IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 5 February 2024, accepted 12 February 2024, date of publication 16 February 2024, date of current version 23 February 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3366517

RESEARCH ARTICLE

Combining Reinforcement Learning With Genetic Algorithm for Many-To-Many Route Optimization of Autonomous Vehicles

SUNHYUNG YOO[1], HYUN KIM[2], AND JINWOO LEE[1]
[1]School of Built Environment, The University of New South Wales, Sydney, NSW 2052, Australia
[2]Department of Computer Science and Information Engineering, Korea National University of Transportation, Chungju 27469, South Korea

Corresponding author: Jinwoo Lee (brian.j.lee@unsw.edu.au)

This work was supported by the Korea Agency for Infrastructure Technology Advancement (KAIA) Grant funded by the Ministry of Land, Infrastructure and Transport under Grant RS-2021-KA161756.

ABSTRACT This study introduces an approach for route optimization of many-to-many Demand-Responsive Transport (DRT) services. In contrast to conventional fixed-route transit systems, DRT provides dynamic, flexible, and cost-effective alternatives. We present an algorithm that integrates DRT with the autonomous shuttles at Korea National University of Transportation (KNUT), allowing dynamic route modifications in real-time to accommodate incoming service calls. The algorithm is designed to take into account the shuttle's current position, the destinations of passengers already on board, the current locations and destinations of individuals who have requested shuttle services, and the remaining capacity of the shuttle. The algorithm has been developed to combine genetic algorithms and reinforcement learning. The performance evaluation was conducted using a simulation model that emulates KNUT's campus and the adjoining local community area. The simulation results show significant improvements in two key metrics, specifically the 'Request to Pick-up Time' and 'Request to Drop-off Time' during high-demand periods over the single-shuttle operation. Additional simulation test with random service requests and stochastic passenger walking distances showed the potential adaptability across different settings.

INDEX TERMS Autonomous vehicles, intelligent transportation systems, machine learning, machine learning algorithms, public transportation, transportation.

I. INTRODUCTION

In light of the recent advancements in autonomous vehicle technology, there is growing interest in exploring alternatives to traditional driver-accompanied fixed-route public transport [1]. Autonomous shuttle bus services are a particularly appealing option due to their relatively low ongoing costs, as there is no need to hire a driver. Deploying autonomous shuttle services in low-demand corridors or areas can significantly enhance the transport viability for their residents, effectively bridging gaps in the first and last mile of transit trips [2].

Numerous pilot experiments in several countries, including Australia, Austria, Estonia, Finland, France, Germany, Norway, Switzerland, Sweden, The Netherlands, and US, have demonstrated substantial advancements in technology and acceptance by citizens [3], [4]. It should be noted that the overall publicly available documentation of the autonomous shuttle bus pilots is not sufficient to provide a complete framework, but it looks like most pilots were conducted on fixed routes [5].

Traditional fixed-route systems, due to their static nature, face various inefficiencies in operation. They rely on

The associate editor coordinating the review of this manuscript and approving it for publication was Frederico Guimarães.

predetermined schedules and routes, leading to an underutilization of capacity during off-peak hours and overcrowding during peak times [6]. Additionally, accommodating unexpected demand surges, events or fluctuations in passenger distribution can pose a significant challenge [7]. These limitations result in passenger inconvenience, longer travel times, and high operational costs [8]. Due to the diminishing cost-effectiveness, the quality of local public transport services inevitably declines, leading to a detrimental impact on service users and a subsequent decrease in ridership.

Addressing the shortcomings of traditional fixed-route systems, many-to-many Demand-Responsive Transport (DRT) services serve as a dynamic and flexible alternative. These services run on flexible routes in response to real-time demand, enhancing passenger convenience by reducing wait times and promptly adjusting to unexpected changes in demand. Their capacity to adapt to demand fluctuations allows them to avoid the inefficiencies often associated with fixed-route systems, such as capacity underutilization and overcrowding.

In this research, we propose a new methodology specifically designed to optimize routes within a many-to-many Demand-Responsive Transport (DRT) framework. This approach is tailored for autonomous shuttle buses and caters to passengers with diverse origins, destinations, and preferences. Coordinating the operations of many-to-many DRT services, which encompass autonomous shuttle buses and cater to passengers with varying trip origin and destination locations, and preferences continues to present a challenge [9]. The immediate objective of this study is to embed and assess the proposed algorithm within the specific KNUT environment, laying down a foundational understanding of the algorithm's potential.

Three simulation scenarios are prepared to validate our proposed methodology. Scenario 1 tests a single autonomous shuttle bus's capacity to handle all service requests within 20 minutes, pinpointing peak demand periods. Scenario 2 builds on this, deploying additional vehicles during these peak times to evaluate service efficiency improvements. Scenario 3 introduces unpredictability, with random service requests and varied passenger walking distances, testing the algorithm's adaptability across fluctuating demand and diverse needs. Through these scenarios, this study seeks to provide a solution that enhances the efficiency, adaptability, and overall effectiveness of DRT operations.

II. RELATED WORKS

Demand-Responsive Transport (DRT) initially emerged in the 1970s as an innovative solution to extend the coverage of conventional fixed-route networks, particularly for groups like the elderly and individuals with special needs and in areas with lower demand densities [10], [11]. Enabled by interactive and intelligent information platforms such as smartphone applications or websites, DRT services are expected to grow in popularity [12]. The feasibility and effectiveness of DRT systems have been tested by pilot programs conducted in many countries such as Australia [13], UK [14], and US [15].

As a form of paratransit, one-to-one DRT services are often used as a mobility solution for specific demographics, such as older adults, people with disabilities, or individuals in areas underserved by traditional public transport [16]. While this mode of service contributes to social inclusion, it often involves high costs due to individualized service routes, long waiting times, and high vehicle-kilometer costs [17]. In the United States, the total operating expenses of paratransit service exceeded 1.2 billion dollars with only 173 million dollars collected in fares, as reported by the American Public Transit Association [18].

One-to-many DRT services, acting like a feeder bus service, function as connecting services between low-density areas and traditional fixed-route transit services [19]. This service model requires carefully coordinated schedules and efficient vehicle dispatching to ensure an acceptable level of service quality [20]. Efficient vehicle routing is a key problem, as these services are dynamic and have to adjust in real-time to passenger demand and traffic conditions. The dynamic nature of the vehicle routing problem in both one-to-many DRT services and the logistics sector are similar. Koç et al. provided a comprehensive review of existing studies on the simultaneous pick-up and delivery routing problem [21]. However, most existing algorithms for this problem only consider the transport of goods from a specific depot to clients and vice versa.

Many-to-many DRT services can offer flexible transit services, accommodating multiple pickup and drop-off points. One of the most significant challenges in implementing efficient many-to-many DRT services is determining an optimal route for vehicles. The task of managing variable pick-up and drop-off locations, times, and customer preferences creates a complex web of routing possibilities that makes efficient and timely service provision a challenging task. The issue of routing becomes even more complex when considering the need to minimize the cost and maximize the utilization of vehicles while providing a satisfactory level of service for customers. This difficulty is particularly prevalent in many-to-many DRT services where the routes are not fixed and can change dynamically based on demand. Moreover, factors such as traffic congestion, service delays, and varying travel demands further complicate the routing process. While advanced information technology platforms and intelligent dispatching and matching systems have been introduced to enhance operational efficiency, these tools still face limitations and challenges. Therefore, in the context of many-to-many DRT services, research into more robust and efficient routing algorithms, that can quickly adapt to dynamic changes in demand and other variables, remains a critical area of study.

III. METHODOLOGY

Coordinating autonomous shuttle buses to serve passengers with diverse starting points and destinations introduces a multi-dimensional problem that consists of several critical aspects. The foremost consideration is the dynamic nature of demand. Passengers may request a shuttle bus service at any time, from any location, with each requiring drop-off at a different destination. This ever-changing demand presents a considerable challenge for autonomous shuttle buses. The operating system must consistently adapt to real-time requests while maintaining efficient route management.

Routing is also a significant challenging task. Given the range of starting points and destinations, determining optimal routes to minimize travel times and distances is essential. Unlike fixed-route buses that continuously follow the same circuit, an autonomous shuttle bus must possess the capability to dynamically create and adjust its route in real time. This adaptability needs to account for new passenger pick-ups and drop-offs as they occur. Capacity management is another crucial factor. As autonomous shuttle buses have limited capacity, it's critical for the service to ensure that this capacity is not exceeded while attempting to meet as many requests as possible. This requirement implies that the algorithm must consider not only the shortest route but also the appropriate number of passengers to be loaded at each point along the route.

The factor of uncertainty can introduce significant complexity. For instance, there's a chance that passengers who have requested service may not actually be at their stated pick-up points when the shuttle arrives. This circumstance raises the question of whether the shuttle should wait for the passenger, potentially inconveniencing other passengers, or proceed to maintain service efficiency.

To provide an effective and efficient autonomous shuttle bus service, this study formulates two key algorithms, as illustrated in Fig. 1. The process of service provision initiates when a user submits a service request. Upon receipt of this request, the system activates the 'On-Demand Boarding Guidance Algorithm'. This algorithm undertakes the responsibility of locating the boarding stops for passengers. Factors considered include the location of the service request, its proximity to the ODD, and the operational efficiency. Once the boarding location has been determined, the system then generates an Estimated Time of Arrival (ETA) for the pick-up.

Assuring access to high-quality shuttle position data is paramount, as it directly influences the successful deployment and optimal functioning of the algorithm. The proposed algorithm assumes access to high-quality shuttle position data, a crucial element for its successful implementation and optimal performance.

Following the initial step, the request and ETA data are processed by an additional algorithm of the 'Route Optimization Algorithm'. This algorithm is tasked with determining the most efficient bus route, and as a part of this process, it calculates the Estimated Times of Arrival (ETA) at the drop-off points for all active service requests. In the case of multiple buses are in operation, the most appropriate bus from the fleet to fulfill the request is examined as well.

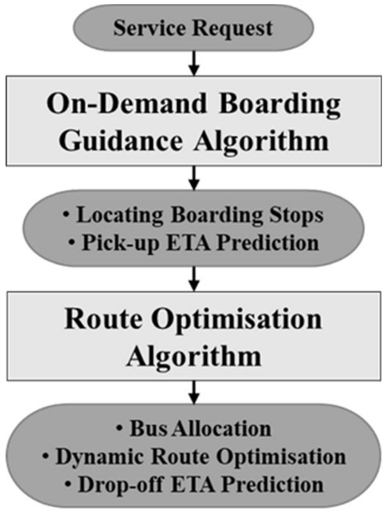


FIGURE 1. Demand-responsive shuttle bus service process.

A. ON-DEMAND BOARDING GUIDANCE ALGORITHM

The patron's interest in bus services significantly diminishes when boarding stops are located more than a five-minute walking distance [22], [23]. Moreover, as walking distance increases, the accurate estimation of arrival times at pick-up points becomes increasingly challenging, thereby complicating the provision of optimal dynamic routing in real-time. Given the average adult walking speed of five km/hr., a distance of 400 meters is calculated for a five-minute walk. Hence, in the cases of the walking distance exceeding 400m, an alert is issued declaring the location non-serviceable.

To avoid frequent stops of the autonomous shuttle bus, when a stop location within the same link already exists, new service requests are directed to the previously established stop. This strategy facilitates efficient and prompt service. Likewise, passengers intending to travel to the same or nearby destinations within a link can be directed towards a singular drop-off point. The process of the on-demand boarding guidance algorithm is presented in Fig. 2.

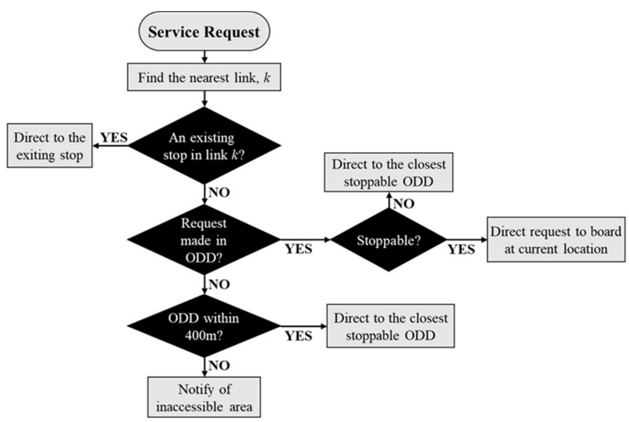


FIGURE 2. On-Demand boarding guidance algorithm.

In the case of unexpected events, such as inclement weather, road works, or accidents, the safety protocols mandate an immediate halt to services, prioritizing the well-being and security of both passengers and other road users. Similarly, when facing sudden demand spikes that the vehicle capacity cannot accommodate, our system steadfastly upholds safety and service reliability. It promptly issues notifications to passengers with excess service requests, clearly communicating that we are unable to accommodate their service call currently.

B. ROUTE OPTIMIZATION ALGORITHMS

This study aims to address a complex problem where it responds to service requests that originate from various pick-up points (OL_1, OL_2, OL_3, …, OL_k) and are destined for various drop-off points (DL_1, DL_2, DL_3, …, DL_k). The problem also requires considering the current location of the vehicle (L_{veh}) to determine the most optimal route option. For example, when there is one service request yet to be picked up and a new service request comes in, there are a total of six possible routes. In this context, the vehicle's location always takes precedence, and for the same service request, boarding always precedes alighting. The detailed routes for two simultaneous service requests are as follows:

$$\begin{aligned}
1_{\text{st}} &: (L_{veh} \rightarrow OL_1 \rightarrow DL_1 \rightarrow OL_2 \rightarrow DL_2),\\
2_{\text{nd}} &: (L_{veh} \rightarrow OL_1 \rightarrow OL_2 \rightarrow DL_1 \rightarrow DL_2),\\
3_{\text{rd}} &: (L_{veh} \rightarrow OL_1 \rightarrow OL_2 \rightarrow DL_2 \rightarrow DL_1),\\
4_{\text{th}} &: (L_{veh} \rightarrow OL_2 \rightarrow OL_1 \rightarrow DL_1 \rightarrow DL_2),\\
5_{\text{th}} &: (L_{veh} \rightarrow OL_2 \rightarrow OL_1 \rightarrow DL_2 \rightarrow DL_1),\\
6_{\text{th}} &: (L_{veh} \rightarrow OL_2 \rightarrow DL_2 \rightarrow OL_1 \rightarrow DL_1)
\end{aligned}$$

As the number of simultaneous service requests increases, the number of possible route options likewise surges exponentially. Depending on the number of simultaneous service requests, the available route options are as follows: 1 route option for 1 simultaneous service request, 6 route options for 2 simultaneous service requests, 90 route options for 3 simultaneous service requests, 520 route options for 4 simultaneous service requests, 113,400 route options for 5 simultaneous service requests, and a 7,484,400 route options for 6 simultaneous service requests.

Utilizing a simple Exact method to identify the shortest path between links and derive the optimal route invariably results in long computation time and necessitates large storage space for learning. Previous studies on the vehicle routing problem used the Reinforcement Learning (RL)-based algorithms, which provide prompt service based on pre-experienced dynamic routes [24], [25], [26], [27]. For instance, Lu et al. demonstrated an innovative lane-level traffic control approach in connected vehicle environments, optimizing traffic flow and safety using an RL algorithm [28]. RL algorithms have also been applied in other fields, including chemical process design [29] and network traffic management [30], showcasing their broad applicability beyond the transport sector. Nonetheless, the accumulation of learning data, engendered from a variety of scenarios (states), mandates substantial storage capacity. Further complicating matters, the time associated with loading this copious learning data induces constraints in the real-time guidance of dynamic routes.

To evaluate the effectiveness and practicality of the RL-only model, we analyzed its performance over a range of training episodes. The environment of the KNUT campus (Section IV-A) was utilized, featuring a single shuttle bus agent. The models were trained across three incremental demand scenarios: up to 3 simultaneous random service requests, 4 to 6 simultaneous random service requests, and 7 to 10 simultaneous random service requests. Performance metrics were gathered every 1,500 episodes. Notably, the model was programmed to randomly select actions when encountering inexperienced states during the exploitation phase. The results are presented in Fig 3.

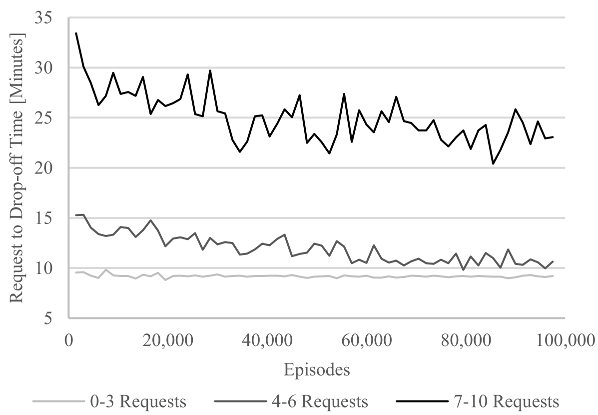


FIGURE 3. On-Demand boarding guidance algorithm.

In the low-demand scenario (0-3 requests), the RL-only model demonstrated a proficient learning curve, converging after approximately 25,000 episodes. This indicated a strong capability for the model to rapidly adapt and make effective decisions under conditions of low demand. However, when the number of simultaneous service requests increased to 4-6 request scenario, the model did not converge even after 100,000 episodes. The issue was even more severe in the higher demand scenario (7-10 requests). The performance was not stable, characterized by significant fluctuations throughout the training episodes. The inconsistent outcomes indicate that the RL-only model might face challenges in efficiently handling high service demands. The RL approach requires extensive learning periods to adapt to a wide range of scenarios and deduce the most efficient route plan, especially in high-demand environments.

To address these challenges, we integrated the Genetic Algorithm (GA) into our approach [31]. GA, a metaheuristic inspired by the principles of natural evolution, is adept at rapidly generating a range of potential solutions [32]. It operates by creating a diverse 'population' of solutions and then iteratively applies evolutionary processes, such as selection, crossover, and mutation. This approach enables GA to explore a wide array of solutions effectively, making it particularly efficient in identifying routes that are reasonably optimal in a short time frame, especially in high-demand situations.

While RL excels in incrementally refining solutions through iterative interaction with the environment, it may not be the most time-efficient method for immediate decision-making in dynamic and high-demand contexts. GA's ability to quickly assess and propose multiple potential solutions becomes essential. In our combined RL and GA approach, GA serves as a rapid and effective initial problem-solver in scenarios not adequately explored by the RL agent. This integration forms a powerful synergy, enhancing route optimization in high-demand scenarios. The hybrid approach combining Reinforcement Learning and the Genetic Algorithm is presented in Table 1, explaining the general flow of the proposed method.

In the RL algorithm, the Q-value formula is used to evaluate the fitness, and it is defined as follows:

$$Q(L_n, L_{n+1}) = Q(L_n, L_{n+1}) + \alpha[r_n + \gamma Q(L_{n+1}, a) - Q(L_t, L_{t+1})] \quad (1)$$

where,

$Q(L_n, L_{n+1})$	the desirability of choosing the route between link *n+1* to follow link *n*;
r_n	the current reward;
a	all the available links between link *n+1* to follow link *n*;
α	learning rate;
γ	discounting factor.

To elaborate, the algorithm doesn't just prioritize the most convenient or nearest service request. Instead, by setting the discount factor (γ) to 0.95, the algorithm also weighs in the potential rewards of future actions. Additionally, the learning rate, α, determines how much new information will overwrite old information. If α is set to 0, the Q-values are never updated, thus nothing is learned. If α is set to 1, the new information completely overwrites the old information. The decay function of the learning rate will be applied in this study to allow the model to explore widely at the early stages of learning. As the learning progresses and the Q-values become more reliable, the learning rate decays, thus the agent relies more heavily on the learned Q-values for decision making. This ensures a balance between exploration and exploitation.

$$\alpha_k = e^{-k \cdot \lambda_\alpha} \quad (2)$$

where,

α_k	learning rate at episode k;
k	The number of current episode k;
λ_α	the decay rate for learning rate α.

TABLE 1. Pseudo-code of Route Optimization Algorithm.

Algorithm: Route Optimization Algorithm
(Hybrid Approach - Reinforcement Learning and Genetic Algorithms)

```
////// Q-learning phase //////
Initialize Q-table with all zeros
N_epi = [Total number of episodes]
for episode from 1 to N_epi do
  current_state = initialize shuttle_location
  service_requests = [call_1, call_2, ..., call_n]
  while not all service_requests served do
    if random (0, 1) < exploration_prob then
      action = random_action(available_actions(current_state))
    else
      action = argmax(Q-table[current_state])
    end if
    next_state, reward = take_action(action)
    Update Q-table using Q-value formula
    current_state = next_state
  end while
end for

////// Genetic Algorithm (GA) Phase //////
N_routes = [Number of routes in GA]
N_gen = [Number of GA generations]
Initialize a list of Routes for GA_population
for i from 1 to N_routes do
  route = generate_route_using_Q-table()
  add route to GA_population
end for
for generation from 1 to N_gen do
  Sort GA_population by fitness
  parents = top 50% of GA_population
  offspring = []
  for i = 1 to (N_routes / 2) do
    parent_1 = random_choice(parents)
    parent_2 = random_choice(parents)
    child = crossover(parent_1, parent_2)
    mutate(child)
    add child to offspring
  end for
  GA_population = parents + offspring
end for

////// Proposing the best route //////
best_route = route with maximum fitness in GA_population
return best_route
```

The reward value (r_n) can be calculated as the sum of the immediate reward and the global reward, according to the following mathematical formula.

$$r_n = \beta\left(\frac{1}{N \cdot d(L_n, L_{n+1})}\right) + (1 - \beta)R_{global} \quad (3)$$

where,

β	a parameter controlling the balance between the local and global actions;
N	the number of existing service requests to be served;
$d(L_n, L_{n+1})$	the cost of travelling from L_n to L_{n+1};
R_{global}	the global reward (= average fitness of the last M solutions).

The parameter β is used to balance between immediate, local actions and global actions. Immediate actions refer to actions that yield a reward directly related to the current state, such as the cost of travelling from L_n to L_{n+1}, while global actions are more concerned with the overall states, represented by R_{global}. The parameter β with a decay function controls the balance between these two types of rewards. Initially, at the start of the learning process, β is set to a small value to give more weight to immediate rewards, as the agent doesn't have much information about the global environment. This encourages the agent to explore the environment and learn about the immediate effects of its actions. As the agent gains more experience and learns more about the environment, β gradually increases through a decay function, shifting the balance towards global rewards.

$$\beta_k = 1 - e^{-k \bullet \lambda_\beta} \tag{4}$$

where,

β_k a parameter of balancing between the local and global actions at episode k;
k The number of current episode k;
λ_β the decay rate for parameter β.

As previously mentioned, the vast array of potential route options in Many-to-many DRT, particularly during peak demand periods, may challenge the feasibility of using Q-values derived from reinforcement learning. To address this issue, a select number of routes, informed by learned Q-values, will serve as the initial population for a Genetic Algorithm. This algorithm then follows standard GA procedures, where new generations are produced through crossover and mutation processes [32].

In each generation, offspring chromosomes are created by selectively combining elements from parent chromosomes. This process generates candidate vehicle route options, which are then compared with the existing parent chromosomes. Through iterative refinement, the algorithm proposes an optimal route. The quality of the proposed route is evaluated using a fitness function, ensuring its effectiveness.

$$F(Route) = \sum_{k=1}^{N} (t_k^{walk} + t_k^{wait} + t_k^{board} + t_k^{in_{veh}} + t_k^{alight}) \tag{5}$$

where,

$F(Route)$ the fitness function of a bus route in GA
t_k^{walk} the walking time of service request k from origin to boarding location;
t_k^{wait} the waiting time of service request k at boarding stop;
t_k^{board} the time taken to board for service request k;
$t_k^{in\_veh}$ the in-vehicle time from boarding stop to alighting stop for service request k;
t_k^{alight} the time taken to alight for service request k.

Currently, the algorithm prioritizes minimizing costs from the user's perspective in route selection, as the costs related to service provision (like vehicle operating and administrative costs) are yet to be determined. The impact of a passenger's service request on the system's overall efficiency is quantified using a specific mathematical formula in the fitness function.

In cases where multiple autonomous shuttle buses are concurrently operational, and a new service request is initiated, the determination of the most suitable vehicle to the request becomes paramount. When a new request is received, the algorithm undertakes a comprehensive evaluation to determine the most suitable vehicle to fulfill this request. After computing the fitness function for each vehicle, considering the inclusion of the new service request, the algorithm compares these sums across all vehicles. The selection of the vehicle to handle the new request is based on which one has the lowest fitness function sum, as this indicates a route that efficiently minimizes disruptions to service times. Choosing the vehicle with the most efficient route ensures that the new request is accommodated in the most optimal way, keeping the disruption to other passengers at a minimum and upholding the overall efficiency of the service.

IV. SIMULATION

This study selected a specific study area. By focusing our research within this defined service zone, we test and refine our algorithms using data reflective of a realistic operational environment.

A. STUDY AREA

For this study, we have selected the Chungju campus of the Korea National University of Transportation (KNUT) and the adjacent local community area as the testbed. This study area was chosen deliberately, as it provides a controlled environment, facilitating the initial stages of testing and validation for our proposed algorithm. By initiating our experimentations in this specific context, we aim to address and pre-emptively identify any potential scalability issues, establishing a solid foundation for future adaptations of our algorithm to larger and more complex network settings.

According to the KNUT's statistical yearbook [33], 6,372 students are currently enrolled, and 917 staff members on the campus. The total area under study is 777,363 m^2, which encompasses the 629,083 m^2 campus and the local community area of 148,280 m^2.

Given that not all areas are drivable for an autonomous shuttle bus, this study specifically designates those paths that are suitable for autonomous shuttle operation. These paths, known as the Operational Design Domain (ODD), are particularly designed to function for autonomous shuttle buses. Fig. 4 represents the study area where service requests can be placed, along with the ODD defined for the autonomous shuttle buses. The ODD includes parallel running roads allowing bidirectional traffic. This layout enables the shuttles to cater to passenger pick-ups and drop-offs from either direction.

FIGURE 4. Study area.

There is a total of 6 left/right turning intersections, 3 roundabouts, and 2 U-turn points. Based on the locations of the turning points, the ODD was partitioned into a total of 32 distinct links. Subdividing the ODD allows for the efficient management of multiple service requests. Excluding a 50m radius around the turning points, autonomous shuttle buses can stop and allow passengers to board or alight. The composition of the link lengths used in the simulation is listed in Table 2.

TABLE 2. Link Lengths within the Testbed.

Link #	Length (m)	Link #	Length (m)	Link #	Length (m)
1, 2	290	13, 14	310	23, 24	140
3, 4	150	15, 16	150	25, 26	220
5, 6	260	17, 18	390	27, 28	220
7, 8	300	19, 20	250	29, 30	100
9, 10	330	21, 22	170	31, 32	60
11, 12	420				

In the simulation, the average driving speed of the autonomous vehicles is set at 25 km/h. Additional time delays have also been factored into the model to ensure a realistic representation of a typical journey. Specifically, a delay of 10 seconds is incorporated at turning points such as U-turns or left/right turns. Moreover, the model accounts for the time taken for passenger boarding and alighting. A static delay of 30 seconds is presumed every time the vehicle comes to a stop for passengers boarding or alighting. In addition to this, a time allowance of 5 seconds per passenger is made for the processes of boarding or alighting. The autonomous shuttle bus in the simulation is modelled with a maximum capacity of 15 passengers. While these parameters and assumptions enable a structured and thorough simulation, they can bring limitations that could impact the generalizability of our results. The settings used in our model, such as the vehicle speed and time delays, are based on general standards, but they may not encapsulate all real-world variations. Hence, future research and applications of our findings should consider these limitations and take into account the potential variability in real-world scenarios.

B. SERVICE REQUEST (DEMAND) DATA

The spatial-temporal data used in this simulation is based on the usage patterns of shared electric bicycles currently available around the KNUT. The initial rental location of the bicycle becomes the service request location, with the closest link in the ODD designated as the pick-up point. Similarly, the bicycle's return point is presumed to be the final destination, and the nearest link within the ODD becomes the drop-off point.

The data period for shared bicycle rental usage covers a full year, from 14 September 2021 to 13 September 2022. Bicycle usage instances where both the rental and return locations fall within the same ODD link are excluded, as these would entail passengers getting on and off at the same links. Additionally, the highly variable weekend rental data was omitted. This leaves a total of 120,255 shared bicycle rental history expected for the autonomous shuttle bus service.

After identifying the appropriate links for each service request point, we calculate the walking distance to the pick-up point to determine whether service can be provided within a 5-minute (400m) on foot. The distribution of service requests is detailed in Table 3. All service requests could be accommodated within a 400m walking distance of a pick-up point, except two in the link 27/28.

In the simulation, the walking time was estimated by taking into consideration the average walking speed of older adults, which is typically around 1.3 m/s [34]. This provides sufficient time for passengers of all ages to comfortably reach their designated boarding locations. To maintain the efficiency of our operation, any passengers who do not arrive at the boarding point within the allocated time are categorized as no-show service requests. This mechanism ensures that our service continues to run smoothly, accommodating the majority while addressing the challenges of timely passenger coordination.

In the training phase, episodes commence with the shuttle bus agent randomly placed, and service requests are generated to reflect real-world scenarios. These requests, including their origins and destinations, are based on historical data from shared bicycle rentals. This approach ensures a realistic representation of service demand, reflecting actual spatial and temporal patterns rather than a uniform campus-wide distribution.

TABLE 3. Spatial Information of Service Requests.

	Walking distance [m]						
Link#	0-100	100–200	200-300	300-400	400-500	500-600	Total
1/2	8,025	20,281	2,666	1			30,973
3/4	175	2,002					2,177
5/6	2,415	11	1				2,427
7/8	4,893	2					4,895
9/10	3,254	4,472					7,726
11/12	2	5	4				11
13/14	7,074	6,477	8				13,559
15/16	11	26					37
17/18	495	5,381	6	1			5,883
19/20	3,088	3,160					6,248
21/22	435	585					1,020
23/24	2,902	2,947					5,849
25/26	23	6,320	3				6,346
27/28	8,253	18,045	44			2	26,344
29/30	915	4,831					5,746
31/32	987	27					1,014
Total	42,947	74,572	2,732	2		2	120,255

Each training episode spans a 30-minute interval, shifting the start time by 5 minutes for each subsequent episode. For example, following an 8:00 to 8:30 interval in one episode, the next uses data from 8:05 to 8:35. This approach helps replicate the dynamic nature of service demands throughout the day. An episode is terminated once all passengers have been served and delivered to their destinations. This method allows us to replicate the spatial and temporal patterns observed in the bicycle rental history. Time periods without shared bicycle rental history are excluded, resulting in a total of 45,817 unique episodes for one complete experiment. Through conducting 1,000 such experiments, the system accumulates experience and progressively improves its performance.

For validation, three simulation scenarios were devised. In Scenario 1, we examine the case where a single autonomous shuttle bus is tasked with handling all service requests. This scenario allows us to identify any time points at which the service demands exceed the single vehicle's capacity to deliver services within the 20-minute timeframe. We selected this timeframe as a benchmark because it matches the maximum walking duration across the campus. Additionally, it corresponds to the time taken for a round trip by the shuttle across the campus. Our hypothesis is that user interest in the shuttle service might wane if the travel time exceeds 20 minutes. Following the results of Scenario 1, Scenario 2 is developed. In Scenario 2, additional vehicles are deployed during the problematic time periods identified in Scenario 1. This round of simulation assesses the improved efficiency of the service with the added vehicle.

In our study, we did not compare the shuttle service's simulation results with the shared bicycle rental history due to the distinct nature of their usage patterns. Bicycle rentals are often used for leisure or exercise, which significantly differs from the primarily transport-focused purpose of our shuttle service. Our main goal was to ensure that the shuttle service offered a time-efficient alternative to walking, thereby meeting the campus-wide need for accessible transportation. Consequently, to test our algorithm under more unpredictable conditions, we designed Scenario 3.

Scenario 3 introduces additional complexity to our simulations. Additional service requests were generated randomly with a 5% chance of occurrence, and all service requests were assigned random walking distances. These distances followed a normal distribution based on bicycle rental records (an average = 76.7 & a standard deviation = 29.6). This scenario is crafted to simulate an environment with fluctuating demand and varying passenger needs, providing a robust test for the adaptability and resilience of our algorithm across different operational environments.

C. RESULTS

The simulation outcomes were assessed using two key indicators: 'Request to Pick-up Time' and 'Request to Drop-off Time'. 'Request to Pick-up Time' represents the time interval from the moment of a service request to the moment a passenger successfully boards the autonomous shuttle bus. Additionally, 'Request to Drop-off Time' represents the total time duration from the service request until the passenger reaches their final drop-off point.

Results from the first scenario are presented in Fig. 5 with daily average values. There were two periods of high volumes of service requests: from October to December 2021 and from March to June 2022. In contrast to the RL-only model (Fig. 3), our hybrid approach algorithm, which integrates Genetic Algorithms (GA), was more effective in handling these high-demand service requests. While the RL-only model managed the high-demand service requests with an average 'Request to Drop-off Time' of 23 minutes, the hybrid algorithm model significantly reduced this time to approximately 18 minutes.

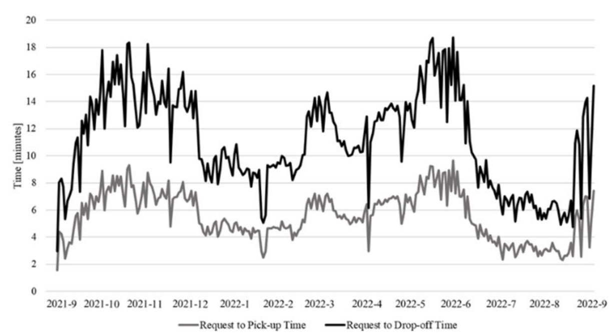


FIGURE 5. Scenario 1 results (daily average).

Fig. 6 presents a detailed comparative analysis, delineating the performance of the Fixed Route, Greedy algorithm, RL-only algorithm and our hybrid approach algorithm. The Fixed Route approach, which involves the shuttle stopping at all links, consistently resulted in higher 'Request to Drop-off Times', reflecting its inflexibility to adapt to varying service demands. The Greedy algorithm showed some improvement, yet it was unable to consistently match the more dynamic demands efficiently. While the RL-only strategy marked an improvement over the Greedy algorithm during low demand seasons, it did not maintain consistently better results, showing variability during peak seasons. In contrast, our hybrid algorithm outperformed the other methods, maintaining lower 'Request to Drop-off Times'. It's evident from the performance trends that our hybrid approach consistently achieves lower 'Request to Drop-off Times', demonstrating its robustness in optimizing shuttle routes more efficiently than the conventional methods.

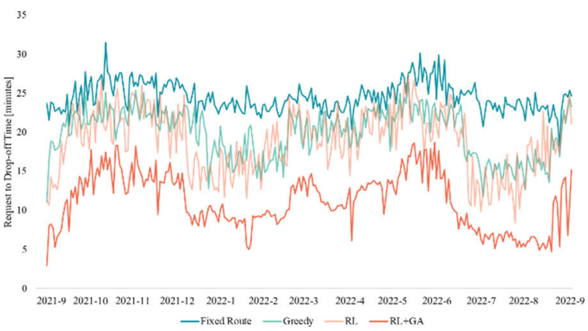


FIGURE 6. Performance comparison of fixed route, greedy and RL+GA.

Despite the hybrid algorithm achieving the daily average 'Request to Drop-off Time' within the 20-minute target, a standard deviation of 3.8 minutes indicates notable inconsistencies. The service quality fluctuates, with certain days and times experiencing significantly longer drop-off times than the average. This inconsistency in service times has implications for passenger satisfaction and the reliability perception of the shuttle service.

Table 4 presents the percentages of passengers whose 'Request to Drop-off Times' exceeded 20 minutes. A darker black cell background indicates a higher proportion of passengers experiencing 'Request to Drop-off Times' beyond 20 minutes. Except for the month of June, most passengers were serviced within 20 minutes during the early morning hours (0-8 AM). However, after 8 AM, service requests noticeably increased, leading to a higher frequency of 'Request to Drop-off Times' exceeding the 20-minute threshold. This trend was particularly prominent from October to December 2021 and from March to June 2022, during which over 30% of passengers failed to receive their transport services within 20 minutes. This significant proportion of service delays during these periods indicates a need for additional operational shuttle buses.

In a further simulation of Scenario 2, the proposed vehicle dispatch schedule was implemented and tested. As illustrated in Fig. 7, a substantial reduction in both 'Request to Pick-up Time' and 'Request to Drop-off Time' was recorded. This improvement is attributable to the concurrent operation

TABLE 4. Percentage of Passengers with 'Request to Drop-off Times' Over 20 Minutes.

Hour	Service Request Period												
	21 Sep	21 Oct	21 Nov	21 Dec	22 Jan	22 Feb	22 Mar	22 Apr	22 May	22 June	22 July	22 Aug	22 Sep
0-1	0.0	6.9	1.0	7.0	0.0	1.1	0.0	8.4	4.2	**14.7**	0.0	0.0	2.8
1-2	0.0	1.2	1.1	4.1	0.0	0.0	0.0	2.0	4.5	**11.4**	0.0	0.0	0.0
2-3	0.0	9.9	0.0	3.3	0.0	0.0	0.0	1.7	1.7	**6.8**	0.0	0.0	0.0
3-4	0.0	3.8	0.5	2.0	0.0	0.0	0.0	4.2	0.8	**4.3**	0.0	0.0	0.0
4-5	0.0	0.5	2.3	5.7	0.0	0.0	0.0	1.5	3.6	**13.0**	0.0	0.0	0.0
5-6	0.0	1.6	0.0	0.0	0.0	0.0	0.0	0.0	0.0	**3.7**	0.0	0.0	0.0
6-7	0.0	1.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	**5.3**	0.0	0.0	0.0
7-8	2.3	0.0	1.6	0.0	0.0	0.0	0.0	0.0	0.0	**2.4**	0.0	0.0	0.0
8-9	**9.4**	**17.8**	**20.3**	**16.6**	**14.7**	**11.5**	**18.5**	**13.4**	**30.1**	**24.4**	3.1	2.9	**27.8**
9-10	**6.9**	**17.1**	**15.3**	**16.8**	**14.7**	**12.8**	**14.8**	**14.6**	**21.0**	**25.5**	5.2	3.1	**16.2**
10-11	**4.5**	**22.1**	**18.4**	**22.3**	**0.0**	**1.3**	**20.3**	**16.6**	**27.1**	**32.3**	2.5	2.2	**20.0**
11-12	**7.5**	**24.8**	**22.0**	**23.0**	**15.7**	**15.6**	**20.8**	**19.4**	**31.5**	**29.3**	3.0	1.5	**22.6**
12-13	**17.3**	**29.0**	**35.7**	**30.6**	**18.8**	**17.1**	**26.9**	**27.1**	**38.4**	**30.9**	6.6	8.8	**31.4**
13-14	**17.8**	**31.3**	**32.2**	**26.5**	**10.6**	**11.7**	**23.7**	**25.0**	**29.8**	**32.1**	4.8	1.1	**21.7**
14-15	**13.5**	**34.8**	**33.1**	**30.8**	**2.4**	**5.3**	**26.6**	**25.7**	**32.4**	**35.3**	1.6	9.4	**26.3**
15-16	**10.5**	**32.5**	**30.2**	**28.5**	**4.4**	**3.5**	**24.6**	**25.4**	**34.1**	**34.1**	5.2	4.6	**24.5**
16-17	**18.9**	**33.9**	**36.7**	**26.6**	8.5	6.7	**25.1**	**23.3**	**32.6**	**30.8**	3.2	1.7	**20.9**
17-18	**19.7**	**33.7**	**34.7**	**28.2**	7.6	7.2	**24.3**	**22.2**	**32.5**	**33.2**	4.3	0.4	**21.8**
18-19	**18.0**	**37.5**	**40.3**	**29.8**	15.1	16.6	**27.2**	**25.2**	**32.6**	**29.8**	6.6	4.4	**21.4**
19-20	**8.3**	**30.2**	**32.6**	**27.7**	1.4	8.3	**22.3**	**21.8**	**29.8**	**28.7**	1.0	2.5	**19.1**
20-21	**10.5**	**28.0**	**28.1**	**22.9**	2.3	3.8	**16.7**	**18.2**	**24.9**	**27.4**	0.8	0.0	**9.8**
21-22	**11.3**	**24.8**	**24.1**	**20.4**	0.4	4.1	**12.0**	**15.1**	**20.4**	**21.1**	2.6	0.0	**6.4**
22-23	**5.7**	**24.0**	**19.8**	**15.6**	0.0	0.0	**12.7**	**17.2**	**20.2**	**22.1**	0.0	0.0	**11.9**
23-24	**5.8**	**17.0**	**17.7**	**16.5**	0.0	0.0	**5.1**	**12.7**	**19.1**	**20.2**	0.0	0.0	**10.8**

of two vehicles during peak demand periods. In the month of June, the operation of a single vehicle resulted in an average 'Request to Drop-off Time' of over 18 minutes. However, the strategic dispatch of an additional vehicle effectively decreased the average 'Request to Drop-off Time' to 11 minutes.

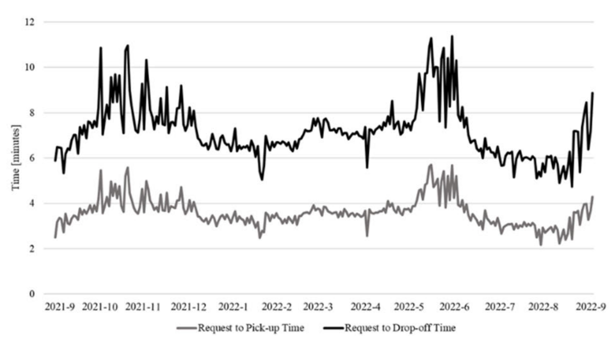


FIGURE 7. Scenario 2 results (daily average).

Our findings indicate the significance of dynamic demand patterns in shaping operational strategies. During peak periods, over 30% of passengers experienced 'Request to Drop-off Times' beyond 20 minutes from the result of simulation Scenario 1. This observation reinforces the need for transit agencies to consider temporal variations in demand while planning their services. The introduction of an additional shuttle during peak demand periods in Scenario 2 not only reduced the average 'Request to Pick-up Time' and 'Request to Drop-off Time', but also decreased the variation in these metrics. The standard deviation of 'Request to Pick-up Time' was reduced from 2.7 in Scenario 1 to 1.8 in Scenario 2, and the standard deviation of 'Request to Drop-off Time' was reduced from 3.8 to 2.4. This reduction in variability suggests a more consistent and reliable service. By implementing our proposed method, agencies could better align their vehicle deployment with demand patterns, thereby improving service efficiency and user experience.

In response to more fluctuating demand, Scenario 3 was designed to evaluate our algorithm. Fig. 8 showcases the actual daily counts of randomly generated service requests under this scenario. From October 2021 to December 2021, a high level between 10 and 15 random service requests was generated daily. A notable spike in the random service requests was also observed around June 2022. There exists a distinct correlation between days of higher original demand and days with more counts of random service requests.

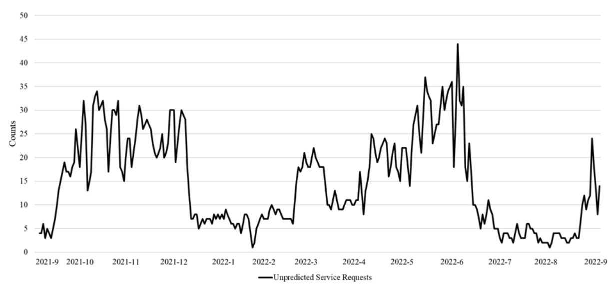


FIGURE 8. Randomly generated service requests (daily count).

Consistently, the metrics 'Request to Pick-up Time' and 'Request to Drop-off Time' were used to evaluate the simulation results of Scenario 3. These results are illustrated in Fig. 9. Compared to Scenario 2, a noticeable difference was observed during periods with a high level of random service requests. During these periods, there was an increase in both 'Request to Pick-up Time' and 'Request to Drop-off Time'. The introduction of a stochastic variable relating to the walking distance for all service requests in Scenario 3 could have added a layer of intricacy to the service dynamics, potentially causing service delays.

When faced with a surge in random service requests, requires more time to process and fulfill these requests. Such findings found that the algorithm's performance constraints and its adaptability in handling a large volume of unpredictable demands. Notably, even if the top level of random service requests were generated in June 2022, the 'Request to Drop-off Time' was managed under 17 minutes. During this month, two vehicles were deployed for the whole day. This data is instrumental for refining the algorithm, particularly in strategizing vehicle deployment to guarantee optimal service delivery.

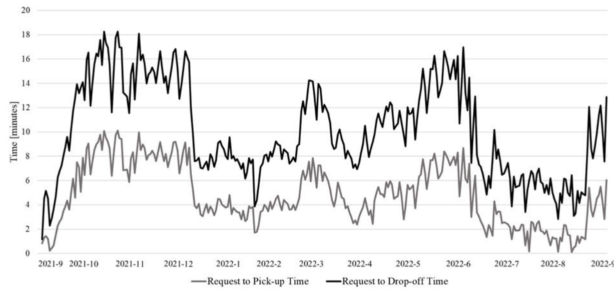


FIGURE 9. Scenario 3 results (daily average).

V. DISCUSSION AND CONCLUSION

The increased interest in autonomous vehicles has stimulated their exploration as a potential solution for many-to-many Demand-Responsive Transport (DRT) services. In this research, we devised an operational system for autonomous shuttle buses to coordinate service requests from an array of passengers, each with differing pick-up and drop-off points. The aim was to enhance efficiency and passenger satisfaction while maintaining cost-effectiveness.

Our study developed two algorithms. The 'On-Demand Boarding Guidance Algorithm' was designed to manage service requests based on their location and closeness to operational areas, ensuring convenience and viability. The 'Route Optimization Algorithm', a more intricate component, employed a hybrid model that combined a Genetic Algorithm (GA) with Reinforcement Learning (RL) to establish the most efficient route considering multiple simultaneous service requests. These algorithms were crucial in the development of an operational system that could optimally manage the dynamic demand patterns that characterize many-to-many DRT services.

Our simulations were conducted within the Chungju campus of the Korea National University of Transportation (KNUT) and the adjoining local community area. Over 30% of passengers experienced 'Request to Drop-off Times' beyond 20 minutes with the deployment of a single shuttle. The periods with high service demand necessarily require additional vehicles during peak times. When additional shuttles were deployed during high-demand periods, substantial improvements in both 'Request to Pick-up Time', which was maintained below 4 minutes, and 'Request to Drop-off Time', kept under 12 minutes. This enhancement could be attributed to our algorithm's ability to efficiently manage new service requests. As new requests came in, the algorithm determined the most suitable shuttle based on a fitness function. The shuttle with the lowest fitness function sum was selected to service the new request, thus optimizing efficiency for both new and existing passengers. This strategic approach not only validated our methodology but also highlighted the adaptability of our shuttle allocation system in meeting dynamic service demands.

Our research has significant implications for transit agencies, urban planners, and policymakers interested in harnessing autonomous vehicles to meet public transit needs. The proposed approach is expected to boost the service efficiency of autonomous shuttle buses in a many-to-many DRT context, making them a viable solution for the future public transport system. As we progress towards implementing higher levels of autonomous vehicles (SAE levels 4 and 5) in urban mobility, the potential risks of cyberattacks in connected and automated vehicles were critically analyzed by previous studies [35], [36], [37]. The vital role of artificial intelligence was emphasized by Nascimento et al. to enhance autonomous vehicle safety [38]. Our approach, combining reinforcement learning and genetic algorithms, aligns with this progression but necessitates improved against cyber threats in future works.

Our experiments were conducted using historical data from shared rental records. While this data provided a substantial basis for our simulations, it may not fully align with the demand patterns associated with autonomous shuttle bus services. Recognizing this potential discrepancy, we have undertaken additional scenario analyses, introducing randomly generated service requests and a variety of walking distances to boarding locations. When faced with a peak of random service requests in June 2022, the algorithm efficiently managed the 'Request to Drop-off Time' within 17 minutes. Despite these efforts to demonstrate the adaptability of our algorithm, it is imperative that future research continues to test and validate the applicability of our algorithm across different operational contexts.

This study does come with certain limitations. Future studies could seek to develop models that are capable of learning and adapting in response to real-time changes in demand and circumstances. It is essential to confirm the algorithm's effectiveness and robustness in future studies, particularly in situations where data quality might vary. Additionally, it is important to note that our study is conducted within the unique context of the KNUT campus, characterized by an absence of local roads, leading to stable and easily predictable vehicular traffic conditions. Nonetheless, for applications in more complex urban environments, future research should focus on developing robust models capable of learning and adapting in real-time to fluctuating traffic conditions and other unforeseen events to maintain a consistent level of service efficiency.

Incorporating prior information about service request patterns at different Origin-Destination Descriptions (ODDs) could significantly enhance the route optimization process. Future studies should explore the integration of historical demand data, which could provide the algorithm with a predictive edge, allowing for proactive rather than reactive route planning. This could be particularly beneficial in optimizing routes by anticipating high-demand locations and times, thereby improving service reliability and efficiency.

REFERENCES

[1] J. Cregger, M. Dawes, S. Fischer, C. Lowenthal, E. Machek, and D. Perlman. (2018). *Low-Speed Automated Shuttles: State of the Practice Final Report*. [Online]. Available: https://rosap.ntl.bts.gov/view/dot/37060

[2] M. Mahmoodi Nesheli, L. Li, M. Palm, and A. Shalaby, "Driverless shuttle pilots: Lessons for automated transit technology deployment," *Case Stud. Transp. Policy*, vol. 9, no. 2, pp. 723–742, Jun. 2021.

[3] M. Azad, N. Hoseinzadeh, C. Brakewood, C. R. Cherry, and L. D. Han, "Fully autonomous buses: A literature review and future research directions," *J. Adv. Transp.*, vol. 2019, pp. 1–16, Dec. 2019.

[4] C. Iclodean, N. Cordos, and B. O. Varga, "Autonomous shuttle bus for public transportation: A review," *Energies*, vol. 13, no. 11, p. 2917, Jun. 2020.

[5] J. Ainsalu et al., "State of the art of automated buses," *Sustainability*, vol. 10, no. 9, 2018, Art. no. 3118.

[6] M. W. Savelsbergh and M. Goetschalckx, "A comparison of the efficiency of fixed versus variable vehi," *J. Bus. Logistics*, vol. 16, no. 1, p. 163, 1995.

[7] A. L. Erera, M. Savelsbergh, and E. Uyar, "Fixed routes with backup vehicles for stochastic vehicle routing problems with time constraints," *Networks*, vol. 54, no. 4, pp. 270–283, Dec. 2009.

[8] C. M. Novoa, "Static and dynamic approaches for solving the vehicle routing problem with stochastic demands," Ph.D. dissertation, Dept. Ind. Syst., Lehigh Univ., Pennsylvania, PA, USA, 2005.

[9] A. Bucchiarone, S. Battisti, A. Marconi, R. Maldacea, and D. C. Ponce, "Autonomous shuttle-as-a-service (ASaaS): Challenges, opportunities, and social implications," *IEEE Trans. Intell. Transp. Syst.*, vol. 22, no. 6, pp. 3790–3799, Jun. 2021.

[10] L. Davison, M. Enoch, T. Ryley, M. Quddus, and C. Wang, "Identifying potential market niches for demand responsive transport," *Res. Transp. Bus. Manage.*, vol. 3, pp. 50–61, Aug. 2012.

[11] M. Enoch, S. Potter, G. Parkhurst, and M. Smith, "Why do demand responsive transport systems fail?" in *Proc. Transp. Res. Board 85th Annu. Meeting*, Washington, DC, USA, 2006.

[12] L. Ferreira, P. Charles, and C. Tether, "Evaluating flexible transport solutions," *Transp. Planning Technol.*, vol. 30, nos. 2–3, pp. 249–269, Apr. 2007.

[13] N. Walker, P. Harbutt, and J. Morris, "Addressing access and mobility issues in rural and regional Victoria," in *Proc. 27th Australas. Transp. Res. Forum (ATRF)*, Adelaide, SA, Australia, 2004.

[14] L. Davison, M. Enoch, T. Ryley, M. Quddus, and C. Wang, "A survey of demand responsive transport in great Britain," *Transp. Policy*, vol. 31, pp. 47–54, Jan. 2014.

[15] X. Yan, X. Zhao, Y. Han, P. V. Hentenryck, and T. Dillahunt, "Mobility-on-demand versus fixed-route transit systems: An evaluation of traveler preferences in low-income communities," *Transp. Res. A, Policy Pract.*, vol. 148, pp. 481–495, Jun. 2021.

[16] R. Cervero, *Paratransit in America: Redefining Mass Transportation*. Westport, CT, USA: Greenwood, 1997.

[17] J. D. Nelson, S. Wright, B. Masson, G. Ambrosino, and A. Naniopoulos, "Recent developments in flexible transport services," *Res. Transp. Econ.*, vol. 29, no. 1, pp. 243–248, Jan. 2010.

[18] L. Fu, "A simulation model for evaluating advanced dial-a-ride paratransit systems," *Transp. Res. A, Policy Pract.*, vol. 36, no. 4, pp. 291–307, May 2002.

[19] F. Cavallaro and S. Nocera, "Flexible-route integrated passenger–freight transport in rural areas," *Transp. Res. A, Policy Pract.*, vol. 169, Mar. 2023, Art. no. 103604.

[20] A. Anburuvel, W. U. L. D. P. Perera, and R. D. S. S. Randeniya, "A demand responsive public transport for a spatially scattered population in a developing country," *Case Stud. Transp. Policy*, vol. 10, no. 1, pp. 187–197, Mar. 2022.

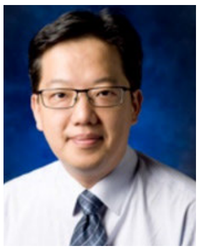
[21] Ç. Koç, G. Laporte, and İ. Tükenmez, "A review of vehicle routing with simultaneous pickup and delivery," *Comput. Oper. Res.*, vol. 122, Oct. 2020, Art. no. 104987.

[22] A. El-Geneidy, M. Grimsrud, R. Wasfi, P. Tétreault, and J. Surprenant-Legault, "New evidence on walking distances to transit stops: Identifying redundancies and gaps using variable service areas," *Transportation*, vol. 41, no. 1, pp. 193–210, Jan. 2014.

[23] J. Chia, J. Lee, and M. Kamruzzaman, "Walking to public transit: Exploring variations by socioeconomic status," *Int. J. Sustain. Transp.*, vol. 10, no. 9, pp. 805–814, Oct. 2016, doi: 10.1080/15568318.2016.1156792.
[24] R. S. Sutton and A. G. Barto, *Introduction To Reinforcement Learning*, vol. 135. Cambridge, MA, USA: MIT Press, 1998.
[25] S. Yoo and J. B. Lee, "Revising bus routes to improve access for the transport disadvantaged: A reinforcement learning approach," *J. Public Transp.*, vol. 25, 2023, Art. no. 100041.
[26] J. J. Q. Yu, W. Yu, and J. Gu, "Online vehicle routing with neural combinatorial optimization and deep reinforcement learning," *IEEE Trans. Intell. Transp. Syst.*, vol. 20, no. 10, pp. 3806–3817, Oct. 2019.
[27] S. Yoo, J. B. Lee, and H. Han, "A reinforcement learning approach for bus network design and frequency setting optimisation," *Public Transp.*, vol. 15, no. 2, pp. 503–534, Jun. 2023.
[28] W. Lu, Z. Yi, Y. Gu, Y. Rui, and B. Ran, "TD3LVSL: A lane-level variable speed limit approach based on twin delayed deep deterministic policy gradient in a connected automated vehicle environment," *Transp. Res. C, Emerg. Technol.*, vol. 153, Aug. 2023, Art. no. 104221.
[29] A. Khan and A. Lapkin, "Searching for optimal process routes: A reinforcement learning approach," *Comput. Chem. Eng.*, vol. 141, Oct. 2020, Art. no. 107027.
[30] M. S. Sheikh and Y. Peng, "Procedures, criteria, and machine learning techniques for network traffic classification: A survey," *IEEE Access*, vol. 10, pp. 61135–61158, 2022.
[31] A. Seyyedabbasi, R. Aliyev, F. Kiani, M. U. Gulle, H. Basyildiz, and M. A. Shah, "Hybrid algorithms based on combining reinforcement learning and metaheuristic methods to solve global optimization problems," *Knowl.-Based Syst.*, vol. 223, Jul. 2021, Art. no. 107044.
[32] J. H. Holland, "Genetic algorithms," *Sci. Amer.*, vol. 267, no. 1, pp. 66–73, 1992.
[33] KNUT. (2022). *Statistical Yearbook*. [Online]. Available: https://www.ut.ac.kr/prog/schulStatsDta/kor/sub01_02_07/list.do
[34] R. W. Bohannon, "Comfortable and maximum walking speed of adults aged 20–79 years: Reference values and determinants," *Age Ageing*, vol. 26, no. 1, pp. 15–19, 1997.
[35] M. Girdhar, Y. You, T.-J. Song, S. Ghosh, and J. Hong, "Post-accident cyberattack event analysis for connected and automated vehicles," *IEEE Access*, vol. 10, pp. 83176–83194, 2022.
[36] U. Alvi, M. A. K. Khattak, B. Shabir, A. W. Malik, and S. R. Muhammad, "A comprehensive study on IoT based accident detection systems for smart vehicles," *IEEE Access*, vol. 8, pp. 122480–122497, 2020.
[37] J. Cui, G. Sabaliauskaite, L. S. Liew, F. Zhou, and B. Zhang, "Collaborative analysis framework of safety and security for autonomous vehicles," *IEEE Access*, vol. 7, pp. 148672–148683, 2019.
[38] A. M. Nascimento, L. F. Vismari, C. B. S. T. Molina, P. S. Cugnasca, J. B. Camargo, J. R. de Almeida, R. Inam, E. Fersman, M. V. Marquezini, and A. Y. Hata, "A systematic literature review about the impact of artificial intelligence on autonomous vehicle safety," *IEEE Trans. Intell. Transp. Syst.*, vol. 21, no. 12, pp. 4928–4946, Dec. 2020.

SUNHYUNG YOO received the Ph.D. degree from The University of New South Wales (UNSW) with a focus on transportation planning and optimization. He has since furthered his commitment to the field as a Research Associate with UNSW. His primary research concentrates on the intricate optimization of bus routes, aiming to enhance accessibility for the transport disadvantaged. His research interests include urban transit planning, sustainable mobility solutions, and the integration of technology in public transport.

HYUN KIM is currently a Professor with the Department of Transportation and Energy Convergence, Korea National University of Transportation, and the Director of the Transportation and ICT Convergence Research Center. His research interests include technology, ITS, autonomous vehicles, AI, and big data and its interaction with mobility.



JINWOO (BRIAN) LEE received the Ph.D. degree in civil engineering from the University of Toronto, Canada. He is currently an Associate Professor in city planning with The University of New South Wales, Sydney. He led a number of research and industry projects on transport planning in Australia and Canada. His research interests include transport planning, travel behavior, active transport, integrated transport infrastructure, and land use planning. His current research focuses on exploring how to improve transport planning practices by integrating transport infrastructure and land use planning and enhanced understanding of travel behavior for sustainable and loveable cities.