.131	2.468	4.599	999	2.131	2.468	4.599	999
24	25	999		10000	1.314	1.321	2.634	999	1.314	1.321	2.635	999

Author Manuscript

This article is protected by copyright. All rights reserved.

Table 3 Sensitivity analysis on variation of UE's disutility factor

rho	UE dmd	SO dms	T dmd	UE(T) t*v	UE(R) t*v	UE(ex) dmd	SO(T) t*v	SO(R) t*v	SO(ex)d md	UE dmd R	SO dmd R	UE avg time	SO avg time	total time	Total R Vol on the net
0.01	200.244	196.179	396.423	111134.	1350.79	54.947	139092.	1086.34	69.072	145.297	127.107	9.297	8.547	2437.14	272.404
0.010	185.504	210.918	396.422	98965.2	1311.01	48.876	151329.	1198.72	75.141	136.628	135.778	9.596	8.829	2509.74	272.406
0.010	170.997	225.426	396.423	86008.4	1187.32	42.453	164210.	1239.20	81.567	128.544	143.859	9.237	8.614	2426.52	272.403
0.010	156.95	239.471	396.421	74865.3	1064.22	36.938	175283.	1284.85	87.087	120.013	152.385	8.868	8.432	2349.07	272.397
0.010	143.591	252.831	396.422	65070.5	949.734	32.093	185004.	1334.53	91.927	111.499	160.904	8.518	8.294	2284.27	272.403
0.010	130.999	265.383	396.382	55995.5	851.718	27.6	193952.	1385.32	96.38	103.4	169.003	8.237	8.197	2237.04	272.403
0.010	119.082	277.075	396.157	47370.8	768.534	23.325	202100.	1435.78	100.433	95.758	176.642	8.026	8.128	2204.31	272.4
0.010	108.693	287.729	396.422	40158.3	695.876	19.751	209805.	1480.85	104.267	88.942	183.463	7.824	8.072	2176.72	272.405
0.01	099.013	297.408	396.421	33187.7	635.316	16.293	216762.	1523.54	107.727	82.721	189.681	7.68	8.032	2158.86	272.402
0.01	090.42	306.095	396.515	27637.9	579.013	13.543	222484.	1564.41	110.571	76.877	195.525	7.532	8.001	2143.43	272.402
0.01	182.313	313.795	396.108	22400.0	527.993	10.947	226903.	1606.38	112.762	71.366	201.034	7.398	7.991	2134.37	272.4

Remarks: rho refers to disutility factor of UE that is: ρ ; dmd: demand; T dmd: total demand; UE(T) t*v: sigma of time multiplied by traffic volume across all the links including the dummy links pertaining to UE;
UE(R) t*v: sigma of time multiplied by traffic volume across only real links (excluding the dummy links) pertaining to UE; UE(ex)dmd, excessive demand pertaining to UE

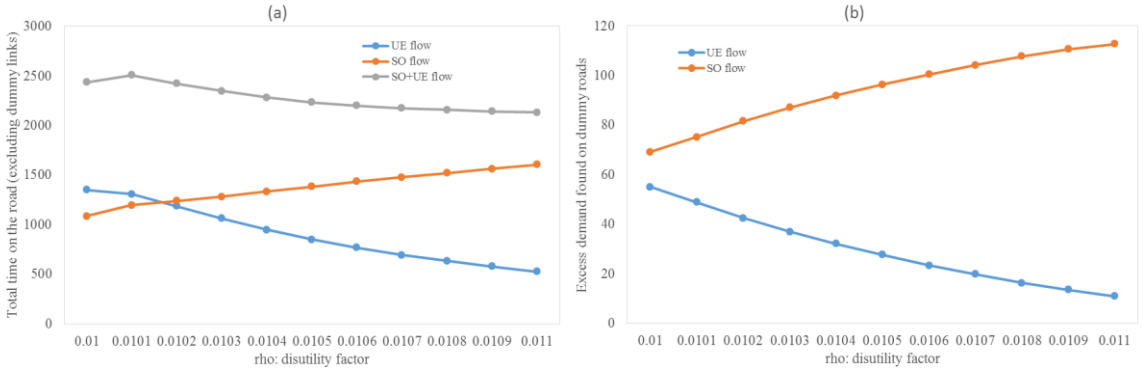


Figure 3 Sioux-Falls, results of the disutility factor (a) total time on the network (excluding dummy links), (b) Excess

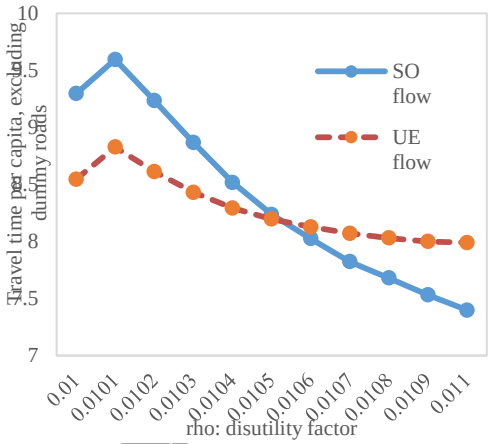


Figure 4 Sioux-Falls, variation of the result with respect to UE's disutility factor: average of travel time on the real network, excluding dummy roads

This article is protected by copyright. All rights reserved.



Figure 5, Melbourne's central area, user equilibrium hourly traffic flow (at-capacity roads are red colored),

References

Aashtiani, H.Z., 1979. The multi-modal traffic assignment problem. PhD dissertation, Massachusetts Institute of Technology.

Aashtiani, H.Z., Magnanti, T.L., 1981. Equilibria on a congested transportation network. *SIAM Journal on Algebraic Discrete Methods* 2, 213-226.

Aashtiani, H.Z., Poorzahedy, H., 2004. Braess' phenomenon in the management of networks and dissociation of equilibrium concepts. *Transportation planning and technology* 27, 469-482.

Adeli, H., Karim, A., 2000. Fuzzy-wavelet RBFNN model for freeway incident detection. *Journal of Transportation Engineering* 126, 464-471.

Adeli, H., Karim, A., 2005. *Wavelets in intelligent transportation systems*. John Wiley & Sons.

Adeli, H., Samant, A., 2000. An adaptive conjugate gradient neural network–wavelet model for traffic incident detection. *Computer-Aided Civil and Infrastructure Engineering* 15, 251-260.

Anderson, J.M., Nidhi, K., Stanley, K.D., Sorensen, P., Samaras, C., Oluwatola, O.A., 2014. *Autonomous vehicle technology: A guide for policymakers*. Rand Corporation.

Babazadeh, A., Aashtiani, H., 2005. Algorithm for equilibrium transit assignment problem. *Transportation Research Record: Journal of the Transportation Research Board*, 227-235.

Babonneau, F., Vial, J.-P., 2008. An efficient method to compute traffic assignment problems with elastic demands. *Transportation Science* 42, 249-260.

Bagloee, S.A., Sarvi, M., 2015. Heuristic Approach to Capacitated Traffic Assignment Problem for Large-Scale Transport Networks. *Transportation Research Record: Journal of the Transportation Research Board*, 1-11.

Bagloee, S.A., Sarvi, M., Patriksson, M., 2016a. A Hybrid Branch-and-Bound and Benders Decomposition Algorithm for the Network Design Problem. *Computer-Aided Civil and Infrastructure Engineering*.

Bagloee, S.A., Tavana, M., Asadi, M., Oliver, T., 2016b. Autonomous vehicles: challenges, opportunities, and future implications for transportation policies. *Journal of Modern Transportation* 24, 284-303.

Bar-Gera, H., 2010. Traffic assignment by paired alternative segments. *Transportation Research Part B: Methodological* 44, 1022-1046.

Bar-Gera, H., Boyce, D., 1999. Route flow entropy maximization in origin-based traffic assignment, *14th International Symposium on Transportation and Traffic Theory*. Elsevier Science, Oxford, U.K., Jerusalem, Israel, pp. 397–415.

Beckmann, M., McGuire, C., Winsten, C.B., 1956. *Studies in the Economics of Transportation*, Yale University Press, New Haven, CT.

Bell, M.G., Shield, C.M., Busch, F., Kruse, G., 1997. A stochastic user equilibrium path flow estimator. *Transportation Research Part C: Emerging Technologies* 5, 197-210.

Bennett, L.D., 1993. The existence of equivalent mathematical programs for certain mixed equilibrium traffic assignment problems. *European Journal of Operational Research* 71, 177-187.

Boyce, D., 2014. Network equilibrium models for urban transport, in: Fischer, M.M., Nijkamp, P. (Eds.), *Handbook of Regional Science*. Springer Berlin Heidelberg, pp. 759-786.

Brooke, A., Kendrick, D., Meeraus, A., 1996. *GAMS Release 2.25: A user's guide*. GAMS Development Corporation Washington, DC.

Brouwer, L.E.J., 1912. Über Abbildung von Mannigfaltigkeiten. *Math. Ann.,71* Berichtigung ebd. S. 1912; 598, 97-115.

Castillo, E., Calvino, A., Sánchez-Cambronero, S., Lo, H.K., 2013a. A multiclass user equilibrium model considering overtaking across classes. *IEEE Transactions on Intelligent Transportation Systems* 14, 928-942.

This article is protected by copyright. All rights reserved.

Castillo, E., Calviño, A., Sánchez-Cambronero, S., Nogal, M., Rivas, A., 2013b. A percentile system optimization approach with and without path enumeration. *Computers & Operations Research* 40, 2711-2723.
Castillo, E., Nogal, M., Calviño, A., Rivas, A., Lo, H.K., 2013c. A model for continuous dynamic network loading problem with different overtaking class users. *Journal of Intelligent Transportation Systems* 17, 328-350.
Chen, A., Zhou, Z., Ryu, S., 2011a. Modeling physical and environmental side constraints in traffic equilibrium problem. *International Journal of Sustainable Transportation* 5, 172-197.
Chen, B.Y., Lam, W.H.K., Sumalee, A., Shao, H., 2011b. An efficient solution algorithm for solving multi-class reliability-based traffic assignment problem. *Mathematical and Computer Modelling* 54, 1428-1439.
Correa, J.R., Schulz, A.S., Stier-Moses, N.E., 2004. Selfish routing in capacitated networks. *Mathematics of Operations Research* 29, 961-976.
Cottle, R.W., 1964. Nonlinear Programs with Positively Bounded Jacobians, *Department of Mathematics*. University of California (Berkeley, CA, 1964).
Dafermos, S., 1982. The general multimodal network equilibrium problem with elastic demand. *Networks* 12, 57-72.
Dafermos, S.C., 1972. The traffic assignment problem for multiclass-user transportation networks. *Transportation Science* 6, 73-87.
Dafermos, S.C., Sparrow, F.T., 1969. The traffic assignment problem for a general network. *Journal of Research of the National Bureau of Standards B* 73, 91-118.
Dresner, K.M., Stone, P., 2007. Sharing the Road: Autonomous Vehicles Meet Human Drivers, *IJCAI*, pp. 1263-1268.
Facchinei, F., Pang, J.-S., 2003. *Finite-dimensional variational inequalities and complementarity problems*. Springer.
Ferrari, P., 1997. Capacity constraints in urban transport networks. *Transportation Research Part B: Methodological* 31, 291-301.
Ferris, M.C., Meeraus, A., Rutherford, T.F., 1999. Computing Wardropian equilibria in a complementarity framework. *Optimization Methods and Software* 10, 669-685.
Ferris, M.C., Munson, T.S., 2014. PATH 4.7. *Gams Corporation. Washington*.
Ferris, M.C., Pang, J.-S., 1997. Engineering and economic applications of complementarity problems. *Siam Review* 39, 669-713.
FHWA, 2002. Status of the Nation's Highways, Bridges and Transit: Conditions and Performance. . *US Department of Transportation* Washington, DC,.
Fisk, C.S., Boyce, D.E., 1983. Alternative variational inequality formulations of the network equilibrium-travel choice problem. *Transportation Science* 17, 454-463.
Florian, M., Morosan, C.D., 2014. On uniqueness and proportionality in multi-class equilibrium assignment. *Transportation Research Part B: Methodological* 70, 173-185.
Folsom, T.C., 2011. Social ramifications of autonomous urban land vehicles, *IEEE International Symposium on Technology and Society*.
Gabriel, S.A., Conejo, A.J., Fuller, J.D., Hobbs, B.F., Ruiz, C., 2012. *Complementarity modeling in energy markets*. Springer.
Gao, Z., Wu, J., Sun, H., 2005. Solution algorithm for the bi-level discrete network design problem. *Transportation Research Part B: Methodological* 39, 479-495.
Gartner, N.H., 1980a. Optimal Traffic Assignment with Elastic Demands: A Review Part I. Analysis Framework. *Transportation Science* 14, 174-191.
Gartner, N.H., 1980b. Optimal Traffic Assignment with Elastic Demands: A Review Part II. Algorithmic Approaches. *Transportation Science* 14, 192-208.
Ghosh-Dastidar, S., Adeli, H., 2003. Wavelet-Clustering-Neural Network Model for Freeway Incident Detection. *Computer-Aided Civil and Infrastructure Engineering* 18, 325-338.
Harker, P.T., 1988. Multiple equilibrium behaviors on networks. *Transportation Science* 22, 39-46.
Harker, P.T., Pang, J.-S., 1990. Finite-dimensional variational inequality and nonlinear complementarity problems: a survey of theory, algorithms and applications. *Mathematical programming* 48, 161-220.
Hartman, P., Stampacchia, G., 1966. On some non-linear elliptic differential-functional equations. *Acta mathematica* 115, 271-310.
Haurie, A., Marcotte, P., 1985. On the relationship between Nash–Cournot and Wardrop equilibria. *Networks* 15, 295–308.
Hearn, D., 1980. Bounding flows in traffic assignment models. *Technical report* Research Report 80-4.
Hearn, D., Ribera, J., 1980. Bounded flow equilibrium problems by penalty methods, *Proceedings of IEEE International Conference on Circuits and Computers*, pp. 162-166.

This article is protected by copyright. All rights reserved.

Huang, H.-J., 2002. Pricing and logit-based mode choice models of a transit and highway system with elastic demand. *European Journal of Operational Research* 140, 562-570.
Inouye, H., 1987. Traffic equilibria and its solution in congested road networks, *IFACIFIPIFORS CONFERENCE ON CONTROL IN*.
Intriligator, M.D., 1971. *Mathematical optimization and economic theory*. Siam.
Karim, A., Adeli, H., 2002a. Comparison of fuzzy-wavelet radial basis function neural network freeway incident detection model with California algorithm. *Journal of Transportation Engineering* 128, 21-30.
Karim, A., Adeli, H., 2002b. Incident detection algorithm using wavelet energy representation of traffic patterns. *Journal of Transportation Engineering* 128, 232-242.
Karim, A., Adeli, H., 2003. Fast automatic incident detection on urban and rural freeways using wavelet energy algorithm. *Journal of Transportation Engineering* 129, 57-68.
Larsson, T., Patriksson, M., 1992. Simplicial decomposition with disaggregated representation for the traffic assignment problem. *Transportation Science* 26, 4-17.
Larsson, T., Patriksson, M., 1995. An augmented Lagrangean dual algorithm for link capacity side constrained traffic assignment problems. *Transportation Research Part B: Methodological* 29, 433-455.
Larsson, T., Patriksson, M., 1999. Side constrained traffic equilibrium models— analysis, computation and applications. *Transportation Research Part B: Methodological* 33, 233-264.
Larsson, T., Patriksson, M., Rydergren, C., 2004. A column generation procedure for the side constrained traffic equilibrium problem. *Transportation Research Part B: Methodological* 38, 17-38.
Li, S.E., Deng, K., Zheng, Y., Peng, H., 2015. Effect of Pulse-and-Glide Strategy on Traffic Flow for a Platoon of Mixed Automated and Manually Driven Vehicles. *Computer-Aided Civil and Infrastructure Engineering* 30, 892-905.
Lo, H.K., Szeto, W., 2002. A methodology for sustainable traveler information services. *Transportation Research Part B: Methodological* 36, 113-130.
Maher, M., Hughes, P., 1995. Estimation of the potential benefits from an ATT system using a multiple user class stochastic user equilibrium assignment model, *Applications of Advanced Technologies in Transportation Engineering (1995)*. ASCE, pp. 700-704.
Marcotte, P., Nguyen, S., Schoeb, A., 2004. A strategic flow model of traffic assignment in static capacitated networks. *Operations Research* 52, 191-212.
Marcotte, P., Wynter, L., 2004. A new look at the multiclass network equilibrium problem. *Transportation Science* 38, 282-292.
Morowati-Shalilvand, S., Mehri-Tekmeh, J., 2013. An extended origin-based method for solving capacitated traffic assignment problem. *Acta Universitatis Apulensis*, 169-186.
Muijlwijk, H., 2012. Static traffic Assignment with Junction Modelling.
Nagurney, A., 2000. A multiclass, multicriteria traffic network equilibrium model. *Mathematical and Computer Modelling* 32, 393-411.
Nagurney, A., 2010. The negation of the Braess paradox as demand increases: The wisdom of crowds in transportation networks. *EPL (Europhysics Letters)* 91, 48002.
Nagurney, A., Dong, J., 2002. A multiclass, multicriteria traffic network equilibrium model with elastic demand. *Transportation Research Part B: Methodological* 36, 445-469.
NHTSA, 2012. 2010 motor vehicle crashes: Overview. *US Department of Transportation, Washington, DC, Research Note DOT HS* 811, 552.
Nie, Y., Zhang, H., Lee, D.-H., 2004. Models and algorithms for the traffic assignment problem with link capacity constraints. *Transportation Research Part B: Methodological* 38, 285-312.
Noriega, Y., Florian, M.A., 2007. *Algorithmic approaches for asymmetric multi-class network equilibrium problems with different class delay relationships*. CIRRELT.
Olarte, R.E., 2009. *A Comparison Between an Origin Based Method and a Nonlinear Complementarity Based Method for Solving the Traffic Assignment Problem*. ProQuest.
Patriksson, M., 2004. Algorithms for computing traffic equilibria. *Netw Spat Econ* 4, 23-38.
Patriksson, P., 1994. *The traffic assignment problem: models and methods*, VSP BV, The Netherlands.
Facsimile reproduction published in 2014 by Dover Publications, Inc., Mineola, New York, NY, USA.
Piao, J., McDonald, M., 2008. Advanced driver assistance systems from autonomous to cooperative approach. *Transport Reviews* 28, 659-684.
Prashker, J.N., Toledo, T., 2004. A gradient projection algorithm for side-constrained traffic assignment. *European Journal of Transport and Infrastructure Research* 4, 177-193.
Roughgarden, T., Tardos, É., 2002. How bad is selfish routing? *Journal of the ACM (JACM)* 49, 236-259.
Ryu, S., Chen, A., Choi, K., 2014a. A modified gradient projection algorithm for solving the elastic demand traffic assignment problem. *Computers & Operations Research* 47, 61-71.

This article is protected by copyright. All rights reserved.

Ryu, S., Chen, A., Xu, X., Choi, K., 2014b. A Dual Approach for Solving the Combined Distribution and Assignment Problem with Link Capacity Constraints. *Netw Spat Econ*, 1-26.
Samant, A., Adeli, H., 2000. Feature extraction for traffic incident detection using wavelet transform and linear discriminant analysis. *Computer-Aided Civil and Infrastructure Engineering* 15, 241-250.
Samant, A., Adeli, H., 2001. Enhancing Neural Network Traffic Incident-Detection Algorithms Using Wavelets. *Computer-Aided Civil and Infrastructure Engineering* 16, 239-245.
Sarvi, M., Kuwahara, M., 2008. Using ITS to improve the capacity of freeway merging sections by transferring freight vehicles. *Intelligent Transportation Systems, IEEE Transactions on* 9, 580-588.
Sarvi, M., Kuwahara, M., Ceder, A., 2004. Freeway ramp merging phenomena in congested traffic using simulation combined with a driving simulator. *Computer-Aided Civil and Infrastructure Engineering* 19, 351-363.
Shahpar, A.H., Aashtiani, H.Z., Babazadeh, A., 2008. Dynamic penalty function method for the side constrained traffic assignment problem. *Applied Mathematics and Computation* 206, 332-345.
Sheffi, Y., 1985. *Urban transportation networks: equilibrium analysis with mathematical programming methods*. Prentice-Hall, Inc., Englewood Cliffs, New Jersey
Simsek, A., Ozdaglar, A., Acemoglu, D., 2005. On the uniqueness of solutions for nonlinear and mixed complementarity problems. Citeseer.
Spiess, H., 1990. Technical note—Conical volume-delay functions. *Transportation Science* 24, 153-158.
Van Vuren, T., Van Vliet, D., Smith, M.J., 1990. Combined equilibrium in a network with partial route guidance, *Traffic control methods. Proceedings of the fifth NG foundation conference, Santa Barbara, Californi*.
Van Vuren, T., Watling, D., 1991. *Multiple user class assignment model for route guidance*.
Xie, J., Xie, C., 2014. An improved TAPAS algorithm for the traffic assignment problem, *Intelligent Transportation Systems (ITSC), 2014 IEEE 17th International Conference on*. IEEE, pp. 2336-2341.
Xie, J., Xie, C., 2015. Origin-Based Algorithms for Traffic Assignment: Algorithmic Structure, Complexity Analysis, and Convergence Performance, *Transportation Research Board 94th Annual Meeting*.
Yang, H., 1998. Multiple equilibrium behaviors and advanced traveler information systems with endogenous market penetration. *Transportation Research Part B: Methodological* 32, 205-218.
Yang, H., Bell, M.G., 1997. Traffic restraint, road pricing and network equilibrium. *Transportation Research Part B: Methodological* 31, 303-314.
Yang, H., Yagar, S., 1994. Traffic assignment and traffic control in general freeway-arterial corridor systems. *Transportation Research Part B: Methodological* 28, 463-486.
Yang, H., Yagar, S., 1995. Traffic assignment and signal control in saturated road networks. *Transportation Research Part A: Policy and Practice* 29, 125-139.
Yang, H., Zhang, X., Meng, Q., 2007. Stackelberg games and multiple equilibrium behaviors on networks. *Transportation Research Part B: Methodological* 41, 841-861.
Yin, Y., Yang, H., 2003. Simultaneous determination of the equilibrium market penetration and compliance rate of advanced traveler information systems. *Transportation Research Part A: Policy and Practice* 37, 165-181.
Yong, L., 2010. Nonlinear complementarity problem and solution methods, *Artificial Intelligence and Computational Intelligence*. Springer, pp. 461-469.
Zhang, G., Chen, J., 2010. Solving multi-class traffic assignment problem with genetic algorithm, *Computational Intelligence and Natural Computing Proceedings (CINC), 2010 Second International Conference on*. IEEE, pp. 229-232.

Appendix A: Proofs of existence and uniqueness:

The inclusion of multiclass into the traffic assignment comes at the cost of facing a nonconvex optimization with asymmetric and non-separable delay functions in which the feasibility and uniqueness of the solutions have yet to be addressed (Florian and Morosan, 2014). In this section, we prove that strict monotonicity of the functions (delay and demand) is enough to guarantee existence and uniqueness of the solutions from a complementarity formulation.

The proof of existence and uniqueness of the solutions for the multiclass traffic assignment are made based on concrete proof developed by (Aashtiani and Magnanti, 1981). To do so, we reframe the proposed NCP formulation to the notations and structures set up by them. Let us first state the complementarity formulation provided by (Aashtiani and Magnanti, 1981) (for the sake of clarity, we have used bold font for the notations used by them):

$$(\boldsymbol{T}_p(\boldsymbol{h}) - \boldsymbol{u}_i)\boldsymbol{h}_p = 0 \qquad \boldsymbol{p} \in \boldsymbol{P}_i,\ \boldsymbol{i} \in \boldsymbol{I} \qquad \text{(A1)}$$

$$\boldsymbol{T}_p(\boldsymbol{h}) - \boldsymbol{u}_i \geq 0 \qquad \boldsymbol{p} \in \boldsymbol{P}_i,\ \boldsymbol{i} \in \boldsymbol{I} \qquad \text{(A2)}$$

This article is protected by copyright. All rights reserved.

$\sum_{p\in P_i} h_p - D_i(u) = 0 \;\; i \in I$ (A3)

$h \geq 0$ (A4)

$u \geq 0$ (A5)

where I is the set of OD pairs; P_i is the set of "available" paths for the OD pair i; h_p is the flow on path p; h is the vector of $\{h_p\}$ with dimension $n1 = \sum_{i\in I} |P_i|$; u_i is the minimum travel time for the OD pair i; u is the vector of $\{u_i\}$ with dimension $n2 = |I|$; $D_i(u)$ is the demand function for O-D pair i, $D_i : R_+^{n2} \to R_+^1$; $T_p(h)$ is the delay time function for path p, $T_p(h) : R_+^{n1} \to R_+^1$; as an special case of additive model we have $T_p(h) = \sum_{a\in A} \delta_{ap} t_a^i(h)$ for all $p \in P_i$ and $i \in I$; δ_{ap} is 1 if link a is in path p and 0 otherwise; and finally $t_a^i(h)$ is the delay function for link $a \in A$ (A is the set of links) and OD pair i (note $t_a^i : R_+^{n1} \to R_+^1$). Aashtiani and Magnanti (1981) established two important theorems setting forward the necessary conditions for the existence and uniqueness of the solutions.

Theorem 5.6. (Existence). Suppose that set of link A constitutes a strongly connected network (that is for every OD pair there exist at least one path), also suppose that for all $p \in P_i$, the $T_p(h) : R_+^{n1} \to R_+^1$ is a continuous function and that, for all $h \in R_+^{n1}$, $T_p(h) > \varepsilon$ for some $\varepsilon \geq 0$. Also suppose that for all $i \in I$, $D_i : R_+^{n2} \to R_+^1$ is a nonnegative continuous function that is bounded from above on the set $\{u \in R_+^{n2} : u_i \geq \varepsilon$ for all $i\}$. Then the traffic equilibrium system (A1-A5) has a solution.

Theorem 6.1. (Uniqueness). Suppose that set of link A constitutes a strongly connected network, also suppose that t, the vector of the volume-delay functions, and $-D$, the vector of the negative of the demand functions, are strictly monotone. Then the arc volumes v and the travel time vector u for the additive traffic equilibrium problem (A1-A5) are unique, and the set of equilibrium path flows is convex.

We can now rewrite our proposed formulation ((12), (13), (15), (16), (20)..(24)) as per the structure of the (A1-A5)

Let us first consider the UE-pertaining equations: we start with equation (12): (i) The first three terms in the bracket (i.e. $t_j^k + t_{ij}(X) + \lambda_{ij}$) can be rewritten as $T_p(h)$ which is the (inflated) travel time of path p from node i to destination k as a function of the flows; (ii) With the same token x_{ij}^k can be rewritten as h_p which is the flow of path p; (iii) Moreover, the set I can be defined as $I = N \times N$. Similarly, t_i^k can be rewritten as u_i, the minimum travel time from node i to destination k, denoted by i. Consider the capacity constraint equations (13): for each road (i, j) we can assume a dummy link connecting i to j associated with $t_{ij}(X) + \lambda_{ij}$ as travel time which is devised to absorb excessive traffic volume ($c_{ij} - \sum_k (x_{ij}^k + \bar{x}_{ij}^k)$). Hence λ_{ij} can be rearranged as: $0 + t_{ij}(X) + \lambda_{ij} - t_{ij}(X)$ which is now analogous to Eq.(12). Moreover, for the previously discussed dummy link, we define a strictly monotone demand function of $t_{ij}(X) + \lambda_{ij}$ such that it yields ε (a very trivial positive value) and it tends to zero as the travel time increases. Equations (15), (16) and (20) ensure flow is conservative at each node and the flows meet the (variable) demand. Therefore, these equations can be replaced with (A3).

In a similar fashion, SO-pertaining equations can also be rewritten following the Aashtiani-Magnanti formulation. Consequently, we are now able to establish existence and uniqueness conditions based on Theorems 5.6 and 6.1: (i) D is a vector of non-negative functions bounded from the above, (ii) both t and $-D$ are vectors of the continuous and strictly monotone functions on the set of travel time u.

It is important to highlight the necessity of strict monotonicity of the demand function to ensure the uniqueness of the solutions. To this end, Marcotte and Wynter (2004) have numerically shown that a merely monotone demand function (fixed demand functions), along with strictly monotone travel time functions, may result in multiple solutions.

Harker (1988) provides shorter proof for the existence and uniqueness of the solutions (for the mixed UE-SO traffic pattern) that also calls for strict monotonicity of the travel time and demand functions. Nevertheless, he assumes the travel times are – merely - functions of their respective links' flows. However, in the context of his work, we in this paper extended the proof, as well as the solution methodology in two aspects: (i) the travel times are functions of the vector of links' flows which enables us to model delay times at intersections or two-

This article is protected by copyright. All rights reserved.

way passages; and (ii) the capacity constraint is also added into the proof and solution method to enhance the realism of the model.

Appendix B: Cournot flow: A key assumption here, is converting the partial SO problem into a UE one based on marginal cost. This transformation is known to be valid when the flow is homogeneous. But, does this remain true in the multiclass context? If one assumes that the flow of the other class is given, it makes sense but, what about solving them simultaneously? This may introduce correlations between the two flow variables in the derivation. To answer these questions, let us expand the mathematical expression of the marginal cost as follows: $\tilde{t}_{ij}(X) = t_{ij}(X) + \bar{x}_{ij} . dt_{ij}(X)/d\bar{x}_{ij}$. For ease of notation, without loss of generality, let us assume $t_{ij}(X) = t_{ij}(\bar{x}_{ij}, x_{ij})$. The derivative term can then be expanded as: $dt_{ij}(X)/d\bar{x}_{ij} = d\bar{x}_{ij}/d\bar{x}_{ij} . \partial t(\bar{x}_{ij}, x_{ij})/\partial\bar{x}_{ij} + dx_{ij}/d\bar{x}_{ij} . \partial t(\bar{x}_{ij}, x_{ij})/\partial x_{ij}$. By assuming $dx_{ij}/d\bar{x}_{ij} = 0$ the equation is shorten to $\partial t(\bar{x}_{ij}, x_{ij})/\partial\bar{x}_{ij}$. The assumption of $dx_{ij}/d\bar{x}_{ij} = 0$ suggests that the changes of the SO flow precipitate no significant reactions from the UE traffic, even though some such reactions may well occur (Bennett, 1993). This is a clear example of what is known as "mixed equilibrium with a Cournot altruistic controller" (Bennett, 1993). The term 'Cournot' is used in a way that is consistent with its standard interpretation in economic theory: a Cournot equilibrium represents a set of output levels such that each firm's output is maximized under the assumption of 'zero conjectural variation, which is defined by Intriligator (1971). In other words, each competing firm assumes that a variation in its own output will not have any effect on its competitors' outputs. In terms of traffic flow, it can be interpreted as follows: some vehicles follow the user equilibrium criteria whilst the remainder are directed by a Cournot altruistic controller whose objective is to minimize the total cost of travel for all traffic on the network. The Cournot controller assumes that while the percentage of the connected vehicles is not significant", there will be no reactions from the traffic operating under the user equilibrium criteria, even though some such reactions may well occur. As the market share of the connected vehicles increases, the above assumption $dx_{ij}/d\bar{x}_{ij} = 0$ will no longer hold true. To this end, some scholars suggest that the problem should be rewritten as a bilevel programming problem in which the controller acts in a 'Stackelberg', rather than a Cournot, manner: when making routing decisions, a Stackelberg controller assumes that a change in the distribution of the traffic under his direction may well provoke some changes in the distribution of all the other traffic on the network, and so he/she seeks to take into account such reactions (Bennett, 1993).

Author Manuscript

This article is protected by copyright. All rights reserved.