

Animating the Crowd Mirage: A WiFi-Positioning-Based Crowd Mobility Digital Twin for Smart Campuses

CHUNHUA CHEN, YUXIN YANG, and HAO YUAN, University of Macau, China

LONGBIAO CHEN, Xiamen University, China

LEYE WANG, Peking University, China

BINGQING QU, BNU-HKBU United International College, China

DINGQI YANG*, University of Macau, China

Understanding crowd mobility is critical for many applications. In this paper, we propose CrowdMirage, a WiFi positioning-based crowd mobility digital twin for smart campuses. Specifically, we first design an end-to-end human mobility trace extraction pipeline from the comprehensive but noisy WiFi connection logs on a university campus. We then design two predictive and simulative models for the crowd flow prediction and simulation tasks, respectively. Considering the particularity of on-campus mobility, we propose a cross-grained crowd flow prediction model to forecast crowd flow at both building and floor levels. For crowd flow simulation, we design a conditional generative model based on conditional diffusion to simulate the crowd flow under given mobility-related contexts that are systematically identified. We evaluate CrowdMirage on two-year WiFi connection logs collected at our university. The results show that CrowdMirage achieves superior performance in both crowd flow prediction and simulation tasks. Our case studies show that CrowdMirage cannot only accurately forecast cross-grained crowd flow across different cases, but also simulate interpretable crowd flow under previously unseen conditions.

CCS Concepts: • **Human-centered computing** → **Ubiquitous and mobile computing**; • **Information systems** → **Data mining**.

Additional Key Words and Phrases: Smart Campus, Digital Twins, Mobility Simulation, Crowd Flow

ACM Reference Format:

Chunhua Chen, Yuxin Yang, Hao Yuan, Longbiao Chen, Leye Wang, Bingqing Qu, and Dingqi Yang. 2024. Animating the Crowd Mirage: A WiFi-Positioning-Based Crowd Mobility Digital Twin for Smart Campuses. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 8, 4, Article 158 (December 2024), 32 pages. <https://doi.org/10.1145/3699792>

1 INTRODUCTION

The study of crowd mobility is essential across various fields, such as urban planning [59], public health [62], social sciences [46], and environmental sustainability [6], and disaster management [65] etc. On one hand, modeling comprehensive crowd flow can support predictive analysis forecasting future crowd flow based on past crowd flow

*Corresponding author: Dingqi Yang

Authors' addresses: [Chunhua Chen](mailto:chunhuachen@um.edu.mo), chunhuachen@um.edu.mo; [Yuxin Yang](mailto:yuxinyang@connect.um.edu.mo), mc15582@connect.um.edu.mo; [Hao Yuan](mailto:hao.yuan@connect.um.edu.mo), hao.yuan@connect.um.edu.mo, State Key Laboratory of Internet of Things for Smart City and Department of Computer and Information Science, University of Macau, Macau SAR, China; [Longbiao Chen](mailto:longbiaochen@xmu.edu.cn), longbiaochen@xmu.edu.cn, Xiamen University, Xiamen, China; [Leye Wang](mailto:leyewang@pku.edu.cn), leyewang@pku.edu.cn, Peking University, Beijing, China; [Bingqing Qu](mailto:bingqingqu@uic.edu.cn), bingqingqu@uic.edu.cn, BNU-HKBU United International College, Zhuhai, China; [Dingqi Yang](mailto:dingqiyang@um.edu.mo), dingqiyang@um.edu.mo, State Key Laboratory of Internet of Things for Smart City and Department of Computer and Information Science, University of Macau, Macau SAR, China.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, or post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

© 2024 Copyright held by the owner/author(s). Publication rights licensed to ACM.

ACM 2474-9567/2024/12-ART158

<https://doi.org/10.1145/3699792>

Table 1. Statistical comparison of datasets

Dataset	#Users/Devices	#Check-ins	#POIs	Sampling Rate	Activity Scale	Duration
GeoLife [75]	182	24.9M	-	1-5s	City	54 months
Lausanne MDC [33]	200	95.4M	-	1-15mins	Country	18 months
Copenhagen CNS [57]	700	7.6M	-	5mins	City	4 weeks
Gowalla [10]	196.6K	6.4M	1.3M	4.59h (median)	Global	21 months
Brightkite [10]	58.2K	4.5M	773K	3.25h (median)	Global	31 months
Foursquare [66]	2.7M	90M	11.1M	22.18h (median)	Global	22 months
CrowdMirage	81.8K	4.2B	8K	27s (median)	Campus	24 months

history [74], providing invaluable insights for making informed decisions and implementing proactive policies. On the other hand, crowd flow simulation [23] generates crowd flow under given scenarios to answer “what-if” queries, which can be used as an effective sandbox for benchmarking mobility-related applications/policies. Both predictive and simulative analyses are critical for crowd mobility studies, which are thus the key ingredients in the development of a so-called mobility digital twin, serving as a holistic digital replica of crowd mobility.

In the current literature, existing works mainly focus on proposing algorithmic solutions addressing either predictive or simulative tasks in an ad-hoc manner without building an end-to-end framework [40]. They mostly resort to existing mobility datasets publicly available online, which are usually collected on a small scale or suffer from low data quality. On one hand, lab-controlled data collection schemes using specific hardware can obtain high-quality but small-scale mobility datasets, which are usually limited to hundreds of users. For example, as shown in Table 1, GeoLife dataset [75] contains detailed trajectories of 182 users in Beijing over a period of more than 4 years, with a pre-defined sampling rate of every 1-5 seconds or every 5-10 meters per point; the Lausanne Mobile Data Challenge (MDC) [33] dataset is limited by 200 users in a country scale with a pre-defined sampling rate of 1-15 minutes per point; the Copenhagen Networks Study (CNS) [57] involves over 700 university students over a period of four weeks, with a sampling rate of GPS every 5 minutes. However, such a small sample size makes it difficult to infer broad mobility patterns or perform downstream tasks. It is also unclear if these samples are representative of the entire population.

On the other hand, user-shared mobility data on social media or mobile networks contains large-scale but low-quality data. For example, the widely used Gowalla [10], Brightkite [10] and the Foursquare dataset [66] include 6.4 million, 4.5 million, and 90 million check-ins at 1.3M, 773K, and 11.1M Point of Interests (POIs) by 196.6K, 58.2K and 2.7M users on a global scale, respectively. However, user voluntarily shared check-ins on social media platforms imply intrinsic biases due to user preferences, and also suffer from severe data sparsity issues, as evidenced by the median time between successive check-ins being over 4 hours, 3 hours, and 22.18 hours for the three datasets, respectively.

These limitations introduce inevitable biases to the downstream predictive and simulative analyses, which thus hinder the development of a holistic mobility digital twin.

In this paper, we design and develop CrowdMirage, an end-to-end mobility digital twin framework for smart campuses. It serves as a digital replica of on-campus crowd flow, which not only senses and monitors the comprehensive crowd flow via WiFi Access Points (APs) deployed on a university campus, but also supports both predictive and simulative analyses of on-campus crowd mobility. We report our experience in identifying and tackling the challenges and issues encountered in extracting crowd flow data from raw WiFi connection data, and in designing both predictive and generative models for the on-campus crowd mobility forecasting and simulation, respectively. Specifically, the design and development of CrowdMirage face the following research challenges and practical issues.

- *How to precisely extract crowd flow from noisy WiFi connection records?* In this study, we use WiFi APs on a university campus to sense crowd flow. Specifically, on the university campus of 1.09 km², over 8,000 WiFi APs have been deployed, covering approximately 93% indoor and 40% outdoor areas of the campus, providing Internet services to over 13,000 students and staff, as well as to guests. Individuals carrying WiFi-equipped mobile devices (e.g., smartphones, watches, bracelets, tablets, and laptops) leave their spatiotemporal “digital footprints” when moving on campus, recorded by the (automatic) connection logs between the devices and the WiFi APs. However, such data is intrinsically noisy to represent crowd mobility, due to the inevitable presence of non-mobile devices (e.g., desktops), multiple devices possessed by the same individual, multiple WiFi SSIDs, and randomized MAC addresses of some devices, etc. Existing studies [12, 33, 57, 69, 73] on WiFi-positioning-based human mobility often provide little or insufficient technical details on the data processing pipeline.
- *How to design practical predictive models for cross-grained on-campus crowd mobility?* Compared to traditional GPS-based mobility traces, WiFi-positioning-based mobility traces cannot only characterize the crowd flow between buildings on campus, but also finer-grained crowd flow between building floors, where some floors of different buildings could be connected via skyways. In this context, modeling the building- and floor-level crowd mobility are both critical for efficient facility management, such as for air conditioning or elevator control optimization. However, crowd mobility often exhibits different spatiotemporal patterns at different levels of granularity, while also having intrinsic correlations across granularities. Existing work on crowd flow prediction mostly focuses on uni-grained crowd flow [25, 36, 64, 70] and thus fails to jointly capture the cross-grained crowd mobility patterns.
- *How to design conditional generative models for the holistic simulation of crowd flow beyond imitation?* The on-campus crowd mobility patterns usually exhibit on one hand obvious spatiotemporal regularities. On the other hand, such mobility patterns indeed depend on various context factors, including but not limited to time in a day, day in a week, academic almanac, weather conditions, and warning signals (for tropical cyclones, rainstorms, or thunderstorms, etc.). However, existing crowd flow simulation techniques mostly resort to statistical or mechanistic models to reproduce previously seen crowd mobility patterns [2, 56] and ignore the mobility-related context factors. Against this background, it is important to design conditional generative models incorporating these context factors for the holistic simulation of crowd flow beyond imitation, i.e., simulating crowd flow under previously unseen scenarios.

To address these issues, we first design an end-to-end crowd mobility flow extraction pipeline based on a thorough analysis of the raw WiFi connection logs to identify and filter out various mobility-irrelevant noises to extract informative on-campus crowd flow data. Afterward, considering the particularity of the on-campus mobility with different location granularities, we propose a cross-grained crowd flow prediction model with a learnable knowledge transfer mechanism between different granularities, so as to effectively forecast future crowd flow at both building and floor levels. Moreover, we formulate the crowd flow simulation problem as a conditional generation problem, and then systematically identify influential context factors and resort to conditional diffusion models for conditional generation. We build a prototype of our mobility digital twin CrowdMirage and evaluate it on the long-term and comprehensive WiFi connection logs collected in our university for two years. Our contributions are summarized as follows:

- We propose an end-to-end robust crowd mobility flow extraction pipeline with detailed data processing and flow estimation steps, which can accurately estimate crowd mobility flow from the raw and noisy WiFi connection logs.
- We propose a cross-grained crowd flow prediction model, which designs a learnable knowledge transfer mechanism between different location granularities, benefiting on-campus crowd flow prediction at both building and floor levels.

- We design a conditional generative model for crowd flow simulation based on conditional diffusion, which systematically identifies and integrates influential mobility-related context factors for the holistic simulation of crowd flow beyond imitation.
- The evaluation using real-world WiFi connection logs shows that CrowdMirage achieves superior performance in both crowd flow prediction and simulation tasks, with 3.35%-9.09% and 8.80%-68.92% improvement over the best-performing baselines, respectively. Our case studies also demonstrate that CrowdMirage can not only accurately forecast cross-grained crowd flow across different cases, but also simulate interpretable crowd flow under previously seen conditions.

The rest of the paper is organized as follows: In Section 2, we present the related work. This is followed by an overview of the CrowdMirage framework in Section 3. The three main components of our study, namely crowd mobility flow extraction, crowd flow prediction, and crowd flow simulation, are detailed along with their design choices in Sections 4, 5, and 6, respectively. Section 7 describes the experiments conducted on these three components, demonstrating the effectiveness of our proposed model. In Section 8, we provide a discussion on potential use cases for both crowd flow prediction and simulation. Finally, Section 9 concludes the paper.

2 RELATED WORK

In this section, we briefly discuss two threads of related work on human mobility analytics and smart campus digital twins.

2.1 Human Mobility Analytics

Human mobility analytics can be dated from 1885, when Ravenstein studied human migration using mobility trajectories extracted from demographic data [50]. Nowadays, various IoT sensors make human mobility trajectory data much more accessible [17, 22]. On the one hand, user-voluntarily-shared mobility data on social media platforms is often used for human mobility studies, such as a widely used Foursquare global dataset consisting of over 90 million check-ins by over 2 million users on over 11 million POIs [67]. However, due to the voluntarily sharing basis, such a dataset has inevitable issues of sparsity and biases [66]. On the other hand, lab-controlled data collection schemes recruit participants for mobility data collection. Due to its practical implementation and the privacy concerns of individuals, the scale of the publicly available mobility traces are often small, which is usually limited to hundreds of users (e.g., 700 students in Copenhagen Networks Study [57], and 200 users in Lausanne Mobile Data Challenge [33]).

According to the mobility analytic problem settings [40], human mobility modeling techniques roughly fall into two types of tasks (i.e., predictive and generative tasks) with two types of data representation (i.e., individual mobility trajectories and crowd flow). First, predictive tasks on mobility trajectories are known as location prediction problems, forecasting the location of an individual based on the user's historical mobility traces [63, 66]. Second, predictive tasks on crowd flow forecast the crowd flow (the number of individuals or vehicles) of locations based on historical crowd flow data [38, 41]. Third, generative tasks on mobility trajectories try to generate synthetic trajectories that are similar to real-world human mobility traces in terms of statistical patterns [23, 39]. Finally, generative tasks on crowd flow generate synthetic flow among locations, mimicking the real-world mobility flow patterns [55, 56]. These mobility modeling tasks have been widely studied to support various smart city applications, such as urban event organization [8], location recommendation [68], crowdsensing [71], urban resource allocation [9], telecommunication [19, 42] and urban dynamic simulation [37], etc.

In this paper, we design and develop CrowdMirage, a crowd mobility digital twin that not only senses and monitors *comprehensive* on-campus crowd mobility via Wi-Fi APs, but also learns to *forecast* future crowd flow and *simulate* crowd flow in unseen scenarios. Note that a few recent works also study comprehensive user mobility traces collected via WiFi APs for a building complex [69, 73] or a university campus [12]. While the first

two works focus on predictive modeling, the second work presents an empirical analysis of on-campus mobility. Our work differs from these works by designing and developing a crowd mobility digital twin for both predictive and simulative analyses of on-campus crowd mobility.

2.2 Campus Digital Twins

A campus digital twin refers to a virtual replica or representation of a physical campus environment that enables monitoring, analysis, and simulation of various aspects of campuses [5]. In the current literature, existing works on campus digital twins can be classified into two main categories according to their focuses, i.e., immersive learning and campus management.

On the one hand, digital twins of learning environments, such as virtual classrooms, labs, or exam rooms, are developed to provide immersive learning to students via virtual reality and augmented reality technologies. In particular, such immersive learning has gained increasing popularity during COVID-19 [32] in the past few years. For example, the University of Kansas School of Nursing designs virtual exam rooms for students to conduct health assessments with patients via question-answering under the supervision of the instructors [4]; Morehouse College designs a so-called metaversity [32] providing students with virtual classroom, labs, and even imagined learning places [7]; the University of Pennsylvania uses digital platform to imitate in-person office hours [34]; the Fisk University developed digital twins for their Cadaver Lab [47].

On the other hand, digital twins of campus infrastructures are developed for efficient campus management, which primarily focuses on the integration of the Building Information Modeling (BIM) paradigm [24]. For example, Heriot-Watt University [60] and the University of Glasgow [49] developed campus digital twins to investigate the potential energy, carbon, and cost savings of building heating systems, for the goal of zero-carbon building/campus management; Eindhoven University of Technology [31] and the University of Texas Austin [35] developed campus digital twins for smart building energy management and control; Rice University [61] developed its Campus Digital Twin on top of ArcGIS with searchable facilities. Moreover, besides the energy consumption data, various IoT sensor data can also be integrated into campus digital twins. For example, Dublin City University developed a 3D version of the digital campus, which integrates data from IoT sensors on footfall, energy, and water usage [24], and provides real-time monitoring of the campus infrastructure. King Abdullah University of Science and Technology developed a campus digital twin to assess transportation emissions on campus via simulated mobility data [1]. The University of Murcia integrated data from various IoT sensors, including energy consumption, building occupancy, room usage, solar energy, and other resource consumption [51] for better campus management. University Ramon Llull integrates building information modeling tools with the Internet of Things-based wireless sensor networks in the fields of environmental monitoring and emotion detection to provide insights into the level of comfort [72]. These existing works mostly focus on the monitoring and control of on-campus infrastructures, thus providing limited intelligence to campus management.

In this paper, we focus on building a crowd mobility digital twin using real-world Wi-Fi-positioning-based mobility data, providing both predictive and simulative analyses of on-campus crowd mobility. To the best of our knowledge, this is the first crowd mobility digital twin of its kind.

3 FRAMEWORK OVERVIEW

The design of our CrowdMirage is shown in Figure 1. First, the crowd mobility flow extraction module extracts both crowd flow (i.e., the number of users surrounding an AP) and Origin-Destination (OD) transition flow (i.e., the number of users transiting from an origin to a destination) from the noisy WiFi connection records, via our mobility trace extraction pipeline designed based on our extensive empirical analysis and investigation of the WiFi connection logs. Second, the crowd flow prediction module designs a cross-grained information transfer mechanism that systematically integrates the building- and floor-level spatiotemporal Graph Neural Networks

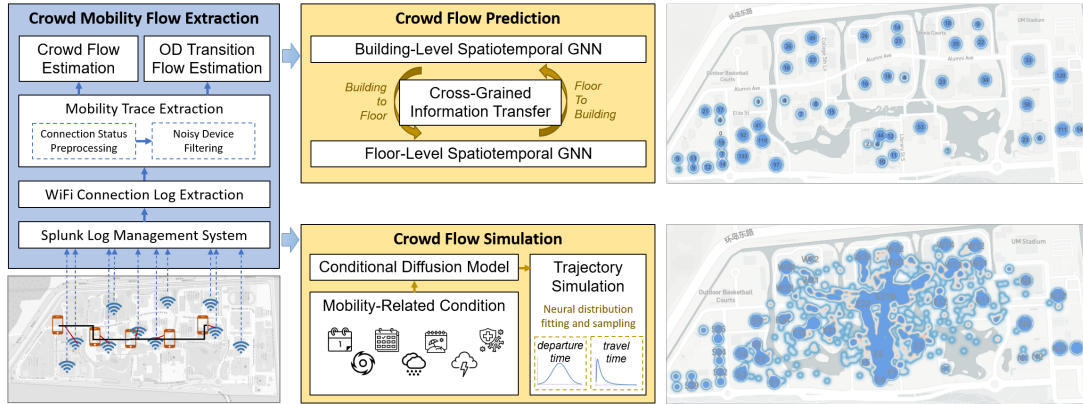


Fig. 1. Overall design of CrowdMirage

Table 2. WLAN Configurations and Policies

Account	Secured_WLAN	eduroam	Public_WiFi
University account	✓	✓	✓
eduroam account	×	✓	✓
Guest	×	×	✓

(GNNs), to benefit the crowd flow prediction on both location granularities. Third, the crowd flow simulation module designs a conditional diffusion model generating both crowd flow and OD transition flow conditioned on given mobility-related contexts, to achieve the generation beyond imitation for previously unseen contexts. The right panel shows the heatmap visualization of the predicted and simulated crowd flow. We present the details of individual modules below.

4 CROWD MOBILITY FLOW EXTRACTION

In this section, we present our approach to extracting crowd mobility flow from the noisy WiFi connection records, discussing how we address the technical issues we encountered.

4.1 Wireless Network Deployment and Configuration

In our university, the Wireless Local-Area Network (WLAN) is deployed with over 8,000 APs, covering 93% indoor and 40% outdoor areas of the campus. The APs support the IEEE 802.11b/a/g/n/ac/ax standard with radio frequencies of 2.4/5 GHz using devices from Aruba Networks. The WLAN service is provided to different users with different accounts following three different policies as shown in Table 2. First, Secured_WLAN can be accessed by university account holders only (students and staff), providing both public internet and university intranet services. Second, eduroam is an international WiFi internet access roaming service for users in research, higher education, and further education institutes; users can access eduroam via their own institute accounts if their affiliated institution has joined the eduroam project. Note that the university account holders can also access eduroam. Third, Public_WiFi provides public guests with account-free and short-term Internet access, for a maximum of 8 hours per day; a guest needs to agree to the terms of use via a captive portal before accessing this network.

Table 3. WiFi log examples

Access Point	Timestamp	Connection Status	Hashed Device MAC	Hashed Account
RC8-GF-12	2023-12-01 15:27:30	Auth request	c4b***c46	a0d***b08
RC8-GF-12	2023-12-01 15:27:30	Auth success	c4b***c46	a0d***b08
RC8-GF-12	2023-12-01 15:27:30	Assoc request	c4b***c46	a0d***b08
RC8-GF-12	2023-12-01 15:27:30	Assoc success	c4b***c46	a0d***b08
RC2-1F-18	2023-12-01 15:50:00	Assoc request	c4b***c46	a0d***b08
RC2-1F-18	2023-12-01 15:50:00	Assoc success	c4b***c46	a0d***b08
SQC3-2F-3	2023-12-01 21:05:02	Disassoc from sta	c4b***c46	a0d***b08
SQC3-2F-3	2023-12-01 21:40:29	Deauth from sta	c4b***c46	a0d***b08

4.2 WiFi Connection Logs

The WiFi connection logs are managed on the Splunk¹ platform. As shown in Table 3, we extract each connection record consisting of five fields: Access Point, Timestamp, Connection Status, Hashed Device MAC, and Hashed Account.

- **Access Point:** A unique label per AP, which specifies its location at three levels, i.e., building, floor, and AP, separated by the “-”, from left to right. For example, “RC8-GF-12” represents the building of Residential College (RC) 8, Ground Floor (GF), AP with the ID 12.
- **Timestamp:** The timestamp for each connection in the format of “yyyy-MM-dd HH:mm:ss”.
- **Connection Status:** For users with accounts, the first-time connection log series would be “Authentication request—Authentication success—Association request—Association success”. For account-free users, the authentication steps do not apply. Afterward, if a connected device is in a fast-roaming mode, it will skip the authentication steps and keep switching between different APs with “Assoc request—Assoc success” until the device completely disassociates or deauthenticates from the WLAN. Otherwise, the connected device would also follow the first-time connection logs when switching to another AP. Besides, associations are not always completed with “Assoc success” but may fail with “Association Failed” due to bad network quality.
- **Hashed Device MAC:** A unique hashed value for the MAC address of a WiFi-equipped device connected to a WLAN.
- **Hashed Account:** A unique hashed value for the user account (if any) that is used to connect to a WLAN. Note that each user account could be used on multiple WiFi-equipped devices.

4.3 Mobility Trace Extraction Pipeline

The extracted WiFi connection logs contain useful information on on-campus human mobility, but also contain inevitable noises. Subsequently, we design a robust human mobility trace extraction pipeline as follows.

4.3.1 Connection Status Pre-Processing. The extracted WiFi connection logs consist of different connection status, which are redundant to represent the spatiotemporal presences of the corresponding devices.

First, we found a large amount of duplicated logs with the connection status “Auth request” and “Auth success”. After digging into the raw connection logs from the Splunk log management system, we found that these duplicated entries are mostly generated by devices with university accounts authenticating to one of the networks eduroam or Secured_WLAN and then quickly shift to the other within the same second; these two networks are deployed on the same AP hardware but with different BSSIDs, resulting in two identical entries in our extracted

¹https://www.splunk.com/en_us/blog/learn/what-splunk-does.html

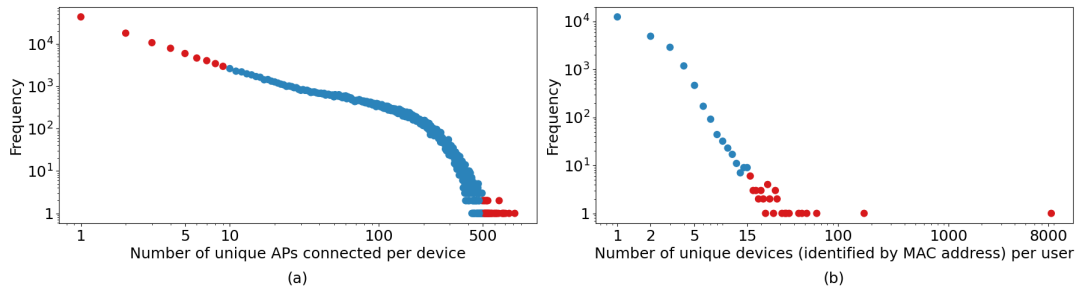


Fig. 2. Data distribution of the raw WiFi logs. (a) The frequency of the number of unique APs connected by each device per week. (b) The frequency of the number of unique devices by each user. The red color indicates the noises being filtered out.

logs (because we only focus on which physical AP hardware a device connects to, while the BSSIDs are not extracted). This shift is due to individual device configuration of auto-connection, depending on the connection priority setup, signal strength and quality. Based on this insight, we first eliminate these duplicated log entries.

Second, we discover that the status of “Auth request” and “Assoc request” are not always reliable to represent the physical closeness of a device to an AP, because they might be followed by “Auth failed” and “Assoc failed”, which could be caused by the WiFi signal strength and quality. Note that the “Auth failed” status could also be caused by wrong credentials, but this rarely happens because the successfully authenticated credentials are often automatically stored by user devices. Therefore, we remove log records with the two connection status.

Third, the “Auth success” and “Assoc success” status indicate that a user device is within a reasonable distance from an AP such that the authentication or association process can be successfully accomplished.

Finally, we observe that “Deauthentication from station” and “Disassociation from station” status are often significantly lagged with an unpredictable delay. For example, when a device moves from one AP to another, the “Disassoc from sta” log from the former AP is often recorded with a timestamp later than the “Assoc success” log at the latter AP, which makes the “Disassoc from sta” status unreliable to estimate the actual departure time of the device from the former AP.

Following the above reasons, we filter the log entries with “Auth success” and “Assoc success” status, and then keep only the log fields including Access Point, Timestamp, Hashed Device MAC, and Hashed Account. Subsequently, the filtered logs record the physical presence of devices close to APs.

4.3.2 Noisy Device Filtering. Based on the filtered logs, we then extract the trajectories of the devices, represented as a sequence of AP-timestamp pairs. However, log records from some devices cannot reflect on-campus human mobility. Through our empirical analysis, we identify the following types of noisy devices and filter out the noisy devices according to the defined criteria.

- **Guest Devices.** The guest devices refer to the devices connecting to the network Public_WiFi without authentication and thus no user account. Our empirical analysis shows that each guest device has connected to a much smaller number of APs (4.03 on average per day) than the devices with accounts (26.4 on average per day). Subsequently, the guest devices have quite limited contributions to the crowd mobility traces and are thus considered not representative of on-campus mobility. Therefore, these guest devices are excluded from our analysis.
- **Non-(or low-)mobile Devices.** Our empirical analysis shows that some devices may attach to a small number of APs over a long period, which could be WiFi-equipped desktops, lab equipment, or smart home equipment that cannot reflect human mobility. According to our empirical analysis on the frequency of

the number of unique APs connected by each device per week as shown in Figure 2(a), we define those devices that have ever connected to less than 10 different APs per week as non-(or low-)mobile devices. These devices represent 30% of the total devices and correspond to 9% of the total logs.

- **Publicly Shared Devices.** We observed in the raw WiFi logs that a few devices may attach to an unreasonably large number of APs in a period, which could be shared handsets for campus security staff that are not representative of individuals' trajectories. According to the frequency of the number of unique APs connected by each device per week as shown in Figure 2(a), we define those devices that have ever connected to over 500 APs per week as publicly shared devices. These devices represent 0.25% of the total devices and correspond to 2% of the total logs.
- **Devices with non-persistent MAC randomization.** We observed that a few accounts are associated with an unreasonably large number (up to 8000) of devices, uniquely identified by hashed device MACs. Our investigation shows that they are probably due to the setting of non-persistent MAC randomization of some devices. For example, smartphones equipped with Android 10/11 [16] or IOS 14 [58] randomize their MAC addresses for different WiFi networks with the persistent MAC randomization policy by default, where a persistent randomized MAC address is used for a WiFi network (a unique SSID). In our studied WLAN, this implies that each device could have three MAC addresses for the three SSIDs, Secured_WLAN, eduroam, and Public_WiFi, respectively. However, the device equipped with Android 12 [16] and above can manually set to a non-persistent MAC randomization mode, which will re-randomize the MAC address at the start of every connection. As we cannot explicitly identify this case in our extracted logs with only hashed device MAC, we thus filter out accounts with a large number of devices (unique MAC addresses). As seen in Figure 2(b), most users possess a small number of devices. In the daily lives of students and staff, the most commonly used devices are mobile phones, tablets, and laptops. We empirically assume each user may have up to 5 mobile devices, corresponding up to $3 \times 5 = 15$ MAC addresses if all are set to persistent MAC randomization. We thus filter out devices with accounts with over 15 hashed device MAC addresses, representing 0.05% of the total accounts and corresponding to 10% of the total logs.

Note that the above noisy devices may overlap, such as a guest device is often a low-mobile device.

4.3.3 Human Mobility Trace Extraction. After the log preprocessing and filtering of noisy devices, we need to extract human mobility traces from device mobility traces. Specifically, an individual often possesses multiple devices, such as a mobile phone, a smart watch/bracelet, or a tablet/laptop. The collective logs of these devices often show "sudden-move" where two devices under the same user account appear at two distant APs at (almost) the same time. Because active devices may not be carried by the user all the time; a student may go to a classroom with her mobile phone while leaving her tablet connected to WiFi at the dormitory. Therefore, we group the WiFi logs with the same Hashed Account and sort out the device with the most connections in one week, which is regarded as the most actively used device of the user in that week. We thus select the logs of this primary device to represent the mobility trace of the user in that week, formulated as a sequence of AP-timestamp pairs. We regard this step as the primary device selection and evaluate its usefulness in the experiments later. We consider the most active device for every week here to take into account the fact that a user may upgrade their most used devices over time.

4.4 Crowd Flow Estimation

Considering the practical use cases (e.g., the frequency of the campus loop shuttle is every 10-15 minutes), we estimate the crowd flow at a 10-minute time granularity based on the extracted human mobility trace. We assume that a user contributes to only one AP in a 10-minute time slot. Then *the crowd flow of an AP can be estimated as the number of users contributing to the AP in a 10-minute time slot*. However, this crowd flow estimation method has to consider the following practical issues raised from our empirical analysis of the dataset.

- Extracted human mobility traces may have connections to multiple APs in one 10-minute time slot. In this case, the AP that has the most connections with the device is considered the contributed AP in the 10-minute time slot.
- Two consecutive logs may also imply a time interval over 10 minutes. In this case, considering the campus size of 1.09 km², if the two continuous WiFi logs are connected with two different APs and the time interval is less than 60 minutes, we assume that the user is moving from one AP to the other. The first half of the interval contributes to the origin AP and the second half contributes to the destination AP. If the time interval is over 60 minutes, we assume that the user is out-of-campus and does not make a contribution to any AP in that period.
- Users may stay at the same locations for a long time, such as attending a class or resting in the dormitory. This might cause the WiFi module of devices to be in a sleeping mode, especially for devices with no apps running in the backend. Therefore, if the two consecutive WiFi logs are connected to the same AP with a time interval of over 10 minutes, we assume the user always contributes to the AP during this time period.

Following the above process, we estimate the crowd flow of i -th AP denoted as x_i and all crowd flow in one 10-minute time slot t can be represented as a vector $\mathbf{X}_t = [x_{1,t}, x_{2,t}, \dots, x_{n,t}]$, where n is the total amount of the APs. Furthermore, following the three-level hierarchy of AP locations “Building-Floor-AP”, the flow in one floor or building can be computed as the sum of flow of all APs located in the floor or building.

4.5 OD Transition Flow Estimation

Besides the crowd flow that characterizes the number of users surrounding an AP, we also estimate the Origin-Destination (OD) Transition Flow that characterizes the number of users transiting from an origin to a destination during each 10-minute time slot, represented as an OD transition flow matrix. The OD transition flow estimation should consider the following practical issues besides those mentioned above for crowd flow estimation.

- As we record the number of transitions from an origin to a destination in each 10-minute time slot, the transition may take over 10 minutes. Therefore, we record each transition into its departure time slot.
- Staying at the same location (the origin is the same as the destination) does not contribute to any OD transition flow.
- To characterize the transition from/to out-of-campus places, we add one extra row/column to the OD transition flow matrix to record these transitions.

Following the above process, we estimate the OD transition flow matrix in the 10-minute time slot t , denoted as $\mathbf{M}_t \in \mathbb{R}^{(n+1) \times (n+1)}$ where $\mathbf{M}_t[o, d]$ is the transition flow from the origin o to the destination d and $(n+1)$ is the number of APs plus one extra row/column for out-of-campus places. Following the three-level hierarchy of AP locations “Building-Floor-AP”, we can compute the transition flow between floors or buildings by aggregating the transition flow between the corresponding AP pairs.

4.6 Data Privacy

In the whole framework design, we implement the following privacy protection mechanisms. First, the WiFi connection logs are extracted from the Splunk log management system directly with hashed device MAC and hashed accounts, without any other metadata about the devices or user accounts. Second, the mobility trace extraction pipeline is implemented and executed on the same intranet of the log management system, without any external access. Third, only the extracted crowd flow and OD transition flow are used to represent on-campus crowd mobility. Finally, the AP label (e.g., “RC8-GF-12” as shown in Table 3) is designed to be able to identify the corresponding building (RC8) and floors (GF) only but not the specific location where the AP (12) is located on the floor. Moreover, we design the following predictive and simulative analysis module considering only the building- and floor-level crowd mobility.

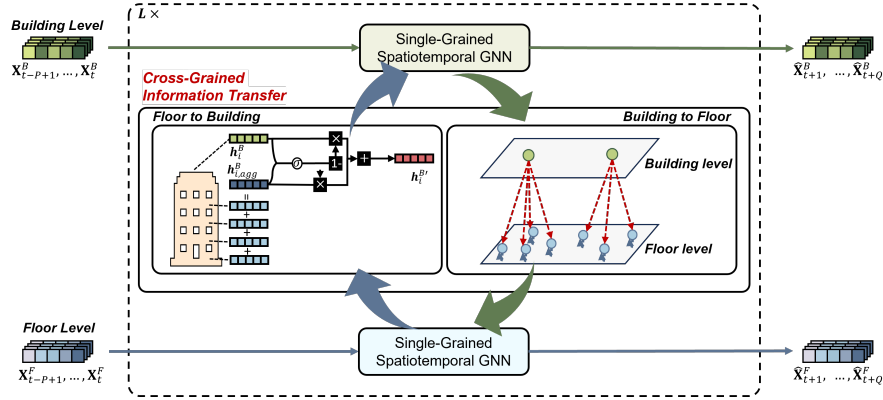


Fig. 3. Our crowd flow prediction model capturing the cross-grained crowd mobility patterns

5 CROWD FLOW PREDICTION

Crowd flow prediction aims to forecast future crowd flow with the input of historical crowd flow data. In our CrowdMirage, we focus on the crowd flow prediction at both the building and floor levels, while considering not only the spatiotemporal dynamics unique to individual levels but also their cross-dependencies. The overview of our proposed model is shown in Figure 3. It first uses two identical but separate single-grained spatiotemporal GNNs designed to capture the respective spatiotemporal dependencies of crowd flow within a single location granularity building or floor, respectively. The cross-grained information transfer module then designs a dual-channel information transfer mechanism, i.e., “floor to building” and “building to floor” channels, each employing unique strategies for effective information transfer. In the following, we first present our problem formulation, followed by the single-grained spatiotemporal GNN and the cross-grained information transfer module.

5.1 Definitions and Problem Formulation

We consider the two-level location hierarchy of crowd flow:

- **Building-Level Crowd Flow.** We denote the set of buildings as $\mathbb{V}^B = \{v_1^B, v_2^B, \dots, v_{N_B}^B\}$, where N_B is the number of buildings. For crowd flow of the i -th building at time slot t , we define it as $x_{i,t}^B \in \mathbb{R}$. As a result, we use $\mathbf{X}_t^B = [x_{1,t}^B, x_{2,t}^B, \dots, x_{N_B,t}^B] \in \mathbb{R}^{N_B}$ to represent all crowd flow observations for all buildings at time t .
- **Floor-Level Crowd Flow.** Similar to the building level, we denote the set of floors as $\mathbb{V}^F = \{v_1^F, v_2^F, \dots, v_{N_F}^F\}$, where N_F is the number of floors, and crowd flow of the i -th floor at time slot t as $x_{i,t}^F \in \mathbb{R}$. Furthermore, we use $\mathbf{X}_t^F = [x_{1,t}^F, x_{2,t}^F, \dots, x_{N_F,t}^F] \in \mathbb{R}^{N_F}$ to represent all crowd flow observations for all floors at time t .

Crowd flow prediction aims to predict the future campus-wide crowd flow (including both building and floor granularity) in the future Q time steps i.e., $\hat{\mathbf{X}}_{t+1:t+Q}^B$ and $\hat{\mathbf{X}}_{t+1:t+Q}^F$, given the historical crowd flow observations in the past P time steps $f(\mathbf{X}_{t-P+1:t}^B$ and $\mathbf{X}_{t-P+1:t}^F$).

5.2 Single-Grained Spatiotemporal GNN

To model the crowd flow within a single granularity (building or floor), we adopt the spatiotemporal GNN proposed in [64]. Here we use two identical but separate spatiotemporal GNNs for the building- and floor-level

crowd flow modeling, with different graph topology designs for building and floor levels. In the following, we first briefly present the spatiotemporal GNN, followed by our graph topology design.

5.2.1 Spatiotemporal GNN. The spatiotemporal GNN [64] consists of a gated temporal convolution and a graph convolution network, designed to capture the respective temporal and spatial dependencies of crowd flow. Without loss of generality, we present one layer below. First, it uses a temporal dilated convolution layer to extract temporal features, and also a gating mechanism to enhance the model’s capacity to learn complex temporal dependencies. Specifically, the gated temporal convolution is represented as follows:

$$\mathbf{H}^t = \tanh(\Phi * \mathbf{F}) \odot \sigma(\Psi * \mathbf{F}) \quad (1)$$

where Φ and Ψ are learnable parameters of convolution filters, \odot is the Hadamard product (element-wise multiplication), $\sigma(\cdot)$ is the sigmoid activation function which determines the ratio of information passed and $*$ denotes the dilated convolution operation. We use \mathbf{H}^t to represent the output of temporal convolution and \mathbf{F} to represent the input after linear transformation in the feature dimension. Afterward, it employ the multi-graph convolution to further capture the spatial dependencies on different graphs as follows:

$$\mathbf{H} = \sigma \left(\bigcup_{\mathbf{A} \in \mathbf{A}} f(\mathbf{A}; \theta_l) \mathbf{H}^t \mathbf{W}_l \right) \quad (2)$$

where \mathbf{A} is the set of graphs; in this work, we design two types of graphs, i.e., a static graph for capturing the location topology \mathbf{A}_{top} and a dynamically adaptive graph for capturing flow correlation \mathbf{A}_{adp} (more details below). \bigcup denotes the sum aggregation function. \mathbf{W}_l is transformation matrix of l -th layer. The function $f(\mathbf{A}; \theta_l)$ is selected as a polynomial function of order l of the graph Laplacian L , which leverages the spectral properties of graphs for feature transformation. The spatiotemporal GNN stacks a few number of such layers. To make the prediction, it incorporates skip connections that integrate features from each layer’s output, followed by two fully connected layers for the final prediction. In the following, we present our design of the static topology graph \mathbf{A}_{top} and a dynamically adaptive graph \mathbf{A}_{adp} .

5.2.2 Graph Construction. We consider two types of graphs for spatial dependencies of crowd flow. Intuitively, the crowd flow of two locations may be strongly associated if they are spatially adjacent. Therefore, we consider forming a two-level unified topology graph based on the physical proximity of spaces. Specifically, at the building level, we employ the Delaunay triangulation method [13], which is an effective approach widely used in Geographic Information Systems (GIS) for analyzing surface morphology. Each building is treated as an independent GPS-located point. These points are interconnected, forming a complex triangular network. A distinctive feature of this network is that the circumcircle of any triangle does not encompass any additional nodes. Utilizing this approach, pathways connecting the buildings within the campus are computed. The resulting building-level topology graph is shown on the left side of Figure 4. Here, the edges of the building-level topology graph are defined as:

$$\mathbf{A}_{top}^B[i, j] = \begin{cases} 1, & \text{if there is a path between } v_i^B \text{ and } v_j^B \\ 0, & \text{otherwise} \end{cases} \quad (3)$$

Subsequently, we consider the explicit hierarchy of the buildings and their contained floors in space, and establish the following rules based on the building-level topology graph: 1) Floors within the same building are deemed “directly accessible”; 2) Two floors from two different buildings connected by a path (on the building graph) or by a skyway, are considered “indirectly accessible”; 3) All other pairs of floors are deemed “inaccessible”. Following these definitions, for all pairs of floors that are “directly accessible”, connectivity is assumed to require only one hop (e.g., floor1-floor2); For all pairs of floors considered “indirectly accessible,” connectivity involves

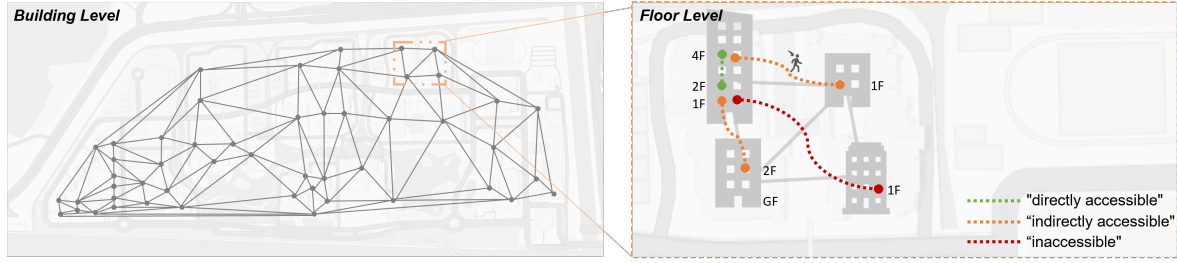


Fig. 4. Topology graphs of building and floor levels

three hops (floor1-building1-building2-floor2). Thus, we define the edges of the floor-level topology graph as:

$$\mathbf{A}_{top}^F[i, j] = \begin{cases} 1, & \text{if } v_i^F \text{ and } v_j^F \text{ are "directly accessible"} \\ \frac{1}{3}, & \text{if } v_i^F \text{ and } v_j^F \text{ are "indirectly accessible"} \\ 0, & \text{if } v_i^F \text{ and } v_j^F \text{ are "inaccessible"} \end{cases} \quad (4)$$

5.2.3 Adaptive graph. Besides the static graph that we built based on the heuristics of spatial adjacency, there may also exist implicit and non-obvious spatial dependencies within the campus that cannot be captured by the above method. To capture these implicit spatial dependencies, we incorporate two self-adaptive adjacency matrices (graphs) at building levels and floor levels, denoted as \mathbf{A}_{adp}^B and \mathbf{A}_{adp}^F , respectively. These matrices are designed to dynamically adjust to the data and are updated through backpropagation during the training process, enabling the discovery of latent spatial dependencies autonomously [64]. Without loss of generality, we take the adaptive graph at the building level as an example. Its self-adaptive adjacency matrix is derived from the learnt node embeddings, defined for source nodes as \mathbf{E}_1^B and for target nodes as \mathbf{E}_2^B , where $\mathbf{E}_1^B, \mathbf{E}_2^B \in \mathbb{R}^{N^B \times c}$. The self-adaptive adjacency matrix is computed as follows:

$$\mathbf{A}_{adp}^B = \text{SoftMax} \left(\text{ReLU}(\mathbf{E}_1^B \mathbf{E}_2^{B^T}) \right), \quad (5)$$

where the ReLU activation function serves to discard non-essential connections by filtering out non-positive weights, thereby focusing on stronger relationships. The softmax function is used for normalization.

5.3 Cross-Grained Information Transfer

Through the single-grained spatiotemporal GNNs, both building and floor nodes have encoded their spatiotemporal representations tailored to their respective granularities. Intuitively, the building-level crowd mobility patterns generally exhibit more macroscopic flow trends with stronger regularities, while at the floor level, the patterns are characterized by more detailed and specific local changes along with more fluctuations and noises. Subsequently, integrating regular spatiotemporal characteristics from the coarse granularity into the fine granularity can mitigate its fluctuations or noises, thus effectively smoothing the data and enhancing the robustness of the final predictions. Meanwhile, incorporating detailed spatiotemporal information from the fine granularity into the coarse granularity can enrich its spatiotemporal information, thus leading to more precise predictions. Based on this intuition, we design a dual-channel information transfer mechanism, with “floor to building” and “building to floor” channels, each employing unique strategies for effective information transfer.

5.3.1 Building to Floor. The information transfer from building to floor is designed to share the information of a building feature with its corresponding floor features. Intuitively, floor nodes that belong to the same building are likely closely related and share valuable spatiotemporal information, although their respective degrees of

correlation with the building may vary. For example, if a building's crowd flow increases over time, certain floors (such as the entrance floor or floors with key facilities) may also exhibit the same pattern, whereas less utilized floors may not. Furthermore, as crowd patterns evolve dynamically over time, the relationships among nodes between building level and floor level also fluctuate with time.

To address these complex cross-grained spatiotemporal patterns, we propose a Hierarchical Graph Attention Network (HGAT) module to effectively share information from the building level to the floor level. We dynamically compute weights based on the feature similarity between nodes at the building level and floor level to determine the extent of influence a building has on its floors and the degree to which each floor retains its characteristics. Specifically, we construct a hierarchical graph that includes nodes at both building and floor levels. As shown in Figure 3, for each floor node, two types of connections are established: 1) an edge linking it to its corresponding building node; and 2) a self-loop. The hidden features for building level nodes are denoted as $\mathbf{H}^B = [\mathbf{h}_1^B, \mathbf{h}_2^B, \dots, \mathbf{h}_{N_B}^B]$, and for floor level nodes as $\mathbf{H}^F = [\mathbf{h}_1^F, \mathbf{h}_2^F, \dots, \mathbf{h}_{N_F}^F]$. Our HGAT is formulated as follows:

$$\mathbf{h}_j^{F'} = \sigma \left(\alpha_{ij} \mathbf{W}_B \mathbf{h}_i^B + \alpha_{jj} \mathbf{W}_F \mathbf{h}_j^F \right) \quad (6)$$

where α_{ij} denotes the attention score of building i to floor j , and α_{jj} denotes the attention score of floor j to itself, which are computed as:

$$\alpha_{ij} = \frac{\exp(\text{LeakyReLU}(e_{ij}))}{\exp(\text{LeakyReLU}(e_{ij})) + \exp(\text{LeakyReLU}(e_{jj}))} \quad (7)$$

$$\alpha_{jj} = \frac{\exp(\text{LeakyReLU}(e_{jj}))}{\exp(\text{LeakyReLU}(e_{ij})) + \exp(\text{LeakyReLU}(e_{jj}))} \quad (8)$$

$$e_{jj} = \mathbf{b}^T \mathbf{W}_F \mathbf{h}_j^F, \quad e_{ij} = \mathbf{a}^T [\mathbf{W}_B \mathbf{h}_i^B \parallel \mathbf{W}_F \mathbf{h}_j^F] \quad (9)$$

where \mathbf{W}_B and \mathbf{W}_F are the weight matrices associated with the linear transformations for each building node and each floor node, respectively, and \mathbf{a} and \mathbf{b} are the weight parameters for attention output, which are jointly trained with the spatiotemporal GNNs.

5.3.2 Floor to Building. The information transfer from floor to building aims to propagate more comprehensive and detailed spatiotemporal information from the floor level to the building level. Intuitively, there is a quantitative hard constraint between the crowd flow at the building level and the floor level, where the total crowd flow of a building equals the sum of the crowd flow of all its floors. Analogically, the spatiotemporal feature of a building can also be collectively represented by the spatiotemporal features of all its floors. Here we follow a ‘‘aggregation and propagation’’ paradigm which uses a gating mechanism to update the spatiotemporal features at the building level. The workflow of our proposed approach is as follows. First, we perform an aggregation operation to compile the features of the floor nodes subordinate to each building node, obtaining the aggregated features for each building node as follows:

$$\mathbf{h}_{j,agg}^B = \bigcup_{i \in \mathbb{I}} \mathbf{h}_i^F \quad (10)$$

where \bigcup represents the aggregation function and \mathbb{I} is the set of floor nodes belonging to building v_j . We apply this operation on all buildings and ultimately concatenate the results to obtain the aggregated building-level hidden features $\mathbf{H}_{agg}^B \in \mathbb{R}^{N_B \times T \times d_B}$. Afterward, we implement a gating mechanism to control the degree of information propagating, adapting to the dynamic nature of crowd patterns. The building-level features $\mathbf{H}^{B'}$ could be updated as follows:

$$\mathbf{H}^{B'} = \mathbf{G} \odot \mathbf{H}_{agg}^B + (1 - \mathbf{G}) \odot \mathbf{H}^B \quad (11)$$

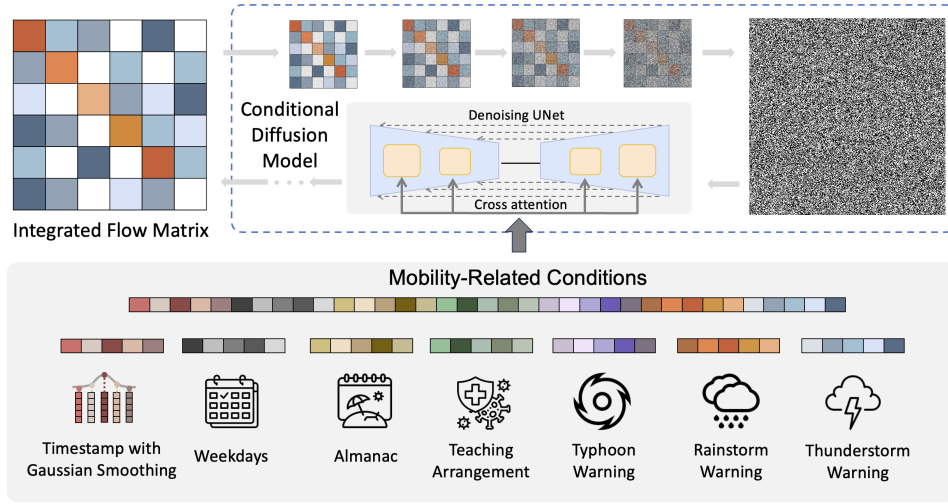


Fig. 5. Crowd flow simulation via conditional diffusion

where the update gate \mathbf{G} is computed as follows:

$$\mathbf{G} = \sigma \left(\mathbf{W}_g \cdot [\mathbf{H}_{agg}^B; \mathbf{H}^B] + \mathbf{b}_g \right) \quad (12)$$

\mathbf{W}_g and \mathbf{b}_g are learnable parameters of the update gate. Note that this gating mechanism dynamically integrates the features aggregated from the floors with the building's own features, thus enriching the feature representation at the building level.

6 CROWD FLOW SIMULATION

Crowd flow simulation aims to simulate on-campus crowd flow under given conditions. CrowdMirage is designed to provide the holistic simulation of on-campus crowd flow in a controllable manner beyond simply imitating the existing mobility patterns. In essence, such mobility patterns depend on various context factors, including but not limited to time in a day, day in a week, academic almanac, weather conditions, etc. In this context, we design our crowd flow simulation module as shown in Figure 5, we resort to a conditional diffusion model for the generative modeling of on-campus crowd and OD transition flow, and use mobility-related contexts to condition the generative model for the holistic simulation of on-campus mobility, which can thus simulate crowd flow under previously unseen scenarios (unseen combinations of different context factors). In the following, we first present our problem formulation, and then briefly introduce the conditional diffusion model, and then describe our detailed condition design, followed by our trajectory simulation to animate the generated crowd and OD transition flow.

6.1 Definition and Problem Formulation

In the crowd flow simulation task, we combine the crowd flow and OD transition flow between N_B buildings by inserting the crowd flow vector $\mathbf{X}_t = [x_{1,t}, x_{2,t}, \dots, x_{N_B,t}]$ into the diagonal entries of corresponding OD transition matrix $\mathbf{M}_t \in \mathbb{R}^{(N_B+1) \times (N_B+1)}$ in the same 10-minute time slot t , resulting in a so-called *integrated flow matrix* $\tilde{\mathbf{M}}_t \in \mathbb{R}^{(N_B+1) \times (N_B+1)}$. The series of all such integrated flow matrices over a time period T is denoted as $\mathcal{M} = \{\tilde{\mathbf{M}}_1, \tilde{\mathbf{M}}_2, \dots, \tilde{\mathbf{M}}_t\}$. Moreover, for each 10-minute time slot t , the mobility-related context factors could have

a significant impact on both the crowd flow and OD transition flow of the corresponding $\tilde{\mathbf{M}}_t$. Therefore, we define the corresponding context \mathbf{C}_t , and the collection of all contexts for \mathcal{M} is denoted as $\mathcal{C} = \{\mathbf{C}_1, \mathbf{C}_2, \dots, \mathbf{C}_T\}$. The crowd flow simulation task is to generate a series of integrated flow transition matrices $\tilde{\mathbf{M}}$ conditioned on the given mobility-related contexts \mathcal{C} . To this end, by considering the integrated flow matrix as an image, we resort to conditional diffusion models for image (integrated flow matrix $\tilde{\mathbf{M}}_t$) generation, conditioned on the given context factor \mathbf{C}_t .

6.2 Conditional Diffusion Models

6.2.1 Diffusion Models. Diffusion Models [28] consists of two inverse processes, i.e., a forward diffusion process and a reverse denoising process. In the diffusion part, the training data will be destroyed step by step through the progressive addition of Gaussian noise. Subsequently, the model learns to recover the destroyed data by the denoising process. Formally, diffusion models can be regarded as a type of latent variable model, as they map data to a latent space through a Markov chain process. This chain gradually adds noise to the data $\tilde{\mathbf{M}}_0$ by the forward process $q(\tilde{\mathbf{M}}_s | \tilde{\mathbf{M}}_{s-1})$ to obtain the approximate posterior $q(\tilde{\mathbf{M}}_{1:S} | \tilde{\mathbf{M}}_0)$ as follows:

$$q(\tilde{\mathbf{M}}_s | \tilde{\mathbf{M}}_{s-1}) = \mathcal{N}(\tilde{\mathbf{M}}_s; \sqrt{1 - \beta_s} \tilde{\mathbf{M}}_{s-1}, \beta_s I), \quad (13)$$

$$q(\tilde{\mathbf{M}}_{1:S} | \tilde{\mathbf{M}}_0) = \prod_{s=1}^S q(\tilde{\mathbf{M}}_s | \tilde{\mathbf{M}}_{s-1}), \quad (14)$$

where $\{\beta_s | s = 1, \dots, S\}$ is the variance schedule and S represents the diffusion steps. At the end of the diffusion steps, the data is converted into pure Gaussian noise. The goal of training a diffusion model is to learn the reverse process by learning $p_\theta(\tilde{\mathbf{M}}_{s-1} | \tilde{\mathbf{M}}_s)$, a Gaussian distribution whose mean and variance need to be calculated. Starting with the pure Gaussian noise $p(\tilde{\mathbf{M}}_S) = \mathcal{N}(\tilde{\mathbf{M}}_S; \mathbf{O}, I)$, the whole denoising process can be presented as:

$$p_\theta(\tilde{\mathbf{M}}_{s-1} | \tilde{\mathbf{M}}_s) = \mathcal{N}(\tilde{\mathbf{M}}_{s-1}; \mu_\theta(\tilde{\mathbf{M}}_s, s), \Sigma_\theta(\tilde{\mathbf{M}}_s, s)), \quad (15)$$

$$p_\theta(\tilde{\mathbf{M}}_{0:S}) = p(\tilde{\mathbf{M}}_S) \prod_{s=1}^S p_\theta(\tilde{\mathbf{M}}_{s-1} | \tilde{\mathbf{M}}_s), \quad (16)$$

The efficacy of diffusion models in generating data is largely attributed to their alignment with the inductive biases inherent in image-like data, particularly when the core neural architecture is configured as a UNet [15, 53].

6.2.2 Conditioning Mechanisms. Diffusion models possess the capability to model conditional distributions. In mainstream conditional generation tasks, the cross-attention mechanism [52] is introduced to the core architecture UNet to guide the generation process in a controllable manner. It is an effective way to handle various condition modalities in the learning attention-based model process. Specifically, the conditions can be encoded into a latent space and jointly trained along with the denoising process.

6.2.3 Integrated Flow Matrix to Image Adaptation. The integrated flow matrix is a square matrix, which encodes both crowd flow and OD transition flow. This matrix can be regarded as a grey-scale image with a unique channel. However, the uint8 image matrix contains intensity values between 0 and 255 meanwhile the max values of the integrated crowd flow transition matrix are far more than 255 for both crowd flow and OD transition flow among all statistics. Therefore, we conduct min-max normalization on the crowd flow and OD transition flow before feeding the matrix to diffusion models.

Table 4. Categories Indexing

Context	Values
Timestamps	144 timestamps
Weekdays	Mon / Tue / Wed / Thu / Fri / Sat / Sun
Almanac	no day-off / examination / holiday / vacation / recess
Teaching arrangement	face-to-face / online / hybrid / suspended
Tropical cyclone	no signal / No. 1 / No. 3 / No. 8 and above
Rainstorm	no signal / Amber / Red / Black / No. 8 and above
Thunderstorm	no warning / warning / No. 8 and above

6.3 Mobility-Related Condition Design

The mobility patterns behind the crowd flow and OD transition flow are highly correlated with their contexts, which could be used for the conditional generation of the integrated crowd and OD transition flow matrices. Considering the practical environments of our studied university campus, we choose the following context factors of seven categories: Timestamps, Weekdays, Almanac, Teaching Arrangement, and Weather Conditions including Tropical Cyclone Warning Signal, Thunderstorm Warning Signal, and Rainstorm Warning Signal as shown in Table 4. These seven context factors are represented as learnable embeddings in a latent space so as to condition the conditional diffusion models by joint training with the integrated flow matrices in UNet. We present the detailed condition design below:

6.3.1 Timestamp. This context factor captures the crowd flow patterns at different time slots of the day. For instance, we usually observe a rush of students moving to the classroom buildings right before classes begin, or a large crowd gathering toward canteens before lunchtime. Therefore, according to our 10-minute granularity of crowd flow, we consider a total number of 144 timestamps for a day, where each timestamp corresponds to a time slot. However, such timestamps might be over-specific, resulting in a sudden shift of generated crowd flow between two neighboring timestamps. To alleviate this issue, we employ the Gaussian weighted averaging [14] to generate smoothed timestamp embeddings. Specifically, for each timestamp k (and its timestamp embedding \vec{e}_k), it generates a smoothed time stamp embedding \vec{t}_k as the weighted average of its neighboring timestamps' embeddings, where the weight is defined by a Gaussian with a learnable bandwidth σ_k :

$$\vec{t}_k = \frac{\sum_i f_k(i) \cdot \vec{e}_k}{\sum_i f_k(i)}, \text{ where } f_k(i) = \frac{1}{\sigma_k \sqrt{2\pi}} e^{-\frac{dist(i,k)^2}{2\sigma_k^2}} \quad (17)$$

where $dist(i, k)$ computes the distance between k and i timestamps in a cycling loop:

$$dist(i, k) = \begin{cases} |i - k|, & \text{if } |i - k| < 72 \\ 144 - |i - k|, & \text{otherwise} \end{cases} \quad (18)$$

6.3.2 Weekdays. This context factor differentiates different days of the week, as the crowd flow patterns are significantly different between these days. For instance, there might be regular crowd flow patterns of students and staff moving on campus due to regular classes and activities during weekdays, while during weekends the crowd flow patterns might be less regular because of the increase in recreational activities.

6.3.3 Almanac. This context factor takes into account the academic calendar, which is highly relevant to on-campus crowd flow patterns. For instance, during public holidays or semester breaks, there is usually a clear drop in the crowd flow on campus; during examination periods, there is often an increase in the crowd flow toward

libraries and study rooms. Considering the length of the days off would impact the leaving intention of campus personnel, we design the following five non-overlapped values for this context:

- No day-off: regular date without day-off
- Examination: exam period at the end of each semester
- Holiday: day-off of one single day
- Vacation: day-offs more than or equal to three days (including weekends)
- Recess: day-offs more than 15 days, such as summer break and winter break of students

6.3.4 Teaching Arrangement. The COVID-19 pandemic has significantly impacted the way universities conduct their classes, which in turn, has led to different crowd flow patterns. During the local epidemic outbreak, following the government's directives, the university modified the course arrangements to comply with the regulations. During the teaching mode transitions from 2022 to 2023, the university adopted the following teaching arrangements to adapt to the changing circumstances:

- Face-to-Face: This is the traditional teaching mode where all students are required to attend classes on campus.
- Online: In this mode, all students take classes online.
- Hybrid: This mode is a blend of face-to-face and online teaching. Most students attend classes on campus, while some infected or quarantined students attend the same classes online.
- Suspended: Due to a severe outbreak of the pandemic on campus or in response to government directives, all classes are temporarily halted.

6.3.5 Tropical Cyclone Warning Signal. A tropical cyclone is also known as a typhoon, which usually comes with heavy wind and rains and could cause destructive disasters, and thus has a strong impact on the crowd mobility patterns. To protect citizens' safety, the meteorological bureau of the local government would issue an early warning to the citizens and take different preventive measures based on the predicted typhoon levels. We consider the following values for this context:

- No.1 Tropical Cyclone Warning Signal: No special measures are taken. This signal serves as a warning that a typhoon is forming, which is regarded as a preliminary alert to prepare for potential severe weather conditions.
- No.3 Tropical Cyclone Warning Signal: This signal may cause a partial suspension of classes (for a certain period) based on the assessment of the situation, such as the wind scale, the forecasted duration, and the potential risks to students and staff, following the announcement of the government.
- No.8 Tropical Cyclone Warning Signal and above: The government takes comprehensive measures to minimize potential damage from the typhoon, including suspending all teaching activities, entering an immediate prevention state, and halting all public transportation.

6.3.6 Rainstorm Warning Signal. Rainstorms can also impact crowd mobility patterns. Following the instructions from the meteorological bureau of the local government, a rainstorm warning signal is issued when heavy rain is observed or predicted. This system operates independently of other warnings such as thunderstorm warning signals or tropical cyclone signals (below No.8). For the No.8 tropical cyclone warning signal and above, the rainstorm and thunderstorm warning signals are not issued anymore because they are by default associated with the tropical cyclone warning signal. Therefore, we consider the following values for this context:

- Amber Rainstorm Warning Signal: Issued when the rainfall is expected to reach about 20mm in an hour or has already reached 20mm in the past hour, and the rain is expected to continue.
- Red Rainstorm Warning Signal: Issued when the rainfall is expected to reach about 50mm in an hour or has already reached 50mm in the past hour, and the rain is expected to continue.

- Black Rainstorm Warning Signal: Issued when the rainfall is expected to reach about 80mm in an hour or has already reached 80mm in the past hour, and the rain is expected to continue.
- No.8 Tropical Cyclone Warning Signal and above: The rainstorm and thunderstorm warning signals are not issued anymore, overwritten by this extreme tropical cyclone warning signal.

6.3.7 Thunderstorm Warning Signal. This signal is issued when a thunderstorm is observed or predicted to occur. Upon the issuance of this signal, appropriate self-preventive measures should be taken, including staying in safe indoor locations and avoiding outdoor activities, which thus impact crowd mobility patterns. Note that similar to the rainstorm warning signals, the thunderstorm warning signal will not be used when a No.8 tropical cyclone warning or above is issued. Therefore, we consider the following values for this context:

- Thunderstorm Warning Signal
- No.8 Tropical Cyclone Warning Signal and above

We extract the above conditions for each integrated flow matrix. Specifically, the almanac and teaching arrangement are available on the university website and the historical record of weather conditions can be found on the Meteorological Bureau website. All the context factors are marked with timestamps aligned with the time slot of the integrated flow matrix. Therefore, each flow matrix is conditioned by a combination of the seven contexts (the concatenation of the seven respective embedding vectors), which are trained together with the UNet in the conditional diffusion model.

6.4 Trajectory Simulation

To dynamically visualize the generated crowd flow, we design a trajectory simulation method to animate the generated crowd flow. Specifically, the generated flow matrices encode the crowd and OD transition flow at a 10-minute time granularity, which is not intuitive to be dynamically visualized as a video playback for example. Therefore, we simulate individual origin-destination trajectories in a finer time granularity of 1 minute from the OD transition flow within each time slot of 10 minutes. To this end, we design a neural network to model to fit the probability distribution of departure time and travel time. More precisely, for each integrated flow matrix, we first conduct a statistical analysis to identify the best distributional assumption for the departure time and travel time, and then use a fully connected neural network to estimate the parameters of the probability distributions based on the context, origin and destination. Subsequently, for each origin-destination flow in the generated flow matrix, we can predict probability distributions of the departure time and travel time, and sample the corresponding amount of trajectories for visualization.

6.4.1 Statistical Analysis. We conduct a statistical analysis of the real trajectories to identify the best distributional assumption for the departure time and travel time. For each origin-destination pair in the integrated flow matrix, we get the real departure time (0-9th minute within the 10-minute time slot) and travel time on 1-minute granularity for all the corresponding trajectories. We then test a wide range of 80 different probability distributions² by fitting each of them to the departure time and the travel time, using the Python package “fitter”. We then select the best distribution according to the sum of the square errors. This statistic analysis reveals that the best distributional assumption for departure time is the truncated-normal distribution with a range constraint of [0, 9], where its PDF is as follows:

$$f(t_{dept} | \mu_{dept}, \sigma_{dept}, a, b) = \frac{\phi\left(\frac{t_{dept} - \mu_{dept}}{\sigma_{dept}}\right)}{\sigma_{dept} \left(\Phi\left(\frac{b - \mu_{dept}}{\sigma_{dept}}\right) - \Phi\left(\frac{a - \mu_{dept}}{\sigma_{dept}}\right) \right)} \quad (19)$$

²<https://fitter.readthedocs.io/en/latest/faqs.html#what-are-the-distributions-available>

Table 5. Statistics of the datasets

Year	2022	2023
#Buildings	53	
#Floors	391	
#WiFi Logs	1,773,027,241	2,395,657,469
Crowd Flow Range	0-1,241	0-1,786
OD Flow Range	0-503	0-530
All Flow Range	0-16,268	0-20,396

where $a = 0$, $b = 9$, $a \leq t_{dept} \leq b$, μ_{dept} is the location parameter, $\sigma_{dept} > 0$ is the scale parameter, ϕ is the probability density function of the standard normal distribution, Φ is the cumulative distribution function of the standard normal distribution. Meanwhile, the best distributional assumption for travel time is found to be the lognormal distribution, where its PDF is defined as follows:

$$f(t_{trav}|\mu_{trav}, \sigma_{trav}) = \frac{1}{t_{trav}\sigma_{trav}\sqrt{2\pi}} \exp\left(-\frac{(\ln t_{trav} - \mu_{trav})^2}{2\sigma_{trav}^2}\right) \quad (20)$$

where travel time $t_{trav} > 0$, μ_{trav} is the location parameter, $\sigma_{trav} > 0$ is the scale parameter.

6.4.2 Neural Parameter Estimation. Based on the above analyses, we use a fully connected neural network to learn to predict the parameters μ_{dept} and σ_{dept} for the truncated-normal distribution of departure time, and the parameters μ_{trav} and σ_{trav} for the lognormal distribution of travel time, based on the input context (as presented in the Section 6.3, the origin and the destination (represented as learnable embeddings)). The network is trained to minimize the negative log-likelihood loss, which is equivalent to the maximum likelihood estimation for fitting probability distributions to data. In the simulation process, for each origin-destination flow in the generated flow matrix, we first predict parameters of the probability distributions of the departure time and travel time and then sample the corresponding amount of trajectories.

7 EXPERIMENTS

In this section, we first present our experimental setup including dataset statistics, baselines, and evaluation protocol, followed by the performance of individual tasks with both quantitative results and case studies.

7.1 Experimental Setup

7.1.1 Dataset. We collected crowd flow data in our university campus for the past two years 2022-2023. Table 5 shows the statistics of our collected data. Note that the flow range is computed over a whole year. All flow here refers to the total populations (sum) of both crowd flow and OD transition flow in a time slot.

To evaluate our crowd mobility flow extraction pipeline, we resort to the user occupation history of our university library, which has the time series of the number of users inside the library's restricted area, where users need to scan their access cards; the number of users is measured by the access card scanning records. This data is collected every 10 minutes over a period of two months. Note that there is a cafeteria and a shop located in the library building but in the public area (no access card needed). Subsequently, the library user occupation data does not include the users in these two places. In other words, the library user occupation data might slightly underestimate the actual users inside the library building. In contrast, our estimated crowd flow based on Wi-Fi connection records covers all these places.

7.1.2 Baselines. We consider the following state-of-the-art baselines for crowd flow prediction and crowd flow simulation tasks:

- *Crowd flow prediction baselines:* **HA** predicts future values using the mean of past observations. **SVR** [18] fits an optimal hyperplane within a tolerance margin, accommodating both linear and non-linear regression. **RNN** [20] utilizes an internal state to process sequential inputs, ideal for time-dependent data. **LSTM** [29] enhances RNNs to remember inputs over long periods, mitigating long-term dependency issues in sequence prediction. **GRU** [11] simplifies LSTMs with a gating mechanism to address the vanishing gradient problem while capturing temporal dependencies. **DCRNN** [36] utilizes a diffusion convolutional recurrent neural network architecture to capture both spatial and temporal dependencies in crowd flow data, enhancing prediction accuracy through sequence-to-sequence learning. **ASTGCN** [25] utilizes a three-branch architecture incorporating attention mechanisms and graph convolutions to analyze multi-scale time features. **STGCN** [70] integrates dual temporal convolutions and a graph convolution to learn spatiotemporal dependencies. **GWNET** [64] employs an adaptive adjacency matrix and stacked dilated causal convolutions to model spatial-temporal dependencies effectively. **D2STGNN** [54] employs a decoupled framework that separates the diffusion and inherent signals in crowd flow data, and integrates a dynamic graph learning module for evolving spatial-temporal dependencies. **STWave** [21] leverages a spectral graph attention mechanism combined with wavelet transform to capture both spatial and temporal dynamics in a unified framework.
- *Crowd flow simulation baselines:* **GM** (Gravity Model) [2] is inspired by Newton's law of Gravitation; it assumes that the magnitude of mobility flow is directly influenced by the populations of the origin and destination, and inversely influenced by the spatial distance separating them. **DGM** [56] (Deep Gravity Model) introduces deep neural networks into GM, better modeling the relationships between geographic features and mobility flows. **CGAN** (Conditional Generative Adversarial Nets) [45] uses a generative model and a discriminative model for adversarial learning with the incorporation of conditional information into the generator and discriminator. **CVAE** [26] (Conditional Variational Auto-Encoder) adapts the generative process by using additional inputs to guide data generation; the encoder and decoder are modified to condition these inputs, enabling controlled data output that meets specific constraints or attributes. **DM** (Diffusion Model) [28] is the standard diffusion model as a baseline for comparison; it is trained to generate the integrated flow matrix without any conditions. **CDMnoGS** (Conditional Diffusion Model without Gaussian Smoothing) is one variant of our crowd flow simulation method without Gaussian smoothing on the timestamp embeddings and using the standard timestamp embedding instead.

7.1.3 *Evaluation Protocols.* Our evaluation protocols for the crowd mobility flow extraction, crowd flow prediction task and crowd simulation task are as follows.

- **Crowd mobility flow extraction.** We extract the estimated crowd flow and OD transition flow of the library building, and compare them with the library user occupation history. Specifically, for crowd flow, we directly measure the Mean Average Error (MAE) between our estimated crowd flow and the user occupation data every 10 minutes. For OD transition flow, we compute the net change in the number of users (total inflow subtracting the total outflow) of the library building every 10 minutes, and compare it against the change in the number of users from the library user occupation history. We experimentally verify the key design choices of our crowd mobility flow extraction pipeline via an ablation study.
- **Crowd flow prediction task.** In the crowd flow dataset, there are 53 nodes (buildings) at the building level and 391 nodes (floors) at the floor level. To ensure that the model can fully capture the typical differences between the academic term and non-term periods, We conduct experiments using real crowd flow data collected in 2023, dividing the dataset into 12 parts, each representing a month, and further splitting each part into 70% for training, 10% for validation, and 20% for testing. We train the model on the training set and select the optimal parameters for each granularity based on the best results of the validation sets at two different granularities. Subsequently, we test the model on the test set. We report the performance on three widely used metrics, i.e., Mean Absolute Error (**MAE**) and Weighted Mean Absolute Percentage Error

Table 6. Test scenarios for crowd flow simulation

Test scenarios	Timestamps	Weekdays	Almanac	Teaching	Weather warning
1. "Semester weekday" (31 days)	1-144	Mon	no day-off	face-to-face	no signal
2. "Semester weekend" (36 days)		Sat	no day-off	face-to-face	no signal
3. "Hybrid teaching" (30 days)		Fri	no day-off	hybrid	no signal
4. "Student recess" (18 days)		Thur	recess	face-to-face	no signal
5. "Suspended course" (1 days)		Sun	no day-off	suspended	no signal
6. "Typhoon" (1 days)		Sun	no day-off	face-to-face	No.8 typhoon signal
7. "Exam&Rain" (1 days)		Tues	examination	face-to-face	Black rainstorm signal
8. "Exam&Thunder" (2 days)		Mon	examination	face-to-face	Thunderstorm signal
9. "Holiday&Hybrid" (1 days)		Wed	holiday	hybrid	no signal

(**WMAPE**). Note that WMAPE weights the absolute percentage errors by the actual values at each data point; this feature enables the comparison of performance across different granularities using this metric. In our experiments, we utilize a multi-step forecasting approach, where we use observations of crowd flow from the past two hours (12 time steps) to predict the crowd flow for each time step within the upcoming two hours.

- **Crowd flow simulation task.** We consider the simulation of daily crowd flow in this task, where we generate 144 integrated flow matrices, each for the corresponding 10-minute time slot. We consider 9 representative conditions as our test scenarios, out of all 206 historical conditions (unique combinations of context factors appearing in the past two years), as shown in Table 6. We generate 144 integrated flow matrices for each of the test cases by feeding the corresponding conditions to our trained model, denoted as $\widehat{\mathcal{M}} = \{\widehat{\mathcal{M}}_1, \widehat{\mathcal{M}}_2, \dots, \widehat{\mathcal{M}}_{144}\}$. Unconditional baselines generate matrices directly without using these conditions. We then compare $\widehat{\mathcal{M}}$ with the ground truth matrices by reporting the following metrics. Jensen-Shannon Divergence (**JSD**) quantifies the similarity between the generated matrices and the ground truth by measuring the divergence between their respective probability distributions; a lower JSD value indicates better performance. Common Part of Commuters (**CPC**) evaluates the overlap in commuter patterns between the generated and actual OD matrices, which serves as an indicator of the model's ability to replicate observed commuting behaviors; a higher CPC value indicates better performance. Meanwhile, we also report two typical error-based metrics Root Mean Square Error (**RMSE**) and Normalized Root Mean Square Error (**NRMSE**). In the evaluation, we report the performance on the crowd flow vectors \mathbf{X} and OD transition matrices \mathbf{M} separately.

7.2 Crowd Mobility Flow Extraction Performance

To evaluate our crowd mobility flow extraction pipeline, we experimentally verify the key design choices via an ablation study. We consider the following five variants ablated from our full crowd mobility flow extraction pipeline: w/o guest devices filtering, w/o non-mobile devices filtering, w/o public devices filtering, w/o randomized MAC devices filtering, and w/o primary device selection. Please refer to Section 4.3 for details. Table 7 shows the results. We observe that our full pipeline consistently outperforms all ablated variants, showing an average improvement of 31.53% and 2.23% on crowd flow estimation and OD transition flow estimation, respectively.

Moreover, we see that the MAE on the crowd flow is larger than the OD transition flow. This is partially because the library user occupation data tends to have a systematic bias of underestimation of the crowd flow, due to the presence of users at the cafeteria and the shop in the public area of the library. To further understand this issue, we plot the estimated crowd flow and the library occupation data on four typical days as shown in Figure 6. First, we observe that compared to the ablated variants, our full pipeline shows the closest results to the

Table 7. Crowd mobility flow extraction performance in MAE(↓)

Flow Estimation	Crowd Flow Estimation	OD Transition Flow Estimation
w/o guest devices filtering	32.38	4.55
w/o non-mobile devices filtering	61.68	4.87
w/o public devices filtering	23.41	4.41
w/o randomized MAC devices filtering	23.26	4.41
w/o primary device selection	31.02	4.49
CrowdMirage	21.24	4.39

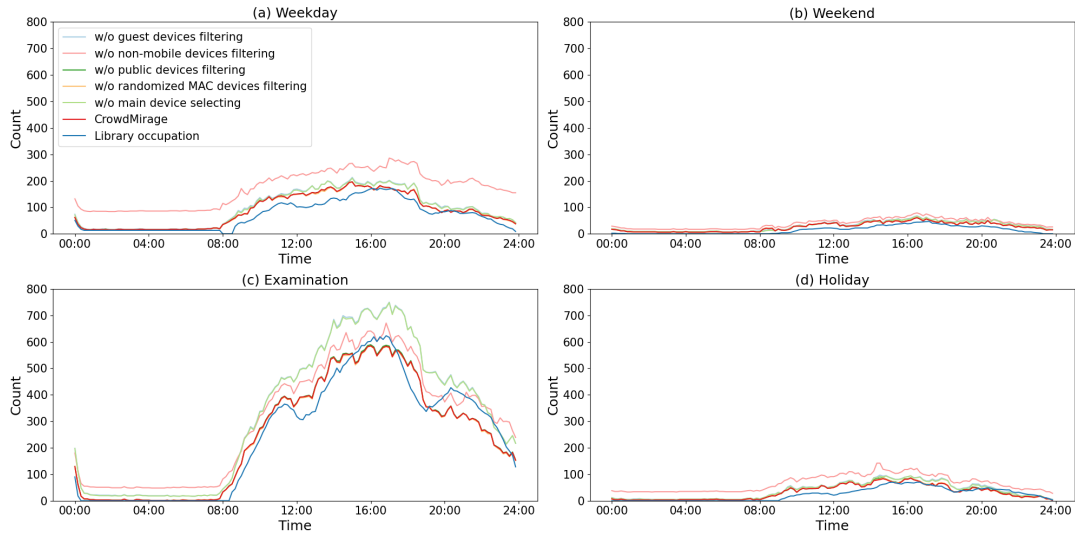


Fig. 6. Comparison between the estimated crowd flow and the library occupation data on four typical days: (a) Weekday, (b) Weekend, (c) Examination, and (d) Holiday.

library occupation data. We also observe that our estimation is generally higher than the library occupation data in the daytime, due to the presence of the underestimation bias of the library occupation data. In contrast, the OD transition flow measures the net change in the number of users, which is less sensitive to this consistent bias.

7.3 Crowd Flow Prediction Performance

7.3.1 Comparison with baselines. Table 8 shows the crowd flow prediction performance of different methods for the future time steps at 10 minutes (1 step), 30 minutes (3 steps), 1 hour (6 steps), and 2 hours (12 steps). We observe that our proposed method achieves the best results at both building and floor levels across different prediction time steps in general. In particular, compared to the GWNET, which serves as our backbone spatiotemporal GNN, our method shows an average improvement of 9.09% and 3.35% on the building and floor levels, respectively; the results validate the effectiveness of our proposed cross-grained information transfer mechanisms, which can indeed benefit the crowd flow prediction on both building and floor levels.

7.3.2 Ablation study. To further validate our proposed cross-grained information transfer module, we conduct an ablation study with the following two variants 1) without the “building to floor” transfer module (denoted

Table 8. Crowd flow prediction performance of different methods

Granularity Method	10 min		30 min		1 hour		2 hour			
	MAE	WMAPE	MAE	WMAPE	MAE	WMAPE	MAE	WMAPE		
Building level	HA	15.93	13.35	15.93	13.35	15.93	13.35	15.93	13.35	
	SVR	11.53	9.60	13.74	11.45	17.06	14.23	23.85	19.97	
	RNN	12.39±0.16	16.23±0.21	12.53±0.27	16.43±0.36	13.06±0.26	17.16±0.34	14.28±0.15	18.86±0.19	
	LSTM	11.35±0.12	14.87±0.16	11.57±0.13	15.18±0.17	11.83±0.15	15.54±0.19	12.71±0.13	16.78±0.17	
	GRU	11.30±0.12	14.81±0.16	11.52±0.09	15.11±0.11	11.89±0.09	15.62±0.12	13.11±0.17	17.31±0.23	
	DCRNN	3.86±0.01	5.05±0.01	6.98±0.01	9.16±0.01	10.21±0.02	13.41±0.03	15.42±0.03	20.37±0.04	
	ASTGCN	4.70±0.05	6.12±0.06	6.31±0.16	8.22±0.09	7.55±0.14	9.85±0.18	9.50±0.22	12.44±0.29	
	STGCN	4.30±0.13	5.60±0.17	6.18±0.08	8.05±0.10	8.07±0.09	10.52±0.11	10.92±0.19	14.30±0.25	
	GWNET	3.85±0.07	5.00±0.09	5.28±0.05	6.88±0.06	6.35±0.07	8.29±0.09	8.09±0.14	10.60±0.18	
	STWave	4.07±0.10	5.30±0.14	5.52±0.10	7.19±0.13	6.80±0.24	8.87±0.31	8.43±0.14	11.04±0.19	
	D2STGNN	3.33±0.06	4.29±0.08	4.82±0.06	6.16±0.07	6.35±0.24	8.13±0.31	8.33±0.18	10.71±0.23	
	CrowdMirage	3.30±0.02	4.30±0.02	4.76±0.01	6.20±0.01	5.97±0.01	7.79±0.02	7.58±0.04	9.92±0.05	
	Floor level	HA	4.05	25.05	4.05	25.05	4.05	25.05	4.05	25.05
		SVR	2.27	13.97	2.73	16.79	3.28	20.19	4.32	26.69
RNN		3.19±0.02	30.79±0.17	3.07±0.01	29.67±0.01	3.09±0.01	29.92±0.12	3.17±0.01	30.90±0.09	
LSTM		2.94±0.01	28.40±0.11	2.92±0.01	28.25±0.05	2.93±0.00	28.44±0.01	2.98±0.01	29.04±0.07	
GRU		2.99±0.19	28.87±0.18	2.93±0.01	28.31±0.09	2.92±0.01	28.29±0.06	2.97±0.00	28.97±0.04	
DCRNN		1.17±0.01	11.32±0.06	1.76±0.00	17.02±0.06	2.29±0.01	22.16±0.08	3.12±0.01	30.42±0.13	
ASTGCN		1.25±0.00	12.00±0.00	1.69±0.02	16.23±0.21	2.02±0.05	19.41±0.48	2.41±0.05	23.33±0.49	
STGCN		1.13±0.01	10.85±0.09	1.61±0.01	15.49±0.08	2.00±0.01	19.27±0.08	2.55±0.02	24.64±0.18	
GWNET		1.07±0.00	10.31±0.02	1.46±0.00	14.00±0.02	1.74±0.01	16.76±0.06	2.07±0.01	20.04±0.07	
STWave		1.27±0.00	12.18±0.02	1.73±0.01	16.63±0.08	2.14±0.01	20.58±0.09	2.71±0.04	26.16±0.37	
D2STGNN		1.25±0.00	11.55±0.09	1.69±0.01	14.99±0.15	2.00±0.01	17.45±0.22	2.41±0.02	20.82±0.36	
CrowdMirage		1.03±0.00	9.86±0.03	1.42±0.00	13.60±0.04	1.69±0.00	16.22±0.03	2.00±0.00	19.32±0.01	

w/o b2f), and 2) the variant without the “floor to building” transfer module (denoted as w/o f2b). The results are shown in Table 9, we observe that our dual-channel information transfer mechanism shows the best results in most cases. Compared to w/o b2f, it shows an average improvement of 2.50% and 2.06% on the building and floor levels, respectively; compared to w/o f2b, it also shows an average improvement of 6.56% and 2.06% on the two respective levels. The results verify our design choices for the dual-channel information transfer mechanism.

7.3.3 Case study. We conduct a case study to show the crowd flow prediction performance with one prediction step (10 minutes in the future) on two typical days (a semester weekday and a recess weekday). The left panel of Figure 7 shows predicted and ground truth crowd flow time series for two typical buildings E6 (a teaching building) and RC9 (a residential college). We observe that the two buildings with different functions yield distinct crowd flow time series over time, while the same weekday within the semester also yields different patterns from the student recess period. Despite these variations, our predicted crowd flow closely resembles the ground truth. Moreover, we also show two snapshots of the predicted crowd flow at the same time 13:00 on the two respective days on the right panel of Figure 7. We observe that the spatial distributions of the crowd flow show a clear difference due to the mobility pattern differences between the semester and recess periods.

Table 9. Ablation study on crowd flow prediction performance

Granularity	Method	10 min			30 min			1 hour			2 hour		
		MAE	RMSE	WMAPE	MAE	RMSE	WMAPE	MAE	RMSE	WMAPE	MAE	RMSE	WMAPE
Building level	w/o b2f	3.58	7.10	4.30	4.88	9.69	6.36	6.17	12.60	8.05	7.71	15.69	10.10
	w/o f2b	3.76	7.63	4.89	5.14	10.26	6.69	6.20	12.57	8.09	7.92	16.39	10.37
	CrowdMirage	3.30	6.91	4.30	4.76	9.63	6.20	5.97	12.09	7.79	7.58	15.70	9.92
Floor level	w/o b2f	1.05	2.52	9.99	1.45	3.35	13.91	1.72	3.98	16.59	2.04	4.60	19.71
	w/o f2b	1.02	2.44	9.82	1.42	3.36	13.61	1.69	3.94	16.29	2.04	4.57	19.67
	CrowdMirage	1.03	2.46	9.86	1.42	3.29	13.60	1.69	3.88	16.22	2.00	4.48	19.32

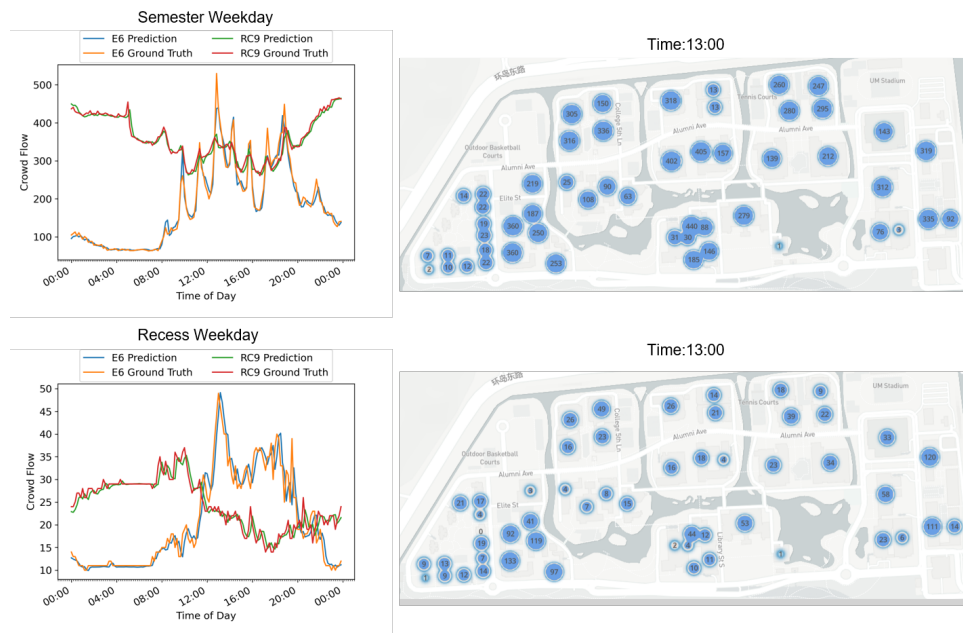


Fig. 7. Case study of crowd flow prediction on two typical days (semester weekday on top and recess weekday on bottom). Left: crowd flow series for two buildings where E6 is a teaching building and RC9 is a residential college. Right: heatmaps of predicted crowd flow at 13:00.

7.4 Crowd Flow Simulation Performance

7.4.1 Comparison with baselines. Table 10 shows the crowd flow simulation performance of different methods. Besides the overall performance on the 9 test scenarios, we also show the results on two typical scenarios 1 and 7, where scenario 1 is a common scenario “Semester weekday” appearing 31 days while scenario 7 is a rare scenario appearing only once. Note that GM and DGM only support the OD transition flow generation because the models only calculate the inflow/outflow probability between different buildings and the diagonal of the OD matrix is designed to encode the staying flow. The results show that CrowdMirage achieves the best overall performance, with an average improvement of 8.80% and 68.92% over the best-performing baselines on crowd flow

Table 10. Crowd flow and OD transition flow simulation performance

Condition	Model	Crowd Flow				OD Transition Flow			
		JSD ↓	CPC ↑	RMSE ↓	NRMSE ↓	JSD ↓	CPC ↑	RMSE ↓	NRMSE ↓
1. “Semester weekday” (31 days)	GM	/	/	/	/	0.0404	0.0626	12.5276	0.0236
	DGM	/	/	/	/	0.0044	0.1476	2.1015	0.0040
	CGAN	0.2539	0.5319	112.1724	0.0628	0.7177	0.0360	17.2060	0.0325
	CVAE	0.1484	0.6870	86.3173	0.0483	0.0157	0.4127	1.6965	0.0032
	DM	0.1963	0.5465	114.6074	0.0642	0.0163	0.2730	2.0257	0.0038
	CDMnoGS	0.1890	0.6962	83.1207	0.0465	0.0740	0.2922	1.6652	0.0031
	CrowdMirage	0.1613	0.7072	83.1027	0.0465	0.0038	0.5192	1.4855	0.0028
7. “Exam&Rain” (1 days)	GM	/	/	/	/	0.0231	0.0354	4.7949	0.0090
	DGM	/	/	/	/	0.2191	0.0454	9.1194	0.0172
	CGAN	0.2140	0.3714	97.8028	0.0548	0.3767	0.0090	27.7273	0.0523
	CVAE	0.3347	0.4271	111.4750	0.0624	0.1073	0.1844	2.5037	0.0047
	DM	0.2064	0.4999	86.3562	0.0484	0.0103	0.2975	1.0665	0.0020
	CDMnoGS	0.2039	0.7963	20.4757	0.0115	0.0649	0.1743	0.8021	0.0015
	CrowdMirage	0.1071	0.8302	22.4006	0.0125	0.0014	0.5225	0.5514	0.0010
Overall	GM	/	/	/	/	0.0305	0.0398	9.4434	0.0178
	DGM	/	/	/	/	0.0629	0.1216	3.1918	0.0060
	CGAN	0.2514	0.5025	93.5117	0.0524	0.6780	0.0185	18.4088	0.0347
	CVAE	0.1862	0.6573	71.8791	0.0402	0.0318	0.3639	1.3250	0.0025
	DM	0.1760	0.5416	99.0566	0.0555	0.0118	0.2875	1.3089	0.0025
	CDMnoGS	0.1573	0.7801	43.3767	0.0243	0.0630	0.2216	1.0261	0.0019
	CrowdMirage	0.1085	0.8007	43.0348	0.0241	0.0020	0.5143	0.7861	0.0015

and OD transition flow simulation respectively. Moreover, for the two typical scenarios 1 and 7, CrowdMirage also achieves superior performance in general, showing its effectiveness in simulating both common and rare conditions.

7.4.2 Ablation study on mobility-related condition design. To experimentally verify our design choices on the mobility-related conditions, we conduct an ablation study on the mobility-related conditions with the following five variants ablated from our full condition design: w/o timestamp, w/o weekday, w/o almanac, w/o teaching mode, and w/o weather. Table 11 shows the results. We observe that CrowdMirage consistently and significantly outperforms all ablated variants, showing an average improvement of 25.14% and 20.04% on crowd flow and OD transition flow simulation respectively. These results strongly support our design choices on the selected mobility-related conditions.

7.4.3 Case study. We conduct several case studies to demonstrate the simulated crowd flow under three different conditions. The first two conditions are previously seen conditions in history, i.e., 1) “Class day without bad weather”: Monday, no day-off, face-to-face teaching, no weather warning signal; and 2) “No-class day with bad weather”: Tuesday, examination, black rainstorm, thunderstorm. We also consider an unseen condition: 3) “Class day with bad weather”: Tuesday, no day-off, face-to-face teaching, black rainstorm, thunderstorm. The case studies are presented in Figure 8 with two snapshots. We plot both the integrated flow matrix and the crowd flow with the simulated OD transition flow on maps via heatmap. We also plot the ground truth of the first two conditions.

Table 11. Ablation study on mobility-related condition design

Ablation	Crowd Flow				OD Transition Flow			
	JSD ↓	CPC ↑	RMSE ↓	NRMSE ↓	JSD ↓	CPC ↑	RMSE ↓	NRMSE ↓
w/o timestamp	0.1321	0.7381	56.2118	0.0315	0.0053	0.3836	1.1007	0.0021
w/o weekday	0.1207	0.7626	53.9037	0.0302	0.0025	0.4909	0.9440	0.0018
w/o almanac	0.1452	0.6862	65.6422	0.0368	0.0035	0.4456	1.0104	0.0019
w/o teaching arrangement	0.1308	0.7346	58.9069	0.0330	0.0022	0.4884	0.8608	0.0016
w/o weather	0.1447	0.7132	65.9981	0.0370	0.0025	0.4847	0.8976	0.0017
CrowdMirage	0.1085	0.8007	43.0348	0.0241	0.0020	0.5143	0.7861	0.0015

First, we observe that condition 1 at 9:57 shows a clear *crowd rush*, which is presented in the heatmap as hot zones of concentrated movement towards 5 teaching buildings, because the regular class time is at 10 while not at 19:20. Moreover, condition 2 shows *no rush* at 9:57 because of the examination period (no-class day). Moreover, at 19:20, we observe a *high crowd flow* in condition 1 and a *low crowd flow* in condition 2, because both no-class day and bad weather decrease the on-campus crowd flow. In all these cases, our simulation closely resembles the ground truth.

Second, we see that for the unseen condition 3, our simulation shows a *low crowd rush* at 9:57. Because the crowd rush reasonably appears on class day, while the bad weather implies a low crowd flow. Moreover, at 19:20, we observe a medium crowd flow (lower than condition 1 but higher than condition 2). Because the bad weather condition reduces the expected high crowd flow on class day. This implies that our method learnt to effectively simulate the crowd rush on class day, while also capturing the effect of extreme weather on the reduction of crowd and OD transition flow.

8 DISCUSSION ON POTENTIAL USE CASES

Recent studies have demonstrated the practical implications of crowd mobility prediction. For instance, a model predictive control strategy for air conditioning systems [43], based on dynamic passenger flow, has shown potential in maintaining indoor temperature stability, providing a faster system response, and saving 13% energy on typical days and 10% energy during cooling seasons. The FI-700 elevator management system [44] uses predicted human flow to enhance comfort and energy efficiency, and has succeeded in reducing average waiting times by up to 20% during busy periods. Moreover, beyond hardware facilities, crowd flow prediction can also prevent crowd congestion by providing early warnings for locations such as gyms, cafeterias, conference halls, or critical public places during events; security resources can be pre-allocated and evacuation measures can be prepared in advance for popular locations, reducing the risk of stampede incidents [30].

On the other hand, crowd flow simulation can serve as an effective sandbox for evaluating crowd flow-related tasks, in particular providing high-quality synthetic crowd flow data for previous unseen scenarios. Specifically, systems mentioned above like the air conditioning and elevator predictive control systems can be tested with simulated crowd flow under extreme or rare scenarios. Furthermore, for a large crowd that poses potential dangers under certain conditions, crowd flow simulation can serve as an effective sandbox for orchestrating evacuation measures and scheduling security resources, providing a test bed with multiple rounds of experiments to enhance public safety. In addition, for studies on crowd behavior analysis [3, 48], and crowd-driven gamification design [27], crowd flow simulation can generate the underlying crowd framework for a more realistic background for these research.

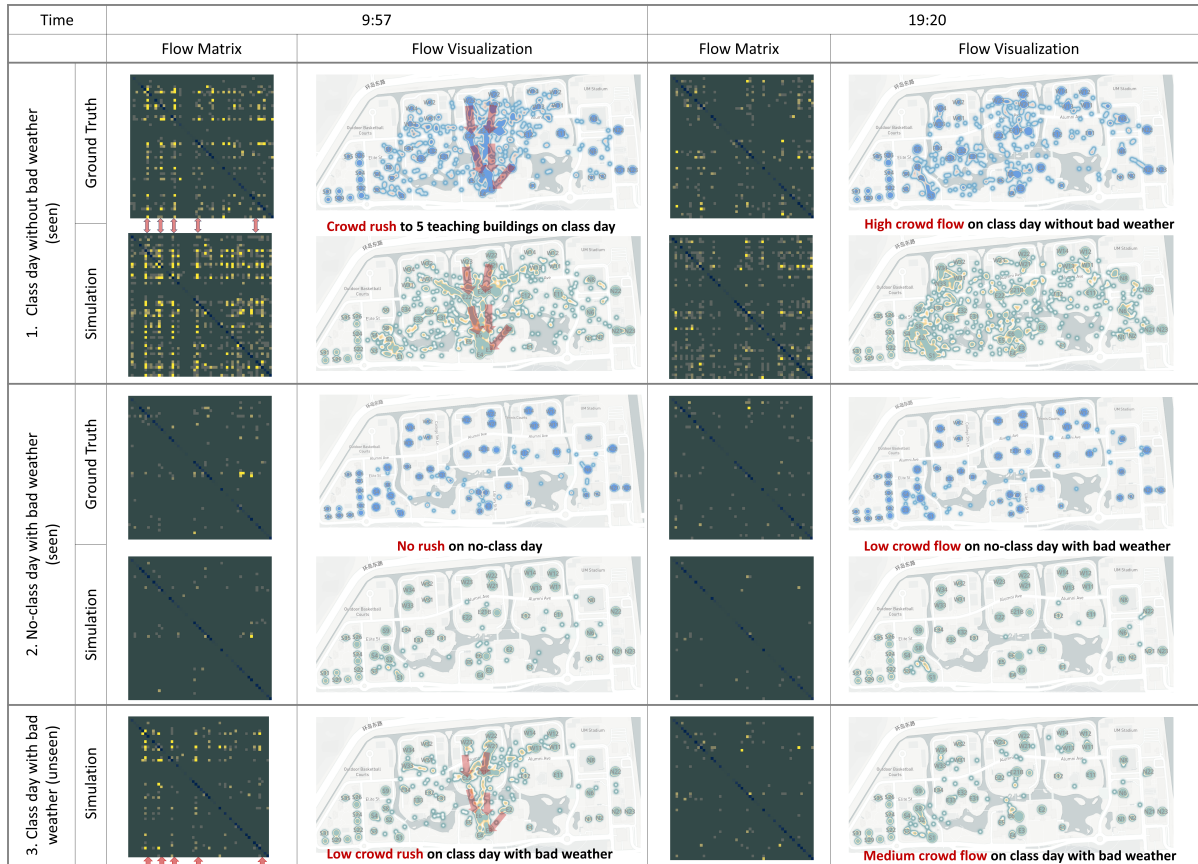


Fig. 8. Case study of crowd flow simulation with three conditions and two snapshots at 09:57 (left) and 19:20 (right).

9 CONCLUSION

In this paper, we propose CrowdMirage, an end-to-end mobility digital twin framework for smart campuses. Specifically, we first design an end-to-end human mobility trace extraction pipeline from the comprehensive but noisy WiFi connection logs. We then propose a cross-grained crowd flow prediction model with a learnable knowledge transfer mechanism between different granularities, so as to effectively forecast future crowd flow on both building and floor levels. Meanwhile, we formulate the crowd flow simulation problem as a conditional generation problem, and systematically identify influential context factors and resort to conditional diffusion models for the conditional generation in a controllable manner. We build a prototype of our mobility digital twin CrowdMirage and evaluate it on the long-term and comprehensive WiFi connection logs collected in our targeted university for the past two years. Results show that our CrowdMirage achieves superior performance in both crowd flow prediction and simulation tasks, with 3.35%-9.09% and 8.80%-68.92% improvement over the best-performing baselines, respectively. Moreover, our case study shows that CrowdMirage cannot only accurately forecast cross-grained crowd flow across different cases, but also simulate interpretable crowd flow under previously seen conditions.

In the future, we plan to investigate the mobility digital twin with trajectory data, so as to provide finer-grained mobility monitoring, prediction, and simulation.

ACKNOWLEDGMENTS

This project has received funding from the University of Macau (MYRG2022-00048-IOTSC) and the Science and Technology Development Fund, Macau SAR (0047/2022/A1, 001/2024/SKL), NSF of China (61802325), the Open Research Project Programme of the State Key Laboratory of Internet of Things for Smart City at the University of Macau (SKL-IoTSC(UM)-2021-2023/ORP/GA06/2022), the FuXiaQuan National Independent Innovation Demonstration Zone Collaborative Innovation Platform (3502ZCQXT2021003), and UIC Research under Grant (UICR0700021-22). This work was performed in part at SICC which is supported by SKL-IOTSC, University of Macau.

REFERENCES

- [1] Azza A.Faiad, Sarah M. Abdel-Ghany, Mohammed Ayachi, and Shehab Ahmed. 2023. City Scale Digital Twins for Mobility Emissions Evaluation. In *2023 International Wireless Communications and Mobile Computing (IWCMC)*. 1166–1171. <https://doi.org/10.1109/IWCMC58020.2023.10182944>
- [2] James E Anderson. 2011. The gravity model. *Annu. Rev. Econ.* 3, 1 (2011), 133–160.
- [3] Ahmad Hoirul Basori, Sharaf Jameel Malebary, Andi Besse Firdausiah Mansur, Andi Tenriawaru, Norazah Yusof, Arda Yunianta, and Omar M Barukab. 2021. Intelligent socio-emotional control of pedestrian crowd behaviour inside smart city. *Procedia Computer Science* 182 (2021), 80–88.
- [4] Kristi Birch. 2022. KU School of Nursing launches “Metaversity” offering virtual, immersive learning. Retrieved Jan 15, 2024 from <https://www.kumc.edu/about/news/news-archive/nursing-metaversity.html>
- [5] Diego M Botín-Sanabria, Adriana-Simona Mihaita, Rodrigo E Peimbert-García, Mauricio A Ramírez-Moreno, Ricardo A Ramírez-Mendoza, and Jorge de J Lozoya-Santos. 2022. Digital twin technology challenges and applications: A comprehensive review. *Remote Sensing* 14, 6 (2022), 1335.
- [6] Noli Brazil. 2022. Environmental inequality in the neighborhood networks of urban mobility in US cities. *Proceedings of the National Academy of Sciences* 119, 17 (2022), e2117776119.
- [7] Rebecca Cairns. 2023. Could the ‘Metaversity’ be the college campus of the future? Retrieved Jan 15, 2024 from <https://www.cnn.com/2023/06/28/americas/metaversity-virtual-reality-morehouse-college-hnk-spc-intl/index.html>
- [8] Longbiao Chen, Jérémie Jakubowicz, Dingqi Yang, Daqing Zhang, and Gang Pan. 2016. Fine-grained urban event detection and characterization based on tensor cofactorization. *IEEE Transactions on Human-Machine Systems* 47, 3 (2016), 380–391.
- [9] Longbiao Chen, Daqing Zhang, Leye Wang, Dingqi Yang, Xiaojuan Ma, Shijian Li, Zhaohui Wu, Gang Pan, Thi-Mai-Trang Nguyen, and Jérémie Jakubowicz. 2016. Dynamic cluster-based over-demand prediction in bike sharing systems. In *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing*. 841–852.
- [10] Eunjoon Cho, Seth A. Myers, and Jure Leskovec. 2011. Friendship and mobility: user movement in location-based social networks. In *Proceedings of the 17th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (San Diego, California, USA) (KDD '11)*. Association for Computing Machinery, New York, NY, USA, 1082–1090. <https://doi.org/10.1145/2020408.2020579>
- [11] Junyoung Chung, Caglar Gulcehre, KyungHyun Cho, and Yoshua Bengio. 2014. Empirical evaluation of gated recurrent neural networks on sequence modeling. *arXiv preprint arXiv:1412.3555* (2014).
- [12] Vedant Das Swain, Jiajia Xie, Maanit Madan, Sonia Sargolzaei, James Cai, Munmun De Choudhury, Gregory D Abowd, Lauren N Steimle, and B Aditya Prakash. 2023. Empirical networks for localized COVID-19 interventions using WiFi infrastructure at university campuses. *Frontiers in Digital Health* 5 (2023), 1060828.
- [13] Boris Delaunay et al. 1934. Sur la sphere vide. *Izv. Akad. Nauk SSSR, Otdelenie Matematicheskii i Estestvennyka Nauk* 7, 793-800 (1934), 1–2.
- [14] Bangchao Deng, Bingqing Qu, Pengyang Wang, and Dingqi Yang. 2024. REPLAY: Modeling Time-Varying Temporal Regularities of Human Mobility for Location Prediction over Sparse Trajectories. *arXiv preprint arXiv:2402.16310* (2024).
- [15] Prafulla Dhariwal and Alexander Nichol. 2021. Diffusion models beat gans on image synthesis. *Advances in neural information processing systems* 34 (2021), 8780–8794.
- [16] Andriod Documentation. 2024. MAC Randomization Behavior. Retrieved March 18, 2024 from <https://source.android.com/docs/core/connect/wifi-mac-randomization-behavior#persistent>
- [17] Christos Doulkeridis, Akrivi Vlachou, Nikos Pelekis, and Yannis Theodoridis. 2021. A Survey on Big Data Processing Frameworks for Mobility Analytics. *ACM SIGMOD Record* 50, 2 (2021), 18–29.

- [18] Harris Drucker, Christopher J Burges, Linda Kaufman, Alex Smola, and Vladimir Vapnik. 1996. Support vector regression machines. *Advances in neural information processing systems* 9 (1996).
- [19] Sijing Duan, Feng Lyu, Huaqing Wu, Wenxiong Chen, Huali Lu, Zhe Dong, and Xuemin Shen. 2024. MOTO: Mobility-Aware Online Task Offloading With Adaptive Load Balancing in Small-Cell MEC. *IEEE Transactions on Mobile Computing* 23, 1 (2024), 645–659. <https://doi.org/10.1109/TMC.2022.3220720>
- [20] Jeffrey L Elman. 1990. Finding structure in time. *Cognitive science* 14, 2 (1990), 179–211.
- [21] Yuchen Fang, Yanjun Qin, Haiyong Luo, Fang Zhao, Bingbing Xu, Liang Zeng, and Chenxing Wang. 2023. When spatio-temporal meet wavelets: Disentangled traffic forecasting via efficient spectral graph attention networks. In *2023 IEEE 39th International Conference on Data Engineering (ICDE)*. IEEE, 517–529.
- [22] Zhihan Fang, Yu Yang, Guang Yang, Yikuan Xian, Fan Zhang, and Desheng Zhang. 2021. CellSense: Human mobility recovery via cellular network data enhancement. *Proceedings of the ACM on IMWUT* 5, 3 (2021), 1–22.
- [23] Jie Feng, Zeyu Yang, Fengli Xu, Haisu Yu, Mudan Wang, and Yong Li. 2020. Learning to simulate human mobility. In *Proceedings of the 26th ACM SIGKDD international conference on knowledge discovery & data mining*. 3426–3433.
- [24] Jaime B Fernandez, Mani Dhingra, Sukanya Mandal, Noel E O'Connor, Aphra Kerr, Kevin McGuinness, Muhammad Intizar Ali, Jamie Cudden, and Kieran Mahon. 2023. Smart DCU digital twin: towards smarter universities. (2023).
- [25] Shengnan Guo, Youfang Lin, Ning Feng, Chao Song, and Huaiyu Wan. 2019. Attention based spatial-temporal graph convolutional networks for traffic flow forecasting. In *Proceedings of the AAAI conference on artificial intelligence*, Vol. 33. 922–929.
- [26] William Harvey, Saeid Naderiparizi, and Frank Wood. 2021. Conditional image generation by conditioning variational auto-encoders. *arXiv preprint arXiv:2102.12037* (2021).
- [27] Brandon Haworth, Muhammad Usman, Davide Schaumann, Nilay Chakraborty, Glen Berseth, Petros Faloutsos, and Mubbasir Kapadia. 2020. Gamification of crowd-driven environment design. *IEEE computer graphics and applications* 41, 4 (2020), 107–117.
- [28] Jonathan Ho, Ajay Jain, and Pieter Abbeel. 2020. Denoising diffusion probabilistic models. *Advances in neural information processing systems* 33 (2020), 6840–6851.
- [29] Sepp Hochreiter and Jürgen Schmidhuber. 1997. Long short-term memory. *Neural computation* 9, 8 (1997), 1735–1780.
- [30] Anders Johansson, Michael Batty, Konrad Hayashi, Osama Al Bar, David Marcozzi, and Ziad A Memish. 2012. Crowd and environmental management during mass gatherings. *The Lancet infectious diseases* 12, 2 (2012), 150–156.
- [31] Digital Twin Lab. 2024. TU/E campus digital twin for smart building management and control. Retrieved Jan 15, 2024 from <https://www.tue.nl/en/research/institutes/eindhoven-artificial-intelligence-systems-institute/digital-twin-lab>
- [32] Luis Alberto Laurens-Arredondo and Lilibeth Laurens. 2023. Metaversity: Beyond Emerging Educational Technology. *Sustainability* 15, 22 (2023). <https://www.mdpi.com/2071-1050/15/22/15844>
- [33] Juha K Laurila, Daniel Gatica-Perez, Imad Aad, Jan Blom, Olivier Bornet, Trinh Minh Tri Do, Olivier Dousse, Julien Eberle, and Markus Miettinen. 2013. From big smartphone data to worldwide research: The mobile data challenge. *Pervasive and Mobile Computing* 9, 6 (2013), 752–771.
- [34] Isaac Lee. 2020. Penn students use digital platform Gather to imitate in-person office hours. Retrieved Jan 15, 2024 from <https://www.thedp.com/article/2020/10/gather-town-penn-cis-virtual-office-hours>
- [35] Nat Levy. 2023. Digital Twin of UT Campus Visualizes Present, Past, Future Energy Needs. Retrieved Jan 15, 2024 from <https://news.utexas.edu/2023/11/09/digital-twin-of-ut-campus-visualizes-present-past-future-energy-needs/>
- [36] Yaguang Li, Rose Yu, Cyrus Shahabi, and Yan Liu. 2017. Diffusion convolutional recurrent neural network: Data-driven traffic forecasting. *arXiv preprint arXiv:1707.01926* (2017).
- [37] Yong Li, Yuan Yuan, Jingtao Ding, and Depeng Jin. 2023. Learning the complexity of urban mobility with deep generative collaboration network. (2023).
- [38] Ziqian Lin, Jie Feng, Ziyang Lu, Yong Li, and Depeng Jin. 2019. Deepstn+: Context-aware spatial-temporal neural network for crowd flow prediction in metropolis. In *Proceedings of the AAAI conference on artificial intelligence*, Vol. 33. 1020–1027.
- [39] Xi Liu, Hanzhou Chen, and Clio Andris. 2018. trajGANs: Using generative adversarial networks for geo-privacy protection of trajectory data (Vision paper). In *Location Privacy and Security Workshop*. 1–7.
- [40] Massimiliano Luca, Gianni Barlacchi, Bruno Lepri, and Luca Pappalardo. 2021. A survey on deep learning for human mobility. *ACM Computing Surveys (CSUR)* 55, 1 (2021), 1–44.
- [41] Yisheng Lv, Yanjie Duan, Wenwen Kang, Zhengxi Li, and Fei-Yue Wang. 2014. Traffic flow prediction with big data: a deep learning approach. *IEEE Transactions on Intelligent Transportation Systems* 16, 2 (2014), 865–873.
- [42] Feng Lyu, Ju Ren, Nan Cheng, Peng Yang, Minglu Li, Yaoxue Zhang, and Xuemin Sherman Shen. 2021. LeaD: Large-Scale Edge Cache Deployment Based on Spatio-Temporal WiFi Traffic Statistics. *IEEE Transactions on Mobile Computing* 20, 8 (2021), 2607–2623. <https://doi.org/10.1109/TMC.2020.2984261>
- [43] Kai Ma, Dan Wang, Wei Wang, Shihao Zhu, and Yuying Sun. 2023. Model predictive control strategy of air conditioning system based on dynamic passenger flow: an airport terminal building case study. In *Building Simulation 2023*, Vol. 18. IBPSA, 3398–3405.

- [44] Tomoaki Maehara, Takahiro Hatori, Satoru Toriyabe, Takefumi Kuragane, and Hisatake Yanagisawa. 2020. FI-700 Elevator Management System Featuring Human Flow Prediction for Comfort of Travel Inside Buildings. *Technologies to Achieve Smarter Mobility* 69, 6 (2020).
- [45] Mehdi Mirza and Simon Osindero. 2014. Conditional generative adversarial nets. *arXiv preprint arXiv:1411.1784* (2014).
- [46] Esteban Moro, Dan Calacci, Xiaowen Dong, and Alex Pentland. 2021. Mobility patterns are associated with experienced income segregation in large US cities. *Nature communications* 12, 1 (2021), 4633.
- [47] University News and Publications. 2021. Fisk University, HTC VIVE, T-Mobile and VictoryXR Launch 5G-Powered VR Human Cadaver Lab. Retrieved Jan 15, 2024 from <https://www.fisk.edu/university-news-and-publications/fisk-university-htc-vive-t-mobile-and-victoryxr-launch-5g-powered-vr-human-cadaver-lab/>
- [48] Alexandre Nicolas and Fadratul Hafinaz Hassan. 2023. Social groups in pedestrian crowds: review of their influence on the dynamics and their modelling. *Transportmetrica A: transport science* 19, 1 (2023), 1970651.
- [49] University of Glasgow. 2019. University of Glasgow - Digital Twins for Net-Zero Campuses. Retrieved Jan 15, 2024 from <https://www.iesve.com/services/projects/15681/university-of-glasgow>
- [50] Ernest George Ravenstein. 1885. The laws of migration. *Journal of the statistical society of London* 48, 2 (1885), 167–235.
- [51] Luis Roda-Sanchez, Flavio Cirillo, Gürkan Solmaz, Tobias Jacobs, Celia Garrido-Hidalgo, Teresa Olivares, and Ernő Kovacs. 2023. Building a Smart Campus Digital Twin: System, Analytics and Lessons Learned From a Real-World Project. *IEEE Internet of Things Journal* (2023), 1–1. <https://doi.org/10.1109/JIOT.2023.3300447>
- [52] Robin Rombach, Andreas Blattmann, Dominik Lorenz, Patrick Esser, and Björn Ommer. 2022. High-resolution image synthesis with latent diffusion models. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*. 10684–10695.
- [53] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. 2015. U-net: Convolutional networks for biomedical image segmentation. In *Medical image computing and computer-assisted intervention—MICCAI 2015: 18th international conference, Munich, Germany, October 5-9, 2015, proceedings, part III 18*. Springer, 234–241.
- [54] Zezhi Shao, Zhao Zhang, Wei Wei, Fei Wang, Yongjun Xu, Xin Cao, and Christian S Jensen. 2022. Decoupled dynamic spatial-temporal graph neural network for traffic forecasting. *arXiv preprint arXiv:2206.09112* (2022).
- [55] Seungjae Shin, Hongseok Jeon, Chunglae Cho, Seunghyun Yoon, and Taeyeon Kim. 2020. User mobility synthesis based on generative adversarial networks: A survey. In *2020 22nd International Conference on Advanced Communication Technology (ICACT)*. IEEE, 94–103.
- [56] Filippo Simini, Gianni Barlacchi, Massimiliano Luca, and Luca Pappalardo. 2020. Deep Gravity: enhancing mobility flows generation with deep neural networks and geographic information. *arXiv preprint arXiv:2012.00489* (2020).
- [57] Arkadiusz Stopczynski, Vedran Sekara, Piotr Sapiezynski, Andrea Cuttone, Mette My Madsen, Jakob Eg Larsen, and Sune Lehmann. 2014. Measuring large-scale social networks with high resolution. *PLoS one* 9, 4 (2014), e95978.
- [58] Apple Support. 2023. Use private Wi-Fi addresses on iPhone, iPad, iPod touch, and Apple Watch. Retrieved August 30, 2023 from <https://support.apple.com/en-us/102509>
- [59] Yongxin Tong, Yuxiang Zeng, Zimu Zhou, Lei Chen, and Ke Xu. 2022. Unified Route Planning for Shared Mobility: An Insertion-based Framework. *ACM Transactions on Database Systems (TODS)* 47, 1 (2022), 1–48.
- [60] Heriot-Watt University. 2020. Heriot-Watt University - Campus Digital Twin. Retrieved Jan 15, 2024 from <https://www.iesve.com/services/projects/14343/heriot-watt-edinburgh>
- [61] Rice University. 2023. Rice University Campus Digital Twin - A 3D model of Rice University Campus Facilities searchable by building. Retrieved Jan 15, 2024 from <https://experience.arcgis.com/experience/4db18397079144e5afa6af9ae48e0ce4/page/Page-1/?views=2D>
- [62] Srinivasan Venkatramanan, Adam Sadilek, Arindam Fadikar, Christopher L Barrett, Matthew Biggerstaff, Jiangzhuo Chen, Xerxes Dotiwalla, Paul Eastham, Bryant Gipson, Dave Higdon, et al. 2021. Forecasting influenza activity using machine-learned mobility map. *Nature communications* 12, 1 (2021), 726.
- [63] Ruizhi Wu, Guangchun Luo, Junming Shao, Ling Tian, and Chengzong Peng. 2018. Location prediction on trajectory data: A review. *Big data mining and analytics* 1, 2 (2018), 108–127.
- [64] Zonghan Wu, Shirui Pan, Guodong Long, Jing Jiang, and Chengqi Zhang. 2019. Graph wavenet for deep spatial-temporal graph modeling. *arXiv preprint arXiv:1906.00121* (2019).
- [65] Takahiro Yabe, Kota Tsubouchi, and Yoshihide Sekimoto. 2017. CityFlowFragility: Measuring the fragility of people flow in cities to disasters using GPS data collected from smartphones. *Proceedings of the ACM on IMWUT* 1, 3 (2017), 1–17.
- [66] Dingqi Yang, Bingqing Qu, and Philippe Cudre-Mauroux. 2020. Location-centric social media analytics: Challenges and opportunities for smart cities. *IEEE Intelligent Systems* 36, 5 (2020), 3–10.
- [67] Dingqi Yang, Bingqing Qu, Jie Yang, and Philippe Cudre-Mauroux. 2019. Revisiting user mobility and social relationships in lbsns: a hypergraph embedding approach. In *The world wide web conference*. 2147–2157.
- [68] Dingqi Yang, Daqing Zhang, Zhiyong Yu, and Zhu Wang. 2013. A sentiment-enhanced personalized location recommendation system. In *Proceedings of the 24th ACM conference on hypertext and social media*. 119–128.
- [69] Xi Yang, Suining He, Bing Wang, and Mahan Tabatabaie. 2021. Spatio-temporal graph attention embedding for joint crowd flow and transition predictions: A Wi-Fi-based mobility case study. *Proceedings of the ACM on IMWUT* 5, 4 (2021), 1–24.

- [70] Bing Yu, Haoteng Yin, and Zhanxing Zhu. 2017. Spatio-temporal graph convolutional networks: A deep learning framework for traffic forecasting. *arXiv preprint arXiv:1709.04875* (2017).
- [71] Zhiwen Yu, Huadong Ma, Bin Guo, and Zheng Yang. 2021. Crowdsensing 2.0. *Commun. ACM* 64, 11 (2021), 76–80.
- [72] Agustín Zaballos, Alan Briones, Alba Massa, Pol Centelles, and Victor Caballero. 2020. A smart campus’ digital twin for sustainable comfort monitoring. *Sustainability* 12, 21 (2020), 9196.
- [73] Qi Zhang, Hengshu Zhu, Peng Wang, Enhong Chen, and Hui Xiong. 2023. Hierarchical Wi-Fi Trajectory Embedding for Indoor User Mobility Pattern Analysis. *Proceedings of the ACM on IMWUT* 7, 2 (2023), 1–21.
- [74] Shiyu Zhang, Bangchao Deng, and Dingqi Yang. 2023. CrowdTelescope: Wi-Fi-positioning-based multi-grained spatiotemporal crowd flow prediction for smart campus. *CCF Transactions on Pervasive Computing and Interaction* 5, 1 (2023), 31–44.
- [75] Yu Zheng, Quannan Li, Yukun Chen, Xing Xie, and Wei-Ying Ma. 2008. Understanding mobility based on GPS data. In *Proceedings of the 10th international conference on Ubiquitous computing*. 312–321.