

# Enhancing Pedestrian Route Choice Models through Maximum-Entropy Deep Inverse Reinforcement Learning with Individual Covariates (MEDIRL-IC)

Boyang Li and Wenjia Zhang

**Abstract**—Understanding pedestrian route choices is pivotal for deciphering individual behaviors and informing decisions in urban planning. Unlike motorized transportation, primarily influenced by the built environment of the origin and destination, pedestrian route choices are also shaped by individual characteristics. Traditional route choice models often employ multi-segment training based on these characteristics, limiting their ability to capture heterogeneity among pedestrians within a unified framework. Additionally, these models frequently rely on oversimplified utility assumptions and neglect the implications of Markov decision sequences, potentially resulting in a misinterpretation of genuine route choice preferences. Addressing these, we introduce the Maximum-Entropy Deep Inverse Reinforcement Learning with Individual Covariates (MEDIRL-IC) framework, which accounts for the stochastic nature of state transitions in Markov decision process. Building on MaxEnt-IRL, MEDIRL-IC uses deep neural networks to separately fit constant individual features and dynamic built environment features before integration, effectively capturing the non-linear reward function in pedestrian scenarios. We further enrich model interpretability using a graph-based causal discovery algorithm, offering insights into complex feature interactions. Validations on a pedestrian mobile signaling dataset affirm MEDIRL-IC’s superior balance between model fit and interpretability. Our contributions present urban planners with a robust analytical tool, facilitating the data-driven design of pedestrian-centric urban landscapes. For research transparency and reproducibility, all codes are accessible at: [https://github.com/BoyangL1/Advanced\\_DeepIRL](https://github.com/BoyangL1/Advanced_DeepIRL).

**Index Terms**—Deep inverse reinforcement learning, Pedestrian route choice, Causal discovery, Cell phone signaling

## I. INTRODUCTION

NON-MOTORIZED travel, such as walking and cycling, significantly enhances urban sustainability [1]. However, the decision-making process of pedestrian behaviors is often complicated [2], including strategic (e.g., by choosing destinations and activities), tactical (by arranging activities and selecting routes), and operational (real-time interactions, e.g., moving or waiting) decisions. Unlike the strategic level’s comprehensive perception of strategy in destination selection

This study was supported by the Natural Science Foundation of China (42171201), and the Shenzhen Municipal Natural Science Foundation (Key Project) (GXWD20201231165807007-20200810223326001). (Corresponding author: Wenjia Zhang.)

B. Li is with the Behavioral and Spatial AI Lab, Peking University Shenzhen Graduate School, Shenzhen, Guangdong 518055, China. (e-mail: liboyang@stu.pku.edu.cn)

W. Zhang is with the College of Architecture & Urban Planning, Tongji University, Shanghai 200092, China, and the Behavioral and Spatial AI Lab, Peking University Shenzhen Graduate School, Shenzhen, Guangdong 518055, China. (e-mail: wenjiazhang@tongji.edu.cn)

[3], and the wayfinding modeling at the operational level, which involves an incomplete perception of the environment [4]. This study focuses on the tactical decision of pedestrian route choice (PRC), which is a sequential decision-making process.

Distinct from driving route choices, which are primarily shaped by road factors and the tendency to choose the shortest path, PRCs prioritize routes that are not only the shortest but also provide more visually appealing sceneries [5]. The sequential decisions in PRC offer individuals more spatial freedom compared to motorized route choices [6]. This freedom is influenced by a multitude of factors including land use adjacent to the road, functional characteristics of the location, personal factors such as age, gender, and individual preferences, as well as intrinsic attributes of the route such as its length and accessibility [7], [8]. Such complexity necessitates the development of models that comprehensively account for this wide array of influences.

In PRC research, traditional discrete choice models (DCM) often fall short due to oversimplified assumptions like linear utility functions and the omission of Markov decision processes [9], [10], leading to inaccuracies in estimating real route preferences. Recognizing these limitations, some studies have pivoted to dynamic discrete choice models (DDCM) [11], enhancing the representation of sequential decision-making [12]. However, DDCMs still face challenges due to increased computational loads. Our research accounts for the variability in individual characteristics and the subtle changes in the built environment along the trajectory. This introduces increased computational complexity within a high-dimensional feature space, making the application of traditional DDCMs less practical.

Given these challenges, researchers have explored deep learning methods for their computational efficiency [13], [14]. Yet, the “black-box” nature of deep learning limits interpretability compared to DCMs, where feature weights directly indicate the importance of decision-making factors. While global interpretation algorithms have been employed to elucidate feature contributions [15], [16], they generally fail to address the complexity of inter-variable relationships. Additionally, while some studies have applied IRL to model route choice issues [17], [18], most of them focus only on the influence of external environmental factors and historical trajectories on route choices, neglecting the role of individual characteristics. This oversight reflects a shortfall in current research, failing to account for decision-maker heterogeneity,

which is crucial in DCM's utility approach that accounts for both choice sets and individual variances [19].

To address the methodological challenges, this research develops a modeling framework called Maximum-Entropy Deep Inverse Reinforcement Learning with individual covariates (MEDIRL-IC). By integrating Markov decision processes and deep neural networks, MEDIRL-IC captures the complex non-linear relationships between multiple factors and individual path preferences, effectively identifying nuanced reward functions. This method stands out by integrating individual characteristics into the IRL framework, marking a novel advancement in PRC analysis. Unlike built environment features, these individual characteristics remain constant across decision states, our study mathematically demonstrates that incorporating these individual characteristics is equivalent to adopting a data stratification strategy. Consequently, individual and built environment features require separate fittings before integration in the neural network architecture. Given the exponential expansion of the feature space with the addition of individual characteristics, our deep network design incorporates separate input and hidden layers for each feature set, merging them in subsequent layers to ensure effective and accurate modeling in high-dimensional feature spaces. Additionally, we integrate graph-based causal discovery algorithms to address the challenge of interpretability in deep learning models. By quantifying causal relationships between variables and utilizing intuitive causal graphs, we can illustrate the specific factors that play pivotal roles in PRC and demonstrate how these factors interact with each other. This approach not only reveals the direct influences on PRC but also uncovers the indirect effects, offering a comprehensive understanding of the multifaceted dynamics involved in pedestrian decision-making processes.

The primary contributions of this paper can be summarized as follows:

- We introduced a maximum-entropy deep inverse reinforcement learning framework called MEDIRL-IC that incorporates individual traveler's characteristics and provided mathematical proofs for the validity of this framework, demonstrating its effectiveness in modeling and understanding individuals' pedestrian route choices.
- We applied a graph-based causal discovery model to interpret the results of deep neural networks. It captures the intricate interactions among high-dimensional features and provides an intuitive visual representation, allowing for a deeper understanding of the factors nonlinearly influencing individuals' PRC decisions.
- Using a large dataset of individual-level pedestrian route trajectories with travelers' socioeconomic features, we validated the proposed framework and compared it with existing models (e.g., R-logit, DNN-PSL, MaxEnt-DIRL), demonstrating that MEDIRL-IC strikes a balance between fitting capability and interpretability, making it a valuable tool for analyzing and predicting individual slow-paced trajectories.

## II. RELATED WORK

### A. Route Choice Decision Modeling

PRC modeling is mainly divided into two approaches [20]: the pattern recognition method, which utilizes deep learning to analyze past pedestrian trajectories for future route prediction, and the behavioral decision-making method, which employs discrete choice models to simulate decision-making processes by building utility functions based on the assumption of rational agents.

The advance in computing power have bolstered the impact of deep learning algorithms on the mission of trajectory prediction via pattern recognition. The introduction of the LSTM model in 2016 [21] for predicting pedestrian trajectories in various environments marked a key development. Notably, research employing temporal neural networks with attention mechanisms has successfully captured sequential travel patterns, leading to personalized routing recommendations for specific Origin-Destination (OD) conditions [14], [22]. Despite their effectiveness in sequence modeling, these deep learning approaches often omit the Markov decision process's complexities, limiting depth in behavioral insights. Moreover, while these models offer a high degree of prediction accuracy, they tend to overlook the intricacies of individual preferences and the nuances of environmental interactions. This limitation makes it challenging to provide a basis for planning decisions.

On the other hand, utility theory has been fundamental in modeling travel decision behaviors, with route choice models primarily framed within the Discrete Choice Model (DCM) approach, including path-based and link-based models [23]. Path-based models, widely adopted for calculating utility values of different paths [24], [25], often struggle with choice set construction due to path overlaps, challenging the Independence of Irrelevant Alternatives (IIA) assumption [26]. To address this, various path generation techniques have been explored [27], [28]. Nevertheless, the parameters of path-based models are sensitive to the number of choice set, as the selection probability of each path is determined based on its attributes (such as time, cost, etc.). Variations in the number of available paths can influence these probabilities, potentially complicating and destabilizing parameter estimation. This issue has led to the development of the link-based recursive logit model [29].

The recursive logit model, a key implementation of DDCM in route choice, advances modeling by treating links as the basic units of analysis and employing the Markov decision process to simulate sequential route choices [10], [30]. It considers both the utility of the current link and the anticipated utility of potentially chosen future links. By recursively computing the choices of paths from the starting point to the destination, it offers a closer approximation to pedestrian decision-making processes and addressing the choice set generation issues present in path-based models [31]. Within the constrained environment modeling framework, several studies have adopted agent-based and recursive search methodologies in navigation graph modeling [4], [32].

Despite these advances, both models are limited by the need for predefined utility function forms that inadequately

capturing the nonlinear impact of environmental and individual characteristics on decision-making, and encounter difficulties with computational complexity and parameter estimation, particularly with large choice sets. These limitations suggest the need for innovative models that more accurately reflect the complex dynamics of pedestrian travel decisions.

### B. Inverse Reinforcement Learning for Route Choice Behavior Modeling

Reinforcement Learning (RL) is a method in machine learning designed for decision-making in sequences, operating within the Markov Decision Processes (MDP) framework [33]. Traditional RL, based on predefined reward functions, has been utilized in travel behavior analysis to deduce optimal routes [34]. However, its forward-learning nature, relying on a known reward function, has limitations, paving the way for the emergence of Inverse Reinforcement Learning (IRL).

IRL diverges from traditional RL by deriving implicit reward functions from observed behaviors rather than relying on predefined preferences [35], [36]. This approach has been extensively applied in travel research, notably by Hidaka et al. [16], who used IRL to analyze pedestrian space perceptions through POIs' appeal. Addressing the dual nature of the reward function, Ziebart et al. [37] introduced maximum entropy inverse reinforcement learning (MaxEnt-IRL) to model travel behavior, effectively capturing the uncertainties in human decision-making. Pan et al. [38] also utilized relative entropy IRL to study taxi drivers' evolving preferences, unveiling diverse strategies among drivers based on their experience. Concurrent research from various scholars suggests that the function form of DDCM aligns with maximum entropy IRL [11], [39]. This alignment bridges traditional economic decision modeling with contemporary machine learning techniques, offering a more sophisticated and adaptable framework for behavioral analysis.

While Maxent-IRL offers a notable advancement, it faces limitations in complex environments due to its predefined reward function and challenges in convergence within large state and feature spaces. Recognizing these issues, Wulfmeier et al. [40] integrated deep neural networks with MaxEnt-IRL, enhancing its adaptability and ensuring smoother convergence, thereby offering a refined tool for analyzing route choices. Liu et al. [41] incorporated the Dijkstra algorithm and neural networks in deep inverse reinforcement learning to understand food delivery riders' trajectories, leading to personalized route recommendations without needing road network data. They further employed graph attention networks to capture inter-segment interrelations, eventually achieving personalized route recommendations based on dynamic traffic conditions [18]. Additionally, Zhao and Liang integrated adversarial learning into the reinforcement learning framework, directly learning from users' travel strategies, pioneering a novel research direction for route choice models [17].

However, the models above frequently overlook the inherent uncertainties in route decisions, focusing primarily on road network factors while underestimating the critical impact of the surrounding built environment on slow-paced travel.

Furthermore, these models struggle with transforming distinct environmental features into comparable vector formats for neural network input, due to varied cardinalities and the insufficiency of simple normalization methods to reflect their relative importance. Additionally, the oversight of individual characteristics in modeling limits the ability to capture the heterogeneous nature of route choices, as many models resort to segmenting individuals into distinct clusters [18], [38], which inadequately integrates these variances into a unified model framework.

In response to these challenges, we introduce a novel model based on MaxEnt-DIRL, enhanced with an inverse learning mechanism and deep learning architecture, merging the strengths of traditional discrete choice analysis with deep neural networks. This model outperforms existed route choice models by better capturing uncertainties in route selection and deeply analyzing the influence of both environmental and individual factors on these choices.

### C. Causal Discovery for Interpretable Deep Learning

The fitting capabilities of deep IRL models surpass those of traditional DCMs when dealing with high-dimensional features, attributed to the complex parametric structure of deep neural networks. However, this complexity renders the predictions from such models difficult to interpret, necessitating the use of interpretation algorithms to decipher the "black box" nature of neural networks. Research has explored relevant algorithms aimed at clarifying how model outputs are influenced by input features. Certain researchers use t-SNE to reduce the dimensionality of input and output features for visualization [42], and perform statistical analysis on the input-output data of specific cluster categories. Inspired by Shapley values from cooperative game theory, the SHAP method helps understand the contribution degree of each input feature to the model output [43], thus revealing the model's logic. For instance, Simini et al. [15] utilized the SHAP method to analyze the impact of geographical factors in deep gravity models.

Deep IRL models, diverging from traditional DNN approaches, seek to uncover the latent reward mechanisms driving decision-making rather than directly modeling observable phenomena. This requires not only identifying the critical input features these models rely on but also understanding how these features interrelate, offering insights into the models' interpretation of complex decision-making processes. To address these complexities in decision-making, Sun et al. [44] leveraged MaxEnt-IRL to create interpretable visual cost maps, facilitating improved comprehension of model decisions from observations, while saliency maps have also been widely adopted for elucidating deep reinforcement learning models' decision processes [45].

To overcome limitations of interpretation algorithms in analyzing variable interactions, we employ graph-based causal discovery for model interpretation, representing data through graphs where nodes and edges signify variables and their causal influences respectively [46]. This method adeptly unravels complex interactions and causalities among features, offer-

ing intuitive graphical visualizations for easier understanding of causal structures [47].

In the realm of travel behavior decision-making, Wanger et al. [48] applied causal discovery to assess the impact of urban morphology on mobility behaviors, highlighting the method's significance in urban planning. Chauhan et al. [49] evaluated causal and predictive models in modeling travel mode choices, identifying advantages based on research goals. Despite their underutilized potential, causal models offer promising new insights for travel decision behavior research [50], enhancing our PRC analyses.

### III. METHODOLOGY

#### A. Problem Formulation

Our hypothesis posits that the Pedestrian Route Choice (PRC) process encompasses two key components - route choice evaluation and trajectory generation. To begin, we divided the urban area into  $250m \times 250m$  grid cells, treating each cell as the fundamental unit for choice sets. Once a destination is determined, pedestrians must make decisions at each cell transition. During this stage, pedestrians consider multiple factors to assess their choices, such as road network characteristics, built environment features, and individual traits. These factors serve as inputs in the reward function, which ultimately evaluates the pedestrian's decisions on the grid. A higher reward value suggests greater attractiveness of a grid cell for the pedestrian's route selection. Subsequently, a comprehensive trajectory is generated to guide the pedestrian's movement within the environment. In the trajectory generation process, key considerations include the reward value of the current grid cell and the distance to the start and end points. This ensures that the resulting path is as short as possible while also traversing grid cells with higher reward values.

The pedestrian's travel decision-making process can be effectively modeled using the Markov Decision Process (MDP) framework. In this framework, we can represent the Pedestrian Route Choice (PRC) as an MDP tuple, denoted as  $(S, A, P_T, R, P_0)$ . Here,  $S$  represents the set of urban spatial grid cells,  $A$  denotes the possible actions that pedestrians can take in each grid cell,  $P_0$  represents the probability distribution of the initial grid,  $P_T(S_j | A_i, S_i)$  describes the probability of a pedestrian transitioning to grid  $S_j$  after taking action  $A_i$  in grid  $S_i$ , and  $R(S)$  represents the reward function that captures the pedestrian's preference for the current state. To approximate the reward function, we can utilize neural networks, which offer the flexibility to capture non-linear and complex relationships. Typically, the reward value can be represented as the output of a neural network, with the input being the grid feature  $s$  and the pedestrian's individual characteristics  $p$ . The reward function is thus given by:

$$R(s) = \lambda(\alpha f_\gamma(s) + \beta f_\theta(p)) = f(s, p) \quad (1)$$

where  $f_\gamma$  and  $f_\theta$  respectively represent neural networks with parameters  $\gamma$  and  $\theta$ . The grid features change with each action taken, while individual characteristics remain constant throughout the decision sequence. Meanwhile, Wulfmeier [40] suggests that linear MaxEnt-IRL can serve as a proxy for

more general deep learning methods. Thus, the backpropagation rules derived from linear maximum entropy IRL can be effectively applied to train neural networks. In the case of IRL, the input tuple is assumed to be MDP/R, implying the absence of an explicit reward function  $R$  in the Markov Decision Process. Instead, a set of  $N$  observed pedestrian trajectories,  $\Xi = \{\xi_1, \xi_2, \dots, \xi_n\}$ , is available, where  $\xi = ((S_1, A_1), \dots, (S_m, A_m))$ .

Applying Gaussian priors and maximum a posteriori estimation, we evaluate the posterior distribution of each subgroup, resulting in specific parameter estimates. From a distribution perspective, parameter estimation for various subgroups can be regarded as modeling each subgroup independently. To encapsulate the variations among subgroups, we leverage the powerful representation capability of deep neural networks, unifying them under a holistic framework. We assume that the parameters for environmental features remain constant, while those for individual features vary across subgroups. To achieve this, we design a two-layer neural network that initially segregates environmental and individual features and then merges them in subsequent layers. This approach not only captures intricate patterns in PRC but also distinctly differentiates fixed feature effects from those varying with individual characteristics. The overall analytical framework is shown in Figure 1.

#### B. Maximum Entropy Inverse Reinforcement Learning Incorporating Individual Information

Based on our assumptions, when pedestrians make route choice decisions, they are influenced not only by the built environment attributes of the urban grid but also by their individual characteristics. To account for this heterogeneity, we introduce the concept of varying reward parameters for different pedestrians in their route choice decisions. In order to calculate these parameters, we can compute the empirical feature values for multiple observed pedestrian trajectories as  $f_\Xi = \frac{1}{|\Xi|} \sum_{\xi \in \Xi} f_\xi$ . Specifically, the average feature representation for a trajectory passing through  $m$  grids can be expressed as:

$$f_\xi = \sum_{i=1}^m f_\gamma(s_i) + f_\theta(p) \quad (2)$$

Assuming that the IRL algorithm has learned a reward function  $R$ , and the RL algorithm has learned a policy function  $\pi$  using this reward function, we can subsequently generate  $n$  trajectories based on this learned policy. We can assume that these generated trajectories adhere to the following constraints, which are satisfied by the real observed trajectories:

$$E[f_\xi | \pi] = f_\Xi \quad (3)$$

From a probabilistic perspective, it can be understood that there exists a latent probability distribution, under which the trajectories are generated. However, if we only consider the constraints of equation (3), multiple possible reward functions might exist for the same trajectory. To address the ambiguity, we derive based on the maximum entropy constraint [37].

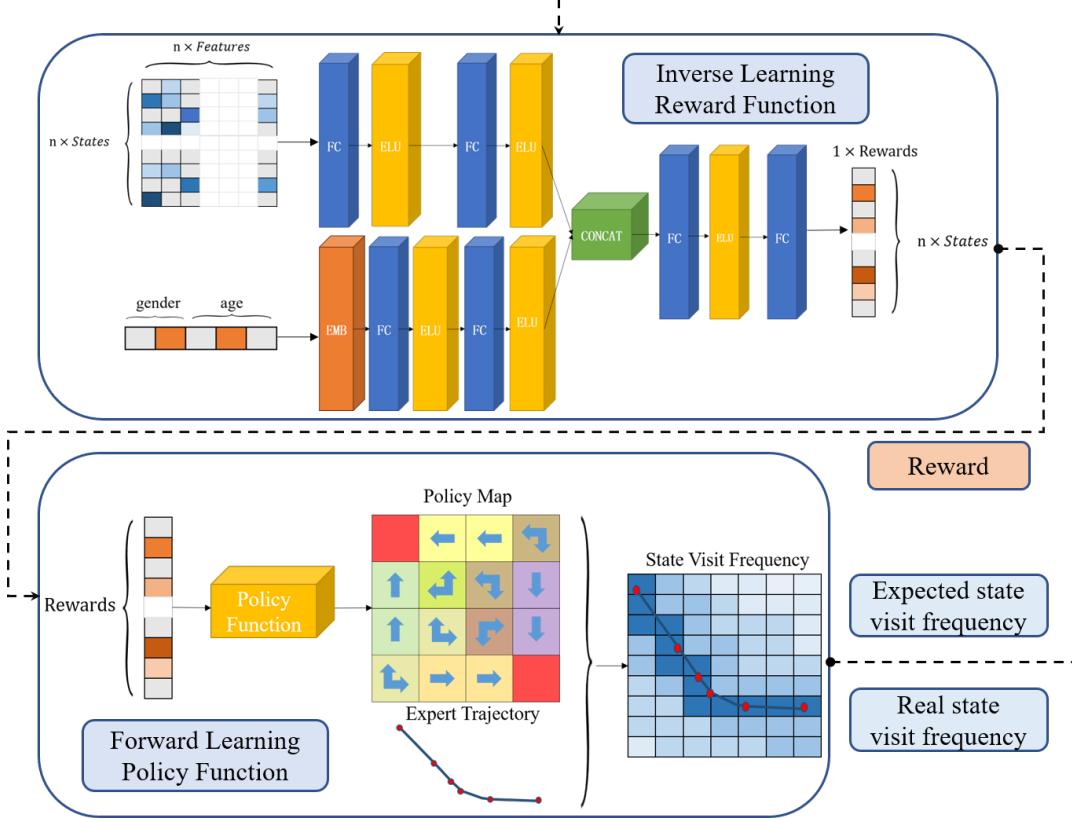


Fig. 1. The overall framework of maximum-entropy deep inverse reinforcement learning with individual covariates

$$\max_{\text{Path } \xi_i} \sum p(\xi_i) \log p(\xi_i) \quad (4)$$

$$\text{s.t. } \sum_{\text{Path } \xi_i} p(\xi_i) f_{\xi_i} = \tilde{f} \quad (5)$$

$$\sum_{\text{Path } \xi_i} p(\xi_i) = 1 \quad (6)$$

Using the method of Lagrange multipliers, we can transform the optimization problem by eliminating the constraints from equations (4)-(6) and expressing it in terms of equation (7):

$$\min L = \sum_{\xi_i} p \log p - \sum_{\xi_i=1}^n \lambda_{\xi_i} (p f_{\xi_i} - \tilde{f}) - \lambda_0 (\sum p - 1) \quad (7)$$

By replacing  $f_{\xi_i}$  with the expression form given in Equation (2), the summation of trajectory features is converted into the summation of state features, represented as  $\sum_{\xi_i=1}^n \lambda_{\xi_i} f_{\xi_i} = \sum_{i=1}^{m \times n} \lambda_{\gamma} f_{\gamma}(s_i) + \lambda_{\theta} f_{\theta}(p)$ , encompassing both environmental features that change with the decision-making process and globally static individual characteristics. Differentiating equation (7) and setting its derivative to zero yields equation (8):

$$\begin{aligned} p &= \frac{\exp(\sum_{\xi_i=1}^n \lambda_{\xi_i} f_{\xi_i})}{\exp(1 - \lambda_0)} \\ &= \frac{\exp(\sum_{i=1}^{m \times n} \lambda_{\gamma} f_{\gamma}(s_i) + \lambda_{\theta} f_{\theta}(p))}{\exp(1 - \lambda_0)} \\ &= \frac{1}{Z} \exp(\sum_{j=1}^{m \times n} \lambda_{\gamma} f_{\gamma}(s_i) + \lambda_{\theta} f_{\theta}(p)) \end{aligned} \quad (8)$$

where  $Z = \exp(1 - \lambda_0)$ . By combining equation (8) with equation (6), we can obtain the generation probability of the trajectory:

$$\begin{aligned} P_{\lambda} &= P(\xi_i | \lambda) \\ &= \frac{\exp(\sum_{i=1}^m \lambda_{\gamma} f_{\gamma}(s_i) + \lambda_{\theta} f_{\theta}(p))}{\sum_{\text{Path } \xi_i} \exp(\sum_{i=1}^m \lambda_{\gamma} f_{\gamma}(s_i) + \lambda_{\theta} f_{\theta}(p))} \end{aligned} \quad (9)$$

The formula for the maximum posterior probability is:

$$p(\lambda | \xi) = \frac{p(\xi | \lambda) p(\lambda)}{p(\xi)} \quad (10)$$

Given the heterogeneity among pedestrians, we assume that the weight vector  $\lambda$  follows a specific distribution, such as the Gaussian distribution. Consequently, the route choice probability for each pedestrian is determined by their unique  $\lambda$  value. Since  $p(\xi)$  has no effect on the parameter estimation of  $\lambda$ , the expression for solving  $\lambda$  is as shown in equation (11):

$$\lambda = \arg \max_{\lambda} \sum_{\xi \in \Xi} \log p(\xi|\lambda) p(\lambda) \quad (11)$$

Let  $L(\lambda) = \sum_{\xi \in \Xi} \log p(\xi|\lambda) p(\lambda)$ , then  $L(\lambda)$  can be expressed as:

$$\begin{aligned} L(\lambda) &= \sum_{\xi \in \Xi} (\log p(\xi|\lambda) + \log p(\lambda)) \\ &= \sum_{\xi \in \Xi} (L_D + L_\lambda) \end{aligned} \quad (12)$$

Differentiating equation (12) yields:

$$\begin{aligned} \nabla L(\lambda) &= \sum_{\xi \in \Xi} \left( \frac{\partial L_D}{\partial R} \frac{\partial R}{\partial \lambda} + \frac{\partial L_\lambda}{\partial \lambda} \right) \\ &= (\mu_D - \mathbb{E}[\mu]) \frac{\partial g(f, \lambda)}{\partial \lambda} + \frac{\partial L_\lambda}{\partial \lambda} \end{aligned} \quad (13)$$

where  $\frac{\partial L_\lambda}{\partial \lambda}$  is the regularization term with respect to parameter  $\lambda$ . Assuming the parameter follows a Gaussian distribution with a mean of 0 and variance of  $\sigma^2$ , i.e.,  $p(\lambda; \sigma) = \mathcal{N}(\lambda; 0, \sigma^2)$ , the regularization term can be expressed as  $\frac{\partial L_\lambda}{\partial \lambda} = -\frac{\lambda}{\sigma^2}$ .

Using the above mathematical derivation, we assessed the posterior distributions for various pedestrian subgroups, deriving their specific parameter estimates. Differences between these groups can be accommodated using parametric distributions, eliminating the necessity for conducting multiple experiments for each group. Our aim is to demonstrate how this methodology identifies distinct groups within a unified framework, distinctly differentiating the effects of static individual attributes from those of built environment features that change with decision progression. By integrating these elements, we provide a robust mathematical foundation that bolsters the reliability and versatility of our model.

### C. Causal Discovery between Environmental Factors and Reward Values

While our MEDIRL-IC approach leverages deep neural networks to model the factors influencing PRC, the lack of interpretability arising from the black-box nature of neural networks poses a challenge when trying to explore the decision-making process in detail. To overcome this challenge and gain a clearer understanding of how environmental and individual factors impact route choice behavior, we utilize causal discovery methods to analyze the results of our model.

Glymour et al. [47] provided a detailed review of causal discovery algorithms, categorizing them into two major groups. The first group is based on conditional independence tests, which includes constraint-based and score-based causal discovery algorithms, such as PC, FCI [51], GES [52]. These algorithms are advantageous for their broad applicability and competence with large-scale samples but struggle to differentiate Markov equivalence classes. The second group, centered on functional causal models (FCM), features algorithms such as LiNGAM [53] and ANM [54]. While this category

can distinguish Markov equivalence classes due to specific data distribution assumptions, it comes with limitations in applicability across some causal relationships and increased computational complexity. These FCM-based methods also exhibit greater sensitivity to noise and are better matched with high-quality data. Chauhan's research [49] further evaluated the performance of these algorithms in uncovering causal relationships in travel choice behaviors. Choosing the appropriate method requires a reasoned analysis based on application scenarios and data characteristics. Considering the PC algorithm's effectiveness in deducing causal relationships and its adaptability to large datasets, this paper uses it for model interpretation.

The PC algorithm begins with a fully undirected graph  $G$ , connecting all features. It iteratively examines nodes  $x$  and  $y$  in  $G$  with an initially empty separation set  $O$ . Based on the neighbors of  $x$  and  $y$ , nodes are progressively added to  $O$ . When  $x$  and  $y$  satisfy a conditional independence relationship, their connecting edge in the skeleton graph is removed [55]. This procedure results in a directed causal graph delineating the causal relationship between features and reward values.

The core of the PC algorithm lies in determining the independence between features. Given variables  $y, x_1, x_2, \dots, x_n$ , when controlling for some of the features  $x_2, x_3, \dots, x_n$ , the correlation between two features  $y$  and  $x_1$  cannot be determined using the Pearson correlation coefficient. At this point, higher-order partial correlation coefficients can be utilized for analysis. When the influence of  $h$  features is excluded, it is called the  $h$ -th order partial correlation coefficient ( $h \leq K - 2$ ), and its expression is shown in equation (14):

$$\rho_{x,y|K \setminus h} = \frac{\rho_{x,y|K \setminus h} - \rho_{x,h|K \setminus h} \rho_{y,h|K \setminus h}}{\sqrt{(1 - \rho_{x,h|K \setminus h}^2)(1 - \rho_{y,h|K \setminus h}^2)}} \quad (14)$$

where  $\rho_{x,y|K \setminus h}$  represents the partial correlation coefficient between  $x$  and  $y$  when conditioned on all variables in  $K$ , excluding  $h$  features in the separation set. To ascertain the conditional independence of  $x$  and  $y$  using  $\rho_{x,y|K}$ , the Fisher Z-Test is applied. Initially, the Fisher Z-Transformation is utilized to convert  $\rho_{x,y|K}$  into a normal distribution, the transformation formula is represented by equation (15):

$$Z_{x,y|K} = \frac{1}{2} \log \left( \frac{1 + \rho_{x,y|K}}{1 - \rho_{x,y|K}} \right) \quad (15)$$

Next, we set up two hypotheses: the null  $H_0(x, y|K) : \rho_{x,y|K} \neq 0$  suggesting that  $x$  and  $y$  are not conditionally independent, and the alternative  $H_1$  asserting their conditional independence. If the calculated test statistic exceeds a threshold derived from significance level  $\alpha$ , we reject  $H_0$ . Consequently, if  $x$  and  $y$  are found to be conditionally independent when controlling for  $h$  features, the link between them can be eliminated.

After the aforementioned procedures, we obtain a partially undirected graph depicting the relationships between features and reward values. Utilizing the d-separation principle, we discern dependency directions among features, transforming

the graph into a partially directed one, shifting from correlation to causality.

It is worth noting that the directed graph established by the PC algorithm might contain undirected edges. If undirected edges occur between features and reward values, their direction can be manually determined based on the research assumptions. For edges within features, judgment can be made based on expert knowledge. This ensures alignment with prior understanding, making the final causal graph consistent with real-world contexts.

## IV. EXPERIMENTS

### A. Pedestrian Trajectory Data

Conventional travel surveys often overlook many types of pedestrian slow-paced activities [56]. These surveys typically have a restricted number of respondents, and the collected background information can often be skewed. Furthermore, monitoring participants over extended durations poses a significant challenge. As a result, researchers find it difficult to analyze the route choice behavior of different pedestrians on a larger scale [7]. Mobile phone signaling data offers an edge over these traditional methods. It provides expansive, large-scale, and unbiased behavioral insights [56], free from recall errors and personal biases [57]. Given the ubiquity of mobile devices like smartphones, the ability to detect human movement and activity patterns temporally and spatially through mobile phone signaling has become increasingly feasible. In this study, we utilize mobile phone signaling data in the city of Shenzhen to reconstruct pedestrians' travel trajectories.

For primarily non-motorized slow travel modes, the route choice behaviors are more significantly influenced by the surrounding built environment and individual preferences. Consequently, grasping the intricacies of PRC becomes pivotal for urban planning, underpinning the creation of sustainable and habitable urban landscapes [58]. Urban centers, exemplified by the Nanshan District in Shenzhen, chosen for this study, boast an extensive network of mobile phone signaling base stations. Such a dense setup augments trajectory sampling frequency, making the analysis of PRC more meticulous. Moreover, the rich variety of functional zones in the area, including commercial sectors, cultural venues, and green spaces, offers a robust canvas to discern the influences of the built environment on PRC.

We filtered urban residents' slow travel trajectories using criteria of speeds under 5 km/h and lengths below 5 km. We further refined our dataset by only considering trajectories commencing and concluding within the Nanshan District, recorded between July 1 and 16, 2019. This filtering yielded 95,014 trajectory entries from 44,164 unique individuals, cumulatively spanning 2,160,763 road nodes, with 3,584 being unique to the Nanshan District. Such comprehensive data is invaluable for in-depth analysis of slow travel behaviors.

Post de-identification and preprocessing, the data was structured into three distinct tables: one for individual attributes, another for origin-destination (OD) pairs, and a third for the specific road nodes each travel covered. These nodes facilitated the depiction of precise routes for each OD pair. Upon

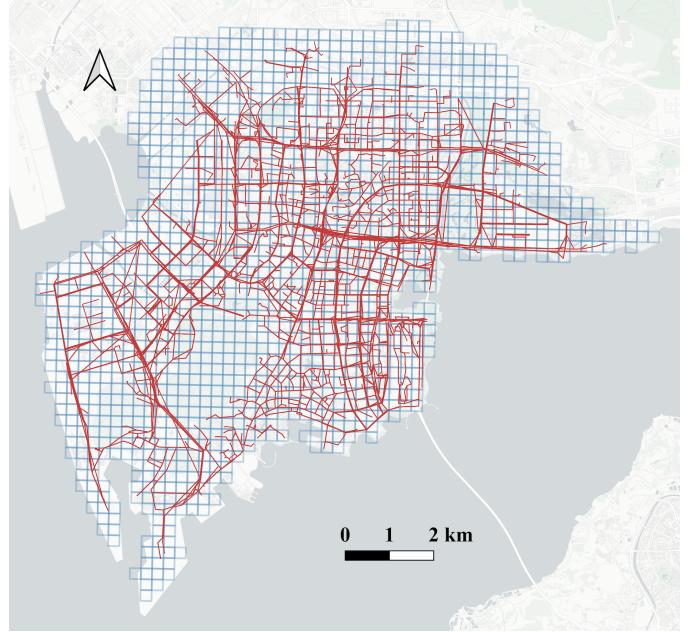


Fig. 2. Research area (Nanshan, Shenzhen) and pedestrian trajectories

alignment with the OpenStreetMap (OSM) road network, we derived full pedestrian trajectories, visualized in Figure 2. Here, red lines depict the filtered pedestrian trajectories, and the blue grids highlight the grid cells of focus in this study.

### B. Feature Selection

The multifaceted nature of pedestrian travel decisions necessitates the careful selection of indicators for modeling, ensuring an accurate representation of their route choice behaviors. Basu et al. [8] synthesized prior research, grouping 105 pertinent factors into three primary categories: socio-demographic factors, built environment factors, and trip-related features. The trip-related features provide a holistic descriptor of the entire route and are not apt for our dynamic choice framework. Therefore, our PRC-centered study will focus exclusively on socio-demographic and built environment factors, without delving into trip-related features.

When delving into socio-demographic factors, our study underscores gender and age as pivotal factors. Existing research indicates a gendered preference: male pedestrians often favor less crowded routes, while females prioritize routes with easily accessible shops [59]. Age presents a more complex relationship with PRC [60]. Some research suggests that younger pedestrians gravitate towards bustling paths leading to commercial zones [61], yet other studies posit an ambiguous connection between age and travel choices [62]. Guided by the "Medium and Long-term Youth Development Plan (2016-2025)" issued from the State Council of China, we categorized our study's data into three age brackets: youth (14-34), middle-aged (35-64), and elderly (65 and above). Given the sparse ownership of mobile communication devices among children, individuals below 14 years are excluded from our analysis. In essence, a nuanced comprehension of these socio-demographic

aspects is vital for sculpting urban spaces that resonate with varied pedestrian demographics.

This study categorizes built environment elements into three major categories:

1) *Road-surrounding land use metrics*: Prior research underscores a marked correlation between residential density and PRC [63]. To delve deeper, this study harnesses population census data in tandem with architectural information sourced from Shenzhen's "Kintsmesh Project". We align this population data with specific residential building identifiers. Subsequently, by aggregating populations within distinct urban grids, we derive the residential density for each, expressed in terms of thousands of individuals per square kilometer.

Leveraging the land use data from Shenzhen, our study calculates both the land use mix and the proportion of open spaces within the city. Previous studies have indicated the potential influence of diverse land use configurations on route preferences [64]. By integrating the land use data with urban grids spatially, we quantify various land use types within each grid. We then employ the Shannon entropy to compute the land use mix, as shown in Equation (16). In this equation,  $A_{ij}$  represents the area proportion of the  $i$ -th land use type within the  $j$ -th grid, and  $N$  stands for the total count of land use types present in that grid. The resulting value for this metric lies between 0 and 1: a score of 0 indicates a uniform land use type, while a score of 1 points to an equitably dispersed mix of all land use types within the grid.

$$E_j = \left( - \sum_i (A_{ij} \ln(A_{ij})) \right) / \ln(N_j) \quad (16)$$

Furthermore, research by Guo and Loo underscores the significant correlation between open spaces and PRC [27]. To delve deeper into this association, our study pinpoints areas designated as open spaces from the land use data. These earmarked spaces are then spatially integrated with the 250m urban grid data, facilitating the calculation of the open space proportion for each grid.

2) *Location features*: Initially, we gathered 521,495 Points of Interest (POI) data entries for Shenzhen from the *Amap* platform. To focus on non-motorized slow travel, we narrowed down to 12 pertinent categories and then tallied the distinct category POI points within each grid. Recognizing the inherent variability in the baseline number of distinct POIs, we adopted the term frequency-inverse weighted frequency (TF-IWF) algorithm for nuanced, weighted computations. A specific feature is accorded greater weight within a grid if its proportion surpasses the average proportion observed across all grids. Let  $N_{d,t}$  denote the count of POI of type  $t$  in grid  $d$ ,  $N_d$  symbolize the aggregate POIs in grid  $d$ ,  $W_c$  represent the total POIs across Shenzhen, and  $W_{c,t}$  signify the count of POI of type  $t$  spanning all grids. The formulation of TF-IWF metric is depicted in equation (17).

$$\text{TF-IWF} = \text{TF} \times \text{IWF} = \frac{N_{d,t}}{N_d} \log \left( \frac{W_c}{W_{c,t}} \right) \quad (17)$$

3) *Road-network features*: Leveraging OSM road network data, our study identifies three critical indicators: intersection

density, distribution density of varied road types, and proximities to the city center and key transportation hubs.

Initially, by employing the "breaking intersecting lines" function in ArcGIS, we establish links between each node and its neighboring roads. This process facilitated the identification of intersections, which were subsequently integrated spatially with the grid. The resulting intersection counts within each grid were then transformed into a density metric, denoted as intersections per square kilometer.

Next, OSM categorizes roads into diverse types, including footways and pedestrian paths. By spatially aligning the road data with the grid, we logged the length proportions of these road categories within each grid. Relying on the previously referenced TF-IWF metric (equation (17)), we assessed the significance of individual road types.

Lastly, to compute distances to the city center and paramount transportation hubs, transportation site data for Shenzhen was extracted from the POI dataset. This dataset encompassed an airport, eight train stations, one coach terminal, and 68 bus stations. We further identified coordinates for two primary city centers and five sub-centers in Shenzhen. Using Shenzhen's road network data in conjunction with ArcGIS, we designated the grid center as the origin and the city center or transportation hubs as destinations. The network analysis tool in ArcGIS enabled us to measure the network distances for each O-D pair, determining the distance from each grid to the nearest vital facility in kilometers. Table I encapsulates the built environment attributes incorporated in our study. Notably, grids with null values were omitted during the calculation of means and standard deviations.

### C. Experiment Design

1) *Model Accuracy Evaluation*: We assess the stability of the training process by tracking the reward difference's convergence over consecutive epochs and measuring trajectory similarity using the average LLR. A stabilized parameter set indicates our reward function's capability to accurately represent pedestrian trajectories. To delve deeper into model accuracy, we produced predictive trajectories via the enhanced A\* algorithm and compared them with real trajectories from the test set. The model's prediction accuracy was assessed using the average DTW distance. We benchmarked our model against three established models: the Recursive Logit (R-logit), the DNN-PSL, and the MaxEnt-DIRL, which only considers built environment features.

2) *Urban Grid Attraction with Reward Function*: In our study, we employed a reconstructed reward function to analyze the route choice behaviors of different pedestrian groups. Within a given area, the reward function was adapted by incorporating both the built environment features and the socio-demographic profiles of pedestrian subgroups. The derived data was depicted in the form of a reward map, highlighting the unique routing preferences inherent to each subgroup. Moreover, by averaging the reward values across all subgroups, we shed light on the spatial distribution of built environment attributes in Shenzhen and assessed their influence on PRC behaviors.

TABLE I  
BUILT ENVIRONMENT FEATURES

<b>Variables</b>	<b>Description</b>	<b>Max</b>	<b>Mean</b>	<b>St.dev</b>
<b>Road-surrounding Land Use Features</b>				
Population Density	Ratio of the number of people in the grid to the area of the grid (thousands of persons/km <sup>2</sup> )	148.87	9.04	15.61
Land Use Mix	Diversity of land-use patterns in a given grid (0-1)	0.95	0.42	0.27
Open Space Ratio	Ratio of the area of public and private open space within the grid (0-1)	1.00	0.48	0.38
<b>Location features</b>				
Attractions	Park, plazas and tourist sites (no.)	5	1.14	0.13
Food & Beverages	Restaurants, cafes and other dining establishments (no.)	170	11.42	9.37
Transportation	Airports, train stations, bus terminals and other transit hubs (no.)	56	4.57	3.45
Sport	Gymnasiums, basketball venues and other recreational venues (no.)	71	2.79	1.68
Public	Public restrooms, payphones, and other community amenities (no.)	20	1.68	0.95
Enterprises	Companies, factories, and agricultural, forestry, fisheries establishments (no.)	231	9.30	9.25
Medical	Hospitals, pharmacies, and other healthcare facilities, including veterinary clinics. (no.)	33	3.69	1.99
Government	Government offices, embassies, prosecutors, and other official institutions. (no.)	40	3.01	1.96
Life	Telecommunication centers, post offices, and other essential service providers. (no.)	111	7.19	5.31
Education & Science	Museums, art galleries, schools, and other educational or research institutions. (no.)	68	4.07	2.92
Shopping	Shopping malls, supermarkets, and other retail outlets. (no.)	16	1.27	0.29
Finance	Banks, ATMs, insurance agencies, and other financial institutions. (no.)	10	1.34	0.17
<b>Road-network Features</b>				
Intersections	Number of intersections within a certain grid (intersections/km <sup>2</sup> )	192	11.87	19.47
Center	Distance to the nearest urban center (km)	40.61	9.89	9.89
Airport	Distance to the nearest airport (km)	98.80	42.35	24.47
Railway	Distance to the nearest railway station (km)	41.89	11.61	9.10
Dock	Distance to the nearest port (km)	85.13	38.51	19.10
Coach	Distance to the nearest coach station (km)	20.98	3.22	4.03
Expressway	Density of expressway arterials, ramps and overpasses (m/ km <sup>2</sup> )	1675.31	247.01	85.17
Cycleway	Density of greenways, non-motorized lanes, park bike lanes (m/ km <sup>2</sup> )	1203.89	133.95	24.41
Suburban Road	Density of suburbs, industrial and mining areas, forest paths (m/ km <sup>2</sup> )	591.62	134.99	12.76
City Branch	Density of residential road, airport road, waterfront road (m/ km <sup>2</sup> )	1888.08	173.24	100.51
Inner Road	Density of living streets, trails, service roads, horse trails (m/ km <sup>2</sup> )	3200.51	220.69	108.48
Main Road	Density of main road, secondary road (m/ km <sup>2</sup> )	3799.80	305.60	223.82
Not Built Road	Density of construction road (m/ km <sup>2</sup> )	1843.42	175.16	44.73
SideWalk	Density of sidewalk, pedestrian street, steps (m/ km <sup>2</sup> )	5093.03	223.62	97.841
Urban Secondary	Density of urban secondary roads, airport vehicle roads (m/ km <sup>2</sup> )	1028.55	152.71	68.42

**3) Causal Mechanism Analysis of PRC Behavior:** To understand the route choice behaviors of different pedestrian demographics, we employed the graph-based causal discovery methods. Initially, we constructed a directional causal graph to visually represent the relationships between features and reward values. This graph aids in pinpointing primary features influencing PRC preferences. We further utilized higher-order partial correlation coefficients to gauge the positive or negative influence of these features. By mapping these coefficients, we created a causal strength heatmap, allowing for precise assessment of feature importance. This methodological approach not only distinguishes the impacts of features across pedestrian demographics but also sets the foundation for a deeper analysis.

#### D. Evaluation Metrics

To validate our model's efficacy, we implement a series of meticulous evaluation methods.

**1) Reward Function Convergence:** We first evaluate the convergence of the reward function during the training progress to ascertain that our model suits the training dataset effectively. Given the unavailability of real reward values for urban grids, we gauged the training convergence by measuring the reward function's difference between successive epochs. The training is deemed complete and halted when the average reward variation dips below 0.01.

**2) Similarity Assessment using Average Log-Likelihood Ratio (LLR):** To underscore our methodology's ability to genuinely mirror pedestrian route preferences, we assess the similarity between predicted trajectories and actual pedestrian paths using the average LLR. If  $X$  stands for the feature attribute of the grid the real pedestrian trajectory traverses and  $Y$  represents the corresponding attribute column of the generated trajectory, the LLR between these columns is defined as equation (18):

$$LLR(x_i) = \frac{\sum_{i=1}^n \left( \frac{\log(P(X=x_i))}{\log(P(Y=x_i))} \right)}{n} \quad (18)$$

Here,  $x_i$  represents the corresponding element in both feature attribute columns, and  $n$  denotes the total count of elements in grid feature  $X$ . The LLR discrepancies for each column are aggregated, and their mean gives insights into modeling precision. An LLR nearing zero denotes the generated trajectory is akin to the real trajectory, while a larger LLR signifies a divergence.

**3) Predictive Trajectory Using Enhanced A\* Algorithm:** Building on the reconstructed reward functions for different pedestrian demographics, we utilize an enhanced A\* algorithm to predict pedestrian trajectories between given origins and destinations. This algorithm's essence lies in calculating each grid's priority as illustrated in equation (19):

$$f(n) = g(n) + h(n) + \lambda \cdot \text{cost}(n) \quad (19)$$

where  $f(n)$  represents the cumulative priority of grid  $n$ ,  $g(n)$  signifies the distance from grid  $n$  to the origin,  $h(n)$

represents the estimated distance from grid  $n$  to the destination, and  $\text{cost}(n)$  denotes the cost of choosing grid  $n$ .  $\lambda$  is the cost's scaling factor, with  $\text{cost}(n) = 1 - \text{reward}(n)$ .

**4) Dynamic Time Warping (DTW) for Trajectory Similarity:** We leverage the Dynamic Time Warping (DTW) distance as a metric to assess the similarity of true and predicted trajectories. DTW quantifies the average discrepancy between the real and the predicted trajectories. The real trajectory is represented by point sequence  $\mu_A = \{a_1, a_2, \dots, a_m\}$  and the predicted trajectory by  $\mu_B = \{b_1, b_2, \dots, b_n\}$ , where each  $a_i$  and  $b_j$  are two-dimensional coordinate points. Subsequently, an  $m \times n$  matrix  $D$  is constructed, where each element  $D(i, j)$  represents the euclidean distance between point  $a_i$  from  $\mu_A$  and point  $b_j$  from  $\mu_B$ , defined as equation (20):

$$D(i, j) = \sqrt{(a_{i,x} - b_{j,x})^2 + (a_{i,y} - b_{j,y})^2} \quad (20)$$

Then, for all points  $i = 1, \dots, m$  and  $j = 1, \dots, n$ , the DTW distance between any two points can be computed iteratively as shown in equation (21), where  $DTW(m, n)$  represents the minimum distance between sequences  $\mu_A$  and  $\mu_B$ .

$$DTW(i, j) = D(i, j) + \min \left\{ DTW(i-1, j-1), DTW(i-1, j), DTW(i, j-1) \right\} \quad (21)$$

**5) Incremental Assignment (IA) Method for Dynamic Traffic Assignment:** To assess the reconstructed reward function's efficacy, dynamic simulations were executed utilizing an Iterative Assignment (IA) traffic assignment algorithm. This methodology is adaptable to both grid-based and conventional road networks. Let  $G = (V, E)$  denote the traffic network, with  $V$  representing the traffic nodes and  $E$  symbolizing the edges. In each iteration, pedestrians make travel decisions aligned with their individual preferences and the current flow of each route. The formula for updating flow is delineated in Equation (22), where the current total flow difference  $F$  is calculated as  $F = \sum_{e \in E} |f_e^{(k+1)} - f_e^k|$ . The iterative update process halts when the difference in allocated flow between two consecutive distributions falls below a predefined threshold, as indicated in Equation (23):

$$f_e^{(k+1)} = f_e^k + \Delta f_e \quad (22)$$

$$\frac{F}{\sum_{e \in E} f_e^k} < \epsilon \quad (23)$$

To evaluate the effectiveness of our model based on traffic assignment outcomes, we employ the Path Overlap Ratio (POR) as a metric to assess the results derived from the IA implementation for dynamic traffic equilibrium [12]. For a given route, let  $p_r$  denote the quantity of actual trajectories traversing it, and  $p_e$  represent the number of trajectories generated by the IA algorithm. Given a road network comprising  $N$  routes, the POR can be formally defined as in Equation (24):

$$POR = \frac{\sum_{n \in N} p_r}{\sum_{n \in N} p_e} \quad (24)$$

### E. Model Configuration

In this paper, based on the standard partitioning strategy, we divided the trajectory data into 70% for the training set, 15% for the validation set, and 15% for the testing set. To meet our model's requirements for discrete states, we transformed the selected experimental trajectories into (State, Action) tuples. Here, "State" represents the index value of the 250-meter grid in Shenzhen, while "Action" denotes the expert trajectory's movement on that grid. The possible actions are: move up, move down, move left, move right, and stay.

The built environment features are expressed using a  $N_{\text{grids}} \times 30$  vector matrix, where  $N_{\text{grids}}$  represents the number of grids in the study area. Each feature is normalized to counteract the effects of varying scales. Individual features, such as gender or age demographics, are expressed using one-hot encoding. Given the absence of a clear ordinal relationship for these features, this encoding method ensures feature independence, thereby preventing any inadvertent numerical biases during model interpretation.

For the training process, we need to set three primary hyperparameters: the learning rate  $\alpha$ , the discount factor  $\gamma$  for the reward function, and the number of iterations  $n$ . To find the optimal parameter combination, we employed a grid search method. The learning rate  $\alpha$  spans the range  $\{0.01, 0.02, \dots, 0.2\}$ , the discount factor  $\gamma$  is within  $\{0.1, 0.2, \dots, 1.0\}$ , and the reward function difference for the termination condition is set to be less than 0.01. We randomly selected 100 trajectories for performance evaluation. The final settings for the hyperparameters were  $\alpha = 0.02$ ,  $\gamma = 0.9$ , and  $n = 30$ . Notably, during the forward learning of the policy function, a value iteration for the entire grid must be performed. Consequently, the training duration is contingent on both the grid count in the study area and the trajectory count. Training was executed on a NVIDIA 4090 graphics card. Each iteration lasted about 50 minutes, with the reward function converging in under 20 iterations.

## V. RESULTS

### A. Evaluation of Model Accuracy and Comparative Analysis

We divided the dataset into six distinct subsets based on individual characteristics. Each of these subsets was employed to update the model parameters during a single iteration, with one epoch covering a complete traversal of the entire dataset. The stabilization of the reward function indicates the model's convergence to a reliable state that can adeptly fit the real trajectories. Another essential criterion for assessing model accuracy involves comparing the feature expectations of grids passed by predicted trajectories to those of the real trajectories. As illustrated in Figure 3, as the training epochs increase, there is a noticeable convergence of the reward function. Simultaneously, the average LLR between the predicted and real trajectories exhibits a steady decline. This implies a growing likelihood that the predicted trajectories under the prevailing reward function closely resemble the real ones.

To assess the validity of the reconstructed reward function, this study generates predictive trajectories using the enhanced A\* algorithm and compares them with real trajectories from

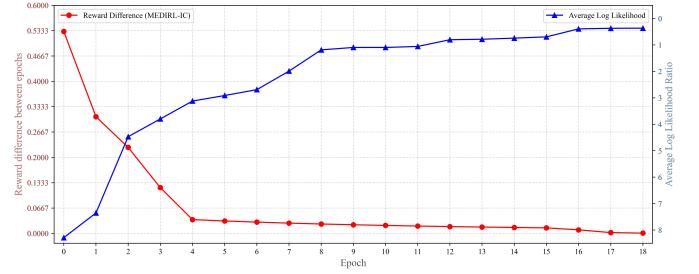


Fig. 3. Convergence analysis of reward difference and trajectory similarity (average log likelihood ratio)

the test set. A positive correlation exists between the consistency of the predicted trajectory and the real ones, signifying a tight fit between the reward function and PRC behavior. For comparative insights, we benchmarked against three distinct models: the Recursive Logit (R-logit) model [10], DNN-PSL model [65], and the MaxEnt-DIRL model, which does not consider individual traveler's attributes. The Recursive Logit model can be considered as an inverse reinforcement learning model that adapts a linear reward function. In contrast, the DNN-PSL model employs a deep neural network to infer the utility function of the choice set, adeptly capturing the nonlinear relationship between features. Compared with the MaxEnt-DIRL model, it is expected that after considering individual characteristics, the MEDIRL-IC model's ability to fit PRC can be improved. The complexity of trajectory evaluation is an aspect that cannot be overlooked when considering individual travel decisions [66], underscoring the necessity for a metric that accommodates both individual variations and trajectory deviations. The predicted trajectories are derived from the demographic's reward map, which inherently integrates individual preferences. To quantify modeling accuracy, we employed the average DTW distance between the predicted and real trajectories. This metric effectively measures similarity by capturing both the shape and sequence of trajectories, offering a detailed view of model performance. The comparative accuracy of the various models is illustrated in Figure 4.

Results from Figure 4 underscore the effectiveness of the MEDIRL-IC model proposed in this study, highlighting its proficiency in accurately replicating authentic PRC behavior. The DTW distance mean and interquartile range (IQR) for the MEDIRL-IC model are 2.92 and 0.77, respectively, which are the lowest among the models compared. This indicates that the MEDIRL-IC model exhibits superior accuracy and stability in trajectory prediction. Outliers in MEDIRL-IC can be attributed to model complexity and data variability. In accordance with the bias-variance trade-off theory, increased complexity reduces bias but elevates variance [67]. Outliers also emerge due to individual heterogeneity, where certain individuals' unique travel preferences significantly deviate from the norms of the group. Compared to MaxEnt-DIRL, our method demonstrates improvements in both the mean and the IQR, suggesting that individual preferences are key decision variables influencing PRC behavior. Notably, our model significantly outperforms the R-Logit model. This superiority

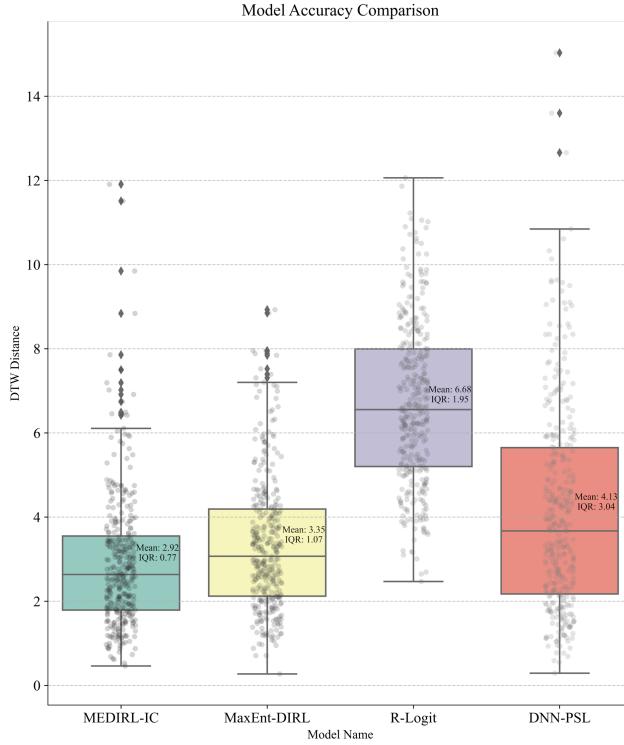


Fig. 4. Comparison of trajectory accuracy across different models

indicates that deep neural networks can considerably enhance the model's feature-fitting capability. When compared with the DNN-PSL model—a discrete choice model rooted in the logit function—our model clearly outperforms. This performance underscores our model's capacity to adeptly discern dependencies between sequential choices, effectively illuminating the nuanced intricacies inherent in complex behaviors and decisions.

Figure A1 offers a visual comparative analysis of the trajectories predicted by various models, taking the PRC behavior of a young male as a representative sample. This illustration compares real trajectories against those generated by each model. For the IRL and Recursive Logit models, we applied the reward function to the urban grid, subsequently employing the enhanced  $A^*$  algorithm to generate trajectories. In the case of the DNN-PSL model, constrained by the choice set, we produced 10 alternative trajectories, selecting the one with the highest utility value.

Transitioning from the detailed visualization of individual trajectories to a broader analysis of dynamic equilibrium, Figure 5 expands the scope by evaluating the collective behavior of multiple trajectories within a real-world road network, employing the same IA method to simulate the impact of individual route choice preferences on traffic saturation. Figure 5(a) illustrates the grid distribution derived from the reconstructed grid reward values, with the blue trajectories representing the pedestrians' actual trajectories. Pedestrians make travel decisions based on their individual preferences. Figure 5(b) depicts the traffic distribution for the road network, where pedestrians aim for the shortest possible route to

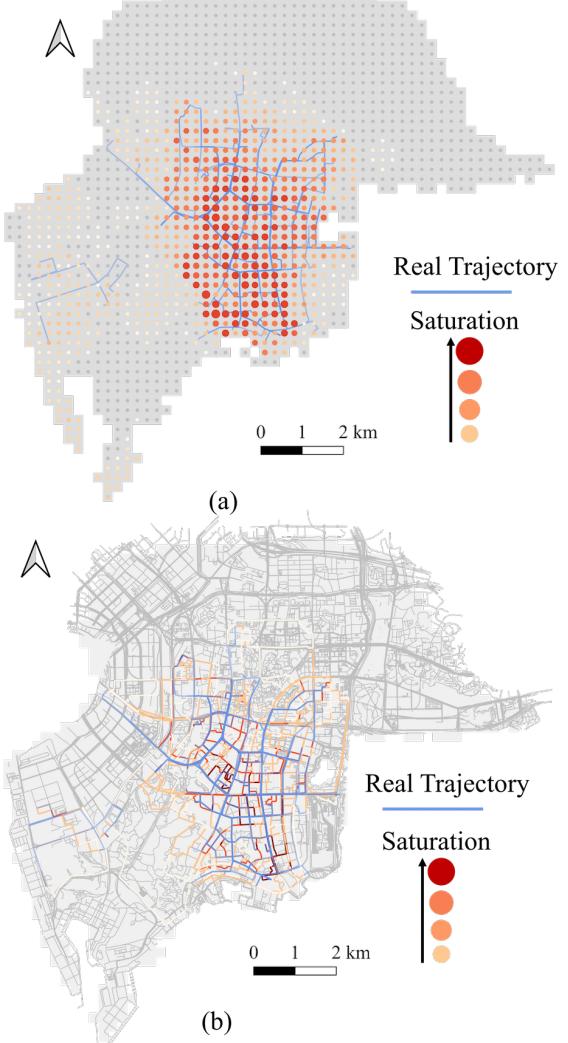


Fig. 5. Comparative results of dynamic traffic equilibrium based on (a) grid reward values and (b) road network architectures

their destinations. The comparison reveals that roads or grids traversed by real-world trajectories exhibit higher levels of traffic saturation in the simulations. The POR of the grid-based equilibrium result is  $3.713 \times 10^{-3}$ , compared to  $8.923 \times 10^{-3}$  for the road-based equilibrium. Such findings suggest that incorporating individual preferences into traffic assignment models provides a more accurate reflection of actual conditions than the traditional approach of optimizing for the minimal cost within a road network.

Upon synthesizing our observations, including assessments of mean error and IQR alongside direct comparisons of trajectory predictions, and dynamic traffic equilibrium analyses across both grid-based rewards and actual road networks, it appears that the trajectories generated by MEDIRL-IC are closely aligned with observed pedestrian movements. This observation suggests the potential benefits of integrating individual characteristics into the modeling framework for pedestrian

travel decisions. Consequently, this provides a solid foundation for subsequent modeling of pedestrian travel preferences through reward functions derived from our model.

### B. Urban Grid Attraction for Pedestrians Based on Estimated Reward Function

The preceding analysis validates the effectiveness of our proposed method, illustrating that the reward functions generated by MEDIRL-IC enable the examination of travel preference disparities across populations with different sociodemographic attributes. Based on the real trajectories, we employed the method proposed in this study to infer the reward function. Subsequently, for any given area, we can integrate its built environment features with the sociodemographic characteristics of pedestrian subgroups into the inferred reward function. This integration is exemplified in Figure 6, with the Nanshan District serving as our case study. The culmination of this process is a reward map that delineates the city grid's attraction in swaying pedestrian route choices. For clarity, we have normalized the reward values between 0 and 1, where a higher value signifies a greater probability of pedestrians opting for that grid in their route decisions.

Figure 6 shows a three-dimensional reward map of the Nanshan District. Employing our model, we calculated the reward values for the same grid across different pedestrian demographic, subsequently documenting the highest reward value along with its associated socio-demographic attributes. In this representation, the grid's color signifies the corresponding individual characteristic, while the elevation captures the magnitude of the reward value.

In Figure 6(a), data is categorized based on age, revealing the spatial distribution of grids associated with high reward values. When paired with the POI distribution analysis of the Nanshan District, it becomes apparent that grids commanding high rewards for middle-aged pedestrians predominantly cluster around areas with high-tech companies. Conversely, the grids that attract younger and elderly age demographics align with recreational spaces, such as parks. A subsequent spatial autocorrelation analysis using ArcGIS further affirms the role of age in shaping PRC. With a Moran's index value of 0.351, a z-score of 25.745, and a p-value below 0.001, there's a pronounced positive spatial correlation between age and grids with high rewards.

Figure 6(b) presents the distribution of high-reward grids based on gender. Given the binary nature of gender, a Moran's index calculation wasn't feasible. As an alternative, we adopted a hotspot analysis. The results reveal that around 85.195% of all grids do not exhibit significant clustering, suggesting that gender does not have a significant impact on PRC.

We could extend the application of the reward function to encompass any district of Shenzhen. By computing the average reward values across various pedestrian demographics, we formulated Nanshan's reward map, as illustrated in Figure 6(c) (see appendix Figure A2 for Shenzhen's reward map). This map resonates strongly with Nanshan's distribution of built environment features. Moreover, it distinctly pinpoints the

district's hotspots, highlighting areas of particular appeal for walking travel. Such insights are invaluable to urban planners, pinpointing zones that require further development to address pedestrians' preferences and requirements.

### C. Causal Effects of Built Environment and Individual Features on PRC Preferences

To enhance the global interpretability of our model, we employ a graph-based causal discovery algorithm for result analysis. By leveraging the PC algorithm to craft a causal directional graph, we are able to intuitively discern the relationships between features and reward values, highlighting the direct influence of these features on PRC preferences. Figure 7 presents a causal directional graph mapping the association among built environment and individual features with PRC preferences.

Figure 7 shows that the 'gender' feature stands isolated with the causal graph. This suggests that gender does not exert direct influence on PRC. On the other hand, the 'age' feature showcases intricate interactions: those of middle-aged demographics seem influenced by features tied to enterprise, while the elderly gravitates towards tourist attraction features. This accentuates the diverse route choice preference across age demographics.

As to built environment features in Figure 7, we discern that factors such as land-use mix, living services, and city branches have a positive association with PRC preferences. Conversely, the open space ratio displays a pronounced negative correlation. Urban areas characterized by diverse land-use often have richer urban functions, enabling pedestrian easy access to various services and resources. The positive influence of city branch on PRC preferences can be attributed to its role in enhancing accessibility, making it easier for pedestrians to reach their destinations. In contrast, a higher open space ratio may signify areas predominantly serving a singular urban function, potentially lacking in entertainment and leisure options, thus explaining the inverse relationship between spatial openness and PRC behaviors.

Our adoption of the graph-based causal analysis not only offers an intuitive visualization of variable interactions but also deepen our understanding of these intricate relationships. For instance, Figure 7 reveals that while population density does not directly influence the PRC preferences, it exerts a mediating effect through its influence on land-use mix. Thus, when we probe the multifaceted interactions among features, it's vital to account for diverse underlying mechanisms. A mere adjustment to one variable might not directly affect PRC preferences, emphasizing the need for a comprehensive understanding. Earlier explanatory algorithms were primarily centered on the effects of modifying a singular feature on PRC preferences, referred to as the reward value. In contrast, graph-based causal discovery method offers a deeper insight into the complex dynamics among features.

Once directional relationships are established, the higher-order partial correlation coefficient serves as a metric for the intensity of causal relationships. This metric facilitates a detailed exploration of whether features exert positive or

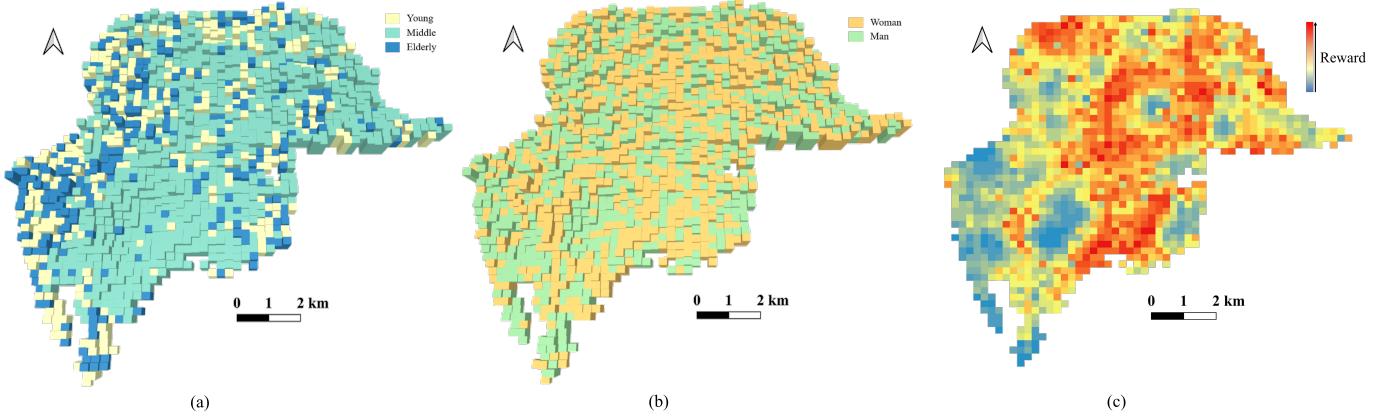


Fig. 6. Distribution of high-reward grids across different attributes (a) Distribution of high reward grids based on age (b) Distribution of high reward grids based on gender (c) Distribution of reward grids according to the average value of individual characteristics

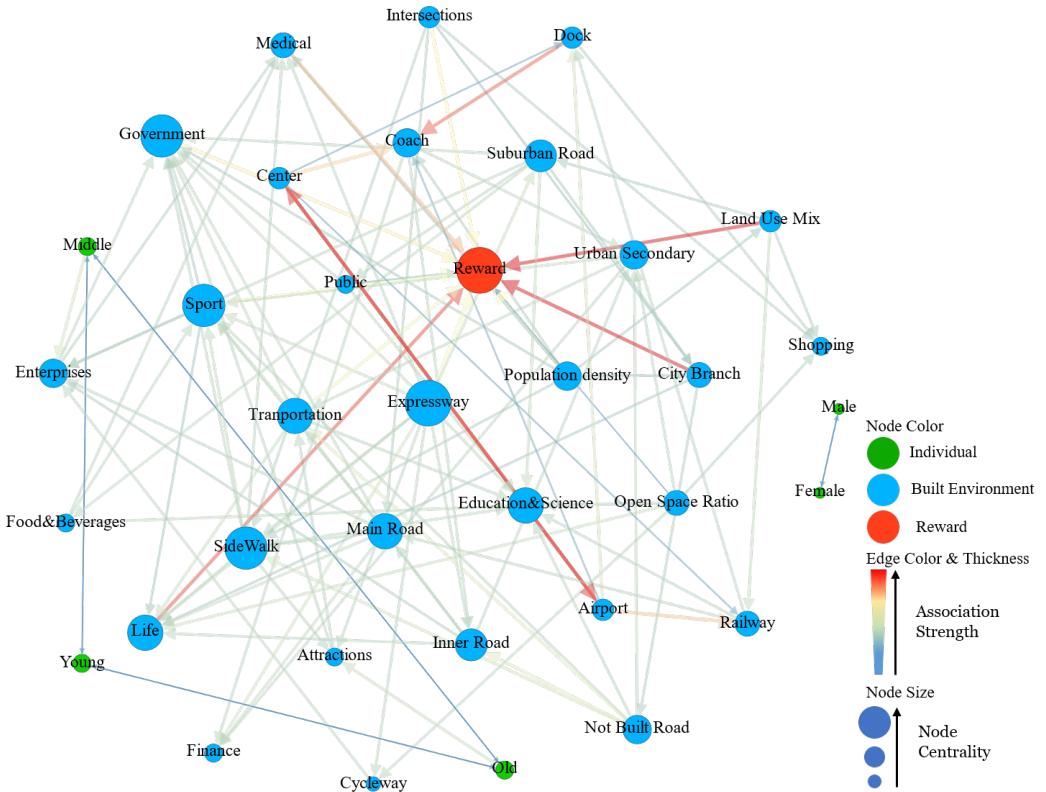


Fig. 7. Causal direction graph (node colors denote feature attributes, node size indicates centrality, edge color and thickness reflect association strength, edge arrows reflect causal direction)

negative influences on PRC preferences. We utilize a heatmap in Figure 8 to visually represent the strength of these causal linkages. Blank regions within the heatmap denote features whose correlation strength didn't meet the significance threshold. A key advantage of this analytical approach is its capacity to quantify the intensity of causal relationships rather than merely establishing their existence.

In Figure 8, the bottom row clearly illustrates the features that significantly impact the reward value, along with their corresponding magnitudes. It is noteworthy that the land-use mix, city branch, and life service exert a potent positive influence on PRC preferences, with causal strengths measured

at 0.97, 0.97, and 0.88 respectively. These values suggest that diverse land use, excellent urban accessibility, and a wealth of living facilities all have a positive relationship with PRC. Moreover, high population density, often indicative of enhanced social interactions and diverse activity choices, also positively correlates with PRC preferences. The positive impacts of medical services, governmental services, and transport services emphasize the pivotal role of public service facilities in enhancing pedestrians' travel satisfaction. Analyzing these key features provides valuable insights, guiding strategies for more effective urban planning and resource allocation.

Conversely, as illustrated in Figure 8, both the open space

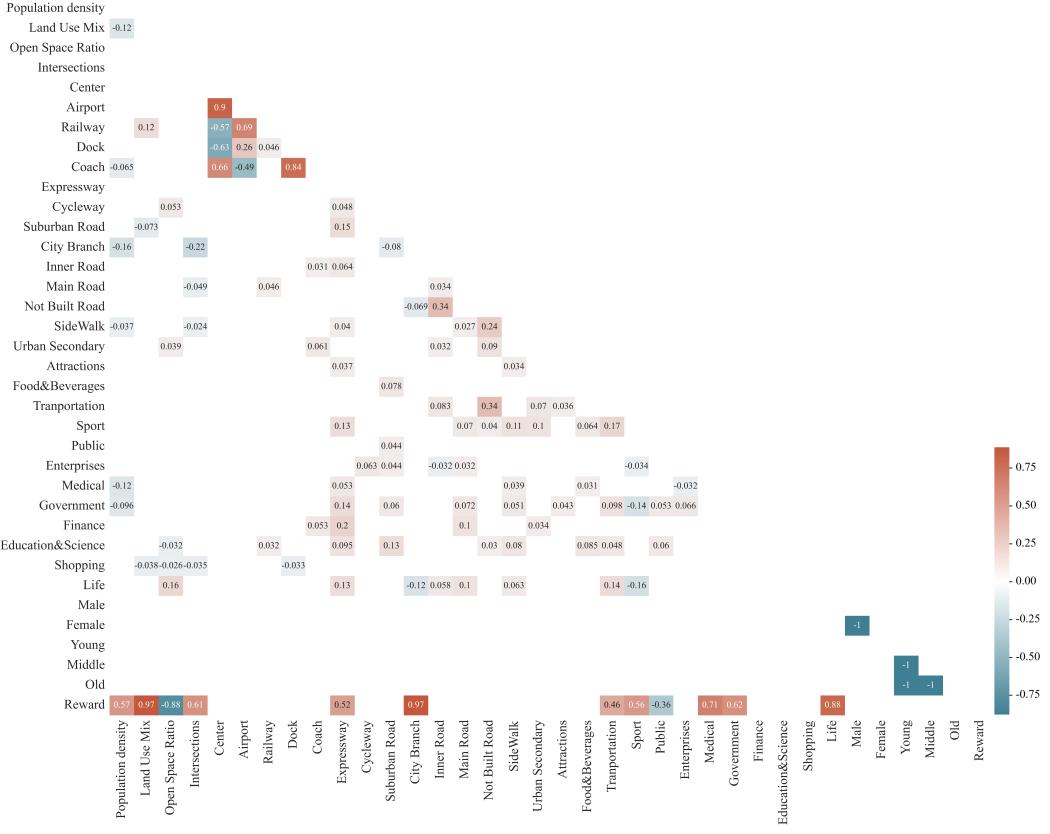


Fig. 8. Causal Strength Graph (red grids denote positive causal relations and blue grids signify negative ones between features)

ratio and public facilities present as features with negative influences. The negative weight for open space ratio suggests that expansive areas of undeveloped urban spaces might limit the availability of recreational and leisure choices for pedestrians, negatively impacting their PRC preferences. Regarding the negative correlation with public facilities, a deeper probe into specific causes could yield strategies to transform this negative effect into a positive influence.

## VI. CONCLUSION AND DISCUSSIONS

In this study, we introduce MEDIRL-IC, an advanced framework that has rarely been employed for analyzing PRC. Our method improves PRC analysis by incorporating the advantages of traditional discrete choice analysis and deep neural networks, enhanced with an inverse learning mechanism. Utilizing mobile signaling trajectory data from Shenzhen, we reveal complex patterns of PRC among different demographics, highlighting the influence of environmental and individual factors. We applied our methodology to ascertain reward functions tailored to diverse pedestrian demographics. This allowed us to delve into how route choice preferences are molded by both environmental and personal factors. With the reward function as our foundation, we generated predictive trajectories for pedestrians leveraging an enhanced  $A^*$  algorithm. Empirical results reveal that our model exhibits superior trajectory prediction accuracy, surpassing baseline models. Furthermore, the integration of a graph-based causal discovery algorithm augments our model's interpretability,

offering a comprehensive understanding of the intricate interplay between features and PRC preferences. Conclusively, our research underscores the pivotal roles played by varied land use, urban connectivity, and comprehensive living amenities in determining PRC behaviors, providing crucial insights for informed urban planning and policy decisions.

This article focuses on providing a unified methodological framework for analyzing PRC among diverse populations. However, it acknowledges specific limitations in its practical applicability. First, despite selecting features that have shown significant effects in previous studies, the constraints of the mobile signaling data set limited our ability to comprehensively encompass all potential factors affecting PRC. Consequently, this study only considers built environment features, individual socio-demographic features, and trip-based features such as road length and density. While our model is highly scalable and can be expanded in the future to incorporate a more comprehensive range of influencing factors. Second, our trajectory dataset only covers the Nanshan District, which may not adequately represent the PRC behaviors elsewhere [68]. Future research could expand the dataset's scope and compare the heterogeneity of PRC behaviors from different regions. Third, our analysis is confined to a grid-based rather than a link-based approach. This design choice, while facilitating modes of travel with higher spatial freedom, such as walking, does not account for the dynamic and complex nature of individual movements and interactions within the road network. Recognizing this, we propose that future efforts should explore

a link-based model to capture more nuanced patterns of route selection. Finally, while our model can effectively capture the intricate non-linear relationships between features, accurately interpreting these relationships remains challenging. Subsequent studies could delve deeper into the spatial causality of environmental features and the temporal causality associated with sequential decisions, to gain a deeper understanding of the underlying mechanisms behind PRC.

## APPENDIX

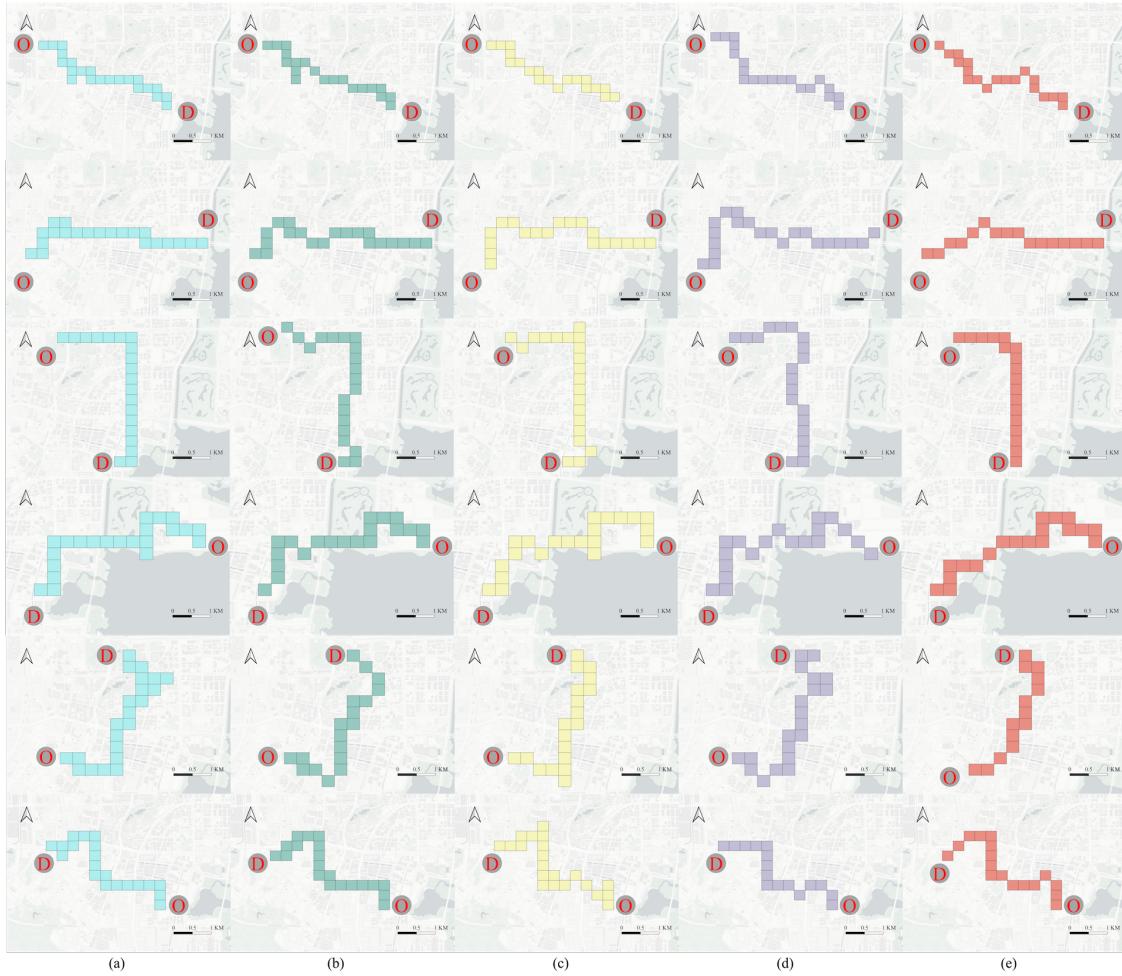


Fig. A1. Comparison between real trajectories and model-predicted trajectories (a) Actual trajectory (b) Trajectory generated by Recursive Logit (c) Trajectory generated by MaxEnt IRL (d) Trajectory generated by MEDIRL-IC (e) Trajectory generated by DNN-PSL.

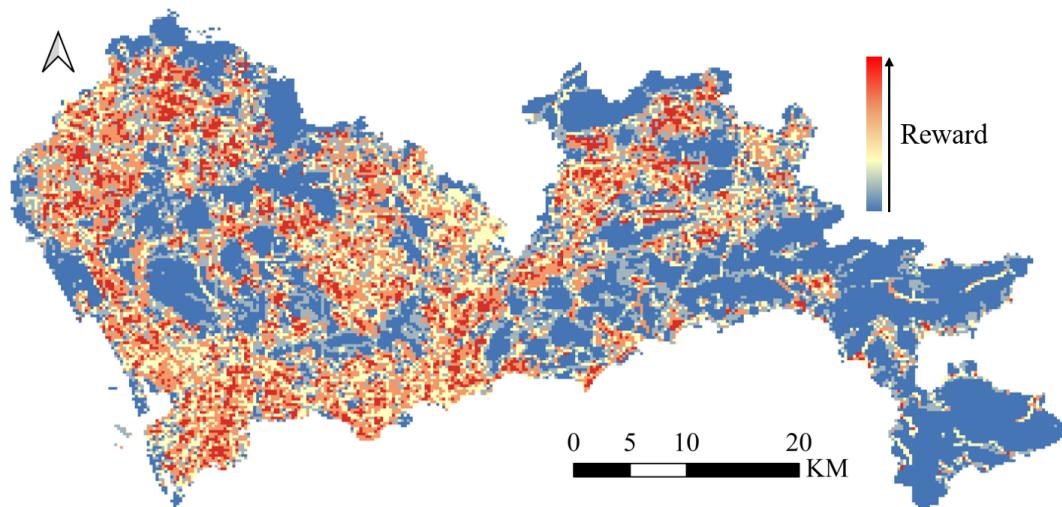


Fig. A2. The average reward map of Shenzhen: identifying hotspots of pedestrians' slow travel

## REFERENCES

- [1] T. A. Litman, "Economic value of walkability," *Transportation Research Record*, vol. 1828, no. 1, pp. 3–11, 2003.
- [2] S. P. Hoogendoorn and P. H. Bovy, "Pedestrian route-choice and activity scheduling theory and models," *Transportation Research Part B: Methodological*, vol. 38, no. 2, pp. 169–190, 2004.
- [3] E. Alvarez and J. G. Brida, "An agent-based model of tourism destinations choice," *International Journal of Tourism Research*, vol. 21, no. 2, pp. 145–155, 2019.
- [4] G. Vizzari, L. Crociani, and S. Bandini, "An agent-based model for plausible wayfinding in pedestrian simulation," *Engineering applications of artificial intelligence*, vol. 87, p. 103241, 2020.
- [5] D. Quercia, R. Schifanella, and L. M. Aiello, "The shortest path to happiness: Recommending beautiful, quiet, and happy routes in the city," in *Proceedings of the 25th ACM conference on Hypertext and social media*, 2014, pp. 116–125.
- [6] S. P. Hoogendoorn and P. H. Bovy, "Dynamic user-optimal assignment in continuous time and space," *Transportation Research Part B: Methodological*, vol. 38, no. 7, pp. 571–592, 2004.
- [7] R. Basu and A. Sevtsuk, "How do street attributes affect willingness-to-walk? city-wide pedestrian route choice analysis using big data from boston and san francisco," *Transportation research part A: policy and practice*, vol. 163, pp. 1–19, 2022.
- [8] N. Basu, M. M. Haque, M. King, M. Kamruzzaman, and O. Oviedo-Trespalacios, "A systematic review of the factors associated with pedestrian route choice," *Transport reviews*, vol. 42, no. 5, pp. 672–694, 2022.
- [9] M. Bierlaire, "Discrete choice models," in *Operations research and decision aid methodologies in traffic and transportation management*. Springer, 1998, pp. 203–227.
- [10] M. Fosgerau, E. Frejinger, and A. Karlstrom, "A link based network route choice model with unrestricted choice set," *Transportation Research Part B: Methodological*, vol. 56, pp. 70–80, 2013.
- [11] S. Ermon, Y. Xue, R. Toth, B. Dilkina, R. Bernstein, T. Damoulas, P. Clark, S. DeGloria, A. Mude, C. Barrett *et al.*, "Learning large-scale dynamic discrete choice models of spatio-temporal preferences with application to migratory pastoralism in east africa," in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 29, no. 1, 2015.
- [12] W. Zhang, D. Lu, H. Liu, and B. Li, "Varying built environment contexts and trip chain decisions: A multinomial-choice gradient boosting decision trees analysis," *Travel Behaviour and Society*, vol. 34, p. 100684, 2024.
- [13] H. Wen, Y. Lin, H. Wan, S. Guo, F. Wu, L. Wu, C. Song, and Y. Xu, "Deeprouute+: Modeling couriers' spatial-temporal behaviors and decision preferences for package pick-up route prediction," *ACM Transactions on Intelligent Systems and Technology (TIST)*, vol. 13, no. 2, pp. 1–23, 2022.
- [14] N. Mou, Q. Jiang, L. Zhang, J. Niu, Y. Zheng, Y. Wang, and T. Yang, "Personalized tourist route recommendation model with a trajectory understanding via neural networks," *International Journal of Digital Earth*, vol. 15, no. 1, pp. 1738–1759, 2022.
- [15] F. Simini, G. Barlacchi, M. Luca, and L. Pappalardo, "A deep gravity model for mobility flows generation," *Nature communications*, vol. 12, no. 1, p. 6576, 2021.
- [16] K. Hidaka, K. Hayakawa, T. Nishi, T. Usui, and T. Yamamoto, "Generating pedestrian walking behavior considering detour and pause in the path under space-time constraints," *Transportation research part C: emerging technologies*, vol. 108, pp. 115–129, 2019.
- [17] Z. Zhao and Y. Liang, "A deep inverse reinforcement learning approach to route choice modeling with context-dependent rewards," *Transportation Research Part C: Emerging Technologies*, vol. 149, p. 104079, 2023.
- [18] S. Liu and H. Jiang, "Personalized route recommendation for ride-hailing with deep inverse reinforcement learning and real-time traffic conditions," *Transportation Research Part E: Logistics and Transportation Review*, vol. 164, p. 102780, 2022.
- [19] D. A. Hensher and W. H. Greene, "The mixed logit model: the state of practice," *Transportation*, vol. 30, pp. 133–176, 2003.
- [20] A. Rudenko, L. Palmieri, M. Herman, K. M. Kitani, D. M. Gavrila, and K. O. Arras, "Human motion trajectory prediction: A survey," *The International Journal of Robotics Research*, vol. 39, no. 8, pp. 895–935, 2020.
- [21] A. Alahi, K. Goel, V. Ramanathan, A. Robicquet, L. Fei-Fei, and S. Savarese, "Social lstm: Human trajectory prediction in crowded spaces," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 961–971.
- [22] B. Wang, Y. Guo, and Y. Chen, "Personalized path recommendation with specified way-points based on trajectory representations," in *2021 17th International Conference on Mobility, Sensing and Networking (MSN)*. IEEE, 2021, pp. 396–403.
- [23] T. Koch and E. Dugundji, "Limitations of recursive logit for inverse reinforcement learning of bicycle route choice behavior in amsterdam," *Procedia Computer Science*, vol. 184, pp. 492–499, 2021.
- [24] J. Broach, J. Dill, and J. Gliebe, "Where do cyclists ride? a route choice model developed with revealed preference gps data," *Transportation Research Part A: Policy and Practice*, vol. 46, no. 10, pp. 1730–1740, 2012.
- [25] G. Menghini, N. Carrasco, N. Schüssler, and K. W. Axhausen, "Route choice of cyclists in zurich," *Transportation research part A: policy and practice*, vol. 44, no. 9, pp. 754–765, 2010.
- [26] C. G. Prato, "Route choice modeling: past, present and future research directions," *Journal of choice modelling*, vol. 2, no. 1, pp. 65–100, 2009.
- [27] Z. Guo and B. P. Loo, "Pedestrian environment and route choice: evidence from new york city and hong kong," *Journal of transport geography*, vol. 28, pp. 124–136, 2013.
- [28] A. Sevtsuk, R. Basu, X. Li, and R. Kalvo, "A big data approach to understanding pedestrian route choice preferences: Evidence from san francisco," *Travel behaviour and society*, vol. 25, pp. 41–51, 2021.
- [29] E. Frejinger, M. Bierlaire, and M. Ben-Akiva, "Sampling of alternatives for route choice modeling," *Transportation Research Part B: Methodological*, vol. 43, no. 10, pp. 984–994, 2009.
- [30] Y. Oyama, "Capturing positive network attributes during the estimation of recursive logit models: A prism-based approach," *Transportation Research Part C: Emerging Technologies*, vol. 147, p. 104014, 2023.
- [31] J. Ding-Mastera, S. Gao, E. Jenelius, M. Rahmani, and M. Ben-Akiva, "A latent-class adaptive routing choice model in stochastic time-dependent networks," *Transportation Research Part B: Methodological*, vol. 124, pp. 1–17, 2019.
- [32] E. Andresen, M. Chraibi, and A. Seyfried, "A representation of partial spatial knowledge: a cognitive map approach for evacuation simulations," *Transportmetrica A: transport science*, vol. 14, no. 5-6, pp. 433–467, 2018.
- [33] R. S. Sutton and A. G. Barto, *Reinforcement learning: An introduction*. MIT press, 2018.
- [34] F. Wei, S. Ma, N. Jia *et al.*, "A day-to-day route choice model based on reinforcement learning," *Mathematical Problems in Engineering*, vol. 2014, 2014.
- [35] A. Y. Ng, S. Russell *et al.*, "Algorithms for inverse reinforcement learning," in *Icmi*, vol. 1, no. 2, 2000, p. 2.
- [36] P. Abbeel and A. Y. Ng, "Apprenticeship learning via inverse reinforcement learning," in *Proceedings of the twenty-first international conference on Machine learning*, 2004, p. 1.
- [37] B. D. Ziebart, A. L. Maas, J. A. Bagnell, A. K. Dey *et al.*, "Maximum entropy inverse reinforcement learning," in *Aaaai*, vol. 8. Chicago, IL, USA, 2008, pp. 1433–1438.
- [38] M. Pan, W. Huang, Y. Li, X. Zhou, Z. Liu, R. Song, H. Lu, Z. Tian, and J. Luo, "Dhp: Dynamic human preference analytics framework: A case study on taxi drivers' learning curve analysis," *ACM Transactions on Intelligent Systems and Technology (TIST)*, vol. 11, no. 1, pp. 1–19, 2020.
- [39] M. Zimmermann and E. Frejinger, "A tutorial on recursive models for analyzing and predicting path choice behavior," *EURO Journal on Transportation and Logistics*, vol. 9, no. 2, p. 100004, 2020.
- [40] M. Wulfmeier, P. Ondruska, and I. Posner, "Maximum entropy deep inverse reinforcement learning," *arXiv preprint arXiv:1507.04888*, 2015.
- [41] S. Liu, H. Jiang, S. Chen, J. Ye, R. He, and Z. Sun, "Integrating dijkstra's algorithm into deep inverse reinforcement learning for food delivery route planning," *Transportation Research Part E: Logistics and Transportation Review*, vol. 142, p. 102070, 2020.
- [42] V. Mnih, K. Kavukcuoglu, D. Silver, A. A. Rusu, J. Veness, M. G. Bellemare, A. Graves, M. Riedmiller, A. K. Fidjeland, G. Ostrovski *et al.*, "Human-level control through deep reinforcement learning," *nature*, vol. 518, no. 7540, pp. 529–533, 2015.
- [43] S. M. Lundberg and S.-I. Lee, "A unified approach to interpreting model predictions," *Advances in neural information processing systems*, vol. 30, 2017.
- [44] R. Sun, S. Hu, H. Zhao, M. Moze, F. Aioun, and F. Guillemand, "Human-like highway trajectory modeling based on inverse reinforcement learning," in *2019 IEEE Intelligent Transportation Systems Conference (ITSC)*. IEEE, 2019, pp. 1482–1489.
- [45] S. Greydanus, A. Koul, J. Dodge, and A. Fern, "Visualizing and understanding atari agents," in *International conference on machine learning*. PMLR, 2018, pp. 1792–1801.

- [46] J. Pearl, *Causality*. Cambridge university press, 2009.
- [47] C. Glymour and K. Zhang, "Review of causal discovery methods based on graphical models," *Frontiers in genetics*, vol. 10, p. 418407, 2019.
- [48] F. Wagner, F. Nachtigall, L. Franken, N. Milojevic-Dupont, R. H. Pereira, N. Koch, J. Runge, M. Gonzalez, and F. Creutzig, "A causal discovery approach to learn how urban form shapes sustainable mobility across continents," *arXiv preprint arXiv:2308.16599*, 2023.
- [49] R. S. Chauhan, C. Riis, S. Adhikari, S. Derrible, E. Zheleva, C. F. Choudhury, and F. C. Pereira, "Determining causality in travel mode choice," *arXiv preprint arXiv:2208.05624*, 2022.
- [50] W. Zhang and K. Ning, "Spatiotemporal heterogeneities in the causal effects of mobility intervention policies during the covid-19 outbreak: A spatially interrupted time-series (sits) analysis," *Annals of the American Association of Geographers*, vol. 113, no. 5, pp. 1112–1134, 2023.
- [51] P. Spirtes, C. N. Glymour, and R. Scheines, *Causation, prediction, and search*. MIT press, 2000.
- [52] D. M. Chickering, "Optimal structure identification with greedy search," *Journal of machine learning research*, vol. 3, no. Nov, pp. 507–554, 2002.
- [53] S. Shimizu, P. O. Hoyer, A. Hyvärinen, A. Kerminen, and M. Jordan, "A linear non-gaussian acyclic model for causal discovery," *Journal of Machine Learning Research*, vol. 7, no. 10, 2006.
- [54] P. Hoyer, D. Janzing, J. M. Mooij, J. Peters, and B. Schölkopf, "Nonlinear causal discovery with additive noise models," *Advances in neural information processing systems*, vol. 21, 2008.
- [55] M. Kalisch and P. Bühlman, "Estimating high-dimensional directed acyclic graphs with the pc-algorithm," *Journal of Machine Learning Research*, vol. 8, no. 3, 2007.
- [56] D. Gillis, A. J. Lopez, and S. Gautama, "An evaluation of smartphone tracking for travel behavior studies," *ISPRS International Journal of Geo-Information*, vol. 12, no. 8, p. 335, 2023.
- [57] M. Okmi, L. Y. Por, T. F. Ang, and C. S. Ku, "Mobile phone data: A survey of techniques, features, and applications," *Sensors*, vol. 23, no. 2, p. 908, 2023.
- [58] W. Zhang and C. Li, "Varying mobility adaptation to covid-19 of residents in urban villages and commodity housing communities: A quasi-experimental analysis in shenzhen, china," *Cities*, vol. 145, p. 104711, 2024.
- [59] V. Wickramasinghe and S. Dissanayake, "Pedestrians' route choice behavior for intra-cbd trips in a developing country," Tech. Rep., 2015.
- [60] W. Zhang, C. Liu, and H. Zhang, "Public acceptance of congestion pricing policies in beijing: The roles of neighborhood built environment and air pollution perception," *Transport Policy*, vol. 143, pp. 106–120, 2023.
- [61] C. Bafatakis, D. Duives, and W. Daamen, "Determining a pedestrian route choice model through a photo survey," Tech. Rep., 2015.
- [62] T.-H. T. Gim and J. Ko, "Maximum likelihood and firth logistic regression of the pedestrian route choice," *International regional science review*, vol. 40, no. 6, pp. 616–637, 2017.
- [63] C. P. Tribby, H. J. Miller, B. B. Brown, C. M. Werner, and K. R. Smith, "Analyzing walking route choice through built environments using random forests and discrete choice techniques," *Environment and Planning B: Urban Analytics and City Science*, vol. 44, no. 6, pp. 1145–1167, 2017.
- [64] K. Ito, T. G. Reardon, M. C. Arcaya, S. Shamsuddin, D. M. Gute, and S. Srinivasan, "Built environment and walking to school: findings from a student travel behavior survey in massachusetts," *Transportation research record*, vol. 2666, no. 1, pp. 78–84, 2017.
- [65] A. D. Marra and F. Cormann, "A deep learning model for predicting route choice in public transport," in *21st Swiss Transport Research Conference (STRC 2021)*. STRC, 2021.
- [66] R. Korbacher, H.-T. Dang, and A. Tordeux, "Predicting pedestrian trajectories at different densities: A multi-criteria empirical analysis," *Physica A: Statistical Mechanics and its Applications*, vol. 634, p. 129440, 2024.
- [67] B. Neal, S. Mittal, A. Baratin, V. Tantia, M. Scilicula, S. Lacoste-Julien, and I. Mitliagkas, "A modern take on the bias-variance tradeoff in neural networks," *arXiv preprint arXiv:1810.08591*, 2018.
- [68] C. Liu and W. Zhang, "Social and spatial heterogeneities in covid-19 impacts on individual's metro use: A big-data driven causality inference," *Applied Geography*, vol. 155, p. 102947, 2023.



**Boyang Li** received his B.E. degree in Remote Sensing Science and Technology from the School of Remote Sensing and Information Engineering at Wuhan University in 2022. He is currently pursuing his master's degree at the School of Urban Planning and Design at Peking University, with a research focus on smart cities and big data. He has a keen interest in the analysis of pedestrian spatial decision-making and spatial cognition based on artificial intelligence.



**Wenjia Zhang** received his Ph.D. in Community and Regional Planning from the University of Texas at Austin. He has held research positions at the University of North Carolina at Charlotte and Peking University Shenzhen Graduate School, and he currently serves as the director of the Behavioral and Spatial AI Lab. Since August 2024, Zhang has been a Full Professor at the College of Architecture and Urban Planning at Tongji University. His research expertise lies in the analysis of urban and regional spatial structures, human mobility behavior, and the cutting-edge integration of big data and AI technologies in urban studies.