Equity-Aware Cross-Graph Interactive Reinforcement Learning for Bike Station Network Expansion

Xi Yang
University of Connecticut
xi.yang@uconn.edu

Suining He*
University of Connecticut
suining.he@uconn.edu

Mahan Tabatabaie
University of Connecticut
mahan.tabatabaie@uconn.edu

ABSTRACT

Thanks to advances in the urban big data, the bike sharing, especially station-based bike sharing, has emerged as the important first-/last-mile connectivities in many smart cities. Bike station network (BSN) expansion recommendation, i.e., recommending placement locations of new stations, is essential for satisfying local mobility demands, enhancing the BSN service quality, and may significantly affect the resource fairness and accessibility of different communities in the neighborhood. Furthermore, the dynamic and complex urban mobility environments make the station placement highly challenging to satisfy the mobility needs.

To ease and facilitate the urban planning with awareness of mobility equity, we have designed and proposed CGIRL, a novel equity-aware **C**ross-**G**raph **I**nteractive **R**einforcement **L**earning approach for BSN expansion recommendation. Specifically, we have designed a novel reward function within our actor-critic reinforcement learning approach, jointly accounting for the local mobility, bike resource distribution equity, and accessibility of different socioeconomic groups to the expanded stations. To capture the policy of station decisions from the BSN deployment, we integrate the location graph's mobility and equity correlations across the city regions as a graph network, and design a novel cross-graph interaction network with embedding attention and sequential dependency that adaptively captures and interactively differentiates the correlations within station placement. Our extensive experimental studies upon a total of 393 (111 new) bike stations from New York City (NYC), Washington D.C. (DC), and Chicago have validated the effectiveness of CGIRL in the equity-aware BSN expansion recommendation.

CCS CONCEPTS

• **Information systems** → **Geographic information systems**.

KEYWORDS

Equity-aware recommendation, bike sharing network expansion, spatio-temporal graph attention, cross-graph interactive reinforcement learning

*Corresponding author

SIGSPATIAL '23, November 13–16, 2023, Hamburg, Germany

ACM ISBN 979-8-4007-0168-9/23/11…$15.00
https://doi.org/10.1145/3589132.3625588

ACM Reference Format:
Xi Yang, Suining He, and Mahan Tabatabaie. 2023. Equity-Aware Cross-Graph Interactive Reinforcement Learning for Bike Station Network Expansion . In *The 31st ACM International Conference on Advances in Geographic Information Systems (SIGSPATIAL '23), November 13–16, 2023, Hamburg, Germany.* ACM, New York, NY, USA, 12 pages. https://doi.org/10.1145/3589132.3625588

1 INTRODUCTION

Thanks to the advances in urban big data, bike sharing, particularly the station-based bike sharing [12], has proliferated in many cities. Considering each bike station as a node and their mutual trips as the edges, the system can form a bike station network (BSN), serving as the important first- and last-mile urban connectivities.

Due to the recent deployment success and facilitating urbanization, the bike sharing service providers and city planners are coordinating the BSN expansion [8, 18], determining the placement locations of new stations in the city regions beyond the existing network. For instance, the Indego bike sharing program in Philadelphia, Pennsylvania has added 25 new stations and 300 bikes in 2021 [1], followed by 30 new stations and 400 bikes in 2022. Similarly, Spain, France, and Romania in Europe are also embracing new bike sharing systems in the post-pandemic age [2]. Conventional BSN expansion largely relies on site survey, crowdsourcing, and public hearings, which can be often time-consuming and labor-intensive [5, 8]. In order to reduce the efforts of the city planners and BSN service providers and enable smart decision-making support, the goal of this paper is to design a data-driven BSN expansion recommendation system, as illustrated in Fig. 1, which enhances decision efficiency, BSN profitability, as well as socioeconomic equity.

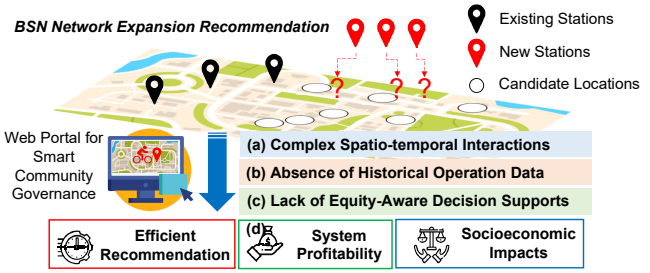


Figure 1: Motivation of BSN expansion recommendation.

Despite some prior efforts [8, 9, 13], it remains to be challenging due to the following three essential technical gaps:

(a) Complex spatio-temporal interactions across city regions: Each city region can demonstrate complex spatio-temporal interactions with others in the city due to urban functions as well as riders' commute purposes and preferences. For instance, the bike riders tend to commute between workplaces and residential

areas during weekdays, while visiting the points-of-interest (POIs) related to the entertainment or shopping activities on weekends [7]. Thus, the recommendation decision of one station's placement can have short-term, long-term, and sequential effects on the subsequent BSN expansions as well as the residents' mobility patterns. Existing approaches may not capture the complex, dynamic, and spatio-temporal interactions across the regions where stations are potentially deployed. With the large-scale mobility system records (e.g., bike trips, ride sharing trips, and human check-ins) and mobility environment (such as POIs) datasets generated from the urban location-based services, there exists an imperative need to design an effective, efficient, and adaptive BSN expansion recommendation mechanism that can ease and facilitate the decision-making process. These, however, have not been thoroughly explored in the prior studies [8, 9, 13].

(b) Mobility demand characterization without historical station placement: One may consider predicting the potential mobility demands for the station placement recommendation. However, for the BSN expansion problem, such prediction methods might not prevail due to the lack of historical bike usage in those candidate station locations. Toward a data-driven BSN expansion design, how to incorporate the heterogeneous information sources, such as POIs and exogenous patterns of other mobility modalities (e.g., taxi/ride sharing pick-ups and human mobile phone check-ins), is essential but challenging for valuating and determining the candidate station locations for effective BSN expansion recommendation and decision-making.

(c) Lack of equity-aware BSN expansion decision support modeling: Due to the nature of the first/last-mile urban connectivities, the results of the BSN expansion recommendation may have profound and complex impacts upon the mobility of various local communities. We note that bike sharing serves as an affordable and essential commute option, especially for historically disadvantaged communities or groups (e.g., low-income, historically underrepresented groups, and minorities), as well as accessibility to groceries, pharmacies, and many other life-essential urban facilities [17]. However, many existing methods largely focus on satisfying the historical mobility demands, while their fairness-agnostic settings might overlook the diverse communities' accessibility gaps and equity needs [26] for the station resources within their data-driven modeling and formulation. These may lead to the BSN expansion decisions biased toward the advantaged communities.

To fill the aforementioned gaps, we propose **CGIRL**, an equity-aware **C**ross-**G**raph **I**nteractive **R**einforcement **L**earning approach for BSN expansion recommendation. Our CGIRL takes in the distributions of existing bike stations, POIs, exogenous mobility datasets (e.g., taxi, ride sharing, and human check-ins), and the socioeconomic attributes in the BSN. We have designed a novel data-driven approach based on reinforcement learning (RL) for BSN expansion recommendation due to RL's efficiency in learning the complex, short-term, and long-term recommendation strategies, and adaptivity to the mobility environment that usually changes dramatically over time. Toward this framework, we have made the following three major technical contributions:

(1) Cross-Graph Interactions for BSN Expansion Policy Learning (Sec. 3): We have designed a cross-graph interaction policy learning mechanism within our RL to capture the correlations across the stations during the sequential process of station placement. The policy network of our CGIRL is based on a cross-graph interaction network (CGIN) that jointly characterizes the global topological and local node features of the formed BSN graph. CGIN captures the global topological features related to the riders' mobility interactions with the POI distributions through a graph convolution (GCN) mechanism to overcome the interaction modeling gap in the challenge (a). In the meantime, CGIN leverages a graph attention (GAT) mechanism to find the interdependencies of BSN with respect to the node features of the location graph, i.e., the local neighborhood characteristics (e.g., socioeconomic factors) and mobility demands (i.e., captured by other transportation platforms and human mobility data), in order to quantify the mobility needs to address the data absence in the challenge (b). Via our graph embedding attention (GEA), CGIRL further learns the cross-graph interactions, and provides the comprehensive and adaptive RL policy for BSN expansion recommendation.

(2) Equity-Aware BSN Expansion Recommendation without Historical Station Placement (Sec. 4): In order to approach the equity-agnostic concerns in the challenge (c), we have designed within our RL a multi-objective reward function named the Equity and Mobility Rewards (EMR) to enable the equity-aware BSN expansion. Our EMR consists of two major components, i.e., the mobility equity reward that concerns the riders' spatial connectivity and proximity, and the mobility demand reward that accounts for the mobility demands based on other transportation platforms and human mobility check-in data. This way, CGIRL provides the equity-aware BSN expansion recommendation that caters for the accessibility and mobility needs by the historically disadvantaged communities.

(3) Extensive BSN Data Analytics and Experimental Studies (Sec. 5): We have conducted extensive experimental studies based on the real-world BSN datasets to evaluate our CGIRL. Specifically, we have conducted multiple case studies upon the BSN expansion in three metropolitan cities in the U.S., i.e., New York City (NYC), Chicago, and Washington D.C. (DC). We have conducted BSN expansion recommendation studies on a total of 393 (111 new) bike stations, with exogenous datasets including 304,570 SafeGraph check-ins, 21,804,736 Uber/Taxi trips, 20,497 socio-economic records, and 42,791 POIs. Our experimental results have validated the effectiveness (e.g., in terms of matching the latent mobility demands) and fairness (e.g., in terms of enabling comprehensive accessibility for disadvantaged communities) of CGIRL in the BSN expansion recommendation, with 42.94% higher rewards on average compared with the baseline approaches.

2 SYSTEM, DATASETS, AND CONCEPTS

2.1 System Overview

We illustrate in Fig. 2 the system framework of CGIRL, which consists of three layers for BSN expansion. Specifically, at (a) data processing layer, given the selected city regions for future station expansion, CGIRL first finds the existing stations from the historical bike trip data, extracts the spatio-temporal and socioeconomic features at the neighborhood of candidate locations and existing stations as the CGIRL's inputs. Then at (b) cross-graph interactive

reinforcement learning layer, an actor-critic RL takes in the spatio-temporal features and identifies the BSN expansion policy. In particular, we form the location graphs and leverage a graph-based policy network (PN) as an actor to learn and capture the station placement policy, which maximizes the Equity and Mobility Reward (EMR). Another critic network (CN) interacts with PN by estimating the EMR from the BSN's mobility features to reduce the training variance. At (c) expansion recommendation layer, given the learned policy, CGIRL recommends the station placements through the web portal to facilitate city planning decision-making and mobility resource allocation processes.

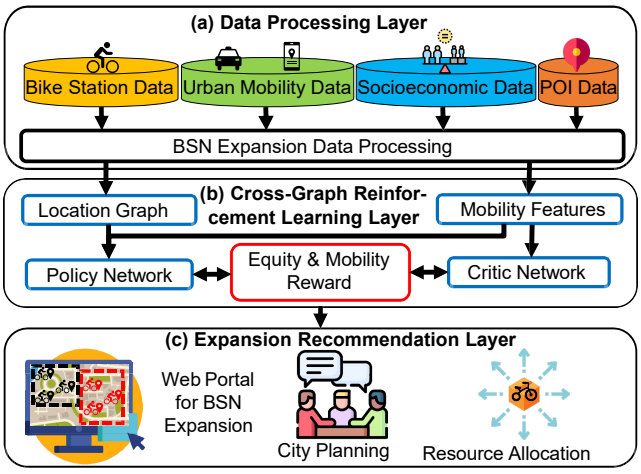


Figure 2: Overview of CGIRL framework.

2.2 Datasets Studied

To enable the BSN expansion recommendation, we develop our CGIRL based on the following real-world datasets:

(1) **Bike Stations**: We have collected BSN operation data of a total of 393 stations in NYC (Citi Bike; 08/2016–09/2016), Chicago (Divvy Bikes; 05/2017–06/2017), and DC (Capital Bikeshare; 03/2013–09/2013). We have identified the existing and new stations as follows for our prototype study: 68 and 71 in NYC, 118 and 25 in Chicago, and 96 and 15 in DC.

(2) **SafeGraph Check-ins.** We have harvested the human check-in data in NYC, Chicago, and DC from SafeGraph. We have identified the aggregate check-ins at the points-of-interest (POIs) and the census block groups from which these check-ins originate. We have collected 244,920 records of NYC, 46,340 records of Chicago, and 13,310 records of DC.

(3) **Uber Pick-ups and Taxi Trips**. We have collected 4,534,327 Uber pick-up records in NYC, 16,477,365 taxi trip records in Chicago, and 7,930,441 taxi trip records in DC.

(4) **Socio-economic Data.** To account for social equity, we have extracted the social ethnicity distribution, per capita income, and population holding bachelor's degrees in 2010 from the U.S. Census Bureau. We collected 11,089 socio-economic records in NYC, 6,205 in Chicago, and 3,203 in DC.

(5) **Points-of-Interest (POIs).** From OpenStreetMap, we have extracted the numbers of POIs with respect to each of the 7 major categories including sustenance, education, transportation, financial, healthcare, entertainment arts & culture, and others. We have collected 29,233 POIs in NYC, 7,955 in Chicago, and 5,603 in DC.

2.3 Problem Settings and Important Concepts

Our studies here consider the urban bike sharing station system as a network (graph), where each station is considered as a node, and their mutual edges are formed by the bike trips. This way, we form the bike station network (BSN) and our approach aims to determine the placement of new stations based on inputs of the historical locations of bike stations and other exogenous mobility, urban environment, and demographic factors. We present the important concepts for the formulation of CGIRL as follows.

(a) Candidate Grid Set. For ease of spatial processing, we divide the target city map into an $H \times W$ grid map, each grid of which is of size $d \times d$ m^2. We find the candidate grid set denoted as Ω^{New}, which is a geographic polygon from the grid map, as a set of the grids where new bike stations are recommended to be placed.

(b) Existing Stations and Candidate Locations. Let $\mathbf{Z}^{\text{Ex}}$ be the set of the existing bike stations. Through real-world data analytics, we have observed that a bike station is usually situated around the road intersection. Thus, we consider in this study the road intersections as candidate locations for the station placement. Let N_0 be the number of all candidate locations for placement of new stations inside the candidate grid Ω^{New}, and let $\mathbf{Z}_0$ be the initial set of all candidate locations, and $|\mathbf{Z}_0| = N_0$. We denote the index of an existing station/candidate location as $i \in \mathbf{Z}^{\text{Ex}} \cup \mathbf{Z}_0$, and the geographic coordinates as $[\text{lon}_i, \text{lat}_i]$. To form the spatio-temporal feature and quantify each reward, we find a $d \times d$ m^2 grid centered at $[\text{lon}_i, \text{lat}_i]$ as the grid neighborhood $\mathcal{N}_i$ for each station or candidate location i.

(c) Spatio-temporal Features. Recall that CGIRL needs to overcome the challenge (b), i.e., the absence of historical bike mobility demand. To estimate the potential mobility demand at each candidate location, we take into account other transportation platforms or mobility modalities, including (a) taxis/ride-sharing demands and (b) human mobile phone check-ins in the historical time period, as the exogenous inputs for CGIRL. Specifically, we find for each existing station/candidate location $i \in \mathbf{Z}^{\text{Ex}} \cup \mathbf{Z}_0$ the aggregate numbers of taxi/Uber pick-ups (demand) and human mobile phone check-ins from SafeGraph in the historical time period within its neighboring grids $\mathcal{N}_i$. We standardize each feature, and form a 2-D vector $\mathbf{F}_i$.

(d) RL Formulation. To ease the above BSN expansion modeling, our data-driven cross-graph interactive reinforcement learning formulation takes into account a total of N^{New} selection steps to recommend the candidate locations for the N^{New} stations in Ω^{New}, and we consider our RL formulation as follows.

- **Action:** At each selection step n, CGIRL makes an action, i.e., selecting a candidate location j to recommend the placement of a new station Γ_n.
- **State:** After imposing the action at step n, our CGIRL observes an updated set of stations $\mathbf{T}_n = \mathbf{Z}^{\text{Ex}} \cup \mathbf{Z}_n^{\text{New}}$, where $\mathbf{Z}_n^{\text{New}} \subset \mathbf{Z}_0$ denotes the new stations selected from the candidate locations $\mathbf{Z}_0$ by the n-th step. Here we have $\mathbf{Z}_n^{\text{New}} = \{\Gamma_1, \Gamma_2, \ldots, \Gamma_n\}$. Then we find the state at step n, denoted as

$\mathbf{S}_n$, based on the spatio-temporal features of the station set $\mathbf{T}_n$.

- **Policy:** The policy π of our RL formulation is defined as the strategy for choosing the candidate location j to place a new station Γ_n at the n-th step.

We note that this prototype considers an actor-critic RL framework consisting of a policy network and a critic network that interact with each other to reduce the variance of policy gradient and improve the training process. The policy network serves as the agent to impose the decision action based on π, and interacts with the emulated urban environment that consists of existing stations. At each selection step n, the emulated urban environment returns the reward $R_n(\Gamma_n)$ (detailed in Sec. 4).

(e) Station Placement Constraints. To prevent placing two stations at an excessively close neighborhood, we consider a spatial constraint that the candidate locations within a $c \times c$ m^2 grid neighborhood centered at a recommended station Γ_{n-1} will be excluded at the next selection step n. This way, we prevent two stations from being too close to each other and ensure the overall service coverage of the BSN [8]. At each step n, we update the candidate location set when a candidate station is placed, and the candidate location is removed from $\mathbf{Z}_0$. In other words, we have the updated candidate location set at the n-th step, $\mathbf{Z}_n = \mathbf{Z}_0 \setminus \left(\mathbf{Z}_{n-1}^{\text{New}} \cup \mathbf{C}_{n-1}\right)$, after deploying a new station $\mathbf{Z}_{n-1}^{\text{New}}$, where $\mathbf{C}_{n-1}$ denotes the set of locations within a $c \times c$ m^2 grid neighborhood of all the new stations $\mathbf{Z}_{n-1}^{\text{New}}$.

(f) Location Graph in BSN. For the ease of integrating the constraints explained above, we consider placement recommendation of the new stations as a sequential process in our cross-graph RL formulation. To preserve the mutual correlations across the existing and new stations in the sequential learning process, at each step n, we form a location graph in the BSN, denoted as $\Phi_n(\mathbf{S}_n, \mathbf{F}_j)$, where the graph nodes consist of the state $\mathbf{S}_n$ and features of one of the candidate locations $j \in \mathbf{Z}_n$, and the edges are given by the correlations between the nodes. Based on the formed graph, we design a cross-graph interaction network (CGIN; see Sec. 3.2) to capture the correlations of station placement with the selection step, and recommend the most likely location for station placement recommendation.

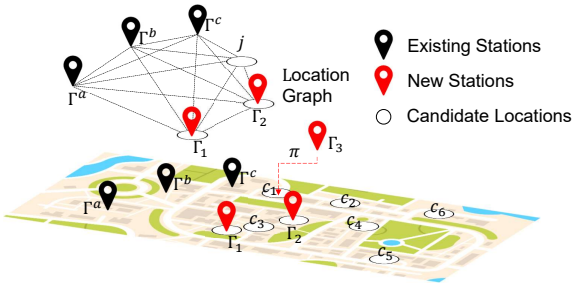


Figure 3: Illustration of our BSN expansion.

(g) Problem Definition. Given the existing bike stations, candidate locations, and their spatio-temporal features, the problem of our BSN expansion recommendation is to determine a set of N^{New} candidate locations within the Ω^{New} to place the corresponding N^{New} bike stations.

Fig. 3 illustrates an example of the BSN expansion process at a time step $n = 3$, where two new stations, $\mathbf{Z}_2^{\text{New}} = \{\Gamma_1, \Gamma_2\}$, are already placed at two candidate locations near the existing stations $\mathbf{Z}^{\text{Ex}} = \{\Gamma^a, \Gamma^b, \Gamma^c\}$, and we form the station set $\mathbf{T}_3 = \mathbf{Z}^{\text{Ex}} \cup \mathbf{Z}_2^{\text{New}} = \{\Gamma^a, \Gamma^b, \Gamma^c, \Gamma_1, \Gamma_2\}$. We form a location graph $\Phi_3(\mathbf{S}_3, \mathbf{F}_j)$ with the state

$$\mathbf{S}_3 = \left\{\mathbf{F}_{\Gamma^a}, \mathbf{F}_{\Gamma^b}, \mathbf{F}_{\Gamma^c}, \mathbf{F}_{\Gamma_1}, \mathbf{F}_{\Gamma_2}\right\}, \tag{1}$$

and features of one of the candidate location j. Using this location graph, CGIRL then finds the third new stations based on policy π at the time step 3, Γ_3, from the 5 candidate locations, $\mathbf{Z}_3 = \{c_1, c_2, c_4, c_5, c_6\}$. Here we exclude c_3 to avoid the overcrowded BSN.

3 CROSS-GRAPH INTERACTIVE RL DESIGNS

3.1 Overview of Model Architecture

Based on the important concepts defined in Sec. 2.3, toward a data-driven BSN expansion recommendation policy, we have designed within CGIRL an actor-critic cross-graph interactive reinforcement learning framework [4], as illustrated in Fig 4. Our CGIRL consists of a policy network (PN) serving as an actor and a critic network (CN) with an Equity and Mobility Reward (EMR) function. The PN characterizes the complex spatio-temporal interactions across regions to recommend locations for the new station, while the CN reduces the training variance [15]. By optimizing the EMR, our CGIRL framework enables an equity-aware decision and mobility demand characterization mechanism given no historical station placement on the candidate locations.

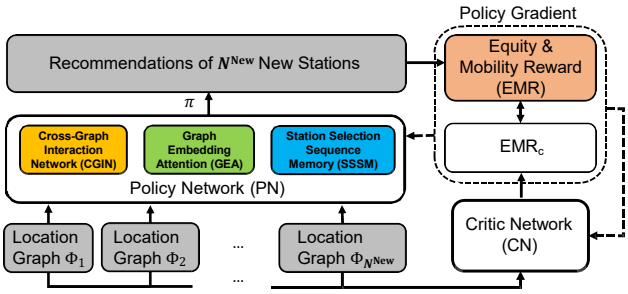


Figure 4: Overview of CGIRL model architecture.

• **Policy Network**. The PN learns the expansion policy π based on the observed state $\mathbf{S}_n$, as well as candidate locations $\mathbf{Z}_n$, at each selection step n. At each step, we form a location graph $\Phi_n(\mathbf{S}_n, \mathbf{F}_j)$, where $j \in \mathbf{Z}_n$. To recommend a candidate location for a placement action at step n, PN estimates the probabilities p_j of selecting every candidate location j based on the spatio-temporal features of the nodes in $\Phi_n(\mathbf{S}_n, \mathbf{F}_j)$. To this end, PN traces and updates the correlations of the expanding graph during the sequential location selection through its three main technical components (Fig. 5):

(1) **Cross-Graph Interaction Network** (CGIN), consisting of a graph convolutional network (GCN) and a graph attention network (GAT), that captures the dynamic cross-graph correlations between the nodes of BSN (Sec. 3.2);
(2) **Graph Embedding Attention** (GEA), which further differentiates the contributions of generated embeddings from GCN and GAT to the resulting BSN expansion policy (Sec. 3.3);

(3) **Station Selection Sequence Memory** (SSSM), which memorizes the selection in the BSN expansion (Sec. 3.4).

Based on the probabilities p_j, the PN samples the candidate locations for BSN expansion during the training phase, or greedy policy during the testing phase. We estimate the EMR with the selected candidate locations.

• **Critic Network.** We have designed a critic network (CN) to jointly estimate the potential EMR based on the spatio-temporal features of the observed state and reduce the training variance [15]. CGIRL trains both CN and PN in a policy-gradient fashion [21] by maximizing our EMR, and finds the equity-aware BSN expansion policy π.

CN leverages the spatio-temporal features of all the stations in the state, i.e., $\mathbf{F}_k$, where $k \in \mathbf{T}_n$, to estimate the reward, EMR_c, which is compared against other EMR values returned from PN to reduce the variance of training. Details of CGIRL training can be referred to Appendix A.1. Here we use three 1D convolutional (Conv1D) layers with a kernel size of 1 to encode the spatio-temporal features of each existing station/candidate location into a hidden state, and then we generate the estimated reward EMR_c by averaging the hidden states over all stations, i.e., $\text{EMR}_c = \text{CN}(\mathbf{S}_n)$.

3.2 Cross-Graph Interaction Network (CGIN)

• **Motivations.** The potential reward including the bike demand of a candidate location is highly correlated with the bike stations in its neighborhood. Therefore, for the location graph $\Phi_n(\mathbf{S}_n, \mathbf{F}_j)$ at each step n, where $j \in \mathbf{Z}_n$, we generate a graph embedding for the candidate location. As illustrated in Fig. 5, we propose two aspects of correlations between nodes i.e., the global topological correlations and the local node feature correlations.

Specifically, the global topological correlations represent the bike riders' mobility trends due to the difference in POI between locations. The local node feature correlations represent the riders' mobility interactions with the neighborhoods' socioeconomic factors and other transportation platforms. These two correlations interact with, and are complementary to each other for CGIRL to comprehensively model and recommend the BSN expansion decisions.

• **Detailed Designs.** To enable this cross-graph interaction network (CGIN), we have designed a graph convolutional network (GCN) [11] to extract the mobility correlations and generate mobility embeddings, and a graph attention network (GAT) [22] to mine the node feature correlations from the node features, including taxi/ride-sharing pick-ups and visitors' check-ins, and generate node feature embeddings. To realize the above, we present the details of GCN and GAT as follows.

(a) Graph Convolutional Network (GCN): To capture the global topological correlation, we first construct an adjacency matrix $\mathbf{A}_n \in \mathbb{R}^{(|\mathbf{S}_n|+1)\times(|\mathbf{S}_n|+1)}$ for the location graph $\Phi_n(\mathbf{S}_n, \mathbf{F}_j)$. We find the pairwise POI correlations among the $(|\mathbf{S}_n| + 1)$ locations of the existing/deployed $|\mathbf{S}_n|$ stations and the new station j. Let $\boldsymbol{\rho}_j$ and $\boldsymbol{\rho}_k \in \mathbb{R}^T$ be the POI vectors of length T for locations j and k, where the t-th element represents the number of POI of category $t \in [1, T]$ (min-max normalized).

In the location graph, we find each adjacency matrix element $\mathbf{A}_n[j,k] = \mathbf{A}_n[k,j]$ ($j \in \mathbf{Z}_n$ and $k \in \mathbf{T}_n$), is given by the cosine similarity of POI distributions, $\boldsymbol{\rho}_j$ and $\boldsymbol{\rho}_k$, in their node neighborhood $\mathcal{N}_j$ and $\mathcal{N}_k$, respectively, i.e.,

$$\mathbf{A}_n[j,k] = \mathbf{A}_n[k,j] = \frac{\sum_t^T \boldsymbol{\rho}_j[t] \cdot \boldsymbol{\rho}_k[t]}{\sqrt{\sum_t^T \boldsymbol{\rho}_j[t]^2} \cdot \sqrt{\sum_t^T \boldsymbol{\rho}_k[t]^2}}. \tag{2}$$

In other words, a larger similarity value between two locations implies higher topological correlations between them. Our GCN further captures such a topological correlation within the formulation of CGIRL.

We obtain $\mathbf{A}_n[k,k']$ for two stations k and $k' \in \mathbf{T}_n$ in the same manner. We set all the diagonal elements of $\mathbf{A}_n$ as zeros. Then, the adjacency matrix $\mathbf{A}_n$ is normalized by its degree matrix, $\mathbf{D}_n$, i.e.,

$$\tilde{\mathbf{A}}_n = \mathbf{D}_n^{-\frac{1}{2}} \cdot (\mathbf{A}_n + \mathbf{I}) \cdot \mathbf{D}_n^{\frac{1}{2}}, \tag{3}$$

where $\mathbf{I}$ is the identity matrix.

Afterwards, we find the embeddings of all nodes at the l-th convolutional layer, $\mathbf{H}_n^{(l)}$, from the $(l-1)$-th layer by

$$\mathbf{H}_n^{(l)} = \tilde{\mathbf{A}}_n \cdot \mathbf{H}_n^{(l-1)} \cdot \mathbf{W}_n^{(l-1)}, \tag{4}$$

where $\mathbf{W}_n^{(l-1)} \in \mathbb{R}^{W^{(l-1)} \times W^{(l)}}$ is a learnable parameter matrix, and at the first (input) layer

$$\mathbf{H}_n^{(1)} = [\mathbf{F}_1, \ldots, \mathbf{F}_k, \ldots, \mathbf{F}_{|\mathbf{S}_n|}, \mathbf{F}_j], \tag{5}$$

and $W^{(1)} = 2$ is the length of the vector $\mathbf{F}_j$ or $\mathbf{F}_k$.

After L convolutional layers, we obtain the topological embeddings of j being the last element of $\mathbf{H}_n^{(L)} \in \mathbb{R}^W$, i.e.,

$$\mathbf{h}_j^{(\text{GCN})} = \mathbf{H}_n^{(L)}[|\mathbf{S}_n| + 1]. \tag{6}$$

(b) Graph Attention Network (GAT): The global topological correlations captured by GCN may not suffice to assist PN in distinguishing between candidate locations. The BSN expansion also has latent interactions with local node features within the location graph, including the socioeconomic factors and the mobility patterns within other transportation platforms and mobility modalities.

We further leverage a GAT to extract the interactions from the local node features, i.e., taxi/ride-sharing demand and human mobile phone check-ins, by the node attention mechanism. The node attention weights help encode the interactions between selected stations in terms of socioeconomic factors as well as shared demand characteristics.

Specifically, we first compute the correlation for the b-th attention head, $e_{j,k}^{(b)}$, between two nodes j and k by

$$e_{j,k}^{(b)} = \text{LeakyReLU}\left(\mathbf{v}^{(b)} \left[\mathbf{W}^{(b)}\mathbf{F}_j \,\middle\|\, \mathbf{W}^{(b)}\mathbf{F}_k\right]\right), \tag{7}$$

where $\mathbf{W}^{(b)} \in \mathbb{R}^{W \times 2}$ and $\mathbf{v}^{(b)} \in \mathbb{R}^{1 \times 2W}$ are the trainable weight matrices at the b-th attention head. We set $W^{(L)} = W$ for the ease of fusion of the embeddings from GCN and GAT.

Then, we apply a softmax function across all k in $\mathbf{T}_n$ to obtain the attention score between k and j, i.e.,

$$\alpha_{j,k}^{(b)} = \frac{\exp\left(e_{j,k}^{(b)}\right)}{\sum_{k \in \mathbf{T}_n} \exp\left(e_{j,k}^{(b)}\right)}. \tag{8}$$

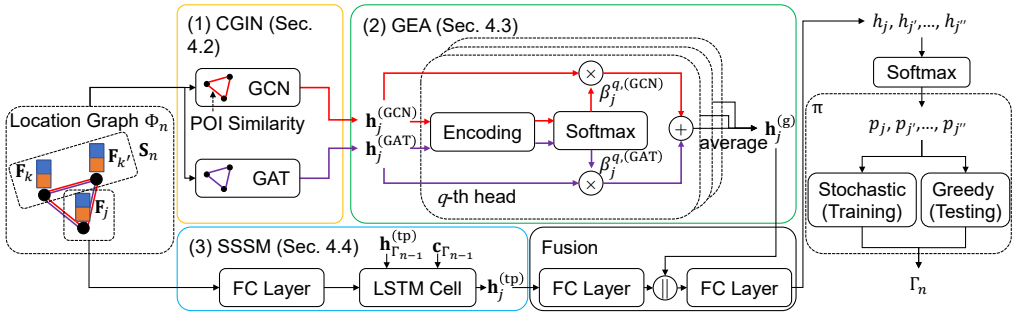


Figure 5: Illustration of the policy network (PN) design in CGIRL. The red lines represent the embeddings for topological correlations, and the purple lines for node feature correlations.

Then, we compute the weighted sum of all features of the state $\mathbf{S}_n$ and obtain the node feature embeddings of j, $\mathbf{h}_j^{(\text{GAT})}$, as the mean over all B attention heads, i.e.,

$$\mathbf{h}_j^{(\text{GAT})} = \frac{1}{B}\sum_{b}^{B} \tanh\left(\sum_{k\in\mathbf{T}_n} \alpha_{j,k}^{(b)}\mathbf{W}^{(b)}\mathbf{F}_k\right). \tag{9}$$

3.3 Graph Embedding Attention (GEA)

• **Motivations.** To capture and differentiate the interactions of the topological and node feature correlations obtained from GCN and GAT, we have further designed a graph embedding attention mechanism as illustrated in Fig. 5(b).

• **Detailed Designs.** Considering the topological embeddings for a candidate intersection j, $\mathbf{h}_j^{(\text{GCN})}$, we encode it into a hidden feature at the q-th attention head, $\epsilon_j^{(q,\text{GCN})}$, by

$$\epsilon_j^{(q,\text{GCN})} = \left(\mathbf{v}^{(q)}\right)^{\top} \cdot \tanh\left(\mathbf{V}^{(q)}\mathbf{h}_j^{(\text{GCN})} + \mathbf{b}^{(q)}\right), \tag{10}$$

where $\mathbf{v}^{(q)} \in \mathbb{R}^V$ and $\mathbf{V}^{(q)} \in \mathbb{R}^{V\times W}$ are trainable parameters. We similarly obtain $\epsilon_j^{(q,\text{GAT})}$ for the node feature embeddings.

Then, we take a softmax function to calculate the attention scores for the topological embeddings, i.e.,

$$\beta_j^{(q,\text{GCN})} = \frac{\exp\left(\epsilon_j^{(q,\text{GCN})}\right)}{\exp\left(\epsilon_j^{(q,\text{GCN})} + \epsilon_j^{(q,\text{GAT})}\right)}, \tag{11}$$

and the feature embeddings, $\beta_j^{(q,\text{GCN})}$, by the same approach.

We obtain the output of the GEA, $\mathbf{h}_j^{(\text{g})}$, as the averaged weighted sum of $\mathbf{h}_j^{(\text{GCN})}$ and $\mathbf{h}_j^{(\text{GCN})}$ over all Q attention heads, i.e.,

$$\mathbf{h}_j^{(\text{g})} = \frac{1}{Q}\sum_{q}^{Q}\left(\beta_j^{(q,\text{GCN})}\mathbf{h}_j^{(\text{GCN})} + \beta_j^{(q,\text{GAT})}\mathbf{h}_j^{(\text{GAT})}\right). \tag{12}$$

3.4 Station Selection Sequence Memory (SSSM)

• **Motivations.** Besides the spatio-temporal features, we note the station placement recommendation of one station will have short- and long-term impacts upon the other stations' decisions. Therefore, we further leverage the long short-term memory (LSTM) as a sequence memory to capture the short-term and long-term impacts.

• **Designs.** Recall that at a selection step $(n-1)$, we select Γ_{n-1} to build a new station. We first adopt a fully connected layer to obtain the hidden feature of the spatio-temporal features of $\mathbf{F}_j$, i.e.,

$$\widehat{\mathbf{F}}_j = \mathbf{U}_1 \cdot \mathbf{F}_j + \mathbf{b}^{u_1}, \tag{13}$$

where $\mathbf{U}_1 \in \mathbb{R}^{W_1\times 2}$ and $\mathbf{b}^{u_1} \in \mathbb{R}^{W_1}$.

Then, we adopt a long-short term memory (LSTM) cell to encode the sequential dependency of decisions at the step n on that of the previous step $n-1$ as illustrated in Fig. 5(c). Specifically, we obtain the sequential hidden state, $\mathbf{h}_j^{(\text{tp})} \in \mathbb{R}^{W_2}$, and cell state, $\mathbf{c}_j \in \mathbb{R}^{W_2}$, for the next step n from $\widehat{\mathbf{F}}_j$ and $\mathbf{h}_{\Gamma_{n-1}}^{(\text{tp})}$ and $\mathbf{c}_{\Gamma_{n-1}}$ of step $n-1$, i.e.,

$$\mathbf{h}_j, \mathbf{c}_j = \text{LSTM}\left(\mathbf{h}_{\Gamma_{n-1}}^{(\text{tp})}, \mathbf{c}_{\Gamma_{n-1}}, \widehat{\mathbf{F}}_j\right). \tag{14}$$

To fuse the sequential hidden state, $\mathbf{h}_j^{(\text{tp})}$, with the graph embeddings, $\mathbf{h}_j^{(\text{g})}$, we first adopt a fully connected layer to generate the updated sequential hidden state $\widehat{\mathbf{h}}_j^{(\text{tp})}$ from $\mathbf{h}_j^{(\text{tp})}$ so that its dimension match the graph embeddings, i.e.,

$$\widehat{\mathbf{h}}_j^{(\text{tp})} = \mathbf{U}_2 \cdot \mathbf{h}_j^{(\text{tp})} + \mathbf{b}^{u_2}, \tag{15}$$

where $\mathbf{U}_2 \in \mathbb{R}^{W\times W_2}$ and $\mathbf{b}^{u_2} \in \mathbb{R}^{W}$ are trainable parameters.

Then, we get the output embedding h_j from the concatenation of $\mathbf{h}_j^{(\text{g})}$ and $\widehat{\mathbf{h}}_j^{(\text{tp})}$ via a fully connected layer, i.e.,

$$h_j = \mathbf{U}_2' \cdot \left(\mathbf{h}_j^{(\text{g})} \,\middle\|\, \widehat{\mathbf{h}}_j^{(\text{tp})}\right) + \mathbf{b}^{u_2'}, \tag{16}$$

where $\mathbf{U}_2' \in \mathbb{R}^{1\times 2W}$ and $\mathbf{b}^{u_2'} \in \mathbb{R}^{1}$ are trainable parameters.

We calculate h_j for all candidate locations $j \in \mathbf{Z}_n$, and the probability of selecting j at step n is

$$p_j = \frac{\exp(h_j)}{\sum_{j\in\mathbf{Z}_n}\exp(h_j)}. \tag{17}$$

We note that during the model training CGIRL samples the candidate location, Γ_n, based on p_j. During the model testing, CGIRL recommends the highest p_j to place a new station, Γ_n, at step n.

4 REWARD DESIGNS FOR RL

Our proposed equity and mobility reward (EMR) function consists of the mobility equity reward (Sec. 4.1), which enables the equity awareness of the BSN expansion, and the mobility demand reward (Sec. 4.2), which accounts for the bike sharing operation meeting the demands of the neighborhoods. Both components will be further integrated into EMR (Sec. 4.3) for equity-aware BSN expansion recommendation.

4.1 Mobility Equity Reward

We take into account the socio-economic attributes (e.g., social ethnicity, income, and education level in our studies) in CGIRL. For each attribute, similar to prior fairness studies [17, 19], we consider

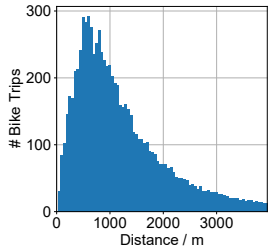


Figure 6: Estimated number of bike trips as the function of geo-distance in NYC.

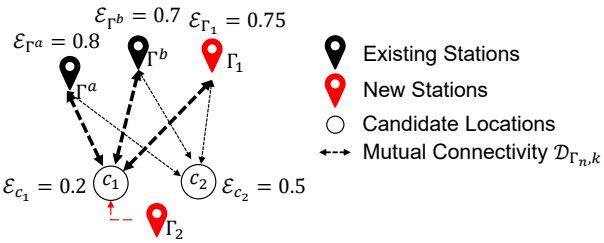


Figure 7: Illustration of Connectivity Equity Reward. Thicker dash lines represent higher mutual connectivity, $\mathcal{D}_{\Gamma_n,k}$.

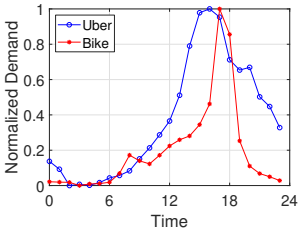


Figure 8: Normalized hourly demands of Uber vs. bike trips in NYC.

the advantaged and disadvantaged communities, denoted as $\mathbf{c}^{(+)}$ and $\mathbf{c}^{(-)}$.

Specifically, respectively for social ethnicity, income level, and education level, CGIRL accounts for the corresponding disadvantaged communities as *(i)* the historical minorities other than Caucasian, *(ii)* the median household income below 80% of the city median income level, and *(iii)* the populations with education levels lower than 40^{th} percentile of the entire city.

(a) Connectivity Equity (CE) Reward. With the above, we first design the connectivity equity $R^{(\text{CE})}$ that accounts for the potential bike riders from different communities, such that the placement of a new station can enhance the spatial connectivities of $\mathbf{c}^{(-)}$ with $\mathbf{c}^{(+)}$. By ensuring the spatial connectivities, the BSN can provide more accessibility to the bike riders from disadvantaged communities in accessing more economic opportunities from the business-centric regions or job centers.

Specifically, for each existing station/candidate location i, we find all the mobile phone check-ins from the SafeGraph dataset within its neighborhood $\mathcal{N}_i$. The anonymized and sanitized check-ins provide the aggregate number of the mobile phone users and their home addresses from different census block groups. Then, we identify from all check-ins the percentage of mobile phone check-ins whose home addresses are from the census block groups where $\mathbf{c}^{(-)}$ predominantly reside, denoted as $\mathcal{E}_i \in [0, 1]$. For instance, when 60 out of 100 check-in users are from the census block groups where $\mathbf{c}^{(-)}$ predominate, we have $\mathcal{E}_i = 60\%$.

We take into account the relationship between the bike trips and the geo-distance of the locations to characterize the spatial connectivity of the city locations. We design a data-driven approach to characterize the spatial connectivity between two communities at locations i and i' by their mutual bike trips. Based on the historical bike trips, we estimate the potential bike trips between two locations by their mutual geo-distance and characterize their spatial connectivity. We let $\mathcal{D}_{i,i'}$ be the number of bike trips between two locations i and i', and $\mathcal{G}_{i,i'}$ be their mutual geo-distance (in meters). Specifically, given all the existing stations $\mathbf{Z}^{\text{Ex}}$, we find their mutual geo-distance pairs, and form the histograms of pairwise distances with each bin width as 50m. Then, we calculate the number of historical average station-to-station bike trips in each bin and use it to estimate the potential number of bike trips for an input distance $\mathcal{G}_{i,i'}$.

For instance, based on such relationship in NYC shown in Fig. 6, we estimate the potential station-to-station bike trips for two locations that are 600m away as 292. We can observe that the average number of bike trips is strongly correlated with the mutual geo-distance with a peak at around 500m, which indicates a stronger bike mobility trend for stations that are around 500m away than stations that are closer or further.

Given the approximated mutual connectivity between the selected candidate location Γ_n and $k \in \mathbf{T}_n$, $\mathcal{D}_{\Gamma_n,k}$, we define the connectivity equity reward for a new station Γ_n, i.e.,

$$R_n^{(\text{CE})}(\Gamma_n) = \frac{1}{|\mathbf{S}_n|} \sum_{k \in \mathbf{T}_n} \left(1 + \exp\left(\phi^{(\text{CE})} \cdot |\mathcal{E}_{\Gamma_n} - \mathcal{E}_k|\right)\right) \cdot \mathcal{D}_{\Gamma_n,k},$$

where $\phi^{(\text{CE})}$ serves as the weight parameter adjusting the sensitivity of $R_n^{(\text{CE})}$. In other words, we consider the CE reward increases given higher $\mathcal{D}_{\Gamma_n,k}$ between the new station's and other locations with the larger social disparity, $|\mathcal{E}_{\Gamma_n} - \mathcal{E}_k|$.

As illustrated in Fig. 7, we select the candidate station location c_1, instead of c_2, to deploy Γ_2 at step $n = 2$, mainly because there exists a larger difference between $\mathcal{E}_{c_1}$ and $\{\mathcal{E}_{\Gamma^a}, \mathcal{E}_{\Gamma^b}, \mathcal{E}_{\Gamma_1}\}$ in addition to larger mutual connectivity. We adopt the exponential function to capture the non-linear relations. We process $\mathcal{D}_{\Gamma_n,k}$ through min-max normalization. By maximizing the CE rewards, the station placement action selects Γ_n to strengthen its connectivities (in terms of bike trips) between the locations with potentially higher social disparity, and help enable improved accessibility to life-critical resources.

(b) Proximity Equity (PE) Reward. In addition to the communities where the check-ins might originate from, we further take into account the proximity of a new station to the existing ones where the disadvantaged community $\mathbf{c}^{(-)}$ predominantly reside. This way, the disadvantaged communities can benefit from shorter walking distances in accessing the mobility resources.

Specifically, we design the proximity equity (PE) reward, $R_n^{(\text{PE})}$, which takes in $\mathcal{P}_{\Gamma_n}$, i.e., the populations in $\mathcal{N}_{\Gamma_n}$, and the inverse of the geo-distances $\mathcal{G}_{\Gamma_n,k^{(-)}}$ to the existing stations $k^{(-)} \in \mathbf{T}_n$ where $\mathbf{c}^{(-)}$ predominantly reside. The PE reward is formally given by

$$R_n^{(\text{PE})}(\Gamma_n) = \left(1 + \mathcal{P}_{\Gamma_n}\right) \cdot \sum_{k^{(-)} \in \mathbf{T}_n} \frac{1}{\exp\left(\phi^{(\text{PE})} \cdot \mathcal{G}_{\Gamma_n,k^{(-)}}\right)}, \tag{18}$$

where $\phi^{(\text{PE})}$ is an adjustable parameter. We also adopt the exponential function to account for the non-linear relations. In other words, the PE reward increases as the new station Γ_n is closer to the densely populated regions with more disadvantaged communities.

4.2 Mobility Demand Reward

Due to the lack of historical bike usage data, it is extremely challenging to estimate the potential bike demand of the new station. Fig. 8 illustrates the normalized average hourly demand of Uber around a bike station against that of bike sharing of that station in NYC. We calculate the average hourly demand of two systems by taking the mean of their demands at each hour. We sample the Uber demand in a 500×500 m^2 grid centered at the bike station as an example for calculation. We can observe a noticeable correlation between the hourly demand of Uber and bike sharing. In addition, we have observed the similar trend between the number of check-ins from SafeGraph and the bike sharing demands. These studies motivate us to leverage the exogenous datasets, i.e., other transportation platforms (such as Uber and taxi rides) and human mobility data, to characterize the potential bike sharing demand.

Specifically, we take into account the taxi/ride-sharing pickups, O_{Γ_n}, and the aggregate number of the mobile phone user check-ins from SafeGraph, Λ_{Γ_n}, at the neighborhood $\mathcal{N}_{\Gamma_n}$ of a new station location Γ_n as the potential mobility demand for the station placement. Then, the mobility reward, denoted as $R_n^{(\text{Mob})}(\Gamma_n)$, for a new station Γ_n is given by

$$R_n^{(\text{Mob})}(\Gamma_n) = \left(O_{\Gamma_n}\right)^{\tau} + \left(\Lambda_{\Gamma_n}\right)^{\gamma}, \tag{19}$$

where we use τ and γ to control the feature scale. In our empirical studies, we have observed that the bike demand demonstrates linear correlations with taxi/Uber pickups and the square root of mobile user check-ins, and we set τ as 1 and γ as 0.5 in our studies.

4.3 EMR Integration

We finally obtain the following multi-objective function of EMR per selection step n, i.e.,

$$R_n(\Gamma_n) = \lambda^{(\text{CE})} R_n^{(\text{CE})}(\Gamma_n) + \lambda^{(\text{PE})} R_n^{(\text{PE})}(\Gamma_n) + \lambda^{(\text{Mob})} R_n^{(\text{Mob})}(\Gamma_n), \tag{20}$$

where $\lambda^{(\text{CE})}$, $\lambda^{(\text{PE})}$, $\lambda^{(\text{Mob})}$, and $\lambda^{(\text{Ac})}$ are the corresponding weight parameters for each component. We note that placement of N^{New} stations at any N^{New} locations in the region Ω^{New} will generate one EMR. In order to differentiate reward scales across different urban locations to facilitate the RL training, from each R_n we subtract a base reward R_n^0, i.e.,

$$R_n^0 = \frac{1}{Z_n} \sum_{j=1}^{Z_n} R_n(j), \tag{21}$$

which is the mean reward across all Z_n candidate locations.

The training process of CGIRL is to determine the BSN expansion recommendation policy π that maximizes the total equity and mobility rewards of all the N^{New} selection steps, i.e.,

$$\text{EMR} = \sum_{n=1}^{N^{\text{New}}} \left(R_n(\Gamma_n) - R_n^0 \right). \tag{22}$$

5 EXPERIMENTAL STUDIES

• **Comparison Schemes.** We compare our CGIRL with the following approaches (details of implementation can be referred to Appendix A.2).

(1) AC: we implement the Ant Colony (AC) as a heuristic approach for the optimization problem of the BSN expansion.
(2) GA: we adapt the Genetic Algorithm (GA) as another heuristic approach for the BSN expansion.
(3) REINFORCE [25]: a Monte-Carlo variant of policy gradient reinforcement learning framework.
(4) DPG [20]: a Deterministic Policy Gradient (DPG) algorithm where we change the policy to be deterministic by selecting the candidate location with the largest probability in the model training phase, and we add a Gaussian noise to the policy for better exploration.
(5) RL-Dense: leverages two fully connected layers to generate the probability of each candidate location from its spatio-temporal features.
(6) RL-Conv2D: leverages 2D convolutional (Conv2D) to capture the correlations of the graph formed by the existing/new stations and a candidate location.
(7) RL-ResNet: leverages two ResNet blocks [6] to capture the correlations of the locations of BSN.
(8) RL-GCN: where we use a GCN layer to generate the probabilities of candidate locations.
(9) RL-GAT: where we use a GAT layer to calculate the probabilities of candidate locations.
(10) SRL [24]: an attention-based sequential actor-critic reinforcement learning (SRL) framework for transportation network expansion.

Table 1: Overall performance comparison of EMRs.

Schemes	NYC			DC			Chicago		
	Ethn.	Inc.	Edu.	Ethn.	Inc.	Edu.	Ethn.	Inc.	Edu.
AC	7.93	12.05	12.32	4.99	6.93	5.53	11.36	15.39	15.15
GA	4.29	7.37	9.33	7.63	7.63	7.70	16.73	16.72	16.72
REINFORCE	8.98	11.29	15.54	6.28	6.28	6.27	15.30	12.86	12.85
DPG	13.77	16.01	22.29	7.49	7.50	7.49	16.19	16.18	16.18
RL-Dense	13.61	17.99	23.76	7.49	7.50	7.49	16.27	16.26	16.26
RL-Conv2D	14.64	18.12	23.88	7.49	7.50	7.49	16.29	16.28	16.28
RL-ResNet	11.05	15.21	21.28	7.58	7.59	7.26	6.66	14.20	6.66
RL-GCN	6.99	4.52	7.35	4.23	4.23	4.22	5.82	5.82	5.82
RL-GAT	15.19	14.44	19.03	7.01	7.01	7.01	16.12	16.35	16.11
SRL	13.66	17.01	22.49	7.45	7.45	7.45	13.97	13.97	13.97
CGIRL	**17.16**	**18.98**	**24.90**	**7.99**	**7.99**	**7.75**	**16.74**	**16.73**	**16.73**

• **Overall Comparison in Rewards.** We have shown in Table 1 the rewards (in terms of EMRs) of all the schemes regarding the three socioeconomic attributes (social ethnicity, income level, and education level) for each city. From the table, we observe an overall better reward of CGIRL for all attributes of all three cities. We can observe the heuristic approaches such as GA and AC cannot fully capture the spatio-temporal mobility, bike resources, and accessibility, thus yielding low rewards compared with other methods. RL-Dense only considers the spatio-temporal features of each candidate location, but fails to capture its interactions with the existing and new stations. RL-Conv2D and RL-ResNet treat the BSN as an image, which may underestimate the correlations across the entire BSN owing to the limited kernel size. REINFORCE's policy network may not differentiate the features of candidate locations, thereof degrading the effectiveness of probability calculation. Furthermore, without a critic network, REINFORCE may generate a large gradient variance.

We can observe that RL-GCN only captures the mobility trends driven by POI differences but fail to consider socioeconomic effects

based on node features. RL-GAT ignores the mobility correlations, degrading their model performance. In addition, all the schemes above fail to account for the effect of the preceding selection steps on the later ones. SRL, despite accounting for orders of the decisions, may not fully preserve the location correlations in the sequential process, thus leading to poorer expansion results. DPG has demonstrated worse performance than CGIRL owing to its poorer exploration of the action space with the nature of deterministic policy than the stochastic policy in CGIRL. CGIRL, however, leverages and differentiates the contributions of different correlations between the candidate locations through the interactions of GCN and GAT within CGIN, and takes into account the selection sequence memory by SSSM, thus enhancing the rewards of expansion decisions.

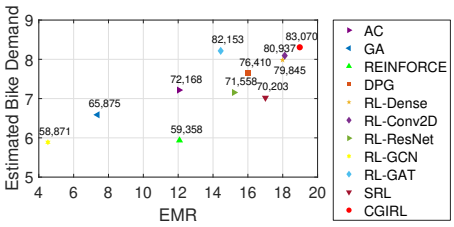


Figure 9: Estimated bike demand against EMRs of all the schemes for NYC (in terms of income level).

• **Relations of EMR with Ground-truth Bike Demand.** We further show the estimated bike demand against EMRs of all the schemes for NYC (in terms of income levels) in Fig. 9. We calculate the estimated bike demand of each selected location as follows. We first find all the ground-truth bike station placements within the 160m×160m region (same as the spatial constraint to avoid duplicated calculation) centered at the location, then we divide the total demand of those ground-truth stations in 2016 by the number of them.

Fig. 9 shows the total estimated bike demand for all the 71 new stations selected by each of the baseline approaches we studied as well as CGIRL. A scheme whose position lies closer to the top right is considered to match more bike demands with better EMR rewards, demonstrating the positive interactions between the reconfigured bike station network and the latent bike mobility needs. From the figure, we validate the effectiveness of EMR designs as in general higher EMR comes with higher estimated bike demand, and we can see that CGIRL demonstrates the potential of matching or bringing more bike usage to the bike sharing system than other approaches.

• **Model Sensitivity Study.** We have studied the sensitivity of CGIRL toward (a) the number of GCN layers (L); (b) the number of attention heads of GAT (B); (c) the number of attention heads of GEA (Q); (d) the size of GCN and GAT embeddings (W); and (e) the size of the hidden feature of each head in GEA (V).

Fig. 10a shows that the EMR improves and then decreases as the number of heads L increases. It may be because increasing the number of layers makes GCN capable of capturing more useful information about the topological correlations between locations, while too many layers capture redundant information leading to worse performance. Therefore, we set $L = 2$. Similar trends can be observed regarding B, W and V in Figs. 10b, 11a and 11b, where EMRs improve at first and then drop as B, W and V increase. We set $B = 3$, $W = 128$ and $V = 128$ accordingly. From Fig. 10c we observe that EMR is the highest when $Q = 1$ and decreases as Q increases, indicating one attention head of GEA is enough to differentiate the contributions of GCN and GAT. Increasing the Q makes the model capture noisy information and degrade the performance.

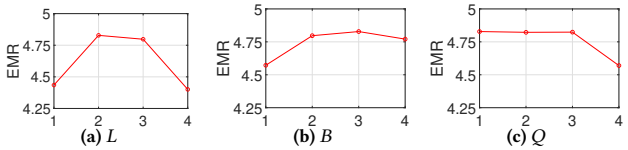


Figure 10: Sensitivity studies of (a) # of GCN layers, L, (b) total # of attention heads of GAT, B, and (c) total # of attention heads of GEA, Q.

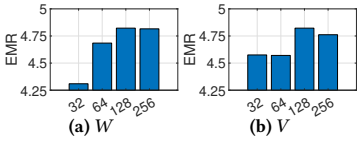


Figure 11: Sensitivity studies of (a) embedding size of the GCN and GAT, W, and (b) hidden size of the GEA, V.

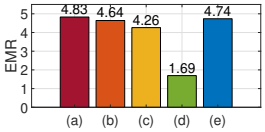
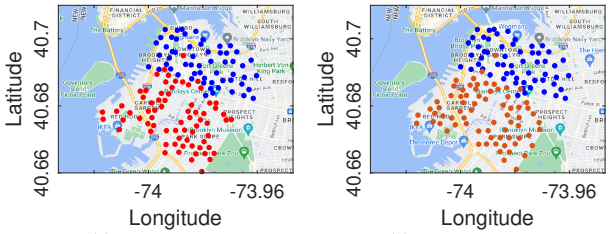


Figure 12: Model ablation study.

• **Ablation Studies on Module Importance.** We have studied the importance of the major design components by comparing (a) the complete design of CGIRL, with its variants excluding the major components: (b) GEA, (c) GCN, (d) GAT, and (e) SSSM. Fig. 12 shows the relative importance of different components of CGIRL upon the rewards. In particular, GAT and GCN have the most contributions to the expansion recommendation, and hence excluding these components significantly degrade the performance. More specifically, a model solely based on GCN (topological correlation) or GAT (node feature correlations) fails to comprehensively account for the rider's mobility trends between the nodes and socioeconomic effects in the BSN. In addition, the two graphs have different impacts upon the policy. Overall, all the components have demonstrated their effectiveness in BSN expansion studies.



(a) Prediction. **(b) Ground-truth.**

Figure 13: Recommended locations (red) by CGIRL, actual new station locations (brown), and the existing stations (blue).

• **Result Visualization.** We show in Fig. 13 the ground-truth locations of the new stations (red dots) and the distribution of new bike stations (brown dots) recommended by CGIRL (in terms of income level) in Brooklyn, NYC as well as the existing stations (blue dots). We note from Fig. 13 that the predicted new stations generally match those placed in real-world, demonstrating the potential of CGIRL in recommending station locations. The resulting root mean square difference between the locations of the placed stations and the actual placement locations is 0.139km.

6 RELATED WORK

• **Decision Recommendation for Urban Planning.** Driven by the urban big location and mobility data, location recommendation techniques have been adopted for the general public in urban

planning decisions such as intelligent transportation service [28]. Various heuristic and optimization approaches, including the ant colony algorithms [27], genetic algorithms [16], and other mathematical programming approaches [23], have been considered for providing the decision supports on the transportation network planning. However, these approaches may not necessarily capture and learn the complex and dynamic BSN expansion processes given the big urban location and mobility data. Different from the above studies, our CGIRL enables a novel location recommendation framework based on cross-graph interactive reinforcement learning for adaptive urban mobility network expansion (BSN in our case). Our comprehensive design with equity, mobility, and accessibility rewards provides efficient recommendations to assist urban city planning and facilitate the BSN expansion decision-making.

• **Reinforcement Learning (RL) for Urban Computing.** Deep RL has been an important approach for various urban computing applications, such as traffic signal control [3] and ambulance dispatch [14]. Ji et al. [10] studied the dynamic taxi route recommendation via deep RL that fuses the multiple spatio-temporal features. To recommend the locations of new metro lines, Wei et al. [24] proposed an attention-based actor-critic deep RL, which did not consider preserving the location correlations without the cross-graph interaction and graph embedding attention designs as in CGIRL. Different from the above works, we have proposed a novel policy network with novel interactions of GCN and GAT. CGIRL captures the correlations between stations in the changing BSN during the sequential selection steps. Unlike [29] where the fairness of transportation recommendations are incorporated as regularizers to improve equity of bike resource allocation, we propose within CGIRL a new reward function that jointly takes into account equity at both the community level (by a connectivity equity reward) and proximity level (by a proximity equity reward), achieving a more comprehensive and equity-aware BSN recommendation framework.

7 CONCLUSION

We propose CGIRL, a novel equity-aware cross-graph interactive reinforcement learning framework for bike station placement decision and recommendation. We have designed a novel reward function within our actor-critic reinforcement learning, jointly accounting for the local mobility, bike resource equity, and accessibility of different socioeconomic groups to the expansion regions. CGIRL takes in the topological and node feature correlations across the city regions as graphs and learns the placement policy of stations from the BSN deployment. Furthermore, CGIRL adaptively captures and differentiates the topological and node feature correlations within station placement through a novel cross-graph interaction network integrating graph convolution and attention. Our extensive experiment studies in three metropolitan cities have validated the effectiveness of CGIRL in automatically determining BSN expansion.

ACKNOWLEDGMENT

This project is supported, in part, by the National Science Foundation under Grant 2303575, 2021 Google Research Scholar Program, and 2021 NVIDIA Applied Research Accelerator Program.

REFERENCES

[1] 2021. *Aiming for Equity, Indego Adding 30 Bike-Share Stations, E-Bikes in 2022.* https://www.nbcphiladelphia.com/news/green/indego-2022-bike-share-expansion/3034356/
[2] 2021. *Bike Share is Rapidly Taking the World by Storm Post-Covid.* https://www.globenewswire.com/news-release/2021/09/30/2306599/0/en/Bike-Share-is-Rapidly-Taking-the-World-by-Storm-Post-Covid.html
[3] Chacha Chen, Hua Wei, Nan Xu, Guanjie Zheng, Ming Yang, Yuanhao Xiong, Kai Xu, and Zhenhui Li. 2020. Toward a thousand lights: Decentralized deep reinforcement learning for large-scale traffic signal control. In *Proc. AAAI.*
[4] Kamil Ciosek, Quan Vuong, Robert Loftin, and Katja Hofmann. 2019. Better exploration with optimistic actor-critic. *arXiv preprint arXiv:1910.12807* (2019).
[5] Greg P Griffin and Junfeng Jiao. 2019. Crowdsourcing bike share station locations: Evaluating participation and placement. *Journal of the American Planning Association* (2019).
[6] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. 2016. Deep residual learning for image recognition. In *Proc. IEEE CVPR.*
[7] Suining He and Kang G Shin. 2020. Towards fine-grained flow forecasting: A graph attention approach for bike sharing systems. In *Proc. ACM WWW.*
[8] Suining He and Kang G. Shin. 2022. Information Fusion for (Re)Configuring Bike Station Networks With Crowdsourcing. *IEEE TKDE* (2022).
[9] Hsun-Ping Hsieh, Fandel Lin, Jiawei Jiang, Tzu-Ying Kuo, and Yu-En Chang. 2021. Inferring Long-Term Demand of Newly Established Stations for Expansion Areas in Bike Sharing System. *Applied Sciences* (2021).
[10] Shenggong Ji, Zhaoyuan Wang, Tianrui Li, and Yu Zheng. 2020. Spatio-temporal feature fusion for dynamic taxi route recommendation via deep reinforcement learning. *Knowledge-Based Systems* (2020).
[11] Thomas N Kipf and Max Welling. 2016. Semi-supervised classification with graph convolutional networks. *arXiv preprint arXiv:1609.02907* (2016).
[12] Yexin Li, Yu Zheng, Huichu Zhang, and Lei Chen. 2015. Traffic prediction in a bike-sharing system. In *Proc. ACM SIGSPATIAL.*
[13] Junming Liu, Leilei Sun, Qiao Li, Jingci Ming, Yanchi Liu, and Hui Xiong. 2017. Functional zone based hierarchical demand prediction for bike system expansion. In *Proc. ACM SIGKDD.*
[14] Kunpeng Liu, Xiaolin Li, Cliff C Zou, Haibo Huang, and Yanjie Fu. 2020. Ambulance dispatch via deep reinforcement learning. In *Proc. ACM SIGSPATIAL.*
[15] Ryan Lowe, Yi Wu, Aviv Tamar, Jean Harb, Pieter Abbeel, and Igor Mordatch. 2017. Multi-agent actor-critic for mixed cooperative-competitive environments. *arXiv preprint arXiv:1706.02275* (2017).
[16] Mahmoud Owais and Mostafa K Osman. 2018. Complete hierarchical multi-objective genetic algorithm for transit network design problem. *Expert Systems with Applications* (2018).
[17] Xiaodong Qian and Miguel Jaller. 2020. Bikesharing, equity, and disadvantaged communities: A case study in Chicago. *Transportation Research Part A: Policy and Practice* (2020).
[18] Susan A Shaheen, Elliot W Martin, Adam P Cohen, Nelson D Chan, and Mike Pogodzinski. 2014. Public Bikesharing in North America During a Period of Rapid Expansion: Understanding Business Models, Industry Trends & User Impacts, MTI Report 12-29. (2014).
[19] Umer Siddique, Paul Weng, and Matthieu Zimmer. 2020. Learning Fair Policies in Multi-Objective (Deep) Reinforcement Learning with Average and Discounted Rewards. In *Proc. ICML.* PMLR.
[20] David Silver, Guy Lever, Nicolas Heess, Thomas Degris, Daan Wierstra, and Martin Riedmiller. 2014. Deterministic policy gradient algorithms. In *Proc. ICML.*
[21] Richard S Sutton, David A McAllester, Satinder P Singh, and Yishay Mansour. 2000. Policy gradient methods for reinforcement learning with function approximation. In *Proc. NeurIPS.*
[22] Petar Veličković, Guillem Cucurull, Arantxa Casanova, Adriana Romero, Pietro Lio, and Yoshua Bengio. 2017. Graph attention networks. *arXiv preprint arXiv:1710.10903* (2017).
[23] Yi Wei, Jian Gang Jin, Jingfeng Yang, and Linjun Lu. 2019. Strategic network expansion of urban rapid transit systems: A bi-objective programming model. *Computer-Aided Civil and Infrastructure Engineering* (2019).
[24] Yu Wei, Minjia Mao, Xi Zhao, Jianhua Zou, and Ping An. 2020. City metro network expansion with reinforcement learning. In *Proc. ACM SIGKDD.*
[25] Ronald J Williams. 1992. Simple statistical gradient-following algorithms for connectionist reinforcement learning. *Machine Learning* (1992).
[26] An Yan and Bill Howe. 2020. Fairness-Aware Demand Prediction for New Mobility. In *Proc. AAAI.*
[27] Zhongzhen Yang, Bin Yu, and Chuntian Cheng. 2007. A parallel ant colony algorithm for bus network optimization. *Computer-Aided Civil and Infrastructure Engineering* (2007).
[28] Yin Zhang, Yujie Li, Ranran Wang, M Shamim Hossain, and Huimin Lu. 2020. Multi-aspect aware session-based recommendation for intelligent transportation services. *IEEE T-ITS* (2020).
[29] Ding Zhou, Hao Liu, Tong Xu, Le Zhang, Rui Zha, and Hui Xiong. 2021. Transportation recommendation with fairness consideration. In *Proc. Springer DASFAA.*

A APPENDIX

We list symbols and the definitions of important concepts of CGIRL formulation in Table 2.

Table 2: List of symbols of important concepts in CGIRL.

Symbols	Definitions
Ω^{New}	Candidate grid set.
N^{New}	The total number of new stations to be placed.
$\mathbf{Z}^{\text{Ex}}$	Set of the existing bike stations.
$\mathbf{Z}_0$, N_0	The initial set and the number of all candidate locations.
$\mathbf{Z}_n^{\text{New}}$	Set of the new stations selected by the n-th step.
Γ_n	Index of the new station selected at step n.
$\mathbf{C}_n$	The set of locations within station placement constraints of all the new stations placed by the n-th step.
$\mathbf{Z}_n$	The updated candidate location set at the n-th step after considering station placement constraint.
$\mathbf{T}_n$	Station set at the step n.
$\mathbf{S}_n$	State at the step n.
R_n	Reward at the step n.
Φ_n	The location graph at the step n.
i	Index of an existing station, a selected new station or a candidate location.
j	Index of a candidate location.
k	Index of an existing station or a selected new station.
$\mathcal{N}_i$	Grid neighborhood of an existing station i.
$\mathbf{F}_i$	Spatio-temporal features of an existing station i.

A.1 Model Training and Testing

• **Gradients for Model Training.** In the model training phase, we run the BSN expansion to place the N^{New} stations for X times and we aim at maximizing the average EMR to increase the probability of selecting locations that generate a greater reward. Specifically, for each episode $x \in [1, X]$, we sample one candidate location for station placement at each step n based on the probability p_j (Eq. (17)), and we obtain a total of N^{New} selected locations of the new stations, $\{\Gamma_1^{(x)}, \ldots, \Gamma_n^{(x)}, \ldots, \Gamma_{N^{\text{New}}}^{(x)}\}$, and the probabilities set of selection, $\mathbf{P}^{(x)} = \left\{p_{\Gamma_1^{(x)}}, \ldots, p_{\Gamma_n^{(x)}}, \ldots, p_{\Gamma_{N^{\text{New}}}^{(x)}}\right\}$. We generate the $\text{EMR}^{(x)}$ of those locations by Eq. (22) and the estimated reward $\text{EMR}_c^{(x)}$ by the critic network.

Given $\text{EMR}^{(x)}$, $\text{EMR}_c^{(x)}$, $\mathbf{P}^{(x)}$, the gradient of the PN is then given by

$$\nabla \mathcal{J}(\theta_p) = \frac{1}{N^{\text{New}} X} \nabla_{\theta_p} \sum_{x=1}^{X} \sum_{n=1}^{N^{\text{New}}} \left(\text{EMR}^{(x)} - \text{EMR}_c^{(x)}\right) \log p_{\Gamma_n}^{(x)}(\theta_p), \tag{23}$$

where the probability $p_{\Gamma_n}^{(x)}$ is a function of the PN parameters θ_p, and we average the gradient over all X episodes to reduce the impact of random parameter initialization on the performance. We note that in Eq. (23) if $\text{EMR}^{(x)}$ is greater than $\text{EMR}_c^{(x)}$, the expanded BSN through the current policy is considered better than the current estimation. Thus, CGIRL updates the parameters θ_p such that the probability of generating such an action (decision) at the episode x increases.

To train the critic network in estimating the EMR, we adopt the Mean Squared Errors (MSEs) between EMR_c and EMR as the objective function. Specifically, we have the gradient of CN as

$$\nabla \mathcal{J}(\theta_c) = \frac{1}{N^{\text{New}} X} \nabla_{\theta_c} \sum_{x}^{X} \sum_{n}^{N^{\text{New}}} \left(\text{EMR}^{(x)} - \text{EMR}_c^{(x)}\right)^2, \tag{24}$$

where $\text{EMR}_c^{(x)}$ is governed by the parameters of the critic network θ_c.

In terms of model testing for recommendation, PN generates a probability p_j for each candidate location j to be selected for station placement at a step n. Taking into account the practical deployment of CGIRL, CGIRL recommends the highest p_j for placing Γ_n in the model testing phase, i.e.,

$$\Gamma_n = \arg\max_j p_j. \tag{25}$$

• **Training Algorithm.** The training of our model for updating θ_p and θ_c is illustrated by Algorithm 1. We first initialize the parameters of the actor (i.e., PN) and critic networks (Lines 1–2). We train the actor and critic for E epochs (Lines 3–10). For each epoch, we run the placement actions for X episodes (Lines 4–7). For each episode, we obtain $\text{EMR}^{(x)}$ of the selected locations and the corresponding probabilities $\mathbf{P}^{(x)}$ by PN (Line 5), and $\text{EMR}_c^{(x)}$ by the CN (Line 6). We then compute the gradient of the PN and CN, respectively, and update the parameters from the gradients (Lines 8–9).

Algorithm 1 Model Training of CGIRL.

Input: Maximum number of epochs E; the number of episodes X; the state $\mathbf{S}_n$; the candidate locations $\mathbf{Z}_n$.

1: Initialize actor parameters θ_p;
2: Initialize critic parameters θ_c;
3: **for** epoch = 1 to E **do**
4: **for** x = 1 to X **do**
5: $\text{EMR}^{(x)}, \mathbf{P}^{(x)} \leftarrow \text{PN}(\Phi_n)$;
6: $\text{EMR}_c^{(x)} \leftarrow \text{CN}(\mathbf{S}_n)$;
7: **end for**
8: Update actor parameters θ_p based on Eq. (23);
9: Update critic parameters θ_c based on Eq. (24);
10: **end for**

A.2 Experimental Settings

• **Evaluation Settings.** We divide a city map into a group of 350m×350m grid areas and then find those grids with new stations as Ω^{New} in our experimental studies. Specifically, we have identified 51 candidate grids with 885 road intersections in Brooklyn, NYC (71 new stations to be placed), 23 candidate grids with 205 intersections in downtown Chicago (25 new stations to be placed), and 15 candidate grids with 162 intersections (15 new stations to be placed).

Based on the average spatial distances to the existing stations, we cluster the 51 candidate grids in NYC into 6 region sets (denoted as Ω^{New}), 23 grids in Chicago into 2 region sets, and 15 grids in DC into 2 region sets. Following the BSN expansion practice in real world [8], we conduct BSN expansion set by set with closer to the

existing stations first. We leverage the new stations already placed at the former region sets as the existing stations to conduct the BSN expansion at the next Ω^{New}'s. Our ablation, sensitivity, and graph interaction studies are conducted upon the reward of placing 11 new stations in the 9 candidate grids closest to the existing stations.

• **Parameter Settings.** We set the default hyperparameters of CGIRL as follows. We set the station neighborhood $\mathcal{N}_i$ as a 350m×350m grid. We set the closest distance between new stations (spatial constraint c) as 160m for NYC, 100m for Chicago, and 100m for DC. We set $\lambda^{(\text{CE})} = 1$, $\lambda^{(\text{PE})} = 0.25$, $\lambda^{(\text{Mob})} = 2$, $\lambda^{(\text{Ac})} = 2$, and $\phi^{(\text{CE})} = \phi^{(\text{PE})} = 10$ for the EMR function in the three cities. We set the number of heads of GAT (B) as 3 and the number of convolutional layers of GCN (L) as 2. The size of the topological and node feature embeddings (W) is 128. For GEA, we set the number of the attention heads (Q) as 1 and the size of the hidden feature (V) as 128. The size of the hidden state of the LSTM (W_2) is 128. We set the number of epochs as 600 for NYC and DC and 1,200 for Chicago. We set the learning rate as 5×10^{-4} (which decays to 1×10^{-4} at the 200^{th} epoch) and Adam as the optimizer.

For AC, we use 100 ant agents and set relative importance of pheromone as 1 and relative importance of heuristic function as 3. For GA, we set the mutation probability as 0.01, and the crossover probability as 0.5. For REINFORCE, the policy network consists of 2 layers of 1D convolutional (Conv1D) and 1 fully connected layer to generate the action from the spatio-temporal features of the state and candidate locations. Each Conv1D layer of REINFORCE has 128 hidden units and the dense layer has 1 hidden unit. DPG adapts the policy network of CGIRL into a deterministic one, and a zero-mean Gaussian noise (with standard deviation of 0.05) is used for exploration. RL-ResNet adopts two ResNet blocks (with 128 filters and 1 filter, respectively). SRL adopts 128 units for the hidden layer of the sequential attention layer. RL-Dense leverages two dense layers with 128 hidden units and 1 hidden unit, respectively. RL-Conv2D leverages 2-layer Conv2D (128 and one 3x3 filters). RL-GCN and RL-GAT leverage the same GCN (2 graph convolution layers) and GAT (3 attention heads) settings as CGIRL (embedding size as 128).

For EMR integration, we note that all baselines will recommend the candidate locations that have the highest probability to be placed with stations. For the baselines (1)–(2), i.e., AC and GA, the EMR is incorporated into their optimization objective. For baselines (3)–(4), the EMR is integrated in Eq. (23) for the policy network. For baselines (5)–(10), we adopt the actor-critic reinforcement learning framework, and the EMR is incorporated into both Eq. (23) and Eq. (24) for the policy and critic networks, respectively.

We have implemented CGIRL and other schemes in PyTorch, and the models are trained and evaluated upon a Linux server with AMD Ryzen 3960X 3.8GHz, 128GB RAM, and 4 × Nvidia GeForce RTX 3090 24GB GDDR6X. For all three cities, the average training time of CGIRL is around 4min per station, while the average testing time (recommendation) is around 9ms per station.

A.3 Discussion

Our current studies focus on bike stations, SafeGraph check-ins, Uber/Taxi pick-ups, and socio-economic and POI data. However, our approach is general enough to accommodate additional features, including the traffic flows/speeds as well as the mobility demands at other public transit systems, which can be further infused through GCN and GAT networks and will be considered in our future work.

We note that the average training time for RL-Conv2D is around 1.24min/station and average testing time is about 6ms/station. While CGIRL may take more computation overhead, the bike station network expansion operations are generally performed as non-real-time tasks. Further computation efficiency enhancement can be achieved through parallel computing and hierarchical coordination, which will be considered in our future studies.