# Optimal Routing for Autonomous Vehicles in a Mixed Congested Network Considering Fairness

Ali Davoodi<sup>1</sup>, Mark Wallace<sup>1</sup>, Daniel Harabor<sup>1</sup> and Dong Ngoduy<sup>2</sup>

*Abstract*— The field of transportation has undergone a revolutionary transformation with the advent of Autonomous Vehicles (AVs). The ability to centrally control the route choices of these vehicles presents a promising opportunity to minimize total travel time within networks and effectively alleviate congestion. Recent studies have modelled Human-Driven Vehicles (HDVs) as selfish drivers who prioritize user-optimum routes, while AVs strive to optimize the system by reducing congestion. However, this routing approach may be perceived as unfair by AV users, leading to potential challenges. In this paper, we propose a heuristic framework to solve the dynamic mixed traffic flow problem using a simulation-based simulator and address this issue by providing equitable paths to AVs. These routes not only result in a significant reduction in the total system travel time (TSTT) but also minimize the need for AVs to make excessive compromises.

## I. INTRODUCTION

In a transportation network, drivers selfishly seek to minimize their individual travel time. However, from the system's perspective, the most desirable traffic flow pattern is known as System Optimum (SO) routing. Under this approach, certain drivers may opt for longer routes compared to the shortest path, but the overall travel time for all drivers is minimized. AVs are a rapidly developing technology that can open up new opportunities to mitigate congestion and improve network performance by centrally coordinating the SO routes for AVs.

Although some work [1] focuses on the 100% presence of self-driving cars in the network, current predictions [2] indicate that it may take until the 2040s or 2050s before they become widespread and affordable, and have a substantial share of the market. Moreover, complete automation may take even longer to accomplish. Thus, we should anticipate the coexistence of both HDVs and AVs on the roads for a considerable period of time. Therefore, the objective is to explore a mixed traffic flow system in which HDVs follow the dynamic user equilibrium (DUE) and are expected to minimize their own travel costs, whereas the AVs are managed to enhance the system performance of the entire network.

Several studies have explored this concept [3][4][5][6], but there is a challenging assumption. Would AV owners allow themselves to be controlled externally by a central agency for the good of the system? Giving the right incentives, such as fuel discounts, toll exemption or tax credits can encourage the AVs to choose SO routes. In addition, system-level traffic control measures may become increasingly feasible with the emergence of AVs, thereby reducing disparities between SO and UE traffic assignments. But are they willing to accept long routes without any preconditions? In other words, the question arises: to what extent are AV owners willing to compromise on their own preferences? Hence, in order to facilitate the adoption of prescriptive routes for selfdriving cars, it is crucial to establish a fair system where the occupants of AVs can be confident that their travel time will not exceed a certain threshold.

This paper aims to explore mixed dynamic user equilibrium - dynamic system optimum (DUE-DSO) traffic assignments that provide equitable route recommendations for selfdriving cars. We develop a heuristic framework that incorporates a queue-based simulator and a Swapping Algorithm, which seeks to impose fairness for SO seekers within the route choice model.

### II. LITERATURE REVIEW

# *A. Mixed traffic assignment models*

Harker [7] made a pioneering contribution to the literature by introducing the concept of multiple equilibrium behaviours on networks. Nevertheless, this field of study is still in its early stages of investigation, and most early works on this topic concentrated on static mixed UE-SO traffic assignment. Bagloee [4] assumed that connected vehicles (CVs) should behave as static SO (ST-SO) users, while other travellers are ST-UE routes and used a logit formulation to model elastic demand. Zhang [6] proposed a bi-level optimisation problem, with the upper level determining the best AV rates for each pair of ODs and the lower level was a multi-class STA problem.

Compared to traditional Static Traffic Assignment (STA) models, the Dynamic Traffic Assignment (DTA) model provides a more detailed means for capturing the interaction between traffic flow propagation and travel time in a temporally coherent manner. DTA has been thoroughly studied since the late 70s, but even more so in the last twenty years since Intelligent Transport Systems (ITS) have become more prevalent as a means of addressing the issues related to traffic congestion (readers should refer to [8] for an overview).

DTA models can be classified into two main categories: simulation-based models and analytical models. In analytical models, the DTA problem is formulated as a mathematical problem (MP), optimal control (OC), variational inequalities (VI) or complementarity model and solved directly using optimization techniques. In these models, optimality is the

<sup>\*</sup>This work was not supported by any organization

<sup>1</sup>Faculty of Information Technology, Monash University, Australia ali.davoodi@monash.edu

<sup>2</sup> Institute of Transport Studies, Monash University, Australia

first priority. On the other hand, simulation-based models are designed to be used in large-scale real-world applications. In these models, the traffic flow is propagated over the network by means of a traffic simulator [9].

Recently, Guo [5] focused on dynamic mixed traffic flow and employed a double queue model (DQM) as a dynamic network loading (DNL) model. They developed a bi-level model in which the routing for HDVs was modeled using the instantaneous dynamic user equilibrium (IDUE) approach, while the routing for AVs was modeled using the SO approach at the upper level. To solve the problem, they proposed a heuristic framework and conducted experiments on two small networks with varying penetration rates of AVs. In another recent study [3], a simulation-based dynamic traffic assignment (SBDTA) framework was introduced to address the challenge of mixed DUE-DSO equilibrium in a scalable network. The framework incorporated a dynamic joint routing and second-best pricing control at the upper level which aimed to encourage AVs to select SO routes by offering toll savings.

## *B. Fairness routing in traffic assignment*

Jahn [10] utilizes the benefits of a constrained shortest path algorithm in routing, which ensures fairness when compared to pure SO routing. In another related study, Angelelli [11] proposed a heuristic algorithm to compute equitable routes using a piecewise linear approximation of the travel time function. However, it is important to note that these works focused on the STA problem. For DTA models, the use of constrained time-dependent shortest path algorithms is computationally expensive. Additionally, the estimation of edge travel time becomes more complex when capturing more realistic features such as spill-back and queuing effects.

## III. RELATED WORKS

Among the related works, Mansourianfar's study [3] is perhaps the most similar to ours, but there are some notable differences. They used the AIMSUN mesoscopic simulator to develop a framework for mixed behaviour traffic flow. To solve the DSO problem, one way is to formulate it as a DUE using path marginal travel times. However, computing path marginal costs are challenging and costly [12][13][14]. Path marginal cost is the change in the total system travel time when one additional unit of flow at time  $t$  goes through path  $p$ . A local approximation method is to calculate link marginal cost (LMC) and then consider path marginal cost as the sum of all links through that path [15] [16]. To compute LMC, Mansurifar used the BPR function suggested by Mahmassani [16]. However, this method does not consider the effect of the queue on LMC. On the other hand, Ghali [15] provides a more accurate formulation for LMCs based on the link cumulative curves. It is shown that the link marginal cost is equal to the time difference between the time when the vehicle enters the link and the earliest time after that when the queue on the link vanishes. In our work, we use DTAlite [17], a queue-based macroscopic simulator, that allows us to calculate LMCs using Ghali's method. Moreover, DTAlite is well-matched and efficient for managing large-scale networks.

## IV. CONTRIBUTIONS

This paper contributes to the literature in two main aspects. Firstly, we have developed a framework to address the mixed dynamic user equilibrium-dynamic system optimum (DUE-DSO) traffic flow problem utilizing the DTAlite simulation platform and Ghali's method for computing LMC. This framework allows for an accurate analysis of the traffic flow dynamics. Secondly, we have introduced a Swapping Algorithm that enhances fairness in route selection compared to pure SO routing. By incorporating this algorithm, we achieve a better balance between individual travel costs and system-wide performance, ultimately leading to more equitable route recommendations.

## V. METHODOLOGY

In this section, we present the framework of our algorithm designed to address the mixed dynamic UE and dynamic fair system optimal (FSO) traffic assignment.

# *A. Notations and parameters*

Consider a directed transportation network denoted by  $G = (N, E)$ , where E represents the set of edges and N represents the set of nodes. The network includes a set of origin-destination (OD) pairs denoted by  $W$ . For each OD pair  $w \in W$  and time interval  $\tau \in T$ ,  $d_{UE}^{w,\tau}$  and  $d_{FSO}^{w,\tau}$  represent the total demands of UE seekers and FSO seekers, respectively. Furthermore,  $TT_e^{\tau}$  represents the timedependent travel time of edge  $e \in E$  at interval  $\tau \in T$ , and  $LMC<sub>e</sub><sup>\tau</sup>$  represents the link marginal cost associated with edge e at interval  $\tau$ . For each OD pair  $w \in W$  and interval  $\tau \in T$  with non-zero demand, there are two sets of positive paths  $P_{UE}^{w,\tau}$  and  $P_{FSO}^{w,\tau}$ , representing the paths for UE and FSO seekers, respectively. For any path p in either  $P_{UE}^{w,\tau}$  or  $P_{FSO}^{w,\tau}$ , let  $h_p$  denote the flow of traffic on that path. The path travel time denoted as  $PTT_p$ , is the sum of the travel times of the edges it traverses. Similarly, the path marginal cost, denoted as  $PMC_p$ , is the sum of the link marginal costs of the edges it passes through. This can be expressed as  $PTT_p = \sum_{e \in p} TT_e^{\tau}$  and  $PMC_p = \sum_{e \in p} LMC_e^{\tau}$ .

# *B. Simulator*

The DTAlite simulator engine [17] utilizes Newell's kinematic wave model [18], which is based on the macroscopic fundamental diagram (MFD) assumption. The simulator effectively captures the forward and backward wave propagation by monitoring cumulative flow counts, both receiving and sending, on the network links. Figure 1 illustrates the simulation procedure. The initial step involves the generation of agents and their corresponding routes based on predefined path sets. Subsequently, these agents are loaded into the network. The DTAlite simulator engine monitors the cumulative receiving and sending flow counts for each network edge. Additionally, for each agent, it records the arrival and departure times associated with each edge of the route. In the final step, the travel time is computed by averaging the time



Fig. 1: Simulation Procedure

experienced by agents in each time interval. Furthermore, the link marginal costs are derived from the cumulative flow counts, employing Ghali's definition [15].

#### *C. Swapping Algorithm*

In simulation-based DTA algorithms, the Method of Successive Averages (MSA) is commonly utilized (Readers can see [19] for other solution algorithms). This method is popular for its use of predetermined step size and its independence from derivative information, making it wellsuited for simulation-based algorithms. In the regular MSA algorithm (Algorithm 1, during the n-th iteration, a portion of the demand, specifically  $1/(n+1)$  (line 1), is shifted from the non-shortest paths to the shortest path  $shp^{w,\tau}$  (line 8). The auxiliary demands, denoted as  $y_p$ , are obtained through an All-or-Nothing assignment strategy, where all positive auxiliary are assigned to the shortest path (lines 3:7). This iterative process aims to improve the flow pattern by reassigning flows onto more optimal routes.

To solve the DSO problem, as mentioned earlier, it can be transformed into the DUE by computing the path marginal costs. However, in order to introduce a fairness condition to the path redistribution process (DFSO), we do not simply select the path with the lowest marginal cost  $l_{pmc}^{w,\tau}$  and transfer flow from other paths to it. Instead (Algorithm 2), we choose the path with the lowest marginal cost among those with reasonable travel times  $PTTp < (1 + \phi)PTTshp^{w,\tau}$ (lines 3:8). If no path satisfies this condition, the shortest path  $shp^{w,\tau}$  is added to the set of paths (lines 9:12). Subsequently, the path flows are updated (line 18)by calculating the auxiliary paths of the current flows (lines 13:17). In this redistribution algorithm, the aim is to prevent flows from being directed towards unfair paths, ensuring a more equitable distribution of traffic.

## *D. Mixed DUE-DFSO traffic flow framework*

Algorithm 3 outlines the Mixed DUE-DFSO algorithm. It begins with an All-or-Nothing assignment (lines 4:13), where all demands are initially assigned to the time-dependent shortest path. The simulator is then loaded with these paths to update travel times and link marginal costs (lines 14). In the main loop, the UE seekers search for new shortest paths (lines 19:21) and redistribute flows using the regular Method of Successive Averages (MSA) algorithm (Algorithm 1)



 $\sqrt{U}$ 



18: end if  $h_p^{n+1} = h_p^n + \alpha (y_p - h_p^n)$ 20: end for

(lines 22). Meanwhile, the FSO seekers search for least path marginal cost (lines 25:26), and redistributed based on Algorithm 2 (lines 27), specifically designed for the DFSO problem. This iterative process aims to achieve a balance between UE and equitable system optimum in the traffic assignment.

#### VI. NUMERICAL RESULTS

In this section, we examine the performance and outcomes of our Mixed DUE-DFSO algorithm on the Sioux Falls network, which is a commonly used benchmark in the literature. This network comprises 24 nodes, 76 edges, and 528 OD pairs, and its characteristics can be found at (http: //www.bgu.ac.il/˜bargera/tntp). The algorithms were implemented in the C++ programming language using Visual Studio code. The simulation consists of 60 assignment intervals, each with a duration of 1 minute, and cumulative flow counts are monitored every 1 second. Demand is loaded uniformly into the network during the first 15 intervals, and the remaining intervals have zero demand to empty the network for comparison of total system travel time (TSTT). The algorithm terminates when the error, *err*, falls below  $\varepsilon_{max} = 10^{-2}$  or reaches the maximum cycle of  $N_{max} =$ 200. Given that the execution time in various scenarios was less than 4 seconds, we have chosen not to include it in the report.

Figure 2 shows the gap between the SO and UE approaches for different variations of the demands, ranging from 10 to 30 thousand. As the vehicular load in the network increases, the gap between the SO and UE flow patterns initially grows and then decreases. In other words, when network congestion reaches excessive levels, the potential for improving the performance of network by transitioning from selfish routing to optimal system routing through a central system diminishes.

The simulation results for the Sioux Falls network at a demand level of 10,000 vehicles are presented in Figure 3. This figure demonstrates how the increase in the presence of AVs leads to a reduction in TSTT for vehicles. The red and blue lines represent pure user UE and SO routing, respectively. The other lines represent FSO routing with varying fairness coefficients  $(\phi)$ . It is evident that as the fairness coefficient increases, the routing approaches move closer to the SO routing, as the system allows for longer paths. Additionally, by employing a fairness coefficient of 20 %, it is possible to achieve solutions that are close to the optimal system. Figure 4 and Figure 5 present the same analyses for demands of 20,000 and 30,000.

Table I evaluates the performance of our swapping algorithm in the Sioux Falls network at a demand of 10,000. The first column represents different assignment behaviours, including SO and Fair System Optimal with various fairness coefficients. In all results, the AVs rate is 100 %. In Column 2, the worst-case path is identified among active paths with a flow of more than 1 vehicle. Similarly, Columns 3 and 4 report the worst-case paths for flow thresholds of more than 2 and 5 vehicles, respectively. Examining these results highlights that our swapping algorithm significantly reduces the worst-case scenarios. Furthermore, as the path flow thresholds increase, the improvement in the worstcase scenarios becomes more pronounced. This indicates that the majority of paths violating fairness conditions have negligible flows. Table II and Table III provide the same analyses for demands of 20,000 and 30,000, respectively.

In Figure 6, the red line represents the average travel time of AVs who seek SO routes at different penetration rates in the system in the Sioux Falls network at the demand of 10000, while the blue line represents the average travel time of HDVs who follow user optimum routes. The purple line shows the average travel time for both classes combined. However, In Figure 7, the same results are depicted, but with AVs using FSO routing and a fairness coefficient of 15 %. Likewise, Figures 8, 9, 10, and 11 present identical analyses for demand quantities of 20,000 and 30,000. These graphs indicate that there is not much difference in average travel time between Scenario DUE-DSO and Scenario DUE-DFSO. However, in the former scenario, AVs are required to choose longer routes and endure longer travel times, especially at

lower penetration rates. On the other hand, HDVs have shorter travel times compared with the DUE-DFSO scenario. Considering that the main objective is to reduce the TSTT and improve network performance rather than reducing the travel time for HDVs, it can be concluded that the proposed algorithm provides equitable routes that significantly reduce the TSTT. The fact that AVs exhibit less self-sacrifice in this scenario enhances its feasibility and practicality.





Fig. 2: Gap between DUE and DSO TSTT

Assignments	Worst Case $(\%)$		
	hp > 1	hp > 2	hp > 5
SO.	43.58	23.11	21.27
FSO $(\phi = 20\%)$	20.73	20	20
FSO ( $\phi = 15\%)$	15.24	15.24	15
FSO ( $\phi = 10\%)$	12.20	10	10
FSO $(\phi = 5\%)$	11.98	6.29	5

TABLE I: Performance of the fairness swapping algorithm, demand=10,000



Fig. 3: TSTT of proposed DFSO-DUE framework for varying fairness  $(φ)$  and different AV ratios in the Sioux-falls network, demand=10,000

TABLE II: Performance of the fairness swapping algorithm, demand=20,000

Assignments	Worst Case $(\%)$		
	hp > 1	hp > 2	hp > 5
SO.	50.67	50.62	45.50
FSO $(\phi = 20\%)$	28.94	28.94	20
FSO $(\phi = 15\%)$	20.25	16.28	15
FSO ( $\phi = 10\%)$	18.19	10.30	10
FSO $(\phi = 5\%)$	13.45	7.56	5.59



Fig. 4: TSTT of proposed DFSO-DUE framework for varying fairness  $(φ)$  and different AV ratios in the Sioux-falls network, demand=20,000.

TABLE III: Performance of the fairness swapping algorithm,demand=30,000

Assignments	Worst Case $(\%)$		
	hp > 1	hp > 2	hp > 5
SO.	52.69	52.69	50.04
FSO ( $\phi = 20\%)$ )	28.94	28.32	20
FSO ( $\phi = 15\%)$ )	20.25	18.46	15.35
FSO ( $\phi = 10\%)$ )	18.19	13.34	10.31
FSO $(\phi = 5\%)$	13.45	10.35	6.93



Fig. 5: TSTT of proposed DFSO-DUE framework for varying fairness  $(φ)$  and different AV ratios in the Sioux-falls network, demand=30,000.



Fig. 6: Avg. travel time for DSO, UE, and mixed flow, demand=10,000



Fig. 7: Avg. travel time for DFSO ( $\phi = 15$ ), UE, and mixed flow, demand=10,000



Fig. 8: Avg. travel time for DSO, UE, and mixed flow, demand=20,000



Fig. 9: Avg. travel time for DFSO ( $\phi = 15$ ), UE, and mixed flow, demand=20,000



Fig. 10: Avg. travel time for DSO, UE, and mixed flow, demand=30,000



Fig. 11: Avg. travel time for DFSO ( $\phi = 15$ ), UE, and mixed flow, demand=30,000

# VII. CONCLUSIONS

Autonomous Vehicle (AVs) are capable of coordinated navigation systems that assign paths. With the right incentives, AVs are willing to choose longer paths to alleviate traffic congestion. However, it is crucial to ensure that the assigned paths are fair. This paper proposes a novel dynamic routing strategy that seeks a sub-optimal solution, balancing the reduction of congestion while also providing equitable routes for AVs.

## **REFERENCES**

- [1] Melson CL, Levin MW, Hammit BE, Boyles SD. Dynamic traffic assignment of cooperative adaptive cruise control. Transportation Research Part C: Emerging Technologies. 2018 May 1;90:114-33.
- [2] Litman T. Autonomous vehicle implementation predictions. Victoria, BC, Canada: Victoria Transport Policy Institute; 2017 Feb 27.
- [3] Mansourianfar MH, Gu Z, Waller ST, Saberi M. Joint routing and pricing control in congested mixed autonomy networks. Transportation Research Part C: Emerging Technologies. 2021 Oct 1;131:103338.
- [4] Bagloee SA, Sarvi M, Patriksson M, Rajabifard A. A mixed userequilibrium and system-optimal traffic flow for connected vehicles stated as a complementarity problem. Computer-Aided Civil and Infrastructure Engineering. 2017 Jul;32(7):562-80.
- [5] Guo Q, Ban XJ, Aziz HA. Mixed traffic flow of human driven vehicles and automated vehicles on dynamic transportation networks. Transportation research part C: emerging technologies. 2021 Jul 1;128:103159.
- [6] Zhang K, Nie YM. Mitigating the impact of selfish routing: An optimal-ratio control scheme (ORCS) inspired by autonomous driving. Transportation Research Part C: Emerging Technologies. 2018 Feb 1;87:75-90.
- [7] Harker PT. Multiple equilibrium behaviors on networks. Transportation science. 1988 Feb;22(1):39-46.
- [8] Chiu YC, Bottom J, Mahut M, Paz A, Balakrishna R, Waller S, Hicks J. Dynamic traffic assignment: A primer (transportation research circular e-c153).
- [9] Peeta S, Ziliaskopoulos AK. Foundations of dynamic traffic assignment: The past, the present and the future. Networks and spatial economics. 2001 Sep;1:233-65.
- [10] Jahn O, Möhring RH, Schulz AS, Stier-Moses NE. System-optimal routing of traffic flows with user constraints in networks with congestion. Operations research. 2005 Aug;53(4):600-16.
- [11] Angelelli E, Morandi V, Speranza MG. Congestion avoiding heuristic path generation for the proactive route guidance. Computers and Operations Research. 2018 Nov 1;99:234-48.
- [12] Shen W, Nie Y, Zhang HM. Path-based system optimal dynamic traffic assignment models: formulations and solution methods. In2006 IEEE Intelligent Transportation Systems Conference 2006 Sep 17 (pp. 1298- 1303). IEEE.
- [13] Oian ZS, Shen W, Zhang HM. System-optimal dynamic traffic assignment with and without queue spillback: Its path-based formulation and solution via approximate path marginal cost. Transportation research part B: methodological. 2012 Aug 1;46(7):874-93
- [14] Doan K, Ukkusuri SV. Dynamic system optimal model for multi-OD traffic networks with an advanced spatial queuing model. Transportation Research Part C: Emerging Technologies. 2015 Feb 1;51:41-65.
- [15] Ghali MO, Smith MJ. A model for the dynamic system optimum traffic assignment problem. Transportation Research Part B: Methodological. 1995 Jun 1;29(3):155-70.
- [16] Mahmassani HS, Peeta S. Network performance under system optimal and user equilibrium dynamic assignments: implications for advanced traveler information systems. Transportation Research Record. 1993;1408:83.
- [17] Zhou X, Taylor J. DTALite: A queue-based mesoscopic traffic simulator for fast model evaluation and calibration. 2014.
- [18] Newell GF. A simplified theory of kinematic waves in highway traffic, part I: General theory. Transportation Research Part B: Methodological. 1993 Aug 1;27(4):281-7
- [19] Ameli M, Lebacque JP, Leclercq L. Cross-comparison of convergence algorithms to solve trip-based dynamic traffic assignment problems. Computer-Aided Civil and Infrastructure Engineering. 2020 Mar;35(3):219-40.