Leveraging Deep Reinforcement Learning and Healthcare Devices for Active Travelling in Smart Cities

S. M. Ahsan Kazmi, Zaheer Khan, Adil Khan, Manuel Mazzara, Asad Masood Khattak

Abstract—**Smart cities are increasingly challenged by population growth and the environmental emissions of urban transportation systems, necessitating sustainable urban planning to improve public health, environmental quality, and overall urban livability. A notable aspect in this context is the underutilization of smart healthcare wearable devices or smart healthcare applications in urban transportation systems. This paper proposes an innovative approach to address these challenges effectively. We formulate a non-convex optimization problem aimed at minimizing environmental emissions within transportation systems while considering resident health goals, travel time constraints, and infrastructure limitations. To achieve this, we employ deep reinforcement learning (DRL), which dynamically selects the optimal traveling mode for residents. This approach aims to optimize environmental outcomes while meeting individualized mobility needs. Moreover, our method integrates smart healthcare technologies to capture real-time data and predict optimal traveling modes. By incorporating real-world health metrics into transportation planning, we enhance decision-making processes and promote active transportation options, contributing to healthier urban environments. Through extensive simulations, we demonstrate the effectiveness of our approach in optimizing traveling decisions and advancing sustainable urban mobility practices. Our DRL-based solution effectively promotes active travel, leading to a significant increase in health-related metrics (like calories burned) and a substantial reduction in gCO2 emissions. Up to 74% of journeys were made using active transportation modes. Cycling is particularly popular, accounting for up to 67% of journeys.**

Index Terms—**Intelligent Transport System, Edge-AI, Healthcare Devices, Environmental Impact, Active Transportation.**

I. Introduction

The world population could grow to almost 11 billion by the end of the twenty first century, a sharp rise from 7.9 billion in 2021 [1]. Among other public services, cities' transportation systems face constant pressure from rapid population growth, necessitating sophisticated policies for climate neutrality, urban resilience and sustainability. This includes addressing social, health, and environmental emissions alongside physical transport infrastructure. As urban populations expand, cities must prioritize enhancing residents' quality of life. Advanced technologies such as the Internet of Things (IoT) and Information and Communication Technology (ICT) are crucial in optimizing resource use, improving transport efficiency, reducing congestion, and mitigating environmental emissions to foster sustainable urban development [2], [3]. One important aspect of making cities' transport systems sustainable is enhancing residents' quality of life, given the widespread impact of sedentary lifestyles and poor dietary habits leading to health issues like obesity, cardiovascular diseases, and mental health disorders [4]. Cities face various health risks such as traffic exposure, air pollution, noise, social isolation, and limited access to healthy food. Additionally, the dependence on private vehicles for daily trips is also among a major factor leading to unhealthy lifestyles [5].

The health and well-being of urban citizens are central to sustainable city planning, significantly influencing population health. Current transport solutions typically focus on optimizing traffic management without adequately considering residents' healthcare needs. Moreover, traditional approaches like expanding road networks face challenges such as rising land costs and spatial constraints. While AI-enabled traffic management [6]–[8], connected vehicle systems [9], and electric vehicles [10], [11] offer partial solutions, urban designs prioritizing vehicular traffic contribute to unhealthy environments, promoting sedentary behavior and higher obesity rates. Thus, a novel solution that addresses both environmental emissions and public health is urgently needed.

The integration of smart wearable healthcare devices (SWH) and innovative mobile applications presents a significant opportunity for supporting healthier lifestyles and enhancing smart urban planning. Despite their potential benefits, there is currently an under-utilization of SWH devices and applications in urban transportation systems. These devices monitor physical activity, promote healthier living, and provide real-time health insights [12], [13]. By incorporating SWH devices into smart cities' transportation solutions, policymakers can develop systems that encourage physical activity, reduce pollution, and enhance overall well-being, thereby advancing sustainable urban mobility practices.

One effective strategy involves integrating AI-driven mobility solutions that promote active transportation modes as integral parts of users' lifestyles, complemented by leveraging SWH devices. These initiatives collectively address urban mobility challenges, environmental sustainability, and public health concerns [14]. Educating residents about these benefits can mitigate health risks linked to air pollution and enhance

S. M. Ahsan Kazmi and Zaheer Khan are with the faculty of Computer Science and creative technologies, University of the West of England, Bristol, UK. E-mails: {ahsan.kazmi,zaheer2.khan}@uwe.ac.uk

Adil Khan is with School of Computer Science, University of Hull, Hull HU6 7RX, UK, E-mail: a.m.khan@hull.ac.uk

Manuel Mazzara is with Innopolis university, Russia, E-mail: m.mazzara@innopolis.ru

Asad Masood Khattak is with the College of Technological Innovation, Zayed University, Dubai, United Arab Emirates, E-mail:asad.khattak@zu.ac.ae

overall well-being. With more than half of the global population residing in urban areas, adopting sustainable urban mobility solutions is critical for maintaining livable, healthy, and economically viable cities. Moreover, multimodal transportation systems decrease reliance on private cars, reduce carbon emissions, and support global efforts toward achieving net-zero emissions [15]. Prioritizing walking, cycling, and public transit also alleviates economic strains associated with traffic congestion and vehicle emissions. However, the lack of strong incentives currently hinders the widespread adoption of these measures. Therefore, developing solutions that empower each user to enhance their health and contribute to societal well-being is essential for fostering sustainable cities. To address the complex transportation challenges in smart cities, we propose a solution that integrates AI, edge computing, and SWH technologies. Our approach utilizes Deep Reinforcement Learning (DRL) to enable adaptive decision-making in dynamic urban environments. The goal is to create a sustainable transportation system that minimizes environmental emissions while aligning with residents' health goals and satisfaction metrics such as comfort and travel time. Using a DRL-based solution, our approach facilitates adaptive decision-making in dynamic urban environments. The primary aim is to establish a sustainable transportation system that minimizes environmental emissions while meeting residents' health goals and satisfaction metrics, such as comfort and travel time. Our model focuses on a single community where an edge server, strategically positioned with roadside units, manages real-time transport data efficiently [16], [17]. This edge server collects and processes diverse metrics crucial for our DRL-based solution, including real-time traffic conditions, environmental emissions linked to different transport modes, SWH device real-time data, and individual user preferences. By addressing these factors comprehensively, our goal is to optimize route planning strategies, encourage the adoption of sustainable transportation modes, and enhance overall environmental sustainability tailored to each user's specific circumstances and health status. Note that our DRL-based approach offers a valuable tool for urban transportation planning by enabling dynamic policy adjustments and rapid policy evaluation. By leveraging real-time data from wearable health devices and transportation sensors, we can optimize transportation systems to reduce congestion, emissions, and promote physical activity. This adaptive approach is particularly valuable in rapidly evolving urban environments where timely policy responses are essential. The DRL model is trained and implemented on the edge server within the community area to ensure rapid responsiveness and operational effectiveness. In summary, our contributions include:

- We propose an innovative system model that formulates a non-convex optimization problem tailored for efficient traveling in smart cities, emphasizing residents' transportation mode choices and health goals. This approach seeks to minimize environmental emissions from traveling while prioritizing health objectives. Traditional methods for tackling these challenges often involve parameter relaxation and iterative approaches to approximate near-optimal solutions.
- To address this, we propose a DRL-based algorithm designed to adapt to the dynamic aspects of the problem. This includes variables such as time of day, capacity of each mode, SWH device metrics, and adherence to user satisfaction targets. The algorithm aims to intelligently select the most suitable transportation mode that balances health objectives and environmental emission reduction goals effectively.
- Through extensive simulations, we ensure convergence and significant promotion of active traveling modes via the proposed solution, with up to 74% of journeys opting for active modes. Additionally, Our solution excels in reducing environmental emissions (gCO2), particularly for shorter distances. Moreover, our approach prioritizes choices that maximize health benefits, minimize environmental emission, and ensure resident satisfaction. This emphasis on health and sustainability is underscored by the popularity of cycling, which accounts for up to 67% of journeys in our simulations.

The remainder of this paper is structured as follows: Section II provides a comprehensive review of the literature. Section III outlines the system model. Our proposed solution is detailed in Section IV. Section V, presents the analysis of numerical results to validate our solution's performance. Finally, Section VI draws conclusions based on our findings.

II. Literature Review

This section provides a comprehensive review of the literature concerning active transportation modes and their impact on smart city environments, focusing on methodologies and technologies that assess their societal benefits. Numerous studies underscore the significance of active mobility in promoting public health and environmental sustainability. For instance, the work in [18] explores the health impacts of policies promoting active travel, proposing health impact assessment models to inform decision-making for healthier urban environments. Similarly, [19] analyzes policy scenarios in Southeast Asian cities, demonstrating that integrated measures can effectively reduce CO2 emissions and enhance public transportation systems through active transportation, highlighting the role of factors like transit accessibility and residential density in promoting walking while reducing reliance on motorized transport. In addition to the aforementioned studies, recent research has increasingly focused on the importance of active transport in various cities, particularly with newly implemented policies aimed at enhancing environmental sustainability. For instance, research modeling scenarios in Porto, Portugal in [20] reveals significant reductions in disability-adjusted life years through increased active transportation, alongside benefits in reducing traffic injuries and air pollution, with cardiovascular diseases showing substantial improvements. These findings underscore active transportation's potential to yield significant health and economic benefits, emphasizing the need to integrate health considerations into urban transport policy planning. However, these studies have not utilized advanced healthcare technologies now available. In contrast, the authors in

[21] introduce a web-based application designed to promote sustainable travel through persuasive tools that recommend transport modes (walking, cycling, public transportation, car) based on user feedback and sustainability factors. While these studies provide valuable insights, they often overlook the crucial aspect of promoting active transportation, such as cycling or walking, which requires consideration of scenarios that benefit overall smart city residents. Active transportation may pose health risks if residents are exposed to adverse environmental conditions, as highlighted in studies assessing the combined impact of physical activity and exposure to fine particulate matter (PM2.5) on all-cause mortality [22]. Indeed, while active transportation leads to significant annual reductions in mortality rates, restrictions on active travel during high PM2.5 days did not mitigate mortality risks and could potentially worsen overall health, particularly in heavily polluted cities. This underscores the importance of revisiting guidelines discouraging active travel during periods of high air pollution, highlighting the complex relationship between physical activity benefits and health risks from air pollution in shaping effective public health policies. Overall these solutions highlight the importance of active transport and demonstrate improved performance but they are often tailored to specific cities, limiting their scalability and broader applicability due to the exponential growth of data. Moreover, these studies frequently overlook the potential benefits of integrating modern solutions, such as machine learning perspectives and healthcare applications, into the transportation sector to develop intelligent solutions adaptable across diverse scenarios. We aim to explore how advanced technologies, particularly machine learning, are leveraged to assess and optimize the impact of active mobility on cities, bridging the gap between academic research and practical urban planning.

Machine learning led approaches have recently gained popularity in the transportation community due to their ability to address numerous practical challenges. For instance, researchers in [6] introduce a novel method that integrates a two-stage Gaussian pseudo-spectral method with a decision tree algorithm to enhance unmanned vehicle parking capabilities and trajectory planning speed, proving effective for complex parking scenarios and laying the groundwork for intelligent planning strategies tailored to various parking environments. Similarly, another study by [7] presents an approach to accelerate deep neural network-based tasks in vehicular edge computing by partitioning and offloading tasks among vehicles and roadside infrastructures, considering heterogeneous computation and communication capacities. While these solutions are promising, they do not capture the dynamic interactions inherent in transportation environments. In transportation scenarios, capturing interactions with the environment is crucial for designing a generalized solution that can be adapted to various situations. Reinforcement learning (RL) is particularly well-suited for this purpose as it inherently learns by interacting with the environment and many studies have applied RL in the transportation domain. For instance, an RL-based solution was proposed in [8] that addresses the challenge of optimizing service caching strategies in intelligent cyber-physical transportation systems using cloud-edge computing to improve cache hit rates. Similarly, another RL-based solution was studied in [23] that optimizes train operations for high speed railways. Another similar RL-based solution was presented for dynamic traffic environments with time-varying communication topologies to reduce collision rates [24]. In [25], the authors propose a route recommendation system to balance fuel consumption, travel time, and air quality. The work in [26] employs RL to understand user preferences in attraction recommendations. Other interesting RL-based solutions within ITS are presented in [27]–[30].

These works present valuable machine learning-led solutions that enhance transportation systems in smart cities. However, most of these studies do not consider the impact of multiple modes of transportation available in smart cities or their impact on the sustainability of the transportation system. To model a realistic urban transport system, it is essential to include multiple modes of travel, such as walking and cycling, which are not adequately addressed in current models. Incorporating these modes can lead to a more practical and comprehensive smart city urban mobility system and support initiatives such as livable neighborhoods or 15-minute cities [31]. For instance, in [32], authors compare different transportation means to reduce travel times and emissions using existing infrastructures. Additionally, multiple modes of transportation can significantly lower costs and improve passenger satisfaction, as explored in [33], [34].

In summary, while existing solutions have addressed various aspects of transportation challenges in smart cities, they often lack a model that represents a practical transport system incorporating multiple modes of transportation with dynamic requirements integrated with heath related applications or SWH devices. There is a need for a system that can analyze environmental impacts while utilizing advanced technologies such as machine learning and modern healthcare applications to align with user needs. By leveraging RL and other advanced techniques, future ITS can promote healthier journeys through walking and cycling, significantly reducing environmental emissions and enhancing overall users' health and well-being.

III. System Model and Problem Formulation

Fig. 1 presents our system model, depicting a smart city transport system that includes a set of residents/users, denoted by the set $\mathcal{R}$, each connected to an SWH device representing their current health status. These residents are traveling via different transport modes, represented by the set $\mathcal{T}$, each characterized by distinct attributes, benefits, and limitations, to capture the diverse mobility options available. The road network is presented as a graph consisting of edges and vertices, represented by sets $\mathcal{E}$ and $\mathcal{V}$, which serve as the foundational framework for our analysis.

Overall, we consider a comprehensive system model tailored to address the complexities of smart city transport systems, integrating health data to enhance user well-being. Our model takes into account critical factors influencing transportation dynamics, such as time-varying traffic conditions, mode-specific environmental emissions, health metrics collected from SWH

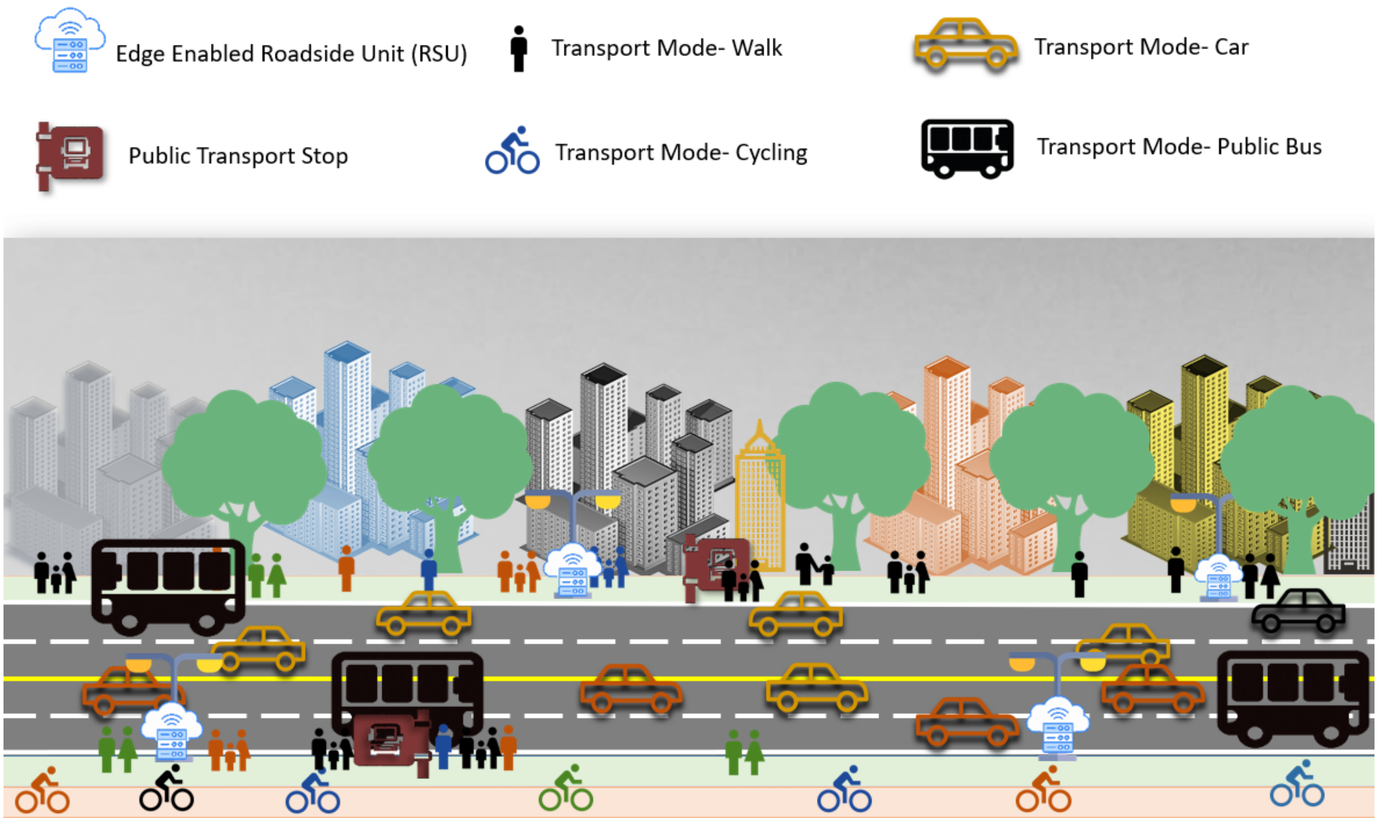


Fig. 1: Urban Transport System: Road Networks, Transport Modes, and Edge-Enabled Healthcare Devices.

devices, and user preferences. Implementing this solution requires real-time or dynamic information within a specific area. While such data could reside in a central cloud, we use edge computing for local storage and processing. Edge computing offers reduced latency and enhanced processing speed, making it ideal for our needs. Accurate and up-to-date information can be gathered through user sign-ins to a dedicated application that tracks traveling modes, preferences, and health metrics, or via sensors placed along routes or integrated into user devices. This data is then transmitted to the edge server, ensuring timely updates. Additionally, Ensuring the security and privacy of user data is paramount. Edge AI allows for the processing of sensitive health and location data closer to the data source, reducing the risk of data breaches associated with central cloud storage. Techniques such as data encryption and secure access controls can be employed to safeguard the information collected. Furthermore, privacy-preserving mechanisms, such as differential privacy, are integrated to ensure that individual user data remains confidential while still allowing for aggregate data analysis. These measures are crucial for maintaining user trust and compliance with data protection regulations. Moreover, by leveraging blockchain technology, we can provide privacy, self-verification, authentication, and authorization of transactions, ensuring that citizen data is handled securely at different administrative and geographical levels. This approach ensures that participatory data is kept and processed locally, enhancing privacy and providing an economic approach for resource utilization in a distributed environment [35]. Our approach can also incorporate additional security and privacy mechanisms without affecting its performance, further enhancing the security and privacy of residents. By addressing these factors, we aim to optimize route planning strategies, encouraging users to choose modes that align with their preferences and health status while also reducing environmental emissions.

A. Road & Modes Modelling

In our work, we represent the road network as a graph consisting of vertices $\mathcal{V}$ and edges $\mathcal{V}$, Each edge $e \in \mathcal{E}$ is characterized by a set of attributes at any time instant t, represented by the tuple $R_e^t = \{l, cp, cl\}$. Here, l is the edge length, cp is the capacity, and cl is the congestion level. These attributes are numerical values or properties that describe the edge's characteristics at a given time. Geographic data, defined by coordinates (latitude and longitude), provides the location of each edge. Similarly, each vertex $v \in \mathcal{V}$ has a set of attributes represented by $R_v^t = \{x, y, r, b\}$, where x and y are the coordinates (latitude, longitude), $r \in \mathcal{R}$ is the set of users present at time t, and b indicates whether the vertex is a bus stop.

Next, we present the available transportation modes considered in this model. The available transportation modes are denoted as $\mathcal{T}$ and include public transit[1], walking, cycling, and cars. Each mode $m \in \mathcal{T}$ is defined by attributes such as travel speed s_m, capacity cp_m, and environmental emissions e_m in

[1]In this work, we only assume public busses as a public transport mode.

TABLE I: Summary of the key notations.

Notation	Definition
$\mathcal{R}$	Set of residents/users
$\mathcal{E}$	Set of edges
$\mathcal{V}$	Set of vertices
$\mathcal{T}$	Set of transportation modes
l	edge length
cp	edge capacity
cl	edge congestion level
b	Set of bus stops
s_m	Travel speed of mode m
cp_m	Capacity of mode m
e_m	Environmental emission of mode m
t_r	Travel time deadline for resident r
gh_r^{min}	Minimum heart rate goal for resident r
gh_r^{max}	Maximum heart rate goal for resident r
m_r	Available transport modes for resident r
$h_r(t)$	Heart rate for resident r at time instant t
$\alpha_{r,m}^e(t)$	Mode m selection for edge e at time instant t
$c_r(t)$	Calorie count for resident r at time instant t
γ	Tuning parameter to balance emission & calorie count

terms of emissions. These attributes are captured in the tuple $R_m^t = \{sp_m, cp_m, e_m\}$. Moreover, we assume that all modes follow the same edges[2], with each edge's properties defined by R_e^t. This modeling allows us to integrate various factors on the road network and transport mode. Next, we elaborate on how users are modeled within the system and specify their requirements.

B. User Modelling

In our model, we represent residents by the set $\mathcal{R}$, where each resident $r \in \mathcal{R}$ has specific requirements and unique characteristics. Each resident r is equipped with an SWH device[3]. These devices can hold personal information, such as weight, height, age, and fitness goals (e.g., taking a certain number of steps or burning a specific number of calories). Typically an SWH device[4] can track factors, such as heart rate, to monitor activities comprehensively, the number of steps taken, and calories burned during an activity. We aim to integrate mobility tasks with health goals, allowing users to travel efficiently while meeting their health objectives. Each resident's requirements are represented by a tuple $R_r^t = \{t_r, gh_r^{min}, gh_r^{max}, m_r\}$, where t_r represents the travel time deadline, gh_r^{min} and gh_r^{max} indicate the minimum and maximum heart rate goal stemming from the SWH, respectively, and m_r represents the available transport modes at time instant t, acknowledging that not all users may have access to every mode. This tuple effectively captures the diverse needs and preferences of urban travellers, considering travel time, and health goals.

The selection of a traveling mode is influenced by numerous external factors, including distance, time of day, weather conditions, available infrastructure, and health metrics. These factors define the constraints or preferences for utilizing different transportation options. For instance, walking may be favored for short distances in pleasant weather, public transit may be optimal during peak hours, and cycling might be chosen for routes with dedicated bike lanes when the weather is favorable. Since these factors are dynamic and can change over time, decisions regarding the best mode of transport must be made for each specific time slot t. We present this decision by the following binary variable if a resident r chooses mode m at time instant t for the edge e:

$$\alpha_{r,m}^e(t) = \begin{cases} 1, \text{if mode } m \text{ selected for edge } e, \\ 0, \text{otherwise.} \end{cases} \tag{1}$$

In this context, the binary variable $\alpha_{r,m}^e(t)$ for each mode $m \in \mathcal{M}$ represents the selection or recommended preference for mode m for edge e during a particular time slot t.

C. Problem Formulation

Our goal is to develop a transportation solution that meets each user's health and latency requirements while minimizing environmental emissions for the collective benefit of society. To achieve this, we formulate a problem focused on reducing overall environmental emissions and maximizing the use of sustainable transportation modes, all while fulfilling users' health and time objectives. Our objective involves addressing several constraints to reach this goal. First, we include a constraint to ensure that each user selects no more than one transportation mode during each time slot t, expressed as:

$$\sum_{m \in \mathcal{T}} \sum_{e \in \mathcal{E}} \alpha_{r,m}^e(t) \leq 1, \quad \forall r \in \mathcal{R}. \tag{2}$$

Secondly, we include a mode availability constraint to ensure that if mode m is not available to residents r at time t, it cannot be selected. This constraint is crucial because certain conditions, such as rain, may render cycling impractical, while walking might be preferred for short distances, or cycling might be optimal for routes with dedicated bike lanes in good weather. We represent these constraints as follows:

$$\alpha_{r,m}^e(t) \leq m_r(t), \quad \forall r \in \mathcal{R}, \forall m \in \mathcal{M}. \tag{3}$$

To integrate real-time traffic data and environmental emissions, we define additional constraints that take into account congestion levels, health metrics, capacity, and required time associated with different modes of transportation. For each mode $m \in \mathcal{T}$ and edge $e \in \mathcal{E}$, we introduce constraints to incorporate these factors:

$$\alpha_{r,m}^e(t) \leq 1 - cl_e(t), \tag{4}$$

$$gh_r^{min}(t) \leq \alpha_{r,m}^e(t) \cdot h_r(t) \leq gh_r^{max}(t), \forall r \in \mathcal{R}, \tag{5}$$

$$\sum_{m} \sum_{r} \alpha_{r,m}^e(t) \leq cp_e(t), \tag{6}$$

$$\sum_{m} \sum_{t} \alpha_{r,m}^e(t) \leq t_r(t), \forall r \in \mathcal{R}, \tag{7}$$

where $cl_e(t)$ represents the congestion level on edge e at time slot t, and $h_r(t)$ represents the instantaneous heart rate, as

[2]In practical settings, each travel mode in smart cities would have different edge properties. However, for simplicity, we have considered a uniform edge in our model and will explore varying edge properties in future work.

[3]We can also use alternative technologies or gadgets such as a smartphone application or smartwatches, capable of monitoring physical activity.

[4]Modern devices come equipped with numerous sensors, including those for environmental impact measurements like air quality (PM2.5, PM10, NOx, etc.). However, for this study, we only consider health-related metrics and will include other environmental metrics in future work.

recorded by the SHW monitors at time slot t. Additionally, cp_e is the capacity of the edge and t_r is the time or latency threshold associated with the user travel time. Then, the utility function is formulated as follows:

$$U_r(t) = \sum_e \sum_m \sum_t \alpha^e_{r,m}(t) \cdot [\gamma c_r(t) + (1-\gamma)\frac{1}{e_m(t)}], \quad (8)$$

where c_r represents the instantaneous calorie count and γ represents the tuning parameter that balances between emission and calorie count. By integrating these constraints and optimizing our utility function, which aims to minimize environmental emissions and maximize the calorie count, our approach seeks to provide sustainable and efficient transportation solutions that involve considering real-time data, health metrics, and various dynamic factors that influence mobility choices. Formally, by combining the aforementioned constraints and utility function, we define our problem as follows:

$$\max_{\boldsymbol{\alpha}} \quad \limsup_{T\to\infty} \frac{1}{T} \sum_{t=0}^{T} \sum_{r}^{\mathcal{R}} U_r(t) \quad (9a)$$

s.t.

$$\sum_{m\in\mathcal{T}} \sum_{e\in\mathcal{E}} \alpha^e_{r,m}(t) \leq 1, \forall r, \quad (9b)$$

$$\alpha^e_{r,m}(t) \leq m_r(t), \quad \forall r, m \quad (9c)$$

$$\alpha^e_{r,m}(t) \leq 1 - cl_e(t), \quad \forall e, \quad (9d)$$

$$gh_r^{min}(t) \leq \alpha^e_{r,m}(t) \cdot h_r(t) \leq gh_r^{max}(t), \forall r, \forall m, \quad (9e)$$

$$\sum_m \sum_r \alpha^e_{r,m}(t) \leq cp_e(t), \forall e, \quad (9f)$$

$$\sum_m \sum_t \alpha^e_{r,m}(t) \leq t_r(t), \forall r. \quad (9g)$$

The optimization problem described in (9) is inherently challenging due to its dynamic nature, which necessitates continuous adjustments to account for changing conditions, and the inclusion of binary decision variables [30], [36]. Traditional optimization techniques struggle with these complexities, often resulting in exponential time complexity and impractical computational demands. To overcome these challenges, we propose leveraging DRL. Unlike traditional methods, DRL is well-suited for dynamic environments, as it can adapt and learn from real-time interactions. DRL agents excel in efficiently navigating complex solution spaces, making them ideal for such optimizing the problem (9). Therefore, we have chosen a DRL based approach (Deep Q-Networks (DQN)) due to their ability to handle complex decision spaces, adapt to non-linear dynamics, and offer practical benefits like computational efficiency and scalability. These algorithms are well-suited for the dynamic and complex nature of urban transportation systems and solving such problems as presented in (9). In the following section, we introduce our solution.

IV. DRL-based Solution Approach

This section outlines our approach to finding the optimal transportation mode ($\boldsymbol{\alpha}$) that aligns with user preferences and minimizes environmental emissions using a DRL agent. The agent operates from a Road Side Unit (RSU), acting as an edge server, collecting essential data such as the user's current location, remaining distance, travel time, health goals, and congestion levels. This data constructs system states, guiding the agent in selecting optimal actions ($\boldsymbol{\alpha}^*$) communicated to users. However, learning the optimal policy for the RL agent can be challenging, particularly when dealing with a vast state space (as per (9)). Therefore, we utilize DRL to address these challenges effectively [30]. In summary, our aim is to aid commuter in a smart city to plan their daily journey to work. Our DRL-based system can be used to determine the most optimal route, not only considering the shortest path but also factoring in real-time data from wearable devices and environmental sensors. By incorporating health data, such as heart rate and step count, the system can suggest routes that promote physical activity, like cycling through a park. Additionally, by considering real-time traffic and air quality information, the system can minimize exposure to pollution and traffic delays. This personalized approach ensures that the recommended route is both efficient and conducive to the user's well-being.

To this end, we model our problem in (9) as a Markov Decision Process (MDP), which involves defining the state space, action space, transition probabilities, and the reward function. This framework helps in modeling the problem dynamics and devising an optimal decision-making strategy. In our model, the state space, denoted by $\mathcal{S}$, represents various elements at a specific time t. Each state $s_t \in \mathcal{S}$ includes the user's current location $u_p(t)$, the user's residual deadline $D_r(t) = t_r(t) - d_r(t)$ (where $d_r(t)$ is the time consumed during the travel), the capacity of the route $cp_e(t)$, the capacity of the mode $cp_m(t)$, the congestion level $cl_e(t)$, and health-related parameters related to heart monitor $h_r(t)$ and calories consumed $c_r(t)$. The state vector at time t is thus represented by:$s_t = \{u_p(t), D_r(t), cp_e(t), cp_m(t), cl_e(t), c_r(t), h_r(t)\}$.

Next, the action space involves the agent selecting actions at each time step, transitioning to the next state until reaching the terminal state. Actions involve selecting the transportation modes ($\boldsymbol{\alpha}$) for traveling. Formally, $a_t^r = \{\boldsymbol{\alpha}(t)\}$, represents a specific action taken by a resident r at time t. The goal is to find a satisfactory policy $\boldsymbol{\pi} : \mathcal{S} \to \mathcal{A}$ that guides the agent's decision-making process, considering mode preferences. The state transition probability defines the likelihood of moving from one state to another based on the actions taken by the agent. It captures the environment's dynamics and determines how the system evolves. The state transition probability is given by: $\mathbf{Pr}\{\boldsymbol{s}_{t+1}|\boldsymbol{s}_t, \boldsymbol{\pi}(\boldsymbol{s}_t)\}$ which defines how the system state s_t evolves to s_{t+1} under the policy π. Finally, each of these transitions results in a reward that provides feedback to the agent about the quality of its actions. Formally, the reward function $R : S * A \to R$, where $R(s, a)$ determines the immediate reward received for taking a specific action in a given state at each time step. In our case, rewards are designed to encourage sustainable transportation choices, promote active transportation, and reduce environmental emissions. The reward function balances the competing objectives of minimizing emissions and maximizing health benefits. It is a weighted sum of two components: one for emissions and one for health benefits. The weights are dynamically adjusted based on real-time data to prioritize either environmental

Algorithm 1 : DRL Algorithm for Sustainable Transportation

Initialize: Set of residents $\mathcal{R}$, modes $\mathcal{T}$, Edges $\mathcal{E}$, Vertices $\mathcal{V}$ attribute tuples for route R_e and R_v, modes R_m, residents R_r;
Hyperparameters: ζ, η, ε_t, ϵ^{Max}, ϵ^{Min}, ϵ^{d} ;
Replay Memory:E_t;
Learning:
Set $\epsilon = \epsilon_{max}$;
Reset;
$t \leftarrow 0$;
repeat
  Observe $\boldsymbol{s}_t$;
  Select ϵ = Random(0,1);
  if $\epsilon < \varepsilon_t$ **then**
    Random action a_t selected;
  else
    $a_t = \arg\max_a Q(s_t|a;\theta)$;
    a_t selected;
  end if
  Action a_t taken, Reward $R(s_t, a_t)$ observed, and transit to next state s_{t+1};
  Store transition $(s_t, a_t, R(s_t, a_t), s_{t+1})$ in buffer E_t;
  Randomly sample minibatch of size J from E_t;
  for $(s_i, a_i, R_i, s'_i) \in \tilde{E}(t)$ **do**
    $a'_i = \arg\max_a Q(s'_i, a;\theta)$;
    Calculate loss $L(\theta_Q)$ via (12) ;
  end for
  Update ϵ periodically as follows:
  $t \leftarrow t + 1$;
  $\epsilon = \epsilon^{\text{Min}} + (\epsilon^{\text{Max}} - \epsilon^{\text{Min}})/\exp(-\epsilon^{d} t)$;
until Satisfy stop criteria, $\forall r \in \mathcal{R}$;
Output: Optimal policy $\boldsymbol{\pi}^*$

sustainability or user well-being. This approach ensures that our optimization model effectively addresses both objectives. Specifically, rewards are directly proportional to our utility $U_r(t)$, reflecting the goal of minimizing environmental emissions and maximizing the calorie count by active transportation. Additionally, a penalty p is subtracted from the reward if the chosen action set fails to meet the time deadline.

Once the MDP is defined, the RL agent's goal is to find an optimal policy, denoted as $\boldsymbol{\pi}$, which maximizes the expected rewards from the initial state. In our problem, this expected reward R represents the anticipated reduction in environmental emissions and promotion of active travel mode for a given state s, based on the action dictated by the policy $\boldsymbol{\pi}(s_t)$ for each resident r. For each policy π, this is represented by a state value function $V_\pi(s)$. Then the optimal policy is found by maximizing this expected value across all possible policies π to get the optimal state value $V_{\pi^*}(s)$. In our problem, presented in (9), since we maximize the utility function, then our reward is directly proportional to this function.

$$V_{\pi^*}(s) = \arg\max_{\boldsymbol{\pi}} \left[\mathbb{E}_{\boldsymbol{\pi}} \left((1-\zeta) \sum_{t=1}^{\Delta T} (\zeta)^{t-1} R(\boldsymbol{s}_t, \boldsymbol{\pi}(\boldsymbol{s}_t) | \boldsymbol{s}_1 \right) \right]. \tag{10}$$

Here, $R(s_t, \boldsymbol{\pi}(s_t))$ represents the expected reward at time t under policy π. The term $(\zeta)^{t-1}$ discounts future rewards, and $(1-\zeta)$ scales the immediate rewards. The objective is to maximize the discounted sum of expected rewards over a finite time horizon ΔT, exploring various policies π to find the optimal policy π^* that maximizes $V_{\pi^*}(s)$.

To find the optimal state-value function $V_{\pi^*}(s)$, we employ Bellman's optimality equation, focusing on maximizing the action-value function $Q(s,a) : S \times A \to \mathbb{R}$. This approach requires comprehensive network statistics for effective solutions. We employ Q-learning in our approach, a model-free method that addresses this by iterative update of the Q-function at each time step t based on interactions with the environment: current state, action, and subsequent state. Given the challenges of updating the Q-tables directly for large problems, we adopt the deep Q-network (DQN) approach. In DQN, the action-value function $Q(s,a)$ is approximated as $Q(s,a) \approx Q(s,a;\theta)$, where θ denotes neural network parameters updated at each time epoch t. Formally, this is presented by:

$$\begin{aligned} Q_t(s,a;\theta) \leftarrow Q_t(s,a;\theta) + \lambda_t \Big[&(1-\zeta) R_t(s,a;\theta) \\ &+ \zeta \max_{a \in \mathcal{A}} Q_{t+1}(s,a;\theta) - Q_t(s,a;\theta) \Big], \end{aligned} \tag{11}$$

where $\lambda_t \in [0,1]$ is a time-varying learning rate. With such an approach we can solve the issue of requiring comprehensive network statistics, However, ensuring stable learning remains a significant challenge in this approach. Stability in learning is crucially managed through a replay buffer in our approach. Replay buffer is a crucial component of our DQN-based approach. The replay buffer stores past experiences, which are randomly sampled during training to break temporal correlations and improve learning stability. This technique offers several benefits, including reduced variance in updates, prevention of overfitting, and accelerated convergence [37]. By leveraging the replay buffer, our model can more effectively learn from diverse traffic scenarios and user behaviors, leading to robust decision-making and policy optimization in urban transportation systems. Then, we store transition represented by a tuple $(\boldsymbol{s}_t,\ \boldsymbol{a}_t,\ E_t(\boldsymbol{s}_t),\ \boldsymbol{s}_{t+1})$ post-epoch to make our solution stable. Then, after each epoch t, the edge server samples random experiences of a size $\tilde{E}$ from the replay buffer of size E to create a mini-batch of size J. This mini-batch is then used to train the neural network and minimize the loss function as follows:

$$\begin{aligned} \mathcal{L}(\boldsymbol{\theta_t}) = \mathbb{E}_{\tilde{E}_t} \Big[&((1-\zeta) R_t(\boldsymbol{s}, \boldsymbol{a}; \boldsymbol{\theta}) + \\ &\zeta \max_{\boldsymbol{a} \in \mathcal{A}} Q_{t+1}(\boldsymbol{s}, \boldsymbol{a}; \boldsymbol{\theta})) - Q_{t+1}(\mathbf{s}, \mathbf{a}; \boldsymbol{\theta_{t+1}}) \Big]^2. \end{aligned} \tag{12}$$

Algo. 1 aims to optimize urban transportation systems by leveraging DRL techniques. The goal is to find an optimal policy that balances environmental sustainability and user health.

The algorithm begins by initializing key parameters and neural network weights, including the set of residents, modes of transportation, edges, vertices, and attribute tuples for routes, modes, and residents. Hyperparameters are also set to control the learning process. A replay memory buffer is created to store past experiences (lines 1-3). The learning phase then commences. An exploration rate is initialized to encourage exploration of different actions. The environment is reset to the initial state, and the algorithm iterates through episodes, representing individual users' journeys (lines 4-7). In each episode, the algorithm starts by selecting the initial state s_0 from the environment and proceeds through epochs, representing users' travel, and an action a_t is selected for the current state s_t using an ϵ-greedy policy that balances exploration and exploitation. The selected action is executed in the environment, resulting in a reward and the transition to the next state. This experience is stored in the replay memory (lines 8-18). Periodically, a minibatch of experiences is randomly sampled from the replay memory and for each experience, the best action a_i' for the next state s_i' is determined. For each experience, the target Q-value is calculated using the current policy, and the loss function between the predicted and target Q-values is computed. The neural network parameters are then updated using gradient descent to minimize the loss (lines 19-23). The exploration rate is gradually decreased over time to encourage exploitation of the learned policy. This process continues until the stopping criteria are met (lines 24-27). Note that the current state s_t transitions to s_{t+1} until stopping conditions are met, such as all users completing their journeys or constraints like travel time, health goals, or infrastructure capacities are violated or exhausted. This iterative process converges to an optimal policy that maximizes long-term rewards and promotes active traveling while reducing environmental emissions for all users $r \in \mathcal{R}$. The final learned policy represents the optimal strategy or policy (line 28).

Next, we discuss the practical implementation issues pertaining to our algorithm. The time complexity grows quadratically with the number of states, and the space complexity is influenced by the replay buffer size and number of hidden units. To address the potential issue of slow training times, we propose a pre-training approach. The model can be initially trained offline on a centralized server with a large number of users, similar to the techniques described in [30]. This pre-trained model can then be adapted for smaller-scale deployments with fewer users by setting extra parameters related to states and actions to null. This approach can reduce the complexity of the proposed approach and improve its scalability.

V. Simulation Results

In this section, we present the results from simulations evaluating the performance of our proposed DRL-based scheme. We start by describing the simulation environment and then proceed to analyze the scheme's effectiveness.

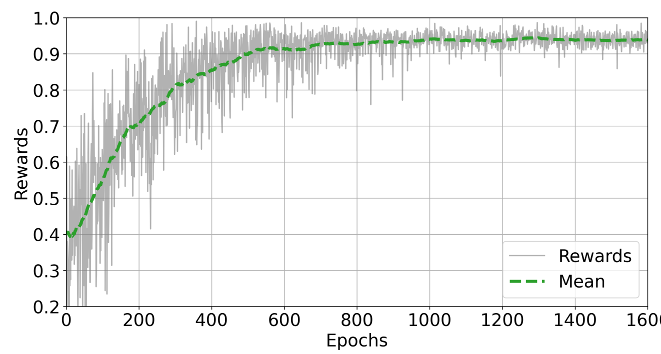
The simulation covers a service area of up to 30 kilometers, where an edge server collects and disseminates real-time information for optimal route planning. Four different travel modes are included, each with distinct attributes that influence user decisions. Mode-specific paths consider factors such as length, maximum capacity, and congestion level, all of which affect travel times. Users are uniformly distributed across random vertices using a Homogeneous Poisson Point Process (HPPP) to ensure equitable distribution. Each user profile includes origin and destination locations, available travel modes, and health goals, particularly focusing on parameters such as heart rate, age, weight, and height, along with travel deadlines. Users also monitor calorie expenditure and heart

TABLE II: Simulation Parameters

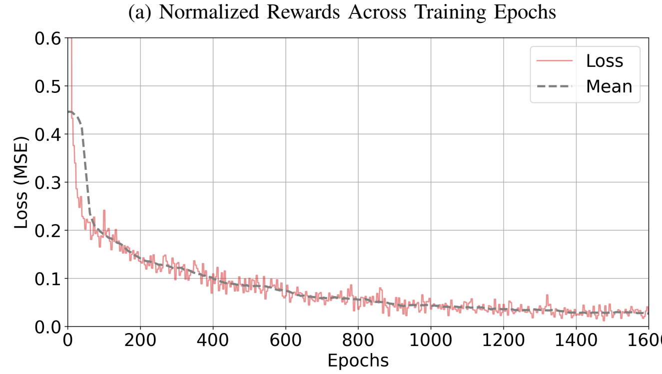
Parameter	**Value**
Service Area	$5 \sim 20$ km
Deadline (t_r)	$10 \sim 30$ min
Number of stops (v_r)	10
Number of bus stops (v_p)	5
Number of users (R)	$20 \sim 100$
Mode emissions (e_m) gCO2 per km	{Car: 160.61; Public Transport: 82; Cycle: 21; Walk: 0}[32]

TABLE III: Hyper-parameters

Parameter	**Value**
Learning Rate (γ)	10^{-3}
Batch Size ($\tilde{E}_t$)	256
Memory Size (E_t)	10^5
Discount Factor (ζ)	0.9
Epsilon Max ($\epsilon^{\max}$)	1
Epsilon Decay (ϵ^{d})	0.0995
Epsilon Minimum ($\epsilon^{\min}$)	0.0001

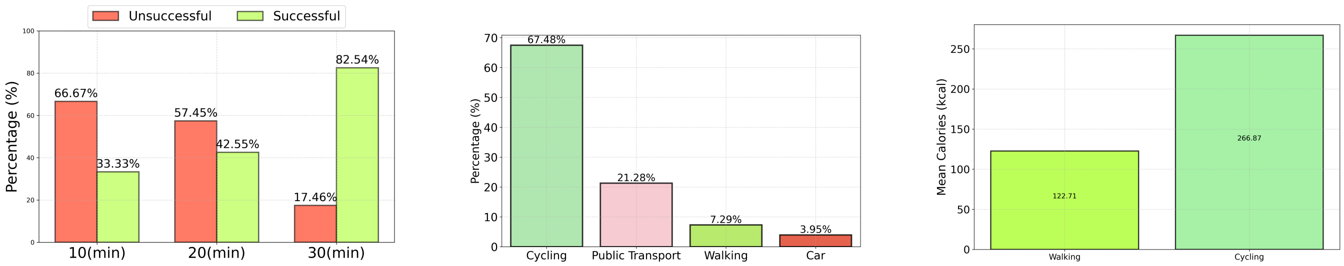


(a) Normalized Rewards Across Training Epochs



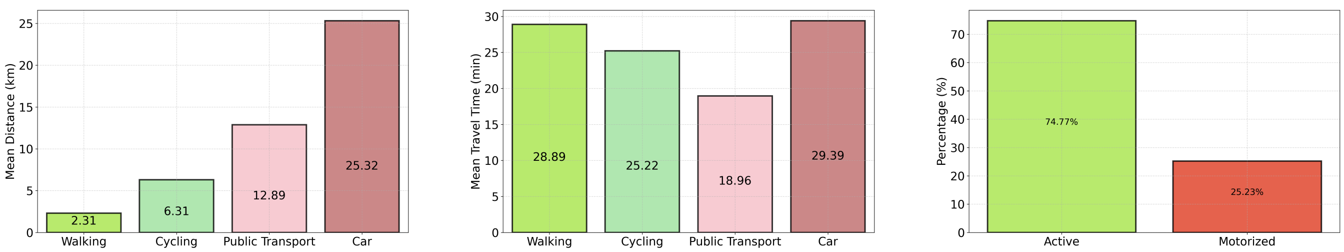
(b) Normalized Loss Across Training Epochs

Fig. 2: Convergence Evaluation of the Proposed Scheme



(a) Journey Completion Ratio. (b) Selected Modes. (c) Mean Calories Expenditure.

Fig. 3: System Performance: Journey Completion Ratio, Transport Modes, and Mean Calories Expenditure.



(a) Mean Distance. (b) Mean Travel Time. (c) Active vs Motorized Journeys.

Fig. 4: System Performance for Travel Modes.

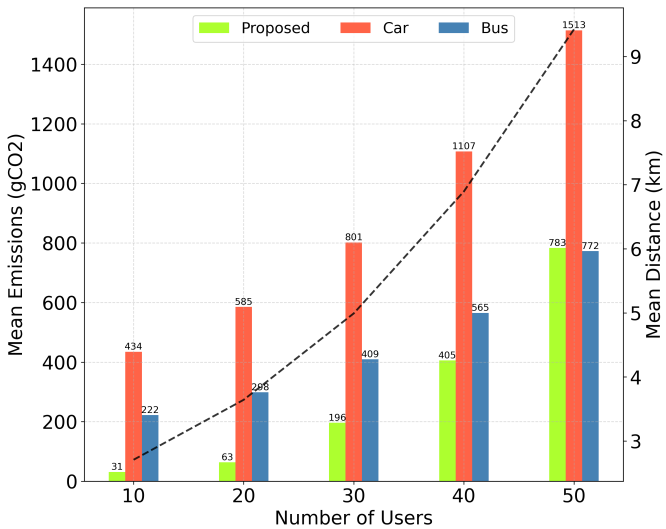
rate using SWH devices[5]. Our approach utilizes key health metrics like heart rate variability, step count, active minutes, to estimate the calories burned and assess the health benefits of different transportation modes as reflected in the reward function. Moreover, we use the Metabolic Equivalent Task (MET) method to estimate calorie expenditure based on these metrics and individual factors. The MET method is used to gauge physical activity intensity. MET values represent the energy cost of various activities compared to resting oxygen uptake. To calculate this, we need to determine the MET values for walking and cycling modes as we have all the other values such as weight. Randomly placed bus stops along some vertices with defined maximum capacities are scattered throughout the network. Buses are scheduled to service these stops every minute, accommodating passengers heading to various destinations with a capacity of 40 residents.

Table II presents our main simulation parameters. Similarly, our DQN network architecture uses three fully connected layers with Rectified Linear Unit (ReLU) activation functions in the hidden layers, each containing 128 neurons. We use Python 3.8 (Anaconda3) with TensorFlow 2.15 and PyTorch 2.4 to implement our DRL based algorithm. We numerically evaluate the results on a physical device with the following specifications: Intel(R) Core(TM) i5-4690 CPU 3.50 (GHz), RAM 32.0(GB), GPU GTX 1060 3(GB). The main hyperparameters and values used for this simulation are presented in Table III.

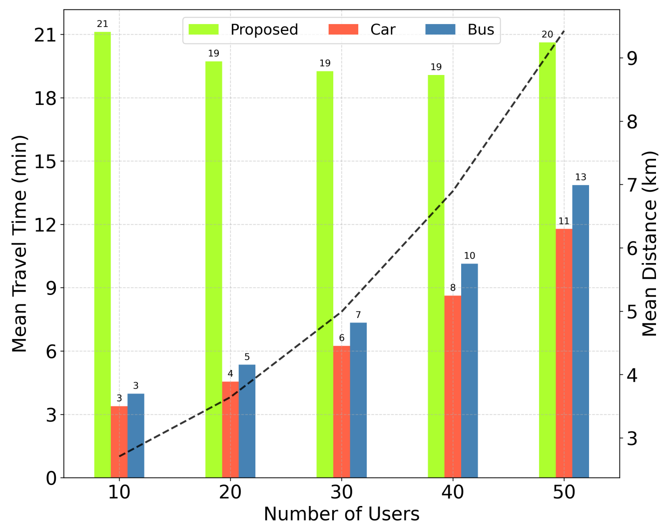
In this simulation study, we examined the convergence of the proposed DRL-based scheme over 1600 epochs for 250 residents, each with unique characteristics. Residents' origins and destinations were uniformly distributed within a 30 km area, across a set of 10 vertices, with 5 designated as bus stops. Each user had a random deadline between 10 to 30 minutes, connected to the edge server via SWH devices. Fig. 2 shows the convergence of the reward and loss functions, with convergence achieved after approximately 850 epochs. Initially, normalized rewards are low (Fig.2a) due to the exploration phase. As episodes progress, policy refinement occurs through network weight updates, reducing the loss function (Fig.2b) and leading to more environmentally and health-friendly traveling modes that meet user deadlines. These results confirm the effectiveness of our approach in achieving a stable policy.

Fig. 3 presents the system performance results, highlighting residents who completed their journeys within deadlines while meeting health requirements, the transport modes used, and the average calories burned. This simulation maintains the same parameters as previous ones but varies user deadlines across three intervals between 10 and 30 minutes. Fig. 3a shows that stricter deadlines (10 minutes) resulted in a lower completion rate (33.33%), while more lenient deadlines (30 minutes) significantly increased successful journeys (82%). With tight deadlines, the DRL agent prioritizes faster modes to meet deadlines and avoid penalties. In contrast, looser deadlines enable a more balanced approach, allowing the agent to choose environmentally friendly options that still meet the relaxed deadlines. For the same simulation, Fig. 3b examines the transportation mode usage among residents who completed their journeys. Cycling is the most popular option (67%), likely due to its efficiency in meeting tighter deadlines. Public transport follows as the second most used mode, emphasizing the necessity of faster travel to adhere to deadlines. This trend supports the earlier observation that strict deadlines prioritize faster modes (cycling and public transport) over slower or less

[5]We would assume this value would be sent to the edge server to track the heart rate and calories for each user to identify the physical activity intensity.



(a) Mean Emissions.



(b) Mean Travel Time.

Fig. 5: System Performance for Number of Users and Mean Distance.

environmentally friendly options like walking or cars. Next, Fig. 3c calculates the mean calories consumed by residents engaging in active journeys, i.e., walking and cycling. We estimated calories per minute using a basic formula, assuming all residents are in the same age group and assigning random weights between 60 and 85 kg. Travel time is calculated via the speed of the mode and the distance between the origin and destination for each resident. MET values for walking ranged from 2.5 to 4 MET (from relaxed to brisk walking) and cycling from 6.5 to 8 MET (based on different cycling speeds associated with the mode cycling). While real data from residents' SWH devices would yield more accurate results, the chosen MET values are well-established in research and accurately predict calorie expenditure [38]. The results indicate that, on average, cycling burns 266 kilocalories (kcal) and walking burns 122 kcal. This higher calorie count for cycling is due to its higher MET value and the preference for cycling to meet deadline constraints encouraging a large number of residents to choose this mode. Our approach promotes modes that maximize calorie consumption and minimize environmental emissions, explaining the popularity of cycling (Fig. 3b).

Fig. 4 demonstrates the positive impact of our DRL-based approach by analyzing mean travel distances, mean travel times, and the distribution of active versus motorized journeys. In Fig. 4a, the mean distances for each mode are: Walking (2.31 km), Cycling (6.31 km), Public Transport (12.89 km), and Car (25.32 km). These values show how deadlines and residents' requirements influence transport modes. Users choose modes based on both distance and deadlines, optimizing efficiency and sustainability. Walking and cycling are preferred for shorter distances, promoting health and environmental benefits. Fig. 4b provides insights into mean travel times across different modes. The notable use of walking and cycling suggests that our approach effectively promotes these active modes over motorized options like public transport and cars. This shift reduces reliance on motor vehicles, yielding multiple benefits: decreased emissions, lower air and noise pollution, and a cleaner, healthier urban environment. Next, Fig. 4c quantifies the sustainability impact, showing that up to 74% of journeys are categorized as active (walking and cycling). This demonstrates the effectiveness of our scheme in encouraging sustainable transportation choices, which alleviate strain on road infrastructure and potentially reduce maintenance and expansion costs. Overall, our approach aligns with the net-zero initiative by promoting active and healthy travels, which balance out the use of motorized transportation resulting in reducing greenhouse gas emissions, and lowering the urban carbon footprint. Encouraging walking and cycling for shorter distances directly supports climate change mitigation and sustainable transport systems.

Fig. 5 analyzes our proposed scheme's performance across varying user numbers and travel distances, evaluating mean environmental emissions (gCO2) and mean travel time (min) with user counts ranging from 10 to 50 and average distances from 1 km to 10 km. Environmental emissions benchmarks are from [32]. We compare our approach to two baselines: "Car" (all travel by car) and "Bus" (all travel by public transport). Fig. 5a shows that environmental emissions increase with user numbers and travel distances. Our proposed approach outperforms both baselines for lower user counts and shorter distances by favoring modes with lower emissions as they promote active modes of travel. However, as user numbers and distances increase, the DRL agent sometimes opts for faster modes like public transport or cars to meet the 30-minute deadline, raising emissions. At 9 km and 50 users, there is a significant shift towards public transport and cars, causing our results to align more closely with the "Bus" baseline. However, our results remain significantly better than the "Car" baseline, as some users continue to choose cycling for these distances if the travel time threshold allows. Next, Fig. 5b illustrates mean travel time. Our approach, while not the fastest, ensures all journeys meet the 30-minute deadline, with slightly longer times often linked to walking and cycling, promoting active lifestyles and sustainability. This strategy enhances the city's transportation system's sustainability and supports the health initiative.

VI. Conclusions

This paper introduces an innovative DRL-based approach to urban transportation systems, designed to balance user health and environmental sustainability. Our DRL method adeptly addresses the dynamic nature of transportation systems by learning an efficient DRL policy for this dynamic problem, outperforming baseline scenarios focused solely on the fastest travel times. Simulation results highlight the significant advantages of our proposed approach, demonstrating substantial reductions in environmental emissions and promoting active travel modes such as walking and cycling over motorized modes whenever possible, enhancing the health and well-being of residents, and encouraging a healthier lifestyle. In future work, we plan to integrate additional healthcare and smart wearable devices (SWH) to capture real-time data, allowing us to predict the optimal travel modes for each journey. We will also consider environmental metrics such as air quality (PM2.5, PM10, NOx, etc.), which are monitored by modern SWH devices, to further enhance our model's environmental impact assessment. Furthermore, we aim to explore multimodal solutions, enabling the combination of different transport modes within a single trip. Additionally, while our current model assumes a uniform edge computing environment for simplicity, future work will address the varying properties of different edges to reflect practical settings more accurately. We also intend to apply our approach at the city level to address practical considerations, helping us design more effective and sustainable transportation solutions for residents.

References

[1] U. N. D. of Economic and P. D. Social Affaris, 2021. [Online]. Available: https://www.un.org/development/desa/pd/content/global-population-growth

[2] T. M. Ho, T. D. Tran, T. T. Nguyen, S. Kazmi, L. B. Le, C. S. Hong, and L. Hanzo, "Next-generation wireless solutions for the smart factory, smart vehicles, the smart grid and smart cities," *arXiv preprint arXiv:1907.10102*, 2019.

[3] G. Dhiman and N. S. Alghamdi, "Smose: Artificial intelligence-based smart city framework using multi-objective and iot approach for consumer electronics application," *IEEE Transactions on Consumer Electronics*, vol. 70, no. 1, pp. 3848–3855, 2024.

[4] K. A. B. Ahmad, H. Khujamatov, N. Akhmedov, M. Y. Bajuri, M. N. Ahmad, and A. Ahmadian, "Emerging trends and evolutions for smart city healthcare systems," *Sustainable Cities and Society*, vol. 80, p. 103695, 2022.

[5] A. Neves and C. Brand, "Assessing the potential for carbon emissions savings from replacing short car trips with walking and cycling using a mixed gps-travel diary approach," *Transportation Research Part A: Policy and Practice*, vol. 123, pp. 130–146, 2019.

[6] P. Liu, Z. Chen, M. Liu, C. Piao, K. Wan, and H. Huang, "Vertical parking trajectory planning with the combination of numerical optimization method and gradient lifting decision tree," *IEEE Transactions on Consumer Electronics*, vol. 70, no. 1, pp. 1845–1856, 2024.

[7] C. Liu and K. Liu, "Toward reliable dnn-based task partitioning and offloading in vehicular edge computing," *IEEE Transactions on Consumer Electronics*, vol. 70, no. 1, pp. 3349–3360, 2024.

[8] H. Yan, X. Xu, M. Bilal, X. Xia, W. Dou, and H. Wang, "Customer centric service caching for intelligent cyber–physical transportation systems with cloud–edge computing leveraging digital twins," *IEEE Transactions on Consumer Electronics*, vol. 70, no. 1, pp. 1787–1797, 2024.

[9] I. Yaqoob, L. U. Khan, S. A. Kazmi, M. Imran, N. Guizani, and C. S. Hong, "Autonomous driving cars in smart cities: Recent advances, requirements, and challenges," *IEEE Network*, vol. 34, no. 1, pp. 174–181, 2019.

[10] D. Sperling, *Future drive: Electric vehicles and sustainable transportation*. Island Press, 2013.

[11] S. A. Kazmi, T. N. Dang, I. Yaqoob, A. Manzoor, R. Hussain, A. Khan, C. S. Hong, and K. Salah, "A novel contract theory-based incentive mechanism for cooperative task-offloading in electrical vehicular networks," *IEEE Transactions on Intelligent Transportation Systems*, vol. 23, no. 7, pp. 8380–8395, 2021.

[12] T. Zhang, X. Zhou, J. Liu, B. Cheng, X. Xu, L. Qi, Q. Tian, and Z. Wan, "Qoe-driven data communication framework for consumer electronics in tele-healthcare system," *IEEE Transactions on Consumer Electronics*, vol. 69, no. 4, pp. 719–733, 2023.

[13] N. Cusack, P. Venkatraman, U. Raza, and A. Faisal, "Smart wearable sensors for health and lifestyle monitoring: Commercial and emerging solutions," *ECS Sensors Plus*, vol. 3, no. 1, p. 017001, 2024.

[14] C. Brand, E. Dons, E. Anaya-Boig, I. Avila-Palencia, A. Clark, A. de Nazelle, M. Gascon, M. Gaupp-Berghausen, R. Gerike, T. Götschi *et al.*, "The climate change mitigation effects of daily active travel in cities," *Transportation Research Part D: Transport and Environment*, vol. 93, p. 102764, 2021.

[15] S. J. Davis, N. S. Lewis, M. Shaner, S. Aggarwal, D. Arent, I. L. Azevedo, S. M. Benson, T. Bradley, J. Brouwer, Y.-M. Chiang *et al.*, "Net-zero emissions energy systems," *Science*, vol. 360, no. 6396, p. eaas9793, 2018.

[16] S. A. Kazmi, T. N. Dang, I. Yaqoob, A. Ndikumana, E. Ahmed, R. Hussain, and C. S. Hong, "Infotainment enabled smart cars: A joint communication, caching, and computation approach," *IEEE Transactions on Vehicular Technology*, vol. 68, no. 9, pp. 8408–8420, 2019.

[17] H. Muslih, S. A. Kazmi, M. Mazzara, and G. Baye, "Cache sharing in uav-enabled cellular network: A deep reinforcement learning-based approach," *IEEE Access*, vol. 12, pp. 43 422–43 435, 2024.

[18] A. De Nazelle, M. J. Nieuwenhuijsen, J. M. Antó, M. Brauer, D. Briggs, C. Braun-Fahrlander, N. Cavill, A. R. Cooper, H. Desqueyroux, S. Fruin *et al.*, "Improving health through policies that promote active travel: a review of evidence to support integrated health impact assessment," *Environment international*, vol. 37, no. 4, pp. 766–777, 2011.

[19] L. D. Frank, M. J. Greenwald, S. Winkelman, J. Chapman, and S. Kavage, "Carbonless footprints: promoting health and climate stabilization through active transportation," *Preventive medicine*, vol. 50, pp. S99–S105, 2010.

[20] P. F. Rodrigues, M. Alvim-Ferraz, F. Martins, P. Saldiva, T. Sá, and S. Sousa, "Health economic assessment of a shift to active transport," *Environmental pollution*, vol. 258, p. 113745, 2020.

[21] D. Esztergár-Kiss, Y. Shulha, A. Aba, and T. Tettamanti, "Promoting sustainable mode choice for commuting supported by persuasive strategies," *Sustainable Cities and Society*, vol. 74, p. 103264, 2021.

[22] G. Giallouros, P. Kouis, S. I. Papatheodorou, J. Woodcock, and M. Tainio, "The long-term impact of restricting cycling and walking during high air pollution days on all-cause mortality: Health impact assessment study," *Environment International*, vol. 140, p. 105679, 2020.

[23] L. Zhang, L. H. U, M. Zhou, and F. Yang, "Elastic tracking operation method for high-speed railway using deep reinforcement learning," *IEEE Transactions on Consumer Electronics*, vol. 70, no. 1, pp. 3384–3391, 2024.

[24] D. Chen, M. R. Hajidavalloo, Z. Li, K. Chen, Y. Wang, L. Jiang, and Y. Wang, "Deep multi-agent reinforcement learning for highway on-ramp merging in mixed traffic," *IEEE Transactions on Intelligent Transportation Systems*, 2023.

[25] A. Sarker, H. Shen, and K. Kowsari, "A data-driven reinforcement learning based multi-objective route recommendation system," in *2020 IEEE 17th international conference on mobile ad hoc and sensor systems (mass)*. IEEE, 2020, pp. 103–111.

[26] Z. Fu, L. Yu, and X. Niu, "Trace: Travel reinforcement recommendation based on location-aware context extraction," *ACM Transactions on Knowledge Discovery from Data (TKDD)*, vol. 16, no. 4, pp. 1–22, 2022.

[27] L. Chen, J. Cao, H. Tao, and J. Wu, "Trip reinforcement recommendation with graph-based representation learning," *ACM Transactions on Knowledge Discovery from Data*, vol. 17, no. 4, pp. 1–20, 2023.

[28] S. A. Kazmi, S. Otoum, R. Hussain, and H. T. Mouftah, "A novel deep reinforcement learning-based approach for task-offloading in vehicular networks," in *2021 IEEE Global Communications Conference (GLOBECOM)*. IEEE, 2021, pp. 1–6.

[29] K.-F. Chu and W. Guo, "Deep reinforcement learning of passenger behavior in multimodal journey planning with proportional fairness," *Neural Computing and Applications*, vol. 35, no. 27, pp. 20 221–20 240, 2023.

[30] S. A. Kazmi and et al., "Computing on wheels: A deep reinforcement learning-based approach," *IEEE Transactions on Intelligent Transportation Systems*, vol. 23, no. 11, pp. 22 535–22 548, 2022.

[31] C. Moreno, Z. Allam, D. Chabaud, C. Gall, and F. Pratlong, “Introducing the “15-minute city”: Sustainability, resilience and place identity in future post-pandemic cities,” *Smart cities*, vol. 4, no. 1, pp. 93–111, 2021.

[32] G. Salazar, J. P. Silva, and B. Ribeiro, “Faster, cheaper, cleaner: Assessing urban mobility,” in *2015 International Conference on Smart Cities and Green ICT Systems (SMARTGREENS)*. IEEE, 2015, pp. 1–6.

[33] M. B. Hariz, D. Said, and H. T. Mouftah, “Mobility traffic model based on combination of multiple transportation forms in the smart city,” in *2019 15th international wireless communications & mobile computing conference (IWCMC)*. IEEE, 2019, pp. 14–19.

[34] A. Lorenz, N. Madeja, and C. Leyh, “A framework for assessing the sustainability impact of intelligent transport systems in the smart city context,” in *2023 18th Conference on Computer Science and Intelligence Systems (FedCSIS)*. IEEE, 2023, pp. 161–169.

[35] Z. Khan, A. G. Abbasi, and Z. Pervez, “Blockchain and edge computing–based architecture for participatory smart city applications,” *Concurrency and Computation: Practice and Experience*, vol. 32, no. 12, p. e5566, 2020.

[36] S. M. A, N. H. Tran, W. Saad, Z. Han, T. M. Ho, T. Z. Oo, and C. S. Hong, “Mode selection and resource allocation in device-to-device communications: A matching game approach,” *IEEE Transactions on Mobile Computing*, vol. 16, no. 11, pp. 3126–3141, November 2017.

[37] C. Zhang, Y. Meng, and V. Prasanna, “A framework for mapping drl algorithms with prioritized replay buffer onto heterogeneous platforms,” *IEEE Transactions on Parallel and Distributed Systems*, vol. 34, no. 6, pp. 1816–1829, 2023.

[38] M. Jetté, K. Sidney, and G. Blümchen, “Metabolic equivalents (mets) in exercise testing, exercise prescription, and evaluation of functional capacity,” *Clinical cardiology*, vol. 13, no. 8, pp. 555–565, 1990.

S. M. Ahsan Kazmi is currently a Senior Lecturer in the Department of Computer Science and Creative Technologies, University of the West of England, Bristol. Prior to this, he worked as an Assistant Professor with the Institute of Information Security and Cyber-Physical System, Innopolis University, Russia for three years. He has also worked as a Postdoctoral fellow, from 2017 to 2018 at the Department of Computer Science and Engineering, Kyung Hee University, South Korea, from where he completed his Ph.D. degree, in 2017. His research interests include applying analytical techniques of optimization, machine learning, and game theory to radio resource management for future networks. He received the best KHU Thesis Award in engineering in 2017 and several best paper awards from prestigious conferences.

Zaheer Khan received the B.Sc. degree in computer science, the M.S. degree in information technology, and the Ph.D. degree in computer science. He has over 20 years of academic research experience. He is currently a Professor of computer science with the University of the West of England (UWE), Bristol, U.K. His primary research addresses intelligent and secure distributed systems, including developing and applying new software processes and methods for systems development by using modern technologies such as blockchain, the IoT, digital twin, cloud/edge computing, and big-data management and analysis. He has been applying his research on digital urban governance and sustainable smart future cities by engaging with many European cities through different international projects, since 2009. He also looks after the smart cities theme of the Computer Science Research Centre (CSRC), UWE. He has extensive experience working on several externally funded large-scale collaborative research and innovation projects.

Adil Khan is currently a seasoned Professor and a Prolific Researcher in machine learning. With a robust background in machine learning, deep learning, and representation learning, he is passionately dedicated to both pedagogy and innovative research in the realm of artificial intelligence. His research journey started in South Korea, in 2006, where he concentrated on human activity recognition through wearable sensors. His groundbreaking discoveries were published in reputable journals and employed by leading technology firms for their healthcare applications. Over his career, he has undertaken more than ten research projects, obtaining substantial funding, and has published in excess of 90 research articles. He has supervised more than two dozen Ph.D. and M.S. students to completion. His expertise and experience are not limited to a single geographic location. His academic career has spanned across various prestigious universities in South Korea, Denmark, Russia, United Arab Emirates, Switzerland, and U.K. These diverse collaborations and experiences have enriched his knowledge and cultural understanding, augmenting his holistic approach toward research and education.

Manuel Mazzara received his Ph.D. degree in Computer Science from the University of Bologna, Italy. He is currently a Professor of computer science with Innopolis University, Russia, with a research background in software engineering, service-oriented architecture, concurrency theory, formal methods, and software verification. He has published many relevant and highly cited papers, in particular in the field of service engineering and software architectures and has collaborated with European and U.S. industries, and governmental and inter-governmental organizations, such as the United Nations, always at the edge between science and software production. The work conducted by him and his team in recent years focuses on the development of theories, methods, tools, and programs covering the two major aspects of software engineering, such as the process side, related to how we develop software; and the product side, concerning the results of this process.

Asad Masood Khattak (Member, IEEE) received the M.S. degree in information technology from the National University of Sciences and Technology (NUST), Islamabad, Pakistan, in 2008, and the Ph.D. degree in computer engineering from Kyung Hee University, South Korea, in 2012. He is currently an Associate Professor with the College of Technological Innovation, Zayed University, Abu Dhabi, United Arab Emirates, in August 2014. He was a Postdoctoral Fellow with the Department of Computer Engineering, Kyung Hee University, South Korea, and later joined the College of Technological Innovation as an Assistant Professor. He is currently leading three research projects, collaborating in four research projects, and has successfully completed five research projects in the fields of data curation, context-aware computing, the IoT, and secure computing. He has authored/coauthored more than 120 journals and conference papers in highly reputed venues. He is serving as a reviewer, a program committee member, and a guest editor for many conferences and journals. He has delivered keynote speeches, invited talks, guest lectures, and has delivered short courses in many universities. He and his team have secured several national and international awards in different competitions.