AI-powered transit simulator: integration between microscopic simulation and machine learning to enhance scalability and accuracy

M. Rezazada[a,*], N. Nassir[a], E. Tanin[a], A. Ceder[b]

[a] Faculty of Engineering and Information Technology, University of Melbourne, Melbourne, Australia

mustafa.rezazada@student.unimelb.edu.au, neema.nassir@unimelb.edu.au, etanin@unimelb.edu.au, ceder@technion.ac.il

[b] Faculty of Civil and Environmental Engineering, Technion-Israel Institute of Technology, Haifa, Israel

* Corresponding author

Keywords: Microscopic simulation, machine and deep learning, public transport operation, bus bunching, demand and supply, intersection management

1 INTRODUCTION

Modeling public transit (PT) operations is a complex task because of various factors such as fluctuating demand and supply variations, external disruptions, diverse behaviors, infrastructure design, and driver behavior. These factors are interconnected, creating a dynamically uncertain system locally with specific data elements and globally across the entire PT network. Extensive research that has been conducted in this field with three-fold categories: scalable and efficient methods such as macroscopic and mesoscopic simulations (Toledo et al., 2010), detailed but isolated analyses like microscopic simulations (Zhang et al., 2008, Johansson et al., 2015), and data-driven approaches such as machine learning and deep learning (ML-DL) (Yazdani et al., 2023).

The first category is effective for modeling the overall traffic and transit system, but often oversimplifies local dynamics by assuming deterministic patterns or theoretical distributions, and thus may not accurately represent real-world complexities (Fernández, 2010). Dynamic day-to-day variations of, for instance, passenger arrival rates, boarding and alighting data, agent interactions, adaptive signal delays and priority control, and driver behavior can significantly impact the PT operational reliability. At the same time detailed analyses, of the second category, with detailed considerations may not fully address how these PT variations impact the reliability of the PT system. That is, it is essential to comprehend how passenger-demand fluctuations at one stop can affect these fluctuations at other stops, the effect of changed demand on reliability, and the potential impacts of bus bunching in one line on other bus lines. Finally, in the third category, data-driven methods show a promise in learning historical patterns but may fail to connect different aspects of the PT network and analyze their interrelationships. For example, proactive prediction of passenger demand at each PT stop could potentially improve reliability measures through optimizing the supply elements.

This study proposes a novel approach that integrates between microscopic simulation and machine learning to provide an AI-powered microscopic holistic model. It is shown that this approach enhances computational efficiency and scalability across larger PT networks. In this work's methodology the machine learning and deep learning (ML-DL) is seamlessly integrated with simulation of real-time with a unified engine operation. It is being powered by AI models, trained on large-scale historical data from extended period of time, that govern certain simulation activities. This integration is not only rationalizing computation by eliminating the need for individual-agent simulation of some secondary components such as cars, signals, etc., but also expands the scalability of microscopic simulation of some primary elements as PT vehicle and passenger movements and their interactions.

2 METHODOLOGY APPROACH

The foundational concept of this work's framework lies in its innovative approach to modeling various components and activities in transit operations. We classify these components into primary and secondary activities. Primary components directly involve vehicle and passenger movements, including vehicle and passenger interactions, dwell and travel times, waiting and queuing behavior, PT stop processes, and route and network geometric design. Microscopic models, such as the Wiedemann-99 Car Following, the Social Force, and the Queueing models, are utilized to explicitly simulate these processes (Johansson et al., 2015, Durrani et al., 2016). The secondary components indirectly impact system reliability, encompassing factors such as demand fluctuations, weather conditions, departure discrepancies, traffic congestion, intersection delays, and public transport priority controls. These are implicitly modeled using ML-DL algorithms trained on extensive real-operation datasets.

The framework is comprised of three core mechanisms or engines: the data processing and fusion engine (DPF-E), the AI engine (AI-E), and the simulation engine (SM-E). The DPF-E automates data retrieval from various sources, processes, integrates, and reshapes data to fit the input required by each AI model. The AI engine hosts multiple specialized ML-DL models that predict and analyze various operational conditions before and during real-time simulation. These models predict passenger demand at each stop, boarding/alighting variability, departure times from terminal for different PT lines, and delays at actuated signalized intersections with priority control procedures. The models promptly and accurately make predictions by learning historical patterns including spatial and temporal variations, compared with traditional methods relying on theoretical distributions or fixed values.

The simulation engine combines all primary components, including physical network, routes, stops, and vehicle and passenger movements and interactions. It receives precise predictions from the AI-E at specified intervals during the simulation and links these processes together. The integration between an AI engine and the simulation enables this accurate representation of historical patterns and variabilities, facilitating the implementation of effective control strategies to mitigate reliability issues like PT vehicle bunching.

3 CASE STUDY

The transit simulator AMATS has been utilized in a case study to assess the operational reliability of Melbourne's Tram Route 96 in Australia. The objectives of the case study are threefold. First, to showcase how AMATS accurately replicates historical patterns and variabilities, such as day-to-day demand fluctuations, dwell time and headway variability in time and space, delay propagation, travel time variabilities, and departure discrepancies, as depicted in Figures 1 and 2. Second, to distinguish between recurrent and non-recurrent reliability issues like bunching under realistic conditions, which include uncertain demand, dispatch irregularities, and external disruptions like weather, signal, and traffic delays at intersections. By modeling the day-to-day variability of underlying factors using AMATS simulation, it was observed that certain bunching patterns commonly occur at specific locations and times on weekdays, as shown in Figure 3. Third, by leveraging the integrated AI-E, the sources of reliability problems were accurately identified, as depicted in Figure 4.

Tram Route 96, being one of the largest and busiest routes in Melbourne, faces challenges such as schedule deviations and bunching, further complicated by its role in connecting multiple attractions and regions through the CBD. The highly uncertain passenger demand makes it challenging for traditional simulators to effectively capture the impact of these fluctuations on transit operation and reliability. These simulators often rely on Poisson and Binomial distributions for boarding and alighting processes, respectively (Toledo, 2010 #90), which fail to adequately capture the sensitivity to external factors such as events and adverse weather conditions which may lead to unanticipated-

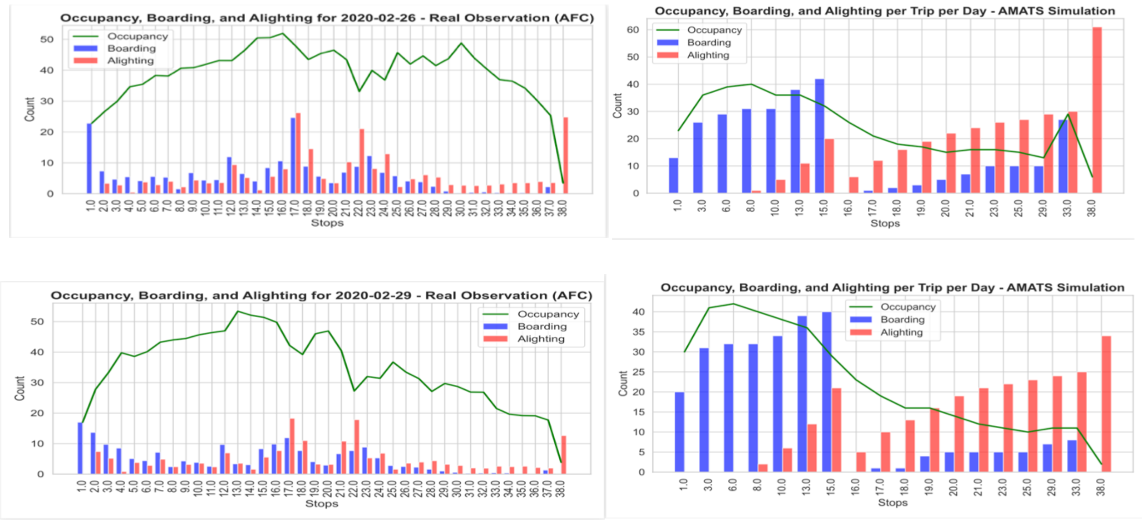

Figure 1 - *Demand fluctuation and load profile (Left: Real observation (AFC), right: AMATS Simulation)*

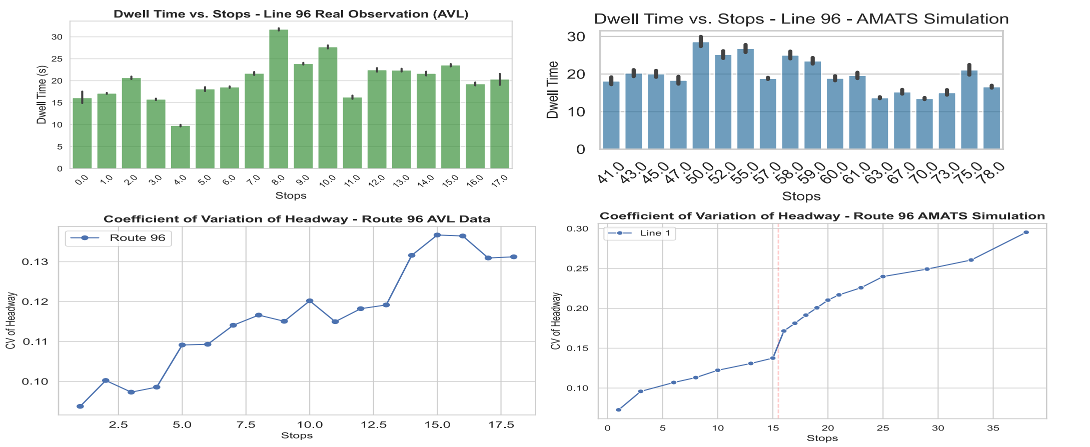

Figure 2 – *Dwell time distribution at different stops (Left: Real observation (AVL), right: AMATS simulation)*

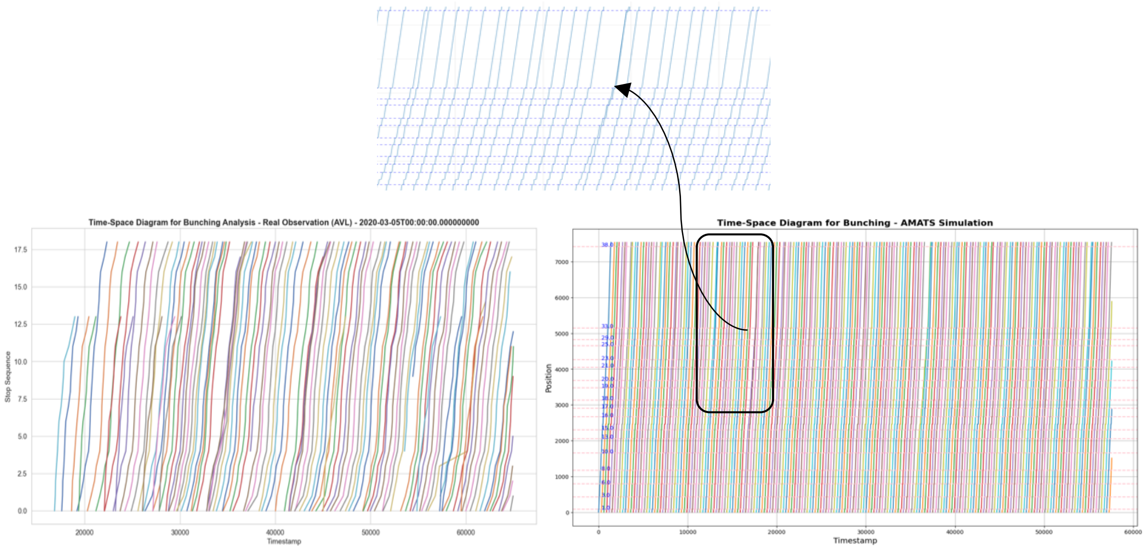

Figure 3 – Headway variability and bunching (Left: AVL data, right: AMATS simulation)

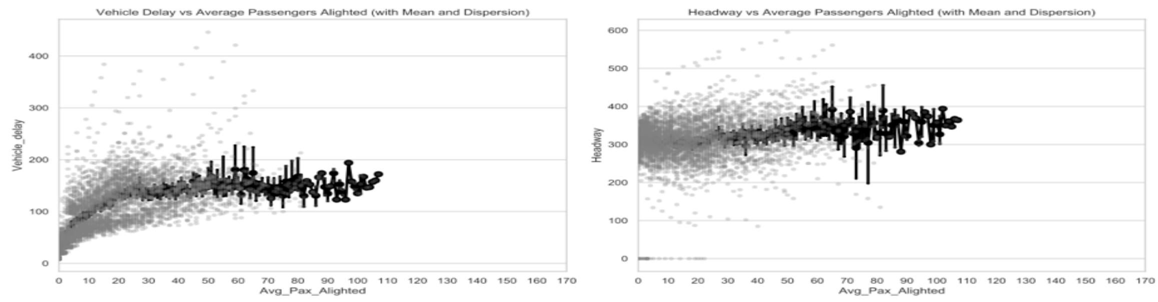

Figure 4 – Correlation of vehicle delay and headway variability with demand (AMATS simulation)

-changes in arrival rate, boarding or alighting at certain stops, affecting dwell time and the whole service performance. AMATS has the capability to capture such variabilities from input data and incorporate them into the microsimulation process where the cascading effects of random factors are modelled and their effect on service performance is estimated and evaluated.

4 DISCUSSION ON APPROACH TESTING

The results demonstrate that the AMATS simulator effectively represents both underlying factors and reliability measures. In Figure 1 the boarding and alighting characteristics and tram's passenger load day-to-day variability are accurately captured. While there is a slight overestimation of demand for some stops compared to AFC data because of the introduction of a weight factor to account for fare evasion, the overall distribution is accurate. The total estimated daily boarding numbers range between 35,000-60,000 with an average of 56,400 compared with the observed value of 56,722 and the Victorian Integrated Transport Model of 58,720; thus, it results with an error of less than 1%.

Measuring dwell time and headway variability is a challenging task. AMATS adeptly captures both their spatio-temporal variability and the mean values with over 90% accuracy as shown in Figure 2. Furthermore, capturing reliability issues, such as bunching under uncertain conditions with fluctuating demand, supply variation, and externalities presents a challenge. By running AMATS over several days, we were able to identify and differentiate between systematic recurrent and non-systematic bunching events. For instance, morning and afternoon peak bunching occurrences were frequently observed on different days, while other bunching incidents varied from day to day and stop to stop, as is illustrated in Figure 3 for a typical weekday. Finally, the results revealed that an increase in demand not only leads to higher vehicle delays at stops but also makes these delays less predictable. This aligns with findings from the literature, where an increase of the number of alighting passengers prolongs dwell time per passenger and thus results with higher and less predictable dwell times (Zhang et al., 2008). AMATS provides a wide range of applications at different scales and resolutions, tailored to specific project requirements.

5 REFERENCES

Durrani, U., Lee, C. & Maoh, H. 2016. Calibrating the Wiedemann's vehicle-following model using mixed vehicle-pair interactions. *Transportation Research Part C: Emerging Technologies,* 67, 227-242.

Fernández, R. 2010. Modelling public transport stops by microscopic simulation. *Transportation Research Part C: Emerging Technologies,* 18, 856-868.

Johansson, F., Peterson, A. & Tapani, A. 2015. Waiting pedestrians in the social force model. *Physica A: Statistical Mechanics and its Applications,* 419, 95-107.

Toledo, T., Cats, O., Burghout, W. & Koutsopoulos, H. N. 2010. Mesoscopic simulation for transit operations. *Transportation Research Part C: Emerging Technologies,* 18, 896-908.

Yazdani, M., Sarvi, M., Asadi Bagloee, S., Nassir, N., Price, J. & Parineh, H. 2023. Intelligent vehicle pedestrian light (IVPL): A deep reinforcement learning approach for traffic signal control. *Transportation Research Part C: Emerging Technologies,* 149, 103991.

Zhang, Q., Han, B. & Li, D. 2008. Modeling and simulation of passenger alighting and boarding movement in Beijing metro stations. *Transportation Research Part C: Emerging Technologies,* 16, 635-649.